

Imthurn, Gabriel

Assessment des Singens in der Sekundarstufe 1. Eine qualitative Interviewstudie

Wien; Zürich : LIT Verlag 2023, 431 S. - (Empirische Forschung zur Musikpädagogik; 12) - (Dissertation, Hochschule für Musik, Theater und Medien Hannover, 2023)



Quellenangabe/ Reference:

Imthurn, Gabriel: Assessment des Singens in der Sekundarstufe 1. Eine qualitative Interviewstudie.
Wien; Zürich : LIT Verlag 2023, 431 S. - (Empirische Forschung zur Musikpädagogik; 12) - (Dissertation, Hochschule für Musik, Theater und Medien Hannover, 2023) - URN: urn:nbn:de:0111-pedocs-333343 - DOI: 10.25656/01:33334; 10.52038/9783643803412

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-333343>

<https://doi.org/10.25656/01:33334>

Nutzungsbedingungen

Dieses Dokument steht unter folgender Creative Commons-Lizenz: <http://creativecommons.org/licenses/by/4.0/deed.de> - Sie dürfen das Werk bzw. den Inhalt vervielfältigen, verbreiten und öffentlich zugänglich machen sowie Abwandlungen und Bearbeitungen des Werkes bzw. Inhaltes anfertigen, solange Sie den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

This document is published under following Creative Commons-License: <http://creativecommons.org/licenses/by/4.0/deed.en> - You may copy, distribute and render this document accessible, make adaptations of this work or its contents accessible to the public as long as you attribute the work in the manner specified by the author or licensor.

By using this particular document, you accept the above-stated conditions of use.



Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Mitglied der


Leibniz-Gemeinschaft

Gabriel Imthurn

Assessment des Singens in der Sekundarstufe 1

Eine qualitative Interviewstudie



Gabriel Imthurn

Assessment des Singens in der Sekundarstufe 1

Empirische Forschung zur Musikpädagogik

herausgegeben von

Prof. Dr. Andreas Lehmann-Wermser
(Hochschule für Musik, Theater und Medien Hannover)

Band 12

LIT

Gabriel Imthurn

Assessment des Singens in der Sekundarstufe 1

Eine qualitative Interviewstudie

LIT

Umschlagbild: © HighwayStarz (Ian Allenden)/Depositphotos.com

Publiziert mit Unterstützung des Schweizerischen Nationalfonds zur Förderung der wissenschaftlichen Forschung.



Gedruckt auf alterungsbeständigem Werkdruckpapier entsprechend
ANSI Z3948 DIN ISO 9706

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.dnb.de> abrufbar.

ISBN 978-3-643-80341-2 (br.)

ISBN 978-3-643-85341-7 (PDF)

ISBN 978-3-643-85344-8 (OA)

DOI: <https://doi.org/10.52038/9783643803412>

Zugl.: Hannover, Hochsch. für Musik, Theater und Medien, Diss., 2023



Dieses Werk ist mit einer CC BY 4.0-Lizenz lizenziert.

© **LIT VERLAG** GmbH & Co. KG Wien,
Zweigniederlassung Zürich 2023
Flössergasse 10
CH-8001 Zürich
Tel. +41 (0) 76-632 84 35 E-Mail:
zuerich@lit-verlag.ch <https://www.lit-verlag.ch>

Auslieferung:

Deutschland: **LIT Verlag**, Fresenstr. 2, D-48159 Münster
Tel. +49 (0) 2 51-620 32 22, E-Mail: vertrieb@lit-verlag.de

Für meine Familie

VORWORT DES HERAUSGEBERS

Der Titel der hier vorgelegten Dissertationsschrift vereint vertraute Begriffe, die wenig Fragen aufzuwerfen scheinen. *Singen* ist wohl in allen Ländern ein integraler Bestandteil des Musikunterrichts. Es mag Unterschiede in der methodischen Vermittlung geben: Stimmbildung und Solmisation etwa sind in verschiedenen Ländern unterschiedlich stark vertreten. Es mag Unterschiede geben in der musikkulturellen Einbettung des schulischen Gesangs: Während dieser in einigen Ländern von einer reichen und weitverbreiteten Chortradition als „usueller Gebrauchspraxis“ (Kaiser) profitiert (wie etwa in den baltischen Staaten oder im alpenländischen Raum), erweisen sich in anderen Ländern eher stilistische und gesangstechnische Einflüsse der Popularmusik als prägend. Bei allen Unterschieden aber bleibt das Singen doch stets die wichtigste Form musikalischer Praxis in der Schule. Warum also diese (in mehr als einer Hinsicht) gewichtige Schrift?

Die Provokation des Buches liegt in der Zusammenstellung mit dem zweiten Nomen im Titel: mit *Assessment*. Auch wenn in dieser Reihe das „denglische“ Kauderwelsch gerne vermieden wird: Hier ist es notwendig, es gibt keine gute Übersetzung. Der englische Begriff meint nämlich ungleich mehr als nur *Bewertung* oder gar *Benotung*. Er bezeichnet viel allgemeiner die Erfassung von (musikalischer) *Leistung* im Sinne eines erreichten Niveaus musikalischen Verhaltens. Aber er ist durch zwei Merkmale von einer bloßen Leistungsmessung unterschieden. Zum einen ist er (meist) von einem pädagogischen Impetus getragen. Was kann *assessment* zu einer Verbesserung des musikalischen Vermögens beitragen? Wie können Lernende von dieser Art des *assessments* profitieren? Davon handelt dieses Buch in wichtigen Teilen, wenn der Begriff des *assessment for learning*, der Leistungserfassung im Interesse des Lernens, in vielen Facetten verfolgt wird.

Assessment ist aber zum anderen auch durch den Versuch geprägt, auf rationaler Basis ein Verständnis für Lern- und Leistungsprozesse zu erreichen. Die verschiedenen Forschungsprojekte zur Kompetenzmodellierung unterschiedlicher Domänen, die seit 2007 in Deutschland durchgeführt worden sind, sind als solche Erkundungen der Struktur musikalischen Lernens zu verstehen. KOMUS, KOPRA-M oder MARKO sind in dieser

Hinsicht Versuche, ex post das Ergebnis (z. B. des Singens) in seiner Struktur und in seiner Stufung zu erfassen.

Das Buch Gabriel Imthurns setzt freilich noch einen anderen Akzent. Nicht das Ergebnis steht im Mittelpunkt oder ist der Ausgangspunkt des forschenden Vorgehens, sondern die (stimmlichen) Voraussetzungen: Was bedeutet etwa die Mutation für die Arbeit mit der Stimme in der Schule? In welchen Sozialformen z. B. kann eine optimale Förderung dann erfolgen? Wie können Rückmeldungen in besonders produktiver, weil akzeptierender Form gelingen?

Solche Fragen verweisen auf das Handlungsrepertoire der Lehrenden. So ist es folgerichtig, dass Gabriel Imthurn eine Interviewstudie unter ihnen durchführt. In einem komplexen Multi-Method-Design wird der Frage nachgegangen, wie Lehrende an Schweizer Schulen mit singenden Kindern bzw. Jugendlichen umgehen. Wie arbeiten sie? Welche Ziele verfolgen sie? Welche Fördermaßnahmen ergreifen sie? Wie erfassen sie den Stand der Schülerinnen und Schüler und wie kommunizieren sie das? Oder in anderen Worten: Welchen Stellenwert hat *assessment* in dieser Arbeit?

In diesen Fragen stecken mehrere provozierende Momente. Zum einen wird damit der Blick auf die vorhandenen Kompetenzen von Lehrkräften gerichtet. Was wissen sie eigentlich über die stimmlichen Gegebenheiten? Welche Fördermöglichkeiten kennen sie? Welche Schwierigkeiten und Möglichkeiten sehen sie dabei für ihre Arbeit? Welchen Kurs verfolgen sie zwischen (selektierender) Benotung und (fördernder) Erfassung? Auch wenn dieses Buch als primär qualitative Studie nur eine kleine Stichprobe erfasst, so ist doch zu befürchten, dass bei Lehramtsstudierenden und Lehrkräften hier große Defizite bestehen. Das gilt für das bei ihnen vorhandene *Handwerkszeug* ebenso wie für die ausformulierten Ziele von Unterrichtenden. Diese Defizite sichtbar gemacht zu haben und zugleich Ziele für die Weiterentwicklung aufgezeigt zu haben, ist eine große Stärke der Schrift.

Dass das Buch dabei für den deutschsprachigen Raum erstmals die vorhandene internationale Literatur sichtet und produktiv macht, ist ein zusätzliches Verdienst der Schrift Gabriel Imthurns. Nach der Lektüre wird deutlicher, was wir noch lernen und in welche Richtung wir uns entwickeln können.

VORWORT DES AUTORS

Die vorliegende Dissertation ist eine biografische Zäsur. Nach mehr als 20 Jahren Tätigkeit als Musiklehrer an mehreren Schulen und auf verschiedenen Leistungsniveaus der Sekundarstufe 1, nach Musikunterricht mit insgesamt mehr als 1500 Schülerinnen und Schülern erfolgte vor einigen Jahren als erster Schritt der Wechsel an die Pädagogische Hochschule der Fachhochschule Nordwestschweiz (FHNW). Mit dem Eintauchen in die akademische Welt – was nach der langen praktischen Arbeit und der in der Schweiz traditionellerweise stark praxisorientierten Ausbildung anspruchsvoll war – wuchs der Wunsch, tiefer zu verstehen, wie sich Musikunterricht konstituiert, welche Stellschrauben diesen beeinflussen und wie er weiterentwickelt werden kann. Letztlich führte dies zum vorliegenden Promotionsvorhaben.

Der bekannte Schweizer Schriftsteller und Literaturwissenschaftler beschreibt seinen eigenen Schreibprozess im Interview wie folgt:

„Zunächst schreibt man, weil man mit irgendetwas nicht fertig wird und weil es einem interessiert und weil man es erst, so geht es mir, beim Schreiben wirklich herausfindet, was man schreiben wollte.“ (Muschg, 2020 [30:25–30:36])

Der Denk-, Forschungs- und Schreibprozess dieser Arbeit war langwierig und gestaltete sich im Sinne des zitierten Schriftstellers, auch weil sich die äußeren Umstände stetig verändert haben. So musste ich in dieser Zeit von meinen Eltern Abschied nehmen, und eine neue berufliche Herausforderung hat viel Zeit und Kraft beansprucht. Nicht zuletzt hat auch die Pandemie ihren Tribut in Beruf und Familie gefordert. Trotzdem bin ich dankbar, nach viereinhalb Jahren das Vorhaben abschließen zu können.

Mein Dank geht in erster Linie an Prof. Dr. Andreas Lehmann-Wermser, der den Schweizer zuerst empfing, dann bestärkt und in der langen Zeit begleitet hat. Trotz erheblicher Distanz zwischen Küttigen (im Kanton Aargau in der Schweiz) und Hannover waren es seine wertvollen Hinweise zu Inhalt, Struktur und Argumentationsweisen und nicht zuletzt seine Ermutigung durchzuhalten, die zum Abschluss dieser Arbeit führten.

Mein Dank geht ebenfalls an meinen Zweitbetreuer Prof. Dr. Helmut Schaumberger. Seine inhaltlichen und strukturellen Hinweise waren in der letzten Phase sehr wertvoll. Zudem möchte ich mich bei Prof. Dr. Judith Hangartner für das offene Gespräch zu meinem autoethnografischen Exkurs bedanken.

Und mein Dank gilt meiner Arbeitgeberin, der Pädagogischen Hochschule der Fachhochschule Nordwestschweiz (PH FHNW), vertreten durch die Direktorin Prof. Dr. Sabina Larcher, die im Rahmen eines Förderprogramms während dreier Jahre eine Arbeitszeit-Entlastung ermöglicht hat. Zudem möchte ich an dieser Stelle auch Prof. Markus Cslovjecssek danken, der in meinen ersten Jahren an der FHNW das Promotionsvorhaben initiiert hat.

Eine ganze Reihe von Personen hat mich in der letzten Phase begleitet, vorangetrieben, hat nachgefragt und mich ermutigt. Dazu gehört auch meine Kollegin Gabriele Noppeney, welcher ein besonderer Dank für das Korrekturlesen der Arbeit gebührt. Danke auch für die Entlarvung der eingeflossenen Helvetismen.

Wenn mich Simone, meine Ehefrau, nicht unterstützt und angetrieben hätte, trotz Verzicht auf gemeinsame Gartenarbeit und dem Spinnen neuer Geschäftsideen für ihren Stoffladen, dann hätte ich es wohl nicht geschafft. Danke für deine Unterstützung. Meine drei Töchter waren glücklicherweise in einem Alter, in dem Verständnis erwartet werden darf, auch ihnen gilt mein Dank. Sie haben in den letzten Monaten jedem fertiggestellten Kapitel entgegengefeibert. Ich bin stolz auf euch!

Küttigen, im Frühjahr 2023

Gabriel Imthurn

INHALTSVERZEICHNIS

Vorwort des Herausgebers	7
Vorwort des Autors	9
Inhaltsverzeichnis	11
Abbildungen	16
Tabellen	17
Abkürzungen	19
Vorbemerkungen zu gendersensiblen Schreibweisen	20
Abstract	21
 A Einführung	 23
 1 Relevante Kontexte des Forschungsfeldes	 28
1.1 Traditionen des schulischen Singunterrichts	28
1.2 Assessment des Singens als Problemfall	33
1.3 Singen im Fokus musikpädagogischer Forschung	37
 2 Gliederung der Studie	 44
2.1 Bereich B: Singunterricht auf dem Prüfstand	44
2.2 Bereich C: Methode und Untersuchung	50
2.3 Bereich D: Darstellung der Resultate	52
2.4 Bereich E: Ausblick	54
 B Singunterricht auf dem Prüfstand	 57
 3 Exkurs: Autoethnografische Rückblende	 58
3.1 Punkt 1: Der Forschende im Forschungsfeld	58
– Präludium: Vorsingen vor der Klasse	58
– Methode: Der Forschende auf seinen eigenen Spuren	59
– Analyse: Inhuman oder doch angemessen?	60
– Fazit: Der eigene Singunterricht als Ausgangspunkt	62
3.2 Punkt 2: Autoethnografie im Diskurs	64
– Interludium I: Eine verpasste Chance	64
– Methode: Die eigene Erleuchtung als Maßstab?	65
– Analyse: Ziel erreicht, oder doch verfehlt?	70
– Fazit: Erleuchtung als Ausgangspunkt für Entwicklung	72

3.3	Punkt 3: Autoethnografie in der Musikpädagogik	72
–	Interludium II: Die Geschichte vom Dankeschön eines Jungen	72
–	Methode: Autoethnografische Verfahren in der Musikpädagogik	73
–	Analyse: Reflektierte und empathische Betreuung	75
–	Fazit: Die adoleszente Stimme und Geschlechtertrennung werden zum Thema	77
3.4	Punkt 4: Fazit aus dem Exkurs	78
–	Postludium: Was passieren kann, wenn schöne Stimmen entdeckt werden ...	78
–	Analyse: Individuelle Förderung	79
–	Fazit: Vom autoethnografischen Exkurs zur Interviewstudie	81
–	Postskriptum: Wie steht es mit dem Zeitgeist?	83
4	Forschungsstand I – Singen in der Adoleszenz	85
4.1	Singen als urmenschliche Tätigkeit	86
–	Singen als anthropologische Konstante	86
–	Konzepte des Singens	88
–	Singen im Lehrplan	89
4.2	Pubertät – Adoleszenz – Jugend	91
4.3	Die adoleszente Mädchenstimme	93
–	Amerikanische Studien	93
–	Hinweise in der deutschsprachigen Literatur	100
4.4	Die adoleszente Jungenstimme	103
–	Die Phasen der Mutation bei Jungen	104
–	Singen während der Mutation	108
–	Jungen in der Identitätskrise	112
4.5	Geschlechtsspezifische Einflüsse beim Singen im Musikunterricht	115
–	Geschlecht als zentrales Merkmal	116
–	Lehrerinnen oder Lehrer?	118
–	Geschlechtsspezifische Unterschiede in der Wahrnehmung des Singens	119
–	Monoedukative Phasen im Musikunterricht	123
5	Forschungsstand II	
	Assessment im Singunterricht	128
5.1	Assessment als holistisches Förder- und Beurteilungsinstrument	129
–	Individualisierung als Ausgangspunkt	129
–	Assessment als holistisches Förder- und Beurteilungsinstrument	132
–	Das Modell von Earl	134
5.2	Assessment in Music Education	140
–	Principles of Assessment	141
–	Assessment of Singing	148
5.3	Assessment des Singens in der deutschsprachigen Musikpädagogik	154
–	Leistungsbeurteilung im Fach Musik und spezifisch im Singunterricht	154
–	Wahrnehmung und Umgang mit Differenz	157
–	Diagnose- und Beurteilungsinstrumente für das Singen	162

6	Exkurs: Rahmenbedingungen des Schweizer Bildungssystems.....	167
6.1	Kantonale Ausprägungen als Resultat des Föderalismus	168
6.2	Ausbildung von Musiklehrkräften in der Schweiz	170
6.3	Leistungsniveaus in der Sekundarstufe 1.....	172
7	Forschungsfragen.....	176
7.1	Eingrenzung des Forschungsdesiderats	176
7.2	Forschungsfragen	177
C	Methode und Untersuchung.....	181
8	Forschungsdesign.....	182
8.1	Forschungsdesign	185
8.2	Auswertungsmethode: Qualitative Inhaltsanalyse nach Mayring	187
	– Entwicklung der Methode	188
	– Primärverfahren: inhaltlich strukturierende Inhaltsanalyse.....	190
	– Sekundärverfahren: Analyseinstrumente	193
8.3	Interviews als Datenbasis.....	195
	– Sample	196
	– Form des Interviews	198
	– Kommunikationszusammenhang	200
8.4	Gütekriterien der QIA.....	202
	– Reliabilität	203
	– Validität	205
9	Untersuchungsmaterial	208
9.1	Auswahlverfahren für das Sample	208
9.2	Erstellung des Interviewleitfadens	211
9.3	Stichprobenplan und Durchführung der Interviews	212
9.4	Transkription und Festlegung der Analyseeinheiten	218
10	Kategoriensystem und Kodierung	220
10.1	MAXQDA als Forschungsplattform.....	220
10.2	Entwicklung des Kategoriensystems und Kodierung des Materials	222
10.3	Reliabilitätsprüfung der QIA.....	226
10.4	Fazit zur Erstellung des Kategoriensystems	231
D	Darstellung der Resultate	233
11	Quantitative Analyse.....	234
11.1	Quantitative Analyse der Kodierung im Überblick.....	234

11.2	Quantitative Analyse des Bereichs Assessment.....	236
11.3	Skalierende Strukturierung der Sozialformen im Singunterricht.....	242
11.4	Fazit	246
12	Vertikale Achse: Perspektive der Lehrkräfte	247
12.1	Porträts der Lehrkräfte.....	248
–	Der Chorleiter bleibt seiner Leidenschaft treu	249
–	Der Schauspieler entwickelt sich vom Animator zum Coach	250
–	Die langjährige Schulmusikerin.....	251
–	Der Schlagzeuger als Energielieferant.....	252
–	Die bewahrende Sängerin in der Waldorfschule.....	254
–	Die Konzertsängerin in der Beobachterrolle	255
–	Der musikalische Generalist mit Hang zur Frustration.....	256
–	Der musiktherapierende Gesangslehrer	257
–	Der Mundartrockner als Motivationstrainer	258
–	Der Alleskönner verlangt viel und erreicht viel.....	259
–	Die junge Sängerin kämpft mit ihren Erwartungen	260
12.2	Zielorientierungen und Überzeugungen.....	261
–	Zielorientierung: Stimmliche Förderung versus Persönlichkeitsförderung... 262	
–	„Das kann ich nicht. Und ich will das auch gar nicht probieren.“.....	268
12.3	Erste Antworten auf Forschungsfragen.....	273
13	Horizontale Achse: Perspektive Assessment.....	277
13.1	Der Einfluss geschlechtsspezifischer Aspekte auf Assessment	277
–	Domänenspezifische Einflüsse des Geschlechts der Lehrkraft	278
–	Wahrnehmung von geschlechtsspezifischen Unterschieden bei Mädchen und Jungen.....	282
–	Der auffällige Stimmwechsel bei einem Mädchen	290
–	PPS – poor pitch singer	293
–	Koedukative und monoedukative Settings	297
–	Die Relevanz geschlechtsspezifischer Einflüsse auf Assessment	301
13.2	Ganzheitliches Assessment als Perspektive	302
–	Analyse der Assessment-Konzeption.....	304
–	Assessment of Learning	308
–	Assessment for Learning	324
–	Assessment as Learning	335
–	Die Relevanz von Assessment im Singunterricht.....	346
13.3	Verknüpfung der verschiedenen Perspektiven	349
–	Entwicklung der Fragestellung und inhaltliche Fokussierung.....	349
–	Etappen des Forschungsprozesses	351
–	Beantwortung der Forschungsfragen	353
–	Ausblick.....	365

E	Ausblick	367
14	Thesen: Assessment im Singunterricht	368
14.1	Wissen zur adoleszenten Stimmentwicklung ist ein wichtiger Faktor in der Begleitung von Jugendlichen	368
14.2	Mit der Umsetzung eines ganzheitlichen Assessment-Modells gelingt individuelle Förderung im Singunterricht	372
	– Assessment for Learning	374
	– Assessment as Learning	376
	– Assessment of Learning	378
14.3	Formative und summative Förder- und Beurteilungsprozesse können mit einem Stimmportfolio adäquat abgebildet werden	380
15	Weiterführende Fragestellungen	386
15.1	Relevanz der adoleszenten Stimmentwicklung im Singunterricht	387
15.2	Kernpraktiken für ganzheitliches Assessment	389
15.3	Gestaltung von Assessment im Singunterricht	391
F	Literaturverzeichnis	393
G	Anhang	431

ABBILDUNGEN

ABBILDUNG 1:	Abgrenzung von Pubertät, Adoleszenz und Jugend.....	92
ABBILDUNG 2:	Abdruck einer Passage zum Stimmwechsel von Mädchen und Jungen (Tosi & Agricola, 1757, S. 149)	102
ABBILDUNG 3:	Ambitus und Tessitura der adoleszenten Stimme nach Welch (2019).....	109
ABBILDUNG 4:	Geschlechtsspezifische Perspektiven im Musikunterricht in koedukativen und monoedukativen Unterrichtssettings.....	118
ABBILDUNG 5:	Assessment-Pyramide nach Earl (2013, S. 43)	140
ABBILDUNG 6:	International principles of assessment in music education (Brophy & Fautley, 2017).....	142
ABBILDUNG 7:	Ablaufmodell induktiver Kategorienbildung und deduktiver Kategorienanwendung (Mayring & Brunner, 2006, zitiert nach Mayring & Fenzl, 2019, S. 640)	191
ABBILDUNG 8:	Modell professioneller Handlungskompetenz – Professionswissen (Baumert & Kunter, 2006, S. 482)	198
ABBILDUNG 9:	Kommunikationsmodell nach Mayring (2015, S. 59)	200
ABBILDUNG 10:	Analyserichtungen der QIA.....	222
ABBILDUNG 11:	Verteilung der Sozialformen im Unterricht des Samples	245
ABBILDUNG 12:	Wirkung von formativem Feedback während dreier Schuljahre in der Sekundarstufe 1	305
ABBILDUNG 13:	Revidiertes Assessment-Modell.....	372
ABBILDUNG 14:	Durchlässiges Assessment-Modell als Basis eines Prozess- und Produktportfolios	380
ABBILDUNG 15:	Forschungsdiesiderate im didaktischen Dreieck (adaptiert nach W. Jank und Meyer, 2014, S. 55)	386

TABELLEN

TABELLE 1:	Liste mit Abkürzungen	19
TABELLE 2:	Phasen der adoleszenten Mädchenstimme nach Gackle (2019, S. 558–559), ergänzt mit dem Ambitus nach Gackle (2000)	97
TABELLE 3:	Phasen der adoleszenten Jungenstimme adaptiert und ergänzt nach Cooksey (2000) und Welch (2019).....	107
TABELLE 4:	Rollen der Lehrkräfte nach Wilson (1996) zitiert nach Earl (2013)	135
TABELLE 5:	Rolle der Lehrkraft in Assessment for, as und of learning nach Earl (2013) sowie Earl und Katz (2006)	137
TABELLE 6:	The Singing Voice Development Measure (©2016 Joanne Rutkowski) (Rutkowski, 2019, S. 635); adaptiert mit deutschen Notennamen	149
TABELLE 7:	Rating Scale von Wise und Sloboda (2008).....	150
TABELLE 8:	Rubric zu Gesangsvortrag nach Holcomb (2019, S. 385)	152
TABELLE 9:	Skala J. Hasselhorn (2015, S. 82)	164
TABELLE 10:	Zahlen zur Sekundarstufe 1 im Kanton Aargau	173
TABELLE 11:	Zahlen zur Sekundarstufe 1 im Kanton Solothurn.....	174
TABELLE 12:	Populäre Referenzlevel von Cohens Kappa (Nurjannah & Siwi, 2017; Xie, 2013).....	204
TABELLE 13:	Stichprobenplan mit der Anzahl der geplanten und durchgeführten Interviews	212
TABELLE 14:	Liste der Interviews mit demografischen Angaben, Erfahrung in Jahren, Informationen zum Leistungsniveau und zum Milieu	213
TABELLE 15:	Reorganisation der Studienteilnehmenden: musikalische versus pädagogische Ausbildung.....	216
TABELLE 16:	Ausgewählte Interviews für die Interrater-Reliabilitätsprüfung	228
TABELLE 17:	Test A – Kappa κ_n in Abhängigkeit der prozentualen Überlappung von Kodiereinheiten	229
TABELLE 18:	Test B – Kappa κ_n in Abhängigkeit der prozentualen Überlappung von Kodiereinheiten	230
TABELLE 19:	Kodierhäufigkeiten der drei Analyserichtungen der QIA.....	235

TABELLE 20:	Kodierhäufigkeiten der Subkategorien von Assessment.....	236
TABELLE 21:	Kodierhäufigkeiten von Assessment for learning	238
TABELLE 22:	Kodierhäufigkeiten von Assessment as learning	239
TABELLE 23:	Kodierhäufigkeiten von Assessment of learning	240
TABELLE 24:	Kodierhäufigkeiten von Sozialformen.....	242
TABELLE 25:	Stimmliche Förderung versus Persönlichkeitsentwicklung: p/P – pädagogisch gebildet (BA/MA), m/M – musikalisch gebildet (BA/MA). Leistungsniveaus: h – progymnasial, m – mittel und t – tief. Die Setzung erfolgte aufgrund der Zuordnungen in Tabelle 15.	263
TABELLE 26:	Widerstand (Ebene Schülerin/Schüler) versus Beziehungspflege (Ebene Lehrkraft)	269
TABELLE 27:	Eigenschaften von assessment of, for und as learning nach Earl (2013, S. 43).....	303

ABKÜRZUNGEN

In der Dissertation wird möglichst auf Abkürzungen verzichtet. Die wenigen verwendeten sind hier aufgeführt.

Abkürzung	
AAL	Assessment as learning
AFL	Assessment for learning
AOL	Assessment of learning
BA	Bachelor
CK	Content Knowledge
DBR	Design Based Research
LP21	Lehrplan 21
MA	Master
PCK	Pedagogical Content Knowlegdge
PK	Pedagogical Knowledge
PPS	poor pitch singer
QIA	Qualitative Inhaltsananlyse

TABELLE 1: Liste mit Abkürzungen

VORBEMERKUNGEN ZU GENDERSENSIBLEN SCHREIBWEISEN

Der Rat der deutschen Rechtschreibung empfiehlt bezüglich einer geschlechtergerechten Schreibweise in seiner ausführlichen Synopse unter anderem, sachlich korrekt zu sein und die Möglichkeit zur Konzentration auf die wesentlichen Sachverhalte und Kerninformationen sicherzustellen (Rat für deutsche Rechtschreibung, 2019). Untersucht wurde eine Vielzahl von Texten und Textsorten sowie die Empfehlungen einiger deutscher Universitäten und verschiedener Ämter in deutschsprachigen Ländern. Der Rat empfiehlt weiterhin die verschiedenen stilistischen und grammatikalischen Strategien, die sich etabliert haben. Dazu zählen insbesondere Paarbildungen, geschlechtsneutrale und geschlechtsabstrakte Ausdrücke und Umformulierungen.

Weitere Geschlechtsformen: Weder im Rahmen der Literaturrecherche noch in den empirischen Daten wurde die Notwendigkeit sichtbar, weitere Geschlechtsformen zu berücksichtigen. Da es sich bei der vorliegenden Studie nicht um eine spezifische Untersuchung im Bereich Gender handelt, wird darauf verzichtet, spezialisierte Schreibweisen wie den Gendergap, das Gendersternchen oder den aktuell in Gebrauch stehenden Doppelpunkt zur Inklusion nicht binärer Personen zu verwenden. Stattdessen wird der aktuellen Empfehlung des Rats der deutschen Rechtschreibung gefolgt.

Lehrerinnen und Lehrer: Im Text wird, wo immer möglich und sofern es sich aus den Interviews erschließen lässt, der Begriff Lehrerin oder Lehrer verwendet. Ist dies nicht möglich, wird entweder eine Paarbildung oder der geschlechtsneutrale Begriff Lehrkraft verwendet. In der Schweiz hat sich zusätzlich der Terminus Lehrperson etabliert. Dieser taucht im Zusammenhang mit den Interviews der Studie auf, wird aber möglichst vermieden.

Schülerinnen und Schüler: Auch bei Schülerinnen und Schülern wird, wo immer möglich, die eindeutige Form verwendet. In der Mehrzahl wird die Paarbildung Schülerinnen und Schüler sowie der geschlechtsneutrale Begriff Jugendliche benutzt oder auf die Begriffe Mädchen und Jungen zurückgegriffen.

ABSTRACT

Assessment des Singens in der Sekundarstufe 1: Eine qualitative Interviewstudie

Die Forschung zum Singen im Musikunterricht der allgemeinbildenden Schule fokussiert auf das Kindesalter. Abgesehen von einigen theoretischen Beiträgen gibt es für die Sekundarstufe 1 im deutschsprachigen Raum keine empirischen Untersuchungen. Insbesondere Diagnose, Förderung und Beurteilung während des Stimmwechsels stellt sich als Forschungsdesiderat heraus. In der vorliegenden Arbeit werden aus dem englischen Sprachraum einerseits wenig rezipierte Arbeiten zur adoleszenten Stimmentwicklung und andererseits das Konzept eines ganzheitlichen Assessments hinzugezogen, um die Perspektive einer heterogenen Stichprobe von Musiklehrern mithilfe einer qualitativen Inhaltsanalyse zu untersuchen. Der erste Schritt in der multiperspektivischen Auswertung des deduktiv und induktiv erstellten Kategoriensystems ist eine skalierende Analyse der eruierten Sozialformen. Neben den mehrheitlich frontalen Unterrichtssituationen finden sich bei mehr als der Hälfte des Samples Aufgabenstellungen für Gruppenarbeiten und insbesondere monoedukative Lehrformen. Weiter kann aufbauend auf Einzelporträts gezeigt werden, dass die in der anglofonen Literatur als wichtig erachtete vocal identity bei einem Teil des Samples intuitiv als wichtiger Teil des Singunterrichts beobachtet und gefördert wird. Zuletzt zeigt sich ein disparates Bild bezüglich der Beachtung des Stimmwechsels und formativer Aspekte von Förderung. Zwar weisen einige Einzelbeispiele auf intuitiv gewachsene formative Fördermodelle hin, insgesamt steht aber das für den schweizerischen Kontext typische Vorsingen als summatives Assessment im Zentrum. In dieser Form der Leistungsbeurteilung wird zwar die Forderung nach Transparenz erfüllt, die Ausgestaltung der Prüfung wird aber als wenig authentisch und die Bewertung als nicht valide beurteilt. Zusammenfassend wird auf der Basis dieser Erkenntnisse abschließend die Idee eines Stimmportfolios skizziert, allerdings mit dem Hinweis, dass empirisch gestützt einerseits die Relevanz der adoleszenten Stimmentwicklung für den Musikunterricht genauer zu untersuchen und andererseits ein Curriculum von Kernkompetenzen für ein ganzheitliches Assessment zu etablieren sei.

Assessment of Singing at the Lower Secondary School: A Qualitative Interview Study

Research on singing in the classroom for general music education focuses on childhood. Apart from some theoretical contributions, there are no empirical studies for the lower secondary school in German-speaking countries. In particular, diagnostic, formative, and summative assessment during the voice change turns out to be a research desideratum. This study draws on little-received work on adolescent voice development and the concept of holistic assessment, both from the English-speaking context, to analyze the perspectives of a heterogeneous sample of music teachers using qualitative content analysis. The first step in the multi-perspective analysis of the deductively and inductively created category system is a scaling analysis of the elicited social forms. In addition to predominantly frontal teaching situations, more than half of the sample exhibited a variety of tasks for group work and in particular monoeducational forms of instruction. Furthermore, individual portraits showed that singer identity, which is considered important in Anglophone literature, and the associated emphasis on personality development, is intuitively observed and encouraged as an important component of voice instruction in a portion of the sample. Finally, a disparate picture emerged regarding the attention to voice change and formative aspects of assessment. Although some individual examples point to intuitively developed formative support models, overall the focus is on the Vorsingen (audition) as a summative assessment, which is typical for the Swiss context. With this form of performance assessment, the demand for transparency is confirmed, but the design of the examination is judged to be less authentic, and the scoring is not valid. In conclusion, based on these findings, the idea of a voice portfolio is outlined, but with the indication that empirically supported, on the one hand, the relevance of adolescent voice development for music teaching should be investigated more closely and, on the other hand, a curriculum of core competencies for a holistic assessment should be established.

A EINFÜHRUNG

„Wir haben jetzt Singen.“

Diesen Satz hört man in Schweizer Schulhäusern der Sekundarstufe 1¹ immer wieder, obwohl eigentlich das Fach Musik gemeint ist. Noch deutlicher wird es im Musikunterricht, wo zu Beginn der Lektion oft eine stereotype Frage in den Raum geworfen wird:

„Singen wir heute?“

Singen wird dabei als Antipode zu allen anderen Tätigkeiten, subsummiert als Theorie, angesehen und präferiert. Was Schülerinnen und Schüler mit *Singen* meinen, soll mit einigen W-Fragen eruiert werden: *Wer* will denn im Unterricht singen? Sind es vor allem Mädchen, wie das im Rahmen von Untersuchungen zu gendersensiblen Musikunterricht vermutet wird (vgl. Gütay & Kreutz, 2010; Hess, 2018), oder doch auch die Jungen? *Was* soll gesungen werden? Sind es aktuelle Hits, wie sie in gängigen musikpädagogischen Zeitschriften angeboten werden oder Schullieder, extra komponiert für Lehrmittel mit didaktischen Hintergedanken (vgl. Beispiele in Detterbeck, 2011)? *Warum* soll gesungen werden? Steht das Gemeinschaftserlebnis im Zentrum oder sind Schülerinnen und Schüler auf der Suche nach einer Erweiterung von Gesangskompetenzen? Könnte es auch sein, dass der Musikunterricht insgesamt als Gegenpol zur intellektuellen Arbeit in den anderen Fächern gesehen wird, singen für das Gemüt? Weiter kann das *Wie* in den Fokus rücken. Meinten die Schülerinnen und Schüler das Singen zu YouTube-Playbacks oder das Singen auf der Bühne mit der brillanten Begleitung der Lehrkraft am Klavier? Letztlich werden sich die verschiedenen Vorstellungen und Wünsche stark unterscheiden und die Frage drückt den allgemeinen Wunsch aus, den Musikunterricht vor allem als Singunterricht zu erleben. Daneben darf nicht unbeachtet bleiben, dass viele Musikzimmer nach wie vor *Singsaal* genannt werden, was zusätzlich dazu beiträgt, den Musikunterricht auf das Singen fokussiert zu denken.

¹ Eine ausführliche Darstellung, was mit der Sekundarstufe 1 in der Schweiz gemeint ist, befindet sich in Kapitel 6.

Mit dem Wechsel der Perspektive hin zu den Empfängern der obigen Frage, nämlich hin zu den Musiklehrkräften, stellt sich die Frage, welche Intentionen und Ziele diese in den Musikunterricht hineintragen und wie sie den beschriebenen Ansprüchen einer Schülerschaft begegnen. Sind sie zufrieden, wenn es gelingt, beim Singen eine wohltuende Atmosphäre zu schaffen, um, wie Gembris (2015) jüngere empirische Studien zitiert, intuitiv „pro-soziale Effekte gemeinsamen Musizierens auf der Basis von Synchronisierungserfahrung“ und „eine positive Wirkung auf die Emotionsregulation“ (S. 12) zu bewirken? Oder arbeiten die Lehrkräfte an intuitiv entwickelten oder durch Lehrpläne gesteuerten Zielen für das Singen? Wie fördern und beurteilen die Lehrkräfte insgesamt und speziell während der Mutationszeit? Gibt es tatsächlich die von Pezenburg und Dyllick (2018) beschriebene Tendenz zu eigenverantwortlichem und individuellem Lernen beim Singen? Auch bei den Musiklehrkräften wird eine große Vielfalt an Vorstellungen und Handlungsweisen zu beobachten sein, nicht nur, aber sicher auch bedingt durch die stark differierenden Ausbildungsgänge und einen nicht unbeträchtlichen Anteil fachfremd unterrichtender Lehrkräfte in der Sekundarstufe 1.

Auch das Musiklehrmittel, das in der Schweiz seit Jahrzehnten in den meisten Musikzimmern zu finden ist, stellt das Singen in den Mittelpunkt. Es beschränkt sich hauptsächlich auf ein Lied- und Songrepertoire, das jeweils an den Geschmack der Zeit angepasst und durch eine stark komprimierte Zusammenfassung von Musiktheorie und Musikgeschichte ergänzt wird. Die erste Ausgabe trug den Namen *MUSIK AUF DER OBERSTUFE* (Gohl & Hux, 1975; Gohl & Löhner, 1997), wurde nach mehr als 20 Jahren von *MUSIK SEKUNDARSTUFE 1* (Bolli et al., 1999) und weitere 18 Jahre später von *MUSIK AKTIV* (Fluri et al., 2017) abgelöst, inhaltlich und konzeptuell aber kaum verändert. Die deutliche Dominanz des Gesangs zeigt sich in der Struktur des Lehrmittels und dem Fehlen von Arbeitsheften für Schülerinnen und Schüler über Jahrzehnte. Die erweiterte Liedersammlung wurde erst mit der dritten Überarbeitung mit zwei Arbeitsheften (Leiprecht et al., 2017; Mattle & Oberhänsli, 2018) ergänzt. Diese beinhalten zwar einen Aufbau von musikalischen Fertigkeiten zusätzlich zum Singen von Liedern, lassen aber nach wie vor Bereiche wie z.B. Musikgeschichte außen vor. Obwohl bereits in älteren Lehrplänen die Curricula deutlich über das Singen hinausgingen (Huber, 2008) und im aktuellen Lehrplan 21 (D-EDK, 2015) ein breiteres Curriculum vorgegeben und in mehreren

aktuellen Lehrmittel auch abgebildet wird (BKS Aargau, 2019), deutet die konzeptionelle Konstanz der langjährigen Lehrmittelreihe auf eine schweizerische Schulmusik-Tradition hin, welche die eingangs gestellte Frage „Singen wir heute?“ plausibilisiert.

Es mutet deswegen aus einer schweizerischen Perspektive seltsam an, wenn Hoos de Jokisch (2014) von der „Wiederentdeckung des Singens im 21. Jahrhundert“ (S. 6) spricht. Tatsächlich zeigen neben mittlerweile allgegenwärtigen Castingshows auch aktuelle Lehrmittel aus Deutschland (z. B. MusiX – Detterbeck, 2011) einen zunehmenden Anteil an gesanglichen Aktivitäten und mit SINGEN IST KLASSE (Schnitzer, 2008a, 2008b), oder GESANGSKLASSE (Bolender & Müller, 2012a, 2012b) liegen Unterrichtskonzeptionen für Chorklassen vor, die das Singen in den Mittelpunkt stellen. Dabei steht der Aufbau von vokalen und musikpraktischen Kompetenzen analog zur beschriebenen Lehrmittelreihe aus der Schweiz sowohl bei Schnitzer als auch bei Bolender und Müller im Vordergrund und folgt im Großen und Ganzen dem musikdidaktischen Konzept des aufbauenden Musikunterrichts nach W. Jank (2013).

Auch B. Jank (2013) widerspricht Hoos de Jokisch indirekt, wenn sie die unterschiedlichen Singtraditionen im Osten und Westen von Deutschland vor der Wende vergleicht und für das Singen eine Einbettung in die Traditionen „der Schule, der Region und des Landes“ (S. 63) verlangt. So vermutet sie nicht nur in Polen und Österreich „durch ein starkes regional-kulturelles Traditionsbewusstsein ein anderes und intensiveres Singen in den Schulen“ (S. 64), sondern sieht auch in den neuen Bundesländern über die Wende hinaus das Singen als „wichtigste musikpraktische Tätigkeit im Musikunterricht“ (S. 64). Auch wenn sie die Schweiz in ihrem Text nicht erwähnt, folgt das Singen auch in der Schweiz einer langjährigen Tradition. Einige aktuelle Untersuchungen liegen vor und werden später referiert (vgl. Huber et al., 2021).

Parallel dazu tragen vermehrt aktuelle theoretische und empirische Beiträge zu einer Fundierung des Singens bei, können aber gleichzeitig auch als Reaktion auf eine lange Durststrecke betrachtet werden, nachdem H.-C. Schmidt (1979) Ende der 1970-Jahre dem Singen in der Schule ein schlechtes Zeugnis ausgestellt hat. Er weist auf das Fehlen eines allgemeinen Liedrepertoires hin und beschreibt dies als „Resultat einer vehementen lieddidaktischen Krise“ (S. 112). Der Autor verweist dabei auf Adorno, der sich kritisch über das Singen geäußert hat: „Nirgends steht geschrieben, daß

Singen not sei.“ (Adorno, 1991) Neben der viel zitierten Kritik Adornos wird auch die Vereinnahmung des Singens im Dritten Reich als Grund für den Verlust der Selbstverständlichkeit des Singens (Niessen, 2008, S. 45) angesehen. Es ist somit aus einer historischen Perspektive nachvollziehbar, wenn aktuell für das Singen vermehrt theoretische und empirische Fundierungen erarbeitet werden.

Antwerpen (2014) beleuchtet z. B. die ästhetischen Bildungspotenziale des Singens und entwickelt auf der Basis von Experteninterviews eine Rahmenkonzeption für eine schulische Gesangspädagogik. Dies kann als normatives Votum zugunsten einer verbesserten Singpraxis gesehen werden. Neus (2017) tritt einen Schritt zurück und diskutiert die Frage, was unter dem Begriff Singen zu verstehen sei. Sie definiert Singen als Modus von Musik und als anthropologische Konstante, wobei physische, psychophysische, individuelle und soziokulturelle Parameter sich gegenseitig beeinflussen (S. 76 f.). Das von ihr entwickelte Perspektivraster zeigt auf, dass „[d]as Singen [...] qua Natur an den Menschen als Kultursubjekt gebunden“ ist und „notwendigerweise und alternativlos im Zusammenhang mit dem einzelnen singenden Menschen betrachtet werden“ (S. 215) muss. Der Beitrag erarbeitet Begründungsperspektiven für das Singen, die für die Untersuchung von Singsituationen als theoretische Grundlage dienen können. Dyllick (2011, 2019) wiederum beschäftigt sich seit einiger Zeit auf einer theoretischen Ebene mit dem didaktischen Potenzial einer schulischen Vokalpraxis, indem sie aus einer konstruktivistischen Perspektive auf Basis von Kersten Reichs Modell der didaktischen Handlungsebenen (Dyllick, 2019, S. 69 ff.) den Einsatz des Singens in verschiedenen Bereichen des Musikunterrichts begründet. Alle drei Beiträge bieten auf ihre Weise Grundlagen, sich dem Singen in der Schule vermehrt auf empirische Weise zu nähern.

Die Frage „Singen wir heute?“ hat für die vorliegende Arbeit den Nährboden geliefert, einen spezifischen Blick auf den Singunterricht der Sekundarstufe I zu richten. Die eingangs geschilderten denkbaren Betrachtungsweisen von Lernenden und Lehrenden auf den Lerngegenstand Singen bilden die drei Unterrichtskomponenten des didaktischen Dreiecks nach Heimann (Bönsch, 2006, S. 18). In der Logik des Modells werden insbesondere auch die Beziehungen zwischen den Komponenten sichtbar, was mit Blick auf obigen Sprachduktus nach weiteren W-Fragen verlangt. Wie schafft es die Lehrkraft, Schülerinnen und Schüler zum Singen zu

bringen, sie zu fördern und entsprechende vokale Kompetenzen zu entwickeln? Welche Einflüsse des Lerngegenstands Singen manifestieren sich in der Phase der Pubertät bei den Jugendlichen? Und wie reagieren diese in ihrer Mutation auf die Anforderungen der Lehrkräfte in ihrem physischen, psycho-physischen, individuellen und soziokulturellen Kontext, um nochmals Neus (2017) zu zitieren? Und nicht zuletzt stellt sich die Frage, welchen Bezug die Lehrkraft mit ihrer eigenen Stimme zum Lerngegenstand Singen hat. Die Vielfalt der Betrachtungsweisen allein im Kontext der Schule zeigt die Komplexität des Themas Singen im Musikunterricht. Es stellt sich notwendigerweise die Frage, welche Perspektive für die Forschungsarbeit gewählt wird, um in der Thematik eine bearbeitbare Fragestellung generieren zu können.

Der seit einigen Jahren im deutschsprachigen Diskurs der Erziehungswissenschaft anzutreffende Begriff *Assessment* (z. B. Winter, 2015; C. Schmidt, 2020) wird als geeignete Basis betrachtet, um die genannten Beziehungen zwischen den Komponenten des didaktischen Dreiecks zu beleuchten. Spezielle Beachtung findet dabei das Modell von Earl (2003, 2013), der mit der Terminologie *assessment for learning*, *assessment as learning* und *assessment of learning* die unterschiedlichen Aspekte von Diagnose, Förderung und Beurteilung zu einem ganzheitlichen Konzept zusammenfasst². Die Perspektive der Forschungsarbeit wird so auf die Kerntätigkeit der Lehrkraft gelegt, auf das Unterrichten, auf die Gestaltung von Lernprozessen, auf die Beurteilungspraxis, welche gerade in der Schweiz mit dem *Vorsingen* eine langjährige Praxis pflegt.

Die folgenden Kapitel grenzen das Forschungsinteresse ein, begründen die Themen, die vorbereitend dargestellt werden, beschreiben den Untersuchungsgegenstand, nehmen Bezug auf die Auswertungsmethode und bieten damit einen Überblick über den argumentativen Aufbau der Studie.

² Gelegentlich wird in Titeln auch *assessment through learning* verwendet, allerdings nicht als eigene Begrifflichkeit, sondern in der Meinung, *Assessment* fördere die Lernmotivation von Schülerinnen und Schülern (z. B. McGinnis, 2018) oder *Assessment* begünstige den Lernprozess der Lehrkräfte (Driscoll, 1998).

1 Relevante Kontexte des Forschungsfeldes

Der Ausgangspunkt für die vorliegende Forschungsarbeit ist das Singen im Musikunterricht in der Sekundarstufe 1. In der Einleitung wurden ausgehend vom didaktischen Dreieck nach Heimann (Bönsch, 2006) mögliche Blickwinkel auf den Gegenstand Singen vorgestellt und mit der gewählten Assessment-Brille zusätzlich auf die Interaktion zwischen Lehrkräften und Lernenden fokussiert. Einleitend zur Gliederung der Forschungsarbeit (vgl. Kapitel 2) werden in diesem Kapitel drei Themenbereiche erörtert, die für die Kontextualisierung der Studie wichtig erscheinen: es sind erstens schweizerische und internationale Traditionen im schulischen Singunterricht, zweitens mögliche Komplikationen bei der Leistungsbeurteilung des Singens und drittens bestehende Forschungskontexte des Singens.

1.1 Traditionen des schulischen Singunterrichts

Singkulturen im Musikunterricht folgen regionalen und nationalen Gepflogenheiten (Niessen, 2008; Ashley, 2015; B. Jank, 2017; Schaumberger, 2020a; Schaumberger et al., 2020). In diesem Abschnitt wird deshalb überlegt, ob eine schweizerische Tradition auszumachen ist (Blanchard, 2021; Marty, 2021) und wie diese in einen historischen und internationalen Kontext eingeordnet werden kann.

Einleitend wird an dieser Stelle definiert, wie die Begriffe Singunterricht bzw. Singen verwendet werden. Die vorliegende Arbeit bewegt sich im Kontext des obligatorischen Schulunterrichts der Schweiz. Die vielfältigen Aktivitäten des Singens finden im Rahmen des regulären Musikunterrichts statt (vgl. dazu auch 6.1) und werden in dieser Arbeit meist zusammenfassend Singen genannt, egal in welcher Gruppenkonstellation diese stattfinden. Gemeint ist z. B. das gemeinsame Singen in der Klasse oder eine Singsequenz in Gruppen, aber auch eine mögliche Einzelbetreuung durch die Lehrkraft. Singunterricht ist verkürzend als Singen im Musikunterricht zu verstehen und meint nicht den teilweise stattfindenden freiwilligen Chorunterricht. Obwohl der untersuchte Unterricht einzelner Lehrkräfte an das Konzept von Chorklassen erinnert, ist dieses Konzept in

der Schweiz nicht als solches etabliert, sondern das Singen zeigt sich nach wie vor als zentraler Gegenstand des Musikunterrichts, wie nachfolgend dargelegt wird. Die genannten Begriffe Singen und Singunterricht grenzen sich von Gesangsunterricht ab, der als außerschulischer Einzelunterricht verstanden wird. Die von B. Jank (2013, 2017) geprägte qualitative Unterscheidung zwischen Singen und Gesang wird in dieser Arbeit später thematisiert (vgl. 4.1).

Kälin (1976) untersucht in der Schweiz Mitte der 1970er-Jahre den Musikunterricht ausgehend von den damals aktuellen Lehrplänen der verschiedenen Kantone und beschreibt eine musikpädagogische Tradition, die hauptsächlich auf das Lied setzt. Im Gegensatz zum Primarschulbereich, wo die Fachbezeichnung „Singen“, „Chant“³, „Gesang“ oder „Gesangsunterricht“ (S. 460) lautet, wird im Bereich der Sekundarstufe 1 das Fach Musik bereits vielerorts „Gesang und Musik“ (S. 460) genannt. In allen 26 Lehrplänen wird Klassengesang genannt, in 4 Lehrplänen ergänzend Einzelgesang. Zusätzlich wird in 15 Lehrplänen gefordert, dass Lieder täglich auch in anderen Fächern gesungen werden sollen (S. 460 ff.). Kälin führt die bestehenden Eckwerte direkt auf die Zielvorstellungen von Kretschmar und Kestenberg zurück:

„Zwar ist bei Kretschmar schon von Musik die Rede, und Kestenberg benennt den Gesangsunterricht⁴ um, doch dominiert noch immer die Übermittlung eines wertvollen Liedgutes. Damit wird das Lied auch zum Übungsfeld beinahe der gesamten musikalischen Unterweisung. Zudem hat alles Praktische in der Musikerziehung den absoluten Vorrang gegenüber dem Theoretischen. Auch soll das Theoretische grundsätzlich im Zusammenhang mit der praktischen Liedarbeit stehen. Vom Lied aus führt der Weg zum musikalischen Kunstwerk.“ (S. 467)

Kälin bemängelt anschließend den jahrzehntelangen Stillstand und verlangt nach einer Weiterentwicklung des Lehrplans, insbesondere „geh[e] es

³ Chant ist Französisch und bedeutet Gesang oder Singen. Die Untersuchung von Kälin (1976) bezog sich auf alle Landesteile der Schweiz. Der bereits genannte Lehrplan 21 (D-EDK, 2015) deckt nur die deutschsprachige Schweiz ab.

⁴ An dieser Stelle ist Singen im Schulunterricht gemeint und nicht Einzelunterricht.

um den Abbau der Dominanz des Liedes und um bessere Berücksichtigung des Werkhörens“ (S. 469). Marty (2021) belegt die damalige Dominanz des Singens mit einer Umfrage aus dem Jahr 1977, die festhält, dass „mehr als die Hälfte der Zeit mit Singen im Klassenverband verbracht wurde [...]“ (S. 82). Setzt man diese Aussagen in Bezug zu der von H.-C. Schmidt (1979) benannten und in der Einleitung bereits dargelegten lieddidaktischen Krise in Deutschland, wird klar, dass im Gegensatz zu Deutschland mindestens bis zu dieser Zeit Singen in der Schweiz eine besondere Stellung im Musikunterricht hatte. Die Analyse von Kälin ist zudem als wichtiger Beitrag zum damaligen Diskurs zur Umgestaltung des Lehrplans für das Fach Musik einzuschätzen. Wie Marty (2021) beschreibt, konnte in diesen Jahren durch eine koordinierte Aktion des SCHWEIZERISCHEN KOMITEES ZUR FÖRDERUNG DER SCHULMUSIK (SKSF) den Kantonen ein Lehrplanentwurf vorgelegt werden, der in den nachfolgenden Jahren meist „als Vorbild oder Anregung bei der Revision“ (S. 83) der kantonalen Lehrpläne diente. Prominent war dabei der gesetzte Titel *Musikunterricht*, der mehrheitlich zur neuen Fachbezeichnung Musik führte und insbesondere die thematische Öffnung des Unterrichts demonstrierte. Marty belegt in seiner Untersuchung zu schulmusikalischen Diskursen, dass trotzdem während einer länger andauernden Phase bis über die Jahrtausendwende hinaus die Bedeutung des Singens zwar kontrovers diskutiert, aber trotzdem überwiegend „als zentraler und nicht als paritätischer Unterrichtsgegenstand betrachtet wurde“ (S. 85). Noch nach der Jahrtausendwende bewertet Kleinen (2006) mit seiner Außensicht die musikpädagogische Situation in der Schweiz als zurückliegend mit der Begründung, „Singen und Musizieren bestimmten den Musikunterricht nahezu vollständig“ (S. 310). Das von Kleinen eingebrachte Argument der Rückständigkeit ist als Vergleich zu Gepflogenheiten der deutschen Musikpädagogik einzuordnen, wirkt aber mit Blick auf das Aufkommen der Chorklassenarbeit bereits wieder überholt. Wie bereits in der Einleitung gezeigt wurde, bestätigt sich die Dominanz des Singens ebenfalls in einer Liedersammlung, die in mehreren Auflagen und unter verschiedenen Titeln langjährig als wichtigstes Musiklehrmittel der Schweiz vertrieben wurde. Letztlich wird auch im aktuellen Lehrplan 21 (D-EDK, 2015) mit dem eigenständigen Kompetenzbereich *Singen und Sprechen* deutlich, welcher Wert dem Singen nach wie vor beigemessen wird.

Aktuelle Einblicke in den schweizerischen Kontext der Sekundarstufe 1 vermittelt Blanchard (2019, 2021) in seinen ethnografischen Studien, die auch das Singen miteinbeziehen. Darin beschreibt er eine Unterrichtskultur, in der *Musizieren* – und damit meint er mehrheitlich Singen – im Zentrum steht, als gemeinsame Tätigkeit orientiert an einem sich entwickelnden Song-Repertoire. *Nichtmusizieren* wird als Gegensatz davon abgegrenzt und umfasst Lerntätigkeiten mit individuellem Lernzuwachs, sei dies musiktheoretisches Wissen oder instrumentales Knowhow. Obwohl das bearbeitete Sample mit fünf Lehrkräften nicht als repräsentativ für die Schweiz erachtet werden darf, zeigt sich mit diesem empirischen Beleg die nach wie vor bestehende Dominanz des Singens. Zusätzlich stellt sich heraus, dass Zweifel an der Qualität des Singens angebracht sind, insbesondere Singen weder als Zentrum des musikalischen Lernens noch als eigener Lerngegenstand betrachtet wird, sondern lediglich als Aktivierungsstrategie dient. Er fasst wie folgt zusammen:

„Dass man singt, ist eine unhintergehbare Setzung des schulischen Musikunterrichts. Auf dieser Basis kann die Frage verhandelt werden, was man singt. Wie man singt, ist indes nicht nur nicht verhandelbar, sondern scheint gar keine Frage zu sein.“ (2019, S. 287, Hervorhebung im Original)

Zusammenfassend kann also gefolgert werden, dass Singen zwar immer noch eine gewichtige Rolle spielt, aber als Lerngegenstand zumindest im vorliegenden Sample in seiner Qualität hinterfragt werden muss. Auf jeden Fall scheint die von Kälin (1976) Jahrzehnte früher hervorgehobene Vermittlung eines „wertvollen Liedgutes“ (S. 467) nicht mehr im Zentrum zu stehen.

Der Blick über die Landesgrenzen hinaus zeigt, dass zumindest im ehemaligen Osten von Deutschland Singen im schulischen Musikunterricht durchgehend gepflegt wurde (Niessen, 2008; B. Jank, 2013) und die Aussage zur lieddidaktischen Krise (H.-C. Schmidt, 1979) hauptsächlich auf den westlichen Teil des Landes zu beziehen ist. Schaumberger (2020a) bestätigt auch für Österreich die prominente Rolle des Singens, welche sich sogar im Namen des Lehrplans der Vorstufe mit dem sprechenden Namen *Singen und Musizieren* widerspiegelt. Die Dominanz wird auch auf weitere Staaten wie England und Italien (Schaumberger et al., 2020) bezogen. Zwar wird an gleicher Stelle von Initiativen einiger Länder zur Stärkung

des Singens gesprochen, z. B. in England durch die Initiative *Sing Up*, aber gleichzeitig wird gerade dort der Zugang zum Singen als kritisch erachtet. Zusätzlich beschreibt Ashley (2015) für England eine Diskrepanz zwischen der hohen Qualität der „English cathedral tradition“ (S. 70) und den teilweise bescheidenen Ergebnissen im Musikunterricht.

Im Gegensatz zu den genannten Ländern, wo Singen entweder eine wichtige Rolle spielt oder wo Initiativen bestehen, dem Singen wieder einen größeren Stellenwert zu geben, präsentiert sich in Schweden eine andere Situation. Johnson (2020) beschreibt einen drastischen Rückgang der Bedeutung des Singens in den schwedischen Lehrplänen zwischen 1955–1980. Als Gründe werden ähnlich wie in Deutschland Nationalismus und zusätzlich das Verschwinden sozialer Kontexte wie z. B. Arbeiterbewegungen genannt, die laut Johnson erst wieder kreierte werden müssten, um dem Singen zuerst in der Gesellschaft und dann auch im schulischen Musikunterricht einen wichtigeren Stellenwert zu geben. Abschließend sei der amerikanische Kontext genannt, in dem Singen in der Sekundarstufe traditionsgemäß in Junior High/Middle-School Choirs (z. B. Robinson & Gesler, 2007) stattfindet. Schülerinnen und Schüler wählen dort aus, ob sie sich Chor, Band oder Orchester anschließen, was für das Singen den Grad an Freiwilligkeit zumindest erhöht und deshalb Vergleiche mit der hiesigen Situation erschwert.

Das Singen scheint also nicht nur in der Schweiz eine durchgehende Tradition zu haben, sondern auch in anderen Regionen und Ländern. Zusätzlich wird dem Singen vermehrt Aufmerksamkeit geschenkt (Lehmann-Wermser, 2008; Hoos de Jokisch, 2014), was verschiedene Singinitiativen und Forschungskontexte (vgl. 1.3) zeigen. Eine weitere schweizerische Eigenheit ist das Vorsingen, das in der Schweiz vielerorts gepflegt wird. Schaumberger (2020a) problematisiert die summative Beurteilung folgendermaßen: „In der musikpädagogischen Literatur sowie in persönlichen Schilderungen werden immer wieder Situationen im Musikunterricht beschrieben, in denen Schülerinnen und Schüler einzeln vorsingen mussten, beurteilt wurden, und gegebenenfalls aufgrund ihrer Singleistung eine schlechtere Note im Zeugnis bekommen haben.“ (S. 223) Wieso dieser Moment des Vorsingens problematisch ist, wird im nächsten Abschnitt untersucht.

1.2 Assessment des Singens als Problemfall

Die Leistungsmessung des Singens bzw. summative Beurteilung wird kontrovers beurteilt. Laut Hammar (2016) findet die Tätigkeit Singen immer als innige Beteiligung der Person statt und die Beurteilung betrifft so auch immer die Person als Ganzes⁵. Im vorliegenden Abschnitt wird diese Problematik beleuchtet und zugunsten einer ganzheitlichen Assessment-Perspektive geöffnet.

„Eine schöne Stimme ist ein Glücksfaktor. Ein Stimmklang, der uns berührt, erzeugt Sympathie, löst angenehme Gefühle aus.“ (Mahlert, 2017) Diesen Satz stellt Mahler an den Anfang seines Aufsatzes *GLÜCK IM MUSIZIEREN* und gibt damit der emotionalen Aufladung des Singens und dem zugehörigen Kommunikationsaspekt ein starkes Gewicht. Gleichzeitig wird damit eine wesentliche Diskussion rund um das gemeinsame Singen im Musikunterricht und im Hinblick auf Förderung und Beurteilung eröffnet, denn das Zitat von Mahler erzeugt für den Kontext Schule einige Komplikationen: Wenn eine schöne Stimme berührt und Sympathien erzeugt, dann stellt sich die Frage, was mit allen anderen Stimmen ist. Gibt es im Umkehrschluss *unschöne* Stimmen und haben Schülerinnen und Schüler mit *keiner* schönen Stimme *kein* Glück? Was passiert in einer Schulklasse, wenn keine Sympathien erzeugt werden, weil eine Stimme nicht berührt, sondern auffällt durch wenig Kontur in der Melodie oder fehlendes Rhythmusgefühl und vielleicht in schwierigen gruppendynamischen Situationen untergeht? Diese Fragestellungen führen direkt zur Problematik der Leistungsbeurteilung im schulischen Singunterricht, wo nicht nur die mögliche Expertenmeinung einer Lehrkraft, sondern auch Schülerinnen und Schüler stimmliche Leistungen wahrnehmen und bewusst oder unbewusst beurteilen. Der Blick in den Musikunterricht ist der Blick

⁵ Eine prägende Konfrontation diesbezüglich gab es im Anschluss an eine Präsentation des Verfassers, bei der die schweizerische Tradition des Vorsingens und die damit verbundene Notengebung als summatives Beurteilungsinstrument erläutert wurde. Die Reaktion einiger Teilnehmenden war von großem Unverständnis geprägt, es wurde von Körperverletzung gesprochen und davon, dass eine solche Form der Leistungsbeurteilung seit den 1970er-Jahren kein Thema mehr sei. Die vorliegende Arbeit leistet einen Beitrag, dieser ablehnenden Haltung produktiv zu begegnen.

in eine soziale Gemeinschaft, wo Sympathien zum Tragen kommen, wo vielleicht schöne Stimmen berühren, aber auch untrainierte Stimmen zu hören sind. Die Stimme als starkes Persönlichkeitsmerkmal (Neus, 2017) wird im Miteinander permanent manifest. Die Beschaffenheit des Singens bezüglich Körperlichkeit, Persönlichkeit und Emotionalität fasst Hoos de Jokisch (2014) mit dem Satz „[i]ch singe, (al)so bin ich“ (S. 6) zusammen. Das bedeutet, dass jede individuelle Leistungsmessung im Singunterricht, sei dies innerhalb eines Gruppensettings oder auch im stillen Kämmerlein⁶, zu einer Einschätzung der Person wird. Diese Problematik gilt es zu überwinden.

Singen kann also nicht losgelöst von der Person auf musikalische Fähigkeiten reduziert werden, sondern bleibt in seiner Ganzheitlichkeit komplex. Sammler (2017) unterscheidet den Aspekt der tatsächlichen Singfähigkeit und der persönlichen Überzeugung zur eigenen Singfähigkeit. Er beschreibt bezüglich der natürlichen Entwicklung bei Kindern, dass „bereits Zweijährige kurze Melodien und Rhythmen durchaus erkennbar nachsingen und Dreijährige beginnen, ihre z. T. ganz eigenen Melodien zu kreieren“ (S. 187 f.). Dalla Bella et al. (2007) haben in ihrer Studie zur Singfähigkeit Erwachsener zeigen können, dass abgesehen von zwei Personen mit selbst deklarerter *tone deafness*⁷ alle Erwachsenen aus einer Gruppe von 65 Personen in der Lage waren, ein einfaches Lied aus der Erinnerung tonal und rhythmisch korrekt zu singen, in langsamem Tempo in gleicher Qualität wie professionell Singende. Für die vorliegende Studie kann deshalb davon ausgegangen werden, dass Schülerinnen und Schüler mehrheitlich in der Lage sind, an einem Singunterricht erfolgreich teilzunehmen bzw. stimmliche Leistung zu erbringen. J. Hasselhorn (2015) zeigt zudem, dass die valide Messung musikpraktischer Fähigkeiten möglich ist, beschränkt sich dabei mit seinen Skalen aber auf rein melodische (und rhythmische) Singfähigkeiten.

⁶ Musiklehrkräfte sind z. T. der Meinung, bloßstellende Momente bei der Beurteilung könnten vermieden werden, wenn die Schülerinnen und Schüler einzeln und abgesondert beurteilt werden. Siehe z. B. Herr Brugger (01) in Kapitel 12.1.

⁷ In der Studie wird bei den beiden Personen aufgrund üblicher Rezeptionsfähigkeiten lediglich ein „pure output disorder“ (Dalla Bella et al., 2007, S. 1188) deklariert.

Die Fähigkeit zu singen scheint also ein messbares anthropologisches Merkmal zu sein (vgl. 4.1). Die Überzeugung, nicht singen zu können, scheint aber bei vielen Erwachsenen und bereits bei Kindern verbreitet zu sein. So werden negative Singerlebnisse im Rahmen von Musikunterricht und allgemein Momente der Scham oft in eröffnenden Kapiteln von Studien und Artikeln zum Thema Singen angeführt (Brunner, 2014; Tellisch, 2015; Schaumberger, 2020a). An dieser Stelle sei nochmals beispielhaft auf Adorno (1991) verwiesen, der selbst einen solchen emotionalen Moment in seiner Kinder- und Jugendzeit beschreibt: „Ich erinnere mich deutlich, wie peinlich es mir war, wenn meine Mutter und deren Schwester – beide Sängerinnen von Beruf – etwa auf Bitten meines Vaters im Wald ‚O Täler weit, o Höhen‘⁸ anstimmten.“ (S. 75) Adorno benutzt diese emotional geprägte Aussage als Ausgangspunkt in seiner Argumentation gegen das Singen, die eigene Scham ins Zentrum stellend, obwohl er nicht selbst singen musste, sondern sich lediglich für seine Mutter und ihre Schwester schämte. Legrand (2018) beschreibt ganz ähnlich ein mediales Problem des Singens in heutiger Zeit: „Kinder erfahren sehr früh die Diskrepanz zwischen eigenem Stimmausdruck und der Stimme aus dem Lautsprecher. [...] Diese Diskrepanz zu den eigenen Vorbildern vermöge, Schüler⁹ zu demotivieren.“ (S. 97) Adorno entwickelt seine Ablehnung zum Singen ähnlich wie viele Jugendliche heute, welche ihr eigenes Singen mit der immer perfekten Gesangsstimme aus dem Radio vergleichen.

Beide Beispiele zeigen, wie prägend Singerlebnisse persönlich sein können. Ashley (2015) beschreibt für englische Verhältnisse, dass viele Jungen bereits im Schuljahr sechs und sieben eine *non-singer identity* ausbilden, aus Gründen, die später (vgl. 4.4) ausgeführt werden. Diese inneren Überzeugungen beschreibt Spychiger (2013) mit dem Konstrukt des

⁸ Beginn des Liedtextes von Abschied vom Walde, Op. 59, No. 3 (Felix Mendelssohn/Joseph von Eichendorff).

⁹ Hier ist vom generischen Maskulin auszugehen. Also nicht nur Schüler, sondern auch Schülerinnen sind mit den Vorbildern aus den Medien konfrontiert und können deswegen demotiviert werden. Dies ist deshalb wichtig, weil laut Hess (2015) Jungen dem Singen weniger abgewinnen können als Mädchen und sich deshalb im vorliegenden Text fälschlicherweise darauf schließen ließe, deren Vorbilder seien kein Problem.

musikalischen Selbstkonzepts¹⁰, das in einer Dimension die Überzeugung zu den eigenen musikalischen Fähigkeiten misst, einschließlich des Singens. Zusätzlich zeigten Fiedler und J. Hasselhorn (2020) eine Korrelation mit der Motivation im Musikunterricht. Es gilt deshalb im Singunterricht nicht nur vokale Fähigkeiten zu beachten, sondern auch Überzeugungen miteinzubeziehen, um an den Lerngegenstand Singen heranzukommen und eine Entwicklung zu ermöglichen. Bleibt die Beurteilung des Singens lediglich bei einer punktuellen summativen Beurteilung stehen, werden gerade Schülerinnen und Schüler mit einem negativen Selbstkonzept vor allem in ihren Überzeugungen bestätigt, aber kaum dazu angehalten, an ihrer vokalen Entwicklung zu arbeiten.

In der vorliegenden Arbeit wird aufgrund der Fokussierung auf die Sekundarstufe 1 die beschriebene Herausforderung von Förderung und Beurteilung zusätzlich durch die biologischen Prozesse der Stimmentwicklung während der Pubertät erschwert. Neben der körperlichen Ebene gilt es zusätzlich entwicklungspsychologische Aspekte zu berücksichtigen, insbesondere mögliche Scham und Unsicherheit beim Singen. Die Individuen treffen zudem in der Schule beim Singen aufeinander, was zusätzlich zu einer labilen sozialen Situation führt (vgl. 4.5). Die verschiedenen Aspekte zeigen, wie wichtig ein holistischer und über das gesangstechnische Knowhow hinausgehender Ansatz ist. Die spezifische Herausforderung der Mutation ist auch der Grund, weshalb ein eigenes theoretisches Kapitel zur Stimmentwicklung in der Pubertät in die Arbeit aufgenommen wurde (vgl. Kapitel 4).

Mögliche Lösungsansätze beschreibt Schaumberger (2020a), betont aber gleichzeitig, „dass bei diesem Thema im deutschsprachigen Raum Neuland betreten wird“ (S. 226). Er referiert als mögliche Ausgangspunkte für die Entwicklung von Beurteilungskonzepten ebenfalls (vgl. Einführung zu Bereich A) die aus dem anglofonen Sprachraum kommenden Ideen ganzheitlichen Assessments und nimmt zusätzlich Bezug auf die Schulchöre in den USA, in denen bereits langjährige Bewertungstraditionen vorliegen (vgl. 5.2). Er listet am Ende seines Textes eine ganze Reihe von möglichen Vorkehrungen auf, um ein Beurteilungskonzept für

¹⁰ Die Vorstellung des Konstrukts basiert auf einem Forschungsbericht von Spychiger, Obertz und Gruber (2010) und wurde für die Sekundarstufe 1 von Fiedler (2016) bestätigt.

das Singen zu etablieren. Diese werden später in dieser Arbeit nochmals aufgenommen. Ein wenig enthusiastisch und mit einem normativen Impetus schreibt er: „Für ein solches Konzept des Kompetenzaufbaus im Singen und eine daran gekoppelte Leistungsbeurteilung, die ohne die eingangs zitierten bloßstellenden Momente auskommt, lohnt es sich mutig einzutreten.“ (S. 245) Das von Earl (2013) vorgeschlagene ganzheitliche Assessment-Konzept, das diagnostische, formative, summative sowie selbstevaluative Assessmentsschritte beinhaltet, erweitert das Setting von punktuellen Leistungsmessungen hin zu einem Konzept von *Förderung und Beurteilung*, in dem zusätzlich zu den stimmtechnischen Zielen auch negative Selbstkonzepte oder die Stimmentwicklung in der Pubertät auf einer breiten Ebene angesprochen werden können. Ziel der vorliegenden Arbeit ist es nun, anhand empirischer Daten zu prüfen, inwiefern sich solche ganzheitlichen Praktiken im Feld bereits etabliert haben. Ob sich in der musikpädagogischen Forschung bereits Hinweise zeigen und an welche Forschungsergebnisse angeknüpft werden kann, wird nun im letzten Abschnitt dieses Kapitels untersucht.

1.3 Singen im Fokus musikpädagogischer Forschung

Im ersten Abschnitt wurde herausgearbeitet, dass Singen zumindest in der Schweiz nach wie vor im Zentrum steht, aber Zweifel an der aktuellen Ausrichtung des Lerngegenstands angebracht sind (Blanchard, 2019). Im zweiten Abschnitt wurde argumentiert, dass Förder- und Beurteilungsprozesse die Gestaltung des Singunterrichts immer erschweren, weil der Lerngegenstand und die Person insbesondere während der Pubertät nicht voneinander getrennt werden können. Es gilt nun in diesem Abschnitt herauszufinden, inwiefern bereits Konzepte und didaktische Modelle existieren, die das komplexe Zusammenspiel in Förder- und Beurteilungsprozessen abbilden und über reine Singinitiativen wie *Primacanta* (Graefe-Hessler & Jank, 2015), *Belcantare* (B. Jank & Buschmann, 2013) oder *Singen macht Sinn* (Forge & Gembris, 2012) hinausgehen.

Auf der einen Seite zeigt sich ein Forschungsschwerpunkt zur Kinder- und Jugendstimme (z. B. Fuchs, 2019a) rund um das Leipziger Symposium

für Kinder- und Jugendstimme¹¹ und im Primarschulbereich gibt es einige Arbeiten zur Stimmdiagnose (Bojack-Weber, 2012; Greuel et al., 2007; Greuel & Horst, 2011), die berücksichtigt werden können. Für die vorliegende Arbeit mit dem Fokus auf Assessment wird zudem die Frage relevant, ob Lehrkräfte individuelle Förderung in Betracht ziehen und ob es ihnen gelingt, ihren Unterricht aufgrund von Diagnosen adaptiv zu gestalten. Das von Bolender und Müller (2012a) vorgelegte Lehrmittel Gesangsklasse sowie die Auslegeordnung von Hoene und Thurman (2011) in Umgang mit Heterogenität im Musikunterricht bieten seit einiger Zeit Vorlagen für individuelle Zugänge. Zuletzt interessiert in dieser Arbeit auch die Frage, wie die musikpädagogische Forschung das Handeln der Lehrkräfte aufgreift. In diesem Abschnitt wird deshalb ein kurzer Blick auf Projekte geworfen, die sich mit Wissensbeständen (Puffer & Hofmann, 2017) und Individualkonzepten (Niessen, 2014) von Lehrkräften befassen, um dann zum nächsten Kapitel überzuleiten, worin die Gliederung der vorliegenden Arbeit vorgestellt wird. Spezifisch für das Singen in der Sekundarstufe 1¹² im regulären Schulunterricht liegen im deutschsprachigen Raum keine Untersuchungen vor. Deutlich mehr Aufmerksamkeit erhält die mutierende Jungenstimme.

In den jährlich erscheinenden Publikationen des bereits erwähnten Leipziger Symposiums für Kinder- und Jugendstimme¹³ finden sich sowohl stimmphysiologische (Fuchs, 2010), stimpädagogische (Fuchs, 2019a, 2019b; Peterson, 2019), als auch entwicklungspsychologische Themen (Fuchs, 2018; Kalmbach, 2010). Der Phoniater Fuchs schreibt oft thematisch fokussiert auf Knabenchöre, die auf professionellem Niveau singen, was den Vergleich mit dem regulären Musikunterricht zwar befruchtet, aber z. T. auch erschwert, da der Grad an Freiwilligkeit im Gegensatz zur Schule deutlich höher ist. Die Fokussierung auf Stimmbildung und künstlerischen Ausdruck wird von Oberschmidt (2016) auch für Chorklassenmodelle bezüglich Adaptionsmöglichkeit auf einen regulären Mu-

¹¹ Webseite: <https://www.uniklinikum-leipzig.de/einrichtungen/kinderstimme> (abgerufen am 27.07.2021).

¹² Diese Aussage bezieht sich auf die 7.–9. Klasse (vgl. auch Kapitel 6 zu den schweizerischen Verhältnissen).

¹³ <https://www.logos-verlag.de/cgi-bin/engtransid?page=/Buchreihen/stimme.html> (abgerufen am 02.08.2021).

sikunterricht problematisiert. Interessanterweise ist es aber genau Fuchs (2015), der aus der Optik dieser Knabenchor-Settings einen ganzheitlichen Blick von Musiklehrkräften erwartet und sogar die Forderung stellt, dass diese eine minimale Anamnese der Stimmen von Schülerinnen und Schülern machen, also nicht nur den Klassenklang kennen, sondern auch die Klangqualität einzelner Stimmen begleiten und überwachen¹⁴.

In eine ähnliche Richtung tendiert Brunner (2014). Er interessiert sich für ganzheitliche Förderprozesse und untersucht Diagnose- und Förderstrategien von Lehrkräften der Grundschule für das Singen. Er stellt fest, dass im Singunterricht zwar „manche Bereiche offensichtlich gut ‚laufen‘ und im Abgleich mit der Literatur intuitiv richtig umgesetzt werden“ (S. 265), aber gezielte methodische Verfahren kaum zum Einsatz kommen, obwohl diese bekannt sind. Auch Lehmann-Wermser (2008) vermutet wenig zielgerichtetes Tun der Lehrkräfte beim Singen:

„Die Popularität des Singens darf allerdings nicht darüber hinweg täuschen, dass viele didaktische Fragen offen bleiben. In den derzeit gängigen Modellen bleibt Singen merkwürdig unverbunden neben anderen Inhalten stehen. Zu vermuten ist auch, dass in den subjektiven Theorien von Lehrerinnen und Lehrern darüber wenig begründete Annahmen formuliert werden, weil Singen quasi naturwüchsig zum Musikunterricht gehört.“ (S. 90)

Diese didaktische Unbedarftheit, die beide Autoren monieren, beschrieben als intuitives Handeln oder urwüchsiges Geschehen, zeigt sich wiederum in der Untersuchung von Blanchard (2019), wenn er beschreibt, dass sich Singen zwar an den Idealen einer Chorarbeit orientiert, aber die Lehrkräfte in seinem Sample lediglich das basale Ziel anstreben, „dass die Schüler_innen überhaupt singen“ (S. 263). Singen bleibt dabei, wie Lehmann-Wermser es formuliert, merkwürdig unverbunden. Abgesehen von intuitiven Zugängen macht Brunner (2014) in seiner Untersuchung keine etablierten Diagnose- und Fördermethoden aus:

¹⁴ Dieser Artikel war der Ausgangspunkt, die Entwicklung des Basischecks Stimme (Imthurn, 2020) zu starten. Dieser hat das Ziel, den diagnostischen Ansatz für formative Settings mit einem Set von Tests zu begünstigen.

„Testverfahren, die als zuverlässige Diagnoseinstrumente¹⁵ dienen könnten, um eine individuelle Förderung zu ermöglichen, haben bislang ebenfalls keinen Eingang in die Unterrichtspraxis gefunden. Dabei könnten solche Testverfahren für eine aussagekräftige Überprüfung des Lernfortschritts hilfreich sein, nicht im Sinne einer Benotung, sondern als motivationale Förderung.“ (S. 265 f.)

Brunner bewertet also das Potenzial der Diagnose ähnlich wie Fuchs mit seiner Forderung nach Anamnese als positiven Schritt in einer Weiterentwicklung des Singunterrichts und verortet Diagnose als Motivationselement im Rahmen formativen Assessments (vgl. 5.3). Auch Bojack-Weber (2012) bestätigt diesen Effekt in ihrer Untersuchung zur Singfähigkeit von Grundschülerinnen und -schülern.

Mit einem möglichen Einsatz von diagnostischen Tools stellt sich auch die Frage nach den Settings im Singunterricht. Greuel (2006) erwartet im Anschluss an diagnostische Verfahren eine „innere[] Differenzierung von Musikunterricht“ (S. 15), um die individuellen Potenziale der Stimmen adäquat zu fördern. Aus dem Kontext der referierten Artikel lässt sich aber schließen, dass Singen meist als Klassenaktivität gedacht wird, wobei gemeinsam mit der Lehrkraft im Rahmen einer frontalen Unterrichtssituation gesungen wird. In der bereits erwähnten Studie von Blanchard (2019) zeigt sich sogar eine klare Tendenz, dass Musiklehrkräfte sich beim Singen als Chorleiter sehen, und mit dem Einsingen auch den Gepflogenheiten eines Chorsettings folgen. Es zeigt sich so eine Vermischung von chordidaktischen und musikpädagogischen Ansätzen, die von Oberschmidt (2016) kritisch hinterfragt wird. Er zitiert Neus (2017)¹⁶: „Es herrscht keine Klarheit darüber, was das Singen als Musikhandlung im Musikunterricht leisten soll [...]“. (S. 200 f.) Die von Greuel erwartete innere Differenzierung zeigt sich einerseits als Querstand zum Chorsetting, das bei Lehrkräften verankert zu sein scheint, und andererseits zeigt sich große

¹⁵ Das von Brunner angesprochene Testverfahren (Greuel et al., 2007) wird später referiert.

¹⁶ Im Druckexemplar wird das Jahr 2017 angegeben, in Online-Bibliotheken teilweise das Jahr 2016. Auf jeden Fall muss Oberschmidt bereits 2016 Zugriff auf den Text gehabt haben.

Unklarheit, wie sich der Lerngegenstand Singen im Musikunterricht generell konstituieren könnte.

Die großen Singinitiativen erwarten eine allgemeine Stärkung des Sings, messen den Erfolg an stimmphysiologischen Merkmalen wie der Vergrößerung des Ambitus und einer längeren Phonationsdauer (z. B. Gütay & Kreutz, 2010), eine Tendenz zur Individualisierung kann aber nicht erkannt werden. Das gemeinsame Singen steht in den genannten Initiativen trotz anderer Implikation im Titel an erster Stelle. Graefe-Hessler und B. Jank (2015) schreiben: „*Primacanta – Jedem Kind seine Stimme* zielt darauf, jedem Grundschulkind den lust- und qualitätsvollen Gebrauch seiner Singstimme zu ermöglichen.“ (S. 219, Hervorhebung im Original) Obwohl sich im Bericht zeigt, dass jedes Kind profitieren konnte und auch individuelle Potenziale sichtbar wurden, wird bei dieser Singinitiative nicht erwähnt, dass individuelle Bedürfnisse erfasst und bearbeitet wurden, im Gegenteil war es Teil des Konzepts, die Lieder gemeinsam in Großgruppen zu lernen. Die Vermutung liegt also nahe, Singen als *Singen in der Gruppe* zu denken. Auch Pezenburg und Dyllick (2018) vermuten beim Singen eine traditionell frontale Unterrichtssituation, sehen aber Tendenzen zur Individualisierung. Sie schreiben in ihrem Aufsatz SINGEN ALS DIDAKTISCHES HANDLUNGSFELD am Ende:

„Daneben wird auch zunehmend nach Möglichkeiten gesucht, die im Bereich des Sings traditionell häufig frontalen Unterrichtssituationen aufzubrechen, das Singen stärker zu kontextualisieren und dabei – im Sinne moderner Lerntheorien – verstärkt eigenverantwortliches und individuelles Lernen zu ermöglichen [...]“. (S. 302)

Dyllick (2011) referiert hier ihren aus konstruktivistischer Perspektive geschriebenen Aufsatz, in dem sie eine stärkere Gewichtung von Eigenverantwortung und individueller Passung des Lernangebots fordert, im Gegensatz zur beobachteten didaktischen Monokultur des Sings. Dies resultiert später in einer theoretischen Arbeit, in der sie Singen auf der einen Seite als Motor musikalischen Lernens beschreibt (z. B. bei der relativen Solmisation), aber auch einzelne Stimmen aus dem Chor-Setting herauslöst (z. B. in Castingshows) und Schülerinnen und Schüler dazu anhält, ein Portfolio zur eigenen Stimmentwicklung zu führen (Dyllick, 2019, S. 129 ff.). Singen wird so individualisiert und aus dem kollektiven

Gemeinschaftstun herausgelöst¹⁷. Pezenburg und Dyllick beziehen sich auch auf Oberschmidt (2016), der „das Entdecken und Erspüren der ganz individuellen Ausdrucksmöglichkeiten des Singens“ (S. 15) gefördert wissen will. Neben diesem konstruktivistischen Ansatz für eine Individualisierung des Singens nennt Hoos de Jokisch (2014) auch den Selbstdarstellungstrieb der heutigen Popkultur als Türöffner für individuelles Singen.

Für die vorliegende Arbeit zeigen sich Anknüpfungspunkte bei diagnostischen Verfahren und in konstruktivistischen Ansätzen, welche individualisierendes Unterrichten theoretisch begründen. Zum notwendigen fachdidaktischen Knowhow und zu Überzeugungen, wie Singen unterrichtet werden sollte, ist wenig bekannt. Fachdidaktische Wissensbestände wurden in den letzten Jahren von Puffer und Hofmann (2016, 2017) in der Falko-Studie (Krauss et al., 2017) untersucht. Aufbauend auf die Theorie von Shulman (1986, 1987) wurde für Musiklehrkräfte ein Testinstrument entwickelt, in dem *pedagogical content knowledge* (PCK) valide abgefragt werden kann. Für die vorliegende Studie sind die spezifisch auf Gesang ausgerichteten Items aber zu wenig detailliert. Die nachfolgend durchgeführte LEHET-Studie von Puffer und Hofmann (2021) konnte nachweisen, dass mithilfe von Videovignetten fachdidaktische Wissensbestände für das Singen aufgebaut werden können. Das entwickelte Instrument kann sich als wichtiger Baustein für den Aufbau von Diagnosefähigkeiten von Musiklehrkräften erweisen, was im Ausblick (vgl. Bereich E) nochmals aufgegriffen wird. Beide Studien waren allerdings quantitativ ausgerichtet und hatten nicht das Ziel, die individuellen Konzepte von Lehrkräften sichtbar zu machen.

Seit längerer Zeit forscht Niessen zu Individualkonzepten von Musiklehrkräften (2006, 2014, 2016). Sie verwendet leitfadengestützte Interviews und die Grounded Theory (Glaser et al., 1998) als Auswertungsmethode, um eine Theorie der Individualkonzepte von Musiklehrkräften zu entwickeln. In ihrer Arbeit fällt insbesondere auf, dass in Individualkonzepten

¹⁷ An dieser Stelle sei noch der Hinweis platziert, dass Dyllick mit ihrem theoretischen Ansatz zurück zum Musikunterricht der 1970er-Jahre in der Schweiz (Kälin, 1976) führt, in dem, wie bereits dargelegt, das Lied als Ausgangspunkt musikalischen Lernens diente. Neu ist dann allerdings, dass das Lernen individuell stattfindet und nicht, wie Oberschmidt (2016) schreibt, mechanistisch unterwiesen wird.

der biografische Bezug zu positiven wie negativen Vorbildern hervorsteicht und die eigene Erfahrung jeweils als prägendes Element herausgearbeitet werden kann. Sie bezweifelt, bei Musiklehrkräften typische Merkmale zu finden, eher sei von Heterogenität auszugehen. Als Fazit schlägt sie vor, „auch für Lehrer_innen eine ‚Fehlerkultur‘ zu etablieren, in der ‚missglückter‘ Unterricht vor allem als Chance zum Umlernen und produktive Irritation, nicht aber als Beleg eines Scheiterns verstanden wird“ (Niessen, 2014, S. 7).

In der aktuellen Untersuchung zur PROFESSIONALISIERUNG VON KINDER- UND JUGENDCHORLEITERN erarbeitet Schaumberger (2020b) ein Kompetenzmodell für Chorleiterinnen und Chorleiter, das mit einer qualitativen Inhaltsanalyse auf der Basis von Experteninterviews überprüft wird (S. 382 ff.). Obwohl die Arbeit nur bedingt auf die Verhältnisse der Schule übertragen werden kann, ergeben sich doch Schnittmengen, z. B. im Vergleich mit einzelnen Lehrkräften der Schweiz, die ebenfalls eine Ausbildung für Chorleitung haben und zudem in der Sekundarstufe 1 arbeiten.

Zu Beginn des Kapitels wurde die spezifische Situation in der Schweiz beschrieben, wobei klar wurde, dass die Singtradition in der Schweiz zwar nach wie vor stark ist, dies aber kein Beleg für eine hochstehende Singpraxis zu sein scheint. Weiter wurde herausgearbeitet, dass sich bezüglich der Leistungsbeurteilung Herausforderungen aufgrund der engen Verbindung der Persönlichkeit mit der Singstimme zeigen, die zusätzlich durch die Stimmentwicklung in der Pubertät erschwert werden. Im letzten Abschnitt wurde herausgearbeitet, dass die Idee, im Singunterricht individuelle Entwicklung adäquat zu adressieren, nur wenig entwickelt ist. Zudem gibt es zum Singen in der Sekundarstufe 1 kaum Studien, an welche angeschlossen werden kann. Die Beschreibung individueller Förder- und Beurteilungsprozesse in der Sekundarstufe 1 stellt sich damit als Forschungsdesiderat heraus. Diese Lücke soll mit leitfadengestützten Interviews bearbeitet werden. Die notwendigen Wissensbestände und der methodische Aufbau der vorliegenden Arbeit ist Gegenstand des nächsten Kapitels.

2 Gliederung der Studie

In Kapitel 2 wird die Gliederung der Forschungsarbeit dargelegt. Nach der bereits erfolgten Einführung (A) folgen die vier Bereiche Forschungsstand (B), Methode und Untersuchung (C), Darstellung der Resultate (D) und Ausblick (E). Dieser Aufbau entspricht der üblichen Darstellung von Forschungsergebnissen. Im Gegensatz dazu konstatieren Strübing et al. (2018), dass Forschungsprozesse qualitativer Arbeiten zwingend einer „iterativ-zyklischen Prozesslogik“ (S. 85) folgen und durch die Darstellungslogik von Forschungsberichten nur bedingt abgebildet werden können. Diese Herausforderung machte sich auch in der vorliegenden Arbeit bemerkbar, weil Theorie, Empirie und Methode sich gegenseitig angeregt und auch herausgefordert haben. So wurde im Verlauf der Inhaltsanalyse die Notwendigkeit erkennbar, theoretische Elemente weiter zu vertiefen und anschließend wiederum die Analyse des empirischen Materials anzupassen. Weiter wurde mit dem autoethnografischen Exkurs zu Beginn des Forschungsstandes ein erster Forschungszyklus in die Logik dieser Arbeit eingebaut, denn Theorie (Kapitel 4 & 5) schließt sich an Beobachtung und Analyse an (Kapitel 3) und führt weiter zu Empirie (ab Kapitel 8). Die *iterativ-zyklische Prozesslogik* zeigt sich in der vorliegenden Arbeit immer dann, wenn auf zukünftige Kapitel verwiesen wird, oder eine Argumentation zu einem zurückliegenden Kapitel vervollständigt und revidiert wird. Allerdings kann das iterative Vorgehen nicht vollständig transparent gemacht werden, weil die Darstellungslogik dies erschwert und die sprachliche Ausgestaltung zu komplex würde. Die Gliederung der vier anschließenden Bereiche wird nun linear in vier Abschnitten ausgeführt.

2.1 Bereich B: Singunterricht auf dem Prüfstand

Die Darlegung des Forschungsstandes zum Thema Assessment im Singunterricht erfolgt in zwei Kapiteln, die mit zwei Exkursen gerahmt werden. Eröffnet wird der Bereich B mit der bereits erwähnten Autoethnografie, worin ausgewählte Episoden aus dem langjährigen Unterrichtsgeschehen des Forschenden dargestellt und analysiert werden. Darauf folgen im ers-

ten theoretischen Kapitel ausgewählte Aspekte des Singens, um auf die damit verbundenen spezifischen Herausforderungen im Assessment hinzuführen (vgl. auch 1.2). In Kapitel 5 wird der Begriff Assessment theoretisch fundiert und als Förder- und Beurteilungskonzept auf das Singen bezogen. Der nachfolgende Exkurs beschreibt einige Eigenheiten des schweizerischen Bildungssystems und ermöglicht damit für *Ausserschweizer*¹⁸ einen Blick auf die regionalen Besonderheiten. Abschließend folgen die Forschungsfragen dieser Studie. In einem kurzen Abriss werden nun die einzelnen Kapitel skizziert.

Mit der autoethnografischen Darlegung der Position des Forschenden (Kapitel 3) wird für eine deutschsprachige Studie ein untypischer Einstieg gewählt. Obwohl Orgass (2017) sich deutlich zugunsten einer unverfälschten Wissenschaftslogik von autoethnografischen Verfahren distanziert, wird mit diesem Kapitel eine methodische Exploration in der deutschsprachigen Musikpädagogik gewagt. Das entstehende Bild eines sich entwickelnden Singunterrichts, gezeichnet durch retrospektive Unterrichtsbeschreibungen, analytischen Betrachtungen in der Ich-Form sowie Überlegungen zur Methode der Autoethnografie, dient als thematische Landkarte und bereitet inhaltlich die beiden theoretischen Kapitel vor. Dass Autoethnografie nicht zwingend abzulehnen ist, wie Orgass dies argumentiert, zeigt ein Blick über die Sprachgrenzen hinaus, wenn z. B. de Vries (2012) seine musikpädagogische Laufbahn in einem Zwiegespräch zwischen sich und seinem jüngeren Selbst reflektiert und gleichzeitig einen Beitrag zum Methodendiskurs liefert.

In Kapitel 4 werden verschiedene Aspekte der adoleszenten Stimmentwicklung dargelegt. Der Begriff *adoleszente*¹⁹ Stimme bzw. adoleszente Mädchen- und Jungenstimme wird in der vorliegenden Forschungsarbeit aus der anglofonen Literatur übernommen (z. B. Sweet, 2018; Gackle, 2019; Welch et al., 2019; Freer, 2020). Der Begriff *adolescent voice*

¹⁸ Ausserschweizer [sic! kein scharfes s in der Schweiz] sind immer die anderen. Der Begriff wird sowohl im deutschsprachigen Teil des Kantons Valais als auch im Kanton Schwyz verwendet, um die Schweizer auf der anderen Seite der Alpen bzw. außerhalb des Kantons zu bezeichnen, notabene den weitaus größeren Teil. Für den Begriff sind keine gegenderten Versionen bekannt.

¹⁹ Der Begriff *adolēscēns* stammt aus dem Lateinischen und bedeutet heranwachsend, jung oder jugendlich bzw. als Substantiv junger Mann, junges Mädchen.

development wird in englischsprachigen Ländern flächendeckend benutzt, um die stimmliche Entwicklung während der Pubertät zu beschreiben. Die Setzung dieses Begriffs auch für die deutsche Sprache zielt darauf ab, den Begriff *Mutation* zu umgehen, weil dieser meist nur mit dem männlichen Geschlecht assoziiert wird. Der Terminus *adoleszente Stimme* beschreibt geschlechtsneutral Jungen- und Mädchenstimme und öffnet damit thematisch hin zu den Mädchen, die in der deutschsprachigen Literatur zum Thema vernachlässigt werden²⁰.

Einleitend wird in Abschnitt 4.1 überlegt, wie Singen als Tätigkeit insgesamt und in der Schule eingeordnet werden kann. Dazu wird zuerst Singen als anthropologische Konstante definiert, wobei sich eine Diskrepanz zwischen der Fähigkeit zu singen und dem zugehörigen Selbstkonzept zeigt. Weiter wird auf Neus (2017) zurückgegriffen, um verschiedene konzeptuelle Ebenen des Singens zu unterscheiden und diese abschließend auf den aktuellen Lehrplan zu beziehen.

In der Literatur zur adoleszenten Stimmentwicklung werden nicht nur stimmliche Phänomene, wie die Veränderung von Ambitus und Klangfarbe, oder entwicklungspsychologische Aspekte thematisiert, sondern auch ihre Auswirkungen auf das gemeinsame Singen in der Gruppe. Aus diesem Grund wird in Abschnitt 4.2 zuerst überlegt, wie diese Bereiche konzeptuell geordnet werden können. Dazu wird die Darstellung von King (2013) hinzugezogen, in der vorgeschlagen wird, biologische Prozesse mit dem Begriff Pubertät zu verknüpfen, für entwicklungspsychologische Momente den Begriff Adoleszenz²¹ zu verwenden und gruppendynamische Prozesse mit dem soziologisch verwendeten Terminus Jugend zu untersuchen. Es resultieren zwei Ebenen, einerseits die individuelle Entwicklung während der Pubertät sowie der Adoleszenz und andererseits die Einbettung dieser Entwicklung in die soziale Situation im Singunterricht. Die nachfol-

²⁰ An dieser Stelle sei Joanne Rutkowski gedankt, die mich 2019 an der ISAME7 in Florida auf die entsprechende Literatur hingewiesen hat.

²¹ Hier zeigt sich eine gewisse Unschärfe in der Abgrenzung zum Begriff *adoleszente Stimme*. Im anglofonen Raum werden biologische und entwicklungspsychologische Elemente der Stimmentwicklung im Begriff subsummiert. In der vorliegenden Studie wird dies so übernommen, außer wenn eine Abgrenzung notwendig wird, so z. B. bezüglich der Einflüsse auf Selbstkonzepte (Gender), die von biologischen Prozessen unterschieden werden können.

genden Abschnitte besprechen jeweils die individuelle Ebene der Mädchen und Jungen, gefolgt von einem Abschnitt zum gemeinsamen Singen.

In der deutschsprachigen Literatur zeigt sich bei beiden Geschlechtern eine Lücke in der Rezeption anglofonen Literatur. So fällt auf, dass im musikpädagogischen Bereich bei Jungen vereinfachend von Praemutation, Mutation und Postmutation (z. B. A. Mohr, 1997, 2014) gesprochen wird, obwohl seit den 1970er-Jahren Untersuchungsergebnisse aus Amerika vorliegen, die ein differenzierteres Bild zeichnen (Cooksey, 1977a, 2000). Ein erster Beitrag, der das Tor hin zur anglofonen Literatur öffnet, wurde vom amerikanischen Musikpädagogen Freer und dem deutschen Lehrmittelautor Detterbeck vorgelegt (Freer & Detterbeck, 2019)²² und der Sammelband *SINGING WITH CHILDREN: INTERNATIONAL PERSPECTIVES* (Sandt, 2020) führt die Tendenz hin zu einer internationalen Sichtweise weiter²³. Im Hinblick auf die Beschreibung und Evaluation formativer und summativer Assessment-Prozesse im empirischen Teil dieser Studie scheint es zentral, die Entwicklung der adoleszenten Stimme bei Mädchen und Jungen zu verstehen und die bereits erwähnte anglofone Literatur dazu einzubeziehen. Dies führt je zu einem Abschnitt zu Mädchen (vgl. 4.3) und zu Jungen (4.4). In beiden Abschnitten wird gezeigt, dass nicht nur durch die Pubertät ausgelöste Veränderungen beschrieben werden müssen, sondern immer auch Elemente der entwicklungspsychologischen Seite eine Rolle spielen und so in der Schule zur Herausforderung werden.

Der Begriff Jugend führt in Abschnitt 4.5 zu einer weiteren Perspektive, denn das Zusammentreffen der Geschlechter beim gemeinsamen Singen erfordert zumindest in Anteilen die Beachtung des Genderdiskurses²⁴.

²² Mit dem Titel *WENN DIE STIMME ‚BRICHT‘* (Freer & Detterbeck, 2019) bleiben die Autoren aber immer noch thematisch auf das Geschlecht der Jungen fixiert.

²³ Ein interessanter Umstand sei an dieser Stelle angefügt. Die Literaturlisten deutsch- und englischsprachiger aktueller Artikel zur adoleszenten Stimme zeigen kaum Überschneidungen. So zitieren z. B. Freer (2020) aus Amerika und Fuchs (2020) aus Deutschland mit ihren Artikeln zur adoleszenten Jungensstimme, notabene beide im selben Band erschienen (Sandt, 2020), vollkommen unterschiedliche Literatur.

²⁴ In dieser Studie werden geschlechtsspezifische Merkmale mehrheitlich als durch hormonelle Prozesse ausgelöste biologische Phänomene betrachtet, dies im Bewusstsein, dass damit auch Selbstfindungsprozesse ausgelöst werden, welche das soziale Geschlecht Gender betreffen. Die aktuelle Erweiterung der

Zusätzlich drängt sich aus dem autoethnografischen Kapitel die Frage auf, ob koedukative und monoedukative Formen des Unterrichtens beim Singen sinnvoll sein könnten. Das Kapitel wird deshalb mit geschlechtsspezifischen Aspekten des Singens, beleuchtet aus verschiedenen Blickwinkeln, abgeschlossen. Im Verlauf des Kapitels 4 wird mit der Beschreibung wichtiger Phänomene des Singens eine Basis für die Betrachtung von Förder- und Beurteilungsprozessen gelegt, die im nachfolgenden Kapitel thematisiert werden.

Das Kapitel 5 widmet sich dem zentralen Begriff der vorliegenden Arbeit: *assessment*. Der Begriff ist vor allem aus dem anglofonen Raum bekannt und wird für jegliche Form von Bewertung und Beurteilung angewendet, seien dies Sachen, Prozesse oder Personen. Der Duden gibt für den deutschen Begriff *Assessment* drei Bedeutungen an: Steuerveranlagung, Bewertung und Assessment-Center, wobei der letzte Begriff als „psychologisches Testverfahren, bei dem jemandes Eignung (besonders für eine Führungsposition) festgestellt werden soll“²⁵, definiert wird. Die vorliegende Arbeit verwendet das Wort *Assessment* in einem ganzheitlichen Sinn, wie dies aus dem anglofonen Raum bekannt ist. Für den Singunterricht werden so einerseits Maßnahmen beschrieben, die im Rahmen einer formativen Beurteilung eine diagnostische Einschätzung ermöglichen, sei dies durch die Lehrkraft oder die Lernenden selbst, aber auch Leistungsmessungen, die zu Noten führen.

Aufgrund der noch wenig etablierten holistischen Wortbedeutung wird im ersten Abschnitt (vgl. 5.1) der Begriff nicht nur sprachlich und inhaltlich vertieft, sondern auch bezüglich der zugrunde liegenden erziehungswissenschaftlichen Konzepte untersucht (Winter, 2015; C. Schmidt, 2020). Fokussiert wird aber vor allem auf das *Assessment-Modell* von Earl (2013), der *Assessment* auf das Lernen bezieht, und die Bereiche *assess-*

Diskussion mit Blick auf Transkinder wurde in den Interviews nicht thematisiert und wird deshalb in dieser Studie nicht berücksichtigt.

²⁵ Duden-Online für die Begriffe *Assessment* und *Assessment-Center* (<https://www.duden.de> – abgerufen am 27.08.2021). In Gesprächen wird immer wieder deutlich, dass diese Bedeutung stark verankert ist und eine holistische Betrachtungsweise nicht vorausgesetzt werden darf.

ment for learning (AFL), assessment as learning (AAL) und assessment of learning (AOL)²⁶ unterscheidet.

Im zweiten Abschnitt (vgl. 5.2) folgt eine Fokussierung auf den Musikunterricht. Dabei dienen Veröffentlichungen englischsprachiger Literatur in Form von Monografien (z. B. Fautley, 2010) und Handbüchern (Brophy, 2019a, 2019b) sowie zahlreiche Zeitschriftenartikel als Grundlage. Zuerst wird die *philosophy of assessment* (Mantie, 2019) diskutiert und mit den acht Prinzipien für Assessment (Brophy, 2019c) abgeglichen. Dies dient als Grundlage für eine Fokussierung auf Assessment-Methoden des Singens. Neben dem Einsatz von *Rubrics*²⁷ (vgl. Whitcomb, 1999; Panadero & Jonsson, 2013; Schneider et al., 2019) und dem Einbezug softwarebasierter Tools (z. B. Demorest et al., 2017) wird auch der kommerzielle Aspekt von Assessment-Lösungen (z. B. MusicFirst²⁸) besprochen. Allerdings müssen die daraus entstehenden Möglichkeiten auch kritisch beleuchtet werden, da die Entwicklung nicht selten vor dem Hintergrund der Legitimation von Musikunterricht stattfindet (Mantie, 2019).

Im dritten Abschnitt (vgl. 5.3) wird wieder auf den deutschsprachigen Diskurs fokussiert und untersucht, inwiefern Lehrkräfte auf eine ganzheitliche und individualisierende Unterrichtsgestaltung im Singunterricht vorbereitet sind. Die Untersuchung von Brunner (2014) zu diagnostischen Fähigkeiten von Lehrkräften mit Fokus auf den Singunterricht wurde bereits erwähnt (vgl. 1.3), wird aber noch inhaltlich vertieft. Wichtige Hinweise kann auch die Diskussion zur Wahrnehmung von Heterogenität im Musikunterricht (Linn, 2017) sowie die Einschätzung zur Leistungsbeurteilung des Singens von Schaumberger (2020a) liefern. An dieser Stelle zeigt sich aber auch das Forschungsdesiderat, da kaum Studien zum Singen in der Sekundarstufe 1 vorliegen.

Die Klammer um die beiden theoretischen Kapitel wird mit einem Exkurs in Kapitel 6 zu den Eigenheiten des Bildungssystems der Schweiz geschlossen. Dies wird als notwendig erachtet, zumal die Schulstufe der

²⁶ Die verschiedenen Bereiche von Assessment werden im Dokument mit ihren Abkürzungen (AFL, AAL, AOL) verwendet (vgl. Tabelle 1, S. 19).

²⁷ Der Begriff *Rubric* meint vielfältige Formen von (Kompetenz-)Rastern (Stevens & Levi, 2013) und wird in dieser Arbeit nicht übersetzt, sondern in der englischen Form belassen. Dies wird später begründet (vgl. 5.2).

²⁸ <https://www.musicfirst.com/assessment/> – abgerufen 05.11.2021.

Sekundarstufe 1 sich deutlich von anderen deutschsprachigen Ländern unterscheidet. Zusätzlich zeigen sich große Unterschiede durch das föderalistische System der Schweiz, in dem die Kantone die Hoheit über die Ausgestaltung der Schule und deren Inhalte haben und auch die Anstellungsbedingungen bestimmen. Im Exkurs wird zuerst gezeigt, wie in den letzten Jahren in einem politischen Prozess eine strukturelle und inhaltliche Vereinheitlichung im Schulsystem durchgesetzt wurde. Zudem wird dargestellt, wie die Veränderungen in der Ausbildungssituation im Anschluss an die Gründung von pädagogischen Hochschulen nach 2000, kombiniert mit dem allgemeinen Mangel an Lehrkräften, zu einer heterogenen Berufsgruppe von Musiklehrkräften führte. Im letzten Abschnitt wird die nach wie vor unterschiedliche Ausgestaltung von Leistungsniveaus und die unterschiedliche Stundendotation in den verschiedenen Kantonen skizziert. Der Exkurs dient dazu, zusätzlich zu der Diskussion der regional-nationalen Singtradition der Schweiz (siehe 1.1) auch die relevanten Merkmale der Bildungslandschaft im Hinblick auf die Ausgestaltung der Interviewstudie zu verdeutlichen.

Das letzte Kapitel des Bereichs B beinhaltet die Forschungsfragen für die empirische Studie (vgl. Kapitel 7). Das Forschungsdesiderat wird in fehlenden Studien zum Singen in der Sekundarstufe 1 im deutschsprachigen Raum deutlich, insbesondere auch bezüglich der Thematik der individuellen Förderung und Beurteilung. Vorhandene Studien sind eher auf die Grundschule ausgerichtet (z. B. Forge & Gembris, 2012) oder streifen das Thema Singen innerhalb einer anderen Thematik (z. B. Tellisch, 2015). Die Forschungsfragen zielen somit auf die Beschreibung und Analyse von Praktiken des Singunterrichts im Sinne einer Situationsanalyse und deren zugrunde liegenden Zielvorstellungen ab. Fokussiert werden Praktiken der Diagnose, Förderung und Beurteilung, immer unter Berücksichtigung der adoleszenten Stimmentwicklung und eines individualisierenden Zugangs zu den Jugendlichen.

2.2 Bereich C: Methode und Untersuchung

Der Bereich C widmet sich der Forschungsanlage und ihren Gütekriterien, der Auswahl des Untersuchungsmaterials und dem konkreten Forschungsprozess.

Das Kapitel 8 dient mit der Entwicklung des Forschungsdesigns als Scharnier zwischen Theorie und Empirie. In einem ersten Schritt wird hergeleitet, wie sich die vorliegende Arbeit in die empirische Sozialforschung eingliedert. Die Überlegungen von Gläser und Laudel (2010) zur Produktion von Wissen werden zusammen mit den von Strübing et al. (2018) vorgestellten Gütekriterien qualitativer Sozialforschung als Basis für den strukturellen Aufbau der vorliegenden Arbeit vorgestellt. Im Abschnitt 8.1 wird das Forschungsdesign erläutert. Zusätzlich zur Wahl von Interviews als Datenbasis wird auch der Rückbezug auf die Erfahrung des Forschenden im autoethnografischen Exkurs argumentiert. Die Interviews dienen in Anlehnung an Niessen (2006) dazu, „die Überzeugungen der Lehrerinnen und Lehrer sowie das damit verbundene unterrichtliche Handeln abzubilden“ (S. 142). Im Abschnitt 8.2 wird zuerst die qualitative Inhaltsanalyse (QIA) nach Mayring und Fenzl (2019) in einen historischen Zusammenhang gestellt. Anschließend wird die strukturierende Inhaltsanalyse (Schreier, 2014) als Grundlage der vorliegenden Arbeit definiert und beschrieben. Zusätzlich werden verschiedene Analysetechniken vorgelegt, die sich für die vorliegende Arbeit als sinnvoll erweisen. So zeigt sich die skalierende Inhaltsanalyse als wertvoll für die Evaluation von Sozialformen und die typenbildende Inhaltsanalyse bezüglich der Analyse und dem Vergleich von Einzelfällen. In Abschnitt 8.3 wird die Strategie des Samplings und die Form der Interviews hergeleitet sowie deren Kommunikationszusammenhang problematisiert. Dazu wird das von Mayring (2015) entwickelte Kommunikationsmodell expliziert und mit Blick auf die vorliegende Arbeit ausgeführt. In Abschnitt 8.4 werden Gütekriterien dargestellt, die für die vorliegende Arbeit relevant sind. Zusätzlich zu den generellen Ansprüchen an Forschung (Flick, 2019) werden die Ansprüche an die qualitative Inhaltsanalyse bezüglich Reliabilität und Validität besprochen.

In Kapitel 9 wird die Auswahl des Untersuchungsmaterials bis hin zur Erstellung des Datenkorpus beschrieben. Zuerst wird dazu ein Auswahlverfahren für Musiklehrkräfte entwickelt, das sich auf das Modell professioneller Handlungskompetenz von Baumert und Kunter (2006) bezieht und auf Varianzmaximierung (Patton, 2015) abzielt. Im zweiten Abschnitt werden die Rekrutierung der Interviewpartnerinnen und -partner sowie die Durchführung der Interviews beschrieben. Die Analyse der Ausbildungsbiografien führt im Anschluss zu einer Justierung der zuvor hergeleiteten

Profile entlang pädagogischer und musikalischer Ausbildungselemente. Im letzten Abschnitt wird mit der Erstellung der Transkriptionen die Überführung der Interviews in den Datenkorpus dargelegt.

Der Kern der qualitativen Inhaltsanalyse nach Mayring (2015) ist das Kategoriensystem. In Kapitel 10 wird der Prozess der Entwicklung transparent gemacht. Im ersten Abschnitt wird ein möglicher Einfluss der Software MAXQDA diskutiert (Rädiker & Kuckartz, 2019), um anschließend im zweiten Abschnitt die Entwicklung des Kategoriensystems zu beschreiben. Ausgehend von theoretischen Überlegungen wird die Strukturierung des Untersuchungsmaterials entlang deduktiv erstellter und induktiv dazugewonnener Kategorien beschrieben. Anschließend wird der schrittweise auf den ganzen Datenkorpus ausgeweitete Kodierprozess argumentiert und der Prozess der inhaltlichen Verdichtung mit Summaries erklärt. Im dritten Abschnitt wird die Intercoder-Reliabilitätsprüfung beschrieben. Das Kategoriensystem mit seinen Summaries zu einzelnen Kategorien ist danach Grundlage verschiedener Analysemethoden aus dem methodischen Baukasten (Schreier, 2014; Stamann et al., 2016) der QIA.

2.3 Bereich D: Darstellung der Resultate

In den drei Kapiteln des Bereichs D werden die Resultate der qualitativen Studie in drei Schritten aufbereitet. Zuerst werden die Kodierhäufigkeiten des Kategoriensystems einer quantitativen Untersuchung unterzogen. Im nächsten Schritt entstehen aus einer vertikalen Perspektive kurze Porträts zu den Singkulturen der interviewten Musiklehrkräfte. Diese dienen dann als Basis, um im Licht der Forschungsfragen Typiken herauszuarbeiten (Kuckartz, 2016) und auf Biografie und Ausbildung zu beziehen. Schließlich werden entlang des Kategoriensystems geschlechtsspezifische Einflüsse und Assessment-Praktiken des Singunterrichts vertieft analysiert.

Die quantitative Analyse des Kategoriensystems in Kapitel 11 bezweckt auf der einen Seite eine Einschätzung, ob und wie sich die deduktiv und induktiv erarbeiteten Kategorien im Sample verteilen und welche Unterschiede sich zwischen den Teilnehmenden der Studie zeigen. Weiter werden die Kodierhäufigkeiten der Assessment-Kategorien vertiefter untersucht, um einschätzen zu können, inwiefern sich ganzheitliches Assessment in den lokalen Singkulturen zeigt. Diese Analyse dient im An-

schluss als Ausgangsbasis für die qualitative Auswertung. Zusätzlich wird in diesem Kapitel eine skalierende Strukturierung der Sozialformen im Singunterricht entwickelt, um einen Überblick zu erhalten, wie stark sich individualisierende Formen in den lokalen Singkulturen zeigen.

In Kapitel 12 werden auf der Grundlage des Kategoriensystems und weiterer Forschungsbelege die Zielvorstellungen und die Förder- und Beurteilungsstrategien der Interviewpartnerinnen und -partner in Porträts zusammengefasst und mit repräsentativen Überschriften charakterisiert. Die vertikale Perspektive wirft einen Blick auf Einzelpersonen und ermöglicht die Gruppierung von Lehrkräften entlang gemeinsamer Zielvorstellungen und ähnlicher Unterrichtserlebnisse. Als Ergebnis werden zwei Polaritäten beschrieben, auf der einen Seite die Ausrichtung des Singens auf musikalisch-vokale Ziele oder persönlichkeitsbildende Momente und auf der anderen Seite Widerstand seitens der Schülerinnen und Schüler, der im Gegensatz zu einem empathischen Miteinander in der Klassengemeinschaft steht.

In Kapitel 13 erfolgt eine Analyse des Kategoriensystems aus einer vertikalen Perspektive entlang der beiden Hauptthemen der vorliegenden Arbeit. Im ersten Abschnitt (vgl. 13.1) werden die Interviews in Bezug auf die adoleszente Stimmentwicklung und weitere geschlechtsspezifische Phänomene untersucht. Dazu gehören die Wahrnehmung und der Umgang mit der adoleszenten Stimme durch die Lehrkräfte ebenso wie ihre eigene geschlechtsspezifische Vorbildrolle. Aufgrund der gewonnenen Erkenntnisse erfolgt im zweiten Abschnitt (vgl. 13.2) eine Analyse der Förder- und Beurteilungsstrategien. Dazu wird zuerst eine Analyse der Konzeption der Assessment-Prozesse durchgeführt. Diskutiert wird insbesondere die im Sample stark dominierende Verknüpfung summativer Beurteilung mit anschließendem Feedback, welche als „formative use of summative assessment“ (Fautley, 2010, S. 198) beschrieben wird und nicht mit dem Modell von Earl (2013) übereinstimmt, das formative Einschätzung und entsprechende Intervention (AFL) als Grundlage für die Leistungsbeurteilung (AOL) vorsieht. Aufgrund dieser Analyse werden die summativen Beurteilungsprozesse (AOL) zuerst genauer analysiert, spezifisch in Bezug auf individualisierende Formen und Transparenz in der Leistungsbeurteilung. Mit Rückbezug auf die adoleszente Stimmentwicklung wird auch untersucht, ob den damit verbundenen Phänomenen Rechnung getragen wird. Bezüglich Förderprozessen (AFL) zeigt sich eine große Breite

an Methoden für Klassensettings, die allerdings zugunsten einer Fokussierung auf individualisierende Zugänge nur minimal in die Analyse einbezogen werden. Im Zentrum stehen dagegen der Umgang mit Differenz und die Suche nach diagnostischen Elementen, wobei nicht nur stimmliche Phänomene, sondern auch Selbstkonzepte in den Blick genommen werden. Die von Earl ebenfalls ins Feld geführte metakognitive Ebene (AAL) wird aufgrund spärlicher Belege vor allem bezüglich des Unterrichtsklimas untersucht.

Der letzte Abschnitt (vgl. 13.3) fasst die Resultate der QIA entlang der Forschungsfragen zusammen, wobei auch die Ergebnisse der Kapitel 11 und 12 berücksichtigt werden. Im Abgleich mit der besprochenen Literatur und dem vorgestellten Assessment-Modell für den Singunterricht (vgl. 5.1) werden Elemente von Diagnose und Förderung (AFL), Leistungsbeurteilung (AOL) und Selbstevaluation (AAL) eingeordnet und gewürdigt.

2.4 Bereich E: Ausblick

In der vorliegenden Arbeit wird ein Bogen gespannt von der Praxis des Forschenden im autoethnografischen Exkurs hin zum Ausblick, in dem Erkenntnisse aus der empirischen Studie für die Praxis gebündelt werden. Zudem wird auch ein Bogen gespannt von der noch wenig etablierten Strategie einer ganzheitlichen Leistungsförderung und -beurteilung im Singunterricht hin zu Forschungsdesideraten, welche sich aus der vorliegenden Arbeit ergeben.

Der erste Bogen führt in Kapitel 14 zu drei Thesen für ein gelingendes Assessment im Singunterricht. In einer ersten These werden Ideen gebündelt, wie die adoleszente Stimmentwicklung und geschlechtsspezifische Unterschiede im Singunterricht einbezogen werden können. In einer zweiten These wird überlegt, wie ganzheitliches Assessment beim Singen unter Berücksichtigung der Erkenntnisse aus der Interviewstudie implementiert werden kann. In einer abschließenden These wird die Idee eines Stimmportfolios etabliert, das insbesondere die Verbindung von Stimme und Person sowie interindividuelle Unterschiede adressiert und damit Potenzial hat, ein adäquates Entwicklungs- und Beurteilungsinstrument zu werden.

In Kapitel 15 rücken die Forschungsdesiderate ins Zentrum, die im Laufe der vorliegenden Arbeit entstehen. Der explorative Charakter der

qualitativen Interviewstudie führt zur Beschreibung und Analyse von Assessment-Strategien im Licht theoretischer Konzepte und ist als Beitrag zum Diskurs zur Leistungsbeurteilung im Singunterricht zu sehen, wie ihn Schaumberger (2020a) anregt, beschränkt sich aber auf die Sekundarstufe 1. Insbesondere die Relevanz der Phänomene adoleszenter Stimmentwicklung im Musikunterricht und der Einsatz von Assessment-Strategien inkl. der Portfolio-Idee müssten in Nachfolgestudien im deutschsprachigen Raum untersucht werden.

B SINGUNTERRICHT AUF DEM PRÜFSTAND

Die Gliederung dieses Bereichs wurde in Kapitel 2.1 ausführlich argumentiert. Singunterricht auf dem Prüfstand beinhaltet nicht nur den Forschungsstand zum Singen in der Adoleszenz und zum Assessment des Singens, sondern sondiert auch die Erfahrung des Forschenden im autoethnographischen Teil und das bildungspolitische Umfeld mit dem Exkurs zu den Rahmenbedingungen in der Schweiz. Die vier Perspektiven beleuchten wesentliche Aspekte des Singens in der Sekundarstufe 1 und führen zu den Forschungsfragen, die ihrerseits zum empirischen Teil überleiten.

3 Exkurs: Autoethnografische Rückblende

3.1 Punkt 1: Der Forschende im Forschungsfeld

Präludium: Vorsingen vor der Klasse

Wir befinden uns in der Bezirksschule²⁹ eines Seitentals der Aare um die Jahrtausendwende. In der Aula der 100-jährigen Schule steht ein Stuhlkreis rund um den Flügel. Auf der Bühne, die sich mit drei Treppenstufen direkt an den Raum anschließt, steht ein Mikrofon. Der Lehrer startet: „Guten Morgen. Heute machen wir, wie angekündigt, das Vorsingen. Steht auf, wir singen gemeinsam die vier Lieder noch einmal durch.“ Ein kleines Murren geht durch die Runde, aber gemeinsam wird das Repertoire nochmals geprobt. Anschließend stellt sich der erste Schüler ans Mikrofon und singt Vers und Chorus seines Liedes. Er erhält einen Szenenapplaus, während sich die nächste Schülerin auf der Bühne vorbereitet. Der Lehrer schreibt eine Zahl auf die Klassenliste und fragt nach dem nächsten Song. Der Applaus ist mal lauter, mal leiser, der Gesangsvortrag mal versierter, mal bemühter. Gelacht wird nie, Widerstand gibt es kaum. Es reicht gerade, um in 45 Minuten der ganzen Klasse eine Plattform zu geben und das Vorsingen abzuschließen. Nachdem die Klasse gegangen ist, überlegt sich der Lehrer, ob die auf der Klassenliste verzeichneten Bewertungen stimmig sind und korrigiert da und dort eine Viertelnote³⁰ nach unten oder oben. In der nächsten Musikstunde werden die Resultate verkündet: „Wer möchte seine Note nicht öffentlich hören?“ Einzelne strecken auf, erfahren

²⁹ Die Bezirksschule ist im Kanton Aargau das höchste Leistungsniveau der Sekundarstufe 1. Ein ausführlicher Exkurs zum schweizerischen Schulsystem erfolgt in Kapitel 6.

³⁰ Noten zwischen 4–6 sind genügend, wobei 6 einer sehr guten Bewertung entspricht. Üblicherweise werden Noten in Abständen von Vierteln erteilt, z. B. 5.25 (gleich 5-) oder 4.75 (-5).

die Note im Anschluss privat, für den Rest der Klasse wird die Bewertung öffentlich verlesen.

Methode: Der Forschende auf seinen eigenen Spuren

In diesem Kapitel werden ausgewählte Episoden der Unterrichtspraxis aus der Biografie des Forschenden untersucht. Autoethnografie dient dabei als Methode, um retrospektiv die langjährige Tätigkeit des Musiklehrers darzustellen und zu analysieren. Die vier Feldnotizen wurden mit Blick auf Förderung und Beurteilung des Singens ausgewählt und berichten sowohl von exemplarischen Episoden, die sich in mehreren Klassen so abgespielt haben (vgl. Präludium & Postludium), aber auch von spezifischen Erlebnissen, die entscheidende Momente abbilden (vgl. beide Interludien). Um dieses Kapitel methodisch einzubetten, werden die Ansprüche an Autoethnografie (Ellis & Bochner, 2000; Ellis, 2004; Anderson, 2006; Ellis et al., 2010; Holman Jones et al., 2016b) und der Diskurs zu evokativer und analytischer Autoethnografie (Anderson, 2006; Ellis & Bochner, 2006) ebenso dargestellt und diskutiert wie der marginale Diskurs zu dieser Methode in der musikpädagogischen Forschung des deutschsprachigen Raums (Orgass, 2017; Schulten & Lothwies, 2017; Dartsch et al., 2018). Abschließend wird der vorbereitende Charakter dieses Kapitels hin zu den beiden theoretischen Kapiteln begründet. Dabei dienen die mannigfaltigen Erscheinungsformen und das literarische Flair autoethnografischer Texte (Ellis & Bochner, 2000) als Vorbild, um Feldnotizen, Analysen und Betrachtungen zur Methodik gleich einem Rondo hintereinanderschalten.

Zuvor wird in diesem Abschnitt aber begründet, wieso dieser Exkurs notwendig erscheint und weshalb Autoethnografie als Methode gewählt wird. Lohmeier (2018) unterscheidet aus Sicht der Anthropologie zwischen „emischer und etischer Position zum Forschungsfeld“ (S. 30), wobei der Forschende in emischer Position zum Feld gehört und in etischer Position außerhalb des Feldes positioniert ist. Beide Positionen haben laut Lohmeier Vor- und Nachteile, beeinflussen aber gleichermaßen „den Forschungsprozess, einschließlich des Vorgehens bei der Datenauswertung“ (S. 29). Er bemängelt zudem, dass in qualitativen Studien „das Verhältnis des Forschers zum Forschungsgegenstand, zum Forschungsfeld oder

zum Forschungsthema nur selten reflektiert“ (S. 29) wird³¹. Der autoethnografische Exkurs nimmt diese Kritik auf und ermöglicht in erster Linie die Reflexion der emischen Position des Forschenden im Feld der Musiklehrkräfte. Er lenkt zudem den Blick mit der Auswahl der Feldnotizen bereits auf die Forschungsfragen und plausibilisiert damit die Thematik der vorliegenden Arbeit. Im Gegensatz zu *practitioner research*, bei dem die eigene musikpädagogische Praxis bewusst zum Forschungsfeld gemacht wird (Buchborn & Malmberg, 2013), ermöglicht die autoethnografische Methode den retrospektiven Zugriff auf das Reservoir an Erinnerungen, die prägenden Erlebnisse sowie verändernde Momente. Abschließend ist zu erwähnen, dass die Wahl der Methode auch als Exploration einer im deutschsprachigen musikpädagogischen Kontext kaum verwendeten Methode zu verstehen ist.

Analyse: Inhuman oder doch angemessen?

Ich gehe nun zurück zur Situationsbeschreibung zu Beginn des Kapitels und schreibe in der Ich-Form, um meine jetzige Position als Fachdidaktiker sichtbar zu machen. Im Gegensatz dazu habe ich mein Erleben als Musiklehrer aus einer beobachtenden Außenperspektive beschrieben. Bevor ich auf den Inhalt eingehe, ordne ich die beschriebene Situation ein. Diese Form der Leistungsbeurteilung habe ich mehrere Jahre durchgeführt, insbesondere in der ersten Hälfte meiner Zeit als Musiklehrer. Meist diente das aktuelle Repertoire als Basis für eine leistungsrelevante Beurteilung. Das Vorsingen gehörte unhinterfragt dazu und wurde weder vonseiten der Schülerinnen und Schüler noch der Eltern infrage gestellt.

Welche Perspektive nehme ich nun ein, um meine eigene Praxis zu analysieren? Mit dem normativen Blick durch eine fachdidaktische Brille ist Gesprächsbedarf auszumachen. Es gibt offensichtlich keine Beurteilungskriterien, Noten werden in Expertenmanier gesetzt. Die Leistung wird öffentlich erbracht und durch den Applaus auch öffentlich bewertet. Möglicherweise besteht so die Gefahr, dass die Beliebtheit von Schülerinnen

³¹ Diese Aussage steht in einer gewissen Diskrepanz zur Tradition der Writing Culture (Clifford & Marcus, 1986), in der die Forderung nach Beschreibung der eigenen Perspektive vermehrt ins Zentrum rückt. Es zeigen sich hier regionale Unterschiede, die später nochmals thematisiert werden.

und Schülern das Maß des Erfolgs bestimmt. Mit dem Verlesen der Noten wird eine weitere öffentliche Einteilung vollzogen und es besteht keine Möglichkeit, sich mit der eigenen Leistung auseinanderzusetzen. Darüber hinaus könnte sowohl der Akt des Singens als auch die Bekanntgabe der Note als bloßstellend empfunden werden. Dabei ist noch nicht berücksichtigt, dass die Bewertung des Singens als zulässiges Mittel in der Musikpädagogik hinterfragt werden könnte (vgl. 1.2). Schaumberger (2020a) weist bezüglich Leistungsbeurteilung darauf hin, dass allgemein akzeptierte Bewertungskriterien für das Singen in der Schule fehlen und bis anhin jede Lehrkraft ihren eigenen Weg sucht, wie eben im Beispiel beschrieben.

Eine allzu kritische Bewertung dieser Unterrichtssituation verstellt aber den Blick auf eine gelebte Realität, eine Unterrichtskultur im Kontext einer lokalen Schulkultur rund um die Jahrtausendwende. Die eher kleine Schule mit ungefähr 180 Schülerinnen und Schülern pflegte eine außergewöhnliche Performance-Kultur. Im Schulhausflur konnten oft Gruppen angetroffen werden, die zusammen eine Szene spielten, sei dies im Fach Deutsch, Französisch oder Englisch. Ein attraktives und sehr aktives Theaterfreifach plante regelmäßig zusammen mit meinem Schulchor Aufführungen, nicht nur als Schulanlass, sondern auch in Zusammenarbeit mit Gesangsvereinen als Kulturanlass im Dorf. So kann im Rückblick die Gesangspräsentation vor der Klasse als akzeptabler Rahmen einer performanceförderlichen Schulhauskultur verstanden oder mit dem heutigen Vokabular als Förderung der Auftrittskompetenz bezeichnet werden. Damals war ich als Lehrer mit dem Setting zufrieden, heute sehe ich als Fachdidaktiker Gesprächsbedarf, weiß aber nicht, wie in zwanzig Jahren die beiden Positionen bewertet werden.

Das außergewöhnliche Setting wurde mir erst bewusst, als ich nach dem Wechsel an eine andere Schule das zu Beginn beschriebene Verfahren nicht mehr anwenden konnte, weil der Widerstand in der Schülerschaft zu groß war. Die im schulischen Gesamtsystem etablierte Auftrittskompetenz war an der neuen Schule nicht vorhanden, sondern musste zuerst aufgebaut und trainiert werden. Der Wechsel an eine andere Schule wurde so zum Wendepunkt in meiner Laufbahn als Musiklehrer³². Von einem vielleicht

³² Zu diesem Wechsel hat nicht nur der Schulwechsel beigetragen, sondern auch eine Auszeit, in der ich mich vertieft mit meinem Unterricht auseinandersetzen konnte. Ein Besuch in Bremen bei Dr. Andreas Lehmann-Wermser und

unreflektierten, aber zumindest gelingenden Musikunterricht in einer positiven und auftrittsförderlichen Schulkultur vollzog sich mit der Zeit ein Wechsel hin zu reflektierten Unterrichtssettings mit aufbauenden Strukturen.

Fazit: Der eigene Singunterricht als Ausgangspunkt

In der Analyse der ersten Feldnotiz trifft die vom aktuellen Zeitgeist geprägte Analyse des Fachdidaktikers auf die Beschreibung einer Unterrichtssituation, die Jahre zurückliegt. Die von der Lehrkraft positiv erlebte Bewertungskultur wirkt durch die Einbettung in die Kultur der Schule plausibel, wird aber vom Fachdidaktiker in verschiedener Hinsicht hinterfragt. In dieser Situation stellt sich die Frage, wer die Deutungshoheit erhalten soll. Zinnecker (2000) schreibt aus ethnografischer Sicht zu dieser speziellen Situation im pädagogischen Bereich: „Gebührt die ethnographische Autorität in Sachen Schule dem externen Sozialwissenschaftler, sei er Anthropologe, Soziologe oder Psychologe, dem Erziehungswissenschaftler, dem Didaktiker, dem Lehrer, oder etwa dem Schüler?“ (S. 383) Zinnecker macht mit seiner Frage deutlich, dass in der Betrachtung oder Wertung von Unterrichtssituationen die Perspektive eine wesentliche Rolle spielt und stellt infrage, ob dies zwingend der Blick aus der Forschung, oder wie oben der Blick der Fachdidaktik sein muss. Bezogen auf die vorliegende Arbeit stellt sich konkret die Frage, ob der Blick des Lehrers, der zu Beginn des Kapitels retrospektiv die Unterrichtssituation als gelungene Leistungsbeurteilung darstellt, oder der Blick des Fachdidaktikers, der später im Text die Feldnotiz eher kritisch einordnet, mehr Gewicht erhalten soll³³? Wenn später in dieser Arbeit die Interviews durch die Brille von Assessment analysiert werden, darf die Frage der Deutungshoheit nicht unterschätzt werden, weil jede Lehrkraft aus einer lokalen und biografisch geprägten Perspektive heraus argumentiert. Um dies entsprechend wertzuschätzen, werden diese Sing- oder Schulkulturen später in kurzen Porträts

Dr. Anne-Katrin Jordan – ENTWICKLUNG UND VALIDIERUNG EINES KOMPETENZ-MODELLS IM FACH MUSIK: WAHRNEHMEN UND KONTEXTUALISIEREN VON MUSIK (Jordan, 2012) – hat mich maßgeblich beeinflusst.

³³ Zusätzlich macht das Zitat von Zinnecker deutlich, dass in diesem Exkurs der Blick von außen und die Perspektive der Lernenden ausgeblendet werden.

gewürdigt, im Bewusstsein, genau wie in der obigen Analyse eine externe Deutung vorzunehmen (vgl. 12.1). Der Wechsel der Perspektive hin zur Analyse von Assessment-Praktiken erfolgt dann nicht mit dem normativen Blick des Fachdidaktikers, sondern mit dem Anspruch einer thematisch geleiteten Situationsanalyse.

In der autoethnografischen Methode verliert die Frage der Deutungshoheit aber vorerst an Wichtigkeit, weil die Darstellung evokativer Momente ins Zentrum rückt. Stahlke Wall (2016) beschreibt das gemeinsame Anliegen autoethnografischer Arbeiten folgendermaßen:

„[...] they all refer in some way to a *systematic* approach using ethnographic strategies, the *linking* of personal experience to social, cultural, and political issues, and a *critique* of certain discourses within a cultural context with a vision and hope for change [...].“ (S. 5, Hervorhebung im Original)

Im vorliegenden Fall bezieht sich die Systematik auf die Auswahl von prägenden Situationen des Singunterrichts der forschenden Person, deren Einbettung in einen schulischen Kontext und das persönliche Erleben innerhalb dieser Entwicklung. Die Hoffnung, mit der vorliegenden Arbeit zu einer positiven Veränderung beizutragen, ist im Ausblick in den Thesen zum Assessment des Singunterrichts (vgl. Kapitel 14) präsent. Inwiefern sich Autoethnografie als wissenschaftliche Methode etablieren konnte, wird im nächsten Abschnitt dargelegt. Zuvor wird aber eine weitere Szene aus dem Unterricht beschrieben, diesmal aus einer späteren Zeit, rund um die 2010er-Wende.

3.2 Punkt 2: Autoethnografie im Diskurs

Interludium I: Eine verpasste Chance

Die Schülerin ziert sich, will nicht vor ihren Mitschülerinnen vorsingen. – Sie hat sich vier Jahre lang versteckt, hat abgewiegt, sich wenn möglich gedrückt. Sich selbst zu präsentieren vor anderen, auch wenn es nur die Mädchen waren, kam einfach nicht infrage. Sie hat sich in ihrer Position kaum bewegt in dieser Zeit, kaum Schritte gemacht und hat Unterstützung nicht zugelassen. – Doch der Reihe nach. Der Lehrer bemüht sich mittlerweile, in den vier Schuljahren auf der Oberstufe Auftrittskompetenzen aufzubauen. Die Schülerinnen und Schüler starten im ersten Jahr mit dem Vorsingen in größeren Gruppen, trainieren im zweiten und dritten Jahr das Vorsingen zu zweit vor dem eigenen Geschlecht, Jungen vor Jungen und Mädchen vor Mädchen, und werden dazu regelmäßig stimmlich coacht. Die Geschlechtertrennung während der Performance ist der Kompromiss, den der Lehrer mit den Schülerinnen und Schülern in der neuen Schule ausgehandelt hat. Während der vier Jahre wechseln sich Phasen von Klassensingen auf der Bühne, Gruppensingen mit Coaching und individueller Betreuung ab. Abschließend wird im vierten und letzten Jahr der Bezirksschule ein individueller Song mit YouTube-Playback geprobt und aufgeführt, wiederum geschlechtergetrennt und begleitet von einem mehrmaligen Coaching. Viele freuen sich darauf, es gibt gegenseitige Bewunderung, motivierende Kommentare und auf jeden Fall viel Applaus. Doch nun zurück zu der erwähnten Schülerin. Sie kann der ganzen Sache wenig abgewinnen und spielt mit dem Gedanken, sich der Aufgabe mit dem Karaoke-Auftritt zu widersetzen. Die Schülerin hat immer eine Ausrede bereit, wenn es ums Üben geht. Der Lehrer motiviert fortlaufend, fragt nach, zeigt aber auch die Konsequenz der Verweigerung auf, nämlich eine ungenügende Note. Eine Alternative wird nicht angeboten. Der Tag der Performance naht und der Lehrer befürchtet, dass die Schülerin sich weigert. Aber das trifft nicht ein. Die Schülerin gibt sich einen Ruck, singt allein zum Playback. Ihre Befürchtung, die eigene Leistung wäre grotten-schlecht, erweist sich als falsch. Ihre Leistung ist ordentlich und alle sind erleichtert. Anschließend insistiert die Schülerin auf eine zweite Chance, wird aber vom Lehrer abgewimmelt mit der Begründung, sie hätte so viele

Chancen gehabt, jetzt sei es zu spät, seine Geduld sei abgelaufen. Der Lehrer ist überzeugt, damit eine klare Linie vorzugeben.

Methode: Die eigene Erleuchtung als Maßstab?

Nachdem unter Punkt 1 die Notwendigkeit diskutiert wurde, sich im Forschungsprozess sichtbar zu machen, wird nun die gewählte Methode besprochen. Ellis et al. (2010) beschreiben Autoethnografie entlang ihrer etymologischen Bedeutung. Die Autoethnografie ist „darum bemüht, persönliche Erfahrung (auto) zu beschreiben und systematisch zu analysieren (grafie), um kulturelle Erfahrung (ethno) zu verstehen“ (Ellis et al., 2010, S. 345). Die Methode wird einerseits auf autobiografische Ansätze mit einer literarischen Prägung und andererseits auf ethnografische Verfahren mit einer wissenschaftlichen Orientierung zurückgeführt³⁴. Ellis bezeichnet Autoethnografie zudem nicht nur als Forschungsmethode, sondern auch als deren Produkt.

Das Feld der Autorinnen und Autoren, die sich zu Autoethnografie äußern, ist überblickbar (Ellis & Bochner, 2000; Anderson, 2006; Denzin, 2006; Stahlke Wall, 2016; Holman Jones et al., 2016b). Der Methodendiskurs innerhalb dieses Feldes zeigt Divergenz (Denzin, 2006, S. 420) und wird teilweise polemisch ausgetragen, was wiederum als Merkmal der Autoethnografie verstanden werden kann. Unter sozialwissenschaftlich Forschenden wird Autoethnografie aber eher kritisch bewertet, wie Ellis et al. (2010) selbst ausführen und sich später auch bezüglich Musikpädagogik zeigen wird.

Als zentrale Figuren können Ellis und Bochner³⁵ identifiziert werden (Stahlke Wall, 2016). Sie schreiben nach einer ersten Etablierungsphase der Autoethnografie im *HANDBOOK OF QUALITATIVE RESEARCH* (Denzin & Lincoln, 2000) den Artikel *AUTOETHNOGRAPHY, PERSONAL NARRATIVE, REFLEXIVITY: RESEARCHER AS SUBJECT* (Ellis & Bochner, 2000). Auf der einen Seite grenzen sie darin die Methode konkret von Ethnografie ab,

³⁴ An dieser Stelle gilt es anzumerken, dass die deutschsprachige Ethnografie in den Erziehungswissenschaften eher einen szientistischen Weg geht und englischsprachige Texte durchaus literarische Ansätze pflegen (Geimer, 2011).

³⁵ Carolyn Ellis und Arthur Bochner sind liiert und waren an der University of South Florida tätig. Beide sind mittlerweile emeritiert.

insbesondere mit dem Argument, die eigene Person als Phänomen zu sehen und Geschichten, also „evocative personal narratives“ (S. 737), zu schreiben. Dies wird als Gegensatz zu einer reflexiven Ethnografie (realist ethnography) gesehen, in welcher „the researcher’s personal experience becomes important primarily in how it illuminates the culture under study“ (S. 736). Auf der anderen Seite wird im Artikel deutlich, wie unterschiedlich sich die Zugänge in den zehn vorangegangenen Jahren (vor 2000) entwickelt haben. Sie schreiben: „meanings and applications of autoethnography have evolved in a manner that makes precise definition and application difficult.“ (S. 739) Ellis und Bochner können so als Promotoren betrachtet werden, die eine wenig definierte Methode im wissenschaftlichen Diskurs etablieren wollen. Das zunehmende Interesse an der Methode belegt Stahlke Wall (2016) mit der starken Zunahme von publizierten Texten. So wurden zwischen 1990 und 2003 durchschnittlichen 5–10 autoethnografische Texte publiziert und nach 2003 jährlich ungefähr 35 Texte (S. 2 f.).

Die Autorin und der Autor widersetzen sich zudem im oben erwähnten Handbuchartikel der traditionellen formalen Gestaltung und schreiben bewusst keinen Essay, sondern wählen die Form eines fingierten Telefongesprächs. Sie begründen folgendermaßen:

„But in this piece we want to *show*, not just tell *about* autoethnography. [...] The conventions militate against personal and passionate writing. These books are filled with dry, distant, abstract, propositional essays.“ (Ellis & Bochner, 2000, S. 734, Hervorhebung im Original)

Die Wahl des Schreibstils führt nun zum Kern der Autoethnografie, die von Spry (2001) als Widerstand gegen „Grand Theorizing and the facade of objective research“ (S. 710) beschrieben wird. Der Widerstand gegen akademische Gepflogenheiten folgt der Überzeugung, nicht objektiv und wertfrei sein zu können. Eingetreten wird für eine Gesellschaft, die gerecht, demokratisch und egalitär ist. Die Schreibenden sind sich bewusst, damit pädagogisch und politisch zu sein (Denzin, 2006, S. 422; Holman Jones, 2005). Auf dem Weg dazu sollen nicht trockene, distanzierte und abstrakte Essays geschrieben werden, sondern leidenschaftliche und persönliche Geschichten. Diese Haltung ist als Reaktion auf die „*crisis of confidence*“ (Ellis et al., 2010, S. 345, Hervorhebung im Original) im Postmodernismus der 1980er-Jahre einzuordnen (vgl. auch Stahlke Wall, 2016,

S. 2). Die ersten Autoethnografinnen und Autoethnografen bewegen sich in einem intellektuellen Umfeld, wo „postmodern, poststructuralist, and feminist writers“ (Holman Jones et al., 2016a, S. 18) sich gegen die wissenschaftlichen Standards auflehnen:

„Viele dieser Wissenschaftler/innen wendeten sich der Autoethnografie zu, da sie sich um eine positive Antwort auf die Kritik an den kanonischen Vorstellungen, was Forschung ist und wie Forschung betrieben werden sollte, bemühten. Im Besonderen wollten sie sich auf Möglichkeiten konzentrieren, bedeutsame, zugängliche und sinnhafte Ergebnisse hervorzubringen, die auf persönlicher Erfahrung gegründet und respektvoll gegenüber fremder Erfahrung ist (Ellis & Bochner 2000).“ (Ellis et al., 2010, S. 345)

Ellis und Bochner (2006) sehen in ihrer Methode eine Überwindung des „kolonialistischen“ (S. 436) Forschungsbetriebes, der mit der Art und Weise des Schreibens die Machtverhältnisse zwischen Wissenden und Unwissenden zementiert (S. 436). Anderson (2006) begrüßt den Enthusiasmus, mit dem Ellis und Bochner die Methode forcieren, sieht aber die beschwörende, suggestive und emotionale Art der Texte als Gefahr für eine produktive Einbettung in die Traditionen der Sozialwissenschaften. Evokative Autoethnografie „may have the unintended consequence of eclipsing other visions of what autoethnography can be and of obscuring the ways in which it may fit productively in other traditions of social inquiry“ (S. 374). Er skizziert deshalb Schlüsselmerkmale für eine analytische Autoethnografie (S. 378), die einen besseren Anschluss an akzeptierte Methoden bieten sollen. Anderson schlägt vor, im Forschungsprozess von den Daten und Interpretationen erster Ordnung zu den „more abstract, transcontextual, second-order constructs of social science analysis“ (S. 381) zu kommen. Er verlangt eine größere Abstraktion, mehr als das Hervorheben eines Umstands in einer Geschichte und schreibt: „I use the term analytic to point to a broad set of data-transcending practices that are directed toward theoretical development, refinement, and extension.“ (S. 387)

Die Suche nach einer Anschlussfähigkeit an bestehende Kontexte sozialwissenschaftlicher Forschung wurde denn auch von Ellis und Bochner (2006) nicht unwidersprochen hingenommen. In der von ihnen präferierten bildhaften Sprache starten sie ihre Replik mit einem Beitrag aus den

News, den sie kommentieren: „His words capture the misery so many people [...] are going through. They pull me into the immediate shock he is feeling [...].“ (S. 430) Sie betonen, wie bewegt sie von der persönlichen Geschichte des Geschädigten eines Wirbelsturms sind. Dies führt dann wiederum in Form eines Gesprächs zwischen Ellis und Bochner zur Argumentation, dass die Lesenden von autoethnografischen Texten dazu gebracht werden sollen, „to care, to feel, to empathize, and to do something, to act“ (S. 433). Die Forschenden wiederum müssen dazu „vulnerable and intimate“ (S. 433) sein. Ellis und Bochner erwarten von Forschenden, sich durch ihre Texte verletzlich zu machen, um die Lesenden emotional zu berühren und sogar zum Handeln zu bewegen. Der von Anderson geforderte analytische und theoretisierende Zugang ist für Ellis und Bochner nicht kompatibel mit der ursprünglichen Idee der Autoethnografie, denn mit der Forderung nach Anschlussfähigkeit korrumpiere Anderson die Widerständigkeit, welche die beiden fordern.

So wehren sie sich auch dagegen, von *evokativer Autoethnografie* als methodischem Begriff zu sprechen, vielmehr sei die Evokation das Ziel: „[T]he term ‘evocative’ appeared because readers of autoethnography recognized that its evocative quality was one of the characteristics that made autoethnography a distinctive genre of ethnographic writing.“ (436) Ellis und Bochner definieren das Eintreten der Evokation als Qualitätsmoment. Douglas und Carless (2016) gehen sprachlich noch einen Schritt weiter, indem sie autoethnografische Texte Epiphanien³⁶ nennen. Der *erleuchtende* Charakter dient ihnen als Legitimation der Autoethnografie als Forschungsmethode:

„This kind of striking experience or epiphany can, in our experience, cut through the clutter of day-to-day life as a researcher, academic, or student. In so doing it has the potential not only to inspire others but also to help legitimize autoethnography as a valid, useful, and important way of doing social research.“ (Holman Jones et al., 2016b, S. 92)

Diese Art der Legitimation erinnert an ein Perpetuum mobile, das seine Energie aus sich selbst gewinnt. Auch Strübing et al. (2018) beschreiben

³⁶ Dieser in der christlichen Religion gebräuchliche Begriff umschreibt eine Erscheinung und wird im übertragenen Sinn als Erleuchtung benutzt.

für die Tradition der qualitativen Sozialforschung eine schriftliche Performanz als Gütekriterium, allerdings nicht um Erleuchtung zu evozieren, sondern um Erkenntnisse zu vermitteln. Neben Anderson (2006) gibt es noch weitere Autorinnen, die einen eher analytischen Zugang zu Autoethnografie suchen (Chang, 2008; Snyder, 2015), insgesamt bleiben aber Ellis und Bochner prägend. So sind sie zusammen mit Holman Jones und Adams auch maßgeblich für das *HANDBOOK OF AUTOETHNOGRAPHY* (Holman Jones et al., 2016b) verantwortlich (Stahlke Wall, 2016, S. 5). In diesem Handbuch fassen in einer Art Synthese Anderson und Glass-Coffin (2016, S. 71 ff.)³⁷ fünf wesentliche Merkmale von Autoethnografie zusammen. Diese werden nun abschließend vorgestellt:

- **Sichtbarkeit des Selbst – Visibility of Self:** Der Schlüssel der autoethnografischen Untersuchungsmethode ist die Sichtbarkeit des Selbst in der Forschung: „The ethnographic ‚I‘ is visible as an actor and an agent, acting and reacting visibly to the implications and the consequences of inquiry.“ (S. 71) Dabei wird bewusst Objektivität und Subjektivität vermischt: „[A]utoethnographic inquiry rejects the distinction between the subjective and the objective, seeking, as Davies (1999) puts it, ‚to develop forms of research that fully acknowledge and utilize subjective experience as an intrinsic part of research‘ (p. 5).“ (S. 72) Mit der Sichtbarkeit des Selbst wird in der Autoethnografie die eigene Erfahrung als wertvoll anerkannt und zum Objekt der Forschung erhoben.
- **Starke Reflexivität – Strong Reflexivity:** Durch die Betrachtung und Analyse der eigenen Aktionen und Wahrnehmungen sowie mit Blick und im Dialog mit den anderen kann eine starke Reflexivität entwickelt werden.
- **Engagement:** Als „complete member researcher“ (S. 74) ist der Forschende im Feld ein aktives Mitglied und stellt seine Gedanken dort auch zur Diskussion.
- **Verletzlichkeit – Vulnerability:** Durch die Emotionalität und die Präsentation der eigenen Erfahrungen als *Erleuchtung* (Epiphanie) wird die Verletzlichkeit betont.

³⁷ Zusammenfassung und Übersetzung der Begrifflichkeiten durch den Autor.

- **Unbegrenztheit/Ablehnung der Endgültigkeit und des Abschlusses**
 - **Open-endedness/Rejection of Finality and Closure:** Auf der einen Seite wird betont, dass autoethnografische Texte auf bestimmte Momente bezogen sind, was selbstverständlich auch für viele sozialwissenschaftliche Forschungsprojekte (Gläser & Laudel, 2010) gilt. Auf der anderen Seite wird aber auch der Aspekt betont, die Methode selbst dürfe nicht begrenzt werden.

Auch Stahlke Wall (2016) diskutiert evokative und analytische Autoethnografie und schlägt eine Synthese vor. Er plädiert für eine *moderate Autoethnografie* (moderate autoethnography): „A moderate autoethnography would reconcile the best of these ideas and combine the power of the personal perspective with the value of analysis and theory, so that sociological understanding is advanced in ways it might never have otherwise been.“ (S. 8)

Im vorliegenden autoethnografischen Exkurs wird das subjektive Erleben des Musiklehrers durch die Beschreibung von Vignetten als wichtiges Moment anerkannt. Durch die Analyse aus der wiederum subjektiven Wahrnehmung des Fachdidaktikers wird eine Analyse ermöglicht, die über das evokative Moment von Feldnotizen hinausgeht. Beide Sichtweisen sind jeweils Momentaufnahmen und machen den Forschenden angreifbar, zeigen aber auch sein Engagement im Feld. Das Kapitel positioniert sich damit als moderate Autoethnografie.

Analyse: Ziel erreicht, oder doch verfehlt?

Am Ende der zweiten Feldnotiz zeige ich mich zufrieden mit meiner klaren Linie. Aber ich erinnere mich gut, dass ich damals auch Unsicherheit verspürte. In meiner damaligen Optik hat diese Schülerin nach viel Engagement von meiner Seite endlich, zugegebenermaßen unter Druck, einen Schritt gewagt und sich vor der Klasse präsentiert. Aus ihrer Optik hat sie wohl nach vielen inneren Diskussionen ihre Hemmungen überwunden und sich in die Klassenkultur eingefügt. Unsicherheit verspürte ich, weil ich dieser Schülerin nicht die Chance gegeben hatte, sich in der noch kurzen verbleibenden Schulzeit weiterzuentwickeln. Die Logik der summativen Beurteilung, die alle Schülerinnen und Schüler erst einmal den gleichen Leistungsanforderungen aussetzt, hat mir den Blick verstellt für den individuellen Lernweg und die spezifischen Bedürfnisse der Schülerin.

Die zweite Feldnotiz zeigt eine Beurteilungssituation, die sich stark von der Vorsing-Situation in Punkt 1 (vgl. 3.1) unterscheidet. Die kontinuierliche Arbeit an der Auftrittskompetenz folgt der Kompetenzbeschreibung des Lehrplan 21 (D-EDK, 2015)³⁸ und lässt Raum für individuelles Coaching, Feedback und Eigenverantwortung. Die Idee, in immer kleiner werdenden Gruppen oder sogar allein vor der Klasse zu stehen, wurde von den Schülerinnen und Schülern gut aufgenommen und mit guten Resultaten quittiert³⁹. Unter Berücksichtigung der Effekte der adoleszenten Stimmentwicklung (vgl. 1.2) zeigt sich im vorgestellten Förderkonzept trotzdem eine Lücke. Es fehlt im Unterricht eine kontinuierliche reflexive Ebene, die es den Schülerinnen und Schülern ermöglicht, das eigene Singen von außen zu betrachten und in Bezug zu setzen mit der Persönlichkeitsentwicklung. Mögliche Fragen wären z. B.: Warum will ich nicht vor anderen singen? Warum freue ich mich auf die Bühne? Warum fühle ich mich unwohl beim Singen? Dies wäre z. B. möglich in Form eines Portfolios (Morrison, 2005; Winter, 2015).

Noch einmal zurück zu dieser Schülerin: Wie sich zeigte, hatte diese Schülerin von Anfang an kein Singproblem, sondern eine innere Blockade. Vielleicht hätte ihr mit einer stetigen Selbstreflexion bzw. einer aktiven Bearbeitung ihres Selbstkonzepts (Spychiger, 2013) und der Entwicklung ihres *possible self* (Markus & Nurius, 1986) besser geholfen werden können. Vielleicht hätte ich als Lehrer durch ein Nachdenken über die Heterogenität in der Schülerschaft (Hoene & Thurman, 2011; Linn, 2017) mehr Empathie zeigen können.

³⁸ Der LP21 war in dieser Zeit zwar nicht eingeführt, lag aber in den Grundzügen bereits vor. Als Mitglied der Begleitkommission hatte ich Einsicht. Interessanterweise ist Auftrittskompetenz eingeordnet im Kompetenzbereich Gestaltung und nicht direkt beim Singen und Sprechen.

³⁹ Integraler Bestandteil war jeweils eine Selbstbewertung in Form eines Rasters, das anschließend von mir ausgefüllt wurde. Meistens lagen die Beurteilungen gar nicht so weit auseinander. Und wenn doch, ergab sich die Möglichkeit, dies als Anlass für ein Gespräch zu nehmen, um eine Reflexion bei Schülerinnen und Schülern auszulösen. Selbstbeurteilung wird in der Musikpädagogik eingesetzt (vgl. Biegholdt et al., 2020), wird aber in seiner Wirkung kaum diskutiert.

Fazit: Erleuchtung als Ausgangspunkt für Entwicklung

Die beschriebene Szene in der zweiten Feldnotiz zeigt einen erleuchtenden Moment für die Lehrkraft und damit ein Merkmal evokativer Autoethnografie. Es bleibt aber nicht bei der Erzählung und Einordnung, sondern es wird weiter analysiert unter fachdidaktischen Gesichtspunkten, was an analytische Autoethnografie anschließt. Die autoethnografische Methode ermöglicht an dieser Stelle das Herausarbeiten von Entwicklungsmöglichkeiten. Insgesamt bleibt der Abstraktionsgrad in den Aussagen des Lehrers und den Analysen des Fachdidaktikers aber noch gering. Interessanterweise ist genau dies der Kritikpunkt, der Orgass (2017) zur Ablehnung dieser Methode in der Musikpädagogik bewegt. Dieser Aspekt wird im nächsten Abschnitt ausgeführt, zusätzlich zur Suche nach einem höheren Abstraktionsgrad.

3.3 Punkt 3: Autoethnografie in der Musikpädagogik

Interludium II: Die Geschichte vom Dankeschön eines Jungen

Szenenwechsel. Der gleiche Lehrer unterrichtet an einer kleinen Kreisschule mit Sekundar- und Realschulklassen⁴⁰, den tieferen Leistungsniveaus im Kanton Aargau. In einer Klasse gibt es sieben Jungen und sechs Mädchen, wobei deren vier das Schuljahr im höheren Leistungsniveau wiederholen. Die vier Mädchen singen gerne und gut und hatten bereits ein Jahr Musikunterricht bei diesem Lehrer. Die Jungen sind nicht besonders motiviert mit ihrer Stimme zu arbeiten und im Verlauf des Schuljahres rutschen alle außer einem Jungen in den Stimmbruch, was wiederum nicht dazu beiträgt, die allgemeine Motivation zu verbessern. Während die Mädchen ein neues Lied in Angriff nehmen, arbeitet der Lehrer nun sporadisch mit den Jungen in einem Nebenraum. „Lasst uns nochmals den Chorus von Astronaut⁴¹ singen.“ Schnell etabliert sich in der Jungengruppe eine

⁴⁰ Siehe Kapitel 6.

⁴¹ Astronaut von Sido feat. Andreas Bourani (<https://www.youtube.com/watch?v=WPFLAjmWCtk> – abgerufen am 03.05.2019).

Stimmarbeit, wo diese den aktuellen Stand ihrer Mutation zeigen können. Natürlich sind sie zuerst irritiert, weil sie einzeln vorsingen sollen, aber es funktioniert. Gemeinsam wird aus dem Repertoire oder einem aktuell erarbeiteten Song eine Verszeile gesungen, um diese anschließend individuell zu wiederholen. Die Lehrkraft sucht dazu mit jedem Jungen eine adäquate Stimmlage, um einen gesanglichen Erfolg zu schaffen. „Warte kurz Ralf, ich begleite einen Ton tiefer, dann kannst du alle Töne treffen.“ Bei dieser Gelegenheit wird der Entwicklungsstand der Mutation angesprochen und die Jungen erhalten einen empathischen Kommentar zu ihrer als misslich empfundenen Lage. Am Ende der Stunde bedankt sich ein Schüler beim Hinausgehen spontan. Später im Jahr, bei der unumgänglichen Bewertung für die Zeugnisnote wird mit der Klasse abgemacht, dass ein freier Vortrag aus dem Repertoire in Zweiergruppen vorgesehen ist. Beim Vorbereiten der kleinen Performance wird klar, dass die Jungen ihre Wahl ungeschickt getroffen haben und eine gemeinsame Tonart in den Zweiergruppen nicht funktioniert. Die Jungen sind aber einverstanden, einzeln in ihrer Tonart einen Vers zu singen und abschließend gemeinsam in den Chorus einzustimmen. Dabei nimmt der Lehrer in Kauf, dass der Chorus ungewollt mehrstimmig wird. Die Mädchen setzen sich das eine oder andere Mal für eine eigene Übungsphase ein und erhalten diese auch.

Methode: Autoethnografische Verfahren in der Musikpädagogik

Dieses Kapitel bricht mit dem autoethnografischen Zugang gewollt die traditionelle Struktur einer empirischen Studie der Musikpädagogik auf. In Punkt 3 werden nun autoethnografische Zugänge in der Musikpädagogik untersucht. Nach aktuellem Wissensstand wurden mit einer Ausnahme keine deutschsprachigen autoethnografischen Texte für die Musikpädagogik vorgelegt. Auf den Umstand, dass Autoethnografie im musikpädagogischen Diskurs bis jetzt keine Rolle spielt, verweist neben Roulston (2006) aktuell auch Orgass (2017). Er bezeichnet diese Methode im Rahmen seiner „präventiven Reflexion“ (S. 132) explizit als „Nicht-Forschung“ (S. 120), weil „die Unterscheidung zwischen Konstruktionen erster Ordnung der Beforschten und zweiter Ordnung der Forscherinnen und Forscher“ (S. 133) wegfällt. Aus seiner Sicht fehlen in der Methode die Reflexionslogik und die Distanz zum Objekt der Forschung, die er für die wissenschaftliche Musikpädagogik einfordert. Er folgt damit der Tradition

des „klassischen Modells einer ‚Natur-Wissenschaft‘“ (Zinnecker, 2000, S. 392), wo Biografie und Person von Forschenden möglichst ausgeklammert werden. Orgass formuliert damit eine prinzipielle Kritik und geht bedeutend weiter als Anderson (2006), der lediglich eine analytische Komponente in der Autoethnografie einfordert.

Wie bereits oben erwähnt, konnte nur ein deutschsprachiger autoethnografischer Text gefunden werden. Der Ethnologe Eckhardt (2019) nähert sich in Form von Essays seiner Chorleitungstätigkeit aus verschiedenen Perspektiven und ordnet diese Textsammlung selbst explizit der Autoethnografie zu. Thematisch kann dieser Text nur lose in die Musikpädagogik eingeordnet werden, da darin Chorgesang von Laienchören und nicht musikpädagogische Phänomene als kulturelle Praxen besprochen werden. Interessant ist er deswegen, weil sich der Autor schreibend auf die Suche nach eigenen Positionen macht und sich dabei gegen die, wie er argumentiert, vermeintlich entpersonalisierten theoretischen Wissenschaftskonzepte auflehnt. Musik, in seinem Fall Chorsingen, sei per se als Forschungsgegenstand nicht von „ich mache Musik“ (S. 13) zu trennen. Wie die bereits referierte Aussage von Hoos de Jokisch (2014) „[i]ch singe, (al) so bin ich“ (S. 6) in Bezug auf das Singen, beschreibt Eckhardt eine existenzielle Verbundenheit des Forschungsgegenstands mit der Person. Diese Verbundenheit, so argumentiert er, kann mit der Autoethnografie adäquat adressiert werden. So zeichnet sich sein Text durch wesentliche autoethnografische Methoden nach Anderson und Glass-Coffin (2016) aus: er zeigt sich selbst (*visibility of self*), demonstriert Reflexivität (*strong reflexivity*), engagiert sich als Chorleiter selbst stark für das Chorwesen (*engagement*) und macht sich mit der Darlegung eigener Unsicherheiten verletzlich (*vulnerability*). Sein Text ist zudem stark evokativ geprägt und kann deshalb in die ursprüngliche Methodik von Ellis und Bchner (2000) eingeordnet werden. Im Gegensatz zu Orgass (2017) vertritt er die Meinung, dass eine „Autonomie des Wissenschaftssystems“ (S. 133) nicht als Ziel angestrebt werden muss. Die Autoethnografie von Eckhardt zeigt das Potenzial für ein Forschungsgebiet auf, wo Persönlichkeit und Emotionalität direkt mit dem Forschungsgegenstand, im vorliegenden Fall dem Singen, verknüpft sind.

Im anglofonen Raum dreht Gouzosis (2014) den Spieß ebenfalls um und fordert Geschichten „from the perspectives of music teachers and learners using their own voices“ (S. 4), um der „suffocation orthodoxy“

(S. 3 f.), die seiner Meinung nach die Forschung in der Musikpädagogik konstituiert, zu entkommen. Ausgehend von Ellis (2004) will er mit der Methode der Autoethnografie „a teacher and learner centered sociology, or sociosophy, of music education“ (S. 5) erreichen. Auch er bricht damit aus den traditionellen soziologischen Forschungsmethoden aus, um die Nähe zum beforschten Feld und nicht zuletzt soziale Gerechtigkeit zu suchen, ganz in der Manier der Initianten der Autoethnografie, Ellis und Bochner. Diesen exemplarischen Betrachtungen können eine ganze Reihe von autoethnografischen Arbeiten im musikpädagogischen Forschungskontext nachgereiht werden, sei das im Bereich der Professionalisierung (z. B. Manovski, 2014; Nethsinghe, 2012), oder im Bereich von Musikunterricht (z. B. Frierson-Campbell & Newman, 2008; Thompson, 2015). Bereits erwähnt wurde die Autoethnografie von de Vries (2012), der seine musikpädagogische Laufbahn in einem Zwiegespräch reflektiert.

Wer nach autoethnografischen Texten in der deutschsprachigen Musikpädagogik sucht, wird nicht fündig. Schulten und Lothwesen (2017) erwähnen in ihrer Schrift zu empirischen Methoden in der Musikpädagogik weder ethnografische noch autoethnografische Verfahren. Das Wort *Autoethnografie* findet sich auch im HANDBUCH MUSIKPÄDAGOGIK (Dartsch et al., 2018) nicht. Niessen schreibt aber, dass „ein kaum genutztes Potenzial in ethnografischen Forschungsmethoden stecke“ (S. 430). Knigge und Niessen öffnen die Perspektive weiter hin zu Musikethnologie und Kulturwissenschaften, die „eine Fülle von Anregungen für die Musikpädagogik“ (S. 16) liefern können, was sich z. B. in der bereits mehrfach zitierten Studie HEGEMONIE IM MUSIKUNTERRICHT von Blanchard (2019) zeigt. Seine kulturwissenschaftlichen Standpunkte und Beispiele aus der ethnografischen Studie bieten einiges an Irritationspotenzial auch für die vorliegende Arbeit.

Analyse: Reflektierte und empathische Betreuung

Die beschriebene Betreuungssituation der Jungen zeigt eine weitere Stufe meines Werdegangs auf. Ich bin an meinem dritten längerfristigen Unterrichtsort angekommen, arbeite parallel als Dozent in der Fachdidaktik und

habe begonnen, meinen Unterricht aufgrund von gelesener Literatur⁴² neu zu denken. Mein Selbstverständnis wird in diesen Jahren stark irritiert, als künstlerischer Praktiker, wie Huber (2016) die Musiklehrkräfte in der Schweiz nennt, war mein pädagogisches Knowhow marginal. Zusätzlich war in dieser Lebensphase auch der Nimbus des jungen Lehrers definitiv verloren gegangen. In der Feldnotiz zeigen sich drei inhaltliche Elemente, die mir zentral erscheinen: Erstens pflege ich die Beziehung zu den Jugendlichen, zweitens rückt die adoleszente Stimme in den Fokus, was drittens zu einer partiellen Trennung der Geschlechter führt.

Bei der Arbeit mit den unteren Leistungsniveaus wird schnell klar, dass mein Künstlertum bei Weitem nicht ausreicht, einen guten Musikunterricht zu machen. Schülerinnen und Schüler sind nicht empfänglich für meine musikalischen Qualitätsansprüche und mein gepflegtes Klavierspiel. Sie müssen vermehrt dazu überlistet werden, musikalisch tätig zu werden, auch wenn sie immer wieder fragen, ob in der Stunde gesungen werde. Mit der Zeit wird klar, dass die Pflege der Beziehung gerade auf diesen Stufen von zentraler Bedeutung ist. Als wichtiger Trigger für gelingende Beziehungen diente mir das Bändchen *LOB DER SCHULE* von Joachim Bauer (2008), in dem er zeigt, dass Lehrkräfte dafür verantwortlich sind, die biologischen Motivationssysteme in Betrieb zu nehmen⁴³. Insbesondere seine Verknüpfung von Motivation mit der Funktion der Spiegelneuronen hat mich fasziniert. So ist es laut Bauer nicht möglich, im Unterricht eine distanzierte Beziehung zu pflegen, reagiert wird immer direkt. Aus dieser Einsicht begann ich vermehrt an einem einvernehmlichen Klassenklima zu

⁴² Diese Aussage muss im Licht der Tradition der Ausbildungsgänge für Schulumusik der 1990er-Jahre in der Schweiz gesehen werden. Ich kann mich nicht erinnern, viele fachdidaktische Texte gelesen zu haben oder in der Planung von Unterricht auf fachdidaktische Konzepte Bezug genommen zu haben (vgl. Niessen, 2014). Mein langjähriger Praxisbezug verhalf mir aber trotzdem zu einer Anstellung als Fachdidaktikdozent an einer pädagogischen Hochschule.

⁴³ „Gemeinsam bilden die Leistungsdroge Dopamin, die Wohlfühlsubstanzen aus der Gruppe der Opioiden und das ‚Freundschaftshormon‘ Oxytocin ein geradezu geniales Trio.“ (Bauer, 2008, S. 21) Er schlägt dann eine Brücke zu Ryan und Deci (2000), allerdings ohne diese zu erwähnen, plädiert für soziale Anerkennung und persönliche Wertschätzung (Bauer, 2008, S. 21 f.) und zeigt, dass die biologischen Motivationssysteme über die seelischen Eindrücke aktiviert, aber auch deaktiviert werden.

arbeiten. Zusätzlich zeigte sich dies als wichtiger Motor für den Einsatz ganzheitlicher Assessment-Modelle (Winter, 2015).

Der Umgang mit der adoleszenten Stimme war mir schon immer ein wichtiges Anliegen. Durch meine Fähigkeit, jedes Lied sofort transponieren zu können, suchte ich immer die individuell guten Tonarten, nicht nur für mutierende Jungenstimmen, sondern auch bei der Adaption von Popsongs, die selten eine geeignete Tonhöhe aufweisen. In der Feldnotiz beschreibe ich das Spezielle im Umgang mit der mutierenden Jungenstimme. Obwohl ich in dieser Zeit die Stadien und Stimmumfänge der adoleszenten Jungenstimme noch nicht kannte, habe ich intuitiv darauf bestanden, in einer passenden Stimmlage zu singen. Notwendig war lediglich die Verwendung eines kleinen Ambitus in der gerade aktuellen Stimmlage. Wenn nicht andere Komplikationen auftraten, war es so möglich, mit allen Jungen zu singen. Es wird spannend sein, diesen Umstand noch weiter auszuführen (vgl. 4.4). Als notwendig erwies sich auch die zeitlich limitierte Trennung der Geschlechter. Sobald die Jungen unter sich waren, haben sie sich kaum geweigert, ihre Stimmen zu zeigen.

Im Rückblick ist aus meiner Sicht die Kombination von Empathie und einer individuellen Betreuung in einer monoedukativen Phase einer der Erfolgsfaktoren, dass meine Jungen abgesehen von wenigen Ausnahmen⁴⁴ immer gesungen haben.

Fazit: Die adoleszente Stimme und Geschlechtertrennung werden zum Thema

Die Perspektive des Musiklehrers und die des Fachdidaktikers rücken nun aus biografischen Gründen näher zusammen. Die persönliche Überzeugung, in monoedukativen Settings mit einer empathischen Betreuung die adoleszente Stimme der Jungen adäquat und insbesondere individuell betreuen zu können, wird nicht nur in der Feldnotiz, sondern auch in der begleitenden Analyse sichtbar. Die Beschreibung der evokativen Momente

⁴⁴ Ein Fall ist mir in besonderer Erinnerung. Der Junge hat sich im Unterricht selten beteiligt, sich sogar oft verweigert – aus meiner Perspektive –, und es gelang mir nicht, ihn in die Gruppe einzubinden. Sein Widerstand zeigte sich bei der Entlassungsfeier, wo er im Chor stand, aber demonstrativ das Mitsingen verweigert hat. Das schmerzte.

schließt an die von Ellis und Bochner präferierte Autoethnografie an. Es scheint bereits zu gelingen, die Fragestellung der vorliegenden Arbeit mit Beispielen zu fundieren und die Haltung und Überzeugung des Forschenden darzulegen. Die in dieser Phase gepflegte reflexive Haltung führt zu weiteren Aufgabenformaten, ein Beispiel zeugt davon als vierte und letzte Feldnotiz. Nach Anderson (2006) fehlt aber noch ein Schritt hin zur Abstraktion und Objektivierung. Es muss z. B. gelingen, die beiden Perspektiven insgesamt an didaktische Konzepte anzuschließen, oder sogar in ihrer Gültigkeit zu hinterfragen. Dies gelingt hoffentlich im letzten Punkt.

3.4 Punkt 4: Fazit aus dem Exkurs

Postludium: Was passieren kann, wenn schöne Stimmen entdeckt werden⁴⁵

Eine letzte Episode. Der Lehrer spricht: „Heute singen wir mit dem Mikrofon. Wer will es ausprobieren?“ Ein Geraune geht durch die Klasse. So richtig begeistert will sich niemand zeigen. Doch da gibt es ein Mädchen, das in die Bresche springen will. Der Lehrer ist erstaunt, weil er nicht mit dieser Jugendlichen gerechnet hat, sondern mit der allseits bekannten und extrovertierten Casting-Aspirantin. Nach einem kurzen Einstieg mit einem bereits gut einstudierten Song gibt es nun ein erstes Solo. Das Mädchen erhält eine kurze Instruktion zur Benutzung des Mikrofons und startet dann mit ihrem Vortrag. Sie singt ordentlich und erhält einen tosenden Applaus. Sofort steht die Frage im Raum, wer als nächstes drankommt. Die einen wollen nicht, weil sie ihre Mikrofon-Stimme noch nie gehört haben, die nächsten, weil sie aus Prinzip nicht vor anderen singen wollen. Andere sind aber kurz davor, sich auf das Wagnis einzulassen. Nach einigen Kandidatinnen schließt der Lehrer die Lektion ab und vertröstet die anderen auf das nächste Mal. Eine Jugendliche sagt ihrer Nachbarin, dass sie es sicher das nächste Mal schaffen werde. Tatsächlich fragen die Schüle-

⁴⁵ Der Titel nimmt noch einmal auf Mahler (2017) Bezug (vgl. 1.2), der dem Klang einer schönen Stimme positive Wirkeigenschaften beimisst.

rinnen in der nächsten Musikstunde, ob wieder mit Mikrofon gesungen werde. Der Lehrer nutzt die Motivation und stößt das Mikrofon in der Musikanlage ein. Nach der dritten Schülerin gibt es ein Geschnatter: „Ich habe gar nicht gewusst, dass du so gut singen kannst.“ Während weitere Mädchen singen, breitet sich eine positive Stimmung aus, die auch in der nächsten Singstunde anhält.

Analyse: Individuelle Förderung

Statt wie zuvor nach der Darstellung einer Episode mit einem theoretischen Block weiterzufahren, folgt an dieser Stelle bereits die Analyse der Feldnotiz, um später zusammenfassend einen Gesamtblick auf die Autoethnografie zu werfen und zu den Kapiteln mit dem Forschungsstand überzuleiten.

Die Feldnotiz schildert eine Begebenheit, wie sie während der zweiten Hälfte meiner Tätigkeit als Musiklehrer immer wieder stattgefunden hat. Die positive Dynamik, die durch den Einsatz des Mikrofons jeweils ausgelöst wurde, wurde von mir zuerst nur beobachtet und dankbar aufgenommen und später als Methode systematisch eingesetzt. Die aus der Intuition entstandene Aufgabe bestätigt die Beobachtung von Brunner (2014) und Lehmann-Wermser (2008), die beschreiben, dass Singunterricht oft intuitiv richtig, aber unreflektiert erfolgt. Als äußerliche Auslöser waren sicher die aufkommenden Castingshows hilfreich, wie sie auch die Analysen von Hoos de Jokisch (2014) und Oberschmidt (2016) ins Feld führen. Mittlerweile gibt es in Musiklehrmitteln (z. B. Detterbeck, 2011; Rohrbach, 2019) methodisch aufgearbeitete Anleitungen für die Durchführung von Castingshows, die gerade für die Entwicklung von Peer-Assessment als spannende Aufgabe angesehen werden können. Allerdings müsste sich die Art des Feedbacks von den oft gezeigten „Leider-Nein“-Sequenzen unterscheiden, um die negativen Selbstkonzepte nicht zu zementieren.

In meiner Erinnerung ist der mediale Hype um Castingshows aber nur für sehr wenige Schülerinnen und Schüler relevant gewesen, vielmehr war es die erste Überwindung, sich selbst mit der eigenen Stimme im Lautsprecher zu konfrontieren, die für die Hemmungen verantwortlich war, denn für viele war dieser erste Kontakt mit dem Mikrofon mit einer gewissen Scham verbunden. In der Feldnotiz zeigt sich der behutsame Umgang mit der anspruchsvollen Situation in der Freiwilligkeit und der zeitlichen

Staffelung der Aufgabe. Niemand musste sich einbringen und es gab genügend Zeit, um sich innerlich darauf vorzubereiten. Ich habe zunehmend darauf vertraut, dass sich die Peers gegenseitig motivieren, bis alle ihren ersten Auftritt erfolgreich hinter sich gebracht haben. Grundlage dazu war zunehmend eine von mir gepflegte Unterrichtskultur, wo nicht Fehler, sondern die individuelle Entwicklung im Zentrum stand. Meistens wurde in diesem stark gruppendynamisch geprägten Prozess eine positive Entwicklung in Gang gesetzt. Mit dem Staunen über die bisher unentdeckten Talente fing es meistens an. „Du kannst ja sehr schön singen, das habe ich nicht gewusst.“ Dieser Satz war oft ein Zaubermittel, um das Singen in der Klasse voranzutreiben, gleichzeitig an der Auftrittskompetenz zu arbeiten und vor allem die Eigeninitiative zu fördern. In dieser positiven Atmosphäre konnte sich Peerfeedback und damit verbunden eine Reflexionsebene in den Klassen entwickeln, die im später vorgestellten Konzept von Earl (2013) als *Assessment as learning* beschrieben wird. Die Mikrofon-Methode war bis zum Ende meiner Lehrtätigkeit ein Baustein, um Schülerinnen und Schüler gegen Ende ihres Musikunterrichts auf den solistischen Auftritt mit Karaoke-Playback vorzubereiten (vgl. 3.2). In der Feldnotiz wird zudem von einem monoedukativen Setting gesprochen, was sich bereits in der vorigen Feldnotiz als wichtiger Bestandteil meines Unterrichts herausgestellt hat und durchaus auch aus der Literatur als Erweiterung der Handlungsoptionen für Mädchen und Jungen (Hess, 2018) argumentieren lässt.

Viele meiner Schüler und Schülerinnen haben zu Beginn meiner Lehrtätigkeit nach anfänglicher Skepsis ohne Widerstand mit dem Mikrofon vor der Klasse gesungen. Dabei ging es mir in dieser Zeit nicht in erster Linie um die Bildung von Auftrittskompetenz, sondern darum, auch von den leisen Stimmen eine gut hörbare Probe zu erhalten. Das Argument, die Stimme sei so gut hörbar, hat denn auch immer ziemlich gut funktioniert. In Gesprächen mit befreundeten Musiklehrkräften wurde oft mit Unverständnis reagiert, wenn ich in den letzten Jahren die Mikrofon-Methode präsentierte. Für meine Kollegen aus der Musiklehrergilde stand oft die Angst im Vordergrund, die Jugendlichen würden damit bloßgestellt. Vielfach wurden auch weitere Formen ins Spiel gebracht, so z. B. das Weiterreichen des Mikros, wenn die Klasse gemeinsam singend um den Flügel steht. Es wird spannend sein, auch in den Interviews nach solch intuitiv

entwickelten Aufgabenstellungen zu suchen. Entsprechend wird jeweils am Ende der Interviews nach Highlights gefragt (vgl. 9.2).

Fazit: Vom autoethnografischen Exkurs zur Interviewstudie

Innerhalb des Diskurses zu Autoethnografie zeigt sich eine starke Präferenz für autoethnografische Texte, die ihre Wirkung durch Evokation erreichen und sich sowohl formal als auch inhaltlich explizit als Gegengewicht zu den traditionellen wissenschaftlichen Essays positionieren (Denzin, 2006; Ellis et al., 2010; Holman Jones et al., 2016b). Auf der Gegenseite besteht das Bestreben, die Anschlussfähigkeit der Autoethnografie an die traditionellen Sozialwissenschaften zu entwickeln (Anderson, 2006). Stahlke Wall (2016) plädiert schließlich für eine moderate Autoethnografie, in der sich sowohl das Evokative als auch das Analytische zeigt. Snyder (2015) fasst zusammen: „autoethnography is distinct in that it moves beyond the personal story into an analysis and interpretation of a culture.“ (S. 93)

Das vorliegende Kapitel nimmt den Diskurs trotz prominenter Ablehnung auf und exploriert die Methode mit dem hauptsächlichen Ziel, prägende Momente der Biografie des Forschenden darzustellen, in einen Entwicklungskontext zu stellen und für die Fragestellung der vorliegenden Arbeit nutzbar zu machen. Der autoethnografische Exkurs kann im Sinne von Stahlke Wall als moderate Autoethnografie bezeichnet werden. Die Affinität für die evokative Seite zeigt sich im Bedürfnis, Erlebnisse des Forschenden aus dem Singunterricht nachzuerzählen. Sie wurden in Feldnotizen komprimiert, z. B. das solistische Singen im monoedukativen Setting oder die Mikrofon-Methode. Die anschließenden Analysen zeigen auf der einen Seite die Skepsis gegenüber dem alleinigen Erzählen von Geschichten und auf der anderen Seite die Notwendigkeit, die Episoden zu diskutieren und in die Entwicklung einer Berufsbiografie einzuordnen. Feldnotizen und Analyse können in erster Linie als Daten erster Ordnung (Orgass, 2017) betrachtet werden. Zum Teil wird mit dem jeweiligen Fazit, aber sicher mit der nachfolgenden Fokussierung auf die Themen der Forschungsarbeit ein Abstraktionsgrad gesucht, wie ihn Anderson für die analytische Autoethnografie fordert.

So gilt es abschließend in diesem Kapitel zu überlegen, welche Implikationen dieser Exkurs für die Gesamtanlage dieser Forschung hat. Um das evokative Moment dieses Exkurses und die persönliche Bedeutung

der Themenwahl zu unterstreichen, werden die nächsten Abschnitte in der ersten Person beschrieben. Für mich zeigen sich drei Bereiche, die ich als Fazit präsentieren will:

- Die Darstellung von vier Episoden aus meinem Singunterricht und deren Analyse sollen mithelfen, die Narrative der Interviewpartnerinnen und -partner sorgfältig weiterzuverarbeiten. Die interviewten Lehrkräfte werden von evokativen Momenten berichten und insgesamt ihre Sing- und Schulkultur darstellen. In den Porträts können diese erleuchtenden Momente durchscheinen und in autoethnografischer Manier bereits zu *Erkenntnissen* führen (Ellis & Bochner, 2006). Zudem sollen die Porträts so gestaltet werden, dass die Deutungshoheit (Zinnecker, 2000) möglichst bei den Interviewpartnerinnen und -partnern verbleibt, immer im Bewusstsein der eigenen Interpretationsleistung, aber mit dem Willen, die Vielfalt der Singkulturen lebendig werden zu lassen.
- Ein erster inhaltlicher Punkt ist die Individualisierung im Singunterricht, die sich vor allem in Hinblick auf die stimmliche Entwicklung und die Auftrittskompetenz ergeben hat. In drei Feldnotizen wird sichtbar, dass ich einen starken Fokus zugunsten individualisierender Formen entwickelt habe: in der Jungengruppe wird persönliches Feedback erteilt, bei der Bewertung von Singleleistungen in Form eines Karaoke-Vorsingens wird individuell gecoacht und bei der Episode rund um das Mikrofonstehen stehen wiederum die persönlichen Leistungen einzelner im Zentrum. Obwohl Singen in der Klasse eine große Wichtigkeit hatte – fast jede Klasse hat zum Abschluss der 9. Klasse⁴⁶ als Klasse auf der Bühne erfolgreich und mit Überzeugung einen Song gesungen – habe ich die persönliche Entwicklung immer als wichtigen Punkt erachtet. Der Gedanke eines holistischen Assessments schimmert bereits in verschiedenen Facetten hervor. Zu Beginn war die individuelle Betreuung eher intuitiv gesteuert, gegen Ende wurde aber bereits das Modell von Earl (2013) mit *Assessment for, as und of learning* in seiner Essenz angewandt. Die Idee der Individualisierung hinterlässt allerdings auch ambivalente Gefühle, da sich mit Blick auf mein Umfeld eine Diskrepanz zu zeigen scheint, denn ich sehe im Musikunterricht noch nicht so viel davon bei meinen Kolleginnen und Kollegen.

⁴⁶ Ende der obligatorischen Schulzeit.

Bis jetzt bestätigen sich meine langjährigen positiven Erfahrungen erst in wenigen musikpädagogischen Texten (z. B. Pezenburg & Dyllick, 2018; Dyllick, 2019; Schaumberger, 2020a).

- Ein zweiter inhaltlicher Punkt zeigt sich in der zunehmenden Beachtung geschlechtsspezifischer Unterschiede im Singunterricht. Insbesondere die Mutation der Jungen hat mich immer beschäftigt. Es war immer mein Ziel, die Jungen beim Singen zu halten. Zudem werden nicht von ungefähr in drei der Episoden Mädchen und Jungen getrennt. Zwar wird in der Literatur die Annahme weiterverbreitet, dass Mädchen einen besseren Zugang zum Singen haben, als Jungen (Lehmann-Wermser, 2002; Oebelsberger, 2009; Hess, 2015; Blanchard, 2019), meine gegenteilige Erfahrung kann aber vielleicht damit erklärt werden, dass ich beiden Geschlechtern mit der oftmaligen Trennung eine unabhängige Plattform gegeben habe. Die Kombination aus Beachtung der adoleszenten Stimmentwicklung und der monoedukativen Settings zeigt sich als zweiter Aspekt in der vorliegenden Autoethnografie.

Beide inhaltlichen Schwerpunkte führen zu je einem anschließenden Kapitel. Für mich wird es interessant sein, den jeweiligen Forschungsstand literaturbasiert zu ergründen, bevor ich das Thema Assessment mit der qualitativen Inhaltsanalyse an den Datenkorpus herantrage. Auf jeden Fall werde ich ein differenzierteres Bild entwickeln können.

Postskriptum: Wie steht es mit dem Zeitgeist?

In dieser autoethnografischen Rückblende führt die Auswahl der persönlichen Erlebnisse zur Beschreibung und Legitimierung eines sich zunehmend individualisierenden Unterrichts. Dabei rücken sowohl die Vorteile gemeinsamen Singens (Gembris, 2015) als auch die zahlreichen Initiativen, die das Singen in der Schule per se stärken, in den Hintergrund. Auch wenn argumentiert werden kann, dass individuelle Betreuung beim Singen als Leerstelle in der Musikpädagogik zu betrachten ist (Pezenburg & Dyllick, 2018), bleibt die Frage offen, ob der Nachvollzug des gesellschaftlichen *Megatrends Individualisierung*⁴⁷ nicht hinterfragt werden

⁴⁷ Das Zukunftsinstitut (<https://www.zukunftsinstitut.de> – abgerufen am 21.12.2021) beschreibt auf ihrer Webseite 12 Megatrends. Einer davon ist Individualisierung.

müsste. Der Blick auf den Bereich Beurteilung zeigt, dass spätestens bei der summativen Beurteilung das Individuum im Zentrum steht und dies nicht erst seit Beginn des genannten Megatrends. So führt der autoethnografische Exkurs zwar zu einer Transparenz bezüglich Haltung und Überzeugung des Forschenden, steht aber durch sein evokatives Momentum auch unter Verdacht, einen normativen Drall zu entwickeln.

Aus diesem Grund werden in den nun folgenden Kapiteln der Lehrer sowie der Fachdidaktiker gebeten, einen Schritt zurückzutreten und sich nach bestem Wissen und Gewissen im Hintergrund zu halten. Die autoethnografische Rückblende ist nun abgeschlossen.

4 Forschungsstand I – Singen in der Adoleszenz

Bereits im ersten Kapitel hat sich gezeigt, dass Assessment im Singunterricht während der Jugendzeit in einem hohen Maß abhängig ist von den Phänomenen der adoleszenten Stimmentwicklung (vgl. 1.2). Die beschriebenen Erfahrungen im autoethnografischen Exkurs (vgl. 3.3) dienen als weiterer Beleg, dem Umgang mit der adoleszenten⁴⁸ Stimme in der Sekundarstufe 1⁴⁹ vermehrt Aufmerksamkeit zu schenken. Zahlreiche Texte (z. B. Brunner, 2014; Tellisch, 2015; Schaumberger, 2020a) beziehen sich zudem auf emotionale Schäden aus der Jugendzeit, die durch unqualifizierte Kommentare zur Stimme entstanden sind. Blanchard (2019) beschreibt sogar die feste Vorstellung von Lehrkräften, dass Jungen schlechter singen können (S. 268 ff.). Dieses Kapitel setzt sich deshalb vertieft mit der adoleszenten Stimmentwicklung von Mädchen und Jungen auseinander und rezipiert insbesondere auch internationale Studien, die bis dato wenig Berücksichtigung im deutschsprachigen Raum gefunden haben. Weiter werden in diesem Kapitel auch die Einflüsse der Stimmentwicklung auf das Gruppengefüge im Singunterricht untersucht, insbesondere das Aufeinandertreffen der Geschlechter und mögliche Gendereffekte. Mit der Aufarbeitung dieser Themen wird einerseits das nächste Kapitel zum Thema Assessment vorbereitet und andererseits eine Grundlage geschaffen, die beschriebenen Effekte später in den Interviews zu adressieren.

Das Kapitel startet mit der Frage, was Singen ist und wie es in den Bildungskontext eingeordnet werden kann (vgl. 4.1). Im zweiten Abschnitt werden die Begriffe Pubertät, Adoleszenz und Jugend in ein Verhältnis gebracht, um die verschiedenen Aspekte des Singens spezifisch auf das Alter der Lernenden der Sekundarstufe 1 beziehen zu können (vgl. 4.2). Nachfolgend fokussiert ein erster Abschnitt auf die Mädchenstimme (vgl. 4.3) und ein zweiter auf die Jungenstimme (vgl. 4.4), um abschließend der Frage nachzugehen, welche geschlechtsspezifischen Einflüsse sich beim Singen im Zusammenspiel von Lehrkräften und Jugendlichen zeigen (vgl. 4.5).

⁴⁸ In Kapitel 2.1 wurde der Begriff *adoleszente Stimme* bereits hergeleitet.

⁴⁹ Die Besonderheiten des schweizerischen Bildungssystems werden in Kapitel 6 dargelegt.

4.1 Singen als urmenschliche Tätigkeit

In der vorliegenden Studie steht das Singen im Zentrum. Bis jetzt wurde Singen lediglich als jegliche Form von vokalen Aktivitäten im Musikunterricht definiert und von Gesangsunterricht als Einzelunterricht abgegrenzt (vgl. 1.1). In diesem Abschnitt wird der Blick noch einmal geöffnet, um Singen als menschliche Ausdrucksweise in einem weiteren Rahmen zu definieren und anschließend als Lerngegenstand auf den Lehrplan zu beziehen.

Singen als anthropologische Konstante

In vielen Texten zum Thema Singen wird betont, wie zentral der Gebrauch der Stimme in allen Kulturen ist. So startet Sundberg (2019) seinen Artikel zur Singstimme in *THE OXFORD HANDBOOK OF VOICE PERCEPTION* (Frühholz & Belin, 2019) mit der Aussage: „*The voice is one of the most common, if not the commonest, and probably also the oldest, of all musical instruments. In all cultures in the world, singing is a regular part of human expression.*“ (S. 117, Hervorhebung im Original) Die Gültigkeit dieser Aussage kann mit verschiedenen Studien belegt werden (z. B. Trehub et al., 2015; Mehr et al., 2019) werden. Die Stimme wird somit in allen Kulturen als Musikinstrument wahrgenommen und Singen als natürliche Tätigkeit erlebt. Bereits an früherer Stelle wurde belegt, dass die Fähigkeit zu singen, fast durchwegs gegeben ist (Dalla Bella et al., 2007)⁵⁰. Singen zeigt sich somit als Modus von Musik und als anthropologische Konstante (Neus, 2017, S. 76). Schülerinnen und Schüler tragen so ihr ureigenes Musikinstrument täglich in die Schule. Allerdings zeigen sich auch Komplikationen, denn viele Menschen und insbesondere auch Schülerinnen und Schüler sind wenig überzeugt von ihrer eigenen Singfähigkeit (Ashley, 2015; Sammler, 2017). Dies wird verständlicher, wenn die starke Verbindung von Person und Singen in Betracht gezogen wird.

So beschreibt Hammar (2016) die Benutzung der eigenen Stimme auch als „[d]ie natürlichste und selbstverständlichste Art, Musik gleichzeitig zu

⁵⁰ Die Studie bezieht sich auf die Qualität des Singens im abendländischen Kulturkontext.

erleben [und] auszuführen“ (S. 14), betont aber das emotionale Erleben, weil „dabei die innigste Beteiligung“ (S. 14)⁵¹ entsteht. In der Aussage von Hammar bleibt Singen nicht nur ein Modus von Musik, sondern wird zu etwas Persönlichem. Fuchs (2010) spitzt dies zu, indem er der menschlichen Stimme beim Singen über das Klingen hinaus weitere Funktionen zuschreibt:

„Sie fungiert als eigenständiger Teil der Kommunikation auch über die rein verbale Äußerung hinaus, indem sie über melodische, dynamische und [sic] Klangparameter zum Beispiel Informationen über die Persönlichkeit des Menschen, seinen Bezug zum Inhalt des gesprochenen Wortes und über seine derzeitige emotionale und psychische Befindlichkeit übermittelt.“ (S. 9)

Singen übermittelt somit Persönlichkeitsmerkmale ebenso wie augenblickliche Befindlichkeiten. Der singende Mensch ist somit nicht nur musikalisch tätig, sondern kommuniziert mit dem Singen seine Emotionen, bewusst oder unbewusst. So wird die menschliche Stimme gerne als *Spiegel der Seele* (Ruzicka, 2019) bezeichnet, wobei die Offenlegung emotionaler und psychischer Befindlichkeiten nicht vollständig kontrolliert werden kann, wie aus dem Zitat von Fuchs oben abzuleiten ist. Auch Neus (2017) betont die Verbindung des Singens mit der Person, wo physische, psychophysische, individuelle und soziokulturelle Parameter sich gegenseitig beeinflussen (S. 76 f.). Im Gegensatz zu einer abstrakten Rechnungsaufgabe wird das Singen in der Schule so immer zu etwas Persönlichem und kann z. B. nicht auf musikpraktische Fähigkeiten (vgl. J. Hasselhorn, 2015) reduziert werden, sondern bedarf einer Herangehensweise mit einem holistischen Ansatz. Bevor der Zugang zum Singen im Lehrplan untersucht wird, werden aber die Überlegungen von Neus zur konzeptuellen Vielfältigkeit des Singens auf die Schule bezogen.

⁵¹ Interessant ist zu sehen, dass beide Autoren, weiter oben Sundberg (2019) und hier Hammar (2016), wie viele andere ihre Aussagen zum Singen nicht belegen. Dies bringt zum Ausdruck, dass Singen etwas Allgemeingültiges hat und sich der wissenschaftlichen Diskussion gewissermaßen entzieht.

Konzepte des Singens

Neus (2017) unterscheidet vier Konzepte, die zuerst aufgrund der Perspektive der vorliegenden Studie auf den Kontext der Schule bezogen werden.

Erstens unterscheidet Neus Singen als „grundsätzliches, menschliches Ausdrucks- und Kommunikationsmittel“ (S. 48) von Gesang als künstlerischem Produkt mit einem Bildungsanspruch und übernimmt dabei die von B. Jank (2013, 2017) gepflegte Differenzierung. In ihrer Darstellung ist Singen der Prozess und Gesang das flüchtige Produkt von Singen. Im Kontext der Schule wird diese Unterscheidung im Singunterricht wirksam, weil auf der einen Seite oft gemeinsam gesungen wird und auf der anderen Seite bei Auftritten das flüchtige Produkt Gesang entsteht. Beim *Vorsingen* werden speziell an das Produkt Gesang musikalische und ästhetische Ansprüche erhoben, wobei die Gefahr besteht, die Stimme lediglich als Musikinstrument zu betrachten und die starke Verknüpfung mit der Person im Prozess Singen auszublenden. Während der adoleszenten Stimmentwicklung ist Singen als Prozess zusätzlich erschwert, was wiederum den Gesang als Produkt beeinflusst. Die von B. Jank geprägte und von Neus wiederaufgenommene Unterscheidung begründet so auch das Dilemma in der Beurteilung von Gesang, das bereits als relevanter Kontext dieser Arbeit deklariert wurde (vgl. 1.2). So stellt sich z. B. die Frage, wie Gesang bewertet wird, wenn die Person *niedergeschlagen* ist oder gerade mitten in der Mutation steckt. Für das Thema Assessment im nächsten Kapitel ist diese Unterscheidung hilfreich, um die Förderung im Prozess und die Beurteilung des Produkts zu unterscheiden. Im Hinblick auf die Interviews kann untersucht werden, ob die Lehrkräfte eher am Prozess Singen oder am Produkt Gesang interessiert sind und wie sich dies in ihrem Unterricht zeigt. In der bereits mehrfach erwähnten Arbeit von Blanchard (2019) zeigt sich klar, dass Singen in seinem Sample nicht als Entwicklungsprozess betrachtet wird und keine gesangstechnischen oder ästhetischen Ansprüche an das Produkt Gesang erhoben werden. Im weiteren Verlauf werden die Begriffe Singen und Singunterricht weiterhin für alle Aktivitäten verwendet und dort, wo es um die Beurteilung geht, jeweils mit dem Terminus Gesang ergänzt.

Neus ordnet als zweites Konzept dem Singen vielfältige Erwartungen an die soziokulturellen Wirkungen, wie z. B. eine positive Wirkung auf Gesundheit, verbesserte Lebensqualität oder eine Reduktion der Gewalt-

tätigkeit in der Gesellschaft zu, ist aber gleichzeitig skeptisch, „ob das Singen in der Lage sein kann, all diesen Implikationen standzuhalten“ (S. 52). In diesen Bereich gehört der lange Diskurs zu den Transfereffekten von Musik bzw. von Singen, die hinlänglich untersucht wurden (Gembris et al., 2001; Schumacher, 2009; Jaschke et al., 2013). Für den Singunterricht sind an dieser Stelle die bereits erwähnten positiven Effekte auf die Emotionsregulierung und die Gruppenbildung (Gembris, 2015) zu erwähnen. Dieser Bereich wird im Singunterricht dann wirksam, wenn Lehrkräfte die positiven Effekte des Prozesses Singen betonen und die Ansprüche an das Produkt weniger gewichten.

Eine dritte Funktion zeigt sich in der Frage, ob das Singen der Erziehung dient, wie dies z. B. im Nationalsozialismus (Niessen, 2008) der Fall war, oder ob Singen der musikalischen Förderung dient, um theoretische Inhalte und musikalischen Geschmack zu etablieren (Neus, 2017, S. 53). Der zweite Anspruch zeigt sich aktuell in der zunehmenden Arbeit mit Chorklassen und z. B. auch in den theoretischen Überlegungen von Dyllick (2019) zur Kontextualisierung des Singens. Für die vorliegende Studie kann möglicherweise eine Verknüpfung mit musikalischen Lernzielen im Lehrplan gemacht werden, wie auch im vierten Bereich, für den Neus die Bildung von kultureller Identität durch das Singen von Liedern nennt.

Die vier Bereiche haben mehr oder weniger Gewicht in der vorliegenden Arbeit. Im letzten Abschnitt wird ein Blick in den Lehrplan geworfen, um die Ansprüche an das Singen aus dem Bildungsraum der Studienteilnehmenden zu deklarieren.

Singen im Lehrplan

Im LP21 (D-EDK, 2015) werden für den Kompetenzbereich Singen folgende Tätigkeitsbereiche allgemein formuliert:

„Die Schülerinnen und Schüler setzen ihre Stimme im musikalischen Lernen vielfältig ein, indem sie eine Beziehung zu ihrer persönlichen Stimme aufbauen und sich im stimmlichen Ausdruck als Individuum erleben. Sie singen sowohl chorisch als auch solistisch im Ensemble und erarbeiten ein vielseitiges Repertoire.“ (S. 5)

Im Zentrum steht der Aufbau einer Beziehung zum persönlichen Instrument Stimme, was durch vielfältige musikalische Aktivität geschehen

soll, insbesondere durch das Singen in verschiedenen Konstellationen und durch den Aufbau eines Repertoires. In den folgenden Kompetenzbeschreibungen (S. 7 ff.) wird in drei Bereichen gearbeitet:

- Individuelle Stimmen sollen sich im ersten Kompetenzbereich in einen Chorklang einfügen und Engagement entwickeln. Ziel für alle ist das mehrstimmige Singen, darüber hinaus wird als weiteres Ziel das solistische Auftreten von Schülerinnen und Schüler genannt.
- Im zweiten Bereich werden klassische Stimmbildungselemente aufgezählt, so die Erweiterung des Ambitus und der Gebrauch verschiedener Register. Ebenfalls wird das sichere Treffen von Tönen und eine Variabilität im Stimmklang angegeben.
- Im dritten Bereich wird die Repertoirepflege genannt. Ziel ist das Singen verschiedener Stile aus verschiedenen Kulturen und Epochen.
- Zusätzlich wird im Kompetenzbereich Gestaltungsprozesse die Entwicklung von Auftrittskompetenz unter anderem für das Singen erwähnt.

Insgesamt zeigen sich in den Kompetenzbeschreibungen kaum Hinweise auf die starke Verknüpfung von der Person mit dem Singen. Zwar soll laut Kompetenzbeschreibung eine persönliche Beziehung zur Stimme aufgebaut werden, genannt werden aber hauptsächlich musikalische Kompetenzen. Ein Vergleich mit den von Neus beschriebenen vier Funktionen des Singens ergibt ein differenzierteres Bild.

In der ersten Funktion unterscheidet Neus zwischen Prozess und Produkt, Singen und Gesang. Für beide Elemente können im LP21 Kompetenzen abgelesen werden. So kann z. B. der Prozess des Übens in der Gruppe oder das Vergleichen von Interpretationsmöglichkeiten gut dem Prozess zugeordnet werden. Andere Kompetenzen zielen dagegen klar auf das Produkt ab, so z. B. die Nennung des Gesangsvortrags. Für die Lehrkräfte ergibt sich damit ein großer Spielraum in der Gestaltung des Singunterrichts. Die in der zweiten Funktion genannten soziokulturellen Wirkungen können im LP21 nicht gefunden werden, die Erziehung zur Musik als dritte Funktion zeigt sich dagegen klar, allerdings bei den musiktheoretischen Kompetenzen, wo das Singen explizit als Zugang erwähnt wird. In der vierten Funktion dient das Singen einer kulturellen Identitätsbildung. Diese wird

im LP21 nicht auf einer lokalen oder nationalen Ebene gefordert, vielmehr ist es das Ziel möglichst breite kulturelle Erfahrungen zu ermöglichen⁵².

Zu Beginn des Abschnitts wurde das Singen als urmenschliche Tätigkeit beschrieben, die gleichzeitig musikalischer Ausdruck und Kommunikation ist. Singen wurde als anthropologische Konstante definiert. Wie Neus beschrieben hat, kann das Singen verschiedene Funktionen einnehmen, ist aber immer als Verschränkung der Person mit ihrem Singen zu sehen. Die Analyse des LP21 zeigt zudem eine Betonung musikalischer Kompetenzen, die sich im Prozess (Singen) und im Produkt (Gesang) manifestieren können, mögliche Komplikationen aufgrund der adoleszenten Stimmentwicklung werden aber ausgeblendet. Im nächsten Abschnitt folgt nun ein spezifischer Blick auf die Jugendlichen in der Pubertät, wo die eigene Wahrnehmung und die Stellung in der Gruppe das Singen stark herausfordern.

4.2 Pubertät – Adoleszenz – Jugend

Die vorliegende Studie bezieht sich auf Jugendliche im Alter von 13 bis 16⁵³ Jahren, die bereits in die Phase der Pubertät eingetreten sind oder gerade in diese eintreten. Wie bereits mehrfach ausgeführt, sind nicht nur die Auswirkungen der hör- und sichtbaren biologischen Prozesse der Pubertät, sondern auch Aspekte der Persönlichkeitsentwicklung und des sozialen Miteinanders im Musikunterricht relevant. Dafür werden in der Literatur bezüglich der Identitätsbildung die Begriffe Adoleszenz und in soziologischen Studien der Begriff Jugend verwendet. In Abgrenzung zum Begriff Pubertät, der in der Wissenschaft einheitlich verwendet wird, werden die Termini Adoleszenz und Jugend laut King (2013) in den verschiedenen

⁵² Als Ausnahme sei der vom Kanton Aargau adaptierte LP21 genannt, wo explizit gefordert wird, das Aargauer Liedgut in der Unterstufe zu pflegen. (<https://ag.lehrplan.ch/102KyNdXe5cFZmfA4rXpFuvRrvGFfm5TB> – abgerufen am 19.08.2021).

⁵³ Im Gegensatz zu manchen deutschen Bundesländern treten die Schülerinnen und Schüler in der Schweiz erst in der 7. Klasse in die Sekundarstufe über. Ausführliche Informationen zum schweizerischen Schulsystem werden im Kapitel 6.1 dargelegt.

Forschungsrichtungen „sehr divergierend und uneinheitlich gebraucht“ (S. 29). Für die vorliegende Studie wird die Setzung von Jung übernommen. Adoleszenz wird mit der individuellen Entwicklung assoziiert und der Begriff Jugend im Zusammenhang mit einer soziologischen Perspektive benutzt (S. 30).

Die drei Themenbereiche werden auf zwei Ebenen dargestellt und zueinander in Bezug gesetzt (vgl. Abbildung 1). Pubertät und Adoleszenz fokussieren hauptsächlich auf die individuelle Entwicklung und die persönliche Wahrnehmung und bilden die erste Ebene. Der Bereich Jugend bezieht sich auf die Wahrnehmung in der Gruppe und das Zusammenspiel der Geschlechter und bildet die zweite Ebene. Dieser Logik folgen die drei nächsten Abschnitte dieses Kapitels. In den beiden ersten Abschnitten wird die Entwicklung der Mädchen- und der Jungenstimme dargestellt, bezogen auf biologische und entwicklungspsychologische Merkmale, und anschließend wird das Singen im Musikunterricht aus der soziologischen Perspektive betrachtet.

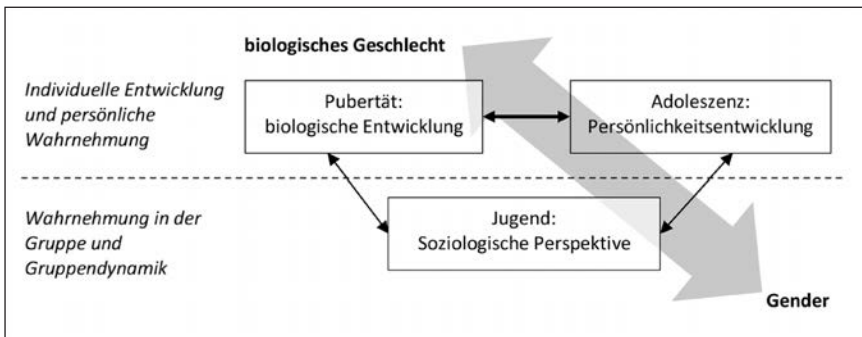


ABBILDUNG 1: Abgrenzung von Pubertät, Adoleszenz und Jugend

Auf einem weiteren untergeordneten Layer wird in der Abbildung 1 das biologische Geschlecht dem Begriff Gender gegenübergestellt. Gender befindet sich rechts unten in der Grafik und weist eine gewisse Nähe zum Begriff Jugend und zur Adoleszenz auf, dies im Gegensatz zum biologischen Geschlecht, das links oben mit dem Begriff Pubertät verknüpft ist. In Abschnitt 4.5 wird das Begriffspaar argumentiert und in den Kontext der Arbeit eingeordnet.

4.3 Die adoleszente Mädchenstimme

Die Literatur zur adoleszenten Mädchenstimme ist überschaubar. Lynne Gackle (Huff-Gackle, 1985) hat sich seit 1985 als Hauptautorin zum Thema im amerikanischen bzw. englischsprachigen Raum etabliert. Dementsprechend schreibt sie mehr als dreißig Jahre später im OXFORD HANDBOOK OF SINGING (Welch et al., 2019) das Kapitel ADOLESCENT GIRLS' SINGING DEVELOPMENT (Gackle, 2019). Der Diskurs zur Entwicklung der Mädchenstimme spannt sich so über mehr als dreißig Jahre und wird im CHORAL JOURNAL⁵⁴ ausgetragen. Spannend ist die stetige Erweiterung der Perspektive, ausgehend von messbaren Phänomenen wie Ambitus und Klangqualität, hin zu einer ganzheitlichen Betrachtung, in der auch die emotionale Entwicklung und die Kontexte des Singens berücksichtigt werden (Sweet, 2015, 2016). Die messbaren Phänomene beziehen sich dabei auf die biologischen Reifeprozesse⁵⁵ der Pubertät und die emotionale Entwicklung auf die psychische Entwicklung in der Adoleszenz, beides Bereiche mit Fokus auf die individuelle Entwicklung (vgl. Abbildung 1). Der Einbezug der Kontexte des Singens wechselt auf die soziologische Perspektive und bezieht sich damit auf den Terminus Jugend und somit auf die Wahrnehmung in der Gruppe (vgl. ebenfalls Abbildung 1). Im nächsten Abschnitt wird das Phasenmodell zur Entwicklung der Mädchenstimme von Gackle (2011, 2019) vorgestellt und mit ausgewählten Studien ergänzt, um anschließend einen Vergleich mit Aussagen der deutschsprachigen Literatur zur adoleszenten Mädchenstimme zu machen.

Amerikanische Studien

Gackle startet 1985 mit einem ersten Artikel zur adoleszenten Mädchenstimme (Huff-Gackle, 1985). Sie nimmt Bezug auf die positiven Effekte der Forschung von Cooksey (z. B. 1977a) zur adoleszenten Jungenstimme

⁵⁴ Das Choral Journal ist das offizielle Publikationsorgan der American Choral Directors Association und wird seit 1959 herausgegeben. Aktuelle Webseite: <https://acda.org/publications/choral-journal> (abgerufen am 12.09.2021).

⁵⁵ Das Referieren von Belegen zu den biologischen Reifeprozessen erscheint an dieser Stelle nicht zielführend, da sich die Arbeit später mit der pädagogischen Seite der Phänomene beschäftigt.

(vgl. 4.4) für die Arbeit von Lehrkräften und Chorleitern und fordert dasselbe für die Mädchen: „Perhaps an in-depth study of the young female voice would yield many of the same helpful answers.“ (Huff-Gackle, 1985, S. 15) Ausgehend von ihrer eigenen Erfahrung als Chorleiterin eines Mädchenchores beschreibt sie als Erste drei verschiedene Entwicklungsstufen der Mädchenstimme (S. 16). In dem im Jahr 1991 erschienenen Artikel *THE ADOLESCENT FEMALE VOICE: CHARACTERISTICS OF CHANGE AND STAGES OF DEVELOPMENT* (Gackle, 1991) führt sie ihr Modell weiter aus und beschreibt vier Stufen der Stimmentwicklung. Der Artikel wird anschließend oft zitiert⁵⁶ und lässt Gackle entsprechend zur Expertin der adoleszenten Mädchenstimme werden. Später beschreibt sie die vier Stufen der Entwicklung neu als vier sich entwickelnde Phasen, im Wissen um den graduellen und individuellen Charakter des Stimmwechsels (Gackle, 2019, S. 557). Abgesehen von kleinen Anpassungen in Sprache und bezüglich der Altersstufen bleibt das Modell von 1991 bis 2019 aber unverändert.

In der Tabelle 2 (S. 97) werden die Eckpunkte der vier Phasen übersetzt und zusammengefasst⁵⁷ wiedergegeben und ergänzt mit dem Ambitus aus einem anderen Artikel von Gackle (2000). Zur Bestimmung der Entwicklungsphasen schlägt Gackle folgende Parameter vor: Bestimmung der indifferenten Sprechstimmlage, des Ambitus (tiefster und höchster singbarer Ton) und der Tessitura⁵⁸ (komfortabel singbarer Bereich); Entwicklung der Register unter Einbezug der Registerübergänge, die allgemeine Stimmqualität (Timbre) und den Anteil der behauchten⁵⁹ Tongebung

⁵⁶ Laut Google Scholar wurde der Artikel 132-mal zitiert, Stand 12.09.2021.

⁵⁷ Übersetzung durch den Autor. Verzichtet wurde auf die Darstellung der Zugehörigkeit zu Stimmlagen, die im Text von Gackle der Verknüpfung mit Chorarbeit geschuldet ist. Zudem wurde auch die Angabe von Frequenzangaben in Hertz für die indifferente Sprechstimmlage weggelassen.

⁵⁸ Im Gegensatz zum Ambitus, der den höchsten und tiefsten singbaren Ton beschreibt, ist die Tessitura der Bereich der Stimme, der angenehm zu erreichen ist.

⁵⁹ In der Phoniatrie spricht man von *Behauchtheit* als Merkmale einer heiseren Stimme (Ptok et al., 2006). Durch den nicht vollständigen Stimmbandschluss beim Phonieren ist ein Luftanteil in der Stimme auditiv wahrnehmbar. Das Phänomen zeigt sich unter anderem während der Pubertät im Rahmen einer physiologischen Phase. Im Gegensatz dazu verwendet die Sprachwissenschaft das Wort *Behauchung*, wenn Konsonanten mit einem Hauchlaut ausgespro-

(Gackle, 2019, S. 559–560). Während Phase 1 die noch unveränderte Kinderstimme beschreibt, beziehen sich die Phasen 2A und 2B auf die einsetzende Pubertät. Unterschieden wird zwischen der Phase 2A, in der eine körperliche Entwicklung der sekundären Geschlechtsmerkmale einsetzt, und der Phase 2B mit dem Start der ersten Menarche. Gackle beschreibt für die Phase 2A eine einsetzende Behauchtheit und eine mögliche Einschränkung des Ambitus im oberen Bereich. Stärker hörbar sind in der Phase 2B Heiserkeit, Unbehagen beim Phonieren insbesondere mit der Randstimme und eine variable Tessitura. In der Phase 3 nimmt die Behauchtheit und Heiserkeit ab, die Stimme gewinnt an Resonanz und wird flexibler und agiler, soll aber noch nicht in eine Stimmgattung eingeteilt werden, sondern vielfältig in der *treble voice*⁶⁰ eingesetzt werden. Das Fehlen der Angaben zum Ambitus im aktuellen Übersichtsartikel von Gackle kann darauf zurückgeführt werden, dass dieser sich im Verlauf der adoleszenten Stimmentwicklung nur unwesentlich verändert (vgl. Abbildung 3). Gackle vergleicht aber ganz bewusst die Phänomene mit der Mutation der Jungen: „Much like their male counterparts, young female singers present characteristic symptoms of vocal change.“ (S. 551) Und ebenfalls analog zu den Jungen beobachtet sie, dass die Mutation in unterschiedlich starken Ausprägungen wahrgenommen wird und sehr individuell verläuft (S. 561).

Im weiteren Verlauf werden einige Studien genannt, welche die Arbeit von Gackle aufgenommen und weiterentwickelt haben. Siple (1994) reagiert mit einer Studie, in der er mit einem Pre-/Posttest die Auswirkungen von Stimmbildung in Kombination mit der Vermittlung von Wissen zum Stimmwechsel untersucht. Er testet zu Beginn und Ende mit einem Fragebogen das vokale Selbstkonzept und mit Stimmaufnahmen die Qualität des Gesangs. Als Ausgangspunkt bezieht er sich auf den Artikel von Gackle (1991), indem er pointiert auf ihre Argumentation verweist: „While the change in females is not as dramatic as in males, it can be traumatic.“ (Siple, 1994, S. 35) Damit verweist er bereits auf die psychologische Ebene der Mutation. In der Studie wurden vier Gruppen gebildet:

chen werden. In der vorliegenden Arbeit wird der Terminus *Behauchtheit* verwendet.

⁶⁰ Die *American Choral Directors Association* definiert *treble voice* als Sängerinnen und Sänger im Alter von acht bis sechzehn Jahren. Treble Voice wird in Abgrenzung zu einer Fixierung auf die Register Sopran und Alt verwendet.

Phasen der adoleszenten Mädchenstimme

Nr. Name der Phase	Indifferente Sprech- stimmlage <i>Tessitura</i> <i>Ambitus</i>	Alter
Charakteristiken der Stimme		
.. <i>Kommentar</i>		
1 Vorpubertär: Unveränderte Stimme	c'-dis'	8-10
– klare/flötenähnliche Qualität	erweitert: (12)	
– ähnlich wie die Jungenstimme in diesem Alter	a-f'	
– keine offensichtlichen Registerbrüche	<i>d'-d''</i>	
– flexibel/agil	b-f'' (a'')	
<i>Kommentar: Abhängig von anderen physiologischen Entwicklungen (z. B. Brustentwicklung, Menarche) kann diese Phase bis zum Alter von 12 dauern.</i>		
2A Prämenarchal: Start der Mutation	ais-cis'	11-13
– Tonumfang der Stimme ungefähr gleich wie in Phase 1	erweitert:	
– aber stärkere Behauchtheit des Tons über den gesamten Tonumfang	ais-d'	
– Registerbrüche (Registerwechsel) erscheinen manchmal rund um f'-ais'	<i>d'-d''</i>	
– Verlust des oberen Tonumfangs	a-g'' (a'')	
– einige haben Schwierigkeiten im unteren Tonbereich		
<i>Kommentar: Erste Anzeichen der körperlichen Reifung zeigen sich (Brustentwicklung, Größenzunahme, andere sekundäre Geschlechtsmerkmale).</i>		
2B Postmenarchal: Höhepunkt der Mutation	a-c'	12-15
– Heiserkeit im gesamten Ambitus	erweitert:	
– Verkleinerung des Ambitus	g-d'	
– Verlust einiger Töne im oberen Register	<i>h-c''</i>	
– Registerübergänge erscheinen bei f'-ais' und auch bei d''-fis''	a-f'' (a'')	
– tiefere Noten können einfacher produziert werden		
– ‚Illusion‘ von Alt-Qualität		
– Probleme oder Unbehagen beim Singen (oder Phonieren)		
– Tessitura ist variabel, verschiebt sich nach oben oder unten ans Ende des Registers		
– Häufiges Auftreten von Registerbrüchen und Behauchtheit		
– Wenn die Stimme diese Phase zu verlassen beginnt, kann sich Vibrato entwickeln		

3	Junge erwachsene Frauenstimme	g–h	14–17
–	Die Klangfarbe entspricht in etwa der Qualität junger Erwachsener.	erweitert: fis–c’	
–	Bei der Stimmqualität wird mehr Reichtum wahrgenommen.	a–g’’	
–	Tonumfang nimmt charakteristisch zu.	(g) a–a’’	
–	größere Konsistenz in den Registern	(h’)	
–	abnehmende Behauchtheit		
–	Registerwechsel zwischen d’–fis’ ist deutlicher		
–	Natürliches Vibrato kann in der Stimme auftreten		
–	größere Flexibilität und Agilität in der Stimme		
–	Lautstärke und Resonanzfähigkeit steigen		
<i>Kommentar: Stimmen können phasenweise fast jede „treble voice“⁶⁰ singen (solange es für die Sängerin bequem ist und auf die Gesundheit der Stimme geachtet wird). In Chorszenarien kann diese Stimme verwendet werden, um Stimmgruppen zu „färben“, abhängig von der gewünschten Klangqualität in einem Musikstück. Zusätzlich zur Förderung musikalischer Fertigkeiten genießen die meisten Schülerinnen den Wechsel zwischen den Stimmen aufgrund der Herausforderung und weil es Interesse weckt. Der Schwerpunkt sollte jedoch auf dem stimmlichen Komfort liegen, ohne dass die zugewiesene Stimmlage belastet oder angespannt wird.</i>			

TABELLE 2: Phasen der adoleszenten Mädchenstimme nach Gackle (2019, S. 558–559), ergänzt mit dem Ambitus nach Gackle (2000)

In einer Kontrollgruppe fand keine Intervention statt, je in einer Gruppe wurde Stimmbildung oder Wissen zum Stimmwechsel vermittelt und in einer vierten Gruppe erfolgten beide Interventionen. Nur in der letzten Gruppe kann er „a significant difference in the subjects’ attitudes toward their singing voices“ (S. 36) feststellen, eine vokale Verbesserung aber in keiner der Gruppen. Er kann so zeigen, dass die Beschäftigung mit dem Selbstkonzept zur eigenen Stimme, er nennt es *vocal self-image*, für den Zugang zum Singen wichtig ist. Damit wird deutlich, dass das Bewusstsein für die Phänomene des Stimmwechsels für die psychologische Entwicklung von Bedeutung ist. Die von Siple benutzten Reflexionsfragen werden in den Thesen für das Assessment des Singens (vgl. Kapitel 14) noch einmal aufgenommen. Phillips (1995) betont in einer weiteren

Anschlussstudie, dass Mädchen vor dem Abschluss ihres Stimmwechsels nicht in ein Register eingeteilt werden sollten (S. 25), und bestätigt damit die Agilität der Stimme in der Phase 3. Im Gegensatz zu Gackle ist er aber der Meinung, dass der Ambitus grundsätzlich zwei Oktaven (g-g'') betragen sollte, nämlich dann, „when they can shift easily from lower to upper vocal registers“ (S. 25). In einer australischen Studie wird ebenfalls gefordert, die Mädchen nicht in Sopran und Alt einzuteilen, sondern abwechselnd in beiden Stimmen singen zu lassen, um den vollen Ambitus zu entwickeln (Gend et al., 2017). Dies kann als Appell gelesen werden, trotz der Mutation die Stimme in einem möglichst großen Ambitus zu fördern, was der rein deskriptiven Beschreibung von Gackle per se nicht widerspricht. Gackle (2019, S. 560) zitiert zwei weitere Studien, die Einzelheiten der von ihr vorgeschlagenen Phasen untersuchten. Bottoms (1995) bestätigt das Phänomen des behauchten Klangs in Phase 2a und 2b mit einer Überprüfung von Klangbeispielen von 97 Probandinnen durch Lehrkräfte und Chorleiter, unterlässt aber eine Altersangabe der Studienteilnehmerinnen, wie Jones (1998) zurecht bemängelt. Williams (1990) vergleicht die Phasen 2A und 2B, wobei keine signifikanten Unterschiede in den Messparametern gefunden wurden, außer der Tatsache, dass in Phase 2B eine Zunahme der Mutationsphänomene zu beobachten ist. In einer aktuelleren Studie bestätigen Willis und Kenny (2011) die von Gackle beschriebenen Phänomene, fügen aber einen Zusammenhang mit dem Gewicht der untersuchten Mädchen hinzu und beschreiben zusätzlich psychische Auswirkungen, nämlich den durchgehenden Verlust des Vertrauens in die eigene Stimme: „It was also noted that all girls in the study (even those who had vocal training) lost confidence in their voice-use ability during this crucial year (age twelve to thirteen years).“ (234) Wiederum erscheint damit eine psychologische Komponente, die in der Phasenbeschreibung von Gackle nicht erwähnt wird, aber bereits in einem Artikel 2006 und dann 2011 in ihrem Lehrbuch *FINDING OPHELIA'S VOICE, OPENING OPHELIA'S HEART: NURTURING THE ADOLESCENT FEMALE VOICE* (Gackle, 2006, 2011) ausgeführt wird. Die Referenz im Titel an Shakespeare's Hamlet nimmt Bezug auf die selbstzerstörerischen Kräfte der Protagonistin Ophelia, die vor allem anderen gefallen möchte. Gackle beschwört die Wichtigkeit, in der schwierigen Zeit der Pubertät den ganzen Menschen zu betrachten: „The voice is an integral part of the total person, inextricably tight to thoughts and emotions.“

(Gackle, 2006, S. 35) In einer phänomenologischen Studie bestätigt Sweet (2015) die von Gackle antizipierten Schwierigkeiten:

„Singing experiences fostered a myriad of emotions for the female singers, from fear and humiliation to frustration and embarrassment, from confidence and pride to feeling powerful. Emotional reactions strongly encouraged or discouraged students’ moment-to-moment involvement in singing.“ (S. 80)

In der Studie wird der Verlust von Selbstvertrauen, den Willis und Kenny (2011) beschreiben, bestätigt und konkretisiert. Sweet beschreibt für alle ihre Interview-Partnerinnen das emotionale Karussell, das Gackle (2019, S. 553) ganz den hormonellen Veränderungen der Pubertät zuschreibt. Zusätzlich zu den Effekten der Stimmveränderung und den begleitenden Schwankungen und Einflüssen auf das Selbstkonzept bringt sie den Kontext des Singens ins Spiel: „The contexts of ‘where’ and ‘with whom’ greatly influenced the female singers’ willingness to sing.“ (Sweet, 2015, S. 81) Sie verlässt damit die Ebene der persönlichen Entwicklung einzelner Personen und bezieht die auftauchenden Schwierigkeiten zusätzlich auf das soziale Setting (vgl. den Begriff Jugend in Abbildung 2). Gerade die Aspekte, *wo* und *mit wem* gesungen werden soll, scheinen von zentraler Bedeutung zu sein, wobei zusätzlich beachtet werden muss, dass sich amerikanische Studien meist auf Chorsettings in Middle- und Highschools beziehen, was auf Freiwilligkeit hindeutet. Sweet betont, dass trotz erworbener Fähigkeiten die Mädchen ihre Stimme zurückhalten, egal ob sie korrekt singen können oder nicht, nur für den Fall „they sang something incorrectly“ (S. 81). Das soziale Setting muss also in der Betrachtung der Stimmentwicklung für ein musikpädagogisches Setting mitgedacht werden.

Auch in der englischen Literatur zur adoleszenten Stimme finden sich Hinweise zu den Mädchen. Ashley (2015) referiert in seinem Werk zum Singen in der Sekundarstufe 1 hauptsächlich die Studien von Gackle, fügt der Diskussion aber zwei spannende Punkte hinzu. Erstens merkt er an, dass „teachers (and too often choir directors) assume that girls’ voices do not change“ (S. 102), und zweitens begründet er den Widerstand gegen die Benutzung der Randstimme mit „popular vocal models in which only modal voice singing is used, with frequent excursion into belt zone“ (S. 102). Auch Welch (2019) bezieht sich auf Gackle, fügt aber ergänzend an, dass

im Rahmen ihres Phasenmodells große interindividuelle Unterschiede zu beachten sind⁶¹.

Obwohl die oben genannten Studien mit überschaubaren Samples den Anspruch an Repräsentativität nicht einlösen können und insgesamt ein Mangel an Studien zur adoleszenten Mädchenstimme in der musikpädagogischen Literatur (Sweet, 2015, S. 73) beklagt wird, kann die Mutation der Mädchen auf verschiedenen Ebenen beobachtet werden. Ganz offensichtlich gibt es durch biologische Reifeprozesse hörbare Stimmveränderungen im Bereich der Klangqualität und bezüglich des Ambitus, dies vor allem in der postmenarchalen Phase 2B. Zusätzlich sind negative Effekte auf das vokale Selbstkonzept zu beobachten, was zu Zurückhaltung und Unbehagen bei der Phonation führen kann. Nicht zuletzt scheint das soziale Setting einen Einfluss auf das Singen zu haben, zu beobachten sind vor allem Zurückhaltung aufgrund der persönlichen Unsicherheit.

Hinweise in der deutschsprachigen Literatur

Im Gegensatz zur englischsprachigen Literatur finden sich in der deutschsprachigen Literatur nur wenige und meist vage Äußerungen zum Thema Stimmwechsel der Mädchenstimme. Im Standardwerk *DIE STIMME* von Richter (2013) wird an zwei Stellen Bezug zum Stimmwechsel der Mädchen genommen, einmal bezüglich des kaum wahrgenommenen Stimmwechsels, da die mittlere Sprechstimmlage nur wenig abfällt (S. 188), und ein zweites Mal bezüglich des Selbstwerts, der bei Mädchen nicht gleich wie bei Jungen durch den Stimmwechsel tangiert wird (S. 163). A. Mohr (1997, 2014) widerspricht Richter insofern, dass er die Senkung der Sprechstimmlage infrage stellt, ohne diese zu belegen, ist aber bezüglich der Wahrnehmung des Stimmwechsels gleicher Meinung. Der kaum auffällige Stimmwechsel wird dadurch begründet, dass die Stimmlippen nur wenig und relativ koordiniert wachsen, was die Anpassung der Steuerung wenig erschwert. Er stellt weiter fest, dass es zu Tonproduktionsstörungen in Form von Heiserkeitsphasen kommen kann und durch das körperliche Wachstum die Stimme körperbetonter und dunkler wird. Generelle Tonum-

⁶¹ Largo und Czernin (2017, S. 19) bestätigen aus medizinischer Sicht die großen interindividuellen Unterschiede bezüglich des Eintritts in die Pubertät und bezüglich der Länge der Pubertätsphase.

fangsveränderungen verneint er und stellt fest: „Manche Mädchen können nach Beendigung der Mutation höher singen als vorher, manche tiefer, und bei vielen bleibt der Stimmumfang gleich.“ (S. 25) Beide Autoren nehmen keinen Bezug auf das Modell von Gackle (2019) und stehen zu zentralen Punkten im Widerspruch: Richter verneint im Allgemeinen hörbare Phänomene und sieht auch keine Einschränkungen bezüglich des Selbstkonzepts und A. Mohr beschreibt zwar Phasen von Heiserkeit, sieht aber weder ein Sinken der Sprechstimmlage noch eine Einschränkung des Ambitus.

Ebenfalls sehr kurz nimmt Bergen (1995) Stellung, spricht lediglich von einem behauchten Stimmklang, der auf fehlende Kraft zum Stimmlippenschluss zurückzuführen sei. Immerhin scheinen die alten Gesangsmeister Tosi und Agricola (1757), die er zitiert, darauf geachtet zu haben, dass nicht nur *Mannspersonen*, sondern auch *Frauenzimmer* ihre Stimme während der Mutation schonen, nämlich solange, „bis die erlangte tiefere Stimme wieder anfängt in Ordnung zu kommen“ (S. 149), wie sie in ihrer Gesangslehre ausführen (vgl. Abbildung 2). Ausführlicher geht Mecke (2007) diesen Phänomenen im Rahmen ihrer Studie zum Stimmwechsel in der deutschen Chorpraxis des 18. und 19. Jahrhunderts nach. Sie nimmt zwar Bezug auf die Phasen der Stimmentwicklung nach Gackle, lehnt aber deren Einsatz ab (S. 75), weil sie selbst Stimmwechsel als deutliche Veränderung des Stimmumfangs und der Sprechstimmlage definiert hat (S. 35, vgl. auch Abbildung 3, S. 109). So kommt sie zum Schluss, dass Mädchen keinen Stimmwechsel haben, was aus der Sicht der vorgelegten Studien als Trugschluss abgelehnt werden muss.

In einem aktuellen Artikel zum Thema Stimmtraining während der Mutation nennt der deutsche Phoniater Fuchs (2020) als Phänomene der Mutation einen rauen und luftigen Stimmklang und eine Einschränkung der Entwicklungsmöglichkeit der Stimme für beide Geschlechter. Ebenfalls erwähnt er mögliche Registerbrüche und betont, dass „the young man in particular has to accept his new vocal pitch as part of his personality before his vocal capacity increases again“ (S. 21). Er weist damit explizit auf die Auswirkungen der Mutation nicht nur bezüglich der Stimme hin, sondern auch auf die psychische Entwicklung in der Adoleszenz, sieht aber nachfolgend bei den Jungen eine „greater vulnerability“ (S. 21) als bei den Mädchen. Fuchs verneint die Existenz von Studien zur „biopsychosocial relationships in the mutation“ (S. 20) und nimmt auch bezüglich der Mutation der Mädchen nicht Bezug auf die Resultate amerikanischer

Studien. Tatsächlich gibt es im deutschsprachigen Raum keine Studien zur adoleszenten Mädchenstimme. Die intensive interdisziplinäre Forschung zum Stimmwechsel beschränkt sich auf die Mutation der Jungen⁶². Der anglofone Diskurs ist im deutschsprachigen Raum weder rezipiert noch in eigenen Studien nachvollzogen worden.

Einzig Freer und Detterbeck (2019) machen in einer deutschsprachigen musikpädagogischen Zeitschrift eine Aussage zur adoleszenten Mädchenstimme: „Im Unterstufenchor macht sich diese Veränderung meist im Laufe der 6. und besonders der 7. Jahrgangstufe recht deutlich bemerkbar. Der Klang der Mädchenstimmen verliert an Glanz, erscheint mitunter hauchig und weniger tragfähig.“ (S. 7) Die emotionale Seite wird aber wiederum nicht erwähnt.

Die Durchsicht der gesangspädagogischen Standardliteratur von A. Mohr (1997, 2014), Richter (2013) und Bergen (1995) zeigt, dass im deutschsprachigen Raum kein besonderes Bewusstsein für die Phänomene der adoleszenten Stimmentwicklung bei Mädchen vorhanden ist. Auch

Die jungen Mannspersonen, müssen zu der Zeit, wenn sie die Stimme verändern, nicht zu vielem, viel weniger starken Singen angestrengt werden; bis die erlangte tiefere Stimme wieder anfängt in Ordnung zu kommen. Ein gleiches ist bey jungen Frauenzimmern, fast in eben dem Alter, da sich bey dem männlichen Geschlechte die Stimme verändert, und zwar bey gewissen Umständen, welche die Frau Mama am besten wird bemerken können, zu beobachten. Widrigenfalls können bey beyden Geschlechtern, die sonst guten Stimmen leicht, wo nicht gar verlohren, doch wenigstens verderbet werden.

ABBILDUNG 2: Abdruck einer Passage zum Stimmwechsel von Mädchen und Jungen (Tosi & Agricola, 1757, S. 149)⁶³

⁶² In der Reihe KINDER UND JUGENDSTIMME von Michael Fuchs wird das Thema z. B. nicht aufgenommen (<https://www.logos-verlag.de/cgi-bin/engtransid?page=/Buchreihen/stimme.html&lng=deu&id=-> abgerufen am 21.09.2021).

⁶³ In lateinischer Schrift: „Die jungen Mannspersonen, müssen zu der Zeit, wenn sie die Stimme verändern, nicht zu vielem, viel weniger starkem Singen angestrengt werden; bis die erlangte tiefere Stimme wieder anfängt, in Ordnung zu kommen. Ein gleiches ist bey jungen Frauenzimmern, fast in eben dem Alter, da sich bey dem männlichen Geschlechte die Stimme verändert, und zwar bey gewissen Umständen, welche die Frau Mama am besten wird bemerken

Fuchs (2020) nimmt keinen Bezug auf Resultate amerikanischer Studien, ist sich aber des Mangels an Studien zur weiblichen Mutation bewusst (S. 30). Der Fokus auf Untersuchungen zu Knabenchören (z. B. Peterson, 2019) oder auch das erst wieder erwachende Interesse am Singen im allgemeinen Musikunterricht (Hoos de Jokisch, 2014; Oberschmidt, 2016) kann als Begründung für diese Lücke dienen, vermag aber nicht wirklich zu erklären, wieso die Forschung von Gackle (2019), Sweet (2015), Ashley (2015) und Welch (2019) aus dem anglofonen Raum nicht im deutschsprachigen Raum rezipiert wurde. Ein ähnliches Bild zeigt sich bei der adoleszenten Jungenstimme, wobei eine Annäherung der anglofonen und deutschsprachigen Positionen zu beobachten ist (Freer, 2020; Fuchs, 2020).

4.4 Die adoleszente Jungenstimme

Im Gegensatz zur Entwicklung der adoleszenten Mädchenstimme ist die Entwicklung der adoleszenten Jungenstimme ein Thema, das in der deutschsprachigen Musikpädagogik größere Beachtung findet. Die aktuelle Forschung ist allerdings seltsam unverbunden mit der Forschung im anglofonen Raum, was aber nicht immer so war. So bezieht sich der amerikanische Forscher Cooksey (2000) in seinen Arbeiten der späten 1970er-Jahre auf Frank und Sparber (1970a, 1970b), die in Wien eine Langzeitstudie zur Stimmentwicklung mit 5000 Kindern zwischen 7 und 14 Jahren machten. Diese wiederum nehmen aber später keinen Bezug auf die amerikanische Forschung, sondern veröffentlichen lediglich eine auf die Mutationsstimme ausgerichtete Liedersammlung (Sparber & Frank, 1980), in der stark komprimiert die Entwicklungsstufen Prämutation, Mutation und Postmutation dargestellt werden. Diese Einteilung wird in der deutschsprachigen musikpädagogischen Literatur meist ohne Beleg weitertradiert (z. B. A. Mohr, 1997; Mecke, 2007; Fuchs, 2020).

Das Modell von Cooksey (2000) beschreibt die Entwicklungsphasen des Stimmwechsels ausführlicher und wird in einem ersten Abschnitt dargelegt. Im Gegensatz zum vorigen Abschnitt wird allerdings auf eine

können, zu beobachten. Widrigenfalls können bey beyden Geschlechtern, die sonst guten Stimmen leicht, wo nicht gar verlohren, doch wenigstens verderbet werden.“

ausführliche Darstellung verzichtet, weil der Forschungsstand aktuell in mehreren Artikeln dargestellt wurde (Welch, 2019; Harrison & Williams, 2019; Freer, 2020). Im zweiten Abschnitt wird der Frage nachgegangen, ob in der Phase der Mutation gesungen werden soll oder nicht. Dieser Disput kann als Haupttreiber für die Untersuchungen zur adoleszenten Jungenstimme identifiziert werden, wobei „the voice-break traditionalists [...] and the progressive public school educators“ (Cooksey, 2000, S. 720) ihre Differenzen lange nicht ausräumen konnten, weil sich empirische und anekdotische Evidenz vermischten (S. 720). Der Abschnitt legt die Sichtweisen dar und zeigt den Stand der aktuellen Argumentation. In diesem Bereich finden sich auch einige Arbeiten aus dem deutschsprachigen Raum (Peterson, 2019; Fuchs, 2020). Im dritten Abschnitt interessiert die Frage, weshalb viele Jungen aufhören zu singen oder Mühe haben, ihren Stimmwechsel zu bewältigen (Ashley, 2015).

Die Phasen der Mutation bei Jungen

In der anglofonen Literatur zum Stimmwechsel der Jungen beziehen sich alle aktuellen Literaturreviews (Welch, 2019; Harrison & Williams, 2019; Freer, 2020) auf eine Langzeitstudie von Cooksey (2000), die ihrerseits auf dessen Feldstudien aus den 1970er-Jahren (Cooksey, 1977a, 1977b, 1977c, 1978) basiert. Cooksey präsentiert dabei ein Modell mit sechs Phasen, das im Anschluss tabellarisch aufgelistet wird (vgl. Tabelle 3). Ergänzend werden Ambitus und Tessitura von Jungen und Mädchen in Abbildung 3 visualisiert.

Cooksey bezieht sein Modell explizit auf die Studien von Frank und Sparber (1970a, 1970b), die in der Entwicklung der Jungenstimme drei Phasen unterscheiden: Prämutationsstimme, Mutationsstimme und Postmutationsstimme⁶⁴. Frank und Sparber schließen mit ihrer großangelegten Longitudinalstudie an einige ältere Studien (Mecke, 2007, S. 36 ff.) an und bestätigen die Resultate von Naidt et al. (1965), die in einer Untersuchung

⁶⁴ Frank und Sparber haben zusätzlich die Kinderstimmen in die Register Sopran und Alt eingeteilt. Dieser Umstand wird in der vorliegenden Arbeit nicht thematisiert, weil nicht auf die Kinderstimme fokussiert wird. Allerdings zeigt sich hier ebenfalls eine Differenz zur amerikanischen Haltung, welche die unmutierten Stimmen insgesamt als *treble voice* bezeichnen (vgl. Fußnote 59).

von 150 Jungen ein erstes dreiphasiges Modell entwickelten. Cooksey (2000) nennt verschiedene Nachfolgestudien, die seine Resultate aus den 1970er-Jahren bis zur Jahrtausendwende bestätigen. Mit der interkulturellen Studie von Cooksey und Welch (1998) wird das Modell auch außerhalb von Amerika überprüft und bestätigt. Die Akzeptanz des Phasenmodells von Cooksey zeigt sich schließlich in der durchgehenden Rezeption im anglofonen Raum.

Das Modell von Cooksey (2000, S. 721 ff.) umfasst sechs Stufen⁶⁵ der Stimmentwicklung (vgl. Tabelle 3), die aufgrund folgender Parameter eingeteilt werden: Gesamtstimmumfang, Tessitura, Stimmqualität (Maß der Einengung, Behauchtheit und Formantreichtum), Entwicklung der Register und mittlere Sprechstimmlage. Die Phase 1 betrifft die Prämutation⁶⁶, in der die Kinderstimme noch keine Anzeichen von Veränderung zeigt. Die anschließenden Phasen 2A, 2B und 2C betreffen die Mutation. Wichtige Aspekte in dieser Phase sind Einschränkungen im oberen Bereich, eine stetige Absenkung der Sprech- und Singstimme sowie eine stabilere Produktion von Tönen im unteren Bereich. Zusätzlich zeigen sich in Phase 2B Falsett und Pfeifregister. In Phase 3A erscheint zwar zuerst die neue Stimme, sinkt aber nochmals ab und ist zu Beginn noch stark reduziert im Ambitus. Erst in Phase 3B wird die Stimme konsolidiert, der Ambitus nimmt zu, erreicht aber noch nicht den Formantreichtum einer erwachsenen Stimme. Allerdings zeigt sich, dass die Mutation bei Jungen sehr unterschiedlich verläuft:

„In general, age is a poor predictor for establishing voice change stages, with any given age group likely to encompass several stages. It is possible for an individual to pass through all stages of adolescent voice change in twelve months, but it is also possible for this process to be much slower and to last several years.“ (Welch, 2019, S. 524)

⁶⁵ Die Nummerierung ist in verschiedenen Quellen uneinheitlich, ebenfalls die Angaben zu Tessitura und Ambitus. In der vorliegenden Arbeit wird Bezug genommen auf Welch (2019, S. 524) und Freer (2020, S. 7). Die Nummerierung wurde adaptiert (vgl. Fußnote 56). Die Inhalte wurden durch den Autor übersetzt.

⁶⁶ Im Gegensatz zu Frank und Sparber (1980) bedeutet hier Prämutation vor der Mutation und nicht Phase 2A.

Welch bezieht sich hier auf eine von Cooksey und Welch (1998) gemeinsam durchgeführte Studie, die eine Fülle von Einzelfällen zeigte, die sich schwer einordnen ließen. Der Eintritt der Mutation ist individuell ebenso verschieden, wie die Gesamtdauer und die Länge der einzelnen Phasen. Peterson (2019) beschreibt aus der Optik des Stimmbildners eines Knabenchors in einer aktuellen Studie die Mutation von 48 Jungenstimmen. Auffällig in seinem Sample sind auch die interindividuellen Unterschiede und die Beobachtung, dass einige Knaben ohne Einschränkung

Phasen der adoleszenten Jungenstimme

Nr.	Name der Phase – Charakteristiken der Stimme	Indifferente Sprechstimm- lage	Alter
		<i>Tessitura</i> Ambitus	
1	Prämutation: Unveränderte Stimme – klare Stimme – wenig Anzeichen von Behauchtheit	c' <i>cis'–ais'</i> <i>a–fis''</i>	
2A	Frühe Mutation: Mittlere Stimme I – Reduktion des Ambitus – Veränderung des Timbres der oberen Töne – Instabilität der gesungenen Töne, insbesondere im oberen Bereich – tendenziell höhere Anstrengung beim Singen – Tonqualität wirkt bemüht, angestrengter und behauchter	ais <i>ais–g'</i> <i>gis–c''</i>	(10) 12–13
2B	Mutation: Mittlere Stimme II – Stimmumfang sinkt um eine Terz – weitere Reduktion des Ambitus – hohe Töne eher instabil, tiefe Töne eher stabil – Falsett und Pfeifregister erscheinen	gis <i>gis–f'</i> <i>f–gis'(a')</i>	12–14
2C	Höhepunkt der Mutation: Mittlere Stimme IIA – Stimmumfang sinkt um eine weitere Terz – weitere Reduktion des Ambitus – hohe Töne eher instabil, tiefe Töne eher stabil – größte Instabilität der Töne und wenig klarer Klang	fis <i>fis–d'</i> <i>d–gis'</i>	13–14

3A Postmutation – Start der Stabilisierung: Neue Stimme (neuer Bariton)	d	14–15
– Ausklingphase	<i>dis–fis</i>	
– Stimmumfang sinkt um eine weitere Terz	H–d'(dis')	
– weitere Reduktion des Ambitus		
3B Postmutation – Bereinigung und Entwicklung: Junge erwachsene Männerstimme	H	15–17
– Expansion des Ambitus	<i>H–gis</i>	(18)
– Stimme wird klarer und weniger behaucht, hat aber noch nicht den Obertonreichtum der erwachsenen Stimme erreicht	G–d'	

TABELLE 3: Phasen der adoleszenten Jungenstimme adaptiert und ergänzt nach Cooksey (2000) und Welch (2019)

während des Stimmbruchs in der hohen Lage weitersingen konnten. Bevor im nächsten Abschnitt vertieft wird, ob Singen während der Mutation adäquat ist, wird die Rezeption des Modells in der deutschsprachigen Musikpädagogik besprochen.

Das Modell von Cooksey wird in der deutschsprachigen Literatur kaum erwähnt. So bezieht sich A. Mohr (1997, 2014) in seinem Buch zur Kinderstimmbildung auf das Modell mit den drei Phasen nach Frank und Sparber (1970a), ebenso wie Fuchs (2020), der aktuell ohne Beleg darauf Bezug nimmt. Selbst Mecke (2007) bezieht in ihrer Darstellung zum Stimmwechsel das sechsstufige Modell von Cooksey nicht mit ein, was möglicherweise ihrem historischen Fokus auf die deutsche Chorpraxis geschuldet ist. Interessant liest sich die Beschreibung des Stimmwechsels im Lehrmittel von Schnitzer (2008a), der sich zwar nicht auf Studien bezieht, aber wesentliche Merkmale der Phasen nach Cooksey beschreibt, dem Lehrmittel geschuldet mit einem didaktischen Fokus (S. 29 ff.). Einen direkten Bezug zu Cooksey machen Freer und Detterbeck (2019) in einer musikpädagogischen Zeitschrift. Sie beschreiben zuerst die drei Phasen nach Frank und Sparber, um dann als erste auch das sechsstufige Modell von Cooksey zu referieren, allerdings ohne die beiden Modelle im Text zu verbinden.

Singen während der Mutation

Ob Singen während der Mutation bei Jungen angemessen ist, wurde lange Zeit kontrovers diskutiert. Unterschiedliche Überzeugungen dazu stellt Freer (2020) nach wie vor bei Musiklehrkräften fest: „Music teachers collectively hold a variety of beliefs regarding the appropriateness of singing instruction during the period of male adolescent voice change.“ (S. 3) Der Diskurs ist alt und lässt sich auf zwei Anspruchsgruppen zurückführen. Auf der einen Seite stehen Chöre, vor allem im Bereich der Kirchenmusik, die auf die spezifische Klangqualität der Kinderstimme angewiesen sind, und auf der anderen Seite steht die Schule, welche im regulären Schulunterricht ihre Klientel nicht auswählen kann. Cooksey (2000) zitiert ein Memorandum der Londoner Bildungsadministration aus dem Jahr 1933, das sich gegen das Stimmbruch-Phänomen stellt:

„The ‚broken voice‘ is the outstanding problem of the boys’ secondary school. But it will be less difficult to deal with the problem if two simple truths are understood and acted on: (a) A boy’s voice never breaks. Physiologically the vocal cords lengthen at an age that varies with individuals. [...] (b) Singing and speaking are essentially the same process. During the changing period it is essential that voices not be strained; which means that talking and singing (within a restricted force and compass) are both definitely desirable activities (cited in McKenzie, 1956, p. 11).“ (S. 719)

Für das Singen und Sprechen wird aus medizinischer Sicht die natürliche Entwicklung hervorgehoben. So wie niemand in Betracht zieht, mit dem Sprechen aufzuhören – immer unter Vermeidung einer übermäßigen Anstrengung – soll auch weiterhin gesungen werden⁶⁷. Es wird aber darauf verwiesen, dass die adoleszente Jungenstimme eine pädagogische Herausforderung ist. Und ein Blick auf aktuelle Artikel zeigt, dass sich mittlerweile auch in der ersten Anspruchsgruppe die medizinische Sichtweise durchzusetzen

⁶⁷ Die Diskussion rund um das Singen während der Mutation zeigt sich auch im täglichen Umgang in der Gesellschaft. Sie kommt immer dann auf, wenn sich Musiklehrkräfte als solche zu erkennen geben. Sofort wird ein Beispiel erzählt, wo ein Musiklehrer (selten sind es Musiklehrerinnen) jemanden bittet, auf das Singen zu verzichten und in die hintere Reihe zu treten.

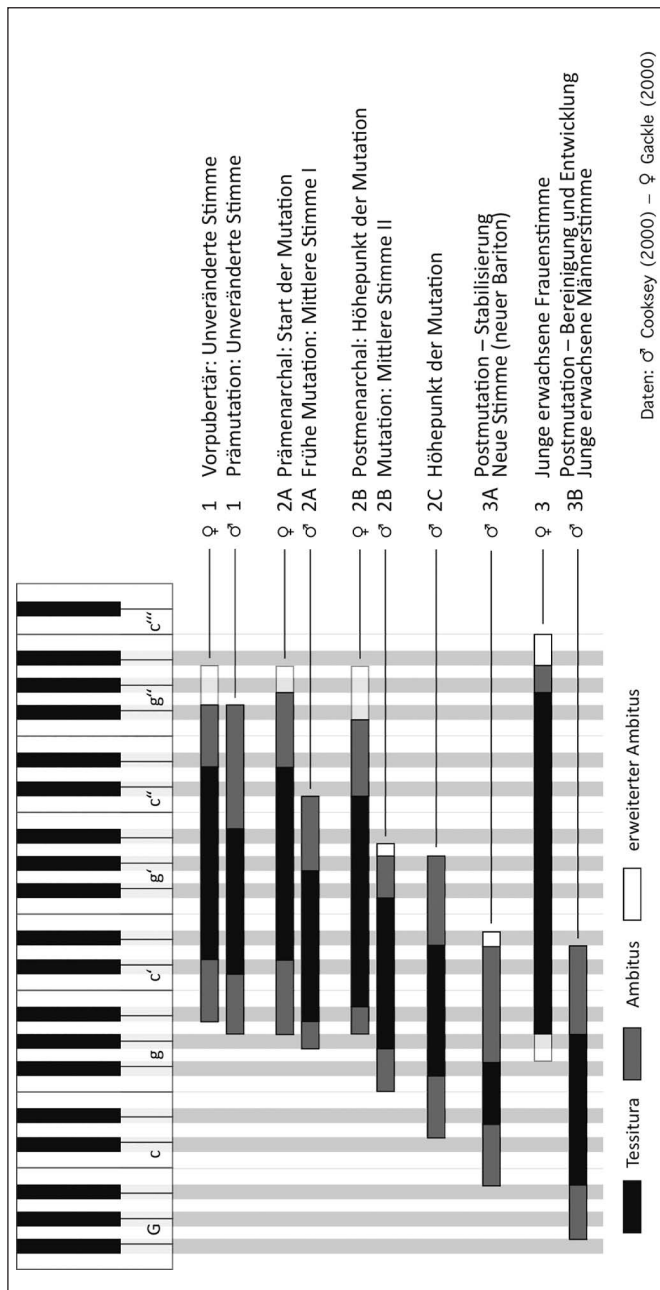


ABBILDUNG 3: *Ambitus und Tessitura der adoleszenten Stimme nach Welch (2019)*⁶⁸

Übersetzung und Adaption durch den Autor. Die Inhalte stammen von Gackle (2000) und Cooksey (2000). Die Nummerierung wurde von Gackle übernommen und auf die Phasen nach Cooksey übertragen. Es zeigen sich so die Hauptphasen Prämutation, Mutation, Postmutation, die von Frank und Sparber (1970b) geprägt wurden und im deutschsprachigen Raum gebräuchlich sind.

beginnt und der *Stimmbruch* nicht mehr als Auslöser für eine Gesangspause argumentiert wird (Fuchs, 2020).

Allerdings sind bis heute Knabenchöre mit der Mutation herausgefordert, denn die natürliche Entwicklung der Jungenstimme setzt dem Mitsingen im Sopran oder Alt auf eine natürliche Weise ein Ende. Zusätzlich ist der Stimmklang einer adoleszenten Stimme im Chorklang aus künstlerischer Sicht nicht mehr erwünscht. Der Stimmarzt Fuchs (2020), der die Leipziger Thomaner medizinisch betreut, fordert deshalb eine Absenz im Chorgesang, schlägt aber vor, während der Mutationszeit in kleinen Gruppen wöchentlich mehrere kurze Singtrainings zu machen, um die Wahrnehmung der eigenen Stimmentwicklung entlang der tiefer werdenden Sprechstimmlage bewusst zu erleben und akzeptieren zu lernen: „In particular, the young male voice succeeds in finding and accepting its new vocal home.“ (S. 29) Sein Anliegen ist der anschließende Wiedereinstieg ins Chorwesen, die vorgeschlagene Form der Betreuung verweist aber auch auf psychologische Faktoren, wenn er für das Weitersingen als Begründung zusätzlich das Kennenlernen, die Akzeptanz und den Zugang zur eigenen Stimme nennt.

Ein anderes Vorgehen wählt Peterson (2019), der die Jungen in der Sopranstimme weitersingen lässt, weil insbesondere trainierte Jungenstimmen in einer ersten Phase ihre Kinderstimme weiterhin einsetzen können, da in dieser physiologischen Phase die Beweglichkeit des Kehlkopfs noch groß ist. Peterson schlägt als Stimmbildner aber vor, diesen Einsatz unter Beobachtung der Jungenstimme und in Absprache mit dem künstlerischen Leiter abzustimmen, insbesondere um eine mögliche Kollabierung des Systems zu vermeiden. Der Wunsch der Jungen, möglichst lange im Chor mitsingen zu können und die große stimmliche Belastung stehen in dieser Zeit allerdings im Widerspruch zu der fragilen stimmlichen Konstitution. Harrison und Williams (2019) vermuten in diesem Zusammenhang, dass das Phänomen des „sudden collapse of the singing voice would explain the reported phenomenon of overnight voice change“ (S. 537). Sie empfehlen deshalb ähnlich wie Fuchs, „that the main body of the singing with a rapidly growing and vulnerable voice is within the fundamental comfort zone at the lower end of the pitch range“ (S. 537). Diese Empfehlung führt direkt zu der zweiten Anspruchsgruppe, die sich nicht an einem Konzert-repertoire messen lassen muss, sondern das Ziel hat, die Jungen während der Mutation weiterhin zum Singen zu bewegen.

Die Empfehlungen sowohl von Fuchs wie auch von Harrison und Williams bieten insbesondere für die Schule einen ersten Hinweis, wie mit der adoleszenten Stimme der Jungen umgegangen werden kann. Das Singen in komfortablen Lagen entlang der Sprechstimmlage scheint sich mittlerweile durchzusetzen, denn auch in der Studie zur adäquaten Tonhöhe für Jungen während des Stimmwechsels von Williams et al. (2021) wird vorgeschlagen, im tiefen komfortablen Bereich, also der Tessitura, zu singen. Sie empfehlen weiter, dies entlang des Systems von Cooksey zu tun: „The Cooksey system of five stages may be a useful guide for the majority of teachers and singers. It is relatively simple to implement, and it is unlikely to be misinterpreted.“ (S. 586 f.) Freer (2020) zählt weitere Elemente auf, die zu einem förderlichen Umgang während des Stimmwechsels führen: Songs⁶⁹ mit limitiertem Ambitus in einer adäquaten Tonlage, rhythmisches Singen in schnellen Tempi, Melodien mit Tonschritten statt Sprüngen und Arrangements mit Ostinati und Bordun-Klängen. „In other words, fit the song to the singers, not the singers to the song.“ (S. 8) Er begründet seine Empfehlungen insbesondere damit, die muskuläre Anpassungsleistung während der Mutation nicht zu strapazieren. An dieser Stelle sei auf den autoethnografischen Diskurs verwiesen, wo die Anpassung der Tonlage auf die individuellen Bedürfnisse von Jungen in einer Feldnotiz beschrieben wurde (vgl. 3.3) und intuitiv als Methode eingesetzt wurde (vgl. Brunner, 2014).

Zusammenfassend kann gesagt werden, dass gerade im regulären Schulunterricht der Titel dieses Abschnitts als Aufforderung gesehen werden darf: Singen während der Mutation! Es gibt keine Studien, welche Singen während der Mutation als schädlich einstufen. Vielmehr scheinen es praktische Gründe oder Unwissen zu sein, die Chorleitende und Musiklehrkräfte

⁶⁹ Wienhausen (2016) bespricht mögliche Stimmprobleme aufgrund von Pop- und Musicalgesang bei Kindern und Jugendlichen. Dies ist auf der einen Seite spannend, weil gerade die Literatur zu Knabenchören sich kaum auf die oft lauten Töne der Vollstimme beziehen, und auf der anderen Seite notwendig, weil gerade in der Schule Popsongs beliebt sind. Er führt die Angst vor den Klangqualitäten von Belt und Scream auf die ungeklärte Funktionsweise zurück und kommt zu dem Schluss, dass „die Angst vor lauten Popqualitäten in Bezug auf die Kinderstimme ungerechtfertigt“ (S. 123) ist, vorausgesetzt es besteht eine „effiziente Balance zwischen Atemstütze und Stimmführung“ (S. 123).

dazu bewegen, die Jungen fernzuhalten oder auszuschließen, wie auch Freer (2020) zumindest für die Musiklehrkräfte konstatiert: „[...] others lack foundational knowledge about the physiological, psychological and sociological changes that affect adolescent boys and their singing voices.“ (S. 3) Er fordert später im Text: „Because male adolescent singers experience changes in both range and the sensation of vocal coordination, music teachers need to take these issues into account when selecting repertoire and rehearsal methods.“ (S. 6) Er bezeichnet es als spezifische Aufgabe von Musiklehrkräften, mit geeigneten Methoden der adoleszenten Stimm-entwicklung zu begegnen. Auch Kalmbach (2010) sieht eine Verantwortung bei der Lehrkraft:

„Betreuung ist hier dringend angesagt, wird aber selten angeboten. Dass Stimmebeobachtung während der Mutation die weitaus beste Möglichkeit der nachhaltigen Stimmpflege ist, wird von alten Lehrmeinungen verneint. Die häufig verordnete Zwangspause signalisiert aber Stillstand, Kompetenzverlust, Unfähigkeit. Derart Gebrandmarkte verabschieden sich oft in diesem Stadium von ihrer Singstimme als Ausdrucksmittel. Auch hier sind die Knaben und jungen Männer meist die Verlierer.“ (S. 135)

Statt der Zwangspause schlägt Kalmbach Stimmebeobachtung und später im Text Einzel- und Gruppentrainings vor, aber insbesondere Engagement für die im gemischten Umfeld untergehenden Jungen (S. 141). Kalmbach formuliert hier – vielleicht auf etwas polemische Art – ein Defizit bei den Jungen. Zielvorgabe ist die Bildung von Selbstvertrauen. Die psychologischen Aspekte der Entwicklung dürfen also nicht vergessen gehen, wie sich bei Fuchs (2020) bereits in Ansätzen gezeigt hat.

Jungen in der Identitätskrise

Harrison und Williams (2019) stellen fest, dass die negative Einstellung zum Singen bei Jungen schon lange ein bekanntes Phänomen ist. Sie zitieren dazu einen Text aus dem Jahr 1946, in dem sich Lehrkräfte der Sekundarstufe 1 darüber betroffen zeigen, dass „most boys enter high school with negative attitudes towards vocal music and music education in general“ (S. 538). Sie machen dieses Phänomen vor allem an der männlichen Identitätsfindung fest: „The insecurity surrounding the physical attributes

of the voice, itself a symbol of changing from boyhood to manhood, causes boys to cease singing for fear of ridicule about their sexuality.“ (S. 539) In dieser Phase der rasanten körperlichen Entwicklung hören einige Jungen auf zu singen, weil sie ihr Singen nicht mehr mit ihrer neuen männlichen Identität vereinen können. In einer ganzen Reihe von Studien wurde das Singen bzw. Nichtsingen von Jungen untersucht (Harrison, 2007), wobei sich immer wieder die stereotype Vorstellung zeigte, dass Singen etwas für Mädchen sei. In den 1990er-Jahren verlagerte sich die Diskussion sozialer Phänomene im Singunterricht in den Genderdiskurs (vgl. Green, 1993, 1997), der im nächsten Abschnitt erneut aufgegriffen wird.

An dieser Stelle wird nun lediglich die Frage der Identitätsbildung vertieft, fokussiert auf die Bildung von *vocal identity* (Ashley, 2015). Er betrachtet den Prozess der Identitätsbildung als zentralen Faktor während der adoleszenten Stimmentwicklung. Das Auseinanderdriften von „musical identity“ und „vocal identity“ (S. 113) wird als Problemzone definiert, da sich das vokale Selbstbild und die performativen Möglichkeiten nicht decken. So wird es für einen 12-Jährigen schwierig, den favorisierten Rocksong, mit dem er sich identifiziert und der zu seinem musikalischen Selbstbild passt, adäquat nachzusingen. Ashley wendet sich deswegen auch gegen Castingsshow von Jugendlichen, in welchen diese ihre Vorbilder nachahmen. Vielmehr schlägt er vor, die Jugendlichen im Klassenchor singen zu lassen, wo auch unsichere Stimmen in einer gewissen Anonymität ausprobiert werden können⁷⁰. Dieser Vorschlag bietet zwar eine Erklärung für das Phänomen des Ausstiegs aus dem Singen und zeigt auch auf, wie die *schwierige Phase* überwunden werden kann, erscheint aber aufgrund des Vorschlags der Anonymisierung eher defizitorientiert, da die schwierige Phase ausgehalten oder überdauert werden soll.

Freer (2015) bietet mit dem Einbezug des Konstrukts der *possible selves* eine lösungsorientierte Strategie. Er bezieht sich auf Markus und Nurius (1986), die das Konstrukt folgendermaßen beschreiben:

⁷⁰ Harrison und Williams (2019) sehen (wie auch andere) in Castingsshow mögliche Motivationsmomente. Die Behauptung Ashleys (2015, S. 115), in den Shows würde mehr Wert auf Performance gelegt als auf das Singen, kann zwar nachvollzogen werden, die in den Shows auftretenden Naturtalente gibt es aber auch in den Schulklassen. Ihre vocal identity ist oft erstaunlich nahe bei den Originalen.

„Possible selves represent individuals' ideas of what they might become, what they would like to become, and what they are afraid of becoming, and thus provide a conceptual link between cognition and motivation.“ (S. 954)

Freer beschreibt verschiedene Stadien, entlang derer die erdachte Zukunft als Sänger zuerst entwickelt und dann realisiert werden kann. In seiner internationalen Studie zeigen sich aber Voraussetzungen, die bei der Lehrkraft vorhanden sein müssen, wobei weniger die kulturellen Gegebenheiten, als vielmehr die individuellen Unterschiede zum Tragen kommen. Als erstes nennt er Interesse an der adoleszenten Stimmentwicklung und geeignete Methoden dieser zu begegnen, als zweites sind es inspirierende musikalische Erfahrungen auf hohem Niveau und als drittes das soziale Umfeld vor und während der adoleszenten Stimmentwicklung, in welchem Modelle für mögliche *future selves* vorhanden sind (Freer, 2015, S. 103). Die Möglichkeit, eine Zukunft für die eigene Stimme zu sehen, beruht nach Freer also auf einer hohen Bereitschaft der Lehrkräfte, die Jungen durch die adoleszente Stimme zu führen, was sich mit der Aussage Kalmbachs (2010) weiter oben deckt: „Betreuung ist dringend angesagt.“ (S. 135) Und als weiterer wichtiger Punkt sind ältere Jungen als vokale Vorbilder gefragt, die in jahrgangshomogenen Klassen weniger zu finden sind, desto mehr in kulturellen Veranstaltungen der Schule, wenn Klassen älteren Jahrgangs oder der Schulchor auftreten.

Freer (2020) bringt eine weitere Betrachtungsweise ins Spiel. Er referiert narrative Forschung, welche die Jungen fragt, wie sie sich das Singen vorstellen. Die Jungen erwarten folgende Elemente im Unterricht:

- musikalische Erfahrungen, die ihren Fähigkeiten entsprechen,
- Wettbewerb in der Klasse und Lernspiele,
- Wissen zur Physiologie des Stimmwechsels,
- Fähigkeiten erwerben, statt Repertoireübungen,
- fähigkeitsbasiertes Feedback von Lehrkräften und Peers und
- ein „challenging, highly-rhythmic, multi-voiced repertoire“ (S. 10).

Zusammenfassend kann der Erwerb von Fähigkeiten, spezifisches Wissen und Wettbewerb als möglicher Zugang zum Singen genannt werden. Zusätzlich bevorzugen Jungen Rollenmodelle von vier bis fünf Jahre älteren Jungen, in denen sie sich als *possible selves* in der Zukunft vorstellen können. Als Autor verschiedener Studien zum Thema zieht er das Fazit, dass „boys want to be treated as the young men they are becoming“ (S. 11).

In der Entwicklung von Jungen zu Männern werden Schlüsselfaktoren manifest, die es zu beachten gilt. Die Jungen müssen in der Entwicklung eines vokalen Selbstkonzepts unterstützt werden, indem sie ihren Stimmwechsel bewusst erleben und sich mit Vorbildern auseinandersetzen können. Ergänzend hilft das Wissen zum Stimmwechsel und Feedback von Lehrkräften und Peers. Das Modell von Cooksey (2000) kann dabei Lehrkräften und auch den Jungen als Orientierungspunkt dienen. Trotzdem bleibt der Stimmwechsel eine Herausforderung, weil die individuelle Entwicklung im sozialen Setting passiert und ein geschützter Raum nicht immer geboten werden kann. Dabei machen sich gruppendynamische Phänomene bemerkbar, die im nächsten Abschnitt untersucht werden.

4.5 Geschlechtsspezifische Einflüsse beim Singen im Musikunterricht

Die Betrachtung der stimmlichen Entwicklung von Mädchen und Jungen in den beiden vorhergehenden Abschnitten fokussierte hauptsächlich auf die individuelle Entwicklung (vgl. Abbildung 1) von Schülerinnen und Schülern. In diesem Abschnitt wird nun eine soziologische Perspektive eingenommen, um die Wahrnehmung des Singens in der Schule zu untersuchen.

Mit Blick auf die vorliegende Studie wird das Klassengefüge als Gruppe definiert, bestehend aus Jugendlichen beiderlei Geschlechts sowie einer Lehrkraft. Das gemeinsame Tun im Singunterricht konstituiert sich aus dem Curriculum der Schule, das Musik bzw. Singen als Lerngegenstand definiert. Diese Gruppe ist aber unfreiwillig im obligatorischen Schulunterricht vereint (Cslovjecssek, 2009) und unterscheidet sich einerseits von freiwilligen Jugendchor-Angeboten, die bei Schaumberger (2020b) in den Fokus rücken, aber andererseits auch von (halb-)freiwilligen Angeboten, wie z. B. die Chorklassen (vgl. Schnitzer, 2008a; Gütay, 2012). Beide Angebote konstituieren sich aus dem gemeinsamen Anliegen des Singens. Gerade bei unfreiwilligem Musikunterricht ist dieses Ziel nicht per se vorhanden, sondern muss etabliert werden. Damit rücken Motivation und gruppendynamische Effekte verstärkt in den Vordergrund, wie auch Schaumberger (2020a) vermutet, wenn er ein Kompetenzmodell fordert, das „Kompetenzzuwächse im Singen bei gleichzeitiger Berücksichtigung

der motivationalen, volitionalen und sozialen Kompetenzen“ (S. 244) für die Schule fordert. Die Nähe zum Kompetenzmodell von Weinert (2001) ist unübersehbar, macht aber gerade beim Singen Sinn, weil weniger stark auf kognitive Aspekte wie in anderen Fächern fokussiert wird.

Aufgrund der körperlichen und psychischen Entwicklung, die gerade während der Pubertät auf das Geschlecht fokussiert, rückt im Blick auf die Gruppe die Kategorie Geschlecht verstärkt in den Fokus. Es gilt deshalb zu überlegen, welchen Einfluss die geschlechtsspezifischen Phänomene der adoleszenten Stimmentwicklung auf das gemeinsame Singen im Musikunterricht und spezifisch auf Förderung und Beurteilung haben. In den folgenden Abschnitten wird deshalb zuerst Geschlecht als Kategorie definiert und spezifische Perspektiven im Zusammentreffen von Mädchen und Jungen sowie Lehrerinnen und Lehrern benannt. Dazu gehören mögliche Einflüsse des Geschlechts der Lehrkraft auf den Musikunterricht im Allgemeinen und spezifisch auf das Singen. Eine weitere Perspektive beschäftigt sich mit der Wahrnehmung des Lerngegenstands durch Mädchen und Jungen und die Wahrnehmung der Jugendlichen durch die Lehrkräfte. Zuletzt werden Hinweise auf mono- und koedukative Unterrichtsphasen untersucht.

Geschlecht als zentrales Merkmal

Die Kategorie Geschlecht ist in der Gesellschaft ein zentrales Merkmal, wie Rüttgers (2016) schreibt: „Geschlecht ist weit mehr als eine Eigenschaft unter mehreren, sondern die zentrale Kategorie für die jeweilige Identität und die Wahrnehmung anderer Menschen, mit der zugleich auch unterschiedliche Zuschreibungen verbunden sind.“ (S. 17) In den beiden vorangegangenen Abschnitten (vgl. 4.3 & 4.4) wurde in aller Selbstverständlichkeit auf geschlechtsspezifische Phänomene von Mädchen und Jungen sowie deren Verhaltensweisen Bezug genommen. In keiner der zitierten Studien war der Begriff *Gender* im Sinne des sozialen Geschlechts (Rüttgers, 2016; Hess, 2018) ein zentrales Thema, weder in eher medizinisch geprägten Artikeln (Fuchs, 2020), noch in musikpädagogisch ausgerichteter Literatur (Sweet, 2015; Harrison & Williams, 2019; Welch, 2019; Freer, 2020). Das einzige Thema, das jeweils in der Kategorie Gender eingeordnet wird, betrifft den Zugang zum Singen, der mittlerweile ausgehend von Überlegungen von Green (1997) im überschaubaren Gender-

diskurs der Musikpädagogik besprochen (Ashley, 2009; Hess, 2015) und später diskutiert wird. Die zentrale Kategorie Geschlecht zeigte sich somit einerseits in der selbstverständlichen Betonung biologischer Reifeprozesse im dualen System von weiblich und männlich und andererseits als Genderphänomen in der Zuschreibung des Singens als weibliche Tätigkeit.

Es ist nicht die Absicht dieser Studie, bezüglich der Kategorie Geschlecht einen Beitrag zu der zunehmend politisierten Diskussion um Gender zu leisten⁷¹. Auch soll auf keinen Fall die Wichtigkeit eines achtsamen Umgangs mit Transkindern und -jugendlichen heruntergespielt werden⁷², vielmehr soll die Basis gelegt werden, die Möglichkeiten eines geschlechtergerechten Musikunterrichts auszuloten, um dies später mit der gelebten Wahrnehmung der Studienteilnehmenden abgleichen zu können. Für die vorliegende Studie wird somit Geschlecht als *zentrale Kategorie* gesetzt, womit die Interaktion zwischen den Geschlechtern zum Thema wird. Es wird darauf verzichtet, die Kategorie Geschlecht auf Biologie oder soziale Konstruktion festzulegen und sprachlich wird auf geschlechtsspezifische Aspekte verwiesen.

Die schematische Aufstellung der Protagonisten des Musikunterrichts (vgl. Abbildung 4) unter dem Blickwinkel der Kategorie Geschlecht verdeutlicht die Möglichkeit unterschiedlicher Perspektiven innerhalb der Gruppe. Neben Lehrerinnen und Lehrern sind es gemischte und geschlechtshomogene Teilgruppen, die miteinander interagieren, oder sich auf die Eigenwahrnehmung beziehen können. Der Verzicht auf weitere Optionen aus dem Bereich Transgender ist vor allem dadurch begründet, dass dazu im deutschsprachigen Genderdiskurs der Musikpädagogik keine Untersuchungen existieren und das Thema sich erst in den Schulen zu manifestieren beginnt⁷³.

⁷¹ Damit bewege ich mich gerade im Umfeld Pädagogischer Hochschulen auf Glatteis, weil der Bezug auf das biologische Geschlecht seine Rolle zunehmend verliert, wie z. B. der Ratgeber zu Diversity der Pädagogischen Hochschule der FHNW zeigt (Berger Kofmel, 2015, S. 27).

⁷² Siehe z. B. die Position von Valseccki von der Beratungsstelle Transgender Network Switzerland in Hunziker (2020).

⁷³ Gespräche mit Studierenden zeigen allerdings eine Tendenz, dass sich hier ein Forschungsdesiderat öffnet. So argumentierte eine Studentin im Rahmen einer Diskussion zu mono- und koedukativen Settings im Singunterricht, dass sie keine Gruppenbildung aufgrund des Geschlechts machen würde, und ein

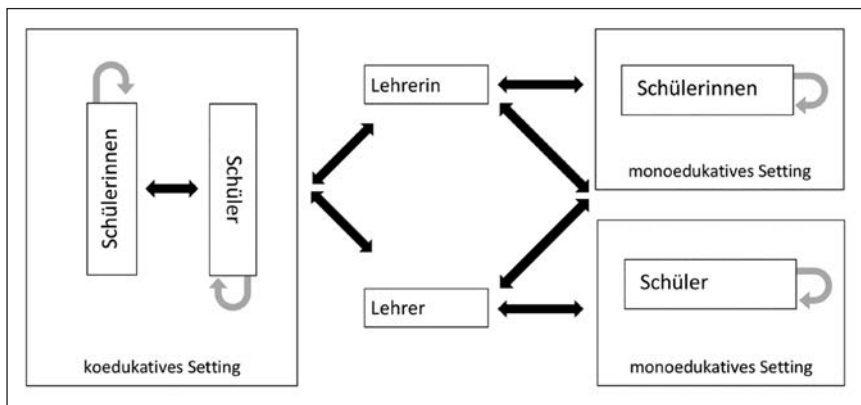


ABBILDUNG 4: *Geschlechtsspezifische Perspektiven im Musikunterricht in koedukativen und monoedukativen Unterrichtssettings*

Lehrerinnen oder Lehrer?

Es stellt sich also als erstes die Frage, ob das Geschlecht der Lehrkraft einen Unterschied macht. Die Metastudie von Helbling (2015) zum Einfluss des Geschlechts der Lehrkräfte auf den Bildungserfolg hat 42 Studien mit 2,4 Millionen Schülerinnen und Schülern in 41 Ländern und 8 Jahrzehnten untersucht, wobei die Mehrheit der Studien die Kompetenzen in Mathematik und Sprache erforschten. Nur eine der untersuchten Studien weist das Fach Musik explizit aus. Die Studie aus der Schweiz von Lupatsch und Hadjahr (2011) zeigt mit einem Sample von $n=827$, dass Lehrer im Gegensatz zu Lehrerinnen signifikant bessere Noten im Fach Musik erteilen und dass Lehrkräfte im Allgemeinen bessere Noten für Mädchen erteilen. Die in der Studie aufgestellte „These, dass Lehrerinnen Jungen benachteiligen würden, [kann] nicht gehalten werden“ (S. 198). In der Studie geht es aber nur um die Notengebung, also um summative Beurteilungsprozesse und nicht um formative Beurteilung. Zwar werden bei Helbling (2015) Studien erwähnt, die bei MINT-Fächern auf eine positive Wirkung von

Student berichtete von einer jugendlichen Person, die sich keinem Geschlecht zuordnen lassen will.

Lehrerinnen auf Schülerinnen schließen lassen, eine allgemeine Tendenz zur Vorbildrolle eines bestimmten Geschlechts zeigen sich aber nicht. So schlägt er vor, höchstens domänenspezifisch zu argumentieren und nicht „die Wichtigkeit gleichgeschlechtlicher Rollenmodelle für den Schulerfolg von Mädchen und Jungen per se herauszustellen“ (S. 8). Studien, welche die Rolle des Geschlechts von Lehrkräften domänenspezifisch für das Fach Musik untersuchen, gibt es keine.

Geschlechtsspezifische Rollenmodelle werden allerdings in einigen Texten für das Singen eingefordert, wurden aber bis dato noch wenig erforscht (Schaumberger, 2020b, S. 385). Freer (2020) schreibt: „Boys do need their teachers to be role models in many ways, however, and one of the most important is as a vocal model.“ (S. 11) Er reduziert das nicht nur auf die Lehrkraft, sondern schreibt auch älteren Schülern diese Möglichkeit zu. Er beschreibt zudem die Herausforderung für Lehrerinnen, die bei tiefen Tönen von den Jungen nicht in der Tonhöhe, sondern im Klang imitiert werden, was zu einer weiteren Oktavierung nach unten führt (S. 11). Dieses Phänomen wird als Timbral-Translationsfehler bezeichnet, der „beim Nachsingen berücksichtigt werden muss, da Menschen Töne akkurater nachsingen können, die ihnen vorgesungen werden, als solche, die sie von einem Klavier [...] abnehmen müssen“ (Merrill, 2020, S. 84). Der Unterschied zwischen männlicher und weiblicher Stimme wird entsprechend nicht nur aufgrund von „auditive[n] Informationen wahrgenommen, sondern auch [über] Muskelspannungen“ (S. 84) und entsprechend imitiert. Kalmbach (2010) beschreibt ebenfalls die Wichtigkeit von gleichgeschlechtlichen Rollenmodellen, nicht nur bei Jungen, sondern auch bei Mädchen, was üblicherweise in Klassen nur durch einen organisatorischen Aufwand erreicht werden kann. Neben der wenig diskutierten Rolle des Geschlechts der Lehrkraft bestimmen aber vor allem die Überzeugungen der Lehrkräfte zu den Fähigkeiten der Schülerinnen und Schüler das Unterrichtsgeschehen, wie sich im nächsten Abschnitt zeigen wird.

Geschlechtsspezifische Unterschiede in der Wahrnehmung des Singens

Bei der Betrachtung der adoleszenten Stimmentwicklung wurde in diesem Kapitel auf die spezifische Entwicklung von Mädchen (vgl. 4.3) und Jungen (vgl. 4.4) fokussiert, ohne die soziologischen Aspekte in dieser

Entwicklung zu beachten. Treffen die beiden Geschlechter zusammen, zeigt sich das feminine Image des Singens als Herausforderung für die Jungen. Dieses Thema zeigte sich bereits bezüglich der vokalen Identitätsbildung als relevant, wird aber im Rahmen des Genderdiskurses nochmals zugespitzt. Zudem werden die Vorstellungen der Lehrkräfte zum Singvermögen von Jungen und Mädchen im Singunterricht bedeutsam.

Oebelsberger (2009) provoziert in ihrem Artikel mit dem Titel SINGEN IST MÄDCHENSACHE ebenso wie Hess (2015) in DAS IST DOCH NICHTS FÜR ECHTE KERLE! mit stereotypen Aussagen zum Singen. In beiden Titeln werden in erster Linie die Vorstellung der Jungen zum Singen angesprochen. Harrison und Williams (2019) zeigen, dass seit mehr als einem Jahrhundert ein Forschungsfokus auf diesen Vorstellungen liegt. So wurde die Abneigung der Jungen und Männer gegen das Singen immer wieder untersucht (S. 537). Sie zitieren als Begründung den femininen Klang der nicht mutierten Jungenstimme: „The unchanged male voice sounds, to the untrained ear, like a female voice, and therefore singing with this voice can be considered feminine.“ (S. 538) Auch Ashley (2006) untersucht das Phänomen „You sing like a girl“ (S. 193) mit Bezug auf die englische Choral-Tradition und kommt zu dem Schluss, dass das feminine Image für die Entwicklung eines positiven Selbstkonzepts zum Singen problematisch ist: „If questions of sexuality and gender identity associated with the voice result in boys not singing during this period, then males are excluded from the opportunity of developing a life-long and life-enriching interest in singing.“ (S. 199) Mögliche Maßnahmen zugunsten der Jungen wurden bereits besprochen (vgl. 4.4), erhalten aber durch den Genderdiskurs eine neue Bedeutung.

Hess (2015, 2018) zeigt, dass auch im deutschsprachigen Raum das Fach Musik, im Speziellen die körperbetonten Elemente Singen und Tanzen, mit einem femininen Fachimage verbunden wird. Zudem zeigen Busch und Kranefeld (2012), „dass Jungen schon zu Beginn der zweiten Klasse ein deutlich weniger positiv ausgeprägtes Selbstkonzept in Bezug auf das Singen aufweisen als Mädchen“ (S. 19). Die Autorinnen und Autoren vermuten genau wie die bereits zitierten Autoren aus dem anglofonen Raum internalisierte Geschlechtsstereotype als mögliche Faktoren. Obwohl Hess (2015) dafür plädiert, die Jungen nicht aus den Augen zu verlieren und „unkonventionelles männliches Verhalten zu ermöglichen“ (S. 333), betont sie doch, dass „die geschlechtlich geprägten Fachimages

ein Spiegel gesellschaftlicher Hierarchien sind und damit letztlich hegemoniale Männlichkeitsvorstellungen stabilisiert werden“ (S. 333). Konkret verlangt sie für „den Musikunterricht, [...] dass er sich jeder Form der Stereotypisierung enthalten muss und das Geschlecht nicht durch explizite Thematisierung dieser Kategorie dramatisiert werden sollte“ (S. 330). Tatsächlich würde dies bedeuten, dass die spezifische Betreuung von Jungen, weil sie Jungen sind und in der adoleszenten Stimmentwicklung spezielle Bedürfnisse haben, zu vermeiden ist. McBride und Palkki (2020) untersuchen Chormaterialien und Methodenbücher im anglofonen Raum bezüglich Genderaspekten und konkretisieren die Kritik an diesen hegemonialen Männlichkeitsvorstellungen. Sie beschreiben die Untersuchungen zur adoleszenten Jungenstimme als „narratives that could be perceived as misogynistic and homophobic“ (S. 416). Sie bezeichnen die Bemühungen um die adoleszenten Stimmen der Jungen und deren Identitätssuche als Androzentrismus und damit als Herabminderung des Weiblichen. Diese Leseart steht im Widerspruch zum Anliegen der referierten Autorinnen und Autoren (z. B. Sweet, 2018; Freer, 2020; J. Williams et al., 2021), welche eine adäquate Förderung von Mädchen und Jungen einfordern. Hess (2018) ist sich des Querstands zwischen Jungenförderung und Genderpädagogik bewusst und lenkt den Blick auf die Verantwortung der Lehrkräfte, die ein „Verhalten jedweder Art zulassen und solange als gleichberechtigt anerkennen und wertschätzen [sollten], wie andere Menschen dadurch nicht in der Verwirklichung ihrer Bedürfnisse eingeschränkt werden“ (S. 212). Was dies für den Musikunterricht bedeuten kann, wird im nächsten Abschnitt diskutiert. Zuvor wird überlegt, welche Überzeugungen bei Lehrkräften zu beobachten sind.

In der Befragung englischer Lehrkräfte durch Green (1997) wird deutlich, dass eine Mehrheit der Lehrkräfte bei den Mädchen größeres Potenzial sieht, eine kleine Gruppe bei beiden Geschlechtern, aber niemand davon ausgeht, dass die Jungen besser singen könnten. In einem etwas älteren Artikel begründet Green (1993) dies mit den bereits besprochenen stereotypen Rollenbildern der Kinder:

„[...] both boys and girls tend to restrict themselves or find themselves restricted to certain musical activities and instruments for fear of intruding into the other sex's territory, where they might stand accused of a sort of musical transvestism.“ (S. 248)

Sie spricht dabei nicht nur das feminine Image des Singens, sondern auch die geschlechtstypische Wahl von Instrumenten an (vgl. auch Lehmann-Wermser, 2002; Harrison & Williams, 2019). Die Lehrkräfte in der Studie von Green beziehen sich in ihrer Argumentation auf die geschlechtsspezifischen Rollenbilder der Kinder und Jugendlichen, übernehmen dabei die beobachtete Haltung ihrer Schülerinnen und Schüler in ihren Zuschreibungen von Fähigkeiten. Tatsächlich zeigt sich diese Haltung auch aktuell noch in der mehrfach zitierten Studie von Blanchard (2019), in der die Musiklehrkräfte den Jungen vokale Fähigkeiten in geringerem Maß zutrauen. Repräsentative Studien, welche die Überzeugungen von Lehrkräften zur geschlechtsspezifischen Leistungsfähigkeit des Singens belegen, liegen nach aktuellem Wissensstand nicht vor. Allerdings wurden die Ansprüche an den Musikunterricht bezüglich des Umgangs mit Mädchen und Jungen verschiedentlich formuliert.

Bereits vor zwei Jahrzehnten forderte Lehmann-Wermser (2002) einen Musikunterricht, „wo Mädchen und Jungen ihre spezifischen Sichtweisen, ihre Fähigkeiten und Schwächen wiederfinden können. Ein Unterricht, in dem Mädchen auch am Schlagzeug sitzen können, ist gut; besser ist einer, in dem auch Jungen lernen sich singend auszudrücken“ (S. 17). Lehmann-Wermser nimmt dabei Bezug auf die Jungen, die Unterstützung benötigen, ihren Ausdruck beim Singen zu finden⁷⁴. Oebelsberger (2009) fordert später selbst, „ein generalisierendes Sprechen über Mädchen und Jungen“ (S. 84) aufgrund der Genderperspektive zu unterlassen, tendiert in ihren Ausführungen aber wie Lehmann-Wermser selbst zu stereotypisierenden Aussagen. So behauptet sie, Jungen würden sich zunehmend dem Singen verweigern, einen groben Stimmeinsatz pflegen und seien nicht bereit, „ihre eigene Stimme differenziert in Dynamik und Ausdruck einzusetzen“ (S. 86). Sie stellt dann die Forderung auf, dass die Musikerziehung hier Entscheidendes leisten könne, indem Jungen wieder vermehrt Ausdruck gewinnen und auf der anderen Seite Mädchen ihre starke Stimme entdecken sollten (S. 86). Einerseits zeigen sich in ihrem Statement ihre eigenen

⁷⁴ Diese Argumentation könnte zu einer Genderinitiative zugunsten der Jungen führen: Statt Physik für Mädchen, Singen und Tanzen für Jungen. Tatsächlich wäre es spannend zu sehen, was passieren würde, wenn mit der gleichen Intensität, wie ein mädchentauglicher Physikunterricht entwickelt wird, das Singen bei Jungen gefördert würde.

Zuschreibungen zu Mädchen und Jungen und andererseits macht sie deutlich, dass die Lehrkräfte auf die Vorstellungen der Jungen und Mädchen bezüglich des Singens einwirken sollen. Dies wird auch von Siedenburg (2015) mit Bezug auf Green gefordert: „[I]m Musikunterricht [spielen] auch die Vorstellungen eine Rolle, die Lehrer_innen von den musikalischen Fähigkeiten und Interessen der Mädchen und Jungen haben.“ (S. 13)

Wie dies geschehen kann, wird z. B. von Busch und Kranefeld (2012) vorsichtig im Sprachgebrauch des Genderdiskurses formuliert:

„Der Wirkung internalisierter Geschlechtsstereotype wäre möglicherweise durch *insensibles, doing*‘ und *un-doing gender*‘, einen situations- und fallspezifischen Umgang mit den strukturellen Benachteiligungen von Jungen beim Singen zu begegnen: Lehrerinnen und Lehrer könnten ihre Vorerwartungen an die Leistung von Jungen und Mädchen und ihren Umgang mit dem Singen überprüfen.“ (S. 20, Hervorhebung im Original)

Busch und Kranefeld schlagen vor, die Jungen durch geschlechtsspezifische Aktivitäten zu fördern, sehen sogar eine strukturelle Benachteiligung der Jungen und vermuten diese in den Zuschreibungen der Lehrkräfte. Als mögliche Form geschlechtsspezifischer Methoden beim Singen werden monoedukative Phasen während der adoleszenten Stimmentwicklung beschrieben.

Monoedukative Phasen im Musikunterricht

Bei der Betrachtung geschlechtsspezifischer Phänomene im Umgang mit Jugendlichen tritt im Bereich der Schule unweigerlich der Diskurs zu Mono- und Koedukation⁷⁵ in den Vordergrund. Pro und Contra werden zuweilen heftig diskutiert, aber selten empirisch belegt (vgl. Ludwig, 2003). Da in Amerika der Diskurs zu Single-sex-schooling (Crawford-Ferre & Wiest, 2013) vor allem polemisch aufgeladen ist, erscheint eine Diskussion

⁷⁵ Monoedukation, oder Seedukation, bedeutet das Trennen einer Gruppe zu Unterrichtender nach einem bestimmten Merkmal. Meist wird der Begriff als Gegenpol von Koedukation verwendet und beschreibt das Trennen nach Geschlecht. Im anglofonen Raum wird von *single-sex schools* gesprochen und teilweise der Begriff *gender-segregation* verwendet.

amerikanischer Literatur an dieser Stelle nicht sinnvoll. Mittlerweile wird im Rahmen von *reflexiver Koedukation* die Trennung der Geschlechter für bestimmte Phasen gepflegt, wobei Graff (2013) berechtigterweise einwendet, dass damit Koedukation das Prinzip und Monoedukation die Ausnahme sei. Sie schreibt:

„Aber geschlechterbewusste Erziehung sollte Pädagogik nicht länger mit Koedukation gleichsetzen, sondern sie vielmehr als das behandeln, was sie ist: eine der möglichen Organisationsformen im Hinblick auf Geschlecht. Mädchen und Jungen können gemeinsam oder getrennt unterrichtet werden, je nach Thema, Alter oder Ziel. Es scheint sinnvoll, beide Formen nicht als sich gegenseitig ausschließende Prinzipien zu verstehen, sondern sie als zwei mögliche pädagogische Settings zu nutzen.“ (S. 92)

Der Fokus soll demnach nicht mehr auf der Frage liegen, welches Setting ein besseres akademisches Resultat hervorbringt (z. B. Crawford-Ferre & Wiest, 2013), oder ob Mädchen im MINT-Bereich gestärkt werden sollen, wie das z. B. Budde et al. (2016) referieren. Stattdessen soll überlegt werden, welche der beiden Organisationsformen für die Ziele des Unterrichts förderlicher sind.

Meier (2008) untersucht in der einzigen vorhandenen Interventionsstudie zu Koedukation im Musikunterricht die Leistungsunterschiede in geschlechtergetrennten und gemischten Klassen. Er kann zwar keine statistischen Signifikanzen für den kognitiven Bereich des Musikunterrichts feststellen, sieht aber Potenzial in offenen Aufgaben für soziale, emotionale und motorische Bereiche bei einem monoedukativen Setting. Abschließend vermutet er in einer seiner neun Thesen, dass Monoedukation für die Entwicklung von Musikalität förderlich ist, insbesondere bei Mädchen das Selbstbewusstsein gefördert werden kann und es weniger Blockaden zum anderen Geschlecht gibt (S. 261 f.). Leider fehlen in der Studie Aufgaben zum Thema Singen, es kann lediglich vermutet werden, dass performative Elemente, die zum Resultat der Studie beigetragen haben, auf das Singen übertragen werden können.

Frauke Hess (2015) beschreibt für das Tanzen und Singen ebenfalls positive Effekte in geschlechtshomogenen Gruppen, insbesondere um schamhafte Momente auszuschließen: „Durch temporäre Geschlechtertrennung werden geschützte Räume eröffnet, in denen alle individuell mit

erweiterten Handlungsoptionen experimentieren können.“ (S. 330) Erfahrungen aus der Praxis zeigen, dass der Lösungsansatz geschlechtshomogener Gruppen sinnvoll ist. So beschreibt Kalmbach (2010) den Unterricht in einer neu gegründeten Singklasse folgendermaßen:

„Da der Unterricht dieser Singklassen von einer Lehrerin und einem Lehrer erteilt wird, gibt es zahlreiche Möglichkeiten der gemischtgeschlechtlichen und getrennten Unterweisung. So wird das Selbstbewusstsein jeder Gruppe ausgeprägt und gestärkt. So ist dies schon ein vollgültiger dualer Ansatz. Die Diskussion um eine geschlechtersensible Erziehung liegt ohnehin im Trend: Koedukation und Seedukation jeweils da, wo es Sinn macht.“ (S. 133)

Auch wenn dies ein Beispiel aus einer Chorklasse ist und lediglich die erfahrungsbasierte Sicht von Kalmbach beschreibt, bestätigt dieser Ansatz aus der Praxis die Resultate der oben genannten Studien. Auch bei Schaumberger (2020b) werden Chorleitende zitiert, die eine zeitweilige Trennung der Geschlechter befürworten. Ashley (2015) beschreibt für englische Verhältnisse das Diamond-Modell, in welchem in niedrigeren Schulstufen in gemischten Gruppen gesungen wird, in der Sekundarstufe 1 in geschlechtergetrennten Gruppen und anschließend wieder in gemischten Gruppen. In seiner Untersuchung zeigte sich, dass es mit diesem Modell gelingt, Jungen beim Singen zu halten: „This, with few exceptions, was the approach favoured by those members of the practitioner group who were most successful in engaging boys with singing.“ (S. 175) Er dramatisiert diese Aussage sogar: „It has been a near-universal experience for those running previously boy-only choirs to find that they lose almost all their boys once the choir is opened to girls.“ (S. 176) Auch Orton und Pitts (2019) kommen in einer empirischen Studie zu dem Schluss, dass Jungen in geschlechtshomogenen Gruppen profitieren können: „[T]he current study shows an additional benefit of single sex singing for boys in providing safe opportunities for them to learn ‚what my voice is like now‘.“ (S. 49)⁷⁶

⁷⁶ Interessanterweise praktiziert die Braunschweiger Domsingschule ein Modell, das die Trennung der Geschlechter auf die 2.–4. Klasse vorverlegt, um Mädchen- und Jungenfreundschaften zu lancieren. Diese helfen anschließend ab der 5. Klasse in gemischten Chorgruppen, dass sich die Kinder auf ihre Peers

Vor dem Hintergrund dieser Resultate macht es Sinn, die Bedürfnisse von Jungen und Mädchen ernst zu nehmen und Entwicklungsmöglichkeiten in einem sicheren Rahmen zu schaffen. Doing Gender bedeutet dann eine Inszenierung des Geschlechts in geschlechtshomogenen Gruppen (Siedenburg, 2015, S. 13) mit dem Ziel, den Jungen Raum zum Kennenlernen der neuen Stimme zu ermöglichen bzw. den Mädchen Raum für die Entwicklung von Selbstsicherheit zu geben. Und der Effekt wäre ein Undoing Gender, weil sich Knaben in der weiblich attribuierten Tätigkeit Singen entwickeln und ausleben können. Ashley (2015) betont aber, dass neben dieser organisatorischen Maßnahme Haltung und Vorstellungen der Lehrkraft von zentraler Bedeutung sind:

„The attitude and values of the teacher seem to play a more pivotal role in achieving gender equity and boys’ participation than do more tangible ‚fixes‘ such as the management of classroom layout or the application of specific subject knowledge about adolescent voices.“ (S. 187)

Ashley weist zurecht darauf hin, dass ein geschlechtergetrennter Unterricht beim Singen nur dann Vorteile bringt, wenn den Jungen zugetraut wird, dass sie singen können. Werden Jungen schlechte Singfähigkeiten zugeschrieben, führen monoedukative Phasen nur dazu, die herausfordernde Gruppe abzuspalten und sie damit weiter zu diskriminieren.

In diesem Kapitel wurde gezeigt, dass Singen eine anthropologische Konstante ist (Sundberg, 2019), die im Rahmen des Curriculums der Schule auch mit Bildungszielen verknüpft ist (D-EDK, 2015). In der Sekundarstufe 1 zeigt sich als Herausforderung die adoleszente Stimmentwicklung (Ashley, 2015). Um diese Entwicklung auf stimmlicher und psychologischer Ebene besser zu verstehen, wurden wenig rezipierte Studien aus dem anglofonen Raum hinzugezogen und deren entwickelte Phasenmodelle dargelegt (Cooksey, 2000; Gackle, 2019; Welch, 2019). Sowohl die Jungen als auch die Mädchen durchlaufen diese Phasen zu individuell unterschiedlichen Zeitpunkten, in unterschiedlicher Dauer und unterschiedlich intensiv. Dies macht den Umgang im Klassenverband an-

und nicht auf das andere Geschlecht fokussieren können. (So berichtet auf einer Tagung von Gerd Münden, Kantor am Dom in Braunschweig – <https://www.braunschweigerdom.de/braunschweigerdomsingschule> – abgerufen am 12.09.2021.)

spruchsvoll, weil sich die Gruppe sehr heterogen entwickelt. Zusätzlich zeigen sich während dieser Zeit Herausforderungen im Zusammenspiel der Geschlechter. Einerseits verbinden Jungen während ihrer Identitätsfindung Singen mit einem femininen Image. Das Konzept der *possible selves* wird dabei als Möglichkeit gesehen, Jungen über Rollenvorbilder ein positives Selbstkonzept zu vermitteln (Freer, 2015). Andererseits sind auch die Mädchen mit ihrem Stimmwechsel herausgefordert, erleben ähnliche Phänomene wie Jungen und sind in dieser Phase insbesondere wenig überzeugt von ihrer Selbstwirksamkeit (Sweet, 2018). Auch dieser zweite Aspekt wird in Singsequenzen in der Sekundarstufe 1 wirksam. Im Rahmen des Genderdiskurses wird zudem zurückhaltend von möglicherweise positiven Wirkungen geschlechtshomogener Unterrichtsphasen gesprochen.

Der Forschungsstand zum Singen dient nun als Grundlage für die Überlegungen zur ganzheitlichen Förderung des Singens im Musikunterricht.

5 Forschungsstand II

Assessment im Singunterricht

Der Fokus der vorliegenden Arbeit liegt auf dem Assessment des Singens. Der Bereich Singen wurde bereits ausführlich referiert, insbesondere um die Herausforderungen für das Assessment deutlich zu machen. Der Begriff Assessment selbst blieb aber bis jetzt undefiniert, abgesehen von dem Hinweis (vgl. 2.1), dass *assessment* im anglofonen Sprachraum eine deutlich offenere Bedeutung hat als im deutschen Sprachgebrauch und insbesondere im pädagogischen Bereich dazu verwendet wird, die Lern- und Beurteilungsprozesse und die damit verbundene Rolle der Lehrkräfte ins Zentrum zu stellen und neu auszurichten. Dabei spielen auch normative Aspekte mit, teilweise implizit, aber auch explizit, weil mit der Verwendung des Begriffs allgemein Schulentwicklung in Verbindung gebracht wird. Terhart (2014) beschreibt den aktuellen Wandel folgendermaßen:

„Der gegenwärtig zu beobachtende Wandlungsprozess des von Lehrern vollzogenen schulischen Prüfens und Beurteilens besteht [...] vielmehr darin, die Schülerbeurteilung von ihrem bislang dominierenden Charakter eines summativen, abschließenden Urteils über die Leistung eines Schülers in Richtung auf eine das Lernen begleitende, formative Unterstützung des Lernens zu transformieren.“ (S. 884)

Terhart macht deutlich, dass ein Wechsel weg von der Dominanz summativer Beurteilung hin zu formativer Förderung stattfindet. Die vorliegende Arbeit untersucht im empirischen Teil diesen Wandel spezifisch für den Singunterricht in der Sekundarstufe 1, indem die Idee eines ganzheitlichen Assessments als Perspektive auf die Interviews mit Musiklehrkräften projiziert wird. In Kapitel 5 wird dazu ein Bogen gespannt von der lerntheoretischen Verortung des Begriffs hin zu den Herausforderungen der Leistungsbeurteilung beim Singen, die bereits in Ansätzen als relevanter Kontext (vgl. 1.2) problematisiert wurden.

Im ersten Abschnitt (vgl. 5.1) wird der Begriff Assessment lerntheoretisch verortet und als holistisches Förder- und Beurteilungsinstrument etabliert. Dabei wird auch die Dichotomie von *formativ* versus *summativ* thematisiert (Lau, 2016) und abschließend das Assessment-Modell von Earl

(2013) eingeführt. Im zweiten Abschnitt (vgl. 5.2) folgt die Fokussierung auf den Musikunterricht, wobei vor allem Aktivitäten aus dem amerikanischen Bereich zur Sprache kommen, weil die postulierte Ganzheitlichkeit sich im deutschsprachigen Raum noch wenig etablieren konnte (Lehmann-Wermser, 2019). Im dritten Abschnitt (vgl. 5.3) folgt der Blick auf die deutschsprachige Community mit Blick auf die notwendigen Kompetenzen für ganzheitliche Assessment-Prozesse im Singen. Insbesondere fallen dabei fehlende Diagnosefähigkeiten bei Lehrkräften (Brunner, 2014) und fehlende Konzepte für die Leistungsbeurteilung (Schaumberger, 2020a) auf.

5.1 Assessment als holistisches Förder- und Beurteilungsinstrument

Individualisierung als Ausgangspunkt

Der Begriff *individuelle Förderung* hat sich in Deutschland nach dem Bekanntwerden der PISA-Resultate zu Beginn des Jahres 2002 flächendeckend als Anforderung an die Bildung etabliert (Klieme & Warwas, 2011, S. 805). Seither gilt in der öffentlichen Bildung individuelle Förderung unhinterfragt als Schlüssel zur Qualitätssicherung und ist verbindlich in allen Bundesländern (M. Hasselhorn et al., 2019, S. 376). Auch in der Schweiz wird individuelle Förderung als Konsequenz der heterogenen Volksschule als Selbstverständlichkeit erachtet (D-EDK, 2014, S. 10). Individuelle Förderung kann somit als pädagogischer Konsens angesehen werden⁷⁷, wobei

⁷⁷ Ein pädagogischer Konsens ist immer auch auf den Zeitgeist zu beziehen. Gerade die Resultate von groß angelegten Studien, wie PISA oder COACTIV lösen nicht nur weitere wissenschaftliche Studien aus, sondern setzen auch eine ganze Reihe von politischen Prozessen zur Förderung der Bildung in Gang. Ob und in welcher Form individuelle Förderung in weiteren 20 Jahren immer noch als zentrale Anforderung erachtet wird, kann kaum vorausgesagt werden. Gehofft werden kann allerdings, dass zur Wirksamkeit bis dann Daten über die Studien von Hattie (2009) hinaus und besonders für das Fach Musik vorliegen.

unklar ist, ob dies für den Musikunterricht ebenfalls gilt (vgl. im Anschluss die Diskussion zu Heterogenität von Linn, 2017).

Das Bewusstsein für individuelle Unterschiede von Lernenden hat bereits zwischen 1960 und 1980 zu einer „Umorientierung unterrichtlicher Strategien geführt“ (M. Hasselhorn et al., 2019, S. 380) und wird in der pädagogischen Psychologie als Perspektivenwechsel von direkter Instruktion mit der Sachlogik von Lehrinhalten hin zu adaptivem Lernen mit Blick auf die individuellen Voraussetzungen beschrieben (vgl. M. Hasselhorn & Gold, 2017, S. 269).

Aufgrund dieser Umorientierung entstanden verschiedene Unterrichtskonzepte wie z. B. Mastery Learning (variable Lernzeit), Adaptive Teaching (individueller Lernweg) oder Scaffolding (von Fremd- zu Selbststeuerung). Für alle genannten Konzepte bildet individuelle Diagnostik die Grundlage und „beruht auf [...] konstruktivistischen Annahmen über Lernprozesse“ (S. 385). Wie wichtig die Umorientierung der Überzeugungen zu Lernprozessen ist, zeigte sich in der COACTIV-Studie, die bezüglich der lerntheoretischen Ansichten von Mathematiklehrkräften zeigt, „dass von den konstruktivistisch geprägten Lehrkräften ein höheres Maß an kognitiver Aktivierung und konstruktiver Unterstützung ausging, was sich als leistungsförderlich erwies, und dass Lehrkräfte mit transmissiven Überzeugungen in geringerem Maße kognitiv aktivierend voringen, was sich für die Lernentwicklung als nachteilig herausstellte“ (S. 391). Der Wechsel hin zu adaptiven Lernprozessen beruht damit auf der Beachtung der individuellen Voraussetzungen und Lernstrategien der Lernenden und basiert auf konstruktivistisch ausgerichteten Überzeugungen der Lehrkräfte zum Lernen. Diese lerntheoretische Ausrichtung zeigt sich im anschließenden Abschnitt als Anknüpfungspunkt und Basis ganzheitlicher Assessment-Prozesse.

Die pädagogische Psychologie liefert zudem die Grundlagen zur Begründung interindividueller Unterschiede des Lernens. So beschreiben M. Hasselhorn und Gold (2017) in ihrem INVO-Modell fünf Bereiche von individuellen Lernvoraussetzungen, wobei sie sich in vier Bereichen auf „das Ende der 1980er Jahre von Pressley, Borkowsky und Schneider (1989) skizzierte Modell der ‚guten Informationsverarbeitung‘ [...], das sogenannte GIV-Modell“ (S. 65) beziehen:

- Aufmerksamkeits- und Arbeitsgedächtnisfunktionen, verantwortlich für Aufnahme und Verarbeitung von Informationen,

- Umgang mit und Qualität von Vorwissen,
- Nutzung und Regulation von Lernstrategien,
- motivationale Dispositionen und Selbstkonzepte und ergänzend
- Volition und lernbegleitende Emotionen. (S. 66 f.)

Diese unterschiedlichen Voraussetzungen für das Lernen dienen später als Ausgangspunkt, um Unterrichtsmaßnahmen bezüglich einer individuellen Förderung des Singens zu beurteilen. Den Blick auf das daraus abzuleitende Handeln von Lehrkräften formulieren M. Hasselhorn et al. (2019) folgendermaßen:

„Zusammenfassend lässt sich ‚individuelle Förderung‘ als pädagogisches Handeln unter konsequenter Berücksichtigung personaler Lern- und Bildungsvoraussetzungen definieren (vgl. Klieme und Warwas 2011, 808) oder auch als ‚pädagogische Handlungen, die mit der Intention erfolgen, die Begabungsentwicklung und das Lernen jedes einzelnen Kindes zu unterstützen, unter Aufdeckung und Berücksichtigung seines je spezifischen Potenzials‘ (Solzbacher und Behrens 2014, 566).“ (S. 378 f.)

Eine Unterrichtsgestaltung unter Berücksichtigung von individuellen Unterschieden der Lernenden strebt somit das Ziel an, Potenziale bestmöglich zu entwickeln. Unbestritten sind somit die individuellen Voraussetzungen von Schülerinnen und Schülern, wie damit umgegangen werden soll, aber schon. So scheint individuelle Förderung im Rahmen von Klassenunterricht nicht besonders effektiv zu sein. Hattie (2009) errechnet für den Begriff *individualized instruction* lediglich eine Effektstärke von $d=0.23$, was eine schwache positive Auswirkung anzeigt⁷⁸. Interessant sind aber die vielen von ihm genannten besonders wirkungsvollen Bereiche, die auf metakognitiven Aspekten des Lernens beruhen, wie z. B. *meta-cognitive strategies* mit $d=0.60$ oder Instruktionsstrategien wie *cognitive task analysis* mit $d=1.29$. Damit wird deutlich, dass sich insbesondere die Reflexion der Lernenden über das Lernen in der beschriebenen Umorientierung im pädagogischen Kontext als wichtig erweist. Formatives Assessment zeigt

⁷⁸ Die Daten zu Hatties Effektstärken wurden einem aktuellen Dokument (250+ Influences on Student Achievement, 2017) entnommen. Die Effektstärke $d=0.40$ wird dabei als Referenzpunkt zur Einschätzung von Effektstärken betrachtet. Effektstärken über $d=0.40$ übersteigen die natürliche Entwicklung.

sich dabei als richtungsweisende Anleitung für Lehrkräfte oder als Denkmodell, das direkt auf individualisierenden Lehr- und Lernvorgänge Bezug nimmt.

Assessment als holistisches Förder- und Beurteilungsinstrument

C. Schmidt (2020) verweist in ihrer Arbeit zu formativem Assessment zuerst auf die begriffliche Unschärfe, die „unter anderem auf das Vorherrschen unterschiedlicher Paradigmen und der damit verbundenen lerntheoretischen Annahmen über Lernen, Leistungsbeurteilung und Unterricht zurückzuführen (Filsecker & Kerres, 2012) [ist], wobei insbesondere konstruktivistische, kognitivistische sowie interaktionistisch-soziokulturelle Lehr-Lern-Theorien relevante Bezugstheorien darstellen“ (S. 8). Diese Unschärfe zeigt sich in vielen Texten zu Assessment, die entweder direkt darauf verweisen (Bennett, 2011; Mantie, 2019), den Begriff nicht diskutieren (Dixson & Worrell, 2016; N. Frey et al., 2018), oder wie Brophy (2019c) auf deren „philosophical foundation“ bzw. auf „primary beliefs“ (S. 919) verweisen. C. Schmidt selbst bezieht sich wie die meisten Autorinnen und Autoren auf Scriven (1967, 1972), der als erster auf den Unterscheid zwischen formativer und summativer Evaluation verwiesen hat, und nennt dann vier Bezugstheorien, die in formativem Assessment als lerntheoretische Basis in Betracht kommen (C. Schmidt, 2020, S. 14 ff.):

- konstruktivistische Lerntheorien, die auf die notwendige Konstruktionsleistung der Lernenden und deren individuelle Zugänge fokussiert,
- kognitivistische Sichtweisen, welche „Metakognition und Selbstwirksamkeit als zentral für die Entwicklung selbstregulierten Lernens beschreib[en]“ (S. 15),
- interaktionistische und soziokulturelle Perspektiven, die Lernen als sozialen Prozess sehen, wo die Stimulation durch geeignete Aufgaben und das Zusammenspiel mit den Peers erfolgt,
- und entwicklungspsychologische Theorien und Modelle, die aufbauende Lernschritte betonen. „Eine entwicklungsorientierte Diagnostik geht demnach (wie die entwicklungsorientierte Didaktik) davon aus, dass sich die Lernentwicklung aller Schülerinnen und Schüler in einer weitgehend ähnlichen Abfolge (Phasen, Stufen) vollzieht, wobei von einer natürlichen Variabilität der Entwicklung (Entwicklungsheterogenität) auszugehen ist (Heimlich & Wember, 2015).“ (S. 16)

Trotz der wenig konturierten Anbindung des Begriffs Assessment an bestehende Lerntheorien zeigt sich in der pädagogischen Literatur, dass formatives Assessment *en vogue* ist (Bennett, 2011; Lau, 2016). So setzt sich auch Earl (2013) vehement für formatives Assessment ein, wobei der Eindruck entsteht, formatives Assessment sei summativem Assessment vorzuziehen. Auch Winter (2015) suggeriert mit dem Titel seines Buches LERNDIALOG STATT NOTEN, summative Formen seien zugunsten formativer Formen zu überwinden. Dieser Eindruck wird von Bennet (2011) auf eine Metastudie von Black und Wiliam (1998, 2010) zurückgeführt, die formativem Assessment eine hohe Effektstärke zuschreibt⁷⁹ und oft als Begründung für den Einsatz von formativem Assessment zugezogen wird. Die Dominanz von formativem Assessment entspricht allerdings nicht der ursprünglichen Idee von formativer und summativer Evaluation.

Lau (2016) weist in ihren Überlegungen darauf hin, dass Scriven (1967) den formativen und summativen Aspekt als zwei verschiedene, miteinander verbundene Rollen von Evaluation definiert hat, gleich wie Bloom et al. (1971), die mit der Idee, vor dem summativen Assessment jeweils verschiedene Stufen von formativem Assessment zu durchlaufen, ein ganzheitliches Lernsetting skizzierten. Die Intention von formativem Assessment liegt bei Scriven und Bloom in der Vorbereitung und Leistungssteigerung summativer Tests. Dixson und Worrell (2016) argumentieren ebenfalls zugunsten einer neutralen Betrachtungsweise und zitieren zur Unterscheidung die Definition der American Educational Research Association (2014):

„Formative assessment involves gathering data for improving student learning, whereas summative assessment uses data to assess about how much a student knows or has retained at the completion of a learning sequence.“ (Dixson & Worrell, 2016, S. 153)

Diese Definition deckt sich inhaltlich mit der ursprünglichen Idee von Scriven und Bloom. Lau (2016) zitiert anschließend verschiedene Studien, welche die Dominanz formativen Assessments argumentativ zu überwinden versuchen. In einem der Artikel wird kritisch darauf hingewiesen, dass

⁷⁹ Bennet äußert sich allerdings kritisch zu dieser Metastudie, weil die verwendeten Studien zu divers gewesen seien.

Lernende immer von summativem Assessment angetrieben werden und dass „the effect it has on students is stronger than the positive effects of formative assessment“ (S. 517). Trotzdem wird vorgeschlagen, Curriculum, Lehrmethoden und die verschiedenen Ebenen von Assessment aufeinander abzustimmen, immer unter der Prämisse, dass Lernen individuell erfolgt. Dies kann als Warnung gelesen werden, formatives Assessment als Wundermittel zu sehen und dabei zu vergessen, dass in summativem Assessment ein starker Antrieb liegt⁸⁰.

Zusammenfassend zeigt sich formatives und summatives Assessment als ganzheitliches Gebilde, das nicht per se getrennt oder gegenseitig ausgespielt werden kann. Die Betonung von formativem Assessment ist in erster Linie der normativen Mission geschuldet, den Wandlungsprozess hin zu stärkerer Lernbegleitung, wie ihn Terhart (2014) zu Beginn des Kapitels beschrieben hat, zu beschleunigen. In dieser Richtung argumentiert auch Earl (2013), dessen Modell zu Assessment im Anschluss vorgestellt wird.

Das Modell von Earl

Earl (2003, 2013) veröffentlichte sein Assessment-Modell unter dem Titel *ASSESSMENT AS LEARNING, USING CLASSROOM ASSESSMENT TO MAXIMIZE STUDENT LEARNING* zuerst 2003 und dann in einer überarbeiteten Version 2013. Obwohl dieses Buch nicht an wissenschaftlichen Standards orientiert ist, sondern mit der Forderung eines lernzentrierten Schulunterrichts einen normativen Anspruch hat, ist seine Rezeption mit 1688 bzw. 1749⁸¹ Zitierungen beachtlich. Das Modell erhielt in den vergangenen 20 Jahren zunehmend Beachtung in der Bildungsadministration als Basis für Schulentwicklung in Primar- und Sekundarstufe (z. B. Kanada: Earl & Katz, 2006) sowie in der tertiären Bildung (z. B. Irland⁸²). Der Titel bezieht sich

⁸⁰ In dieser Diskussion wird auch sichtbar, dass formatives Assessment auf eine intrinsische Motivation abzielt und summatives Assessment extrinsische Motivation bedeutet. Der Autor betont dann die Kraft extrinsischer Motivation.

⁸¹ Zählung nach Google Scholar – 1688 abgerufen 15.09.2021, 1749 abgerufen 20.12.2021.

⁸² <https://hub.teachingandlearning.ie/resource/expanding-our-understanding-of-assessment-and-feedback-in-irish-higher-education>,

insbesondere auf die Einführung von *assessment as learning* (AAL) als Ergänzung zum Begriffspaar *formative* und *summative* Assessment und bezieht metakognitive Aspekte des Lernens mit ein. In diesem Abschnitt werden seine zentralen Forderungen für ein ganzheitliches Assessment erörtert.

Earl verweist zuerst auf die vielfältigen und sich teilweise widersprechenden Rollen der Lehrkräfte: „Many purposes for assessment support one another and sometimes compete or conflict with one another.“ (S. 36) Insbesondere die Diskrepanz zwischen Förderung und Selektion, die sich einerseits aus dem Anspruch des Individuums in der Klasse und andererseits aus dem Anspruch der Gesellschaft an die Schule ergibt, zeigt die Ambivalenz der Rollen. Earl bezieht sein Rollenmodell auf eine Arbeit von Wilson (1996) und nennt fünf Ausprägungen: *Mentor/in*, *Guide*, *Programmleiter/in*, *Buchhalter/in* und *Berichterstatter/in*⁸³ (vgl. Tabelle 4). Diese Ausprägungen beschreiben die unterschiedlichen Rollen der Lehrkraft in Bezug auf die verschiedenen Anspruchsgruppen. Insbesondere die Differenz zwischen der Mentorin, welche die Schülerinnen und Schüler individuell unterstützt, und dem Berichterstatter, der die Leistung nach außen kommuniziert, weist auf Rollenkonflikte hin. Dies wird

Rolle	Ziele
Als Mentor/in	gibt die Lehrkraft jedem Lernenden Feedback und Unterstützung.
Als Guide	sammelt die Lehrkraft diagnostische Informationen, um die Gruppe durch die anstehende Arbeit zu führen.
Als Programmleiter/in	nimmt die Lehrkraft Anpassungen und Überarbeitungen der Unterrichtspraktiken vor.
Als Buchhalter/in	führt die Lehrkraft Aufzeichnungen über die Fortschritte und Leistungen der Schülerinnen und Schüler.
Als Berichterstatter/in	informiert die Lehrkraft Eltern, Schülerinnen und Schüler sowie die Schulverwaltung über Fortschritte und Leistungen der Schülerinnen und Schüler.

TABELLE 4: Rollen der Lehrkräfte nach Wilson (1996) zitiert nach Earl (2013)

(abgerufen am 15.09.2021).

⁸³ Um die Sprache nicht unnötig umständlich zu gestalten, werden die Wortpaare abwechselnd in weiblicher und männlicher Form verwendet.

deutlich, wenn die Mentorin sehr erfolgreich Lernen in Gang setzen konnte, dies aber im Rahmen der Selektion trotzdem zu einer ungenügenden Bewertung führt. Die genannten Rollen werden in Tabelle 4 mit konkreten Zielen ergänzt und finden sich im Anschluss in verschiedenen Bereichen

Assessment for, as, of learning

(adaptiert nach Earl, 2003; Earl & Katz, 2006; Earl, 2013)

AFL assessment for learning

Die Rolle der Lehrkraft besteht darin,

- den Unterricht auf die angestrebten Ergebnisse abzustimmen,
 - die besonderen Lernbedürfnisse von Schülerinnen und Schülern sowie Gruppen zu ermitteln,
 - dementsprechend Materialien und Ressourcen anzupassen,
 - differenzierte Lehrstrategien und Lernangebote zu entwickeln, um einzelne Schülerinnen und Schüler in ihrem Lernen zu unterstützen,
 - und Schülerinnen und Schülern direktes Feedback zu geben und Instruktion zu erteilen.
-

AAL assessment as learning

Die Rolle der Lehrkraft besteht darin,

- gute Praktiken zu modellieren und vorbildhafte Beispiele zu zeigen,
 - die Fähigkeit zur Selbsteinschätzung vorzuleben und zu vermitteln,
 - Schülerinnen und Schüler bei der Festlegung von Zielen und der Überwachung ihrer Fortschritte anzuleiten,
 - gemeinsam mit den Lernenden klare Kriterien für eine gute Praxis zu entwickeln,
 - die Lernenden bei der Entwicklung interner Feedback- oder Selbstüberwachungsmechanismen anzuleiten, um sich mit der Mehrdeutigkeit und Ungewissheit vertraut zu machen, die beim Lernen von Neuem unvermeidlich sind,
 - regelmäßige und herausfordernde Gelegenheiten zum Üben zu bieten, damit die Schülerinnen und Schülern zu selbstbewussten, kompetenten Selbsteinschätzern werden können,
 - die metakognitiven Prozesse der Schülerinnen und Schüler sowie ihren Lernprozess zu beobachten und anschauliches Feedback zu geben und
 - ein Umfeld zu schaffen, in dem es für die Schülerinnen und Schüler sicher ist, Risiken einzugehen, und Unterstützung leicht verfügbar ist.
-

AOL assessment of learning

Die Rolle der Lehrkraft besteht darin,

- eine Begründung für die Durchführung einer bestimmten Lernbeurteilung zu einem bestimmten Zeitpunkt zu geben,
 - eindeutige und definierte Ziele des Lernens zu nennen,
 - Aufgabenstellungen zu erstellen, die es den Lernenden ermöglichen, ihre Kompetenzen und Fähigkeiten nachzuweisen,
 - mehrere Alternativen zur Beurteilung anzubieten,
 - kommunizierte und belastbare Beurteilungskriterien zu benutzen,
 - transparent darzulegen, wie diese Kriterien interpretiert werden,
 - das Bewertungsverfahren zu beschreiben und
 - Strategien für den Fall bereitzuhalten, dass es zu Unstimmigkeiten bei den Entscheidungen kommt.
-

TABELLE 5: Rolle der Lehrkraft in Assessment for, as und of learning nach Earl (2013) sowie Earl und Katz (2006)

des Assessment-Modells von Earl (2013) bzw. Earl und Katz (Earl & Katz, 2006) wieder. Beschrieben werden nun die Bereiche *assessment for learning*, *assessment as learning* und *assessment of learning*.

Die Tabelle 5 zeigt jeweils zusammenfassend die Aufgaben der Lehrkräfte in den verschiedenen Bereichen.

Assessment for Learning

Das Ziel von AFL ist die adaptive Gestaltung von Lernprozessen durch die Lehrkraft: „Assessment for learning is designed to give teachers information that will allow them to modify the teaching and learning activities in which students are engaged to differentiate and understand how individual students approach their learning.“ (S. 37) Dabei wird davon ausgegangen, dass Schülerinnen und Schüler individuelle Lernwege gehen, in diesen Wegen aber aufbauende Lernschritte durchlaufen. Die Lehrkraft findet mit AFL nicht nur heraus, was die Lernenden können, sondern sie erhält Informationen, „how, when, and whether students use what they know“ (S. 38). Durch diese Innensicht werden Vorkonzepte, Wissenslücken und Verunsicherung sichtbar und durch eine spezifische Betreuung bearbeitbar. Im interaktiven Prozess während der ganzen Lernphase wird nicht nur der Wissensstand, sondern auch die Motivation im Lernprozess sichtbar und damit bearbeitbar. AFL bezieht sich auf konstruktivistische Lerntheorien

und entwicklungspsychologische Theorien und Modelle, wie sie zuvor ausgeführt wurden.

Die Lehrkraft kann in dieser Phase den individuellen Lernbedarf ihrer Schülerinnen und Schüler durch drei Bereiche abschätzen, erstens durch das Wissen zum Lernkontext aufgrund des Curriculums, zweitens aufgrund des persönlichen Wissens über die Lernenden und drittens auf der Grundlage formativer Assessments, also spezifischer Beobachtungen und Messpunkte während des Unterrichts (vgl. Tabelle 5). Die Lehrkraft fungiert dabei auch als Buchhalter, indem sie mit Arbeitsinstrumenten wie Checklisten, Arbeitsblättern, Portfolios und weiteren für das Fach adäquaten Artefakten arbeitet. Als zentrale Kompetenz werden aber von Earl diagnostische Fähigkeiten vorausgesetzt, um AFL wirksam werden zu lassen (S. 38 f.).

Assessment as Learning

Earl beschreibt AAL als Teilmenge von AFL, wobei AAL als „a process of developing and supporting metacognition for students“ (S. 39) dient. AFL wird so gestärkt und erweitert, denn die Lernenden sind nicht nur „contributor to the assessment and learning process, but the critical connector between them“ (S. 39). Der Fokus bei AAL liegt somit auf den Schülerinnen und Schülern, die ihr Lernen reflektieren, mit den Ungewissheiten des Lernens umgehen können und sich so weiterentwickeln. Zudem wird von Schülerinnen und Schülern erwartet, ihr Lernen zunehmend eigenverantwortlich zu gestalten⁸⁴. Lerntheoretisch wird damit eine kognitivistische Perspektive sichtbar.

Die Rolle der Lehrkraft besteht darin, Assessment so im Unterricht einzubauen, dass Lernende die notwendigen Fähigkeiten entwickeln können, ihr Lernen selbstkritisch zu analysieren. Schülerinnen und Schüler müssen

⁸⁴ Die Wirksamkeit solcher Mechanismen wurde bereits referiert (Hattie, 2009). Die Antwort in didaktischer Hinsicht wurde von Hattie (2012) mit dem Begriff Visible Learning geprägt und hauptsächlich im Aufgabenbereich der Lehrkräfte verortet (AFL). Neu schwenkt die Diskussion zusätzlich auf die metakognitive Ebene (AAL), wie ein aktueller Buchtitel mit Beteiligung von Hattie zeigt: DEVELOPING ASSESSMENT-CAPABLE VISIBLE LEARNERS (N. Frey et al., 2018).

dazu eine komplexe Kombination von Fähigkeiten, Haltungen und Verhaltensweisen entwickeln. Dies stellt sich nicht spontan ein, sondern muss vorgelebt werden (Earl & Katz, 2006, S. 42). Eine besondere Bedeutung erhält ein angenehmes Lernklima, sodass Risiken angstfrei eingegangen werden können (vgl. Tabelle 5).

Assessment of Learning

Earl beschreibt AOL als dominierende Art des Assessments in der Schule. Der summative Charakter zeigt die relative Position der Lernenden in der Gruppe und wird üblicherweise nach einer Lernphase als Abschluss durchgeführt, um das Erreichen von Lernzielen zu überprüfen und Rechenschaft gegenüber Eltern und Schulbehörden zu garantieren. Oft werden die Resultate mit Buchstaben und Zahlen symbolisiert. Aufgrund von AOL werden Laufbahnentscheide getroffen, was diese Form der Bewertung besonders anfechtbar macht. So betreiben Lehrkräfte oft einen großen Aufwand, um die Resultate von Prüfungen zu legitimieren.

Die Rolle der Lehrkraft besteht darin, Transparenz bezüglich der Beurteilungsstrategie herzustellen und belastbare Kriterien für ihre Assessments (vgl. Tabelle 5) zu definieren. Earl verweist einerseits auf die breite Akzeptanz summativer Beurteilung, nennt aber auch Bedenken bezüglich der Validität dieser Form von Assessment: „[T]here will be more attention to how grades are calculated and how well they actually reflect what they are taken to mean.“ (S. 40 f.) Damit spricht er die oft fehlenden psychometrischen Teststandards in summativen Bewertungen an. Die Frage der Validität wird später (vgl. 5.2) unter dem Aspekt Qualität in den *principles of assessment* (Brophy, 2019c) nochmals aufgenommen.

„Getting the Balance Right“ (Earl, 2013, S. 41)

Abschließend betont Earl, dass die richtige Balance der drei Bereiche AFL, AAL und AOL zu finden sei. Dabei wird noch einmal der normative Charakter seines Assessment-Modells sichtbar, da er den Bereichen AFL und AAL eine dominante Rolle zuweist, wie sich in Abbildung 5 zeigt: „The major focus is on classroom assessment that contributes to student learning, by the teacher (for learning) and by the student (as learning).“ (S. 43) Insgesamt erwartet er, dass AFL und AAL den Unterricht steuern und AOL

dann stattfindet, wenn „summative description is important for students and for others, as a milestone or rite of passage“ (S. 44). In der vorliegenden Arbeit steht nicht die Gewichtung der einzelnen Bereiche im Vordergrund, obwohl es sicher spannend sein wird, diese im Datenkorpus zu untersuchen, sondern die inhaltliche Fokussierung und die Ansprüche an die einzelnen Elemente. Zudem wurde bis dahin der ganzheitliche Anspruch des Modells in den Vordergrund gestellt.

Inwiefern sich das beschriebene Assessment-Modell im anglofonen Raum in der Musikpädagogik wiederfindet, wird im nächsten Abschnitt untersucht.

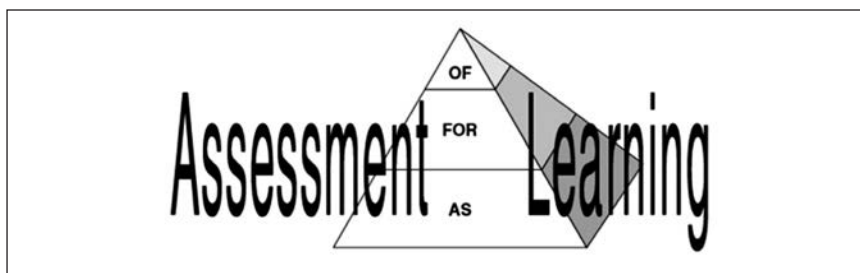


ABBILDUNG 5: Assessment-Pyramide nach Earl (2013, S. 43)

5.2 Assessment in Music Education

Der englische Titel dieses Abschnitts verweist auf die internationale Ausrichtung mit Betonung englischer und amerikanischer Autoren. Seit der Herausgabe des Standardwerks *ASSESSMENT IN MUSIC EDUCATION* von Fautley (2010) bis zum Erscheinen der Handbücher *THE OXFORD HANDBOOK OF ASSESSMENT POLICY AND PRACTICE IN MUSIC EDUCATION* (Brophy, 2019a, 2019b) und *THE OXFORD HANDBOOK OF PHILOSOPHICAL AND QUALITATIVE ASSESSMENT IN MUSIC EDUCATION* (Elliott et al., 2019) sind knapp zehn Jahre verstrichen. Die breite Diskussion zu Assessment führte zu einer internationalen Konturierung des Begriffs für den Musikunterricht und zu Assessment-Prinzipien (Brophy, 2019c), die in diesem Abschnitt als Basis für die Fokussierung auf Assessment im Singunterricht dienen. Im Zentrum stehen dabei nicht theoretische Überlegungen, wie sie C. Schmidt (2020) gemacht hat, sondern die praktischen Auswirkungen auf den Musikunterricht, wie Mantie (2019) sie ebenfalls in seinem Literaturreview beobachtet: „The music education assessment literature, in my reading, is heavy

on prescription and certainty, and light on theorization and provisionality.“ (S. 33) Im zweiten Teil des Abschnittes werden ausgewählte Instrumente für die Beurteilung des Singens beschrieben.

Principles of Assessment

Brophy (2019c) deklariert als *State of the Art* am Ende der beiden Handbücher (Brophy, 2019a, 2019b) acht Prinzipien für das Assessment in der Musikpädagogik⁸⁵. Diese Prinzipien gründen auf drei Überzeugungen, die als „philosophical foundation“ (Brophy, 2019c, S. 919) bzw. als fundamentale Wahrheiten für das Lernen von Musik gesetzt werden:

- „1. All Students are inherently musical. Every person has the capability to experience, learn, and engage with music meaningfully to the extent possible.
2. Assessment of music learning is best when it is ongoing and not episodic. Music learning progresses over time and the assessment of that learning is a continuous process.
3. Assessment reflects the multidimensional nature of music learning. Music is a complex, multidimensional human experience. The assessment of music learning must reflect those characteristics.“ (S. 919)

Obwohl die Aussagen einen normativen Impetus haben und darin wenig Vorläufigkeit (*provisinality*) zu finden ist, lassen sie sich doch an die bereits ausgeführten Inhalte in Kapitel 4 und 5 anschließen. So wird in der 1. Aussage davon ausgegangen, dass alle Schülerinnen und Schüler musikalisch sind und an musikalischen Lernprozessen teilnehmen können. Dies kann insbesondere für das Singen als gegeben betrachtet werden, da Singen als anthropologisches Merkmal beschrieben wird (Neus, 2017; Sundberg, 2019). Auch bezüglich der Singqualität konnte gezeigt werden, dass die meisten Menschen in der Lage sind, Melodien akkurat zu intonieren

⁸⁵ Die *International Principles for Assessment in Music Education* bilden eine gemeinsam erarbeitete Erklärung der Organisation *The International Symposium on Assessment in Music Education (ISAME)*, die seit 2007 existiert und das Thema Assessment als Forschungsgemeinschaft für die Musikpädagogik betreut (Brophy, 2019c, S. 918).

(Dalla Bella et al., 2007). In der 2. Aussage wird musikalisches Lernen als kontinuierlicher Prozess beschrieben, um daraus die Notwendigkeit eines stetigen Assessments herzuleiten. Mantie (2019) beschreibt die natürliche Anlage von Musikunterricht als fortlaufendes Assessment: „Indeed, day-to-day instruction in music involves in situ informal assessments und evaluation without which good teaching would not be possible.“ (S. 35) Der Fokus auf einen kontinuierlichen Assessment-Prozess deckt sich mit dem Modell von Earl (2013), das den formativen Elementen und somit den Lernprozessen mehr Gewicht beimessen möchte. Zuletzt macht Aussage 3 deutlich, dass die drei Überzeugungen nicht nur als Grundlagen für das Assessment im Musikunterricht, sondern auch als Forderungen zu sehen sind. Erwartet wird, dass sich der multidimensionale Charakter des Musiklernens im Assessment widerspiegelt. Für das Assessment im Singunterricht erwies sich ein mehrperspektivischer Blick auf biologische, entwicklungspsychologische und geschlechtsspezifische Momente während der adoleszenten Stimmentwicklung als zentral. Assessment soll somit die multidimensionalen Gegebenheiten des Singens nicht nur beachten, sondern auch in den verschiedenen Modi von Assessment widerspiegeln.



ABBILDUNG 6: *International principles of assessment in music education (Brophy & Fautley, 2017)*

Die im Anschluss vorgestellten Prinzipien für Assessment im Musikunterricht (Brophy, 2019c, S. 918 ff.) reflektieren die bereits dargelegten gegensätzlichen Ansprüche von formativem und summativem Assessment. Im Sinne einer Fokussierung werden die acht Prinzipien nach Möglichkeit auf die Thematik des Singens bezogen. Die Abbildung 6 zeigt im Überblick die acht Bereiche (Brophy & Fautley, 2017).

Shared Language⁸⁶: „Assessment in music education should be discussed using commonly accepted definitions of assessment, measurement, and evaluation.“ (Brophy, 2019c, S. 919)

Im ersten Prinzip wird eine gemeinsame Sprache gefordert, wobei nicht nur sprachübergreifend auf unterschiedliche Bedeutungen der verwendeten Termini rund um Assessment verwiesen wird, sondern auch auf die unterschiedlichen Bedeutungszuweisungen innerhalb einer Sprache. Im deutschsprachigen Raum ist zudem zu beachten, dass der Begriff insgesamt wenig etabliert ist, wie bereits ausgeführt wurde (Lehmann-Wermser, 2019; C. Schmidt, 2020).

Das Verständnis von Assessment wird in anglofonen Ländern zusätzlich von zwei Phänomenen beeinflusst. Mit *accountability* wird beschrieben, dass Schulen und Lehrkräfte verantwortlich gemacht werden für den Erfolg der Schülerinnen und Schüler. Im Rahmen von *high-stake testing* führt dies mit üblicherweise standardisierten Tests zu „school performance-based funding allocations, and, in some cases, the continuation of teachers' employment“ (Brophy, 2019c, S. 909). Mantie (2019) beschreibt in dieser Hinsicht die Gefahr, dass Assessment nicht zugunsten des Lernens der Schülerinnen und Schüler durchgeführt wird, sondern „assessment of teacher and teaching“ (S. 38). Nach Mantie werden die beiden Phänomene aber in gewissen Kreisen der amerikanischen Musikpädagogik als notwendige Maßnahmen zur Legitimation des Kernfachs Musik anerkannt. Für den Musikunterricht sieht er darin allerdings eine Gefahr:

„Many districts are unlikely to invest the necessary resources to assess holistically and will resort to the least expensive and easiest solution: paper and pencil tests on simple-to-score items such as note names.“ (S. 40)

Testungen im Fach Musik zur Evaluation von Lehrkräften oder gar zur Allokation von Mitteln werden sicher in der Schweiz, wohl aber auch in den anderen deutschsprachigen Ländern nicht durchgeführt. Allerdings wird sich zeigen, dass die im Zitat beschriebene basale summative Beurteilung

⁸⁶ Aufgrund der internationalen Ausrichtung der Prinzipien wird auf eine Übersetzung verzichtet.

nicht weit weg ist von den Gepflogenheiten in Deutschland (Lehmann-Wermser, 2019), wie im nächsten Abschnitt ausgeführt wird.

Quality: „Assessments developed for music education must adhere to internationally accepted norms for validity, reliability, and fairness, and focus on student learning.“ (Brophy, 2019c, S. 920)

Brophy spricht in diesem Prinzip nicht mehr von Assessment als Ganzem, sondern von einzelnen Assessments, also konkreten Beurteilungssituationen. Er ordnet unter Validität Maßnahmen ein, die sicherstellen, dass der Inhalt der beurteilten Materie vollständig abgebildet wird, die Beurteilung auf die Fähigkeiten der Lernenden zugeschnitten sind und die Kriterien eine gemeinsame Interpretation erlauben. Reliabilität wird dann erreicht, wenn die Resultate von Assessments in mehreren Durchläufen die gleichen Resultate liefern (S. 910 ff.). Ein großer Teil der englischen Assessment-Literatur beschäftigt sich mit *performance assessment* (z. B. Ciorba & Smith, 2009; Fautley, 2010; Pellegrino et al., 2015), der Entwicklung entsprechender *Rubrics* (Whitcomb, 1999; Schneider et al., 2019) und den Möglichkeiten des Einsatzes von Technik (Massoth, 2019). Für den Bereich Singen lohnt sich ein Blick auf das Assessment in amerikanischen Schulchören. Holcomb (2019) skizziert den Einsatz von *Rubrics* im Chor als gängige Methode der Beurteilung, verlangt aber deren Pilotierung und Justierung auf eine Normalverteilung (S. 384), um Validität und Reliabilität der Instrumente zu gewährleisten. Diese Forderung zeigt sich auch bei Brophy, der auf die entsprechenden Standards psychometrischer Tests verweist. Die *Rubrics* enthalten ausformulierte Qualitätsbeschreibungen in verschiedenen Stufen, ähnlich den Kompetenzrastern⁸⁷, die im deutschsprachigen Raum Verwendung finden. Er unterscheidet diese von *rating scales*, welche die Resultate in Zahlen ausdrücken. Entsprechende Beispiele werden im nächsten Abschnitt besprochen.

⁸⁷ In der vorliegenden Arbeit wird der Begriff *Rubric* aus zwei Gründen nicht mit Kompetenzraster übersetzt. *Rubrics* können in ihrer Form sehr differieren und müssen nicht zwingend wie Kompetenzraster ausformulierte Niveaubeschreibungen für erreichte Ziele haben (Stevens & Levi, 2013). Der zweite Grund liegt in der Abgrenzung zum Diskurs der Kompetenzorientierung und zu Kompetenzmodellen (z. B. Jordan, 2012), der an dieser Stelle nicht aufgegriffen werden soll.

Purpose: „Assessments in music education must have a clear purpose, identity what is being assessed, and define clearly how the results will be used to improve student music learning.“ (Brophy, 2019c, S. 921)

Brophy bezieht sich in diesem Prinzip unmittelbar auf die Idee von AFL nach Earl (2013). Allerdings räumt er auch ein, dass die Qualitätsanforderungen bezüglich Validität und Reliabilität, die unter *Quality* gefordert wurden, bei formativen Assessments möglicherweise nicht eingehalten werden können. AFL und damit verbunden auch AAL ist als Prozess gedacht, bei dem nicht das valide Resultat eines Assessments im Zentrum steht, sondern der Prozess selbst. An dieser Stelle wird das Spannungsverhältnis zwischen formativen und summativen Bewertungsmethoden bzw. der Zielkonflikt zwischen Förderung und Rechenschaftspflicht noch einmal deutlich. Holcomb (2019) beschreibt beispielsweise sogenannte „Model Cornerstone Assessments (MCAs)“ (S. 385), in denen eine Reihe von prozessorientierten Aufgaben zu lösen sind „with opportunities for revision through reflection and peer assessment“ (S. 385). Die Lernenden wählen ihr Repertoire selbst, erstellen Probepläne, üben und reflektieren ihren Lernprozess. Sobald sie bereit sind, nehmen sie ihre Performance auf und erstellen eine Selbstbeurteilung. Begleitend dazu beurteilt die Lehrkraft in den verschiedenen Phasen mit entsprechenden *Rubrics* den Fortschritt. Dieses Modell vereint die verschiedenen Aspekte von Assessment. Die Lernenden werden metakognitiv durch „critical thinking“ und „student accountability“ (S. 383) herausgefordert (AAL) und auf der Seite durch die Lehrkraft zuerst formativ (AFL) und abschließend summativ (AOL) beurteilt.

Operability: „Methods used in music education assessment must be clear, simple to understand, and easy to implement and use.“ (Brophy, 2019c, S. 922)

Dieses Prinzip beschreibt die Durchführbarkeit von Assessment. Einerseits bezieht sich Brophy auf die äußeren Umstände wie die Anzahl der Lektionen und die Vorgaben von Curricula, andererseits wird auch das individuelle Assessment in größeren Gruppen thematisiert. Er zitiert Holcomb (2019), der bei amerikanischen Schulchören teilweise individuelles Feedback beobachtet. Holcomb empfiehlt neben den bereits erwähnten *Rubrics* den Einsatz von Aufnahmegeräten als effizientes Hilfsmittel für Assessments (vgl. MCA weiter oben, S. 383 ff.).

Alignment: „Assessments must align with the curriculum for which they are developed and in which they are operationalized.“ (Brophy, 2019c, S. 923)

In diesem Prinzip wird die Passung zum Curriculum gefordert, wobei dies in Teilen als Querstand für die vorliegende Arbeit angesehen wird. Nicht nur, aber gerade im Kontext der Schweiz mit dem LP21 (D-EDK, 2015) finden interindividuelle Unterschiede und die Herausforderungen der adoleszenten Stimmentwicklung in Curricula wenig Beachtung. Im Sinne einer ganzheitlichen Förderung und Beurteilung wird deshalb in diesem Rahmen die Forderung aufgestellt, Inhalte auch über curriculare Forderungen hinaus ganzheitlich zu betrachten.

Authenticity: „Assessment in music education must be authentic and appropriate for the context in which it is administered.“ (Brophy, 2019c, S. 923)

Brophy fordert in diesem Prinzip authentische Beurteilungssituationen. Assessment soll möglichst nahe am Lerngegenstand stattfinden und auch zum Kontext des Ursprungs der Musik passen. In diesem Prinzip werden somit Fragen des Settings von Assessment sowie die inhaltliche Ebene angesprochen. Bezogen auf den Singunterricht wäre zu beachten, dass formative Vorgänge in ähnlichem Rahmen stattfinden, wie später summative Beurteilungen. Schaumberger (2020a) beschreibt z. B. das gemeinsame Auftreten im Chor als logische Beurteilungssituation nach dem Üben in Chorproben. Die Qualität des Auftritts wird vom Publikum oder von einer Fachjury gleichermaßen (AOL), wie auch vom Chor selbst beurteilt (AAL). In der Logik des Prinzips Authentizität wird Üben und Auftreten zu einem Ganzen und ist somit ein authentischer Vorgang. Zu hinterfragen wäre allerdings das bereits einige Male zitierte Vorsingen an Schweizer Schulen, wenn das Üben in der Klasse stattfindet und das Vorsingen in das stille Kämmerlein verlegt wird. Das Setting ändert sich, Publikum ist nur noch die Lehrkraft und die eigene Stimme steht unvermittelt im Zentrum. Dieser Wechsel geht einher mit einem Verlust von Authentizität. Auf der inhaltlichen Seite gilt es zu überlegen, inwiefern beim Klassengesang Pop-songs adäquat sind, wenn der individuelle Stimmklang zugunsten eines

Chorklangs aufgegeben wird oder ohne den Einsatz von Mikrofontechnik gar nicht erst reproduziert werden kann⁸⁸.

Social Justice: „Music curriculum and its associated assessment must be made available to and appropriately adapted for all students.“ (Brophy, 2019c, S. 924)

Brophy formuliert dieses Prinzip für Menschen mit besonderen Bedürfnissen, „both gifted and those with disabilities“ (S. 924). Auf der einen Seite nennt er die in vielen Staaten mittlerweile gesetzlich geregelten Garantien für einen gleichberechtigten Zugang zur Bildung, die im Rahmen von Inklusion zur Anwendung kommen. Andererseits erwähnt er Talentförderungsprogramme, in denen besondere Begabungen gefördert werden. Insgesamt wird mit diesem Prinzip gefordert, auf die individuellen Bedürfnisse aller einzugehen. Aus der Perspektive der adoleszenten Stimmentwicklung könnte auch die Forderung gestellt werden, einen gleichberechtigten Zugang zum Singen für mutierende Stimmen zu gewährleisten.

Valuing: „Assessment should be valued as a means to improve student music learning and guide instruction.“ (Brophy, 2019c, S. 925)

Das letzte Prinzip bezieht sich auf die Bedeutung, die Lehrkräfte dem Assessment des Musikkernens oder in diesem Fall der Entwicklung vokaler Kompetenzen beimessen. Brophy schreibt: „Quite simply, music educators teach what they value, and assess what they value most.“ (S. 925) Er schlägt deshalb vor, die eigenen Bewertungsinstrumente periodisch zu überprüfen. Mit dem autoethnografischen Exkurs hat der Forschende in Ansätzen die eigene Assessment-Praxis evaluiert. Konkrete Schritte, wie dies umgesetzt werden könnte, stellt Brophy allerdings nicht vor.

In den vorgestellten Prinzipien wird auf verschiedenen Ebenen die bereits ausgeführte Spannung zwischen formativen und summativen Beurteilungsprozessen sichtbar. Dieser Querstand kann auch im Fach Musik bzw. im Singunterricht nicht aufgelöst werden. Brophy ist sich dessen bewusst und skizziert einen möglichen Weg, der die ganzheitliche Seite des Assessments

⁸⁸ Als Mikrofontechnik ist an dieser Stelle nicht Effekt der elektronischen Verstärkung gemeint, sondern der spezifische Klang, der z. B. durch den Nahbesprechungseffekt ermöglicht wird, oder ein behauchter Klang, der nur mit Mikrofon trägt.

mitdenkt und die Interaktion zwischen Schülerinnen und Schülern sowie der Lehrkraft berücksichtigt:

„The incompatibility remains, however, between the unique, individualized, open-ended nature of musical responses and any attempt to treat them as a collective for inferential analysis. One way to resolve this is to develop a comprehensive and balanced set of assessments that provide teachers a holistic view of student learning over time: portfolios provide one such method to accomplish this.“ (S. 926)

Die folgende Vertiefung von spezifischen Assessments in der Musikpädagogik im englisch- und deutschsprachigen Raum zeigt Möglichkeiten für den Singunterricht auf, erhebt aber noch nicht den Anspruch an Ganzheitlichkeit oder Ausgeglichenheit.

Assessment of Singing

Im anglofonen Raum haben sich verschiedene Methoden zur Beschreibung stimmlicher Entwicklung etabliert, darunter verschiedene Beurteilungsskalen und *Rubrics*, sowie digitale und analoge Testbatterien. Innerhalb dieser Vielfalt von Perspektiven werden ausgewählte Methoden und Instrumente vorgestellt. Als Auswahlkriterium dient eine Fokussierung auf Elemente, die Potenzial für den Einsatz im Singunterricht aufweisen. Zu Beginn des Abschnitts wird der Zusammenstellung von Rutkowski (2019) gefolgt.

Rutkowski (1990, 2010) entwickelte die Bewertungsskala *The Singing Voice Development Measure* für Kinderstimmen (vgl. Tabelle 6), die nicht Intonation und Ambitus beurteilt, sondern einen Fokus auf den Einsatz der Register und Registerübergänge legt. „Until persons are comfortable using all their vocal registers, and transitioning among these registers, they will not sing in tune consistently.“ (2019, S. 630) Sie argumentiert, dass als erste Aufgabe das Instrument Stimme gelernt werden muss. Die Fokussierung auf die Evaluation des Ambitus, wie sie auch in einigen Studien im deutschsprachigen Raum eingesetzt wird (Gütay & Kreutz, 2010; Forge & Gembris, 2012), wird von ihr abgelehnt, weil „this assessment alone does not provide a complete picture of vocal development“ (2019, S. 634). In ihrem Test singen die Kinder eine kurze Moll-Melodie mit dem Ambitus von cis'-a' und werden in 5 Kategorien und 4 dazwischenliegende Stadien

eingeteilt. *Pre-Singer* und *Speaking Range Singer* können nur mit dem Brustregister (a–c') singen, *Limited Range Singer* und *Initial Range Singer* zusätzlich das Mittelregister (d'–a') bedienen und die *Singer* sind in der Lage, auch das Kopfregeister (>h') einzusetzen. Sie empfiehlt den Einsatz dieser Skala sowohl für die Forschung als auch für den Einsatz im Klassenzimmer. Die Bewertungsskala wurde zwar nicht für die Sekundarstufe 1 entwickelt, gibt aber für zwei Szenarien dieser Stufe Hinweise. Bei den Mädchen werden in der Phase 2A Registerbrüche und ein Verlust der Randstimme beobachtet (Gackle, 2019) und bei Jungen wird von

Level	Name of level	Description
1	Pre-Singer	does not sing but chants the song text.
1.5	Inconsistent Speaking Range Singer	sometimes chants, sometimes sustains tones and exhibits some sensitivity to pitch but remains in the speaking voice range, lower register (usually a to c').
2	Speaking Range Singer	sustains tones and exhibits some sensitivity to pitch but remains in the speaking voice range, lower register (usually a to c').
2.5	Inconsistent Limited Range Singer	wavers between lower and middle registers and uses a limited range when in middle register (usually up to f').
3	Limited Range Singer	exhibits consistent use of initial range (usually d' to f').
3.5	Inconsistent Initial Range Singer	sometimes only exhibits use of limited range, but other times exhibits use of initial range (usually d' to f').
4	Initial Range Singer	exhibits consistent use of initial range (usually d' to a').
4.5	Inconsistent Singer	sometimes only exhibits use of initial range, but other times exhibits use of extended singing range, upper register (sings beyond the lift to upper register: h' and above).
5	Singer	exhibits use of consistent extended range, upper register (sings beyond the lift to upper register: h'-flat and above).

TABELLE 6: *The Singing Voice Development Measure* (©2016 Joanne Rutkowski) (Rutkowski, 2019, S. 635); adaptiert mit deutschen Notennamen

Lehrkräften oft berichtet, dass nach der Mutation der Ambitus eingeschränkt bleibt. Beides sind kritische Momente der adoleszenten Stimm-entwicklung und lassen sich mit dem Vokabular der vorgestellten Skala beschreiben. In eine ähnliche Richtung geht die Logopädin Spital (2004) mit der Einteilung von Kindern in Sprechsänger, Tiefsänger, Mitsänger⁸⁹ und Sänger.

Rutkowski nennt als weiteres Konzept zur Beurteilung vokaler Kompetenzen die Skala von Wise und Sloboda (2008), die als Kriterien die Bewertung der Intonation, des tonalen Zentrums und der Kontur der Melodie beinhaltet. Die Skala startet auf Stufe 1 mit gesprochenen Tönen. Auf Stufe 2 werden einzelne Töne gesungen, aber zusammenhangslos. In Stufe 3–5 wird hauptsächlich die Stabilität des tonalen Zentrums begutachtet. Dieses kann in Stufe 3 nicht eruiert werden, wechselt in Stufe 4 mitten in der Melodie und wandert in der Stufe 5. Ab Stufe 6 bleibt das tonale

Level	Description
1	Singer sings with little variation in pitch, and may chant in speaking voice rather than singing.
2	Words are correct but there are contour errors. Pitches may sound almost random.
3	Singer accurately represents the contour of the melody but without consistent pitch accuracy or key stability.
4	Melody fairly accurate, or mostly accurate within individual phrases, but singer changes key abruptly, especially between phrases (e.g. adjusting higher-lying phrases down).
5	Melody largely accurate, but singer's key drifts or wanders. This may be the result of a mistuned interval, from which the singer then continues with more accurate intervals but without returning to the original pitch.
6	Key is maintained throughout and melody mostly accurately represented, but some errors (notes mistuned sufficiently to be 'wrong').
7	Key is maintained throughout, and melody accurately represented, but some mistunings (though not enough to alter the pitch-class of the note)
8	All melody is accurate and in tune, and key is maintained throughout.

TABELLE 7: *Rating Scale von Wise und Sloboda (2008)*

⁸⁹ Mitsänger bezeichnet eine Person, welche die Melodie dann trifft, wenn korrekt Singende danebenstehen.

Zentrum stabil, es gibt aber noch falsche Töne im Gegensatz zu Stufe 7, wo diese nur noch unrein intoniert werden. In Stufe 8 stimmt sowohl die Intonation als auch das tonale Zentrum.

Dieses Modell lässt sich in der Sekundarstufe 1 für die summativen Bewertungsszenarien einsetzen. Im Gegensatz zur Skala von Rutkowski wird kein Zusammenhang mit dem Gebrauch der Register hergestellt, es bleibt bei einer Beschreibung des aktuellen Standes und im Wesentlichen bei einer Analyse auf der Ebene Tonhöhe. Bei beiden Skalen fehlt die Evaluation von Rhythmus, Klangqualität und Atmung. Obwohl mit entsprechendem Fachwissen aus diesen Skalen der Handlungsbedarf abgeleitet werden kann, sind die Skalen von Rutkowski sowie Wise und Sloboda für die Selbstbeurteilung von Schülerinnen und Schülern nicht geeignet, weil der notwendige Entwicklungsbedarf sich hinter der Fachterminologie verbirgt.

Für eine ganzheitliche Beurteilung von Gesangsvorträgen setzt Holcomb (2019) auf *Rubrics*: „performance criteria can be combined and assessed within one holistic rubric, that allows the overall performance to be quickly assessed.“ (S. 384) Allerdings weist er darauf hin, dass die Reduktion auf wenige Bereiche und eine Begrenzung der Qualitätsstufen das Feedback zu unspezifisch werden lassen. In seinem Beispiel (vgl. Tabelle 8) wird dies deutlich. Zwar werden neben Melodiekontur und Intonation ebenfalls Themen wie Atmung und Klang abgedeckt, die Formulierung der vier Anspruchsniveaus von *aufstrebend (1)* bis *fortgeschritten (4)* ist aber nicht besonders präzise. Zudem wird der Rhythmus nur am Rand über die Phrasierung angesprochen. Leider belegt Holcomb an dieser Stelle den Prozess der Validierung des Rasters, die er selbst fordert, nicht (S. 384)⁹⁰. Einen Beleg für die Effektivität und Reliabilität entsprechender *Rubrics* liefern Ciorba und Smith (2009) in ihrer Untersuchung der Beurteilung von Soloauftritten von Sängerinnen und Sängern. Im Gegensatz zu den bereits beschriebenen analogen Instrumenten ist das nachfolgende SSAP digital konzipiert.

⁹⁰ An dieser Stelle wird nicht auf die Problematik der Formulierung von *Rubrics* eingegangen. So wird in Tabelle 8 deutlich, dass in tiefsten Niveaus nicht erreichte, sondern fehlende Kompetenzen formuliert werden. Stevens und Levi (2013) schreiben dazu: „yet we also try to avoid overly negative or competitive labels.“ (S. 39 f.)

Demorest et al. (2015) entwickelten das *Seattle Singing Accuracy Protocol (SSAP)* (Demorest & Pfordresher, 2015) als webbasiertes Tool für die Einstufung der Singfähigkeit. Dabei werden ausgehend von einer angenehmen Stimmlage zuerst Imitationsübungen gemacht, um anschließend ein Lied mit und ohne Text zu singen. An dritter Stelle wird getestet, wie das Tonunterscheidungsvermögen ist, um Defizite in der Tonhöhenverarbeitung zu eruieren. Zum Schluss werden demografische Daten erfasst und Fragen zum musikalischen Selbstkonzept beantwortet. Der Test wurde validiert (Pfordresher & Demorest, 2020) und kommt vor allem als For-

Level	fortgeschritten (4)	kompetent (3)	in Entwicklung (2)	aufstrebend (1)
Die Schülerin ... / Der Schüler ...				
Klang	singt stets mit weitem, resonantem Ton in allen Registern.	singt meist mit weitem, resonantem Ton.	hat Schwierigkeiten, einen resonanten Ton zu erzeugen; singt manchmal luftig, unkoordiniert und hat Probleme, sich in den Registern zurechtzufinden.	erzeugt keinen resonanten Ton, singt unkoordiniert und luftig oder ohne Raum.
Atem	singt durchgehend mit ausreichender Stütze und stilistischer Phrasierung.	singt unregelmäßig mit gestütztem Atem; atmet manchmal an unpassenden Stellen.	hat Schwierigkeiten, mit Stütze zu singen; hat etwas Schulter-, Nacken- oder Kiefferverspannung, atmet an unpassenden Stellen.	singt nicht mit Stütze; versteht kaum oder gar nicht, wann/wie korrekt geatmet wird.
Melodiekontur und Intonation	singt alle Töne korrekt und mit perfekter Intonation.	singt die meisten Töne korrekt; gelegentlich ein wenig unsauber.	singt einige Töne richtig; hat Schwierigkeiten, in der Tonart zu bleiben oder Sprünge zu bewältigen.	singt nur wenige Töne richtig und intoniert unsauber.

TABELLE 8: Rubric zu Gesangsvortrag nach Holcomb (2019, S. 385)⁹¹

⁹¹ Übersetzung durch den Autor.

schungsinstrument zum Einsatz. Der Einsatz eines Instruments dieser Art könnte insbesondere für das Selbst-Assessment (AAL) oder in Kombination mit einem entsprechenden Feedback der Lehrkraft (AFL) interessant sein. Entsprechende Überlegungen wurden durch den Forschenden bereits gemacht (Imthurn, 2020).

Cohen (2015) entwickelt seit 2008 die *AIRS Test Battery of Singing Skills (ATBSS)*, die eine Gruppe von bereits getesteten und validierten Aufgaben vereint, so z. B. die Messung des Ambitus, das Singen bekannter und unbekannter Lieder und das Kreieren eines Melodieabschlusses. „What is unique about the ATBSS is the examination of all of these together and with the aim of doing this systematically across the lifespan, across levels of musical training, and across cultures.“ (S. 241) Obwohl dieser Test für die Forschung konzipiert wurde, zeigen sich auch darin Möglichkeiten für einen didaktisch orientierten Einsatz. Cohen erweitert mit diesem Format die Ebene der validierten Tests und Skalen um kreative Aspekte. So wird unter anderem die Aufgabe gestellt, ein Lied zur Farbe Gelb oder zu einem Apfel zu singen, oder aus dem Stegreif zu erfinden. Zudem ist in der Idee, einen Test über die Lebensspanne öfters durchzuführen, der Portfolio-Gedanke, der später in dieser Arbeit ausgeführt wird (vgl. 14.3), angelegt.

Im anglofonen Raum werden mittlerweile auch digitale Assessment-Systeme für den Singunterricht angeboten. So betreibt die Firma MusicFirst⁹² eine interaktive Lernumgebung, die zu Gesangsaufnahmen Feedback zu Tonhöhe und Rhythmus gibt und ein Evaluationssystem für Musiktheorie und Gehörbildung ermöglicht: „Once a student submits a subject-matter assessment, a score is automatically generated and recorded in the Digital Gradebook. Performance proficiency assessments are scored manually, and then automatically recorded in the gradebook.“⁹³ Der Fokus des Werbetextes liegt insgesamt eher auf summativem Assessment. Allerdings bieten Systeme dieser Art einen wesentlichen Baustein für individuelle Feedbacks im Rahmen von AAL und AFL.

⁹² MusicFirst ist eine Division der Wise Music Group (<https://www.musicfirst.com/about-us/> – abgerufen am 5.11.2021).

⁹³ Webseite von MusicFirst zu Assessment (<https://www.musicfirst.com/assessment/> – abgerufen 05.11.2021).

5.3 Assessment des Singens in der deutschsprachigen Musikpädagogik

Zu Beginn der vorliegenden Arbeit wurde das Assessment des Singens allgemein und besonders für die Zeit der adoleszenten Stimmentwicklung problematisiert (vgl. 1.2). Dabei wurde auf verschiedene Initiativen zur Singförderung in der Grundschule verwiesen (z. B. Forge & Gembris, 2012). Obwohl die individuellen Stimmparameter im Rahmen dieser Forschungsprojekte als Indikatoren für deren Erfolg erfasst wurden, entwickelte sich daraus keine allgemeine Tendenz zu einer individuellen Förderung. Einzig Bojack-Weber (2012) und Greuel et al. (2007) sehen didaktisches Potenzial in ihren entwickelten Testinstrumenten. Zudem wurde mangelnde Diagnosefähigkeit für das Singen festgestellt (Brunner, 2014) sowie auf die Arbeit von Puffer und Hofmann (2017, 2021) verwiesen, in der die Entwicklung entsprechender fachdidaktischer Kompetenzen untersucht wird. Für die Sekundarstufe 1 lassen sich abgesehen von theoretischen Überlegungen (Dyllick, 2019) nach aktuellem Stand keine Studien oder Instrumente finden, welche eine Individualisierung anstreben. In diesem Abschnitt wird zuerst überlegt, wie der Begriff Assessment in der deutschsprachigen Musikpädagogik einzuordnen ist, um anschließend Hinweisen zu folgen, inwiefern Differenz und Heterogenität wahrgenommen werden und welche Auswirkungen sich diesbezüglich in der Unterrichtsgestaltung zeigen. Zuletzt werden einzelne Ansätze von Diagnose und Beurteilung untersucht. Beide Bereiche sind für ganzheitliches Assessment elementar.

Leistungsbeurteilung im Fach Musik und spezifisch im Singunterricht

Lehmann-Wermser (2019) wirft aus einer internationalen Perspektive einen Blick auf Assessment im Musikunterricht in Deutschland. Er stellt fest, dass nur wenige Publikationen zum Thema vorliegen. Zu nennen wäre der Band zu *LEISTUNG IM MUSIKUNTERRICHT* (K. Mohr & Schäfer-Lembeck, 2008) oder die Arbeit von Malmberg (2012) zu spezifischen Leistungskriterien in Projektarbeiten. Auch in der Literatur für angehende Lehrkräfte, z. B. in der *MUSIK DIDAKTIK* von W. Jank (2013), sieht Lehmann-Wermser im knappen Kapitel zu Evaluation und Leistungsbeurteilung (Niessen, 2013) nur wenig Hilfestellungen zum Thema. Er sieht Evidenz, dass Noten

im Musikunterricht in Deutschland auf kognitiven Inhalten wie Musiktheorie und Musikgeschichte sowie allgemeinem Interesse und Mitarbeit beruhen. Weiter betont er das Fehlen von Untersuchungen zur Durchführung von Bewertung und zu den damit verbundenen Überzeugungen⁹⁴:

„There have been no German publications on marking practices in music in the 15 years prior to the time this chapter was written. Subsequently, we do not know the distribution of grades or the different areas, such as music appreciation, singing, or composing, that teachers cover in music lessons. There are no studies on teachers’ knowledge about assessment techniques or related assumptions and beliefs.“ (S. 240)

Lehmann-Wermser nimmt mit *marking* in erster Linie auf die summative Beurteilung Bezug. Er zeigt indessen, dass in Zeugnissen neben den Schulnoten bereits Bewertungen in Form von Kompetenzrastern in Gebrauch sind (S. 241), die auf die erreichten Stärken fokussieren und Kommentare ermöglichen: „Especially the ‚comments‘ that conclude these report cards are used by teachers to give individual feedback and make recommendations for future musical engagement.“ (S. 42) Die mit wenig Support von Forschung entwickelten Kompetenzraster werden aber kritisiert, weil die Beschreibungen der Level mehrdimensional und uneinheitlich sind und weil sie Kompetenz, Motivation und Persönlichkeitsmerkmale vermischen. Diese Kritik deckt sich mit den bereits zitierten Aussagen von Holcomb (2019), der darauf hinweist, dass holistische Assessments dieser Art zu wenig spezifisch und deshalb nicht effektiv seien. Als weiterer Punkt müsste überlegt werden, ob formatives Feedback am Ende des Semesters zielführend ist. Fautley (2010) gewinnt dieser Praktik, er nennt es „formative use of summative assessment“ (S. 198), trotz verdrehter Reihenfolge Positives ab, wie übrigens auch Lehmann-Wermser. Beide schätzen den Informationsgehalt deutlich positiver ein als bei Schulnoten. Dem ist sicher beizupflichten, im vorgestellten Modell nach Earl (2013) entwickelt individuelles Feedback sein Potenzial aber im Vorlauf (AFL) von summativem Assessment (AOL, vgl. 5.1). Beurteilungsbögen dieser Art

⁹⁴ Das Fehlen von Untersuchungen kann auch für die Schweiz bestätigt werden. Die Studie von Lupatsch und Hadjahr (2011) zu geschlechtsspezifischen Aspekten der Notengebung wurde an anderer Stelle referiert (vgl. 4.5), hat aber nicht den Prozess der Notengebung, sondern die Resultate an sich untersucht.

müssten somit als Standortbestimmung eingesetzt werden, um individuellen Förderbedarf zu ermitteln. Der abschließende Einsatz als summatives Bewertungsinstrument wäre dann gerechtfertigt.

Im Text führt Lehmann-Wermser anschließend zwei Studien an (J. Hasselhorn & Lehmann, 2015; Jordan, 2012), die sich an psychometrischen Teststandards orientieren, aber in der musikpädagogischen Forschungsgemeinschaft zu harschen Reaktionen führte: „Even after some years now the idea of modeling and assessing competencies has not been broadly accepted.“ (Lehmann-Wermser, 2019, S. 244) Die schlechte Akzeptanz führt Lehmann-Wermser auf die Überzeugungen von Musikpädagoginnen und -pädagogen zum Fach Musik zurück, „partly because the unique nature of music is basically unstandardized and ambiguous“ (S. 238).

Stern (2020) diskutiert die Leistungsbewertung im Musik- und Kunstunterricht in Österreich und verweist dabei zuerst auf die bereits zitierten Autoren Black und Wiliam (1998) sowie Hattie (2009), um die lernförderliche Wirkung von individuellem Feedback zu betonen. Im nicht zu lösenden Dilemma zwischen formativer und summativer Beurteilung schlägt er vor, die summative Beurteilung zwecks Selektion auf die Schnittstellen zu beschränken. Für die lernförderliche Beurteilung empfiehlt er fünf Qualitätskriterien: formative Lerndiagnosen und Feedback, vielfältige Beurteilungsinstrumente zugunsten eines Gesamtbildes, Orientierung an Kompetenzen⁹⁵, Anerkennung und Ermutigung unter Berücksichtigung individueller Voraussetzungen sowie den Einbezug von Selbstbeurteilung. Damit wird hauptsächlich der aktuelle Diskurs rezipiert, insbesondere die Ebene von AFL und AAL. Gerade für die künstlerischen Fächer bestätigt sich die in dieser Arbeit vorgelegte Perspektive auf ganzheitliche Förderung und Beurteilung.

Spezifisch mit der Leistungsbeurteilung des Singens setzt sich Schaumberger (2020a) auseinander. Zu Beginn des Textes skizziert er Singen in erster Linie als Auftritt in der Gruppe, der entsprechend kollektiv beurteilt wird. Im Gegensatz zu diesen motivierenden Gemeinschaftserlebnissen beschreibt er losgelöste und lediglich der summativen Leistungsbeurteilung geschuldete Singanlässe als problematisch: „Ein (Vor)

⁹⁵ Stern nimmt im Text Bezug auf die Diskussion zu Kompetenzorientierung, erwähnt aber entsprechend validierte Modelle für die Musikpädagogik nicht.

Singen kann aber auch zu einer negativen, in bestimmten Fällen auch traumatischen Erfahrung geraten, wenn es [...] als reine Leistungsmessung ohne förderliche Rückmeldung und möglicherweise sogar als Druckmittel eingesetzt wird.“ (S. 223) Schaumberger beobachtet aber eine zunehmend positive Einstellung zur Leistungsmessung durch die Einführung von Chorklassen und bezieht sich auf die beiden Lehrmittel von Bolender und Müller (2012a) sowie Schnitzer (2008b), die durch eine stetige individuelle Leistungsbeurteilung negative Rückmeldungen per se vermeiden wollen. Im weiteren Verlauf bezieht er sich analog zu der vorliegenden Arbeit auf einige anglophone Quellen und referiert den Begriff Assessment.

Auch Schaumberger schließt mit einigen Empfehlungen ab: Auswertung bestehender Kompetenzmodelle sowie der Forschungsergebnisse zur Messung von Fertigkeiten, Einbezug der Erkenntnisse von Lehr- und Lernforschung bezüglich Leistungsmessung, Entwicklung von Instrumenten zur Leistungsmessung, Einbezug „der motivationalen, volitionalen und sozialen Kompetenzen“ (Schaumberger, 2020, S. 244), Abstimmung der Leistungsbeurteilung auf die Schulstufen sowie eine „nachhaltige Verankerung der Ergebnisse in den Curricula für Pädagoginnen- und Pädagogenbildung sowie in der Fort- und Weiterbildung“ (S. 244). Insbesondere der Bezug auf die Herausforderungen der unterschiedlichen Schulstufen und die Nennung von Personal- und Sozialkompetenzen nach Weinert (2001) zeigen, wie viel Gewicht er einer ganzheitlichen Betrachtungsweise bei der Leistungsbeurteilung beimisst.

Die Autoren Stern (2020) und Schaumberger (2020a) betonen im Gegensatz zu Lehmann-Wermser (2019) die Möglichkeiten und die Wichtigkeit von formativen Beurteilungsansätzen. Die vorliegende Arbeit nimmt dieses Anliegen auf und vertieft die genannten Ansätze durch die Darstellung des Assessment-Konzepts nach Earl (2013) und der Assessment-Prinzipien nach Brophy (2019c). Darüber hinaus wird mit der Interviewstudie ein Blick auf die aktuelle Beurteilungspraxis geworfen.

Wahrnehmung und Umgang mit Differenz

In Kapitel 4 wurde deutlich, dass Singen ein anthropologisches Merkmal ist, die Entwicklung aber interindividuell verschieden ist und neben der stimmlichen Veränderung immer auch die Persönlichkeit betrifft. In der Logik von formativem Assessment wird im Unterricht auf diese Unterschiede

Bezug genommen, einerseits durch Maßnahmen, diese Unterschiede zu eruieren und den Unterricht entsprechend zu adaptieren (AFL, diagnostisches Assessment), und andererseits durch Hilfestellungen, wie Schülerinnen und Schüler ihre interindividuellen Lernprozesse reflektieren können (AAL). Dies bedingt von Musiklehrkräften in erster Linie diagnostische Fähigkeiten und Methoden, um auf Heterogenität zu reagieren.

Linn (2017) beschäftigt sich in seiner Dissertation mit Überzeugungen von Lehrkräften zu Heterogenität. Die Untersuchung legt nahe, dass Heterogenität in verschiedenen Bereichen wahrgenommen wird, aber nur marginal zur Adaption des Unterrichts führt: „Eine diagnostische Begleitung der individuellen Entwicklungen oder eine gezielte Beobachtung von personenbezogenen Stärken oder Kompetenzen kommen nur sehr selten vor.“ (S. 173) Dies deckt sich mit den Aussagen Brunners (2014), der für den Bereich Singen kaum diagnostische Fähigkeiten gefunden hat. Obwohl die Untersuchung von Linn sich nicht auf das Singen fokussiert und im Musikunterricht der Unterstufe durchgeführt wurde, liefert sie Einsichten, wieso Lehrkräfte kaum formatives Assessment einsetzen. Linn eruiert drei Bereiche, in denen Lehrkräfte nur unzureichend auf heterogene Klassensituationen eingehen können. Erstens wird keine unterrichtsbegleitende Diagnose gemacht, die „möglichst viele individuelle musikalische Voraussetzungen und Kompetenzen der Schülerinnen und Schüler erfasst“ (Linn, 2017, S. 231), zweitens fehlt „regelmäßige und flexible Binnendifferenzierung“ (S. 231) und drittens vermissen die Lehrkräfte „individualisierte Lernangebote, Materialien oder Aufgaben“ (S. 231). Er ergänzt, dass die kleine Stundendotation für das Fach, fehlendes Methoden-Knowhow und unzureichende Unterrichtsmaterialien die Lehrkräfte davon abhält, individuelle Förderung im Musikunterricht umzusetzen.

In Linns Studie äußert sich lediglich eine Lehrerin zum Singen. Sie nimmt zwar große Differenzen bezüglich der gesanglichen Fähigkeiten der Schülerinnen und Schüler wahr, betont aber, „dass die individuell verschiedenen Kompetenzen und Voraussetzungen der Schülerinnen und Schüler beim gemeinsamen Singen zugunsten des Gemeinschaftsgefühls in den Hintergrund treten“ (S. 170). Obwohl die positiven Effekte gemeinsamen Singens bekannt sind (Gembris, 2015), widersprechen die referierten Empfehlungen Rutkowskis (2019) dieser Lehrkraft. Sie empfiehlt, die Stimmentwicklung innerhalb des Klassenunterrichts individuell zu fördern, um jedem Kind die Möglichkeit zu geben, das ganze Instrument Stimme, also

Kopf-, Mittel- und Bruststimme sowie die Registerübergänge, benutzen zu können. Die genannte Lehrerin reduziert das Singen auf den Gemeinschaftsaspekt, verzichtet dabei aber auf die Förderung von Kindern mit stimmlichen Defiziten. Abschließend stellt Linn somit fest, dass Differenz zwar wahrgenommen wird, dies im Musikunterricht aber nicht zu einer Differenzierung oder Individualisierung führt.

Die Wahrnehmung von Heterogenität kann auch dazu führen, dass Singen im Musikunterricht zu einer Option wird. So beschreibt Cslovjecsek (2009) mit dem Widerspruch *alle müssen wollen* sowohl ein grundsätzliches Problem der Schule als auch die Tatsache, dass nicht alle motiviert sind zu singen. Er argumentiert zuerst, dass die Freiheit der Kinder nur darin besteht, „mehr oder weniger aktiv, motiviert und interessiert oder aber geistig abwesend zu sein“ (S. 82 f.), was echte Partizipation gerade im Musikunterricht infrage stellt. Am Beispiel einer Unterrichtssequenz zeigt er auf, wie diese Spannung im Bereich Singen aufgelöst werden kann. Er lässt Schülerinnen und Schülern die Wahl zu singen oder auf das Singen zu verzichten. Die *Unwilligen* werden gebeten, die Sequenz zu beobachten und zu protokollieren. Trotz der Wahlmöglichkeit bleibt damit das grundsätzliche Problem der eingeschränkten persönlichen Freiheit bestehen, da auch die Aufgabe des Protokollierens nicht zwingend im Sinne der *Unwilligen* sein muss. Cslovjecsek vertieft aber nicht diesen Widerspruch, sondern stellt mit Bezug auf das Modell der multiplen Intelligenzen nach Gardner (1983) in der Singsequenz vielfältige disziplinäre Bezüge her. So haben Schülerinnen und Schüler beispielsweise die Möglichkeit, sich auf den Puls in der Musik zu konzentrieren, der mit der logisch-mathematischen Intelligenz verbunden ist, oder die körperlich-kinästhetische Intelligenz in der Bewegung zu genießen. Die von Cslovjecsek genannten Bezüge zeigen das vielfältige Erlebnispotenzial der gemeinsamen Singaktivität, schaffen aber mit dem möglichen Verzicht auf das Singen neue Probleme. Wie in den meisten Curricula wird auch im LP21 (D-EDK, 2015) der Erwerb vokaler Kompetenzen angestrebt. Obwohl die persönliche Wahlfreiheit durch dieses Bildungsziel beschnitten wird, müsste im Sinne des Lehrplans in einer Singsequenz angestrebt werden, dass alle Beteiligten mitmachen. Zudem stellt sich die Frage, ob der von Cslovjecsek gemachte Bezug auf das Modell von Gardner genügend tragfähig ist, da dieses zwar in pädagogischen Kreisen gerne zitiert wird, aber

als Konstrukt den Standards der Psychologie nicht genügt (Rost, 2008)⁹⁶. Die Frage der Freiwilligkeit beim Singen spitzt Eckhardt (2019) zusätzlich zu, indem er die Autonomie der Sängerinnen und Sänger im Chor infrage stellt. Trotz freiwilliger Teilnahme unterstellen sich Sängerinnen und Sänger in Chören dem Diktat der Leitenden. Eckhardt vergisst aber, dass neben dem Grundbedürfnis Autonomie gerade Verbundenheit und das Erleben von Kompetenz (Ryan & Deci, 2000) im Chor oder dem gemeinsamen Singen gestärkt werden können.

Linn (2017) konnte zeigen, dass Lehrkräfte heterogene Fähigkeiten von Schülerinnen und Schülern wahrnehmen, aber kaum darauf reagieren. Das von Cslovjcek (2009) eingebrachte Modell muss aufgrund des curricularen Anspruchs zwar infrage gestellt werden, weist aber auf unterschiedliche Zugänge der Lernenden zum Singen hin. Einen weiteren Schritt unternehmen Rieche et al. (2018), indem sie in einem Sample mit Lehramtsstudierenden die Fähigkeit, dysfunktionale Überzeugungen bei Schülerinnen und Schülern im Musikunterricht zu erkennen, untersuchen. Für das Singen wäre beispielsweise die verbreitete Überzeugung, nicht singen zu können (vgl. non-singer; Ashley, 2015), zu nennen. Sie stellen fest, dass nur die Hälfte der angehenden Lehrkräfte dysfunktionale Überzeugungen als Problem erkannte und ein Drittel gar keine Probleme wahrnahm. Zudem wurden meistens nur die Auswirkungen erkannt:

⁹⁶ Rost (2008) zeigt auf, dass das Modell der multiplen Intelligenzen zwar im „pädagogisch-psychologischen Feld auf breite Akzeptanz“ (S. 97) stößt, aber „keine Evaluationen der von Gardner propagierten Umsetzung der MI-Theorie in die pädagogische Praxis vorliegen, die den üblichen methodischen Standards der Psychologie genügen“ (S. 98). So wird die pädagogisch-psychologische Sichtweise z. B. auch von Spychiger (1997) für die musikalische Intelligenz entlang der von Gardner aufgestellten kategorialen Kriterien für Intelligenz (Rost, 2008, S. 99) ausgebreitet, ohne Belege der Messbarkeit zu liefern. Sie verweist lediglich auf die von Gardner selbst aufgestellte Devise, mit der neuen Betrachtungsweise von Intelligenz auch die Frage der Messbarkeit zu überwinden, indem statt gemessen, beobachtet wird (Spychiger, 1997, S. 75). Die Frage der Messbarkeit wird von Largo, bekannt durch die Longitudinalstudien zur kindlichen Entwicklung (Largo, 2010, 2017a), überwunden, indem er sich zwar auf Teile des Modells von Gardner bezieht (Largo, 2017b), aber das Wort Intelligenz durch das Wort Kompetenz austauscht.

„Häufig wurden jedoch motivationale und emotionale Probleme erkannt, welche als Folgen der Überzeugung angesehen werden können. Dies deutet darauf hin, dass angehende Lehrkräfte die Auswirkungen ungünstiger Überzeugungen eher wahrnehmen als die Überzeugungen selbst.“ (S. 17)

Im Vergleich zur Wahrnehmung von Heterogenität von Fähigkeiten, wie sie sich bei Linn (2017) vornehmlich zeigen, wird bei Rieche et al. sichtbar, dass die Wahrnehmung von Fähigkeitsüberzeugungen noch weniger ausgeprägt ist. Hoene und Thurman (2011) belegen exemplarisch am Beispiel von „Antonia“, die beim Singen mit der Nachbarin tuschelt und nicht mitsingt, dass die Fähigkeitsüberzeugungen vielfältig sein können und zusätzlich mit äußeren Umständen, welche über Selbstkonzepte hinausgehen, zu rechnen ist. So verweisen sie bezüglich Ersterem auf die Möglichkeit, dass „Antonia“ ihre Stimme nicht mag oder Angst hat, die Töne nicht zu treffen. Als externe Faktoren werden z. B. ein langweiliges Repertoire oder die Anforderung, ein zu wenig geübtes Lied auswendig zu singen, genannt (S. 34). Das Beispiel zeigt, wie komplex gerade beim Singen der Umgang mit individuellen Bedürfnissen von Lernenden ist. Lehrkräfte sind im Singunterricht zudem nicht nur mit „Antonia“, sondern einer ganzen Reihe von Kindern oder Jugendlichen herausgefordert. Dies führt zuweilen zu Überforderung und zu problematischen Reaktionen.

Tellisch (2015) untersucht in ihrer Arbeit die Qualität von Interaktionen im Musikunterricht, indem sie Menschenrechte als „normativen Rahmen für die Gestaltung der pädagogischen Prozesse“ (S. 10) benutzt. Ein Drittel ihrer untersuchten Videovignetten (von insgesamt 213) zeigen Situationen aus dem Bereich Singen, wobei sie knapp 50 % als positiv und anerkennend, ca. 20 % neutral und etwas mehr als 20 % verletzend einschätzt. Tellisch beschreibt die Videovignetten aus verschiedenen Perspektiven und kommt in ihrer Arbeit zu dem Schluss, die Vorbildwirkung und das Fachwissen der Lehrkraft seien Voraussetzungen zu einem „human-anerkennenden Musikunterricht“ (S. 252):

„Versteht sich der Musiklehrer beim Singen im Musikunterricht im Sinne eines Mit-Gestalters von Unterricht mit den Tätigkeiten von mitmachen, ermuntern, anregen, ermutigen und beraten (vgl. Nimczik 1991:88) und versucht er, sich in seine Schüler hineinzuversetzen, sind wichtige Grundvoraussetzungen für eine anerkennende

Beziehungsgestaltung zwischen Lehrer und Schüler in diesem Bereich des Musikunterrichts gelegt.“ (S 252 f.)

In der zusammenfassenden Beschreibung werden wesentliche Aspekte formativen Assessments genannt. Insbesondere die Beziehungsgestaltung wurde bis dato nur im autoethnografischen Exkurs als wichtiges Element beschrieben (vgl. 3.3). Im Hinblick auf formative Rückmeldung weist Tellisch auf eine bisher unberücksichtigte Problematik hin. In der Analyse der Videovignetten zeigt sich in allen fünf beschriebenen Szenen (S. 215 ff.) Widersprüchliches, insbesondere wenn Kinder oder Gruppen eine individuelle Behandlung erhalten. Sie zeigt am Beispiel eines Jungen, wie das Lob für einen richtigen Ton das Potenzial zur Diskriminierung der anderen Kinder beinhaltet. Diese anderen Kinder erhalten kein Lob, weil ihre richtigen Töne als selbstverständlich erachtet werden. Dieses Beispiel erhält mehr Brisanz, wenn das Talent in der Klasse für die naturgegebene schöne Stimme (vgl. 4.1) gelobt wird und der Rest nicht. Ähnlich wie beim Thema Gender, wo eine Ambivalenz zwischen *doing* und *undoing gender* ausgemacht werden konnte, könnte an diesem Beispiel von Tellisch *doing* und *undoing diversity* diskutiert werden. Die Arbeit von Tellisch zeigt auf, dass der Umgang mit individuellen Bedürfnissen anspruchsvoll ist und insbesondere formative Prozesse in AFL möglicherweise als Ungleichbehandlung und deshalb diskriminierend verstanden werden können.

Als Fazit zeigt sich einerseits ein Bewusstsein für Heterogenität und interindividuelle Unterschiede, aber andererseits auch ein Unvermögen, darauf einzugehen. Zusätzlich muss in Betracht gezogen werden, dass das Eingehen auf Differenzen potenzielle Unterschiede betont und damit neue Ungerechtigkeiten etabliert werden, insbesondere wenn der Unterricht als frontale Situation organisiert ist. Im Rahmen dieser Überlegungen könnte in einem erweiterten Sinn auch Diversität oder Inklusion im Singunterricht thematisiert werden (vgl. B. Jank & Bossen, 2017; Herzog, 2018), was aber den Rahmen dieser Arbeit sprengen würde.

Diagnose- und Beurteilungsinstrumente für das Singen

Im ersten Teil des Abschnitts wurde die Wahrnehmung von Differenz und die damit verbundenen Überzeugungen diskutiert. Im zweiten Teil geht es nun darum, bereits entwickelte Methoden für die Diagnose und Beurteilung im Rahmen von AFL und für die objektive Bewertung in AOL zu

diskutieren. Im vorhergehenden Abschnitt wurden Möglichkeiten des Assessments bezüglich Registerbenutzung, akkurater Intonation und Melodiekontur sowie der Einsatz eines *Rubric* beschrieben. In diesem Abschnitt werden nun zwei Instrumente aus dem deutschsprachigen Raum beleuchtet. Als Einstieg wird aber die Sichtweise eines Stimmarztes zur Betreuung von Kinder- und Jugendstimme dargestellt.

Fuchs (2015) betont in seinem Artikel *DIE HEISERE KINDER- UND JUGENDSTIMME* zwar die Notwendigkeit, Diagnosen durch medizinisches Fachpersonal zu stellen, empfiehlt aber „in der verantwortungsvollen Arbeit mit Kinder- und Jugendstimmen“ (S. 9) Musiklehrkräften, sich mit dem Phänomen von Stimmstörungen zu beschäftigen. Fuchs definiert Heiserkeit als „Leitsymptom einer Sprech- und Singstimmstörung“ (S. 6), aber nicht als Krankheit. Insbesondere den regelmäßigen Kontakt zu den Schülerinnen und Schülern sieht Fuchs als Chance, mögliche Stimmveränderungen zu beobachten, wobei analytisches Hören und entsprechendes Training vorausgesetzt werden (S. 9). Fuchs setzt in seinen Ausführungen diagnostische⁹⁷ Fähigkeiten voraus, geht aber nicht darauf ein, dass Lehrkräfte in der Regel mit vielen Kindern gleichzeitig arbeiten müssen. Trotzdem erwartet er, dass Auffälligkeiten wahrgenommen und bearbeitet werden. Mit Blick auf die adoleszente Stimmentwicklung bedeutet dies auf der einen Seite die Notwendigkeit eines Aufbaus von CK und PCK im Umgang mit Stimmveränderung und möglichen Stimmschwierigkeiten, aber auf der anderen Seite auch das Schaffen von Raum während des Unterrichts für individuelle diagnostische Sequenzen im Unterricht.

Greuel et al. (2007) haben in einem älteren Projekt ein entsprechendes diagnostisches Instrument für die Einschätzung der Singstimme eintretender Schülerinnen und Schüler entwickelt. Ziel ist das Erfassen von drei Parametern: Qualität der Randstimmfunktion, Ambitus und Liedraum⁹⁸. Die Autoren nehmen auch Bezug auf möglich Ängste der Kinder und erwarten eine feinfühlige Herangehensweise der Lehrkräfte. Die „musikpädagogische Stimmdiagnose“ (S. 93) hat nicht die Funktion eines Tests,

⁹⁷ Diagnose wird hier im pädagogischen Sinn verwendet, mit dem Ziel, Schülerinnen und Schüler ganzheitlich zu betreuen und zu fördern.

⁹⁸ Die Autoren nehmen keinen Bezug auf den Begriff Tessitura, der bei Gackle (2019) und Welch (2019) für den gut singbaren Bereich gebraucht wird und dasselbe bedeutet.

„sondern eines zielgerichteten, fachlich fundierten, methodisch geregelten und reflexiv kontrollierten Kommunikations- und Verstehensprozesses“ (S. 93). Damit bewegen sich Greuel et al. eindeutig im Bereich AFL. Die Resultate und weitere Bemerkungen zu Atmung, Stimmklang, Artikulation und Resonanz werden jeweils auf einem A4-Blatt pro Klasse notiert. Das Verfahren gleicht mit der Beachtung der Randstimmfunktion der Beurteilungsskala von Rutkowski (1990)⁹⁹. Als Resultat der flächendeckenden Stimmbeurteilung schlagen Greuel und Horst (2011) die anschließende Einteilung der Kinder in verschiedene Anspruchsgruppen vor, um spezifische Bedürfnisse in der Stimmbildung aufgreifen zu können. Durch eine punktuelle Neubeurteilung soll es so möglich sein, in die nächste Gruppe zu wechseln. Im Praxistest ist es mit dieser Methode gelungen, die vokalen Kompetenzen der Hälfte der Kinder in den unteren Gruppen zu verbessern.

J. Hasselhorn (2015) beschäftigt sich mit der Messbarkeit musikpraktischer Fähigkeiten mittels Skalen ohne Anspruch an einen Einsatz für formatives Assessment. Er validiert eine 6-stufige Skala zur Messung gesangspraktischer Fertigkeiten (vgl. Tabelle 9) und bezieht sich dabei auf zwei amerikanische Skalen (Newlin, 2004; Hornbach & Taggart, 2005), die Intonation, melodische Kontur und das tonale Zentrum beachten. J. Hasselhorn zeigt, dass eine Bewertung mit dieser Skala gute Reliabilitätswerte erreicht. In der Studie untersuchte er allerdings nicht nur den

Wert	Beschreibung
1	Das Kind singt das Lied nahezu oder völlig korrekt.
2	Das Kind singt das Lied mit einiger Genauigkeit und beginnt in der vorgegebenen Tonart.
3	Das Kind singt das Lied mit einiger Genauigkeit, beginnt aber in einer anderen als der vorgegebenen Tonart oder moduliert innerhalb des Liedes.
4	Das Kind singt den melodischen Umriss mit deutlich falschen Tonhöhen.
5	Das Kind singt das Lied mit einem anderen melodischen Umriss als das Lied besitzt.
6	Eine sinnvolle Bewertung ist nicht möglich (kein Signal oder Kind nicht ernsthaft mit dem Lösen der Aufgabe beschäftigt).

TABELLE 9: Skala J. Hasselhorn (2015, S. 82)

⁹⁹ Ob dies bewusst so gemacht wurde, geht aus dem Text nicht hervor und erschließt sich auch nicht aus einem nachfolgenden Text (Greuel & Horst, 2011).

Bereich Singen, sondern auch die Bereiche Musizieren und Rhythmusproduktion, um insgesamt einen Test zu entwickeln, der in der Lage ist, musikpraktische Fähigkeiten in verschiedenen Stufen unter Einhaltung klassischer Testszenarien zu messen. J. Hasselhorn empfiehlt zwar den Einsatz der Skala zum Singen für den Unterricht, um summativ zu beurteilen, ist sich aber bewusst, dass z. B. Stimmtechnik damit nicht abgebildet wird (S. 96).

In diesem Abschnitt wurden Ansätze zu formativem und summativem Assessment im deutschsprachigen Raum zusammengetragen. Das vorgestellte Diagnoseinstrument (Greuel, 2006; Greuel & Horst, 2011) kann als Instrument von AFL eingeordnet werden, konzentriert sich allerdings auf wenige stimmliche Parameter. Betreffend valider Beurteilung von Stimmkompetenzen (AOL) konnte lediglich die Arbeit von J. Hasselhorn (2015) hinzugezogen werden. Ein Kompetenzmodell zum Singen, vergleichbar dem Kompetenzmodell zum Bereich Hören und Kontextualisieren (Jordan, 2012), das mehr als nur die Intonation und das Singen von melodischen Konturen berücksichtigt, gibt es bis dato nicht.

In all diesen Arbeiten werden motivationale und volitionale Aspekte wenig berücksichtigt, ebenso wenig wie die psychologischen Aspekte der Stimmentwicklung. Ob sich die Fragebogen zur Erfassung von Selbstwirksamkeit (Carmichael, 2014; Carmichael & Harnischmacher, 2015) oder Motivation (Harnischmacher & Hörtzsch, 2012) zu diagnostischen Zwecken bzw. zur Grundlage von Reflexionen anpassen lassen und in einem didaktischen Setting verwendet werden könnten, müsste untersucht werden.

In diesem Kapitel wurde das Assessment-Modell von Earl (2013) als möglicher Weg zu ganzheitlichen Förder- und Beurteilungspraktiken lerntheoretisch verankert und mit Blick auf mögliche Kategorien für die QIA beschrieben. Aus dem anglofonen Raum wurde auf die auf Musikpädagogik zugeschnittenen Assessment-Prinzipien nach Brophy (2019c) rezipiert und einige Instrumente der Stimmbeurteilung vorgestellt. Dabei ist aufgefallen, dass weniger die theoretische Verortung als vielmehr die Schaffung von praxistauglichen Instrumenten im Zentrum steht. Im letzten Abschnitt zeigte sich, dass in der deutschsprachigen Musikpädagogik die Messung von musikpraktischen Fähigkeiten zwar untersucht wird, diese Testszenarien aber kontrovers aufgenommen werden. Mit Blick auf *accountability* und *high-stake testing* kann dies nachvollzogen werden.

Die Empfehlungen von Stern (2020) und Schaumberger (2020a) zur Leistungsbeurteilung bestätigen hingegen das Forschungsdesiderat, mit dem sich die vorliegende Arbeit beschäftigt. Bevor die Forschungsfragen formuliert werden, werden in einem Exkurs zum schweizerischen Schulsystem dessen Eigenheiten erläutert.

6 Exkurs: Rahmenbedingungen des Schweizer Bildungssystems

Der inhaltliche Exkurs beschäftigt sich mit den Eigenheiten des Schweizer Bildungssystems und den regionalen Besonderheiten der Sekundarstufe 1, wobei ein spezifischer Blick dem Fach Musik und der Ausbildung von Musiklehrkräften gilt. Die Notwendigkeit dieses Exkurses begründet sich in erster Linie durch die Differenz zu einigen Bundesländern sowie zu Österreich, die den Wechsel in die Sekundarstufe 1 bereits nach der 4. Schulstufe vollziehen und sich auch in Bezug auf die Anstellung von Lehrkräften gänzlich unterscheiden¹⁰⁰. Zusätzlich zeigen sich auch innerhalb der Schweiz in den 26 Kantonen und den vier Sprachregionen Besonderheiten. Die differierenden Rahmenbedingungen beruhen in erster Linie auf dem föderalistischen Bildungssystem der Schweiz, das die Hoheit über den obligatorischen Bereich der Schule den Kantonen überlässt und trotz Bemühungen zur Vereinheitlichung der Curricula und einer Harmonisierung der Schulstrukturen große Unterschiede aufweist.

Im ersten Abschnitt werden die Bemühungen zur Vereinheitlichung skizziert und es werden die Auswirkungen auf die Anstellungsbedingungen durch die Verschiebung der Anstellungshoheit von den Kantonen hin zu den lokalen Schulleitungen thematisiert. Im zweiten Abschnitt wird die Ausbildungssituation der Musiklehrkräfte umrissen und die Unterschiede von Mono- und Mehrfachlehrkräften problematisiert. Ausgehend davon werden mögliche Berufsbiografien aufgelistet, um das heterogene Feld der Musiklehrkräfte deutlich zu machen. Im letzten Abschnitt werden die Spezifika der kantonalen Ausgestaltung der Sekundarstufen 1 beschrieben, wobei sich zeigt, dass Unterschiede in der Menge des Musikunterrichts und in der Verteilung der Geschlechter zu finden sind.

¹⁰⁰ Dieses Kapitel ist vor allem an die Leserschaft gerichtet, die mit dem schweizerischen Schulsystem nicht vertraut ist.

6.1 Kantonale Ausprägungen als Resultat des Föderalismus

Das Bildungssystem der Schweiz ist geprägt durch den Föderalismus. Die Hoheit über die Bildung liegt traditionell bei den Kantonen (Heinzer & Hangartner, 2016, S. 32). Im 19. und 20. Jahrhundert haben sich so 26 verschiedene Schulsysteme etabliert. Im Jahr 2006 wurde durch das Schweizer Volk ein Verfassungsartikel angenommen, der „die Kantone durch die Bundesverfassung (BV) verpflichtet, in der Zusammenarbeit untereinander sowie mit dem Bund für hohe Qualität und Durchlässigkeit des Bildungsraumes Schweiz zu sorgen (Artikel 61a Absatz 1 BV)“ (EDK, 2015, S. 3). Die Kantone erarbeiteten anschließend die INTERKANTONALE VEREINBARUNG ÜBER DIE HARMONISIERUNG DER OBLIGATORISCHEN SCHULE (HARMOS-Konkordat) (Plenarversammlung der EDK, 2007) mit dem hauptsächlichen Anliegen, „die Ziele des Unterrichts und die Schulstrukturen“ (S. 1) sowie die Durchlässigkeit der Schulsysteme zu harmonisieren.

Obwohl dem Konkordat nur 15 Kantone beitraten¹⁰¹, wurden die wesentlichen Ziele des Verfassungsartikels flächendeckend umgesetzt (EDK, 2015, S. 26). Strukturell bedeutet dies, dass die Schülerinnen und Schüler insgesamt nach 8 Jahren Kindergarten und Primarschule für drei Jahre in die Sekundarstufe 1 (EDK, 2015, S. 15) übertreten. Die Sekundarstufe 1 bildet mit den Klassen 7–9¹⁰² den Abschluss der obligatorischen Schulzeit. Die Schülerinnen und Schüler sind in dieser Phase in der Regel in einem Alter zwischen 12–16 Jahren. Trotz Harmonisierung der strukturellen Vorgaben ergibt sich aber weiterhin ein heterogenes Bild bezüglich der Ausgestaltung der Sekundarstufe 1. Erläuterungen dazu folgen im dritten Abschnitt, ein Überblick über das gesamte Bildungssystem der Schweiz befindet sich im Anhang.

Nach der strukturellen Vereinheitlichung der Volksschule wurden nationale Bildungsziele (EDK, 2015, S. 19) beschlossen und sprachregional getrennt einheitliche Lehrpläne entwickelt. In der deutschsprachigen Schweiz wurde der Lehrplan 21¹⁰³ mittlerweile in allen Kantonen mit wenigen marginalen Anpassungen eingeführt. Das föderalistische System

¹⁰¹ https://edudoc.educa.ch/static/web/arbeiten/harmos/beitritt_harmos_kantone_d.pdf (abgerufen am 07.06.2019).

¹⁰² Die Schulstufen werden immer noch ohne Kindergarten gerechnet.

¹⁰³ <https://www.lehrplan21.ch> (abgerufen am 07.06.2019).

der Schweiz führte zu einigen kantonalen Abstimmungen aufgrund von Referenden¹⁰⁴ zur Einführung der Lehrpläne, die aber ausnahmslos zugunsten des neuen Lehrplans entschieden wurden (Hansen et al., 2018, S. 12).

Eine weitere Eigenheit des föderalistischen Systems ist die Hoheit der Kantone über die Anerkennung der Diplome, was einen direkten Einfluss auf die Anstellungssituation der Lehrkräfte hat. Die Schulleitungen vor Ort haben seit dem Jahr 2000 zunehmend Aufgaben der kantonalen Verwaltungen übernommen. Mittlerweile liegt die pädagogische Aufsicht über die Lehrkräfte bei den Schulleitungen und zunehmend werden auch Anstellungen durch diese (Heinzer & Hangartner, 2016, S. 47) vorgenommen, meist noch unter der Aufsicht von lokalen Schulbehörden. Zwar werden die Anstellungsbedingungen meist kantonal vorgegeben, die Überprüfung der Qualifikationen der Lehrkräfte bleibt aber Schulleitungssache. Zudem wird in den Medien regelmäßig darüber berichtet, dass aufgrund des zunehmenden Mangels an Lehrkräften oft unqualifiziertes Personal vor den Klassen steht (z. B. Birrer, 2017), was sich später in dieser Studie in der Auswahl des Samples offenbart (vgl. Kapitel 9).

Die Bemühungen um Vereinheitlichung führten dazu, dass im deutschsprachigen Teil des schweizerischen Bildungssystem mittlerweile alle Schülerinnen und Schüler im 7. Schuljahr in die Sekundarstufe 1 wechseln und der Unterricht inhaltlich durch den Lehrplan 21 (D-EDK, 2015) gesteuert wird. Ob dieser neue Lehrplan in Bezug auf Inhalt und Konzept nach der kurzen Einführungszeit schon in die Klassenzimmer hineingetragen werden konnte, darf bezweifelt werden. Er ermöglicht aber einen einheitlichen Blick auf gewünschte Ziele aus Sicht der Bildungsträger und kann als Projektionsfläche für Vorstellungen der untersuchten Lehrkräfte dienen. Abgesehen davon wird auch die Heterogenität im Feld der Musiklehrkräfte einen Einfluss auf die Umsetzung haben, ausgelöst durch sich verändernde Anstellungsmodalitäten und eine veränderte Ausbildungslandschaft.

¹⁰⁴ Das Referendum wird ergriffen, um eine Volksentscheidung zu einem Gesetzesentwurf zu erzwingen.

6.2 Ausbildung von Musiklehrkräften in der Schweiz

Criblez und Lucien (2007) beschreiben den Übergang einer sehr heterogenen Ausbildungssituation mit 145 Ausbildungsstätten in den 1990er-Jahren hin zu 18 Lehrerbildungsinstitutionen im Jahr 2006, angegliedert an Universitäten, Fachhochschulen oder in autonomer Organisation. Lehmann (2007) kommt zu dem Schluss, dass durch die Harmonisierung der Unterrichtsmodelle zwar die Unterrichtsbefähigung insgesamt gestärkt werden konnte, die Wahl- bzw. Abwahlmöglichkeit bestimmter Fächer aber zu neuen Problemen geführt hat. Er schreibt:

„Viele Behörden und Schulleitungen tun sich schwer mit eingeschränkt einsetzbaren Lehrpersonen und kleinere Schulen fürchten sich davor, keine passenden Lehrpersonen zu finden. Dies kann dazu führen, dass Lehrpersonen ein Fach unterrichten, für das sie keinerlei Ausbildung haben. Dies steht in grossem Widerspruch zur Forderung nach einem kompetenten und professionellen Unterricht in allen Fächern.“ (S. 27)

Diese Aussage zeigt bezüglich der Ausbildung von Musiklehrpersonen zwei Aspekte auf, die in einem ambivalenten Verhältnis stehen. Einerseits wird in Schulen nach Lehrkräften verlangt, welche möglichst breit einsetzbar sind und andererseits sollen diese Lehrkräfte kompetent und professionell unterrichten. Diese Ambivalenz zeigt sich direkt in den verschiedenen Ansätzen der Ausbildungsstätten für Musiklehrkräfte. Die pädagogischen Hochschulen bilden zugunsten einer breiten Einsetzbarkeit für die Sekundarstufe 1 Lehrkräfte in üblicherweise 3–4 Fächern aus, was gezwungenermaßen zu weniger Tiefe in der fachlichen Ausbildung führt. Diese Lehrkräfte bringen gerade für kleine Schulen die gewünschte Breite mit, verfügen aber im Fach Musik meist über ein weniger hohes musikalisch-künstlerisches Niveau. Die Musikhochschulen dagegen bilden Monofachlehrkräfte auf hohem fachlichen Niveau aus, die neben ihrer Tätigkeit als Schulmusiker idealerweise auch als Künstler unterwegs sind, wie Hofstetter (Hochschule für Musik FHNW; Basel) in einem Interview ausführt (Henning, 2013). Diese Lehrkräfte bieten zwar eine hohe Professionalität, haben aber als Lehrkräfte der Sekundarstufe 1 durch die begrenzte Dotation der Musikstunden meist wenig Kontakt mit Schülerinnen und Schülern und sind damit pädagogisch stärker herausgefordert.

Scheidegger (2008) zeichnet den zu Beginn des Abschnitts erwähnten Wechsel hin zu pädagogischen Hochschulen für die Zentralschweiz nach. Er bestätigt die anspruchsvolle Situation für Monofachlehrpersonen: „Allein als Monofachlehrperson auf der Sekundarstufe I zu unterrichten, war schwierig, und es waren denn auch meist Personen, die bereits eine Lehrerausbildung absolviert hatten und beruflich verankert waren, die das Studium der Schulmusik I aufnahmen.“ (S. 74) Er bedauert den Systemwechsel der Hochschule hin zu mehr Fächern und die damit verbundene Verkürzung der Ausbildung für Schulmusik 1, sieht aber vor allem eine Einschränkung möglicher Tätigkeiten über die Schule hinaus, z. B. als Chor- und Kursleiter.

Der Wechsel weg von Mono- hin zu Mehrfachlehrkräften führt aktuell zu einer breiten Palette biografischer Hintergründe von Musiklehrkräften in der Sekundarstufe 1, auch bedingt durch den allgemeinen Lehrkräftemangel. Die nachfolgende Aufzählung von Profilen verdeutlicht die heterogene Situation:

- *Gymnasiallehrkräfte*: Ausbildung für Schulmusik 2, mit einer zusätzlichen Vertiefung in Chorleitung, Gesang, Musikwissenschaft, Musik und Bewegung oder auf einem Instrument. Unterricht an Gymnasien der Stufen 10–13 (z. T. 7–13 oder 9–12), aber auch in den verschiedenen Niveaus der Sekundarstufe 1.
- *Schulmusiker in der Sekundarstufe 1*: Ausbildung Schulmusik 1 zusätzlich zu einer generalistischen Ausbildung für diese Stufe oder zu einem instrumentalen Masterdiplom. Unterricht auf allen Niveaus der Sekundarstufe 1, oft aber an Progymnasien (Aargau: Bezirksschule; Solothurn: SEK-P¹⁰⁵).
- *Mehrfachlehrkräfte*: Ausbildung mit den neuen Profilen der pädagogischen Hochschulen, ausgebildet für 2–4 Fächer. Unterricht in allen Leistungsniveaus der Sekundarstufe 1.
- *Generalisten*: Ausbildung an den ehemaligen Seminaren für die Sekundarstufe 1, meist mit einem kleinen musikalischen Anteil, sogenannte Zehnkämpfer. Musikunterricht in der eigenen Klasse, bei Erfolg oft auch in anderen Klassen.

¹⁰⁵ Die beiden Beispiele betreffen das Untersuchungsgebiet der Studie, im Kanton Bern heisst die Stufe z. B. Spez-Sek. In weiteren Kantonen wird der progymnasiale Unterricht von Gymnasiallehrkräften erteilt (z. B. Luzern und Zürich).

- *Sängerinnen und Sänger*: Masterausbildung für Gesang, ohne schulmusikalische Ausbildung. Anstellung über Beziehungen oder aufgrund von Mangel an Lehrkräften.
- *Instrumentalistinnen und Instrumentalisten*: Masterausbildung am Instrument, keine vokale Ausbildung, keine schulmusikalische Ausbildung. Anstellung über Beziehungen oder aufgrund von Mangel an Lehrkräften.
- *(Angehende) Lehramtsstudierende*: häufig unterrichten Studierende bereits während des Studiums, teilweise sogar bereits vor Aufnahme eines Studiums. Die Anstellung erfolgt aufgrund von Mangel an Lehrkräften.

6.3 Leistungsniveaus in der Sekundarstufe 1

Wie bereits dargelegt, liegt die Hoheit über die Schule bei den Kantonen. In der Sekundarstufe 1 werden meist verschiedene Leistungsniveaus angeboten. Für die Kantone Aargau und Solothurn wird an dieser Stelle ausführlich über die verschiedenen Niveaus berichtet, da die Befragten der Interviewstudie vor allem aus diesen Kantonen stammen. Zusätzlich werden einige Informationen zu Privatschulen geliefert, da eine interviewte Lehrerin aus einer Waldorfschule stammt.

Die Harmonisierung der Schulsysteme hat im Kanton Aargau zu einer Verlängerung der Primarschule von fünf auf sechs Jahre geführt und die Oberstufe von vier auf drei Jahre verkürzt. Nach wie vor gibt es aber eine Einteilung in drei Niveaus. Die Bezirksschule ist als progymnasiales¹⁰⁶ Leistungsniveau aufgestellt mit Anschlussmöglichkeit an die Kantonsschulen, wie die Gymnasien im Aargau heißen. Rund die Hälfte der Schülerinnen und Schüler aus der Bezirksschule entschließen sich aber für eine anspruchsvolle Berufslehre. Die Sekundarschule ist auf ein mittleres und die Realschule auf ein tiefes Leistungsniveau ausgerichtet. Für beide Typen sind im Anschluss Berufslehren mit unterschiedlichen Ansprüchen vorgesehen. Die Kleinklassen bestehen aus wenigen Schülern mit starkem

¹⁰⁶ In der Schweiz wird das progymnasiale Niveau als höchstes Leistungsniveau der Sekundarstufe 1 bezeichnet, im Wortsinn als Vorstufe zum Gymnasium, das in Klasse 9–12 bzw. 10–13 besucht wird. Die Gymnasien heißen in verschiedenen Kantonen auch Kantonsschulen.

Förderbedarf im intellektuellen und sozialen Bereich. Mit der Einführung der integrativen Schulung wurden viele dieser Klassen aufgelöst und in die bestehenden Realklassen integriert¹⁰⁷. Die Zahlen in der Tabelle 10¹⁰⁸ belegen die Verteilung der Jugendlichen auf die verschiedenen Leistungsniveaus und den prozentualen Anteil an Mädchen. Dabei ist auffällig, dass mit abnehmendem Leistungsniveau der Anteil an Mädchen sinkt.

An der Bezirksschule wurden bis 2020 in den drei Schuljahren durchgehend zwei Lektionen Musik unterrichtet, ab Sommer 2020 wurde im mittleren Schuljahr auf eine Lektion verkürzt und in den unteren Niveaus werden in drei Jahren insgesamt 4 Lektionen Musik unterrichtet. Das Angebot wird mit einem freiwilligen Chorangebot ergänzt. Als einziger Kanton kennt der Kanton Aargau in der Sekundarstufe 1 zusätzlich den freiwilligen Instrumentalunterricht. Das Angebot besteht aus einer 15-minütigen Lektion pro Woche, ausgerichtet von lokalen Musikschulen und meist ergänzt mit einem kostenpflichtigen Anteil.

Abkürzung			Musik-Lektionen verteilt auf 3 Jahre	Schülerinnen & Schüler pro Leistungsniveau	Anteil Mädchen
Bez.	1.–3. Bezirks-	7.–9. Klasse	2/1(2)/2	39 %	53.7 %
Sek.	1.–3. Sekundar-	7.–9. Klasse	2/1/1	35 %	47.2 %
Real	1.–3. Realschule	7.–9. Klasse	2/1/1	23 %	43.2 %
KK	Kleinklasse	1.–9. Klasse	2/1/1	3 %	38 %

TABELLE 10: Zahlen zur Sekundarstufe 1 im Kanton Aargau

Traditionsgemäß dauert die Primarschule im Kanton Solothurn sechs Jahre. Bei der Harmonisierung wurde einzig der Zugang zur Kantonschule (Gymnasium) angepasst. Neu gibt es die regionalen Sek P-Schulen, die einen Übertritt an die Kantonsschule bereits nach dem 8. Schuljahr ermöglichen. Im Gegensatz zum Kanton Aargau ist in dieser Stufe keine anschließende Berufslehre vorgesehen. Die beiden unteren Niveaus sind aufgeteilt in die SEK E mit erweiterten Anforderungen, die zu anspruchsvollen

¹⁰⁷ Die Entscheidung der Einführung wurde auf der Ebene der Kommune (Gemeinde) getroffen.

¹⁰⁸ https://www.ag.ch/media/kanton_aargau/dfr/dokumente_3/statistik/publikationen/statistikthemen/15_bildung/volksschule/Schulstatistik_2018_2019.zip (abgerufen am 06.06.2019).

Berufslehren führen soll und die Sek B mit Basis-Anforderungen, die entsprechend auf einfachere Berufe vorbereitet. Ähnlich wie im Kanton Aargau wird auf Stufe des Schulstandorts beschlossen, ob eine Sek K (Kleinklasse) geführt wird, oder ob die Schülerinnen und Schüler in der Sek B speziell gefördert werden. Die Zahlen der Tabelle 11¹⁰⁹ zeigen die Verteilung der Jugendlichen in die verschiedenen Leistungsniveaus und den prozentualen Anteil an Mädchen. Obwohl sich die Strukturen im Kanton Solothurn leicht vom Kanton Aargau unterscheiden, zeigt sich bezüglich der Mädchen-Anteile das gleiche Bild: je tiefer das Leistungsniveau, desto weniger Mädchen.

In der Sek P werden in beiden Jahren zwei Lektionen Musik unterrichtet. In den anderen Niveaus wird neben einer einzigen Musiklektion noch das Wahlfach Musik/Chor unterrichtet. Der Instrumentalunterricht wird vom Kanton mit einem Beitrag pro Schülerin/Schüler subventioniert, obliegt aber den lokalen Musikschulen.

Abkürzung			Lektionen verteilt auf 3 Jahre	Schülerinnen & Schüler pro Leistungsniveau	Anteil Mädchen
Sek P	Progymnasium	7.–8.	2-2	16 %	55 %
Sek E	Erweiterte Anforderungen	7.–9.	1-1-1	47 %	50 %
Sek B	Basisanforderungen	7.–9.	1-1-1	36 %	43 %
Sek K	Kleinklasse	7.–9.	1-1-1	1 %	41 %

TABELLE 11: Zahlen zur Sekundarstufe 1 im Kanton Solothurn

In der Schweiz gibt es keine freie Schulwahl. Trotz fehlender Subvention beträgt der Anteil an privaten Bildungsinstitutionen ca. 10 %¹¹⁰ mit einem Anteil von 5 % von Schülerinnen und Schülern. Konsultiert man die Listen mit den bewilligten Privatschulen der Kantone Aargau und Solothurn, wird sichtbar, dass neben einem breiten Angebot für Kindergarten und Primarschule nur wenige Schulen auch die Sekundarstufe 1 bedienen. Bereits seit langem etabliert sind die Rudolf-Steiner-Schulen, die in der Schweiz nach ihrem Begründer benannt werden, anders als in

¹⁰⁹ https://www.so.ch/fileadmin/internet/dbk/dbk-vsa/Schulsystem/Statistik/2019_statistik_oblig_schule.pdf (abgerufen am 06.06.2019).

¹¹⁰ <https://www.bfs.admin.ch/bfs/de/home/statistiken/bildung-wissenschaft/bildungsinstitutionen.html> (abgerufen am 06.06.2019).

Deutschland, wo sie als Waldorfschulen¹¹¹ bekannt sind. Der traditionell musisch geprägte Unterricht beinhaltet auch das Fach Musik, wobei die Ausgestaltung in jeder Schule anders ist. Im Vergleich zu den staatlichen Schulen wird mit Chor- und Orchesterprojekten, der obligaten Eurhythmie und weiteren Musikstunden mehr Wert auf die musikalische Bildung gelegt als in den staatlichen Schulen. Ein weiterer Unterschied ist das Fehlen von Leistungsniveaus in der Sekundarstufe 1.

Für die vorliegende Untersuchung können für diesen Abschnitt zusammenfassend zwei Erkenntnisse genannt werden. Das Fach Musik wird in den Kantonen Solothurn und Aargau in verschiedenen Leistungsniveaus unterrichtet, wobei in den tieferen Stufen weniger Schulstunden zur Verfügung stehen. Dies hat in erster Linie einen Einfluss auf die Quantität der Inhalte des Musikunterrichts. Zweitens sind in den höheren Niveaus tendenziell mehr Mädchen anzutreffen, was sich möglicherweise in Aussagen zu Klassendynamik und Unterrichtsqualität niederschlägt.

¹¹¹ Rudolf Steiner gründete zusammen mit Emil Molt, dem damaligen Besitzer der Waldorf Astoria Zigarettenfabrik, eine Schule für die Arbeiterkinder. <https://www.waldorfschule.de/waldorfpaedagogik/allgemeiner-ueber-blick/was-ist-waldorfpaedagogik> (abgerufen 06.06.2019).

7 Forschungsfragen

Dieses Kapitel schließt den Bereich B mit den Forschungsfragen ab. Bevor diese aufgeführt werden, wird noch einmal rekapituliert, welche Themen in den vier Kapiteln manifest wurden und wie sich darin das Forschungsdesiderat zeigt.

7.1 Eingrenzung des Forschungsdesiderats

Im autoethnografischen Exkurs zeigte sich, dass eine individuelle Betreuung in Ansätzen durch den Forschenden umgesetzt werden konnte. Dabei rückte neben dem Thema Individualisierung zunehmend auch die Adressierung der adoleszenten Stimme in den Fokus. Das Kapitel zur adoleszenten Stimme belegt die Wichtigkeit einer ganzheitlichen Betrachtung der Stimmentwicklung bei Mädchen und Jungen, weil vielfältige und interindividuelle Prozesse den Körper und die Psyche beeinflussen und so gerade im sozialen Miteinander in der Klasse zur Herausforderung werden. Weiter konnte im Kapitel zum Assessment des Singens gezeigt werden, dass Beurteilungsinstrumente in Ansätzen vorliegen und Heterogenität von Lehrkräften wahrgenommen wird, Musiklehrkräften aber kaum Diagnosekompetenzen vorweisen können (Brunner, 2014). Im Exkurs zu den Schweizer Verhältnissen wurde zudem deutlich, dass das Feld der Musiklehrkräfte in der Schweiz heterogen ist und somit auch von divergierenden Kompetenzen und Vorstellungen bezüglich Singunterricht ausgegangen werden muss.

Abgesehen von einigen Studien im Primarschulbereich, die in der Einführung bereits genannt wurden (vgl. 1.3), zeigen sich kaum Studien zum Singen in der Sekundarstufe 1¹¹². Studien zur Individualisierung des Singens in der Sekundarstufe liegen ebenfalls nicht vor und durch die fehlende

¹¹² Die Studie von Gütay und Kreutz (2010) bzw. Gütay (2012) sei hier nochmals erwähnt. Untersucht wurden Kinder der 5. Klasse, die einerseits noch nicht in der adoleszenten Stimmentwicklung und andererseits zu jung für die Sekundarstufe 1 der Schweiz sind. Die Studie von Blanchard (2019) bezieht

Rezeption amerikanischer Studien muss vermutet werden, dass auch das Wissen zur adoleszenten Stimme basal ist, was sich deutlich bezüglich der adoleszenten Mädchenstimme zeigt. Die von Brunner (2014) eingebrachte These, dass intuitiv im Singunterricht viel Gutes gemacht wird, lässt aber auf interessante Erkenntnisse hoffen, weil für die empirische Studie spannende Lehrkräfte gefunden werden konnten.

Im Kapitel zur Gliederung der Studie (vgl. 2.2) wurde bereits argumentiert, dass die empirischen Daten nicht direkt im Musikunterricht gesammelt werden, sondern mittels qualitativer Interviews mit Musiklehrkräften. Entsprechend berücksichtigen die Forschungsfragen diese Perspektiven.

7.2 Forschungsfragen

Im empirischen Teil dieser Arbeit ist folgende übergeordnete Forschungsfrage handlungsleitend:

Welche Unterrichtspraxen zeigen sich in Bezug auf Diagnose, Förderung und Beurteilung von Singunterricht aus der Perspektive der Lehrkräfte in der Sekundarstufe 1, insbesondere unter Berücksichtigung individueller Bedürfnisse von Schülerinnen und Schülern und in Bezug auf geschlechtsspezifische Phänomene?

Diese knapp formulierte Forschungsfrage bezieht sich auf wesentliche Faktoren von Assessment und bezieht zwei Teilperspektiven mit ein. Erstens kann Assessment seine Ganzheitlichkeit nur dann entfalten, wenn individuelle Bedürfnisse adäquat adressiert werden, und zweitens wirkt die adoleszente Stimmentwicklung in hohem Maß auf die Beurteilungs- und Förderprozesse ein. Um der Forschungsfrage gerecht zu werden, wird deshalb immer eine mehrperspektivische Betrachtungsweise angestrebt. Die erste Teilfrage nimmt die Ziele der Lehrkräfte ins Visier, dies als Grundlage für die nachfolgenden Teilfragen, die sich auf die Begriffe Diagnose, Förderung und Beurteilung beziehen. Die Teilperspektiven werden jeweils

sich zwar oft auf das Singen, hilft aber nicht, die vorliegenden Fragen zu beantworten, da der Fokus auf der Konstituierung des Musikunterrichts liegt.

miteinbezogen, ohne diese durchgehend in die Formulierungen zu integrieren.

I. Welche Ziele verfolgen Lehrkräfte in ihrem Unterricht in Bezug auf das Singen und wie kann dies auf eine lokale Schulkultur bezogen werden?

Lehrkräfte gestalten ihren Unterricht auf der Basis von Zielen, die implizit oder explizit handlungsleitend sind (Niessen, 2014). Zudem sind Ausbildung und Biografie der Musiklehrkräfte beeinflussende Größen für die Gestaltung des Singunterrichts, wie auch lokale Schulkulturen, die sich positiv oder negativ auf das Erreichen von Zielen auswirken können. Nicht zuletzt gilt es zu beachten, dass auf verschiedenen Leistungsniveaus unterrichtet wird. Die Teilfrage 1 zielt deshalb darauf ab, die lokalen und individuellen Singkulturen der Interviewpartnerinnen und -partner zu beschreiben, um die anschließende Betrachtung aus der Perspektive von Assessment adäquat einordnen zu können.

II. Welche Vorstellungen bezüglich individueller Förderung können bei Lehrkräften rekonstruiert werden und inwiefern werden dabei geschlechtsspezifische Verhaltensweisen im Unterricht wirksam?

Das Thema Individualisierung ist spätestens seit 2002 ein unhinterfragter Anspruch an die Schule (Klieme & Warwas, 2011). In den letzten Jahren wurden ganzheitliche Assessmentstrategien im deutschsprachigen Raum (Winter, 2015; C. Schmidt, 2020; Schaumberger, 2020a) als geeignete Strategie für gelingende Förderung und Beurteilung beschrieben. Es stellt sich die Frage, ob und in welcher Form sich dieser Anspruch mehr als 10 Jahre später im Musikunterricht manifestiert. Werden individuelle Bedürfnisse im Unterricht adressiert? Spielen dabei geschlechtsspezifische Merkmale, wie Vorbildwirkung oder Gruppendynamische Phänomene eine Rolle? Wird insbesondere von Tendenzen eines monoedukativen Singunterrichts berichtet? Insgesamt stellt sich die Frage, ob individuelle Förderung sich im Singunterricht etablieren konnte. Dabei wird nicht nur der Bereich *assessment for learning*, sondern auch die metakognitive Ebene von *assessment as learning* untersucht.

III. Von welchen diagnostischen Mitteln berichten Lehrkräfte und welchen Einfluss haben diese in der Vorbereitung und Durchführung von Unterricht und im Umgang mit Schülerinnen und Schülern?

Die Teilfrage III bezieht sich auf einen Teilaspekt von *assessment for learning*. Diagnostische Fähigkeiten sind gerade beim Singunterricht von zentraler Bedeutung und sind als Start in eine Unterrichtsphase ebenso wie im Unterrichtsverlauf die Basis individueller und adaptiver Unterrichtsgestaltung (Imthurn, 2020). Untersucht werden soll, ob Lehrkräfte bereits vor oder schon zu Beginn¹¹³ der Umsetzung des kompetenzorientierten LP21 von impliziten oder expliziten Strategien zur Diagnose berichten.

IV. Wie werden Leistungen von Schülerinnen und Schülern gemessen und wie wird darauf vorbereitet?

Die Teilfrage IV bezieht sich auf *assessment of learning* und zielt auf summative Beurteilungsvorgänge ab. Das *Vorsingen* bzw. die *Singnote* ist im Musikunterricht der Schweiz ein gängiges Mittel für die Leistungsbeurteilung und hat einen Einfluss auf Laufbahnentscheide. Spätestens bei der summativen Bewertung von Gesang wird eine individuelle Note gesetzt. Der Anspruch an Transparenz und individueller Passung wird zumindest in den kantonalen Verordnungen der Nordwestschweiz¹¹⁴ flächendeckend verlangt. In dieser Teilfrage interessiert erstens der Aspekt, wie Musiklehrkräfte ihre Beurteilungen durchführen und begründen, und zweitens wird der Frage nach individueller Passung nachgegangen.

¹¹³ Die Interviews für die vorliegende Studie wurden 2017 und zu Beginn 2018 durchgeführt. Der LP21 war bereits 2015/2016 in den Kantonen Baselland und Baselstadt eingeführt worden, stand aber im Kanton Solothurn (2018/2019) und im Kanton Aargau (2020/2021) noch vor der Einführung (<https://www.lehrplan21.ch/stand-der-einfuehrung> – abgerufen am 23.08.2021).

¹¹⁴ „Die Gewichtung [der Beurteilungsbelege] liegt im pflichtgemäßen Ermessen der Lehrpersonen, wobei auf die Entwicklung während der Beurteilungsperiode besonders Rücksicht genommen werden muss.“ (Verordnung über die Laufbahnentscheide an der Volksschule, 2009, § 5, Abs. 2, S. 2).

C METHODE UND UNTERSUCHUNG

Die fehlende Forschung zum Singunterricht in der Sekundarstufe 1 im deutschsprachigen Raum, die fehlende Rezeption von Literatur zur adoleszenten Stimme aus dem anglofonen Raum sowie die noch wenig etablierte ganzheitliche Assessment-Kultur in der Schule (C. Schmidt, 2020; Winter, 2015), sowohl im Fach Musik (Lehmann-Wermser, 2019) als auch im Singunterricht (Schaumberger, 2020a), erfordern für einen ersten Zugang zum Feld eine Forschungsmethode, die eine breite Darstellung von Vorstellungen, Haltungen, Zielen und Methoden von Musiklehrkräften ermöglicht. Mit der vorliegenden Interviewstudie werden die Vorgänge im Musikunterricht aus der Perspektive der Lehrkräfte empirisch verfügbar gemacht, gleichzeitig aber über das Deskriptive hinaus mit dem Fokus auf Assessment und die adoleszente Stimme irritiert und theoretisch reflektiert. Strübing et al. (2018) sehen in dieser sich gegenseitig beeinflussenden Verbindung von Empirie und Theorie den Kern qualitativer Forschungsmethoden. Die theoretische Fundierung, zusätzlich inspiriert von Fragestellungen des autoethnografischen Exkurses, wurde bereits in den vorhergehenden Kapiteln entwickelt. Das Kapitel 8 dient nun als Scharnier, um die qualitative Inhaltsanalyse (QIA) nach Mayring (2015) als geeignete Forschungsmethode für die Untersuchung der Interviews zu begründen und die damit verbundenen Gütekriterien zu verankern. In Kapitel 9 wird die Auswahl der Interviewpartnerinnen und -partner basierend auf dem Modell professioneller Handlungskompetenz nach Baumert und Kunter (2006) sowie die Durchführung und Transkription der Interviews näher beschrieben. Schließlich wird in Kapitel 10 die Entwicklung des Kategoriensystems und der Kodierprozess referiert und die Objektivität des Resultats mittels Interrater-Reliabilitätstest geprüft. Der Bereich C macht insgesamt den strukturierten Forschungsprozess der QIA intersubjektiv nachvollziehbar (Mayring & Fenzl, 2019, S. 633) und liefert die Grundlage für den analytischen Teil (Bereich D) mit der Darstellung der Resultate.

8 Forschungsdesign

Die vorliegende Arbeit versteht sich als Beitrag zum Assessment-Diskurs (Fautley, 2010; Brophy, 2019a, 2019b; Brophy & Haning, 2020) in der Musikpädagogik. Ziel dieses Kapitels ist die Darstellung des Forschungszugangs zur Thematik des Singens über Interviews und deren Auswertung durch eine QIA nach Mayring (2015; Mayring & Fenzl, 2019). Bevor das Forschungsdesign dargelegt und bezüglich möglicher Gütekriterien reflektiert wird, werden die von Gläser und Laudel (2010) genannten Prinzipien für eine verlässliche Produktion von Wissen auf einen aktuellen Diskussionsanstoß von Strübing et al. (2018) zu Gütekriterien qualitativer Sozialforschung bezogen. Dabei zeigt sich eine Weiterentwicklung der von Gläser und Laudel postulierten Linearität des klassischen Forschungsprozesses zugunsten iterativer und sich gegenseitig beeinflussender Verfahren. Das vorgeschlagene Gerüst dient anschließend als Folie für die Begründung einzelner Forschungsschritte.

Gläser und Laudel (2010) leiten für „die Beurteilung und Auswahl von Untersuchungsstrategien und Methoden“ (S. 29) vier Prinzipien für die verlässliche Produktion von Wissen für einen gemeinsamen Theoriekorpus her. Das *Prinzip der Offenheit* stellt sicher, dass neben den erwarteten Forschungsergebnissen auch Unerwartetes berücksichtigt wird. Dies geschieht immer unter der Prämisse *eines theoriegeleiteten Vorgehens* als zweites Prinzip bzw. dem Anschluss an vorhandenes Vorwissen. Das *Prinzip der intersubjektiven Nachvollziehbarkeit* wird durch eine regelgeleitete Forschungsmethode erreicht und mit dem *Prinzip des Verstehens als Basishandlung* als „konstitutive Leistung des Forschers“ (S. 33) vervollständigt. Im weiteren Verlauf skizzieren sie eine lineare Abfolge von Tätigkeiten, beginnend mit dem Formulieren der Forschungsfrage bis hin zur Auswertung und Interpretation der Daten. Für die Forschungspraxis stellen sie aber „Rückkopplungen [...], in denen Erfahrungen aus späteren Phasen aus Untersuchungen genutzt werden“ (S. 36), fest, die möglicherweise die methodologische Strenge untergraben.

Strübing et al. (2018) stellen genau diese Linearität des Forschungsprozesses bzw. dessen methodologische Strenge für die qualitative Forschung grundsätzlich infrage:

„Jeder Versuch, die Linearität standardisierter Prozessmodelle in einer qualitativ-interpretativen Forschungspraxis zu realisieren, muss aus zwei Gründen scheitern: weil Datengewinnung, -analyse und Theoriebildung fortwährend aufeinander verweisende und verwiesene Prozesse sind; und weil die Empirizität, die Methodizität und die Theorizität qualitativer Forschung sich nur in Prozessen wechselseitiger Durchdringung dieser Prozesse entfalten können.“ (Strübing et al., 2018, S. 85)

Die These des Scheiterns der Linearität qualitativer Forschungsprozesse führt in der Argumentation von Strübing et al. zu der Entwicklung fünf übergeordneter Gütekriterien (S. 85 ff.), die den oben genannten Prinzipien von Gläser und Laudel nicht im Grundsatz widersprechen, sondern hauptsächlich abweichend strukturiert sind und einen iterativ-zyklischen Forschungsprozess skizzieren. Als Basiskriterium wird *Gegenstandsangemessenheit* genannt, mit dem gefordert wird, das empirische Feld ernst zu nehmen, aber durch theoretische Ansätze unter Spannung zu setzen. Sie schreiben: „Es geht darum, Teilnehmersichten zuerst zu verstehen und dann durch sie ‚hindurchzusteigen‘.“ (S. 88) Sie erwarten Flexibilität im Umgang mit Methoden sowie Geschicklichkeit in der Datenanalyse und beschreiben den Forschungsprozess als „ein Vorgehen, das insgesamt kreativer und experimenteller sein muss, als es Untersuchungspläne zulassen“ (S. 87). Das iterative Moment zeigt sich hier in der fortwährenden Justierung von methodischen Zugängen im Verlauf der Forschungsarbeit.

Als zweites Gütekriterium nennen sie die *empirische Sättigung*, die durch drei Komponenten gewährleistet wird: erstens wird der Bezug zum Feld und die eigene Positionierung darin als elementar betrachtet, zweitens wird „der Umfang und die Zusammensetzung des Datenkorpus“ (S. 89) als wesentlich angesehen und drittens wird erwartet, dass durch Modifikation während der Datengewinnung sich die Analysetiefe intensiviert und eine Offenheit für die Entdeckung von Neuem besteht. Komplementär dazu wird als drittes Gütekriterium die *theoretische Durchdringung* angeführt. Diese ermöglicht, das empirische Material an den „jeweiligen Fachdiskurs“ (S. 91) anzuschließen und in diesen zu integrieren. Sie schreiben:

„Gute qualitative Forschung entwickelt also statt der dichotom strukturierten Entscheidung zwischen entweder deduktiv oder induktiv verfahrender Forschung eine abduktive Forschungshaltung (Reichertz, 2003),

in der induktive, abduktive und deduktive Modi den Prozess abwechselnd und ineinandergreifend strukturieren und so die iterative Zyklik konstituieren, ohne die qualitative Forschung nicht gelingen kann (Strübing, 2013, S. 128).“ (S. 92)

Die Autoren erwarten für die qualitative Sozialforschung eine abduktive¹¹⁵ Forschungshaltung bzw. kreatives Denken oder das Sehen von Struktur im vermeintlichen Chaos als Ergänzung zu deduktiven und induktiven Elementen (Reichertz, 2013, S. 72 ff.). Erst durch das vierte Gütekriterium, die *textuelle Performanz*, kann ein Forschungsbeitrag seine Wirkung entfalten. Speziell bei qualitativen Studien müssen Forschende in der Lage sein, sinnstiftend zu argumentieren, sodass die Rezipienten die Argumentationsweise nachvollziehen und „zugleich ihre eigenen Schlüsse aus der Lektüre“ (S. 94) ziehen können. *Originalität* als fünftes Güte Merkmal wird dann erreicht, wenn der erzielte Wissensstand das „der Forschung, des Alltagswissens und des Sachwissens der Teilnehmer“ (S. 95) übersteigt.

Die Prinzipien nach Gläser und Laudel (2010) werden mit den fünf Gütekriterien im Allgemeinen bestätigt, unterscheiden sich aber deutlich in der Frage der Forschungslogik. Gläser und Laudel verlangen zwar eine Offenheit bezüglich der Resultate, betonen aber den Anspruch an methodologischer Strenge. Strübing et al. (2018) dagegen erheben das iterative Vorgehen und eine abduktive Forschungshaltung zur Norm. Insbesondere der kreative und geschickte Umgang mit Methoden wird stark betont. Zudem wendet sich Strübing (2017) selbst an anderer Stelle gegen den hegemonialen Anspruch einzelner Methoden. Für die vorliegende Forschungsarbeit gilt es, die postulierte Verzahnung von Theorie, Methode und Empirie darzulegen und zu zeigen, wie die schrittweise annähernde Hinführung zur Analyse und Darstellung der Resultate ausgestaltet ist.

Der Aufbau der Studie wird in Abschnitt 8.1 dargelegt. Dabei steht die Konstruktion des Datenkorpus und die Wahl der Auswertungsmethode im Fokus, wobei zusätzlich der Einbezug eines autoethnografischen Exkurses argumentiert wird. Nachfolgend wird in Abschnitt 8.2 die QIA als Untersuchungsmethode und in Abschnitt 8.3 die Form der Interviews und ihr

¹¹⁵ Reichertz (2013) erklärt Abduktion folgendermaßen: „Die Abduktion sucht angesichts überraschender Fakten nach einer sinnstiftenden Regel, nach einer möglicherweise gültigen Erklärung, die das Überraschende an den Fakten beseitigt.“ (S. 72)

Kommunikationszusammenhang diskutiert, um abschließend in Abschnitt 8.4 über Gütekriterien nachzudenken.

8.1 Forschungsdesign

In diesem Abschnitt stehen Überlegungen zum Datenkorpus im Zentrum, wobei einerseits die Methode der Datenerhebung und andererseits der Ursprung der Daten diskutiert wird. Nachfolgend wird der autoethnografische Exkurs als explorativer Zugang zur Thematik als weitere Datenquelle erläutert sowie die QIA als Auswertungsmethode eingeführt.

In der Forschungsarbeit sollen Diagnose-, Förder- und Beurteilungsprozesse aus der Perspektive der Lehrkraft untersucht werden. Es geht somit nicht darum zu erforschen, wie Schülerinnen und Schüler den Singunterricht und ihre Stimmentwicklung erleben, sondern welche Vorstellungen, Haltungen, Methoden und Ziele sich bei Musiklehrkräften für diesen Unterricht etabliert haben, was dazu führt, dass das unmittelbare Geschehen des Unterrichts nicht in den Fokus genommen wird. Der Verzicht auf einen ethnografischen Zugang durch teilnehmende Beobachtung im Unterricht oder auf dessen videografische Aufzeichnung und anschließende Interaktionsanalyse ergibt sich auch aus weiteren Gründen. So ist der direkte Kontakt mit Schülerinnen und Schülern während der adoleszenten Stimmentwicklung möglicherweise schambehaftet (vgl. 4.5) und die Beobachtung der längerfristigen Entwicklungsprozesse in der adoleszenten Stimmentwicklung bleibt in Momentaufnahmen kaum wahrnehmbar. Zusätzlich werden im Unterricht zwar Tätigkeiten und Interaktionen sichtbar, diese können aber eine Singkultur oder das Individualkonzept von Musiklehrkräften (Niessen, 2006) nicht abbilden und vermögen zudem auch nicht die individuellen Stimmqualitäten zu übermitteln. Für die vorliegende Arbeit wird deshalb als dem Gegenstand angemessen erachtet (Strübing et al., 2018), mittels Interviews zu einem kommunikativen Austausch über den Unterricht zu gelangen. Für Niessen (2017) spielen Interviews in der musikpädagogischen Forschung eine wichtige Rolle. Sie begründet folgendermaßen:

„[M]usikbezogene Erlebnisse, Bedeutungszuweisungen und Erfahrungen [können] mit Hilfe von qualitativen Interviews erfasst werden [...]

und zwar häufig in Facetten, die einer bloßen Beobachtung nicht zugänglich wären. Auch Bildungsprozesse lassen sich nur schwer beobachten, aber sehr wohl erzählen (vgl. u. a. Niessen, 2006).“ (S. 106)

Laut Niessen versprechen Interviews gerade für die Erforschung der beobachteten Entwicklungsprozesse von Schülerinnen und Schülern und für die Beschreibung von Förder- und Beurteilungsmethoden durch Lehrkräfte einen Gewinn. Die nicht beobachtbaren Facetten werden auch von Küsters (2009, S. 20) für die Sozialforschung als wichtiger Grund zugunsten von Interviews aufgeführt. Zudem zeigt die Forschung von Niessen (2006) selbst sowie eine Reihe weiterer Beispiele (z. B. Linn, 2017; Schaumberger, 2020b), dass sich die Erhebung von Daten durch Interviews in der Musikpädagogik etabliert und bewährt hat.

Nach der Wahl der Datengewinnungsmethode ergeben sich drei Anschlussfragen zur Datenherkunft. Erstens muss geklärt werden, wie die zu interviewenden Lehrkräfte ausgewählt werden und wie groß das Sample sein soll (Patton, 2015). Zweitens gibt es konzeptionelle Fragen, die die Form der Interviews betreffen (Helfferich, 2019), und drittens wird als relevant erachtet, unter welchen Voraussetzungen in den Interviews kommuniziert wird (Strübing et al., 2018; Mayring & Fenzl, 2019). Diese drei Aspekte werden später in Abschnitt 8.3 ausgeführt, anschließend an die Beschreibung der Methode, da sich bezüglich des Kommunikationszusammenhangs Ansprüche aus der Methode ergeben. Zusätzlich bestimmen Fragen rund um das Sample und die Datengewinnung maßgeblich mit, ob eine empirische Sättigung nach Strübing et al. (2018) erreicht werden kann.

Ebenfalls als wichtige Komponente der empirischen Sättigung beschreiben Strübing et al. die Positionierung des Forschenden im Feld. Die Wahl eines autoethnografischen Exkurses zur Darstellung dieser Positionierung erfolgte aus zwei Gründen. Erstens ermöglichte die autoethnografische Methode die Darstellung von retrospektiv erstellten Feldnotizen (Ellis & Bochner, 2000) zu relevanten Episoden der Berufsbiografie des Forschenden. Zweitens kann durch die Kontrastierung dieser Rückblenden mit der Reflexion des Fachdidaktikers eine analytische Ebene eröffnet werden, welche die Thematik auf die Fragen der vorliegenden Arbeit zuspitzt. Aus diesem Grund wird der autoethnografische Exkurs auch vor die theoretischen Überlegungen gesetzt. Aus forschungslogischen Gründen werden die Feldnotizen nicht in den Datenkorpus der Interviews aufgenommen,

insbesondere weil das Assessment-Konzept von Earl (2013) die Inhalte des autoethnografischen Exkurses gegen Ende der Berufspraxis bereits beeinflusst hat. Zudem müsste bei der Kombination der Datentypen mit einer Verfälschung der Kodierhäufigkeiten in der QIA gerechnet werden.

Nicht nur die Gewinnung der Daten, sondern auch deren Verarbeitung soll dem Gütekriterium Gegenstandsangemessenheit genügen. Die gewählte Methode muss gewährleisten, dass wesentliche Aspekte der Fragestellung in der Breite und Tiefe erörtert werden können. Dies bedingt eine mehrperspektivische Strukturierung des Datenmaterials entlang der theoretisch erarbeiteten Aspekte und der formulierten Forschungsfragen. Mit der qualitativen Inhaltsanalyse nach Mayring (2015) liegt eine erprobte Methode vor, mit der die Interview-Transkripte systematisch ausgewertet werden können. Das in mehreren Schritten iterativ herzuführende Kategoriensystem ermöglicht den Einbezug theoriegeleiteter Kategorien und lässt trotzdem durch weitere induktiv gesetzte Kategorien die notwendige Offenheit zu, die Gläser und Laudel fordern. Mayring und Fenzl (2019) schreiben: „Das Kategoriensystem (als die Zusammenstellung aller Kategorien) ist das eigentliche Instrumentarium der Analyse. Mit ihm wird das Material bearbeitet und nur die Textstellen berücksichtigt, die sich auf die Kategorien beziehen.“ (S. 634) Für die vorliegende Arbeit wird so ermöglicht, die herausgearbeiteten Themen *Assessment* und *adoleszente Stimm-entwicklung* kategoriengeleitet an den Datenkorpus heranzutragen, darauf zu fokussieren und trotzdem aufgrund zusätzlicher induktiver Schleifen wichtige Zusammenhänge nicht aus den Augen zu verlieren. Die Erstellung des Instruments und der damit zusammenhängende Rückbezug auf theoretische Fragen zeigt zudem das iterativ-zyklische Vorgehen, wie es von Strübing et al. (2018) vorgeschlagen wird. Dieses Vorgehen, die Kritik an der qualitativen Inhaltsanalyse (z. B. Möller, 2012) und der zuweilen als Baukasten beschriebene Pool von Analysewerkzeugen (Schreier, 2014) wird im nächsten Abschnitt vertieft.

8.2 Auswertungsmethode: Qualitative Inhaltsanalyse nach Mayring

In diesem Abschnitt wird die Methode der *qualitativen Inhaltsanalyse* (Mayring, 2015; Mayring & Fenzl, 2019) vorgestellt. Ein kurzer Abriss der

mittlerweile 100-jährigen Geschichte der Methode zeigt die Entwicklung von einer quantitativen zu einer qualitativen Methode, wobei genau dieser Wechsel sich in der Forschungscommunity als Diskussionspunkt herausstellt (Möller, 2012) und zu fortwährender Kritik führt. Anschließend wird das Modell der strukturierenden Inhaltsanalyse als Grundlage weiterer Analysetechniken vorgestellt.

Entwicklung der Methode

Die Analysen des US-amerikanischen Politikwissenschaftlers und Kommunikationstheoretikers Harold Dwight Lasswell zur Notwendigkeit von Propaganda in der Demokratie während der 1920er-Jahre werden als Ursprung der Inhaltsanalyse betrachtet (Möller, 2012, S. 383). Nach Mayring (2015, S. 26 f.) wurde die Entwicklung der Forschungsmethode entscheidend durch die kommunikationswissenschaftliche Fundierung der *School of Journalism* an der Columbia University geprägt. Ziel war es, die Meinungsbildung in den Massenmedien zu ergründen und daraus Nutzen zu ziehen, sei dies in der Auswertung von Kriegspropaganda oder im Rahmen kommerzieller Interessen. Die erste Monografie zum Thema mit dem Namen *CONTENT ANALYSIS* (Berelson, 1952) beschreibt in der Mitte des Jahrhunderts die Methode als „objektive, systematische und quantitative Analyse des manifesten Inhalts von Kommunikation“ (Mayring, 2015, S. 26). Die Analyse von Texten erfolgt dabei in einem Dreischritt: Nach der Bildung eines Kategoriensystems folgt das Durchsuchen des Textes nach relevanten Informationen und die anschließende Kodierung. Die Absicht war immer die Erfassung der Bedeutsamkeit von Informationen und Zusammenhängen aufgrund der Häufigkeit von Kategorien. „Diese ‚klassische‘, quantitativ orientierte [...] Form der Inhaltsanalyse sah sich jedoch bereits frühzeitig einer Kritik ihrer grundsätzlichen Verfahrensweise ausgesetzt.“ (Braun & Ehrenspeck-Kolasa, 2019, S. 178) Im gleichen Jahr kritisiert nämlich Kracauer (1952) die quantitative Inhaltsanalyse in zweierlei Hinsicht. Die Annahme, dass allein das Vorhandensein von Zahlen die quantitative Analyse objektiv und reliabel mache, lehnt Kracauer gleichermaßen ab wie die Idee, nur den manifesten Inhalt zu berücksichtigen. Er fordert, dass ebenso der latente Sinngehalt, die Konnotation und der Kontext von Texten einbezogen werden soll (S. 634 ff.), und zweifelt an der Qualität einiger „impressionistic judgments“ (S. 637), also impressionistisch an-

mutenden quantitativen Analysen: „And these judgments may in fact be more unaccountable than those found in communications studies of a predominantly qualitative nature.“ (S. 637)

Die kritisierten Punkte führen in den folgenden Jahren zur Weiterentwicklung der Inhaltsanalyse, wobei zusätzliche Forderungen diskutiert wurden. Auf der sogenannten *Annenberg-School-Konferenz* im Jahr 1966 wurden insbesondere Fragen der Forschungsqualität diskutiert. Gefordert wurde eine Explizierung des Analysevorgangs und die Klärung des Kommunikationsmodells, das der Datengewinnung zugrunde liegt (Mayring, 2015, S. 27). „Auf diesem Stand der Diskussion ist die Inhaltsanalyse als kommunikationswissenschaftliches Instrument im Wesentlichen stehen geblieben (vgl. Krippendorff 2004; Neuendorf 2002).“ (S. 27) Seit Ende der 1970er-Jahre entwickelte Mayring die QIA als Auswertungsmethode kontinuierlich in den Sozialwissenschaften weiter (Möller, 2012, S. 384; vgl. Mayring, 1982) und Krippendorff (2004) konstatiert eine exponentielle Zunahme der Inhaltsanalyse durch die Verfügbarkeit von Computern und der damit verbundenen verbesserten Verfügbarkeit von Texten.

Möller (2012) stellt trotzdem fest, dass die QIA „als qualitative Methode per se infrage gestellt“ (S. 381) und „sogar ihre Existenz aufgrund forschungsqualitativer Legitimation kritisiert“ (S. 381) wird. Als Gründe nennt sie die fehlende eindeutige Zuordnung zu einem hermeneutischen Verfahren, die anhaltenden Grabenkämpfe rund um die Methode und den Umstand, dass sich die QIA nicht eindeutig der qualitativen oder quantitativen Forschungstradition zuordnen lässt (S. 381). Mayring und Fenzl (2019) nehmen die grundsätzliche Kritik auf: „Gerade in der Soziologie wurde immer wieder behauptet, dass die qualitative Inhaltsanalyse kein elaboriertes oder anspruchsvolles Verfahren sei, da ihre Verfahrensweisen zu wenig theoretisch begründet seien.“ (S. 639) Sie halten der Kritik fehlender Grundlagentheorie entgegen, dass die Verarbeitungsmechanismen in der QIA sich „auf psychologische und linguistische Konzepte alltäglicher Textverarbeitung berufen“ (S. 641). Stamann et al. (2016) begegnen dem Diskurs um die QIA, indem sie die Häufigkeit des Gebrauchs ins Feld führen:

„Qualitativ-inhaltsanalytisches Arbeiten ist innerhalb verschiedenster sozialwissenschaftlicher und erziehungswissenschaftlicher Forschungsrichtungen etabliert; wohl nicht zuletzt deshalb, weil es sich für

viele Fragestellungen und Arbeitskontexte als angemessen und praktikabel erwiesen hat.“ (S. 1)

Beide Begründungen halten einer wissenschaftstheoretischen Diskussion kaum stand, ebenso wenig wie der Verweis von Krüger und Riemeier (2014), welche die QIA ein wenig lapidar als „in der Tradition von Hermeneutik“ (S. 133) stehend einordnen.

Stamann et al. (2016) fordern aber eine präzisere Angabe der Auswertungsmethode als lediglich die Angabe *qualitative Inhaltsanalyse*. Sie schlagen vor, sich konkret auf ein Primärverfahren zu beziehen, das beschreibt, wie das Kategoriensystem gebildet wird, und anschließend Sekundärverfahren zu nennen, welche als Analyse- bzw. Auswertungsverfahren für das Forschungsvorhaben relevant sind. Primär- und Sekundärverfahren sind deshalb Gegenstand der beiden nächsten Abschnitte. Zusammenfassend zeigt sich die QIA als häufig praktizierte Methode, die auf der konstitutiven Leistung des Forschenden, dem Prinzip des Verstehens nach Gläser und Laudel (2010), aufbaut. Obwohl sie sich theoretisch nicht präzise verorten lässt, wird sie als taugliches Instrument in der Sozialforschung eingesetzt.

Primärverfahren: inhaltlich strukturierende Inhaltsanalyse

„Kern der inhaltlich-strukturierenden Vorgehensweise ist es, am Material ausgewählte inhaltliche Aspekte zu identifizieren, zu konzeptualisieren und das Material im Hinblick auf solche Aspekte systematisch zu beschreiben – beispielsweise im Hinblick darauf, was zu bestimmten Themen im Rahmen einer Interviewstudie ausgesagt wird.“ (Schreier, 2014, S. 5)

Schreier beschreibt damit das Basisverfahren von Mayring, das er selbst präzisierend „qualitativ orientierte kategoriengeleitete Textanalyse“ (Mayring & Fenzl, 2019, S. 634) nennt¹¹⁶. Im Gegensatz zu stärker hermeneutisch orientierten Verfahren, in denen ebenso Textstellen ausgewählt und einer Interpretation zugeführt werden, verweist er für die QIA

¹¹⁶ Aus Gründen der Einfachheit wird weiterhin QIA als Abkürzung verwendet.

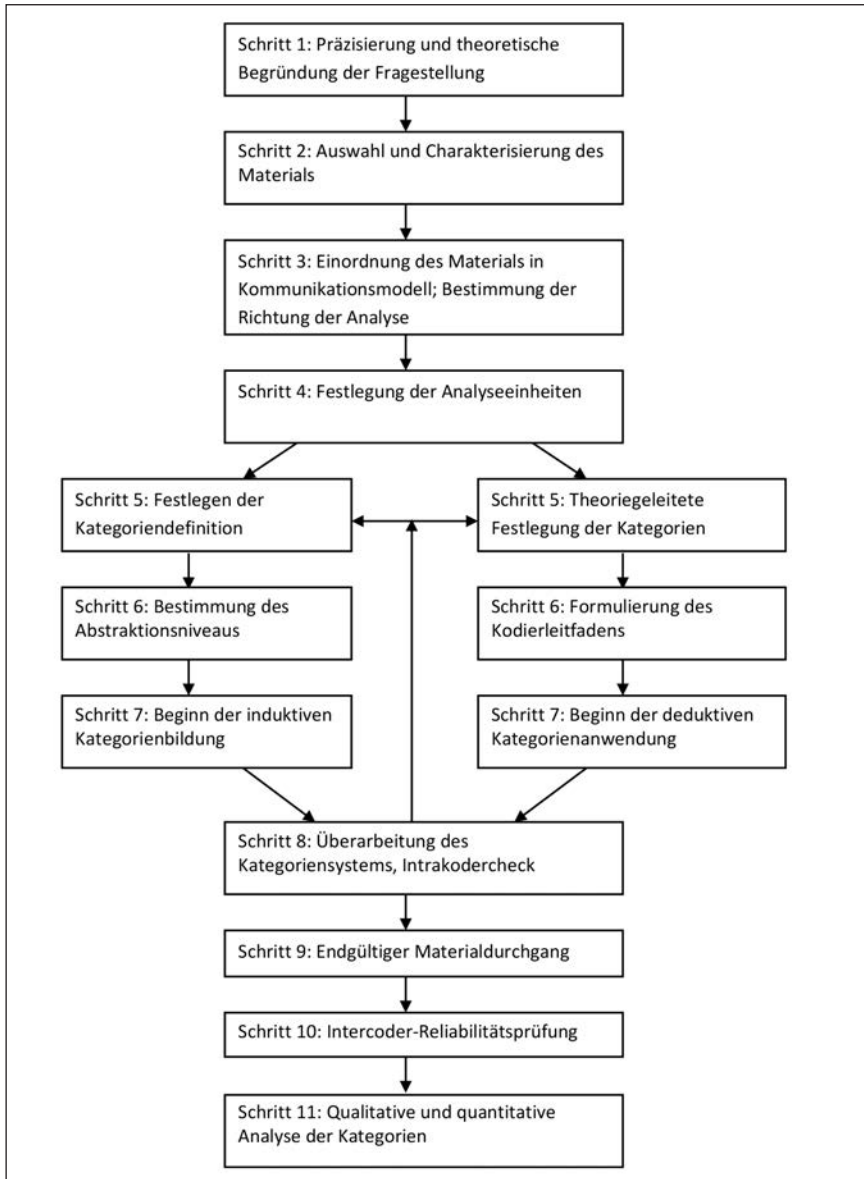


ABBILDUNG 7: Ablaufmodell induktiver Kategorienbildung und deduktiver Kategorienanwendung (Mayring & Brunner, 2006, zitiert nach Mayring & Fenzl, 2019, S. 640)

spezifisch auf die regelgeleitete Vorgehensweise (S. 635), die er einerseits als Abgrenzungsmerkmal und andererseits als Qualitätsmerkmal darstellt.

In der Abbildung 7 wird das Ablaufmodell der QIA nach Mayring und Brunner (2006) dargestellt, zitiert nach Mayring und Fenzl (2019, S. 640). Als erstes fällt auf, dass die vorgegebene Struktur der Methode vor allem den regelgeleiteten Prozess zur Bildung und Überprüfung des Kategoriensystems abbildet. Die Vorarbeiten in Schritt 1 und die Darstellung der Resultate am Ende des Ablaufmodells sind als Schnittstellen zu betrachten. Mayring und Fenzl (2019, S. 335 ff.) beschreiben folgende Schritte:

- **Schritt 1** bezieht sich auf die theoretische Fundierung der Fragestellung und wurde im Bereich B vertieft bearbeitet.
- **Schritte 2 & 3** beziehen sich auf die Auswahl des Samples und die Einordnung der Interviews in den Kommunikationszusammenhang, die im nächsten Abschnitt (vgl. 8.3) ausgeführt und für das vorliegende Material in Kapitel 9 expliziert werden.
- **Schritt 4** bezieht sich auf die Festlegung der Analyseeinheiten. Dazu gehört die Definition der Kodiereinheit, welche die minimal zu kodierende Texteinheit bestimmt. Ebenso wird festgelegt, in welcher Kontexteinheit Kodierungen gesetzt werden sollen, wobei dies in Interviews meist das ganze Interview ist. Zuletzt müssen auch Aussagen gemacht werden, inwiefern Materialteile mehrfach kodiert werden dürfen. Diese Schritte führen zu einem Kodierleitfaden.
- **Schritte 5–7** sind im Ablaufmodell doppelt besetzt. Im rechten Strang werden deduktiv Kategorien hergeleitet und im linken Strang führt der Kodierprozess selbst induktiv zu neuen Kategorien. Ziel dieses Prozesses ist die Erstellung des Kodierhandbuchs mit Definitionen der Kategorien, Ankerbeispielen aus den Transkripten und Abgrenzungshinweisen.
- In **Schritt 8** wird das Kategoriensystem in einem zirkulären Prozess überarbeitet. Bestandteil ist auch ein Intracoder-Check, in dem der Forschende Teile des Datenkorpus rekodiert und das Kodierhandbuch finalisiert.
- **Schritt 9** besteht aus einem endgültigen Durchgang durch das gesamte Material.
- **Schritt 10** bezieht sich auf die Interrater-Reliabilitätsprüfung, in der mittels einer Rekodierung die Bildung des Kategoriensystems überprüft wird (vgl. 8.4). Schritte 4–10 werden in Kapitel 10 für die vorliegende

Arbeit zugunsten der intersubjektiven Nachvollziehbarkeit ausführlich beschrieben.

- **Schritt 11** ist als Scharnier hin zu den Sekundärverfahren angelegt, die im nächsten Abschnitt besprochen werden.

Sekundärverfahren: Analyseinstrumente

Nach der Erstellung des Kategoriensystems in den Schritten 5–10 folgen quantitative und qualitative Analyseschritte in Schritt 11. Sekundärverfahren werden teilweise als eigene Methoden oder alternativ als aufbauende Analyseschritte der inhaltlich strukturierenden qualitativen Inhaltsanalyse beschrieben (Schreier, 2014; Stamann et al., 2016). In der vorliegenden Arbeit zeigten sich neben der verdichtenden und abstrahierenden Analyse der einzelnen Kategorien quantifizierende sowie typisierende Analyseschritte als angebracht. Die drei Auswertungsverfahren werden in der Reihenfolge des Erscheinens in der Darstellung der Resultate (vgl. Bereich D) nachfolgend spezifiziert.

Mayring (2015) strebt mit der QIA eine Überwindung des Grabens zwischen qualitativer und quantitativer Forschung an und nennt dies ein „integratives Methodenverständnis“ (S. 53). Aus diesem Grund schlägt er auch vor, die Kodierhäufigkeiten in einem Kategoriensystem zu analysieren, wie das in quantitativer Analyse der Normalfall ist (Mayring & Fenzl, 2019, S. 634). Er sieht in der bewusst eingesetzten quantitativen Analyse eine Möglichkeit Fallanalysen zu unterstützen und Themen auf ihre Relevanz zu untersuchen. In der vorliegenden Arbeit werden Kodierhäufigkeiten untersucht, um die Relevanz der Themen insgesamt und spezifisch bezüglich einzelner Lehrkräfte abzuschätzen (vgl. 11.1, 11.2). Innerhalb der quantitativen Auswertung der Skalierungshäufigkeiten ergibt sich eine weitere Möglichkeit, nämlich die der skalierenden Strukturierung. Dabei werden Kodierungen meist auf einer Ordinalskala eingeschätzt, um Rangfolgen zu erstellen (Mayring, 2015, S. 106 ff.). Kuckartz (2018) nennt diese Technik evaluative qualitative Inhaltsanalyse. Für die vorliegende Arbeit soll diese Technik dazu verwendet werden, Häufigkeiten von Sozialformen im Singunterricht abzuschätzen (vgl. 11.3), wobei legitimiert werden muss, wie die Skalierung erfolgt.

Mayring (2015, S. 103 ff.) nennt als eine mögliche Analysetechnik die typisierende Strukturierung, Kuckartz (2016) spricht von der

typenbildenden Inhaltsanalyse. Gemeinsam an diesem Verfahren ist der vertikale Blick auf die Fälle entlang der Themen im Kategoriensystem, die in einem zweiten Schritt verglichen und zu Gruppen zusammengefasst werden. Kuckartz schlägt dieses Verfahren nur dann vor, wenn sich eine Gruppierung von Einzelfällen als Möglichkeit abzeichnet. In Kapitel 6.1 wurde die große Heterogenität unter den Musiklehrkräften bereits erläutert, einerseits aufgrund unterschiedlicher Ausbildungen und Berufsbio-graphien, aber auch aufgrund sich verändernder Anstellungsbedingungen. Dies ergibt nicht nur Ansprüche an das Sampling (vgl. 8.3), sondern lässt auch in der Auswertung verschiedene Typisierungen erwarten, insbesondere bezüglich der Forschungsfrage I (vgl. 7.2), die nach den Zielvorstellungen im Singunterricht der Lehrkräfte fragt. Das Ziel der Typenbildung beschreibt Kuckartz folgendermaßen:

„Typenbildung basiert darauf, Ähnlichkeiten zwischen Individuen (Fällen) zu analysieren und diese Fälle entsprechend zu gruppieren, wodurch der Charakter des Singulären zugunsten des Allgemeinen zurückgedrängt wird. Das Ziel der Verallgemeinerung der Resultate qualitativer Forschung ist also ein sehr wichtiges Motiv für Typenbildung.“ (S. 41)

Durch die Typisierung können „Ähnlichkeiten und Unterschiede sowie *ihre übergreifende Struktur*“ beschrieben werden“ (Kelle & Kluge, 2010, S. 84 f., Hervorhebung im Original). Ziel der Typisierung, die im Kapitel 12 ausgeführt wird, ist die Erschließung von Zielvorstellungen, welche sich nicht direkt aus den theoretischen Vorüberlegungen ergeben haben, sondern im Datenkorpus versteckt sind. Strübing et al. (2018) formulieren es so:

„‘Offenheit‘ bedeutet, dass man ein empirisches Phänomen nicht einfach mit soziologischen Fragestellungen konfrontiert, sondern dass man von seiner Exploration das Auftauchen jener Fragen erwartet, mit denen es sich von innen erschließen lässt.“ (2018, S. 86)

Unerwartete Fragestellungen, die sich in der Beschreibung der Einzelfälle und deren Kontrastierung ergeben, immer mit Blick auf Diagnose, Förderung und Beurteilung von Singunterricht, können so wiederum strukturiert herausgearbeitet werden.

Neben den beiden spezifischen Methoden führt die strukturierende Inhaltsanalyse als Primärverfahren vor allem zu einer inhaltlichen Verdichtung der deduktiv gesetzten und induktiv erweiterten Kategorien. In den Forschungsfragen wird nach dem *Wie* gefragt, woraus sich eine „Paraphrasierung und sukzessive Verdichtung von Material“ (Schreier, 2014, S. 14) als sinnvoll erweist. Auch hier beschreibt Mayring (2015) die Art und Weise, wie die Verdichtung vor sich gehen soll (S. 69 ff.). Ziel ist eine zunehmende Abstraktion, wobei dies vor allem im Rahmen der induktiven Verfahren als wichtig erscheint, da bei deduktiven Verfahren in erster Linie Aussagen thematisch gesammelt werden. Im Kapitel 13 werden die Inhalte des Datenkorpus entlang der theoretisch erarbeiteten Inhalte dargestellt, zusammengefasst und kontrastiert. Im Rahmen dieser Analyse wird als weiteres Instrument der QIA die Explikation eingesetzt. „Zu einzelnen interpretationsbedürftigen Textstellen wird zusätzliches Material herangezogen, um die Textstelle zu erklären, verständlich zu machen, zu erläutern, zu explizieren.“ (S. 90) So kann eine erzählte Geschichte in einem Interview auf einer Ebene unklar bleiben, aber durch zusätzlichen Kontext aus Forschungsnotizen oder einem ähnlichen Beispiel aus der Literatur gut zu erklären sein. Dabei dient wiederum die Verbundenheit mit dem Feld als Garant für das Ausbleiben von impressionistisch anmutenden Ausführungen, wie Kracauer (1952) dies zu Beginn des Abschnitts als Kritik von der Inhaltsanalyse formuliert hat.

8.3 Interviews als Datenbasis

In Abschnitt 8.1 wurde als adäquate Form der Gewinnung von Datenmaterial die Arbeit mit Interviews festgelegt, ohne aber über die Interviewpartnerinnen und -partner sowie die geeignete Form der Interviews genauere Überlegungen anzustellen. Zudem wurde bei der Beschreibung der QIA klar (vgl. 8.2), dass das Datenmaterial bezüglich des Kommunikationszusammenhangs beschrieben werden muss, um Fehlinterpretationen zu vermeiden. Diese drei Aspekte bilden den Inhalt dieses Abschnitts.

Sample

Die Einschränkung auf Lehrkräfte, die Musikunterricht erteilen, liegt in der vorliegenden Forschungsarbeit auf der Hand. Weniger klar ist die Strategie der Auswahl von Musiklehrkräften. Patton (2015) schlägt vor, eine „maximum variation sampling strategy“ (S. 283) zu benutzen, um ein möglichst reichhaltiges Datenmaterial zu generieren. In der vorgeschlagenen Strategie werden Unterscheidungskriterien gesucht, die dazu führen, dass

„(1) high-quality, detailed descriptions of each case, which are useful for documenting uniquenesses and diversity, and (2) important shared patterns that cut across cases and derive their significance from having emerged out of heterogeneity“ (S. 283)

entstehen. Mit dieser Strategie soll erreicht werden, einerseits die gewünschten lokalen Singkulturen beschreiben zu können und andererseits verbindende Muster zu erkennen. Kelle und Kluge (2010) formulieren ihre Sampling-Strategie nicht aus der Perspektive von Reichhaltigkeit, sondern fordern Überlegungen zur Verzerrung bedeutsamer Merkmale: „Ausschlaggebend für die Güte einer Stichprobe in der quantitativen Forschung ist also nicht Repräsentativität bezogen auf alle denkbaren Merkmale, sondern die Abwesenheit von Verzerrungen in Bezug auf für die Forschungsfragestellung bedeutsamen [sic] Merkmale.“ (S. 41) Für die vorliegende Studie sind beide Argumentationsmuster spannend, weil ein reichhaltiges Material zum Singunterricht notwendig ist, um in der ganzen Breite Förderung und Beurteilung beschreiben zu können, aber auch entlang der wesentlichen Fragestellungen Lücken und Verzerrungen vermieden werden sollten.

Für die Auswahl des Stichprobenmaterials wird in theoretischer Hinsicht auf das Modell professioneller Handlungskompetenz von Baumert und Kunter (2006) zurückgegriffen, das in wegweisenden Studien (Baumert & Kunter, 2011; Blömeke, 2011) als Grundlage für die Darstellung des Professionswissens diente. Dabei wurde sowohl auf den Kompetenzbegriff von Weinert (2001) als auch auf die Arbeiten von Shulman (1986, 1987) und Bromme (1992) zu den Wissensdomänen zurückgegriffen. Weinert schlägt statt einer einheitlichen Definition des Kompetenzbegriffs eine Reihe von Merkmalen vor, die als Bedingungen für das Vorhandensein von Kompetenz dienen können, ist sich aber bewusst, dass er dabei der

praktischen Nützlichkeit mehr Gewicht gibt als der theoretischen Fundierung innerhalb der Sozialwissenschaften (Weinert, 2001, S. 63). Von Kompetenz spricht er, wenn

- die Voraussetzungen vorhanden sind, erfolgreich komplexen Anforderungen zu genügen,
- diese Anforderungen kognitive, motivationale, ethische, volitionale und soziale Komponenten umfassen,
- das Anforderungsniveau eine gewisse Komplexität übersteigt, also über das Ausüben automatisierter Fertigkeiten hinausgeht, im Wissen, dass dabei die Abgrenzung unscharf ist,
- und wenn viele Lernprozesse stattfinden müssen, die nicht direkt unterrichtet werden können.

Weiter schlägt er vor, Kern- von Metakompetenzen zu unterscheiden, wobei er Kernkompetenzen als ein Bündel von im täglichen Leben benötigten Anforderungen im Arbeitsleben und sozialen Umfeld definiert und von Metakompetenzen dann spricht, wenn über das Wissen der eigenen Kompetenzen gesprochen wird (S. 62f.¹¹⁷).

Für die Einteilung spezifischer Kompetenzen von Lehrkräften in verschiedene Wissensdomänen hat sich das Modell von Shulman (1987) auch in der Musikpädagogik als sinnvoll erwiesen. Puffer und Hofmann (2017) benutzen es z. B. im Kontext der Testung fachdidaktischer Wissensbestände. Shulman schlägt dabei folgende drei Bereiche als kognitive Komponenten des Professionswissens vor: Fachwissen (Content Knowledge – CK), Pädagogisches Wissen (Pedagogical Knowledge – PK) und Fachdidaktisches Wissen (Pedagogical Content Knowledge – PCK)¹¹⁸. CK bündelt die notwendigen fachspezifischen Wissensfacetten, im Fall von Musiklehrkräften sind das vor allem musikalische Fertigkeiten, sei dies am Instrument oder mit der Stimme, sowie bezogen auf Gehörbildung und Theorie. PK subsummiert allgemeine pädagogische Kenntnisse, wie z. B. den Umgang mit Jugendlichen und das Vorbereiten von Unterricht, und PCK sammelt die spezifischen methodischen und pädagogischen Kenntnisse, die wiederum für das Fach Musik notwendig sind. Im Modell

¹¹⁷ Die Liste der Merkmale wurde vom Autor zusammenfassend aus dem Englischen übersetzt.

¹¹⁸ Shulman schlägt weitere Bereiche vor, durchgesetzt haben sich aber die drei genannten.

von Baumert und Kunter (2006, S. 482) in Abbildung 8 werden nun nicht nur die Wissensdomänen nach Shulman dargestellt, sondern auch die Bereiche Überzeugungen und Werthaltungen, motivationale Orientierungen und selbstregulative Fähigkeiten.

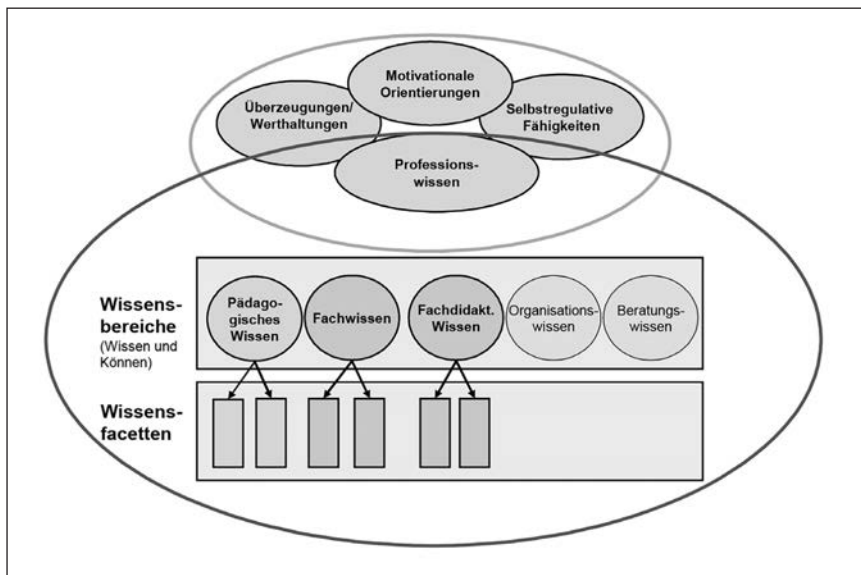


ABBILDUNG 8: Modell professioneller Handlungskompetenz – Professionswissen (Baumert & Kunter, 2006, S. 482)

Werden diese vielfältigen Aspekte auf die Merkmale von Kompetenz nach Weinert bezogen, ergibt sich die Möglichkeit, Eigenschaften von Lehrkräften zu definieren, nach denen bei Interview-Kandidierenden gesucht werden kann. Zusätzlich gilt es auch demografische Kriterien einzubeziehen, auf der einen Seite bezüglich des Geschlechts der Lehrkräfte und auf der anderen Seite bezüglich des Alters und der Erfahrung. Die konkrete Bildung von Profilen für das varianzmaximierte Sampling und die Beschreibung des Samples erfolgt später in Kapitel 9.

Form des Interviews

Nach der Frage, wer interviewt werden soll, gilt es nun zu klären, „wie und mit welcher Begründung das Sprechen (die Textproduktion) der inter-

viewten Person beeinflusst und gesteuert wird“ (Helfferrich, 2019, S. 669). Für die Beantwortung der vorliegenden Forschungsfragen bieten sich leitfadengestützte Interviews an. In diesen wird der Interviewablauf strukturiert und die Themen werden zulasten einer maximalen Offenheit, aber ausgerichtet auf den Forschungskontext vorgegeben. Helfferrich nennt folgende Formen von Stimuli für leitfadengestützte Interviews: „(Erzähl-)Aufforderungen, explizit vorformulierte Fragen, Stichworte für frei formulierbare Fragen und/oder Vereinbarungen für die Handhabung von dialogischer Interaktion für bestimmte Phasen des Interviews.“ (S. 670) Eine größere Offenheit wäre in narrativen Interviews gegeben, da „dem Befragten die Ausgestaltung der vereinbarten Interviewthematik weitgehend“ (Küsters, 2009) überlassen wird. Die von Helfferrich geforderten Erzählaufforderungen weisen eine gewisse Nähe zu narrativen Interviews auf, da sie die Aufgabe haben, „die Befragten zu komplexen, zusammenhängenden, von ihnen selbst gesteuerten Darstellungen an[zuh]regen, das heißt ‚Erzählungen‘ [zu] bieten“ (Gläser & Laudel, 2010, S. 116). Das leitfadengestützte Interview bewegt sich also zwischen einer narrativen Situation, die nach einem thematischen Input wenig Steuerungspotenzial aufweist, und einer Befragung mit vorgegebenen Antworten, die individuelle Themen und spezifische Antworten verunmöglichen.

Für die Planung der leitfadengestützten Interviews wird von Gläser und Laudel gefordert, dass der Leitfaden die Thematik in den Erfahrungshorizont der Interviewten übersetzen muss sowie eigenständige und spezifische Antworten ermöglicht, also über antizipierte Antworten hinausgeht (S. 116). Gefragt ist somit eine geschickte Konstruktion des Leitfadens, welche die Lehrkräfte dazu animiert, ihren eigenen Unterricht zu konturieren, ohne vom Forschenden thematisch zu stark eingeengt zu werden. Gläser und Laudel weisen auch auf die Gefahr hin, dass der direkte Hinweis auf die Untersuchungskonstrukte dazu führen kann, dass keine Informationen oder Fakten mitgeteilt werden, „mit denen man einen sozialen Prozess rekonstruieren kann, sondern im günstigsten Fall die Meinung des Interviewpartners über den Prozess und im ungünstigsten Fall lediglich seine Meinung über die Kausalzusammenhänge des Falls“ (S. 113). Die gefragte Übersetzungsleistung bedeutet für die vorliegende Arbeit, dass die Lehrkräfte z. B. nicht mit dem Konstrukt Assessment konfrontiert werden, da dieses Konzept in der Musikpädagogik kaum etabliert ist, sondern zu Förderung und Beurteilung befragt werden, da dieser Begriff klar konturiert

ist. Dabei muss darauf vertraut werden, dass die Lehrkräfte in der Lage sind, wichtige Unterrichtsaspekte im Rahmen professioneller Unterrichtswahrnehmung zu erkennen (Rieche et al., 2018, S. 3, vgl. noticing). Der verwendete Leitfaden für die vorliegende Arbeit wird in Kapitel 9.2 argumentiert.

Kommunikationszusammenhang

Im Gegensatz zu visuell erfassten Daten in Form einer Videografie oder Forschungsnotizen einer teilnehmenden Beobachtung sind Interviews als direkte Kommunikation konzipiert. Von der Suche der Lehrkräfte für die Interviews bis zur abschließenden Verifizierung von entstandenen Unklarheiten findet der Austausch von Informationen, die Art der Fragestellung und die Reaktion auf Antworten im Rahmen einer Kommunikationssituation statt. Mayring und Fenzl (2019) bieten deshalb für die inhaltsanalytische Forschungsarbeit den Rückbezug auf ein Kommunikationsmodell

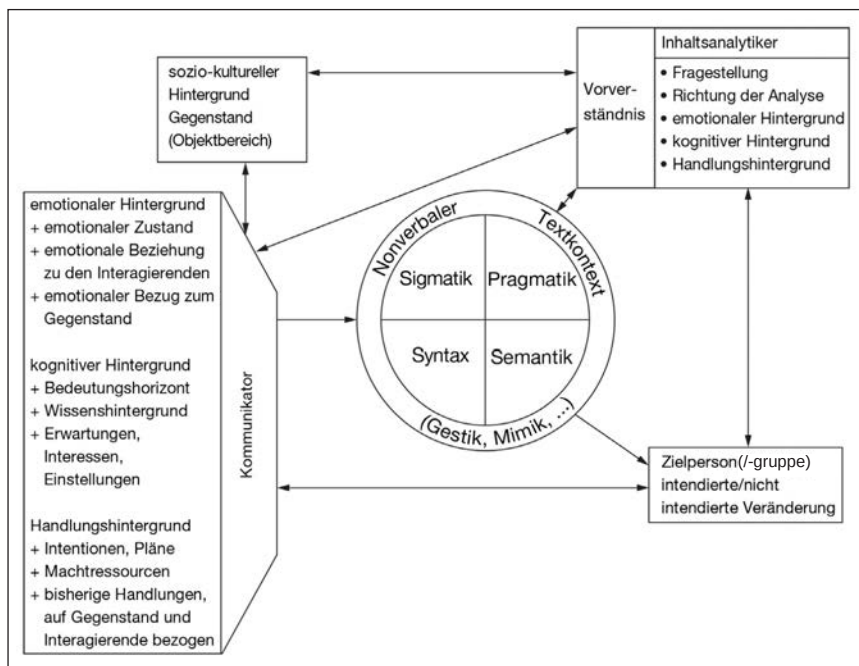


ABBILDUNG 9: Kommunikationsmodell nach Mayring (2015, S. 59)

an, das Mayring (2015, S. 59) aufbauend auf eine Arbeit von Lagerberg (1975) weiterentwickelte und publizierte. Dieses Modell (vgl. Abbildung 9) dient in der Folge als Grundlage für die Einbettung der Interviews in einen Kommunikationszusammenhang. Im Mittelpunkt des Kommunikationsmodells steht das Produkt der Kommunikation, das bei Interviews meist in Form eines Audio-Files bzw. einer im Anschluss erstellten Transkription vorliegt. Die Einordnung dieser vorliegenden Daten ist notwendig, um den Einfluss der verschiedenen Ebenen der Kommunikation sichtbar zu machen. Der manifeste Inhalt, also die Worte und Sätze, die gesprochen und anschließend transkribiert werden, werden während des Interviews begleitet von nonverbaler Kommunikation, wie Gestik und Mimik, die den manifesten Inhalt möglicherweise erklären, umdeuten oder erweitern. Grund dafür sind sowohl der emotionale und kognitive Hintergrund der Interviewten als auch Intentionen und Erwartungen aufgrund bestehender Beziehungen, sowie das soziokulturelle Umfeld. In der vorliegenden Studie ist der Empfänger der Kommunikation und Inhaltsanalytiker dieselbe Person, was die Position dieser Person nochmals in den Fokus rückt, insbesondere weil der Forschende selbst lange Zeit innerhalb des zu untersuchenden Feldes tätig war.

Auch Gläser und Laudel (2010) sowie Helfferich (2019) betonen die Wichtigkeit der Deklaration der Kommunikationssituation im Rahmen der Durchführung von Interviews. Im Gegensatz zu Gläser und Laudel, die den persönlichen Bezug zu den Interviewpartnerinnen und -partnern als kritisch erachten, betonen Strübing et al. (2018), dass der Kontakt zum Feld maßgeblich dafür verantwortlich ist, „welche Auskünfte Informanten [...] bereit sind, über sich preiszugeben“ (S. 88). Trotz der Deklaration verschiedener Einflussmöglichkeiten gehen alle kommunikationstheoretischen Sende-Empfänger-Modelle von der Erkenntnis aus, „dass die Botschaft nicht immer korrekt beim Rezipienten bzw. bei der Rezipientin ankommen muss“ (Röhner & Schütz, 2016, S. 37).

Helfferich (2019) schlägt aus diesem Grund vor, die Rollen in Interviews bewusst zu gestalten, da die Interviewten dieses „Setting, die eigene Rolle und ihr Verhältnis zu der interviewenden Person bzw. zur Wissenschaft“ (S. 670) deuten und entsprechend reagieren. So beschreiben Jeding und Michael (2019) verzerrende Effekte, die im Rahmen der Durchführung von Interviews entstehen können. Die Anwesenheit an sich, das Verhalten und die Erscheinung beeinflussen die Interviewten. Sie

schreiben: „Insgesamt legt der Forschungsstand nahe, dass die Ursachen von Interviewereffekten in der dynamischen Interaktion von Interviewern, Befragten und Befragungsthema liegen.“ (S. 370) Hlawatsch und Krickl (2019) beschreiben einen möglicherweise verzerrenden Effekt auf die Daten von Befragungen aufgrund der Einstellung der Befragten. Einerseits wird die Frage der Kooperationsbereitschaft aufgeführt, die aber bei Interviews, welche auf Freiwilligkeit beruhen, kaum von Bedeutung ist. Andererseits wird angeführt, dass die Teilnehmenden möglicherweise ihre Antworten in der Befragungssituation „in Abhängigkeit von Normen und Erwartungen“ (S. 358) so anpassen, wie sie es „in der Befragungssituation für sozial wünschenswert erachten“ (S. 358).

Die Einbettung in den Kommunikationszusammenhang kann entlang des Kommunikationsmodells von Mayring erfolgen und beinhaltet für die vorliegende Studie die Deklaration der Beziehung zu den Interviewten sowie die Darstellung der soziokulturellen Situation bzw. das schulische Umfeld. Weiter müssen die Interviewsituation und mögliche verzerrende Effekte besprochen werden. Dies geschieht später in Kapitel 9.

8.4 Gütekriterien der QIA

Die Frage der Güte wurde zu Beginn des Kapitels aus zwei Perspektiven in allgemeiner Form erörtert. Sowohl die Prinzipien der Wissensproduktion nach Gläser und Laudel (2010) als auch die Gütekriterien qualitativer Sozialforschung nach Strübing et al. (2018) beschreiben Aspekte der Güte, die es zu berücksichtigen gilt. Auf diese wurde bereits in den vorangegangenen Abschnitten mehrfach implizit und explizit Bezug genommen, immer mit dem Ziel der intersubjektiven Nachvollziehbarkeit (Mayring, 2015, S. 13) der vorliegenden Arbeit. Es stellt sich deshalb die Frage, wie sinnvoll ein eigener Abschnitt zu Gütekriterien in diesem Kapitel ist und auf welche Bereiche spezifisch eingegangen werden muss.

Mayring (2015) stellt für inhaltsanalytische Studien fest, dass „fast vollständig Angaben über die Zuverlässigkeit (Reliabilität) und Gültigkeit (Validität) der erzielten Ergebnisse“ (S. 123) fehlen. Mayring bezieht seine Aussage auf die klassischen Gütekriterien quantitativer Forschung (Krebs & Menold, 2019), die in der Inhaltsanalyse aber bezüglich ihrer Übertragbarkeit auf qualitative Forschung kritisch hinterfragt werden. Zusätzlich

zu der bereits erfolgten Einordnung der Forschungsschritte in die Kriterien nach Gläser und Laudel bzw. Strübing et al. wird die Kritik von Mayring als Anlass genommen, Reliabilität und Validität als Gütekriterien für die vorliegende Arbeit zu diskutieren¹¹⁹.

Reliabilität

Reliabilität bezieht sich auf die Zuverlässigkeit der Messung, die in der QIA hauptsächlich in der Erstellung des Messinstruments (Kategoriensystem) und der anschließenden Kodierung eine Rolle spielt. Zudem ist Reliabilität die Grundlage der Validität, also der Gültigkeit der Messung. Flick (2019) bewertet die klassischen Testverfahren zur Überprüfung von Reliabilität (z. B. Re-Test oder Paralleltest) als problematisch, weil sowohl die Analyse von sprachlichem Material als auch die Größe der Stichprobe eine Vergleichbarkeit erschweren (S. 124). Für die QIA wird als Alternative die Intercoder-Reliabilitätsprüfung vorgeschlagen, in der dieselben Texte mit dem bestehenden Messgerät (Kategoriensystem) auf der Basis des Kodierhandbuchs von einer weiteren Person kodiert werden, um anschließend die Kodierungen zu vergleichen. Mit diesem Vergleich zeigt sich gleichzeitig, ob das Kodierhandbuch verständlich war und zur gleichen Auswahl von Inhalten führt.

Das Resultat wird im Wert Kappa κ_n ¹²⁰ ausgedrückt und als Güte Merkmal einer QIA genannt. Flick schreibt aber: „Eigentlich wird mit diesem Vorgehen Objektivität gemessen, als [sic] die Unabhängigkeit der Ergebnisse von der untersuchenden Person.“ (S. 124) Laut Flick wäre Reliabilität dann gegeben, wenn das Material von der gleichen Person ein zweites Mal kodiert wird, ohne die bereits gesetzten Kodierungen zu kennen. Dieser Interrater-Reliabilitätstest wird selten eingesetzt, ist aber als Baustein in der Entwicklung des Kategoriensystems ansatzweise enthalten (vgl. 8.2). Flick weist zudem auf grundsätzliche Kritik an Reliabilitätsmessungen

¹¹⁹ Ein weiterer Grund leitet sich aus dem Kapitel 5.1 ab, in dem explizit Selbst-Assessment und Reflexion (AAL) als wichtiger Faktor des eigenen Lernens gefordert wird. Der Abschnitt 8.4 nimmt nun im Sinne von *visible learning* (Hattie, 2012) auf ausgewählte Aspekte von Güte Bezug.

¹²⁰ Cohens Kappa ist die Basis verschiedener Berechnungen von Kappa κ_n (Gwet, 2014, S. 22).

hin, da durch die individuelle Interpretation beim Kodieren in komplexen Kategoriensystemen eine große Übereinstimmung nur schwer erreicht ist. Allerdings kann durch einen anschließenden Verständigungsprozess über die unterschiedlich gesetzten Kodierungen das Kategoriensystems auch an Qualität gewinnen (Mayring, 2015, S. 53).

Die Unterschiede individueller Interpretationen sind aber nicht die einzigen Kritikpunkte, die angebracht werden können. So zeigen sich gerade bei der Benutzung der Software MAXQDA zwei Punkte, die angesprochen werden müssen. Die Grundlage der Berechnung des Werts Kappa κ_n basiert auf den Zahlen übereinstimmender und nicht übereinstimmender Kodierungen. Durch die Kodierung von Sinnzusammenhängen ergeben sich trotz des Einsatzes eines Kodierleitfadens Differenzen in der Länge des kodierten Texts. Die Software entscheidet dann aufgrund der prozentualen Überlappung, ob von einer Übereinstimmung ausgegangen werden kann oder nicht. In MAXQDA kann dieser Wert vorgegeben werden. Diese Stellschraube hat einen wesentlichen Einfluss auf den erreichten Wert von Kappa κ_n . Zudem ist die Berechnungsformel von Kappa κ_n nicht unumstritten. So legen Wirz und Kutschmann (2007) dar, dass die Interpretation von Kappa κ_n dann zu Schwierigkeiten führt, wenn einzelne Codes sehr wenig benutzt werden. In der Folge führen tiefe Kodierhäufigkeiten zu verfälschten Resultaten. In MAXQDA wird deshalb eine angepasste Berechnungsmethode nach einem Vorschlag von Brennan und Prediger (1981) benutzt, die diesen Umstand ausgleicht. Trotzdem ist in quantitativen Analysen bei kleinen Kodierhäufigkeiten mit Vorsicht auf den Wert Kappa κ_n abzustützen.

Landis und Koch (1977)			Altman, DG (1991)			Fleiss et al. (2003)		
0.81 – 1.00	excellent	exzellent	0.81 – 1.00	very good	sehr gut	0.75 – 1.00	very good	sehr gut
0.61 – 0.80	substantial	erheblich	0.61 – 0.80	good	gut	0.41 – 0.75	fair to good	mittel- mäßig bis gut
0.41 – 0.60	moderate	mäßig	0.41 – 0.60	moderate	mäßig			
0.21 – 0.40	fair	mittel- mäßig	0.21 – 0.40	fair	mittel- mäßig	< 0.40	poor	schlecht
0.00 – 0.20	slight	leicht	< 0.20	poor	schlecht			
< 0.00	poor	schlechte						

TABELLE 12: Populäre Referenzlevel von Cohens Kappa (Nurjannah & Siwi, 2017; Xie, 2013)

Als weiter Punkt zeigen sich bezüglich der Referenzlevel von Kappa κ_n unterschiedliche Modelle, die auf sprachlicher Basis zu unterschiedlichen Leveln von Güte führen. In Tabelle 12 werden gebräuchliche Referenzlevel dargestellt (zitiert nach Nurjannah & Siwi, 2017; Xie, 2013)¹²¹. Die von Landis und Koch (1977) empfohlene Einteilung in fünf Referenzlevel wurde später von Altmann (1991) übernommen und sprachlich im oberen Bereich leicht umformuliert. Fleiss et al. (2003) verkürzten auf drei Referenzlevel und korrigierten zusätzlich die Werte minimal.

Zusammenfassend zeigen sich nicht nur individuelle Prozesse des Verstehens als Hürde für den Güteaspekt der Reliabilität, sondern auch Vorgaben der Software und der Gebrauch der Sprache. Für die vorliegende Arbeit werden deshalb in Kapitel 10 nicht nur die Arbeitsschritte mit der Software MAXQDA transparent gemacht, sondern es wird auch die Inter-coder-Reliabilitätsprüfung detailliert besprochen. Auf der sprachlichen Ebene werden die Referenzlevel von Landis und Koch präferiert, da *erheblich (substantial)* und *exzellent* deutlicher konturiert sind als *gut* und *sehr gut*.

Validität

Validität im klassischen Sinn wird in der QIA aufgrund der starken gegenseitigen Beeinflussung von Theorie und Empirie (vgl. Strübing, 2018) kritisiert. Der zirkuläre Bezug (Mayring, 2015, S. 125) und die Offenheit der Variablen (Kategorien) erschweren die gewünschte Unabhängigkeit der Empirie. In der QIA werden deshalb erweiterte Qualitätsvorkehrungen getroffen, um valide Resultate zu generieren. Mayring (2015) bezieht sich dazu auf Krippendorff (1980, 2004, 2019), der ähnlich wie Mayring selbst die *Content Analysis* im Lauf der letzten Jahrzehnte stetig weiterentwickelt hat. Für den Bereich der Validität schlägt dieser eine Orientierung am Material, dem Ergebnis und dem Prozess vor (Mayring, 2015, S. 126 ff.):

- Die Gültigkeit des Materials zeigt sich in der Auswahl des Samples als Stichprobengültigkeit und in „der Angemessenheit der Kategoriendefinitionen (Definitionen, Ankerbeispiele, Kodierregeln)“ (S. 126) als

¹²¹ Die Tabelle wurde leicht umgestaltet und ergänzt mit einer deutschen Übersetzung.

semantische Gültigkeit. Die Auswahl des Samples wurde bereits erörtert (vgl. 8.3) und wird nochmals vertieft in Kapitel 9, die Entwicklung des Kategoriensystems und des geforderten Kodierhandbuchs in 10.2 ausgeführt.

- Die Gültigkeit des Ergebnisses weist Krippendorff dem Vergleich mit Untersuchungen ähnlicher Art oder Untersuchungen am selben Material mit anderen Methoden (Korrelative Gültigkeit) sowie durch sich erfüllende Prognosen (Vorhersagegültigkeit) zu. Beide Aspekte sind in der vorliegenden Arbeit schwer erreichbar, erstens weil es keine Untersuchungen zu dieser Thematik gibt, und zweitens weil es kaum Untersuchungen gibt, die den Singunterricht in der Sekundarstufe 1 betreffen. Ausnahmen sind die Arbeiten von Blanchard (2019) und Tellisch (2015), die sich am Rand mit der Thematik beschäftigen und bereits referiert wurden.
- Die Gültigkeit des Prozesses wird erreicht durch bereits durchgeführte erfolgreiche Anwendungen des Messinstruments, was im vorliegenden Fall nicht möglich ist. Die ebenfalls genannten „Erfahrungen mit dem Kontext“ (Mayring, 2015, S. 126 ff.) wurde bereits mit dem autoethnografischen Exkurs (vgl. Kapitel 3) belegt und der Anschluss an „etablierte Theorien und Modelle“ (S. 127) muss sich in der Analyse noch zeigen.

Aspekte der Validität können in dieser QIA an verschiedenen Orten verortet werden¹²² und müssen sich aber nach Strübing et al. (2018) auch in der Performance des gesamten Textes zeigen.

Das Kapitel zum Forschungsdesign wird abgeschlossen mit einer Aussage von Schlömerkemper (2010) zu den Möglichkeiten und Grenzen des Erkennens. Er meint, „dass es um eine Balance geht zwischen einer eher skeptischen Sicht auf die Grenzen des Erkennens und einer dennoch erlaubten Hoffnung auf Ergebnisse, die theoretisch befriedigend und/oder praktisch hilfreich sein können“ (S. 29). Die Beschreibung des For-

¹²² Mayring (2015) ebenso wie Flick (2019) sprechend auch von kommunikativer Validierung, in der die Forschungsergebnisse mit den befragten gemeinsam erörtert werden. Dies wurde für die vorliegende Studie nicht in Betracht gezogen, weil die Perspektive Assessment für die Interviewpartnerinnen und -partner unbekannt war und wohl noch ist. Eine kommunikative Validierung wäre deshalb wegen der fehlenden Basis schwierig geworden.

schungsplans und der notwendigen Schritte zur Qualitätssicherung bezwecken eine Plausibilisierung der vorliegenden qualitativen Interviewstudie, wie auch die folgenden Kapitel, welche die Möglichkeiten und Begrenzungen des Datenmaterials (vgl. Kapitel 9) sowie des Kategoriensystems (vgl. Kapitel 10) evaluieren.

9 Untersuchungsmaterial

Die Art der Interviews und die Vorgaben für den Leitfaden wurde bereits im vorigen Kapitel beschrieben (vgl. 8.3). In diesem Kapitel wird nun zuerst die Auswahl des Samples und die Erstellung des Leitfadens für die Interviews hergeleitet, um anschließend im Sinne einer Quellenkunde (Mayring, 2015, S. 54) die Umstände der Interviews und deren Kommunikationszusammenhang zu beschreiben sowie die Überführung der Audio-Dokumente in für die QIA verwertbare Transkripte zu erläutern. Der Aufbau des Kapitels folgt der Logik der in Kapitel 8.2 beschriebenen Schritte 2 & 3 der QIA.

9.1 Auswahlverfahren für das Sample

Für die Auswahl des Samples wurde im letzten Kapitel auf das Modell professioneller Handlungskompetenz nach Baumert und Kunter (2006) Bezug genommen. Dabei spielen neben den Wissensbereichen CK, PK und PCK nach Shulman (1987) auch Überzeugungen, Werthaltungen und motivationale Orientierungen eine Rolle. Auf der Grundlage dieses Modells und weiterer demografischer Merkmale werden in diesem Kapitel Merkmale zur Auswahl des Samples definiert.

Der Umgang mit Jugendlichen in der Sekundarstufe 1 mitten in der Pubertät ist eine komplexe Anforderung, die Wissen (CK) zur Stimmentwicklung ebenso voraussetzt wie pädagogisches (PK) und fachdidaktisches Wissen (PCK). Die dazu notwendigen Wissensbestände für die adoleszente Stimmentwicklung und die Assessment-Elemente wurden im ersten Teil der Arbeit vertieft und weisen nun darauf hin, die Profile entlang verschiedener Ausbildungswege zu gestalten, dies im Wissen um unterschiedliche Zugänge zu CK, PK und PCK. Zudem zeigen sich die spezifischen Gegebenheiten der schweizerischen Ausbildungs- und Anstellungssituation (vgl. Kapitel 6) als Herausforderung für die Profilbildung, denn in der Praxis findet sich eine heterogene Gruppe von Musiklehrkräften mit unterschiedlichen Berufslaufbahnen, oft nicht legitimiert durch eine Ausbildung, sondern durch die Chance einer Anstellung und dem nachfolgenden

Erfolg beim Unterrichten. Unter Berücksichtigung dieser komplexen Voraussetzungen werden vier Profile von Musiklehrkräften definiert, indem ausgehend von unterschiedlichen Ausprägungen von CK, PK und PCK folgende Gruppen definiert werden:

- **Fachfremd Unterrichtende:** Fachliche (CK) und fachdidaktische Kenntnisse (PCK) sind nicht aufgrund einer Ausbildung zu vermuten, da es an einer schulmusikalischen Ausbildung für die Sekundarstufe 1 fehlt.
- **Generalistisch ausgebildete Lehrerinnen und Lehrer:** Diese Lehrkräfte unterrichten meist alle Fächer in ihrer eigenen Klasse, sei es in der Primarstufe oder in der Sekundarstufe 1. Sie haben eine fundierte pädagogische Grundausbildung (PK) im Rahmen einer Lehrerseminarausbildung erhalten, dabei sind ihnen jedoch nur elementare fachliche (CK) und fachdidaktische Kenntnisse (PCK) für das Fach Musik vermittelt worden.
- **(Fach-)Lehrkräfte:** Diese Personen werden in der Regel Schulmusikerin bzw. Schulmusiker genannt und weisen eine vertiefte Ausbildung für das Fach Musik in der Sekundarstufe I oder II vor. Für diese Lehrkräfte wird im Stichprobenplan vorausgesetzt, dass fachliche (CP) und fachdidaktische Kenntnisse (PCK) sehr gut entwickelt sind. Traditionellerweise haben diese Lehrkräfte eine eher basale pädagogische Grundbildung (PK).
- **Gesangslehrerinnen und -lehrer:** Bei dieser Gruppe von Lehrkräften wird zusätzlich zur Gesangsausbildung eine Ausbildung für das Fach Musik vorausgesetzt, wobei die Hoffnung besteht, auf exzellente Kenntnisse zur Förderung der Stimme zu stoßen.

Die von Weinert (2001) angemahnten sozialen Komponenten professioneller Kompetenz sind beim gemeinsamen Singen (vgl. 4.5) von zentraler Bedeutung und führen zu einem weiteren Merkmal für die Auswahl von Lehrkräften ins Sample. So setzt z. B. der Umgang mit der Mutation von Jungen im sozialen Geflecht einer Klasse eine hohe soziale Kompetenz oder Empathie voraus und ist somit zumindest im Vergleich zum gemeinsamen Durchsingen eines Liedes in der Klasse anspruchsvoller. Dies kann nach Weinert einer Kompetenz zugeordnet werden, die nicht direkt lernbar ist, aber viel Gelerntes voraussetzt. Da diese Kompetenzen nicht direkt messbar sind für die Auswahl des Samples, wurde eine minimale Erfahrung von fünf Jahren als Voraussetzung für die Teilnahme an der Studie

definiert. Lehrkräfte mit weniger Berufserfahrung wurden ausgeschlossen, um einerseits Personen, die nach kurzer Zeit den Schuldienst quittieren (SKBF, 2014, 2018)¹²³, zu eliminieren und andererseits Personen mit wenig Erfolg im Fach Musik auszuschließen, da davon ausgegangen werden kann, dass diese den Unterricht in diesem Fach eher vermeiden oder nicht an der Studie teilnehmen würden.

Als weiteres Kriterium außerhalb des obigen Modells wird das Geschlecht der Lehrkraft definiert, einerseits um Hinweise zu den geschlechtsspezifischen Verhaltensweisen (vgl. 4.5) abzubilden und andererseits, um die Verteilung der Geschlechter unter den Lehrkräften abzubilden. Ebenfalls nicht in diesem Modell berücksichtigt sind die verschiedenen Leistungsniveaus der Schülerinnen und Schüler, die in der Schule unterrichtet werden. Diesbezüglich könnte vermutet werden, dass in höheren Niveaus ein höheres Maß an CK notwendig ist, in tieferen Niveaus mehr PCK und PC. Aussagen zu diesem Merkmal und weiteren Dimensionen, wie verschiedene Schulformen oder Eigenheiten der Schulhauskultur, werden zwar später in den Interviews erwartet, finden aber ihren Niederschlag nicht in der Auswahl der Teilnehmenden.

Für die vorliegende Untersuchung wird in Abschnitt 9.3 ein qualitativer Stichprobenplan erstellt, um ein möglichst breites Spektrum der heterogenen Lehrerschaft abzubilden. Die Matrix enthält die beiden Achsen *Profil der Musiklehrkraft* und *Geschlecht* sowie das Ausschlusskriterium *Erfahrung*. Dies ergibt acht verschiedene Positionen, für die je zwei Interviews geplant wurden, was einerseits genügend Material verspricht und auch den Aspekt der zeitlichen Bewältigung berücksichtigt (Gläser & Laudel, 2010, S. 116 f.). Bevor die Planung und Durchführung der Interviews beschrieben wird, folgen Angaben zur Erstellung des Interviewleitfadens.

¹²³ Entgegen der in Medien verbreiteten Meinung, dass viele frisch ausgebildete Lehrkräfte in der Schweiz wieder aussteigen (z. B. Kunz, 2015) wird im Bildungsbericht Schweiz 2014 (SKBF, 2014) gezeigt, dass nach einem Jahr knapp 90 % den Einstieg in den Lehrberuf in der Sekundarstufe 1 geschafft haben. Nach 5 Jahren wurde in einer Zweitbefragung deutlich, dass ca. 85 % nach wie vor im Lehrberuf tätig sind. Zu der Ausstiegsquote von Musiklehrkräften gibt es keine Zahlen.

9.2 Erstellung des Interviewleitfadens

In Kapitel 8.3 wurden die Voraussetzungen für den Interviewleitfaden erarbeitet, wobei sich hauptsächlich zwei Elemente als zentral erwiesen. Einerseits muss es gelingen, den Befragten authentische Erzählungen zu entlocken, nahe genug an den Forschungsfragen, aber trotzdem offen für die spezifischen Gegebenheiten. Den Musiklehrkräften soll mit einer Erzählaufforderung eine Plattform gegeben werden, ihren Singunterricht möglichst offen und umfassend zu beschreiben. Andererseits ist es nicht zielführend, die Teilnehmenden direkt mit dem Forschungskonstrukt zu konfrontieren, um zu vermeiden, lediglich „die Ansicht eines soziologischen Laien“ (Gläser & Laudel, 2010, S. 113) zu erhalten. Dies gilt im Besonderen für die Perspektive von Assessment, welche in dieser Arbeit an den Datenkorpus herangetragen wird.

Der Leitfaden wurde deshalb aus der Perspektive einer Lehrkraft mit Blick auf die Thematik Diagnose, Förderung und Beurteilung entwickelt. Entstanden ist ein Set von Erzählaufforderungen und Stichworten, das die notwendige Übersetzungsleistung von der Forschung in die Praxis des Musikunterrichts gewährleistet (S. 113). Der Leitfaden liegt ausformuliert und als Mindmap digital vor (siehe Anhang). An dieser Stelle wird er zusammenfassend dargestellt.

In der Einladung für die Studie und während der Einleitungsphase im Interview wird die Thematik Förderung der jugendlichen Stimme kurz erwähnt, aber nicht ausgeführt. In einer ersten Erzählaufforderung wird nach typischen Singsituationen gefragt, um eine möglichst zusammenhängende narrative Phase einzuleiten. Mit dem Fokus auf typische Singsituationen wird zwar nicht direkt nach Förderung und Beurteilung gefragt, aber intendiert, zentrale Ideen und Ziele der Lehrkräfte in den Blick zu nehmen. Ausgehend davon wird die Frage der Förderung mit den Stichworten Aufgabenstellungen und Unterrichtsformen vertieft, wobei jeweils nach motivierenden Unterrichtselementen gefragt wird. Zusätzlich wird mit dem Stichwort Feedback nach fördernden Interaktionsformen gefragt und mit dem Stichwort Zielvorstellungen zu erreichende Kompetenzziele eruiert. Gegen Ende des Interviews wird eine Reflexion über methodische Veränderungen des Unterrichts angeregt, um zum Schluss nochmals eine Erzählphase zu besonderen Highlights zu evozieren. Abschließend stehen im Interview biografische Elemente, wie Ausbildung und außerschulische

künstlerische Tätigkeiten im Zentrum, wenn diese noch nicht im Verlauf des Gesprächs benannt werden. Der Leitfaden wurde nach dem ersten Interview mit Herrn Brugger (01) finalisiert.

9.3 Stichprobenplan und Durchführung der Interviews

Aufbauend auf dem persönlichen Beziehungsnetz¹²⁴ des Forschenden wurden für die vier erstellten Profile Teilnehmende ausgewählt und per E-Mail zu dieser Studie eingeladen. Bei elf angeschriebenen Personen erfolgte eine positive Rückmeldung und es konnte ein Interview geplant und durchgeführt werden, bei vier Anfragen erfolgte keine Reaktion. Die Suche nach weiblichen Lehrkräften hat sich als Herausforderung erwiesen, da sich in der Planungsphase keine weibliche, fachfremd unterrichtende Lehrkraft finden ließ. Die Tabelle 13 zeigt die Anzahl der geplanten und durchgeführten Interviews in der Übersicht.

Die folgende Tabelle 14 zeigt ausgesuchte demografische Daten, wie das Alter und das Geschlecht der Teilnehmenden, die Erfahrung als

	weiblich		männlich		Summe Σ	
	geplant	durchgeführt	geplant	durchgeführt	geplant	durchgeführt
Fachfremd Unterrichtende	0	0	2	2	2	2
Generalistisch ausgebildete Lehrkräfte	1	0	2	2	3	2
Schulmusiklehrpersonen						
– ohne Gesangsausbildung	2	1	3	2	5	3
– mit Gesangsausbildung	3	3	2	1	5	4
Summe Σ	6	4	9	7	15	11

TABELLE 13: Stichprobenplan mit der Anzahl der geplanten und durchgeführten Interviews

¹²⁴ Einige Teilnehmende konnten aufgrund der langjährigen Unterrichts- und Verbandstätigkeit gefunden werden. Weitere Kontakte ergaben sich aus der Aus- und Weiterbildung an der FHNW und aufgrund von Empfehlungen aus dem Arbeitsteam des Forschenden.

Musiklehrkraft in Anzahl Jahren sowie Angaben zu den unterrichteten Leistungsniveaus und der Schule.

Die Interviews wurden in der Reihenfolge ihrer Durchführung nummeriert. Die Namen in der Tabelle entsprechen einem Pseudonym¹²⁵. Der Mittelwert des Alters ist mit 47,8 ($M = 47,8$) relativ hoch, wobei eine Gruppe von drei Teilnehmenden zwischen 27 und 35 Jahren am Anfang ihrer Berufslaufbahn steht und die restlichen Teilnehmenden älter als 46 Jahre sind und sich in der zweiten Hälfte oder am Ende ihrer Laufbahn als Musiklehrkraft befinden. Eine Person sticht mit 67 Jahren heraus. Sie befindet sich zum Zeitpunkt des Interviews bereits im Pensionsalter, ist aber nach wie vor mit einem Teilpensum als Musiklehrkraft tätig. Mit einem Mittelwert von 16,5 Jahren ($M = 16,5$) kann insgesamt von einer großen Erfahrung gesprochen werden. Interessant sind die beiden Personen mit einer Erfahrung von 30 bzw. 31 Berufsjahren, da sie während ihrer ganzen

Interview-Nummer & Pseudonym	Geschlecht	Alter	Erfahrung	Leistungsniveau			Milieu
				hoch	mittel	tief	
01 Hr. Brugger	m	52	22	h			städtisch
02 Hr. Buchser	m	55	12		m	t	städtisch
03 Fr. Fricker	w	49	16	h			ländlich
04 Hr. Hilfiker	m	35	6		m	t	ländlich
05 Fr. Basler	w	47	14	h126	m	t	städtisch
06 Fr. von Rohr	w	37	10	h			städtisch
07 Hr. Rothenburg	m	49	23		m		ländlich
08 Hr. Neuendorf	m	67	30	h	m	t	ländlich
09 Hr. Berner	m	50	10		m	t	ländlich
10 Hr. Leimbacher	m	57	31	h			ländlich
11 Fr. Derendinger	w	28	7	h	m	t	ländlich
Anzahl	4♀/7♂						3 städtisch 7 ländlich
Mittelwert M		47,8	16,5				

TABELLE 14: *Liste der Interviews mit demografischen Angaben, Erfahrung in Jahren, Informationen zum Leistungsniveau und zum Milieu*

¹²⁵ Den Namen werden in der Arbeit je nach Geschlecht Herr und Frau vorangestellt, um erstens eine persönliche Atmosphäre zu kreieren und zweitens Lehrerinnen und Lehrern Wertschätzung entgegenzubringen.

Berufslaufbahn jungen Menschen das Singen beigebracht haben und dies nach wie vor mit Begeisterung tun. Bezüglich der unterrichteten Leistungsniveaus gibt es drei Teilnehmende, die alle Niveaus unterrichten¹²⁶, vier auf dem progymnasialen¹²⁷ Leistungsniveau Unterrichtende und vier Lehrkräfte, die nur auf den beiden unteren Niveaus unterrichten.

Alle Interviews wurden in der Nordwestschweiz durchgeführt, wobei drei im städtischen und acht im ländlichen Raum verortet werden können. Dieser Umstand wird zugunsten einer transparenten Beschreibung des Samples erwähnt, bleibt aber im Verlauf der QIA ohne Relevanz. Die meisten ländlichen Schulorte, die in der Studie vorkommen, liegen in der Nähe einer kleineren Stadt, also in der Agglomeration, was infrage stellt, ob „die Ausdifferenzierung der sozialen und gesellschaftlichen Strukturiertheit [...] zu einer Vielfalt der Orientierungs-, Einstellungs- und Handlungsmuster von Individuen und sozialen Gruppen“ (O. Frey, 2012, S. 503) führt. Im vorliegenden Forschungskontext wird deshalb auf die Analyse von Milieus verzichtet.

Die Durchführung der Interviews fand in der zweiten Jahreshälfte 2017 und im Februar 2018 statt. Die Interviews wurden mehrheitlich in der Schule (Lehrerzimmer oder Unterrichtsraum) durchgeführt, in drei Fällen in der Privatwohnung und in einem Fall in einer Gaststätte. Die Dauer der Interviews war im Schnitt 46,5 Minuten ($M = 46,5$; siehe Tabelle 19, S. 235), wobei das längste 58 Minuten und das kürzeste 32 Minuten dauerte. Die Umstände der Interviews wurden protokolliert.

Die Einordnung in den Kommunikationszusammenhang erfolgt entlang der erarbeiteten Kriterien (vgl. 8.3) zuerst bezüglich der konkreten Interviewsituation und anschließend bezüglich der Beziehung zu den Teilnehmenden der Studie. Die Atmosphäre während der Interviews war durchgehend angenehm, abgesehen von kleinen Irritationen bei einigen vertiefenden Fragen¹²⁸. Speziell erwähnenswert erscheint das Interview mit Herrn Hilfiker (04), in dem thematisch Bezug auf eine vom Forschenden

¹²⁶ Das Interview von Frau Basler (05) setzt sich an dieser Stelle ab, weil das Konzept der Waldorfschule keine Leistungsdifferenzierung in der Sekundarstufe 1 vorsieht.

¹²⁷ Siehe Fußnote 106, S. 172.

¹²⁸ Der Forschende hat es in den Interviews vermieden, Fragen so zu stellen, dass sie als Kritik aufgefasst werden konnten, was aber nicht durchgehend gelang.

durchgeführte Weiterbildung genommen wurde. Dabei berichtet der Lehrer von erfolgreich umgesetzten Elementen. Insgesamt scheint die sozial erwünschte Norm (Hlawatsch & Krickl, 2019, S. 358) in den Interviews aber nicht zu geschönten Antworten geführt zu haben. Die im Kommunikationszusammenhang ebenfalls wichtigen nonverbalen Elemente beschränkten sich auf akustisch wahrnehmbare Elemente der Sprachmelodie und sind deshalb an dieser Stelle vernachlässigbar.

Auch bezüglich der Beziehungsebene können keine Komplikationen ausgemacht werden. Die Teilnehmenden der Studie nahmen den Forschenden als Musiklehrer wahr, womit während der Interviews eine gemeinsame soziokulturelle Basis und damit eine Verständigung auf Augenhöhe erreicht werden konnte, sowohl bei Herrn Brugger (01), mit dem der Forschende durch eine langjährige gemeinsame Chorarbeit verbunden ist, als auch bei Herrn Neuendorf (08), den dieser nur einmal während der Interview-Situation gesehen hat. Der Hang der Schweizer Musiklehrkräfte, den Wissenschaftler in seiner Autorität in der Musikpädagogik anzuzweifeln (Huber, 2023, in Vorbereitung), wird aufgrund der guten Vernetzung bzw. durch den Rapport zum Feld (Strübing et al., 2018) als vernachlässigbar eingestuft, womit davon ausgegangen werden darf, dass die Musiklehrkräfte in den Interviews über wesentliche Elemente ihres Singunterrichts berichteten.

Während der Interviews gab es aber einige Überraschungen bezüglich der Ausbildung der Interviewten, da die zu Beginn erstellten vier Ausprägungen der Professionalisierungswege der Realität nicht standhalten konnte. Beispielsweise haben zwei der Interviewten zwar eine Gesangsausbildung, aber keine Ausbildung für das Fach Musik, obwohl sie unter dieser Prämisse für die Studie eingeladen wurden, wobei eine der beiden Personen eine Ausbildung zur Chorleitung absolviert hat. Weiter stellte sich heraus, dass der fachfremd Unterrichtende eine Ausbildung als Primarlehrer absolviert hat und eine Lehrbefähigung für das Akkordeon hat. Diese Elemente führten dazu, die Ausprägungen der Professionalisierungswege zu überdenken und zusätzlich die in den Interviews erhobenen Daten ein zweites Mal zu verifizieren.

In einem ersten Schritt wurden die aus den Transkriptionen extrahierten Informationen zur Ausbildung durchgängig per Mail bzw. in einem Fall per Telefon überprüft. In einem zweiten Schritt wurde die heterogene Ausbildungssituation analysiert, wobei sichtbar wurde, dass auf der einen Seite

Musikausbildungen auf Bachelor- oder Master-Niveau vorhanden sind und auf der anderen Seite seminaristische Lehrgänge, diese ebenfalls auf BA- und MA-Niveau¹²⁹, insgesamt aber in sehr unterschiedlichen Kombinationen. Bei den ausgebildeten Musikern können Instrumentalisten, Sängerinnen und Sänger sowie Chorleiter, bei den ausgebildeten Lehrkräften Primar- und Sekundarlehrpersonen sowie als spezielle Kategorie die schulmusikalisch Gebildeten ausgemacht werden. Diese Einteilung folgt

	Musikalische Ausbildung			Pädagogische Ausbildung		
	Instrument	CK Gesang	Chor- leitung	PCK Schul- musik	Primar- stufe	PK Sek. 1
Musikalische Ausgebildete						
04 Hr. Hilfiker (35)	MA					
06 Fr. von Rohr (37)		MA				
08 Hr. Neuendorf (67)		MA	MA			
Pädagogisch Ausgebildete						
02 Hr. Buchser (55)						MA
03 Fr. Fricker (49)		X ¹³⁰		BA		MA
07 Hr. Rothenburg (49)			X ¹³¹			MA
Sowohl pädagogische als auch musikalisch Ausgebildete						
01 Hr. Brugger (52)		BA	MA	MA	BA	
05 Fr. Basler (47)		MA		MA		
09 Hr. Berner (59)	BA				BA	
10 Hr. Leimbacher (57)	MA/MA	BA	MA ¹³²	MA		
11 Fr. Derendinger (28)		BA		MA		

TABELLE 15: *Reorganisation der Studienteilnehmenden: musikalische versus pädagogische Ausbildung*

¹²⁹ Um eine bessere Übersichtlichkeit zu erhalten, wurden die vorhandenen Ausbildungen, die meist vor der Bologna-Reform absolviert wurden, der aktuell gebräuchlichen Ausbildungsstufen Bachelor und Master zugeordnet. Abkürzungen: BA – Bachelor-Niveau; MA – Master-Niveau.

¹³⁰ Langjährige Gesangsausbildung ohne Abschluss.

¹³¹ Kursorische Weiterbildungen im Bereich Kirchenmusik auf hohem Niveau.

¹³² Dieser Studienteilnehmer besitzt ein Dirigentendiplom, kein Chorleitungsdiplom.

noch stärker als die erste Profilbildung der Einteilung von Shulman (1987), werden doch CK, PK und PCK in der Einteilung gut sichtbar.

Die Tabelle 15 gibt einen Überblick über die Ausbildungen der Studienteilnehmenden. Interessanterweise ergeben sich aus der neuen Kategorisierung drei ähnlich starke Gruppen. Es sind dies Teilnehmende, die ausschließlich musikalisch oder pädagogisch ausgebildet sind bzw. in beiden Bereichen eine Ausbildung vorweisen können. In der Gruppe der *musikalisch Ausgebildeten* kann Herr Neuendorf (08) eine Chorleiterausbildung vorweisen, was einer schulmusikalischen Tätigkeit in Ansätzen nahekommt. Herr Hilfiker (04) mit dem Instrumentaldiplom und Frau von Rohr (06) mit dem Gesangsdiplom unterrichten ihre Klassen in pädagogischer und schulmusikalischer Hinsicht rein intuitiv. Diese Gruppe könnte am ehesten mit dem vorgängigen Profil der fachfremd Unterrichtenden verglichen werden, obwohl zwei Personen vokal gut ausgebildet sind. In der Einteilung nach Shulman fehlt dieser Gruppe PCK, Herrn Hilfiker (04) sogar CK zum Singen.

In der Gruppe der *pädagogisch Ausgebildeten* haben alle Personen eine generalistische Ausbildung für die Sekundarstufe 1, was auch eine basale Ausbildung für den Musikunterricht beinhaltet. Alle drei haben sich im Verlauf ihrer Ausbildung zusätzlich musikalisch weitergebildet, wobei Frau Fricker (03) als einzige eine schulmusikalische Ausbildung auf BA-Niveau absolviert hat. Diese Gruppe könnte am ehesten mit dem Profil der generalistisch Ausgebildeten verglichen werden, ihnen fehlt CK auf MA-Niveau. In der letzten Gruppe zeigen sich Musiklehrkräfte mit unterschiedlichen *pädagogischen und musikalischen Ausbildungen*, die wenig Passung mit den vorgängigen Profilen aufzeigen. Zum einen haben die fünf Personen unterschiedlich vertiefte musikalische Ausbildungen. Frau Derendinger (11) und Herr Berner (09) können lediglich eine musikalische BA-Ausbildung vorweisen, Herr Leimbacher (10) jedoch mehrere MA-Ausbildungen. Zum anderen gibt es zwei Personen, die als Primarlehrkräfte (BA) ausgebildet sind, und vier Teilnehmende, die eine MA-Ausbildung im Bereich Schulmusik vorweisen können. In dieser Gruppe ist die Ausbildungssituation sehr heterogen.

In der Analyse des Interviewmaterials wird sich zeigen, ob die Unterscheidung in *musikalisch* bzw. *pädagogisch* und *musikalisch-pädagogisch* Ausgebildete zu übereinstimmenden Haltungen und Methoden führt, oder ob die Komplikation bei der Einteilung der Studienteilnehmenden nur der

Ausdruck von Heterogenität ist und sich die Bildung von Profilen für die Forschungsfrage als wenig relevant erweist. Es könnte an dieser Stelle bemängelt werden, dass die unklare Ausbildungssituation besser hätte antizipiert werden müssen. Das Ziel, eine möglichst breite Palette von Ansichten zum Singen im Musikunterricht zu erhalten, wird durch die vorgenommene Überarbeitung der Ausbildungsprofile aber nicht tangiert, vielmehr zeigt sich darin die bereits beschriebene Vielfalt des schweizerischen Bildungssystems und eine Anstellungspraxis, welche die schulmusikalische Ausbildung von Musiklehrkräften offensichtlich nicht als zwingend erachtet.

9.4 Transkription und Festlegung der Analyseeinheiten

In diesem Abschnitt wird der Prozess der Festlegung der Analyseeinheiten beschrieben. Dies entspricht dem Schritt 4 (vgl. 8.2) der QIA und beinhaltet die Beschreibung des Prozesses von der Aufnahme der Interviews bis zur Integration der Daten in MAXQDA.

Die Interviews wurden in vollständiger Länge als Audiodatei aufgenommen. Während des Gesprächs wurden keine Notizen und auch keine Bilder oder Videos erstellt. Kuckartz (2018, S. 165) nennt einige Vorteile von Audio-Aufnahmen gegenüber schriftlichen Notizen, unter anderem die Vermeidung von Verzerrungen durch bereits erfolgte Interpretation während des Notierens und die Möglichkeit direkten Zitierens im Forschungstext. Der ebenfalls von Kuckartz genannte Nachteil einer möglichen Verunsicherung von Interviewten kann aufgrund der durchwegs entspannten Interviewsituationen, die bereits im vorigen Abschnitt erwähnt wurde, ausgeschlossen werden. Als Aufnahmegerät diente ein iPad der 3. Generation mit dem geräteeigenen Mikrofon. Als Aufnahmeprogramm diente die App *Voice Record Pro*. Abgesehen von einer kleinen Unterbrechung von wenigen Minuten während des Interviews mit Herrn Hilfiker (04) aufgrund fehlender Batterieleistung wurden alle Interviews vollständig in einer Gesamtlänge von 511 Minuten aufgenommen.

Die Dateien wurden anschließend auf ein MacBook Pro übertragen, mit iTunes in das Format WAV (16bit/44100kHz) umgerechnet und mit dem Programm F5 Transkription Pro (Version 7.0.1) transkribiert. Dazu wurden zuerst Transkriptionsregeln formuliert, die an dieser Stelle in zusammen-

gefasster Form dargestellt werden. Eine vollständige Übersicht ist digital verfügbar (siehe Anhang). Die Interviews wurden vollständig und wörtlich transkribiert. Der schweizerdeutsche Dialekt wurde möglichst genau in die hochdeutsche Sprache übersetzt, was notwendigerweise zu leicht angepassten Satzstellungen führte. Insbesondere wurden oft Verben zur besseren Lesbarkeit innerhalb der Satzstellung verschoben. Zum selben Zweck wurden an wenigen Stellen Wörter „übersetzt“, so z. B. „lueg“ zu „schau“. Typische Ausdrücke der jeweiligen Schweizer Mundart wurden nicht angepasst und werden im Bedarfsfall in Zitaten mit einer Fußnote erklärt. Krüger und Riemeier (Krüger & Riemeier, 2014, S. 136) weisen auf eine mögliche Beeinträchtigung der Authentizität hin, wenn Dialekt bereinigt wird. In den vorliegenden Interviews kann dies ausgeschlossen werden, da der Forschende die gleiche Mundart spricht. Eine weitere Eigenheit betrifft die oft verwendete Form von eingeschobenen beispielhaften Anweisungen an die in den Interviews nicht vorhandenen Schülerinnen und Schüler zur Verdeutlichung von Handlungen der Lehrkräfte. Diese wurden nicht als direkte Rede transkribiert, sondern einfach als Teilsätze innerhalb der gemachten Aussagen.

Die elf entstandenen Transkripte wurden anschließend in das Programm MAXQDA eingepflegt und einer Anonymisierung unterzogen. Dabei wurden die Nachnamen der Interviewten durch einen Decknamen ersetzt und es wurde jeweils zur Verdeutlichung des Geschlechts Herr und Frau vor die Namen gesetzt. Zusätzlich wurden die Interviews nummeriert und mit dem Alter der Teilnehmenden ergänzt. Somit ist es möglich, die relevanten Daten für die Lesenden dieser Studie nachvollziehbar zu machen.

Das Material dieser QIA besteht somit aus elf Transkripten von als Audiodatei aufgenommenen Interviews. Die einzelnen Interviews werden Analyseeinheit deklariert und vollständig verwendet.

10 Kategoriensystem und Kodierung

Mit der Auswahl des Samples, der Durchführung der Interviews und deren anschließender Transkription wurde die Grundlage geschaffen, in Referenz zu den Forschungsfragen (Kapitel 7) und unter Berücksichtigung des Datenkorpus in einem iterativen Prozess das Kategoriensystem für die qualitative Inhaltsanalyse (QIA) nach Mayring (2015) zu entwickeln. Dieses Kapitel hat die Aufgabe, den zugehörigen Forschungsprozess nachvollziehbar und transparent zu machen. Dazu wird zuerst die digitale Unterstützung durch die Software MAXQDA im Abschnitt 10.1 vorgestellt und diskutiert. Der anschließende Abschnitt 10.2 beschreibt die Schritte 3 sowie 5–10 der strukturierenden Inhaltsanalyse, wie sie von Mayring und Fenzl (2019, S. 638 ff.) vorgeschlagen und in Kapitel 8.2 (vgl. Abbildung 7) bereits dargelegt wurden. Dabei werden einerseits deduktiv drei Hauptkategorien hergeleitet, an das Material angelegt und weiterentwickelt und andererseits in einem induktiven Verfahren aus dem Material heraus weitere Subkategorien entwickelt und ins Kategoriensystem eingepflegt. In aufeinanderfolgenden Phasen wird das Kodierhandbuch (siehe Anhang) erstellt und das ganze Material mehrmals kodiert und rekodiert. In Abschnitt 10.3 wird nachher die Güte der QIA mittels Intercoder-Reliabilitätsprüfung überprüft, um abschließend zur Analyse und Interpretation der empirischen Daten überzuleiten.

10.1 MAXQDA als Forschungsplattform

Die Software MAXQDA war in diesem Projekt in verschiedenen Versionen in Gebrauch, wobei stets mit der aktuellen Version gearbeitet wurde. Die Firma Verbis vertreibt die Software mit dem Anspruch, zu den „weltweit führenden Softwareprogrammen für qualitative und Mixed Methods-Forschung“ (Verbis, 2020) zu gehören. Eine erste Version des Programms wurde unter dem Namen MAX Ende der 1980er-Jahre vorgestellt und seither permanent weiterentwickelt (Rädiker & Kuckartz, 2019, S. v). Der Name MAX nimmt Bezug auf den deutschen Soziologen Max Weber (1864–1920) und beinhaltet zusätzlich die Abkürzung QDA, die für

qualitative data analysis steht. Das Erstellen von Kategoriensystemen und die Kodierung von Text ist seit Beginn der Entwicklung das zentrale Anwendungsfeld der Software.

Der Gebrauch einer Software für Forschungszwecke ist nicht unumstritten. So bemängeln z. B. einige Autoren, dass durch die Programmstruktur „a kind of hidden curriculum“ (Rädiker & Kuckartz, 2019, S. 8) und damit eine spezifische Richtung der Inhaltsanalyse vorgegeben sei, welche die Forschung in ein methodisches Korsett zwingt. Andere Autoren betrachten dagegen Analysesoftware als einen Werkzeugkasten, der in diversen Forschungsmethoden verwendet werden kann (S. 9). In der vorliegenden Arbeit wurde die Tendenz zum Korsett nur am Rande als hinderlich wahrgenommen, da sich der Gebrauch der Software lediglich auf die Hauptfunktionen beschränkte. Der Gebrauch von visualisierenden Analyse-Tools wurde zwar getestet, aber wieder verworfen, weil die methodischen Zugänge und Berechnungsmethoden nicht ermittelt werden konnten. Im Anschluss folgt eine Liste mit den Arbeitsschritten, die in MAXQDA durchgeführt wurden:

- Sowohl die Audio-Daten wie auch die Transkripte wurden nach der Transkription in das Programm MAXQDA eingepflegt.
- Der Prozess der Kategorienbildung, inkl. der Erstellung des Kodierhandbuchs, wurde in mehreren iterativen Zyklen direkt in MAXQDA durchgeführt.
- Ebenso wurden alle ergänzenden Forschungsnotizen den Transkripten direkt zugeordnet sowie zentrale Arbeitsschritte protokolliert.
- Quantifizierbare Daten wurden in Variablen überführt.
- Anschließend wurden die kodierten Textstellen in sogenannten Summaries zusammengefasst, zuerst in den Subkategorien, anschließend auch in den übergeordneten Kategorien.
- Abschließend wurde die Intercoder-Reliabilitätsprüfung mit MAXQDA durchgeführt, darin enthalten war auch der paarweise Vergleich der Kodierungen im Rahmen der Einigungskonferenz.
- In der Analyse konnten einige der eingebauten Auswertungsmethoden hilfreich eingesetzt werden. Die abschließende Auswertung und Interpretation wurde ausgehend von den Summaries und im Abgleich zu den Transkripten außerhalb von MAXQDA vorgenommen.

Die komplette Erfassung des Projekts in einer MAXQDA-Datei ermöglicht den Nachvollzug der Kodierung und Analyse mit dem kostenlosen

Reader¹³³ des Software-Herstellers. Der Zugang zu dieser Datei, und damit zu den Transkripten, der Kodierung und ebenfalls zum Kodierleitfaden, wird im Anhang beschrieben.

10.2 Entwicklung des Kategoriensystems und Kodierung des Materials

Ausgehend von den Forschungsfragen (vgl. Kapitel 7) und den erstellten Profilen des Samplings (vgl. 9.1) wurden zuerst drei Analyserichtungen (Schritt 3) bestimmt, um das Kategoriensystem zu entwickeln. Bevor die verschiedenen Phasen dieses Prozesses besprochen werden, sollen diese Analyserichtungen vorgestellt und in Abbildung 10 visualisiert und aufeinander bezogen werden.

Das narrative Moment innerhalb des Interviewleitfadens (siehe Anhang) mit der Einstiegsfrage nach typischen Singsituationen und der Ausstiegsfrage nach Hoch- und Tiefpunkten zielte als erste Perspektive darauf ab, wichtige und erleuchtende Momente aus der Sicht der Lehrkräfte beschreiben zu lassen und so ihre Perspektive ins Zentrum zu rücken. Die erfahrenen Lehrkräfte hatten so die Möglichkeit, ihre Unterrichtskultur des Singens darzulegen. Diese Momentaufnahmen knüpfen lose an die autoethnografische Methode (Ellis & Bochner, 2000; Ellis et al., 2010) an und schlagen damit einen Bogen zu Kapitel 3, in dem mit dem

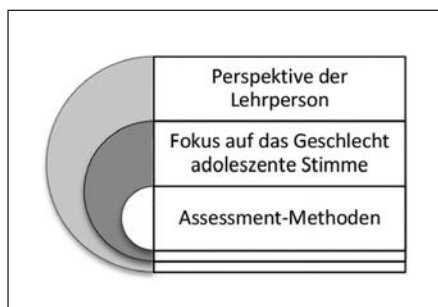


ABBILDUNG 10: Analyserichtungen der QIA

autoethnografischen Exkurs Teile der Unterrichtskultur des Forschenden aufgearbeitet wurde. Die Kategorie *PERSPEKTIVE DER LEHRKRAFT* wurde zwar in ihrer Grundidee deduktiv gesetzt, ihre Subkategorien wurden aber hauptsächlich induktiv aus dem Material heraus entwickelt. Später in Ka-

¹³³ Auf der Webseite von MAXQDA wird der Reader zum Download angeboten: <https://www.maxqda.de/produkte/maxreader> (abgerufen am 20.07.2021).

pitel 12 dienen die Kodierungen dieser Kategorie als Leitfaden zur Entwicklung von kurzen Porträts, um einen vertikalen Blick auf das Untersuchungsmaterial zu werfen und damit die individuellen Unterrichtskulturen der Interviewten lebendig werden zu lassen.

Eine weitere Kategorie wurde deduktiv aus dem Kapitel 4 zum Forschungsstand Singen in der Adoleszenz erschlossen und *FOKUS AUF DAS GESCHLECHT*¹³⁴ genannt. In den Transkripten wurde nach Aussagen zur Entwicklung der adoleszenten Stimme sowie zum Zusammenspiel der Geschlechter beim Singen gesucht, dies als Basis von *content knowledge* (CK) für mögliche Individualisierungstendenzen. Zusätzlich wurden mono-educative Momente in den Blick genommen. Diese zweite Perspektive wurde innerhalb des Leitfadens als Fragenkomplex vorgesehen, einerseits explizit bei der Frage nach dem Umgang mit der Mutation bei Mädchen und Jungen und andererseits implizit auf Nachfrage bei Hinweisen zu Auswirkungen des eigenen Geschlechts auf die Schülerinnen und Schüler.

Der Kern der Inhaltsanalyse liegt aber in der Betrachtung von Assessment-Methoden im Singunterricht und bezieht sich auf die Theorie in Kapitel 5. Ausgehend von den Überlegungen von Earl (2013) wurde *FOR LEARNING, AS LEARNING* und *OF LEARNING* als Subkategorien der Kategorie *ASSESSMENT* deduktiv gesetzt. Im Leitfaden wurde hauptsächlich nach Aufgabenstellungen, Feedback, Förderung und Beurteilung gefragt, der Terminus *Assessment* wurde nicht genannt. Damit sollte eine Beeinflussung der Interviewpartnerinnen und -partner durch das Setzen von normativen Ansprüchen vermieden werden.

Die beiden Kategorien *ASSESSMENT* und *FOKUS AUF DAS GESCHLECHT* führen in Kapitel 12 zu einer horizontalen Analyse des Datenkorpus, um spezifisch für den Singunterricht die Ansprüche des Assessment-Modells von Earl (2013) zu diskutieren. In Abbildung 10 werden die drei Perspektiven mit einem Modell zueinander in Bezug gebracht. Auf der einen Seite wird dadurch sichtbar, dass ausgehend von der Perspektive der Lehrkraft und den Spezifika des Singens zum Assessment, dem Kern der inhaltlichen Analyse, vorgestoßen werden soll. Auf der anderen Seite zeigt sich aber auch eine Durchdringung der Perspektiven, welche die Abgrenzung von

¹³⁴ Die Setzung des Terminus *adoleszente Stimme* erfolgte nach der Bestimmung der Subkategorie *FOKUS AUF DAS GESCHLECHT*, der aus Effizienzgründen nicht mehr justiert wurde.

Kategorien und Subkategorien erschwert und zu Mehrfachkodierungen führt. Dies wird später nochmals thematisiert.

In den folgenden Abschnitten werden nun die verschiedenen Phasen der Entwicklung des Kategoriensystems dargelegt. An dieser Stelle sei darauf hingewiesen, dass das Kategoriensystem aufgrund der Art der vorliegenden Forschungsarbeit nicht in einer Teamkonstellation, sondern individuell entwickelt wurde. Erst in der Interrater-Reliabilitätsprüfung wurde auf eine zweite Raterin zurückgegriffen.

Ausgehend von den drei beschriebenen Hauptkategorien und den hergeleiteten Subkategorien wurden in der ersten Phase die Interviews 02, 04, 06, 10 und 11 ein erstes Mal kodiert. Die Auswahl der Interviews erfolgte intuitiv, ausgehend von einer ersten Einschätzung, welches Material spannend sein könnte. Während dieses Vorgangs wurden weitere Subkategorien auf der Ebene zwei induktiv ergänzt und ebenfalls einige Subkategorien auf einer dritten Ebene eingefügt, insbesondere dort, wo sehr viele Kodierungen vorhanden waren. Im Bereich Assessment wurde z. B. der Bereich Förderung in sieben Subkategorien aufgeteilt, um die verschiedenen Aspekte in der Vielzahl von Kodierungen auseinanderhalten und wieder auffinden zu können. Durch konsequentes Schreiben von Memos wurden die verschiedenen Kategorien inhaltlich verdichtet. Abschließend wurde zu jeder Kategorie eine Definition mit Bezug zur Theorie, notwendige Abgrenzungshinweise und ein oder zwei Ankerbeispiele direkt aus den Interviews eingefügt. Aus diesen Memos wurde das Kodierhandbuch erstellt, das im weiteren Verlauf die Kodierung und Rekodierung unterstützte und als Basis für den Interrater-Reliabilitätstest diente¹³⁵.

Zusätzlich zum Kodierhandbuch entstand der Leitfaden Kodierung (vgl. Anhang), der die Regeln zur Setzung von Codes beschreibt. Neben weniger wichtigen administrativ ausgerichteten Hinweisen ist die Bestimmung des Umfangs von Kodiereinheiten der wesentliche Punkt in diesem Leitfaden. Im Rahmen der strukturierenden Analyse wurde es als sinnvoll erachtet, Sinnzusammenhänge großflächig zu kodieren und deshalb die Kodiereinheiten über einen oder mehrere Satzglieder auszuweiten. So

¹³⁵ In MAXQDA kann zu jeder Kategorie ein Memo verfasst werden. Diese Memos können anschließend zu einem Kodierhandbuch zusammengefasst und in ein Textverarbeitungsprogramm exportiert werden. Das Kodierhandbuch kann über den QR-Code im Anhang eingesehen werden.

wurde z. B. auch die Frage des Interviewers mitkodiert, um ein *ja* oder *nein* in einen Zusammenhang zu stellen. Trotzdem zeigte sich, dass durch die Setzung der drei Analyserichtungen einzelnen Textpassagen mehrere Subkategorien zugeordnet werden mussten. Aus diesem Grund wurde bewusst entschieden, pro Sinneinheit mehrere Kodierungen zuzulassen und die mehrperspektivische Herangehensweise so zu unterstützen. Bereits in dieser Phase zeigte sich, dass sich die Kategorien flächendeckend über die fünf Interviews verteilen.

Nach der Schlussredaktion des Kodierhandbuchs wurde in einer nächsten Phase mit dem Interview 01 ein bisher nicht bearbeitetes Interview kodiert. Das Kodierhandbuch und das redigierte Kategoriensystem erwiesen sich als tragfähig. Lediglich die Subkategorie *DIFFERENZIERUNG* wurde mit einer Subkategorie *DIFFERENZ* ergänzt, schien es doch in diesem spezifischen Interview einen klaren Unterschied zu geben, ob der Lehrer Differenz in seinen Klassen erkennt oder ob er den Unterricht auch differenzierend gestaltet. Anschließend an diesen Testlauf wurde für dieses Interview eine Einschätzung zu den Kodierhäufigkeiten in Bezug auf die drei Analyserichtungen erstellt und erste Summaries für die Kategorien geschrieben. Ebenfalls in dieser Phase erfolgte der Intracoder-Check (Schritt 8), indem die bereits kodierten Interviews nochmals durchgesehen werden, um die Anpassungen im Kategoriensystem nachzuvollziehen¹³⁶. Auch bei diesen Interviews wurden erste Einschätzungen zu den Kodierhäufigkeiten gemacht. In einer letzten Phase wurden die restlichen fünf Interviews 03, 05, 07, 08 und 09 kodiert. Für jede Subkategorie wurde pro Lehrkraft ein Summary erstellt und pro Kategorie zusammengefasst. Diese Arbeit wurde zunächst für ausgewählte Interviews (z. B. 01 und 11) parallel mit dem Erstellen der Porträts in Kapitel 12.1 ausgeführt. In dieser Phase wurde außerdem die Skalierung der Subkategorien *SOZIALFORMEN* vorgenommen. Das Vorgehen dazu wird in Abschnitt 11.3 genauer erläutert.

Im Verlauf der Arbeit mit der Software MAXQDA zeigte sich die Möglichkeit, alle gesammelten Informationen und Notizen direkt in der Datei abzulegen. Zur besseren Auffindbarkeit wurden dazu Hilfskategorien erstellt. In den entsprechenden Summaries wurden Forschungsnotizen, wie

¹³⁶ Es wurde darauf verzichtet, die fünf ersten Interviews noch einmal von Grund auf zu kodieren, wie das von Mayring und Fenzl (2019) vorgeschlagen wird.

Umstände zu Interviews, Einschätzungen und Arbeitsfortschritte pro Lehrkraft, dokumentiert. Zusätzlich zeigte sich die Notwendigkeit, die Daten zu den Berufsbiografien zu verifizieren, was ebenfalls in Hilfskategorien dokumentiert wurde. Alle demografischen Daten wurden in Variablen erfasst, so das Alter der Probanden, das Geschlecht, die Anzahl Jahre Erfahrung und die unterrichteten Schulniveaus.

10.3 Reliabilitätsprüfung der QIA

Nach der schrittweisen Entwicklung des Kategoriensystems und der Kodierung wird mit der Interdecoder-Reliabilitätsprüfung getestet, ob das entwickelte Kategoriensystem intersubjektiv nachvollziehbar ist. Wie bereits beschrieben (vgl. 8.4), kodiert dabei eine externe Person auf der Basis des Kodierhandbuchs und der darin enthaltenen Beschreibungen von Kategorien, Ankerbeispielen und Abgrenzungsvorgaben Anteile des Materials. Dies geschieht in der Erwartung, dass „zwei Personen die gleichen Themen, Aspekte und Phänomene im Datenmaterial identifizieren und den gleichen Kategorien zuweisen“ (Rädiker & Kuckartz, 2019, S. 287). Um die Zweit-Kodierung unter kontrollierten Bedingungen durchzuführen, wird neben dem Kodierhandbuch auch der Kodierleitfaden (Kapitel 16.3) hinzugezogen. Dieser legt fest, wie die Sinneinheiten im Text kodiert werden. Zusätzlich werden die Kodierungen in den Transkripten für die zweite Person entfernt, damit diese unvoreingenommen am Material kodieren kann. Dieser Vorgang wird als wichtiges Gütemerkmal für die Objektivität beschrieben (Mayring & Fenzl, 2019, S. 41), aber in qualitativen Inhaltsanalysen allzu oft mit einem knappen Satz und der Angabe von Kappa κ_n abgehakt. In der vorliegenden Arbeit wurde bereits im Abschnitt zu den Gütekriterien der QIA (vgl. 8.4) der Wert Kappa κ_n diskutiert. In den folgenden Abschnitten wird mit zwei Tests untersucht, wie sich dieser Wert vor und nach der jeweiligen Einigungskonferenz und in Abhängigkeit der Überlappung von Kodiereinheiten verändert, um einerseits zu zeigen, welche Werte für die vorliegende QIA erreicht wurden, aber auch um deutlich zu machen, dass der Wert Kappa κ_n durch einige Parameter beeinflusst wird.

Mit der Zweit-Kodierung wird aber zusätzlich zur Überprüfung der Intersubjektivität auch ein Verständigungsprozess ermöglicht, in dem die

Qualität der QIA in einer sogenannten Einigungskonferenz verbessert werden kann (Mayring & Fenzl, 2019, S. 42). Dazu werden Unklarheiten bei den Definitionen der Kategorien nochmals diskutiert, um anschließend die Kodierung konsensual zu überarbeiten.

Der Vergleich beider Kodierungen führt anschließend mittels Berechnung zum Koeffizienten Kappa κ_n . Die technische Durchführung der Interrater-Reliabilitätsprüfung wird mit der Software MAXQDA durchgeführt. Im Rahmen einer effizienten Abwicklung werden nur Teile des Materials und ausgewählte Kategorien (S. 42) überprüft.

Für die vorliegende Forschungsarbeit konnte innerhalb des Teams eine wissenschaftliche Mitarbeiterin mit erziehungswissenschaftlichem Master¹³⁷ mit der Zweit-Kodierung beauftragt werden. Die Interrater-Reliabilitätsprüfung erfolgte in zwei Phasen. In einer ersten Phase (Test A) wurden fünf Interviews ausgewählt und bezüglich der Kategorie *FOKUS AUF DAS GESCHLECHT* einer Zweit-Kodierung unterzogen. In einer zweiten Phase (Test B) wurden weitere fünf Interviews bezüglich der Kategorie *PERSPEKTIVE DER LEHRKRAFT* ausgewertet. In den folgenden Abschnitten werden die Resultate der beiden Tests präsentiert. In der Tabelle 16 wird gezeigt, welche Interviews für die beiden Tests ausgewählt wurden und wie viele Kodierungen jeweils durch Rater 1 für die entsprechenden Kategorien gesetzt wurden.

Test A

In einem ersten Schritt wurde von den drei Kategorien *PERSPEKTIVE DER LEHRKRAFT*, *FOKUS AUF DAS GESCHLECHT* und *ASSESSMENT* der zweite Bereich gewählt. Dieser umfasst vier übergeordnete Subkategorien mit dreizehn weiteren Subkategorien in einer dritten Ebene, ist thematisch übersichtlich¹³⁸ und in den Interviews gut zu lokalisieren. Kodiert wurden fünf von elf Interviews. Aus der Tabelle 16 erschließt sich, dass vor allem Interviews mit hohen Kodierhäufigkeiten im entsprechenden Bereich ausgewählt

¹³⁷ Wissenschaftliche Mitarbeiterin der Professur Musikpädagogik im Jugendalter (PH FHNW). Master of Science in Education (Erziehungswissenschaft), Universität Bern.

¹³⁸ Siehe Kodierhandbuch im Anhang.

wurden. Weiter wurde mit der Auswahl zweier Lehrerinnen im Pool der Interviews auf die Berücksichtigung beider Geschlechter geachtet.

Bei der ersten Zusammenführung der beiden Kodierungen wurde ein Wert von $\kappa_n = .28$ erreicht, was nach Landis und Koch (1977) einer *mittelmäßigen (fair)* Übereinstimmung entspricht. In der Einigungskonferenz hat sich aber gezeigt, dass verschiedene Umstände zu diesem niedrigen Wert beigetragen haben. Im einleitend durchgeführten paarweisen Vergleich hat sich das Kodierhandbuch als zuverlässig erwiesen, da viele Anpassungen lediglich vorgenommen wurden, um die Größe der Kodiereinheiten anzugleichen und damit den Gesamtzusammenhang innerhalb einer Textstelle besser darzustellen. Nur wenige Kodierungen wurden nach vertiefter Diskussion thematisch korrigiert. Differenzen haben sich vor allem dadurch ergeben, dass Rater 1 im Gegensatz zu Raterin 2 Subkategorien auch doppelt zugeordnet hat, wenn in einer Aussage mehrere Subkategorien angesprochen wurden. Raterin 2 hat dagegen keine überlappenden Kodierungen gesetzt, was im Kodierleitfaden zwar so verlangt, aber nicht umgesetzt wurde. Weiter gab es die eine oder andere Unklarheit aufgrund der fehlenden disziplinären Vertiefung von Raterin 2. Der anschließend an

Test A			Test B	
	Kodier- häufigkeiten Fokus auf das Geschlecht	Auswahl	Kodier- häufigkeiten Perspektive der Lehrkraft	Auswahl
01 Hr. Brugger	7		45	x
02 Hr. Buchser	14		62	x
03 Fr. Fricker	15	x	36	
04 Hr. Hilfiker	17	x	41	
05 Fr. Basler	18		53	
06 Fr. von Rohr	9		39	
07 Hr. Rothenburg	10		50	
08 Hr. Neuendorf	10		51	x
09 Hr. Berner	10	x	48	
10 Hr. Leimbacher	16	x	47	x
11 Fr. Derendinger	29	x	52	x
<i>Total der Interviews</i>		5		5

TABELLE 16: Ausgewählte Interviews für die Interrater-Reliabilitätsprüfung

die Einigungskonferenz erreichte Wert von $\kappa_n=.71$ war deutlich höher und entspricht einer *erheblichen* (*substantial*) Übereinstimmung.

Überlappung	90 %	80 %	70 %	60 %	50 %	40 %	30 %	20 %
Einigungskonferenz								
davor	.19	.21	.24	.28	.38	.44	.53	.61
danach	.61	.64	.66	.71	.72	.73	.75	.77
Differenz	.42	.42	.42	.43	.34	.29	.22	.16

TABELLE 17: Test A – Kappa κ_n in Abhängigkeit der prozentualen Überlappung von Kodiereinheiten

In der Einigungskonferenz wurde zudem ersichtlich, dass die Einstellung zur notwendigen prozentualen Überlappung von Kodiereinheiten einen deutlichen Einfluss auf die Berechnung von Kappa κ_n hat. So wird in MAXQDA eine gemeinsame Kodierung aufgrund der prozentualen Übereinstimmung gesetzt, wobei dieser Wert vor der Berechnung definiert wird. Um diesen Einfluss auf den Wert Kappa κ_n zu evaluieren, wurde Kappa κ_n in Abhängigkeit der prozentualen Überlappung in 10er-Schritten ausgehend von 90 % berechnet. In der Tabelle 17 werden die Werte vor und nach der Einigungskonferenz und die jeweilige Differenz aufgelistet. Vor der Einigungskonferenz steigt der Wert Kappa κ_n stark an je kleiner die Überlappung der Kodierungen ist und die Differenz der beiden Werte von Kappa κ_n unterhalb der 60 %-Überlappung sinkt erkennbar. Zum einen zeigt dies, dass die Übereinstimmung der Kodierung bereits vor der Einigungskonferenz deutlich höher einzuschätzen ist, wenn man lediglich das Übereinstimmen der Fundstellen gewichtet. So wird bei der Überlappung von 20 % ein Wert von Kappa $\kappa_n=.61$ erreicht. Nach der Einigungskonferenz ist Kappa $\kappa_n=.61$ bereits bei einer Überlappung von 90 % erreicht, was zeigt, dass hauptsächlich die fehlende Übereinstimmung in der Kodierstrategie bezüglich der Segmentgröße für den zu Beginn tiefen Wert von Kappa $\kappa_n=.28$ verantwortlich ist. Die Überlegungen deuten weiter darauf hin, dass der Wert Kappa κ_n als reiner Zahlenwert nur dann Sinn macht, wenn auch eine Angabe zur Überlappung besprochen wird, insbesondere wenn wie in der vorliegenden Arbeit die Kodiereinheiten nicht nur Satzglieder, sondern ganze Sinneinheiten beinhalten. Nicht zuletzt zeigt sich hier auch der nicht außer Acht zu lassende Einfluss von Software auf Forschungsergebnisse.

Test B

Ergänzend zum ersten Interrater-Reliabilitätstest wurde in einem zweiten Schritt der Bereich *PERSPEKTIVE DER LEHRKRAFT* mit einer Anzahl von fünf Interviews einer Zweit-Kodierung unterzogen. Die Auswahl der Interviews kann wiederum in der Tabelle 16 nachvollzogen werden und fiel wieder auf Interviews mit hohen Kodierhäufigkeiten im entsprechenden Bereich. Zudem wurden zwei Interviews ein zweites Mal mit einbezogen. Die anschließende Berechnung von Kappa κ_n zeigte mit der Voreinstellung einer Überlappung von 60 % vor der Einigungskonferenz einen Wert von $\kappa_n=.70$, was einer *erheblichen (substantial)* Übereinstimmung entspricht, und erreichte nach dem paarweisen Vergleich von nur zwei Interviews bereits einen Wert von $\kappa_n=.83$. Aufgrund des erreichten Wertes wurden die drei restlichen Interviews nur noch stichprobenartig besprochen.

Auch für den zweiten Test wurde die Berechnung von Kappa κ_n in Abhängigkeit von der prozentualen Überlappung berechnet. Die Werte entsprechen insgesamt einer erheblichen (*substantial*) bzw. einer *exzellenten (excellent)* Übereinstimmung, egal welche prozentuale Abdeckung gewählt wurde. Die Differenz der Werte von Kappa κ_n vor und nach der Einigungskonferenz liegt deutlich tiefer als im ersten Test und untermauert den hohen Wert von Kappa κ_n und damit die große Übereinstimmung in

Überlappung	90 %	80 %	70 %	60 %	50 %	40 %	30 %	20 %
Einigungskonferenz								
davor	.51	.61	.65	.70	.73	.75	.77	.79
danach	.66	.75	.79	.83	.84	.84	.85	.87
Differenz	.15	.14	.14	.13	.11	.09	.08	.08

TABELLE 18: Test B – Kappa κ_n in Abhängigkeit der prozentualen Überlappung von Kodiereinheiten

der Kodierung. Der zweite Test weist damit einen deutlich besseren Übereinstimmungsgrad auf. Auf der einen Seite kann dies mit der besseren Absprache bezüglich der Kodierstrategie begründet werden und auf der anderen Seite darf angenommen werden, dass im zweiten Durchlauf Raterin 2 bereits tiefer mit der Materie vertraut war. Die Interrater-Reliabilitätsprüfung führte im Test A zuerst zu einer mittelmäßigen Übereinstimmung der Kodierungen. Die Analyse der Resultate während der Einigungskon-

ferenz ergab jedoch, dass insbesondere die fehlende Übereinstimmung der Kodierstrategie bezüglich der Größe der Segmente zum mäßigen Resultat führte. Der abschließende Wert von $\kappa_n = .71$ nach der Einigungskonferenz entspricht einer erheblichen Übereinstimmung. Der Test B weist bereits vor der Einigungskonferenz mit dem Wert von $\kappa_n = .70$ eine erhebliche Übereinstimmung auf und kann durch die Einigungskonferenz mühelos auf eine exzellente Übereinstimmung angehoben werden. In beiden Fällen wurde Kappa κ_n in Abhängigkeit von der prozentualen Überlappung der Kodierungen stufenweise berechnet, womit gezeigt werden konnte, dass bereits im Test A eine gute Übereinstimmung angelegt war und im Test B eine sehr überzeugende Übereinstimmung erreicht wurde. Der deutliche Einfluss der prozentualen Überlappung lässt die Diskussion um die Genauigkeit von Berechnungsformeln von Kappa κ_n (Brennan & Prediger, 1981; Wirtz & Kutschmann, 2007) eher in den Hintergrund treten.

10.4 Fazit zur Erstellung des Kategoriensystems

In diesem Kapitel wurde die methodische Durchführung der strukturierenden Inhaltsanalyse und die Evaluation des Kategoriensystems durch den Interrater-Reliabilitätstest belegt. Ausgehend von den Analyseeinheiten wurde das Material durch die Entwicklung des Kategoriensystems strukturiert und durch die Kodierung und Erstellung der Summaries bereits einer ersten interpretativen Verarbeitung unterzogen.

Die Festlegung der Analyserichtung erfolgte zuerst deduktiv und anschließend induktiv, bezogen auf drei Perspektiven: zuerst mit dem Blick der Lehrkräfte auf ihren eigenen Unterricht, dann spezifisch auf die adoleszente Stimmentwicklung und das Zusammenspiel der Geschlechter beim Singen und zuletzt bezüglich Diagnose, Förderung und Beurteilung unter dem Gesichtspunkt von Assessment.

Damit ist die Grundlage für eine vertiefte quantitative und qualitative Analyse und Deskription des Materials und den Rückbezug auf Elemente der Theorie in den nächsten Kapiteln gelegt.

D DARSTELLUNG DER RESULTATE

Der abschließende Schritt der vorliegenden QIA gilt der Analyse des Materials mit Blick auf die Forschungsfragen. Zuerst wird in Kapitel 11 eine quantitative Analyse der Kodierhäufigkeiten erstellt, um eine Einschätzung zu ermöglichen, inwiefern die Thematik der Forschung in den Interviews tatsächlich zum Tragen kommt. Weiter wird die Kategorie Assessment genauer analysiert und mit einer skalierenden Analyse der Unterrichtssettings ergänzt, um die individuelle Verteilung der Sozialformen einzuschätzen. In Kapitel 12 werden die Zielvorstellungen der Lehrkräfte auf der Grundlage des entwickelten Kategoriensystems zusammengefasst. In Form kurzer Porträts werden die zentralen Aspekte der Interviews verdichtet. Dieses Kapitel dient in erster Linie dazu, Unterrichtskulturen fassbar zu machen und eine Basis für die fokussierte Betrachtung von Typiken zu schaffen. Auf dem Weg zum Kern der QIA werden in Kapitel 13 zuerst die Einflüsse der adoleszenten Stimme auf das Assessment beleuchtet, um im zweiten Teil den Blick auf die Assessmentverfahren zu legen. Zuerst wird der Frage nachgegangen, wie Lehrkräfte Förderung und Beurteilung organisieren, um anschließend entlang der drei Bereiche AOL, AAL und AFL zu analysieren, wie die Lehrkräfte ihre Schülerinnen und Schüler individuell fördern und beurteilen. Abschließend wird auf der Grundlage der Analyse Bezug auf die Forschungsfragen genommen.

11 Quantitative Analyse

Obwohl die vorliegende Forschungsarbeit qualitativ ausgerichtet ist, werden im Rahmen einer quantitativen Analyse die Kodierhäufigkeiten der Kategorien und Subkategorien einer einfachen statistischen Analyse unterzogen. Dies geschieht im ersten Abschnitt auf einer allgemeinen Ebene und im zweiten Abschnitt vertieft bezüglich Assessment. Zusätzlich wird im dritten Abschnitt mit einer skalierenden Inhaltsanalyse die prozentuale Aufteilung der Sozialformen im Singunterricht ermittelt.

11.1 Quantitative Analyse der Kodierung im Überblick

Häufigkeit und Verteilung der Kodierungen (Mayring & Fenzl, 2019) ermöglichen Aussagen zu inhaltlichen Tendenzen im Datenmaterial. Aus diesem Grund wird an dieser Stelle eine quantitative Analyse bezüglich der Kodierhäufigkeiten in den drei Kategorien *PERSPEKTIVE DER LEHRKRAFT*, *ASSESSMENT* und *FOKUS AUF DAS GESCHLECHT* vorgenommen. Die Zahlen beziehen sich auf die vom Forschenden vorgenommenen Kodierungen nach Ablauf der Einigungskonferenz der Intercoder-Reliabilitätsprüfung (siehe Kapitel 10.3). In der Tabelle 19 sind die Zahlen aufgelistet, ergänzt mit der jeweiligen Dauer des Interviews.

Insgesamt wurden 1312 Kodierungen verteilt auf 48 Subkategorien vorgenommen, wobei die Kategorie *PERSPEKTIVE DER LEHRKRAFT* mit 579 Kodierungen verteilt auf 16 Subkategorien am meisten berücksichtigt wurde, gefolgt von 569 Kodierungen verteilt auf 23 Subkategorien in der Kategorie *ASSESSMENT* und 164 Kodierungen verteilt auf 9 Subkategorien in der Kategorie *FOKUS AUF DAS GESCHLECHT*. Für die drei Bereiche wurden Mittelwert und Standardabweichung berechnet.

Der Mittelwert der Kodierhäufigkeiten in der Kategorie *PERSPEKTIVE DER LEHRKRAFT* beträgt pro Interview 52.6 mit einer Standardabweichung von 13.0 ($M = 52.6$; $SD = 13.0$). Mit 82 Kodierungen liegen bei Herrn Buchser (02) besonders viele Aussagen zum Thema vor, wobei in diesem Fall aus statistischer Sicht von einem Ausreißer gesprochen wird, da der Mittelwert um mehr als das Doppelte der Standardabweichung überschritten wird. In

der Kategorie *ASSESSMENT* beträgt der Mittelwert der Kodierhäufigkeiten 51.7 mit einer Standardabweichung von 11.6 ($M = 51.7$; $SD = 11.6$). In diesem Bereich gibt es keine Ausreißer. In der Kategorie *FOKUS AUF DAS GESCHLECHT* wird ein Mittelwert von 14.9 und eine im Vergleich zu den anderen Bereichen relativ hohe Standardabweichung von 6.3 ($M = 14.9$; $SD = 6.3$) erreicht. Das Interview von Frau Derendinger (11) gilt als Ausreißer, denn auch hier übertrifft die Häufigkeit mit 30 Kodierungen den Mittelwert um das Doppelte der Standardabweichung.

Die 1312 Kodierungen verteilen sich mit einem Mittelwert von 119.3 ($M = 119.3$) auf die 11 Interviews, wobei die Interviews 02, 10 und 11 überdurchschnittlich viele Kodierungen aufweisen und die Interviews 06 und 07 überdurchschnittlich wenige. Die beiden letztgenannten Interviews waren in ihrer Länge mit 37 und 33 Minuten auch zeitlich sehr kurz, was die niederen Kodierhäufigkeiten erklären könnte. Im weiteren Verlauf muss sich zeigen, ob und in welcher Art die Interviews von Frau von Rohr (06) und Herrn Rothenburg (07) weiterhin auffallen. Bei den Interviews von Herrn Buchser (02), Herrn Leimbacher (10) und Frau Derendinger

	Bereich 1: Perspektive der Lehrkraft	Bereich 2: Assessment	Bereich 3: Fokus auf das Geschlecht	Summe	Inter- view- dauer
01 Hr. Brugger (52)	55	63	7	125	50
02 Hr. Buchser (55)	82	58	14	154	58
03 Fr. Fricker (49)	36	54	16	106	44
04 Hr. Hilfiker (35)	41	52	17	110	32
05 Fr. Basler (47)	53	56	18	127	49
06 Fr. von Rohr (37)	39	35	9	83	37
07 Hr. Rothenburg (49)	50	36	10	96	33
08 Hr. Neuendorf (67)	57	40	10	107	48
09 Hr. Berner (50)	48	44	15	107	55
10 Hr. Leimbacher (57)	52	71	18	141	53
11 Fr. Derendinger (28)	66	60	30	156	52
<i>Summe</i>	579	569	164	1312	511
<i>Mittelwert M</i>	52.6	51.7	14.9	119.3	46.5
<i>Standardabweichung SD</i>	13.0	11.6	6.3	23.5	8.9

TABELLE 19: Kodierhäufigkeiten der drei Analyserichtungen der QIA

(11) muss sich zudem zeigen, ob die hohen Kodierhäufigkeiten tatsächlich auf ein qualitativ reichhaltiges Material verweisen.

Die quantitative Auswertung des Kategoriensystems der drei Kategorien *PERSPEKTIVE DER LEHRKRAFT*, *ASSESSMENT* und *FOKUS AUF DAS GESCHLECHT* belegt mit der breiten Verteilung der Kodierhäufigkeiten auf die 11 Interviews und die drei Hauptkategorien die Resonanz der Thematik beim untersuchten Sample. Zudem deutet die teilweise hohe Standardabweichung auf ein heterogenes Untersuchungsmaterial hin.

11.2 Quantitative Analyse des Bereichs Assessment

In der vorliegenden Forschungsarbeit wird der Blick auf Assessment im Singunterricht gelegt. Die Interviewten wurden nicht darüber informiert, dass das Konzept von *assessment for learning*, *as learning* und *of learning* nach Earl (2013) als Perspektive auf die Interviews angelegt wird. Stattdessen wurde ausgehend vom Leitfaden (siehe Anhang) allgemein nach Förder- und Bewertungsmethoden gefragt, insbesondere weil dies bekannte Termini sind, im Gegensatz zu Assessment (vgl. 5.1). Weiter muss die Art und Weise der Kodierung berücksichtigt werden, da nicht nur Aus-

	Assessment			Summe
	for learning	as learning	of learning	
01 Hr. Brugger	52	5	6	63
02 Hr. Buchser	38	10	10	58
03 Fr. Fricker	34	7	13	54
04 Hr. Hilfiker	33	13	6	52
05 Fr. Basler	38	8	10	56
06 Fr. von Rohr	22	2	11	35
07 Hr. Rothenburg	23		13	36
08 Hr. Neuendorf	16	10	14	40
09 Hr. Berner	33	3	8	44
10 Hr. Leimbacher	46	10	15	71
11 Fr. Derendinger	45	6	9	60
Summe	380	74	115	569
Prozent	66.8	13.0	20.2	100

TABELLE 20: Kodierhäufigkeiten der Subkategorien von Assessment

sagen kodiert wurden, in denen intendierte Assessment-Vorgänge erwähnt wurden, sondern auch Aussagen mit impliziten Hinweisen, um die von Brunner (2014) beschriebene Beobachtung intuitiver Zugänge zu Fördermethoden aufzunehmen.

Aus diesem Grund wird sich erst in der qualitativen Analyse der Aussagen zeigen, wie die vorhandenen Kodierungen gewichtet werden können. Trotzdem gibt es Auffälligkeiten in den Zahlen, die auf quantitativer Ebene bereits Hinweise auf spezifische Subkategorien und einzelne Interviews geben.

Als Erstes fällt in den Kodierhäufigkeiten der Kategorie *ASSESSMENT* in Tabelle 20 auf, dass mit 66.8 % der Kodierungen ein überdurchschnittlicher Teil auf den Bereich AFL fällt, der Bereich AAL hingegen mit 13.0 % und AOL mit 20.2 % der Kodierungen nur wenig Beachtung erhält.

Im Bereich AFL wurden Unterrichtsmethoden kodiert, die Hinweise zu Förderpraktiken beinhalten. Dazu gehören sowohl diagnostische Elemente als auch direktes Feedback sowie allgemeine Fördermaßnahmen in frontalen Settings des Unterrichts. Im Sinne des Hinweises von Mantie (2019), der Musikunterricht *per se* als Assessment darstellt, wurden alle Aspekte allgemeiner Förderung, Zielvorstellungen, das Vorkommen von Feedback und Differenzierung kodiert. Bei der Analyse der Kodierhäufigkeiten von AFL in Tabelle 21 erschließt sich der Grund für die Dominanz dieses Bereichs. Mehr als die Hälfte der Kodierungen (52.4 %) sind bei den Fördermethoden zu finden. Zusätzlich enthält die Subkategorie *FÖRDERUNG* ihrerseits weitere sieben Subkategorien, die verschiedene Kompetenzbereiche, wie z. B. Atmung, tonale und rhythmische Schulung oder Auftrittskompetenz beinhalten. Es zeigt sich hier, dass die Lehrkräfte, die in der Studie befragt wurden, zumindest in quantitativer Hinsicht ein breites Repertoire an Fördermethoden beschreiben. An zweiter Stelle stehen die Kodierungen zu der Subkategorie *FEEDBACK* (16.8 %), gefolgt von der Subkategorie *ZIELVORSTELLUNGEN* (16.1 %). Damit kann belegt werden, dass Zielvorstellungen für die Entwicklung der stimmlichen Fähigkeiten durchgehend vorhanden sind und Feedback im Unterricht als relevantes Element wahrgenommen wird. Die Subkategorie *DIAGNOSE* erhielt mit 3.4 % der Kodierungen aber kaum Beachtung. Nur fünf der Interviewten, also weniger als die Hälfte, erwähnen Elemente von Diagnostik im Singunterricht. Bei acht der Interviewten wurde mit 5.8 % der Kodierungen zwar die Wahrnehmung von *DIFFERENZ* kodiert, aber nur sieben der Interviewten

geben an, im Unterricht zu differenzieren, ebenfalls mit sehr tiefen Kodierhäufigkeiten (5.2 %). Dies deckt sich sowohl mit der Studie von Brunner (2014), die belegt, dass Diagnosefähigkeiten im Singunterricht kaum vorhanden sind, als auch mit der Einschätzung von Schaumberger (2020a), der bezüglich Förderung und Beurteilung Entwicklungsbedarf moniert.

Neben der quantitativen Analyse der einzelnen Subkategorien wird beim Blick auf die Kodierhäufigkeiten der einzelnen Interviews zudem deutlich, dass die Streuung ziemlich groß ist ($SD = 9.3 \%$) und sich die drei Interviews 01, 10 und 11 durch hohe Kodierhäufigkeiten deutlich vom Rest abheben, was als Hinweis auf deutlicher ausgereifte fachliche (CK) und fachdidaktische (PCK) Kompetenzen hinweisen könnte.

Zusammenfassend zeigt sich, dass Fördermethoden, zielgerichtete Arbeit und Feedback in den Interviews ausführlich zur Sprache kommen, der individuelle Zugang zu Schülerinnen und Schülern durch diagnostische Unterrichtselemente, durch eine Wahrnehmung von Differenz und eine darauf abgestützte Differenzierung aber nur wenig entwickelt zu sein schei-

	Zielvorstellungen	Diagnose	Differenz	Differenzierung	Förderung	Feedback	Summe
01 Hr. Brugger	10	1	5	2	28	6	52
02 Hr. Buchser	3	3		1	24	7	38
03 Fr. Fricker	1	2	2		21	8	34
04 Hr. Hilfiker	7		1		19	6	33
05 Fr. Basler	4	5	3		20	6	38
06 Fr. von Rohr	5		1	1	13	2	22
07 Hr. Rothenburg	2		2		16	3	23
08 Hr. Neuendorf	4		3	2	4	3	16
09 Hr. Berner	7			5	16	5	33
10 Hr. Leimbacher	9			6	23	8	46
11 Fr. Derendinger	9	2	5	4	15	10	45
Summe	61	13	22	21	199	64	380
Prozent	16.1	3.4	5.8	5.5	52.4	16.8	100
Mittelwert der Anzahl Kodes pro Interview							$M = 27.2$
Standardabweichung der Anzahl Kodes pro Interview							$SD = 9.3$

TABELLE 21: Kodierhäufigkeiten von Assessment for learning

nen. Die hohe Standardabweichung deutet zusätzlich auf große Unterschiede im Sample hin. Bereits im ersten Teil der quantitativen Analyse sind die Interviews 10 und 11 durch überdurchschnittliche Kodierhäufigkeiten aufgefallen. Im Bereich Assessment sticht nun zusätzlich das Interview von Herrn Brugger (01) hervor.

Im Bereich AAL werden Unterrichtsmethoden kodiert, die auf Selbst-evaluation und Selbstständigkeit der Lernenden sowie Aspekte von Peer-feedback hinweisen. Mit nur 13 % aller Assessment-Kodierungen kann auf eine eher geringe Bedeutung des Themas im vorliegenden Sample geschlossen werden. Wie bereits im Bereich AFL erwähnt, müssen die Kodierhäufigkeiten von AAL in Tabelle 22 zusätzlich zurückhaltend interpretiert werden, da nicht nur intendiertes Handeln zu einer Kodierung geführt hat, sondern auch eine mögliche interpretierte intuitive Anwendung. Zudem ist im Vergleich zum Mittelwert von 5.8 Kodierungen ($M = 5.8$) die Standardabweichung mit 3.8 ($SD = 3.8$) sehr hoch und lenkt den Blick auf

	Unterstützende Umgebung	Selbst- einschätzung	Peerfeedback	Reflexion	Ziele setzen	Eigenverant- wortung	Summe
01 Hr. Brugger	3	2					5
02 Hr. Buchser	1	1		3		5	10
03 Fr. Fricker	3		1	1		2	7
04 Hr. Hilfiker	3	2		3		5	13
05 Fr. Basler		2		4		2	8
06 Fr. von Rohr						2	2
07 Hr. Rothenburg							0
08 Hr. Neuendorf	2	4	2	1		1	10
09 Hr. Berner	2					1	3
10 Hr. Leimbacher	2	2	1	1	1	3	10
11 Fr. Derendinger	3	2				1	6
Summe	19	15	4	13	1	22	74
Prozent	25.7	20.3	5.4	17.6	1.4	29.7	100.0
Mittelwert der Anzahl Kodes pro Interview							$M = 5.8$
Standardabweichung der Anzahl Kodes pro Interview							$SD = 3.8$

TABELLE 22: Kodierhäufigkeiten von Assessment as learning

die Extreme. Im Bereich der Minima fallen in der Tabelle 22 zwei fast leere Spalten und eine fast leere Zeile auf. Das Interview 06 und 07 weisen eine minimale Anzahl von zwei und null Kodierungen auf. Die beiden Interviews sind bereits durch insgesamt tiefe Kodierhäufigkeiten aufgefallen. Im Rahmen der Porträts im Kapitel 12.1 wird sich zeigen, ob die niedrigen Zahlen in AAL auf spezielle Eigenheiten der Interviewten hindeuten. Die Subkategorie *ZIELE SETZEN* wird nur einmal im Interview 10 kodiert. Wie es scheint, ist bei den Lehrkräften dieses Samples der Anspruch, dass Schülerinnen und Schüler sich eigene Ziele setzen, kaum vorhanden.

Insgesamt entsteht im Bereich AAL der Eindruck, dass zumindest in quantitativer Hinsicht die Lehrkräfte wenig Wert auf eigenverantwortliches Handeln und Reflexion legen. Auch Hinweise zu Peerfeedback sind sehr spärlich und weisen zusätzlich darauf hin, dass der Bereich AAL zumindest im vorliegenden Sample für den Singunterricht wenig entwickelt ist.

	Planung der Lernschritte	Transparenz der Interpretation	Formen der Bewertung	Performance	Summe
01 Hr. Brugger			3	3	6
02 Hr. Buchser	3	1	3	3	10
03 Fr. Fricker		2	7	4	13
04 Hr. Hilfiker		4	2		6
05 Fr. Basler		3	2	5	10
06 Fr. von Rohr		3	6	2	11
07 Hr. Rothenburg		3	4	6	13
08 Hr. Neuendorf		2	6	6	14
09 Hr. Berner	1	2	3	2	8
10 Hr. Leimbacher	1	2	4	8	15
11 Fr. Derendinger	1	3	5		9
<i>Summe</i>	6	25	45	39	115
<i>Prozent</i>	5.2	21.7	39.1	33.9	100
<i>Mittelwert der Anzahl Kodes pro Interview</i>					<i>M = 9.1</i>
<i>Standardabweichung der Anzahl Kodes pro Interview</i>					<i>SD = 2.7</i>

TABELLE 23: Kodierhäufigkeiten von *Assessment of learning*

Die Kodierungen von AOL decken den ganzen Bereich summativer Beurteilung ab (vgl. Tabelle 23). Das spezifische des Modells von Earl (2013) ist die Forderung, summative Assessments zu planen und transparent zu beurteilen. Weiter wurden in zwei Subkategorien mögliche Durchführungsformen und Aspekte von Auftrittskompetenz kodiert, damit die Breite der Bewertungspraxis eingefangen werden kann¹³⁹. Mit 20.2 % wurde dieser Bereich zwar öfter kodiert als AAL, aber doch deutlich weniger als AFL. In quantitativer Hinsicht zeigt der Mittelwert von 9.1 ($M = 9.1$) und die Standardabweichung von 2.7 ($SD = 2.7$) eine wenig auffällige Verteilung der Kodierhäufigkeiten. Einzig die Subkategorie *PLANUNG DER LERNSCHRITTE* wurde mit 5.2 % selten kodiert. Auch wenn die Lehrkräfte sich zu Transparenz in der Beurteilung äußern, weist das Fehlen von Kodierungen zum planvollen Vorgehen auf eine Diskrepanz mit dem Assessment-Konzept nach Earl (2013) hin und wirft Fragen bezüglich der Konzeption der Leistungsbeurteilung auf (vgl. 13.2).

In diesem Abschnitt wurde mit einer quantitativen Analyse die Kategorie *ASSESSMENT* untersucht. Ziel war die Suche erster Hinweise für die Beantwortung der übergeordneten Forschungsfrage. Gesucht sind also Hinweise zu Diagnose, Förderung und Beurteilung sowie Formen von Individualisierung. Es konnte gezeigt werden, dass die Lehrkräfte sich ausführlich zu Fördermethoden äußern, aber kaum von Diagnose und Differenzierung im Unterricht sprechen und damit auch wenig Aussagen zu individualisierenden Unterrichtsformen machen. Neben diesem Anhaltspunkt aus AFL zeigen sich auch in AAL geringe Kodierhäufigkeiten. So gibt es wenig Aussagen zu eigenverantwortlichem und reflektiertem Handeln und kaum Hinweise auf Peerfeedback. Beide Elemente sind konstituierende Elemente für einen selbstgesteuerten und individualisierten Unterricht. Die Lehrkräfte berichten in der Kategorie AOL ausführlich über Leistungsbeurteilung und die Form der Rückmeldung, was aber weitgehend fehlt, sind Kodierungen in der Kategorie *PLANUNG DER LERNSCHRITTE*. Dies kann als weiterer Hinweis auf einen Unterricht mit wenig individueller Passung eingeordnet werden, braucht aber wie alle anderen Hinweise eine vertiefte qualitative Analyse.

¹³⁹ Im LP21 werden Kompetenzziele für das Singen definiert, die auf Klangentwicklung und tonales Vorstellungsvermögen abzielen sowie die Auftrittskompetenz fördern.

11.3 Skalierende Strukturierung der Sozialformen im Singunterricht

Im Datenkorpus zeigen sich Bezüge zu unterschiedlichen Sozialformen. Einerseits wurden explizit Aufgabenstellungen für Gruppen oder Beurteilungssituationen in Einzelbetreuung beschrieben und andererseits wurde implizit bei einigen Interviewpartnerinnen und -partnern deutlich, dass Singunterricht vornehmlich im Rahmen einer frontalen Unterrichtssituation stattfindet. Die vorhandenen qualitativen Daten zu Sozialformen, die im Kategoriensystem verfügbar sind, werden nun im Rahmen einer skalierenden Strukturierung (vgl. 8.2) zu einer quantitativen Einschätzung des Einsatzes verschiedener Sozialformen verdichtet. Im vorliegenden Fall wird die Skalierung in Form einer prozentualen Aufteilung vorkommender Sozialformen pro Lehrkraft vorgenommen, dies mittels intervallskalierten Variablen.

Das Verfahren sieht vor, ausgehend vom Kategoriensystem Teile des Materials quantitativ einzuschätzen. Die Subkategorie *SOZIALFORMEN* in der Kategorie *PERSPEKTIVE DER LEHRKRAFT* dient als erster Anhaltspunkt. Die deduktiv erstellten Subkategorien *KLASSENUNTERRICHT*, *GRUPPENARBEIT* und *EINZELBETREUUNG* und der induktiv hergeleitete Spezialfall *EINZELBETREUUNG IM KLASSENUNTERRICHT* wurde immer dann kodiert, wenn die Studienteilneh-

	Klasse	<u>Geschlechtertrennung</u>		Gruppen-	Einzel-
		mono-	ko-	arbeit	betreuung
		edukativ	edukativ		(vor Klasse)
01 Hr. Brugger	12	0	0	1	4 (1)
02 Hr. Buchser	6	1	0	8	0 (2)
03 Fr. Fricker	3	1	0	3	2 (1)
04 Hr. Hilfiker	3	2	3	4	2 (0)
05 Fr. Basler	4	0	2	3	2 (1)
06 Fr. von Rohr	3	0	0	3	1 (0)
07 Hr. Rothenburg	5	0	0	3	0 (0)
08 Hr. Neuendorf	3	2	1	4	2 (0)
09 Hr. Berner	4	1	1	2	1 (3)
10 Hr. Leimbacher	5	6	2	2	6 (3)
11 Fr. Derendinger	6	5	1	5	3 (1)
<i>Summe</i>	<i>54</i>	<i>18</i>	<i>10</i>	<i>38</i>	<i>23</i>

TABELLE 24: Kodierhäufigkeiten von Sozialformen

menden Situationen beschreiben, die auf eine der Sozialformen zutreffen. Als zweiter Ausgangspunkt werden die Kodierhäufigkeiten von mono-educativen und koeducativen Unterrichtsformen der Kategorie *FOKUS AUF DAS GESCHLECHT* miteinbezogen. Diese wurden kodiert, sobald mono-educative Phasen im Unterricht thematisiert wurden. In der Tabelle 24 sind die entsprechenden Kodierhäufigkeiten aufgelistet. Insgesamt ergeben sich damit vier Variablen mit unterschiedlichen Sozialformen: *ARBEIT IN DER KLASSE*, *GESCHLECHTERTRENNUNG*, *GRUPPENARBEIT* und *EINZELBETREUUNG*. Die Hinweise aus der Subkategorie *EINZELBETREUUNG IM KLASSENUNTERRICHT* sind spärlich und werden nicht als einzelne Variable ausgewiesen, da der frontale Charakter der Unterrichtssituation überwiegt. Die Kodierhäufigkeiten beider Bereiche liefern einen zahlenbasierten Hinweis für die intervallbasierte Einschätzung.

Auf dieser Basis und weiteren Forschungsnotizen erfolgt in einem interpretativen Prozess die intervallbasierte Skalierung der Variablen auf eine Gesamtsumme von 100 %. Dazu wurden die kodierten Textstellen auf Unterrichtseinheiten übertragen und außer in einem Fall in 10 %-Schritte umgewandelt. Als Beleg der Vorgehensweise dienen die beiden Interviews 01 und 11. Im Interview mit Herrn Brugger (01) zeigt sich implizit an mehreren Stellen, dass sein Unterricht fast vollständig als *ARBEIT IN DER KLASSE* organisiert ist (z. B. 01:44 ff., 79 ff.) und dass dies nur zwecks einer summativen Beurteilung für eine *EINZELBETREUUNG* aufgebrochen wird (01:144 ff.). Frau Derendinger (11) beschreibt im Interview geschlechtergetrenntem Unterricht für jede zweite Musikstunde (z. B. 11:176 ff.) und für die summative Bewertung Gruppenarbeiten (11:301 ff.). Abbildung 11 zeigt die geschätzte Verteilung der Sozialformen.

Der Blick auf diese Verteilung zeigt bei allen Lehrkräften einen Anteil von mindestens 50 % Singunterricht im Klassenverband. Das gemeinsame Singen wird von allen gepflegt und zeigt sich somit als zentraler Teil des Singunterrichts. Zusätzlich zeigen sich zwei Gruppen von Lehrkräften. Einerseits wird bei vier Lehrkräften sichtbar, dass sie hauptsächlich in der frontalen Unterrichtssituation arbeiten und dies lediglich zwecks einer summativen Leistungsbeurteilung aufbrechen (01, 05, 06, 07). Frau von Rohr (06) und Herr Rothenburg (07) sind bereits in der quantitativen Analyse durch insgesamt wenige Kodierungen aufgefallen und fallen auch hier aufgrund der didaktischen Monokultur, wie Dyllick (2011) es formuliert, auf. Nicht so Herr Brugger (01), der seinen Unterricht bewusst als Chorsituation

gestaltet, und Frau Basler (05), die Schülerinnen und Schüler in der Sekundarstufe 1 stimmlich nicht herausfordern will. Beide Lehrkräfte argumentieren ihre Form des Unterrichts, was in den Porträts (vgl. 12.1) näher ausgeführt wird. In der zweiten Gruppe von Lehrkräften (02, 03, 04, 08, 09, 10, 11) zeigen sich zwei Tendenzen. Zum einen hat sich bei einigen dieser Lehrkräfte ein geschlechtergetrennter Unterricht etabliert und zum anderen konnten Aufgabenformate in Gruppen festgestellt werden, die auch in der summativen Beurteilung zur Anwendung kamen. Die Begründungen zur Entwicklung der beiden Formen werden ebenso in Kapitel 12.1 in den Porträts herausgearbeitet und in Kapitel 13 inhaltlich genauer analysiert.

In der erstellten Tabelle gibt es eine Form von Betreuung, die nicht in die Skalierung integriert wurde. Verschiedene Lehrkräfte berichten davon, Schülerinnen und Schüler außerhalb des Unterrichts individuell zu betreuen, vor allem dann, wenn sich akute Schwierigkeiten zeigen. Als Beispiel dient eine Aussage von Herrn Leimbacher (10):

„Und dann gibt es halt auch noch die Situation, dass ich sage, (flüstert) in der großen Pause, bleib noch schnell hier, oder je nachdem wie der Stundenplan ist, komm schnell, ich warte auf dich, oder Ende Schule, komm wir schauen schnell am Klavier schauen, wo steht deine Stimme, oder und dann allein, passiert plötzlich manchmal wieder oh wunder.“ (10:228 ff.)

Das Beispiel zeigt nicht nur das große Engagement von Herrn Leimbacher, sondern auch den Willen, Schwierigkeiten in einem geschützten Rahmen zu bearbeiten. Dies wird auch von Herrn Brugger (01:276 ff.) so gepflegt, der ansonsten nur im Chorsetting unterrichtet. Die Beispiele zeigen, dass individuelle Betreuung über den intendierten Einsatz als Methode im Unterricht auch am Rande des Unterrichts gepflegt werden kann und später in der Analyse berücksichtigt werden muss.

Die skalierende Strukturierung der Sozialformen zeigt auf der einen Seite Lehrkräfte, die den Singunterricht als frontale Unterrichtssituation gestalten, und auf der anderen Seite jene, die dieses Setting verlassen, um Gruppenarbeiten betreuen zu können oder monoedukative Sequenzen zu gestalten. Inhaltliche Begründungen und Intentionen müssen später auf einer qualitativen Ebene untersucht werden.

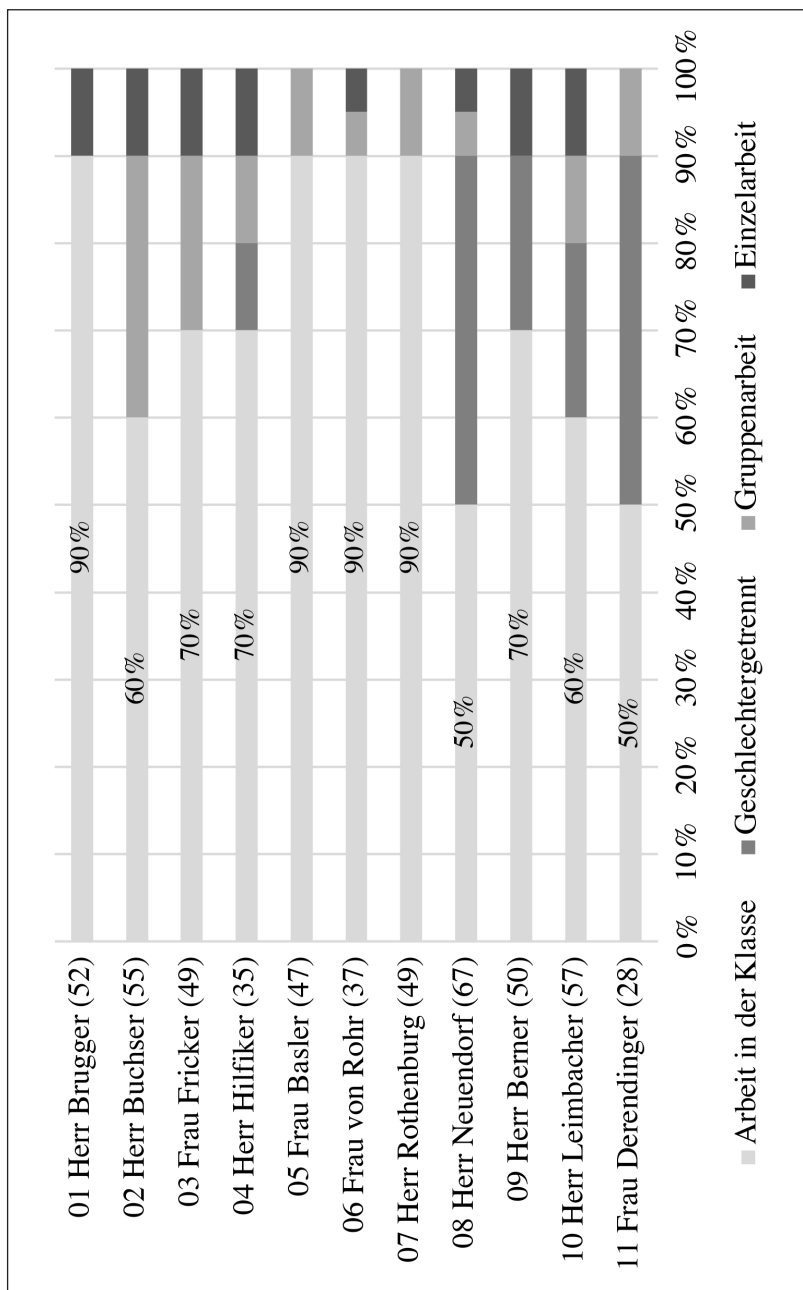


ABBILDUNG 11: Verteilung der Sozialformen im Unterricht des Samples

11.4 Fazit

Zu Beginn des Kapitels konnte gezeigt werden, dass die Mehrheit der Interviews in den Kategorien *PERSPEKTIVE DER LEHRKRAFT*, *ASSESSMENT* und *FOKUS AUF DAS GESCHLECHT* ein reichhaltiges Material bieten. Einzig die Interviews 06 und 07 zeigen niedrigere Kodierhäufigkeiten, im Gegensatz zu den Interviews 02, 10 und 11, die durch sehr hohe Kodierhäufigkeiten auffallen. Die Verteilung der Kodierungen auf die drei Bereiche zeigt zudem durch die zum Teil hohen Standardabweichungen, dass ein heterogenes Sample untersucht werden kann. Die Varianzoptimierung, wie sie Patton (2015) vorschlägt, konnte in der vorliegenden Arbeit somit als erfolgreich zur Generierung von Datenmaterial umgesetzt werden.

Inhaltlich fällt auf, dass in den Interviews zwar eine Vielfalt anhaltungen und methodischen Zugängen zu Förderung und Beurteilung sichtbar werden, das Thema Differenzierung oder Individualisierung sich aber wenig konturiert. Das mag am quantitativen Zugang liegen und wird später in der qualitativen Analyse genauer diskutiert. In der skalierenden Analyse der Sozialformen zeigte sich zudem die Tendenz, die frontale Situation zugunsten weiterer Settings aufzubrechen, was die Frage aufkommen lässt, in welcher Intention dies geschieht und ob Gruppenarbeiten und mono-educative Sequenzen im Singunterricht zu einer Individualisierung beitragen. Im nächsten Kapitel werden die einzelnen Interviewpartnerinnen und -partner in Porträts vorgestellt, um ihren Unterricht nachzuzeichnen sowie Gemeinsamkeiten und Unterschiede herauszuarbeiten.

12 Vertikale Achse: Perspektive der Lehrkräfte

„Zwar übergeht die qualitative Sozialforschung Teilnehmersichten nicht einfach (in Durkheims Tradition) als ‚vorwissenschaftlich‘, sie schließt vielmehr eng an sie an, lässt sie auch als Korrektiv eigener Vorannahmen gelten, nimmt sie insofern in sich auf und lässt auch Leser an ihnen teilhaben. Aber in der analytischen Distanzierung vom Datenmaterial ist die Teilnehmersicht nur ein Ausgangspunkt der Auseinandersetzung, eine Gelegenheit, die eigene Analyse abzusetzen. Es geht darum, Teilnehmersichten zuerst zu verstehen und dann durch sie ‚hindurchzusteigen‘.“ (Strübing et al., 2018, S. 87)

Nach der quantitativen Einordnung des Datenmaterials in Kapitel 11 wird in diesem Kapitel das von Strübing et al. (2018) im obigen Zitat beschriebene Vorgehen aufgenommen, indem die Teilnehmersichten, im vorliegenden Fall die individuellen Sichtweisen der Interviewpartnerinnen und -partner, in Kurzporträts aufgearbeitet werden. Der Blick auf die diversen Unterrichtskulturen und das Nachzeichnen persönlicher Vorstellungen zeigen analog zu Kapitel 3 mit dem autoethnografischen Exkurs individuelle Strategien, die von den teilnehmenden Lehrkräften mit Ausnahmen positiv bewertet werden. Die Porträts würdigen deshalb auf der einen Seite das selbst deklarierte Gelingen und den Umgang mit den Herausforderungen des Singunterrichts der elf Lehrkräfte und konturieren zugleich durch den Forschungsfokus die lokalen Singkulturen bezüglich individueller Zielorientierungen, Diagnose, Förderung und Beurteilung. Wie in Kapitel 8.2 beschrieben, besteht über das Deskriptive hinaus die Erwartung, mögliche Typisierungen (Kuckartz, 2016) zu finden und diese aufgrund der Profilbildung beim Sampling (vgl. 9.1) einzuordnen.

Im Gegensatz zu einem rein hermeneutischen Vorgehen wird im vorliegenden Fall keine Einzelfallanalyse (Kelle & Kluge, 2010) vorgenommen, sondern es werden bestimmte für das Verständnis der Untersuchung relevante Ausschnitte auf einer deskriptiven Ebene zusammengefasst. Das entwickelte Kategoriensystem dient als Ausgangspunkt der Porträts, indem es inhaltlich fokussiert, legitimiert durch das methodische Vorgehen der QIA (vgl. 10.2) und der abschließenden Intercoder-Reliabilitätsprüfung (vgl. 10.3). Wie bereits in Kapitel 9 dargelegt, dient der Ausschluss von

Novizen im Sample dazu, ungefestigte Vorstellungen und wenig etablierte Vorgehensweisen auszuschließen. So bringen in der vorliegenden Studie die Interviewten im Minimum sechs, im Maximum 31 Jahre Erfahrung mit (vgl. Tabelle 14). Die größere Erfahrung ermöglicht es den interviewten Lehrkräften auf der Basis einer breiteren Auswahl von erlebten Situationen zu berichten und möglicherweise weiter entwickelte Begründungsmuster wiederzugeben (Gamoran Sherin & van Es, 2009). Das Kategoriensystem lenkt zwar den Blick insgesamt auf die drei Analyserichtungen *Perspektive der Lehrkraft*, *Fokus auf das Geschlecht* und *Assessment-Methoden*, lässt aber mit der Fokussierung auf einzelne Personen genügend Spielraum für das Auftauchen unerwarteter Fragestellungen (Strübing et al., 2018, S. 86). Als Resultat entstehen so elf auf Singunterricht fokussierte Porträts von Musiklehrkräften in der Sekundarstufe 1.

Im Anschluss an das Verfassen der Porträts im ersten Abschnitt und der damit verbundenen Korrektur eigener Vorannahmen, wie von Strübing et al. gefordert, werden im zweiten Abschnitt auffällige Merkmale aus den Unterrichtsbeschreibungen zu Typiken (Kelle & Kluge, 2010) verdichtet (vgl. 12.2). Dabei steht insbesondere die Frage nach Zielen (vgl. Teilfrage 1) und die damit verbundenen impliziten und expliziten Zielvorstellungen (Niessen, 2014) im Vordergrund. Es wird sich zeigen, dass eine Polarität in der Zielorientierung für das Singen besteht, einerseits Richtung stimmlicher Entwicklung, andererseits in Richtung Persönlichkeitsentwicklung. Zudem berichten einige Lehrkräfte von Widerstand seitens der Schülerinnen und Schüler, andere betonen den empathischen Umgang mit Schülerinnen und Schülern. Abschließend wird im dritten Abschnitt (vgl. 12.3) ein erstes Fazit in Bezug auf die Forschungsfragen gezogen.

12.1 Porträts der Lehrkräfte

Ziel der Porträts ist die Darstellung wesentlicher Überzeugungen und Ziele der Lehrkräfte im Licht der Forschungsfragen und die Suche nach Besonderheiten. Dies erfolgt für jede Lehrkraft mit einem vertikalen Durchgang durch das Kategoriensystem auf der Grundlage der Kodierungen und Summaries. Ebenfalls einbezogen werden die erstellten Notizen bezüglich des

ersten Eindrucks und der quantitativen Einschätzung zu jeder Lehrkraft¹⁴⁰. In den Porträts wird unter anderem beschrieben, welchen Zugang die Lehrkräfte zur eigenen Singstimme haben, welche Ziele sie mit dem Singen in der Schule verfolgen und von welchen Schwerpunkten ihres Singunterrichts sie berichten. Ergänzend dazu wird herausgearbeitet, welche Bezüge die Ausbildung und möglicherweise das Geschlecht zum beschriebenen Unterricht haben. Zusätzlich wird bei einigen Porträts eine soziologische Perspektive eingenommen, da auch die Schulkultur und die Art der Schülerinnen und Schüler als Gegenüber der Lehrkräfte einen Einfluss auf das Funktionieren des lokalen Systems haben. In einer abschließenden Verdichtung werden die Porträts mit prägnanten Überschriften versehen, um eine möglichst lebendige Essenz zu kreieren. Kraemer (2007) schreibt dazu: „Erfahrungsgemäß gibt es nicht den Musiklehrer, sondern vielmehr Persönlichkeiten mit unterschiedlichen Vorlieben, Interessen und Befähigungen.“ (S. 113) Diese Aussage bestätigt sich in den folgenden Porträts auch für das vorliegende Sample.

Der Chorleiter bleibt seiner Leidenschaft treu

Herr Brugger (01)¹⁴¹ ist 52 Jahre alt und arbeitet bereits mehr als 20 Jahre als Musiklehrer an einer progymnasialen Schule¹⁴². Nach seiner Ausbildung zum Primarlehrer (BA) hat er Schulmusik II mit Hauptfach Chorleitung (MA) studiert. Neben dem regulären Musikunterricht betreut er einen Schulchor, der in der Schule einen wichtigen kulturellen Beitrag leistet und mit dem er regelmäßig ins Ausland reist. Außerhalb der Schule ist er für verschiedene Chöre verantwortlich.

Herr Brugger sieht sich selbst als gesangliches Vorbild (01:33 ff.) und beschreibt im Wesentlichen drei Ziele für den Singunterricht: Singen soll für Schülerinnen und Schüler etwas Selbstverständliches sein, ein Gemeinschaftserlebnis und ein entspannendes Moment im Schulbetrieb (01:538 ff.). Seinen Singunterricht gestaltet er bewusst und überwiegend als frontale Unterrichtssituation und das Einsingen zu Beginn jeder

¹⁴⁰ Diese Daten wurden ebenfalls in die MAXQDA-Datei eingepflegt und sind über einen Link im Anhang verfügbar.

¹⁴¹ In den Porträts wird nur zu Beginn auf die Interviewnummer verwiesen.

¹⁴² Dies entspricht in Deutschland der gymnasialen Stufe, vgl. auch Kapitel 6.

Stunde ist mittlerweile Ritual und wichtiges Instrument der klassendynamischen Steuerung, um neben dem Aufwärmen der Stimme mit wenigen gleichbleibenden Übungen und dem damit verbundenen Aufbau der stimmlichen Fertigkeiten den Kontakt zur Klasse herzustellen und ein gutes Klima zu generieren (z. B. 01:68 ff.). Ein weiteres Element ist das Repertoiresingen, wo mit fließenden Übergängen zwischen den Liedern ein Flow entstehen soll, um Energie aufzubauen. Innerhalb des Klassenverbands werden vokale Defizite kaum bearbeitet, da Herr Brugger auf keinen Fall Schülerinnen und Schüler bloßstellen möchte (01:257 f.; 270 ff.). Abgesehen von allgemeinem Feedback (01:167 ff.) gibt es deshalb persönliches Feedback nur in Bewertungssituationen, die er in einer 1-zu-1-Situation (01:148 f.) durchführt. Ebenso wird nur in diesem Feedbackgespräch auf die Mutation der Jungen eingegangen, ein proaktiver Umgang im Klassenunterricht wird aus oben genannten Gründen vermieden.

Insgesamt hat Herr Brugger eine klare Idee, was er wie fördern will und beharrt auf den Zielen, die er hat, weil „es mir wichtig ist, dass sie das checken“ (01:218 f.). Die vokale Förderung steht dabei im Vordergrund und passiert größtenteils in der Anlage einer Chorprobe.

Der Schauspieler entwickelt sich vom Animator zum Coach

Herr Buchser (02) ist 55 Jahre alt und unterrichtet das Fach Musik seit 12 Jahren auf den tieferen Leistungsniveaus. Nach seiner Ausbildung zum Schauspieler und Reallehrer (MA) hat er zuerst während 15 Jahren Klassen des tiefsten Leistungsniveaus als Generalist in allen Fächern, unter anderem auch in Musik, unterrichtet. Seine vokalen und instrumentalen Kenntnisse auf mehreren Instrumenten eignete er sich im Rahmen seiner Ausbildung und autodidaktisch an. Außerhalb der Schule verfolgt er verschiedene musikalische Projekte.

Herr Buchser macht die Erfahrung, dass Schülerinnen und Schüler „eigentlich gerne singen und dass sie [...] durch die Erfahrung mit mir auch gut singen können“ (02:8 ff.). Er erachtet damit die Motivation und die Qualität des Gesangs als grundsätzlich positiv und macht dies fest an seiner eigenen Vorbildwirkung. In den Klassen startet er meist im Plenum mit fröhlichen Liedern als Begrüßungsritual (02:7 f., 53 f.). Trotz mitreißendem Unterrichtsstil treten disziplinarische Schwierigkeiten auf.

Um solchen Unterrichtssituationen und unmotivierten Schülerinnen und Schülern zu begegnen, hat er einen Unterrichtsstil mit einem hohen Anteil an Gruppenunterricht entwickelt. Damit möchte Herr Buchser einerseits die Selbstverantwortung der Schülerinnen und Schüler stärken (02:113 ff.) und andererseits eine bessere Balance für seinen Energiehaushalt erreichen (02:405 f.). Die Aufgabenstellungen in Gruppen führen meist zu einer Performance vor der Klasse. Feedback ist dabei ein wichtiges methodisches Werkzeug und wird einzeln, in der Gruppe, aber auch vor der Klasse gegeben (02:168 ff.). Er betont vor allem die verstärkende Wirkung von Lob: „[...] wenn du sie lobst, gibt es immer noch eine Welle, auf der man noch etwas machen kann.“ (02:109 f.) Zusätzlich hat er auch eine Reihe von reflektierenden Aufgabenstellungen entwickelt, die innerhalb eines längeren Prozesses intrinsische Motivation erzeugen sollen. Er ist sich aber bewusst, dass dazu ein langer Atem notwendig ist (02:415 ff.).

Als ausgebildeter Pädagoge hat er einen Weg gefunden, um den Schülerinnen und Schülern des tiefsten Leistungsniveaus durch das Singen in Gruppen und in der Klasse Selbstmotivation zu initiieren und Auftrittskompetenz zu vermitteln. Das Resultat ist ein Musikunterricht, in dem viel gesungen wird und die Entwicklung der Persönlichkeit im Fokus steht.

Die langjährige Schulmusikerin

Frau Fricker (03) ist 49 Jahre alt und hat mehr als 15 Jahre Erfahrung als Musiklehrerin auf der progymnasialen Stufe. Sie absolvierte eine Ausbildung zur Lehrkraft für die tieferen Leistungsniveaus der Sekundarstufe I (MA) und eine Zusatzausbildung für Schulmusik 1 (BA). Außerdem ist sie aktive Sängerin einer Bluegrass-Band und seit einiger Zeit in Ausbildung zur Schulleiterin.

Frau Fricker betrachtet eine empathische Beziehung zu Schülerinnen und Schülern als zentrale Grundlage für einen erfolgreichen Musikunterricht (03:309 f.). Sie animiert ihre Klassen, indem sie selbst als Vorbild mitsingt (03:79 f.) und gute Laune verbreitet (03:308). Ihr Unterrichtsziel kann vor allem an dem im letzten Unterrichtsjahr stattfindenden Songwriting-Projekt festgemacht werden, die einerseits die entwickelte musikalische Kompetenz und andererseits die Auftrittskompetenz aufzeigen

soll. Die vokale Förderung und das Training der Auftrittskompetenz wird mit regelmäßigen Gruppenarbeiten erreicht (03:172 ff.).

Allerdings nennt Frau Fricker zwei Elemente, welche die angestrebten Ziele ihres Unterrichts infrage stellen. Sie stellt fest, „dass eigentlich eine Schulklasse nie rein tönt“ und findet es „fürchterlich“, dass Schülerinnen und Schüler immer falsch singen (03:27 ff.). Bezüglich der Pubertät beobachtet sie eine starke Hemmung, die eine Entwicklung teilweise verunmöglicht (03:400 ff.). Sie hat aber in ihrer Karriere als Musiklehrerin mehr Nachsicht entwickelt, auch aufgrund der Erfahrung mit ihren eigenen Kindern (03:474 ff.), und hat den Wunsch, im Unterricht mehr individualisieren zu können (03:67 f.). Sie hat dazu bereits einige Ideen ausprobiert, aber nicht konzeptuell in ihren Unterricht eingebaut. Sie nennt z. B. das diagnostische Element, nach den Ferien den individuellen Stand ihrer Schülerinnen und Schüler beim Singen zu erfassen (03:37 ff.), dies aber nicht flächendeckend, sondern nur punktuell. Weiter hat sie in einem Schuljahr einen Teil des Unterrichts monoedukativ unterrichtet (03:26 ff.) und in ihrer Schule Kojen für Gruppenarbeiten eingerichtet (03:192 f.).

Zusammenfassend kann bei Frau Fricker eine gewisse Ambivalenz beobachtet werden: Einerseits fördert sie Schülerinnen und Schüler durch einen empathischen Zugang, fokussiert auf Auftrittskompetenz und setzt punktuell auch diagnostische und individualisierende Elemente um; sie nennt aber andererseits immer wieder Widrigkeiten, die sich negativ in ihrem Unterricht bemerkbar machen, wie z. B. die Aussage, dass Schülerinnen und Schüler immer unrein singen und in der Pubertät gehemmt sind. Dies lässt auf eine gewisse Unzufriedenheit schließen, die vielleicht auch zu ihrer beruflichen Neuorientierung hin zur Schulleiterin geführt hat.

Der Schlagzeuger als Energielieferant

Herr Hilfiker (04) ist 35 Jahre alt und hat sechs Jahre Erfahrung als Musiklehrer an den tieferen Leistungsniveaus der Sekundarstufe 1. Er ist ausgebildeter Schlagzeuger (MA) und als Bandleader in der Jazz-Szene tätig. Er hat weder eine schulmusikalische noch eine pädagogische Ausbildung.

Herr Hilfiker beschreibt sich selbst als stabilen Menschen mit Linie, der nach anfänglich schlaflosen Nächten alles „on the job“ (04:510) gelernt hat. Er arbeitet hauptsächlich mit Liedervorschlägen seiner Schülerinnen und Schüler (04:53 ff.), versucht dabei aber auf alle einzugehen. Er sagt: „im Teenager-Alter ist Musik das ein und alles gewesen, und vor allem welche Musik findest du nicht cool, ist gerade so wichtig gewesen, wie was du cool findest.“ (04:83 ff.) Singen ist aus seiner Sicht Talent-sache (04:223 f.), was dazu führt, dass er keine Bewertung der Intonation vornimmt. „Keine Ahnung wie man das lernen kann noch in diesem Alter, oder irgendwie es ist auch für uns schwierig, oder?“ (07:285 f.) Er stellt infrage, ob Singen gelernt werden kann. Ebenfalls führt er auch seine schlechten Erfahrungen am Gymnasium bezüglich Bewertung an (04:299 ff.) und zieht zum Vergleich seine Bandmitglieder hinzu, die durchgehend nicht singen können, aber hervorragende Musiker sind (07:208 ff.). Herr Hilfiker reduziert gutes Singen damit auf korrekte Intonation mit professionellem Anspruch und ist der Überzeugung, Singen sei nicht lernbar.

Diese Grundannahmen führen bei Herrn Hilfiker zu einem Unterricht, der das Singen auf zwei Ebenen fördert, zum einen über die Arbeit am Text und zum anderen über die Generierung positiver Energie. Die Wichtigkeit guter Aussprache betont Herr Hilfiker im Interview mehrere Male (07:90 ff., 366 ff.) und legt auch in der Vorbereitung von Bewertungssituationen den Fokus auf gute Mundbewegungen und Artikulation (07:90 ff., 248 ff.). Die Singmotivation fördert er zum einen durch die eigene Vorbildwirkung, indem er dynamisch singt (07:365 ff.) und zum anderen steuert er mit kompetitiven Elementen die Klassendynamik (z. B. 07:390 ff.). Das Wort Energie wird mehrere Male sprachlich differenziert im positiven Sinn verwendet: mehr Energie, Energie kommt, Energie entwickeln, energiegeladener und Energie kanalisieren. Sein Ziel hat er dann erreicht, wenn die Schülerinnen und Schüler in einen Flow kommen und er als Lehrkraft überflüssig wird (07:448).

Herr Hilfiker hat trotz fehlender Ausbildung als Schulmusiker und ohne vokale Vorbildung einen Weg gefunden, das Singen in seinem Musikunterricht zu etablieren. Dabei fokussiert er auf die Musik der Jugendlichen und fördert auf der sprachlichen Ebene. Musikalische oder vokale Ziele werden nicht angestrebt, aber wahrgenommen, wenn sie natürlicherweise vorhanden sind (07:422 ff.).

Die bewahrende Sängerin in der Waldorfschule

Frau Basler (05) ist 47 Jahre alt und unterrichtet seit 14 Jahren alle Schulstufen von der Primar- bis zur Sekundarstufe 2 an einer Waldorfschule. Sie ist ausgebildete Sängerin (MA) und hat die Ausbildung für Schulumusik 2 (MA) absolviert. Weiter hat sie eine Ausbildung für Kirchenmusik mit Schwerpunkt Chorleitung. In der Schule betreut sie verschiedene musikalische Projekte und außerhalb der Schule leitet sie verschiedene Chöre und ist solistisch tätig.

Der Musikunterricht beschränkt sich an dieser Schule während der Sekundarstufe 1 neben zwei Lektionen Orchester auf 35–40 Minuten Singen (05:55 ff.). Damit wird dem Umstand Rechnung getragen, dass die Jugendlichen in der Zeit der Pubertät eine Phase der Zurückhaltung durchlaufen. Frau Basler beschreibt dies so: „Ich erlebe teilweise Klassen, die [...] fast aufhören zu singen, [...] die Jungen verstummen wegen Stimmwechsel, die Mädchen verstummen wegen Schüchternheit, also auch eben dieser Wechsel.“ (05:79 ff.) Die Jugendlichen sollen gleichsam durch die Zeit der Mutation und der persönlichen Entwicklung durchgetragen werden (05:62 ff.), dies im Wissen, dass vor dieser Zeit in der Primarstufe und anschließend in der Sekundarstufe 2 intensiv mit der Stimme gearbeitet wird (05:165 ff.). Frau Basler führt ihren Singunterricht hauptsächlich im Klassenverband durch (05:233) und legt trotz der schwierigen Phase großen Wert auf die Singqualität (05:130 ff.) und das musikalische Gestalten (05:309 ff.). Sie benutzt das Einsingen als Einstiegsritual (05:108 ff.) und sucht sich aufgrund ihrer Biografie neben einigen Klassikern aus der Popmusik bewusst klassische Literatur aus (05:582 ff.). Sie sieht sich vor allem als Vorbild für die Mädchen, die sie mit ihrer Stimme durchträgt (05:151 ff.), und wünscht sich für die Arbeit mit den Jungen eine männliche Musiklehrkraft (05:250 ff.). Sie ist dann frustriert, wenn „dieses Gefühl, den Schüler nicht zu erreichen“ (05:561 f.), dominiert und sie nicht eruieren kann, woran es liegt.

Im Singunterricht von Frau Basler werden die Stimmen vor dem Hintergrund des Stimmwechsels ihrer natürlichen Entwicklung überlassen, gleichsam durchgetragen vom aktiven Singen in der Primarstufe hin zum aktiven Singen in der Sekundarstufe 2. Für die Mädchen sieht sie sich selbst in der tragenden Rolle als Vorbild, für die Jungen würde sie gerne auf die Vorbildwirkung eines Mannes zählen.

Die Konzertsängerin in der Beobachterrolle

Frau von Rohr (06) ist 37 Jahre alt, hat auf allen Leistungsniveaus der Sekundarstufe 1 Erfahrung gesammelt und ist seit 10 Jahren auf der gymnasialen Stufe tätig. Sie ist ausgebildete Konzertsängerin (MA), hat aber weder eine pädagogische noch eine schulmusikalische Ausbildung genossen. Neben ihrer Tätigkeit als Lehrkraft tritt sie als Sängerin auf und betreut Gesangschülerinnen und -schüler.

Frau von Rohr setzt sich zum Ziel, mindestens 50% des Unterrichts mit Singen zu gestalten (06:28 f.). Dies passiert hauptsächlich im Klassenverband, außer für die Bewertung, die in Einzel- oder Partnerarbeit durchgeführt wird. Im Interview werden zwei Motivationsfaktoren für das Singen benannt. Zum einen arbeitet Frau von Rohr mit einem Belohnungssystem innerhalb der frontalen Unterrichtssituation. Wenn lauter gesungen werden soll, müssen die Schülerinnen und Schüler aufstehen, führt dies zum Erfolg, dürfen sie sich als Belohnung wieder setzen (06:111 ff.). Zum anderen wird versucht, über eine positive Leistungsbewertung, also gute Noten, die Motivation hochzuhalten (06:347 ff.). Als weiteres Motivationsmoment wird das Mitsingen von Songs auf der Plattform YouTube genannt, das vor allem die Jungen zusätzlich ins Boot holt (06:136 ff.). Wenn das gemeinsame Singen nicht funktioniert, bricht sie die Singsituation ab, obwohl ihr bewusst ist, dass sie damit die Willigen bestraft (06:120 ff.).

Abgesehen von basalen Gesangsübungen nennt Frau von Rohr keine Förderelemente. So wird an einigen Stellen deutlich, dass hauptsächlich die Beobachtung der vokalen Entwicklung im Zentrum steht. Sie berichtet z. B. davon, dass gute Schülerinnen bereits gut sind, wenn sie in ihren Unterricht eintreten (06:209 ff.). Bei Jungen in der Mutation betont sie, auf einen Abzug in der Bewertung zu verzichten (06:326 f.). Zudem rät sie ihren Schülerinnen und Schülern, auf das Vorsingen zu Karaoke-Playbacks zu verzichten, da diese keine Übung darin haben (06:291 ff.). Bei keinem der drei Punkte erwähnt Frau von Rohr methodische Werkzeuge, die zu einer Verbesserung der gesanglichen Leistung führen oder zu einer Überwindung von Hürden beitragen können.

Insgesamt gestaltet Frau von Rohr einen Singunterricht, der auf die vorhandenen Fähigkeiten der Schülerinnen und Schüler aufbaut. Mit basalen Motivationstechniken soll der Eindruck verhindert werden, nicht

oder schlecht singen zu können (06:349 f.), damit auch nach der obligatorischen Schulzeit weiterhin gesungen wird (06:391 ff.).

Der musikalische Generalist mit Hang zur Frustration

Herr Rothenburg (07) ist 49 Jahre alt und arbeitet seit mehr als 20 Jahren als Sekundarlehrer des mittleren Leistungsniveaus. Er ist als Generalist stufenadäquat ausgebildet (MA) und hat sich autodidaktisch ein hohes musikalisches Niveau erarbeitet. Neben seiner Funktion als Klassenlehrer unterrichtet er einige Klassen in Musik. In seiner Freizeit singt er in einem Chor und vertritt dort immer wieder den Chorleiter.

Herr Rothenburg singt in seinem Unterricht viel, nicht nur im Fach Musik, sondern auch im Fach Französisch (07:12 ff.). Er singt dabei mit den Schülerinnen und Schülern teilweise zweistimmig, macht Einsingübungen, die er aus seiner Chorerfahrung kennt (07:47 ff.), und stellt sie in Kanons in Gruppen verteilt im Raum auf (07:39 ff.). Als Ziel für den Musikunterricht beschreibt Herr Rothenburg den Anspruch, auf der Schlussfeier Lieder singen zu können (07:74 f.). Der Fokus des Unterrichts liegt nicht in der Stärkung des Gesangs, sondern von Beginn weg in der Bildung einer Begleitband (07:76 ff.). Es werden zwar viele Lieder gelernt „und manchmal wird der Klang auch ein wenig größer“, aber der Gesang entwickelt sich nicht, sondern „bleibt so ein wenig gleich“ (07:323 ff.). Begeisterung für das Singen zeigt er dort, wo in einer Klasse aufgrund von motivierten Zugpferden mehrstimmig gesungen werden kann (07:339 ff.). In der zweiten Hälfte des Interviews überlagert zunehmend Frustration über die fehlende vokale Entwicklung den enthusiastischen Eindruck der ersten Hälfte. So betont Herr Rothenburg, kein „Sologesangslehrer“ (07:290) zu sein und gibt keine Hinweise, dass er sich als gesangliches Vorbild betrachtet. Im Gegenteil gibt es in seinem Unterricht sogar Zeiten, wo er je nach persönlicher Verfassung (07:479) oder wenn in den Klassen zu viel geschwätzt wird (07:486 ff.) nicht singen mag, oder das gemeinsame Singen sogar abbricht. Er beobachtet großes Schambewusstsein (07:284 ff., 375 ff.) und gibt auch nur im Gruppenunterricht Feedback, damit eine größere Vertrautheit vorhanden ist und sich Schülerinnen und Schüler nicht genieren (07:138 ff.). Er ist zudem frustriert, wenn sie nur dastehen, leise singen und wenig Motivation zeigen (07:320 ff.). Verantwortlich für diese Entwicklung macht er

neben Scham auch die zunehmend schnellere Rotation des Repertoires (07:421 ff.).

Insgesamt entsteht der Eindruck einer rührigen Lehrkraft, die zwar viel mit den Schülerinnen und Schülern singt, aber eine zunehmende Frustration erkennen lässt, weil sie wenig Entwicklung sieht. Die musikalischen Fähigkeiten von Herrn Rothenburg kann er dagegen im Klassenmusizieren und in Klassen, in denen vokale Fähigkeiten und Motivation natürlicherweise vorhanden sind, ausleben.

Der musiktherapierende Gesangslehrer

Herr Neuendorf (08) ist der älteste Teilnehmer der Interviewstudie und kann mehr als 30 Jahre Unterrichtserfahrung auf allen Leistungsniveaus der Sekundarstufe 1 vorweisen. Zusätzlich zu seiner Gesangsausbildung (MA) hat er die Schauspielschule besucht und ein Kirchenmusikdiplom (MA) erworben. Als zweites Standbein betreut er neben seinen Schülerinnen und Schülern Chöre auf professionellem Niveau und ist musiktherapeutisch tätig.

Herr Neuendorf strebt in seinem Musikunterricht zuallererst eine hohe Qualität der Beziehung zu seinen Schülerinnen und Schülern an. Zuerst betont er als zentral wichtiges Mittel das Schaffen einer positiven Atmosphäre, in der „sich das Kind nicht [...] gehemmt fühlt oder bloßgestellt wird, dass es der Umgebung Klasse vertrauen kann, sich zu äußern“ (08:73 f.). Ein besonders hohes Maß an Vertrauen brauchen aus seiner Sicht die Jungen, wobei er insbesondere deren Widerstandspotenzial besonders aufmerksam behandelt (08:133 ff.). Zu diesem Zweck separiert er Mädchen und Jungen und benutzt geschlechtsspezifische Lieder (08:110 ff.). Auf dieser Grundlage startet er stimmbildnerische Aktivitäten. Er meint: „Ich habe als Musiklehrer die Erfahrung gemacht, dass wenn ich sie dort abholen kann, wo sie stehen, habe ich sehr viele Möglichkeiten, [...] an das Material ranzukommen und zu singen.“ (08:53 ff.) An dieser Stelle betont er, nicht als Schulmusiker zu arbeiten, sondern in erster Linie Sänger zu sein und insbesondere auch auf Stimmungen zu achten (08:58 ff.). Er findet es „voll faszinierend, wenn sie sich selber

auch hören“ (08:62 f.) und bringt dabei das Wort „personare“¹⁴³ (08:63) ins Spiel, um zu zeigen, dass in seiner Arbeit die Schwingungen der Stimme auch mit der Entwicklung der Persönlichkeit verknüpft sind. Dazu wartet er immer wieder geduldig auf das Momentum (08:471 ff.), wo Ideen aus der Klasse den Unterricht beflügeln und er in die Rolle des Coaches schlüpfen kann. Die Sorgfalt, mit der er Jugendlichen begegnet, zeigt sich auch im Feedback zum Solo-Gesangsvortrag im letzten Schuljahr, wo er mit Schülerinnen und Schülern die Bewertung in einem Feedbackgespräch auf Augenhöhe aushandelt (08:363 ff.).

Insgesamt scheint Herr Neuendorf ein ausgesprochen empathischer Musiklehrer zu sein, der seine musikalischen Ziele immer mit Blick auf das Schaffen einer lernförderlichen Atmosphäre justiert und dabei den ganzen Menschen im Blick behält.

Der Mundartrock als Motivationstrainer

Herr Berner (09) ist 50 Jahre alt und hat 10 Jahre Erfahrung als Musiklehrer der unteren Leistungsniveaus. Ursprünglich ist er ausgebildeter Primarlehrer (BA) und Instrumentallehrer für Akkordeon (BA). Momentan befindet er sich in der Ausbildung zur Lehrkraft Sekundarstufe 1 für die Fächer Kunst und Musik (MA). In der Schule betreut er einen Schulchor. Außerhalb der Schule ist er als Sänger und Texter in einer Mundart-Band aktiv und verfolgt verschiedene Kleinkunstprojekte.

Obwohl Herr Berner regelmäßig als Sänger auftritt, bezeichnet er sich selbst nicht als Sänger (09:658 ff.) und fokussiert in seinen Zielen für den Musikunterricht auch nicht auf vokale Fortschritte, sondern auf Körperhaltung und Auftrittskompetenz (09:55 ff., 148 ff.). Der Unterricht findet vor allem im Klassenverband statt, außer in Bewertungssituationen, wo geschlechtergetrennt gearbeitet wird (09:216 ff.). Innerhalb der Klasse werden Feedbacks persönlich an Schülerinnen und Schüler gerichtet, was aus seiner Sicht gut akzeptiert wird. Zentral wichtig ist stetiges Ermutigen und „die richtigen Worte [zu] finden“ (09:279 ff., 482 ff.). In seinen Klassen macht Herr Berner einen erschreckend hohen Anteil an

¹⁴³ Der Ursprung des Wortes – *personare* = *hindurchklingen* – verknüpft das Klingen der Stimme direkt mit dem körperlichen Sein.

Brummern (09:390 ff.) aus, bei denen er nur geringe Chancen einer gesanglichen Verbesserung sieht (09:404 ff.).

Herr Berner singt sehr viel in seinem Musikunterricht und fordert seine Schülerinnen und Schüler mit zweimaligem Vorsingen pro Semester außerordentlich heraus (09:189 ff.). Hauptsächlich angestrebt werden persönliche Entwicklungsschritte im Bereich Auftrittskompetenz. Grundlage für das Gelingen ist seine offene Feedbackkultur und sein Engagement für eine wohlwollende Arbeitsatmosphäre (09:513 ff.).

Der Alleskönner verlangt viel und erreicht viel

Herr Leimbacher (10) ist 57 Jahre alt und bereits 31 Jahre als Musiklehrer auf der progymnasialen Stufe tätig. Seine Ausbildung umfasst neben der Ausbildung zum Schulmusiker für die Sekundarstufe 1 + 2 (MA) auch ein Lehrdiplom für Posaune (MA) und ein weiteres Diplom als Dirigent für Blasmusik (MA). Damit ist er in musikalischer Hinsicht von allen Interviewten am breitesten aufgestellt.

Herr Leimbacher sieht die eigene Vorbildwirkung als zentralen Schlüssel seines Erfolgs (10:313 ff.). Er meint zudem: „eine gewisse Fröhlichkeit und Lockerheit und Freude gehört einfach in dieses Fach. Sonst ist es sterbenslangweilig. Und die Musik ist wichtig.“ (10:163 ff.) Die Förderung der Singfreude mit geeigneten Liedern (10:481 ff.) ist ihm ein wichtiges Anliegen, im Gegensatz zum Einsingen, das keine „heilige Kuh“ (10:486) mehr ist. Stimmtechnische Ansprüche gehen aber nicht verloren, sondern werden beispielhaft an Elementen aus den Liedern geübt, z. B. bezüglich der Artikulation (10:488 ff.) oder dem Singen von Intervallen (468 ff.). Sein Unterricht findet im Klassenverband (10:37 ff.), in Gruppen oder geschlechtergetrennt (10:63 ff.) statt, je nach musikalischen Zielen, die er verfolgt. Meist zielt er auf Mehrstimmigkeit ab, von einfachen zweistimmigen Liedern (10:156 ff.) bis zu vierstimmigen Sätzen (10:53 ff.). Er will schnell „zu einem Resultat kommen“ (10:628 f.) und nimmt die Klassen mit viel Enthusiasmus auf seinem Schnellzug mit. Es kann vorkommen, dass er die Stimme einer Gruppe mit Trompete mitspielt, wenn er mit der Qualität des Gesangs nicht zufrieden ist (10:514 ff.), und schöpft dabei immer aus einem reichen Vorrat an Erfahrungen, welcher ihm als sicheres und tragfähiges Netz dient. Er beschreibt es so: „Und jeder Knüpfungspunkt gibt ein

sicheres Trapez, wo du die Klasse [...] auffangen kannst und mit allen musigen¹⁴⁴ kannst“ (10:674 f.). Trotzdem verliert er nie die einzelnen Schülerinnen und Schüler aus den Augen, macht z. B. für Bewertungssituationen unterschiedliche Niveau-Angebote (10:112 ff.) oder löst ein Stimmproblem (10:228 ff.) außerhalb des Unterrichts in der Pause, weil er sich bewusst ist, dass man nie weiß, wie es in einem Menschen drin aussieht (10:452 ff.). Trotzdem fordert er sie auch individuell heraus, indem er sie mit Mikrofon singen lässt, im Wissen um die gute Selbsterfahrung, die sich dabei ergibt (10:431 ff.).

Herr Leimbacher orchestriert seinen Unterricht virtuos, verlangt viel von den Klassen, geht aber trotzdem auf einzelne Schülerinnen und Schüler ein und ist sich bewusst, dass er nur mit viel Empathie zum Ziel kommt.

Die junge Sängerin kämpft mit ihren Erwartungen

Frau Derendinger (11) ist 28 Jahre alt, hat bereits einige Jahre Unterrichtserfahrung auf verschiedenen Leistungsniveaus der Sekundarstufe 1 und arbeitet in einem regionalen Schulzentrum. Nach dem Studium in klassischem Gesang (BA) hat sie Schulmusik II studiert (MA). Neben ihrer Tätigkeit als Musiklehrerin singt sie in verschiedenen Ensembles im Bereich Pop und Jazz.

Frau Derendinger singt in ihrem Unterricht zwar regelmäßig mit (11:64 f.), sieht sich aber nicht direkt als Vorbild. Als Sängerin kann sie die Idee der Schülerinnen und Schüler, das Singen als Pflicht zu empfinden, nicht nachvollziehen, da es für sie „halt [...] ganz etwas anderes ist“ (11:57). Ihr Hauptziel für den Singunterricht sieht sie beim Ausprobieren der Stimme, wobei Mitmachen wichtiger ist als Perfektion (11:375 ff.). Niemand soll auf die Idee kommen können, nicht singen zu können (11:442 ff.). Allerdings sieht sie ihren Einfluss als Lehrerin beschränkt, da sie die Klassen meist nur eine Stunde pro Woche sieht und sie so kaum eine Beziehung aufbauen kann (11:471 ff.).

¹⁴⁴ Das Verb „musigen“ bedeutet Musik machen und wird meist positiv konnotiert.

Im Verlauf der ersten Unterrichtsjahre hat sie ihre im Studium entwickelten Erwartungen an die Möglichkeiten des Singunterrichts nach unten korrigiert (11:571 ff.) und es ist ihr mittlerweile wichtig, realistische Leistungserwartungen an die Schülerinnen und Schüler zu haben: „Weil ich glaube, wenn ich von den Schülern zu viel erwarte, ist einfach das Frustrationsniveau bei ihnen und bei mir relativ hoch.“ (11:578 f.) Sie berichtet z. B. davon, in gewissen Klassenkonstellationen auf das Singen zu verzichten, weil sich die zu investierende Energie nicht lohnt und weil man die Schüler nicht dazu zwingen kann (11:603 ff.). Ebenfalls berichtet sie von Klassen, in denen Anstand und Respekt vollkommen fehlen und sie als Fachlehrerin zu wenig Einfluss hat, um eine Veränderung einzuleiten (11:633 ff.). Als Grund für den Widerstand nennt sie verschiedene Elemente: Scham (11:231 ff.), Druck unter den Peers (11:233 ff.), keine Lust nach der Mutation bei den Jungen (11:157 ff.), allgemein Unlust (11:203 ff.) und Verweigerung (11:241 f., 445). Sie berichtet aber auch von einer positiven Klassendynamik, die durch gemeinsam erlebte, gute Leistung entstehen kann (11:592 ff.).

Insgesamt sieht Frau Derendinger ihre Möglichkeiten in musikalischer Hinsicht und bezüglich der Förderung der Singstimmen beschränkt, da sie aufgrund fehlender Zeitressourcen kaum eine Beziehung zu den Schülerinnen und Schülern aufbauen kann und auch ihren Status als Fachlehrerin kritisch betrachtet. Sie sieht sich oft mit anspruchsvoller Klassendynamik und fehlender Motivation konfrontiert.

12.2 Zielorientierungen und Überzeugungen

Die elf Porträts der Interviewpartnerinnen und -partner zeigen in erster Linie eine große Diversität bezüglich der Vorstellungen zum Singen, den zu erreichenden Zielen und den Herangehensweisen. Trotz dieser Vielfalt zeigen sich auch Gemeinsamkeiten, die in diesem Abschnitt zu Typiken verdichtet werden. Kelle und Kluge (2010) beschreiben als Ziel der Herausbildung von Typiken, dass „Begrifflichkeiten entstehen, mit deren Hilfe die entdeckten Ähnlichkeiten und Unterschiede“ (S. 84) beschrieben werden können. Große Unterschiede bzw. Gemeinsamkeiten zeigen sich bei den Lehrkräften in der Zielorientierung des Singens. Auf der einen Seite wird qualitativ hochwertiger Gesang angestrebt, der aber aus der Optik der

Lehrkräfte nicht in jedem Fall erreicht wird, und auf der anderen Seite wird das Singen als Grundlage für Persönlichkeitsentwicklung genutzt. Dieser Gegensatz wird im ersten Teil des Abschnitts ausgeführt. Weiter zeigt sich ein Widerspruch in der Wahrnehmung der Jugendlichen durch die Lehrkräfte. Von einigen wird Widerstand, fehlende Motivation und Desinteresse der Jugendlichen beschrieben, von anderen die Wichtigkeit der eigenen Empathie in der Beziehungsgestaltung. Dieser Gegensatz manifestiert sich einerseits in der kritischen Wahrnehmung der Jugendlichen durch die Lehrkräfte und andererseits durch deren Anspruch an gelingende Beziehungsgestaltung. Dieser Querstand wird im zweiten Teil des Abschnitts beschrieben.

Im Rahmen der Analyse wird jeweils überlegt, ob und wie sich der Gegensatz zwischen hauptsächlich musikalisch und vornehmlich pädagogisch geschulten Lehrkräften in diesen Typiken widerspiegelt, insbesondere bezüglich der Merkmale CP, PCK und PK und auch bezüglich der motivationalen Orientierungen (vgl. Shulman, 1987; Baumert & Kunter, 2006).

Zielorientierung: Stimmliche Förderung versus Persönlichkeitsförderung

„Aber diesen Anspruch habe ich irgendwo auch nicht, um die weiß-nicht-wie zu Sängern hinzuschleifen, sondern [...] also quasi ihr könnt singen, ihr macht das gut, ihr habt Mut, ihr wisst wie eure Stimme in diesem Sinn funktioniert, ihr habt Routine im Ganzen drin, ehm, und sie machen das Singen gerne, dann bin ich schon sehr, sehr zufrieden.“ (01:389 ff.)

„[...] ich will deine Stimme trainieren, dass du in die Welt hinaus gehen kannst, und eine Stimme hast. Nicht eine Gesangsstimme, auch eine Gesangsstimme, aber eine Stimme.“ (02:129 ff.)

Die beiden Zitate zeigen zwei gegensätzliche Zielorientierungen: Herr Brugger (01) zielt einen routinierten Umgang mit der Stimme an, Herr Buchser (02) will seinen Schülerinnen und Schülern ermöglichen, mit der Stimme Identität zu entwickeln und Gewicht in der Gesellschaft zu erhalten. Die Polarität zwischen vokaler Förderung und angestrebter Persönlichkeitsentwicklung zeigte sich bereits bei der Erstellung des Kategorien-

systems in der Kategorie *ASSESSMENT*, wo für die Subkategorie *ZIELVORSTELLUNGEN* in einer frühen Phase zusätzlich die Subkategorien *FACH*¹⁴⁵, *PERSÖNLICHKEIT* und *GESANG* eingefügt wurden. Bei der Erstellung der Porträts wurde deutlich, dass sich diese Polarität in den Interviews in verschiedenen Abstufungen als Merkmal zeigt, das sich zu einer Typik verdichten lässt und darauf basierend eine Einteilung des Samples in drei Gruppen ermöglicht. Die größte Gruppe (A) vereint sechs Musiklehrkräfte. Diese beschreiben ihre Ambition, in ihren Klassen vokale Kompetenzen zu entwickeln, sind allerdings unterschiedlich erfolgreich. In eine zweite Gruppe (B) fallen drei Lehrkräfte, welche die vokale Entwicklung sorgfältig mit persönlichen Entwicklungsschritten verbinden. In der dritten Gruppe (C) mit zwei Lehrkräften wird das Singen zwar gepflegt, die Qualität steht aber nicht explizit im Fokus, sondern der positive Einfluss des

	Erfahrung	Stimmliche Förderung	Persönlichkeits- bildung	
		dominant (A)	paritätisch (B)	dominant (C)
01 Hr. Brugger (52)	22	PM – h		
02 Hr. Buchser (55)	12			P – mt
03 Fr. Fricker (49)	16		P – h	
04 Hr. Hilfiker (35)	6	M – mt		
05 Fr. Basler (47)	14	PM – hmt		
06 Fr. von Rohr (37)	10	M – h		
07 Hr. Rothenburg (49)	23	P – m		
08 Hr. Neuendorf (67)	30		MM – hmt	
09 Hr. Berner (50)	10			pm – mt
10 Hr. Leimbacher (57)	31		PMMMm – h	
11 Fr. Derendinger (28)	7	Pm – hmt		
<i>Summe</i>		6	2	3

TABELLE 25: Stimmliche Förderung versus Persönlichkeitsentwicklung: p/P – pädagogisch gebildet (BA/MA), m/M – musikalisch gebildet (BA/MA). Leistungsniveaus: h – progymnasial, m – mittel und t – tief. Die Setzung erfolgte aufgrund der Zuordnungen in Tabelle 15.

¹⁴⁵ Die Subkategorie *FACH* wurde dann verwendet, wenn ein Bezug zu *Assessment* hergestellt werden konnte, aber nicht direkt Ziele für *PERSÖNLICHKEIT* oder *GESANG* angesprochen wurden.

Singens auf die Persönlichkeitsentwicklung. Tabelle 25 zeigt die Einteilung der Lehrkräfte in die drei Gruppen, ergänzt mit den Parametern der Ausbildung aus Tabelle 15 sowie den unterrichteten Leistungsniveaus und der Anzahl an Jahren Erfahrung aus Tabelle 14.

In der **Gruppe A** sind sechs Lehrkräfte vereint, welche die stimmliche Förderung implizit oder explizit ins Zentrum stellen, in ihrer Begründung aber unterschiedlich argumentieren: Herr Brugger (01) arbeitet an einem natürlichen Umgang mit der Stimme, wobei er hofft, dass sein eigener natürlicher Umgang ohne „Hemmungen“ (01:35) dazu führt, dass die Schülerinnen und Schüler dies für sich nachvollziehen. Herr Hilfiker (04) und Frau Derendinger (11) fördern zwar das Singen, sind aber beschränkt zufrieden mit der Qualität. So beklagt z. B. Frau Derendinger (11) den Verlust des Kopfreisters bei vielen Mädchen bis zur 9. Klasse (09:258 ff.). Auffällig ist die Beobachterrolle von Frau von Rohr (06) und Herrn Rothenburg (07), die lediglich die natürliche Stimmentwicklung zur Kenntnis nehmen, ähnlich wie Frau Basler (05), die in der Phase der Mutation zwar die Freude am Singen bewahren will, die Zeit der Sekundarstufe aber als Übergangszeit betrachtet, weil die Jugendlichen „verstummen“ (05:79 ff.). Ein Blick auf die Biografie der genannten Lehrkräfte zeigt, dass drei der Lehrkräfte (05, 06, 11) ausgebildete Sängerinnen sind, was die Affinität zu gesanglichen Zielen erklärt, aber nicht, wieso die mit Einzelstimmgebung vertrauten Lehrerinnen keinen größeren Fokus auf das Training der Singstimme während der Mutation legen. Zudem erstaunt, dass keine dieser Lehrerinnen spezifisches Wissen (CK) zur adoleszenten Stimme anführt, oder sich damit vertieft zu haben scheint. Als mögliche Erklärung kann an dieser Stelle die fehlende Rezeption entsprechender Studien (Sweet, 2018; Gackle, 2019; Freer, 2020) angeführt werden (vgl. 4.3).

Neben der beschriebenen Beobachterrolle zeigt sich bei Herrn Rothenburg (07) eine weitere Eigenheit. Er ist in seiner Freizeit zwar ein enthusiastischer Chorsänger, interessiert sich aber in der Schule eher für die Bandarbeit (07:77 f.). Nochmals anders gelagert ist die Überzeugung von Herrn Hilfiker (04) bezüglich der Lernbarkeit des Singens. Er ist der Meinung, Singen sei nicht lernbar (04:83), meint damit aber hauptsächlich die Intonation. Dies muss als Fehlkonzept bezeichnet werden, da Rutkowski (2019) in ihren Studien zeigen konnte, dass hauptsächlich die vollständige Entwicklung des Instruments Stimme (Register/Registerübergänge) für die Intonation zuständig ist, ebenso wie Sadolin (2013), die ungenügende In-

tonation einer Fehlfunktion der Stimme zuordnet. Zudem zeigt eine Studie von Dalla Bella et al. (2007), dass eine Mehrheit der Menschen in der Lage ist, korrekt zu intonieren. Die Aussagen von Herrn Hilfiker (04) werden aus seiner Biografie verständlich, da er weder Gesangsunterricht (CK) noch eine schulmusikalische Ausbildung (PCK) hatte.

Außer Herrn Rothenburg (07) und Herrn Hilfiker (04) sind alle Lehrkräfte in dieser Gruppe im Rahmen einer Masterausbildung vokal ausgebildet, verfügen also über CK, alle berichten von gesanglichen Highlights, beschreiben sich selbst aber nicht als Herbeiführende oder methodisch Anleitende, sondern eher als Beobachtende und Unterstützende. Sie beobachten die vokale Entwicklung ihrer Schülerinnen und Schüler und freuen sich, wenn sich Talente im Unterricht zeigen (z. B. 07: 340 ff.). Nur Herr Brugger (01) scheint mit seinem Konzept der frontalen Chorsituation zufrieden zu sein und münzt den gesanglichen Erfolg auf sich als Vorbild und seine methodischen Ansätze. Er ist ausgebildeter Chorleiter und somit spezialisiert für das Anleiten von Gesang in Gruppen. Insgesamt kann die Zusammensetzung dieser Gruppe nicht schlüssig über die Profilbildung des Samples erklärt werden, denn die Gruppe A vereint Lehrkräfte aus allen drei Profilen (M, P, MP – vgl. Tabelle 25). Allerdings fällt auf, dass vier der Lehrkräfte ihren Unterricht hauptsächlich als frontale Unterrichtssituation organisieren, wie der Vergleich mit der Verteilung der Sozialformen im Unterricht (vgl. Abbildung 11) zeigt. Zudem drängt sich der Vergleich mit dem Sample aus der Studie von Blanchard (2019) auf, in dem ebenfalls deutlich wird, dass Singen zwar zum Unterricht gehört, aber eine stimmliche Entwicklung oder der Aufbau vokaler Kompetenzen nicht im Fokus stehen.

In der **Gruppe C** zeigt sich in der Zielorientierung eine Dominanz von persönlichkeitsentwickelnden Momenten. Herr Buchser (02) hat den Anspruch an die Jugendlichen, „dass [sie] in die Welt hinaus gehen [können], und eine Stimme [haben]“ (02:161), und Herr Berner (09) erwartet, dass in seinem Singunterricht das „Zeigen der eigenen Persönlichkeit“ (09:15) im Hinblick auf ein Anstellungsgespräch¹⁴⁶ gelernt werden kann. Bei beiden Lehrern dient das Singen einem übergeordneten Bildungsziel. Dieses Ziel

¹⁴⁶ Typischerweise wird vor allem in den tieferen Leistungsniveaus der Sekundarstufe 1 im 8. und 9. Schuljahr intensiv nach Lehrstellen gesucht. Siehe auch Kapitel 6.2.

kommt in der Einteilung von Neus (2017) einerseits bei der Unterscheidung von Singen und Gesang nach B. Jank (2017) zum Tragen, indem es den Prozess Singen und nicht das Produkt Gesang in den Fokus nimmt, und andererseits wird mit dem größeren Selbstvertrauen ein Transfereffekt angestrebt (vgl. 4.1). Die beiden Lehrkräfte betonen intuitiv die entwicklungspsychologischen Aspekte der adoleszenten Stimmentwicklung. Beide Lehrer haben eine pädagogische Grundausbildung, eine musikalische Bildung auf BA-Niveau oder tiefer, aber keine Gesangsausbildung. Interessanterweise sind beide als Musiklehrer auf tieferen Leistungsniveaus tätig.

In der **Gruppe B** zeigt sich schließlich eine Verbindung der beiden Pole, was zu einer Verbindung von gesanglichen und persönlichkeitsentwickelnden Momenten führt. Frau Fricker (03) fördert stimmliche Aspekte, arbeitet aber mit ihren Schülerinnen und Schülern an der Auftrittskompetenz und fordert sie mit der Darbietung eines selbst komponierten Songs vor der Klasse ganzheitlich heraus (03:265 f.). Herr Neuendorf (08) bezieht auch beide Pole mit ein, fordert gegen Ende der Schulzeit auf der einen Seite ein selbstgewähltes Projekt von Poetry-Slam bis Videoclip, besteht aber darauf, dass es „etwas Persönliches sein“ (08:339 f.) muss. Er verlangt aber auch einen Einzelvortrag eines wiederum selbst gewählten Liedes. Bei beiden Lehrkräften zeigt sich der Zusammenhang von Singen und Persönlichkeitsentwicklung implizit, im Gegensatz zu Herrn Leimbacher (10), der explizit darauf hinweist:

„Du musst ja, [...] die Psyche pflegen, du musst Persönlichkeit aufbauen. Das ist ja so ein direkter Schlüssel zur Seele, zum Wohlbefinden eines Menschen, wenn er auch mit seiner Stimme etwas anfangen kann. Und ich sage immer wieder, und später im Berufsleben, mit eurer Stimme repräsentiert ihr euch selber, eure Arbeit, eure Ausstrahlung, oder? Und darum wird immer wieder da gearbeitet, [...] einzeln.“ (10:199 ff.)

Auf der Suche nach Gründen für das Gelingen der Verbindung beider Pole kann die Biografie der Lehrkräfte und zusätzlich das Leistungsniveau zur Diskussion herangezogen werden. Herr Neuendorf (08) und Herr Leimbacher (10) haben 30 Jahre unterrichtet, weisen somit eine sehr große Erfahrung auf und sind zudem mit je zwei musikalischen MA-Ausbildungen, Herr Leimbacher (10) zudem mit einem pädagogischen MA unterwegs. Im Gegensatz zu den Lehrkräften der Gruppe C, die Unterricht auf den unter-

ren Niveaus durchführen, sind alle Lehrkräfte der Gruppe B ganz oder zu einem großen Teil auf dem progymnasialen Niveau tätig.

Insgesamt kann der Singunterricht von Frau Basler (05) als Antithese zu der Idee einer Zielorientierung für die Stimmentwicklung gesehen werden. In der Waldorfschule werden bewusst keine Entwicklungsziele für die Stimme bzw. die Persönlichkeit gesetzt. Die Stimmen werden ihrer natürlichen Entwicklung überlassen, durchgetragen vom aktiven Singen in der Primarstufe hin zu anspruchsvollen Gesangsvorträgen im Chor ab der Klasse 10. Einziges Ziel ist das Musizieren, um ästhetische Erfahrungen mit der Stimme zu sammeln (Wallbaum, 2008). Die vokale Entwicklung wird in dieser Schule somit in der Primarstufe gefördert, durch die gleichen Musiklehrkräfte wie in der Sekundarstufe 1 und nach der Mutation wieder aufgenommen.

Zusammenfassend zeigen sich verschiedene Phänomene: In Gruppe A steht das Singen selbst im Vordergrund, wobei der größere Teil der Lehrkräfte verschiedene stimmliche Herausforderungen beschreibt, darauf aber kaum reagiert, sondern einerseits auf die natürliche Entwicklung vertraut und sich erfreut zeigt, wenn Schülerinnen und Schüler zu Vorbildern für andere werden. Lediglich eine Lehrkraft setzt vokale Ziele, die sie auch erreicht. Die Ausbildungen in Gruppe A sind divers. Als gemeinsame Komponente zeigt sich die frontale Unterrichtssituation. In Gruppe B und C wird durch das Singen an einer Persönlichkeitsentwicklung gearbeitet, bei Gruppe C auf tieferem Leistungsniveau, bei Gruppe B auf der progymnasialen Stufe. Zusätzlich wird in Gruppe B auch das Singen gefördert, wobei diese Lehrkräfte neben langjähriger Erfahrung auch bezüglich ihrer musikalischen Bildung punkten können. Einzig Frau Fricker (03) aus der Gruppe B weist eine ähnliche Ausbildung wie die beiden Lehrer der Gruppe C auf. Bei Gruppe B und C wird deutlich, dass die Lehrkräfte neben dem gemeinsamen Singen Unterrichtssequenzen in vielfältigen Sozialformen durchführen. Das Konzept der Waldorfschule verweist auf eine langfristige Planung, die das in der Primarstufe Erreichte bewahrt und später weiter darauf aufbaut. Dies lässt sich aber kaum vergleichen mit dem Unterricht an der Volksschule, da beim Wechsel von Primarstufe zur Sekundarstufe 1 und später an die weiterführenden Schulen die Lehrkräfte in der Regel wechseln.

Insgesamt lassen sich bei der Polarität Gesang – Persönlichkeit keine schlüssigen Hinweise auf einen Zusammenhang mit der Ausbildung finden.

Weder eine musikalische noch eine pädagogische Ausbildung zeigen eine Tendenz für die Betonung vokaler oder persönlichkeitsbildender Ziele. Insbesondere scheint eine professionelle Gesangsbildung nicht zwingend zu einer starken Förderung vokaler Kompetenzen zu führen. Allerdings zeigt sich eine Tendenz für einen gelingenden Singunterricht, wenn die Persönlichkeitsentwicklung mitgedacht oder sogar ins Zentrum gestellt wird. Dies scheint zudem dann der Fall zu sein, wenn die Sozialformen im Unterricht vielfältig eingesetzt werden. Lehrkräfte, die eine Synthese der Ziele erreichen, weisen neben großer Erfahrung auch eine außerordentlich vielfältige Ausbildung auf.

„Das kann ich nicht. Und ich will das auch gar nicht probieren.“¹⁴⁷

Frau Derendinger (11) zitiert im obigen Titel Jungen aus ihren Klassen. Sie nimmt wahr, dass sich Schüler dem Singen ablehnend entgegenstellen und Widerstand ankünden. Herr Neuendorf (08) dagegen beschreibt eine völlig andere Ebene: „man muss Basis schaffen, dass auch Vertrauen da ist, oder weil Singen so etwas Persönliches ist.“ (08:70 f.) Seine Aussage rückt die Notwendigkeit eines empathischen Zugangs der Lehrkraft zu den Jugendlichen ins Zentrum. Verkürzt könnte von einer Polarität zwischen Widerstand und Beziehungspflege gesprochen werden. Allerdings werden so zwei verschiedene Ebenen angesprochen, auf der einen Seite die Ebene der Schülerinnen und Schüler, die aus der Perspektive der Lehrkräfte durch Handlungen und Einstellungen Widerstand ausdrücken, und auf der anderen Seite die Ebene der Lehrkräfte, die Widerstand methodisch zu überwinden versuchen oder im Interview von Vertrauen und Beziehungspflege als Grundlage von gelingendem Singunterricht sprechen.

Für diesen Querstand auf zwei Ebenen können im Sample drei Gruppen gebildet werden: **Gruppe 1** besteht aus Lehrkräften, die Widerstand hinnehmen, darüber frustriert sind, aber keine Maßnahmen ergreifen, **Gruppe 2** aus denjenigen, die von Methoden sprechen, welche sie im Umgang mit Widerstand entwickelt haben, und **Gruppe 3** vereint Lehrkräfte, die Widerstand kaum thematisieren. Die Hinweise für die Gruppenbildung verdichteten sich durch das Schreiben der kurzen Porträts, währenddessen sich

¹⁴⁷ Aussage von Schülern im Wortlaut von Frau Derendinger (11:46 f.).

eine Gruppe von Lehrkräften herauskristallisierte, die frustriert ist über die geringe Beteiligung im Unterricht oder sogar aktiv Widerstand erlebt. Dabei zeigte sich vor allem die Subkategorie *KLASSENDYNAMISCHE PHÄNOMENE* mit ihren Unterkategorien *STEUERUNG DER KLASSENDYNAMIK*, *FEHLERKULTUR*, *AUF SCHÜLER*¹⁴⁸ *EINGEHEN*, *SCHÜLER ZIEHEN SCHÜLER MIT* und *WIDERSTAND* als Fundorte für Hinweise.

Im Rahmen einer Explikation (Mayring, 2015, S. 90) werden vor der Beschreibung der drei Gruppen zusätzliche Begriffe zu den beiden Perspektiven hinzugefügt, um eine inhaltliche Vertiefung für die Handlungsebene der Schülerinnen und Schüler sowie der Lehrkräfte zu erreichen. Widerstand wird mit den Begriffen Desinteresse (z. B. als Heterogenitätsmerkmal: Hoene & Thurman, 2011), Zwang zum Lernen (Cslovjcek, 2009) und Respektlosigkeit (Forge, 2016)¹⁴⁹ ergänzt. Als Gegenpol werden die Begriffe Interesse, Motivation zum Lernen, Kooperation und Respekt gesetzt, inhaltlich gestützt durch das Konzept von Flow (Csikszentmihalyi, 2008) und der Selbstbestimmungstheorie (Ryan & Deci, 2000). Lehrkräfte

Perspektive Schüler/Schülerin ↓		
Desinteresse – Widerstand Zwang zum Lernen – Respektlosigkeit		Interesse – Motivation zum Lernen Kooperation – Respekt
Gruppe 1 06 Fr. von Rohr (37) 07 Hr. Rothenburg (49) 11 Fr. Derendinger (28)	Gruppe 2 02 Hr. Buchser (55) 05 Fr. Basler (47) 08 Hr. Neuendorf (67)	Gruppe 3 01 Hr. Brugger (52) 03 Fr. Fricker (49) 04 Hr. Hilfiker (35) 09 Hr. Berner (50) 10 Hr. Leimbacher (57)
Distanz Gleichgültigkeit – Ablehnung		Beziehungspflege Empathie – Teilnahme
Perspektive Lehrkraft ↑		

TABELLE 26: *Widerstand (Ebene Schülerin/Schüler) versus Beziehungspflege (Ebene Lehrkraft)*

¹⁴⁸ Im Kategoriensystem wurde das generische Maskulinum beim Begriff Schüler benutzt.

¹⁴⁹ Laut Forge beklagen fast 20 % der Grundschullehrer ihrer Studie disziplinarische Schwierigkeiten beim Singen.

nehmen diese Merkmale in ihrem Unterricht vonseiten der Schülerinnen und Schüler wahr und steuern sie im Rahmen von Classroom-Management (Grohé, 2011). Dies kann als Beziehungsgestaltung deklariert werden und mit den Begriffen Empathie und Teilnahme ergänzt werden. Dagegen stehen die Begriffe Distanz, Gleichgültigkeit und Ablehnung. Um die Handlungsebenen von Schülerinnen und Schülern sowie Lehrkräften mit den damit verbundenen Begriffen darzustellen, werden in der Tabelle 26 die drei Gruppen der Lehrkräfte im Sandwich zwischen den beiden Perspektiven gezeigt. Im Vergleich der beiden Ebenen wird deutlich, dass Distanz und Ablehnung vonseiten der Lehrkraft mit Widerstand und Desinteresse der Schülerinnen und Schüler verbunden sind. Im Gegensatz dazu führt Beziehungspflege durch die Lehrkraft zu Kooperation und Interesse. Die Wichtigkeit der Beziehungspflege bzw. die Unmöglichkeit mit dem Gegenüber nicht in Beziehung zu treten, wurde im autoethnografischen Exkurs mit Bezug auf Bauer (2008) kurz ausgeführt (vgl. 3.3) und als Entwicklungsschritt hin zum Einsatz ganzheitlicher Assessment-Prozesse des Forschenden beschrieben.

In **Gruppe 1** fallen drei Lehrkräfte: Frau Derendinger (11) mit einer überdurchschnittlich großen Anzahl an Hinweisen sowie Frau von Rohr (06) und Herr Rothenburg (07), die bereits aufgrund kurzer Interviews und geringer Kodierhäufigkeiten sowie monodidaktischen Unterrichtskulturen aufgefallen sind (vgl. 11.1 & 11.3). Die drei Lehrkräfte stellen im Singunterricht als Teil der Gruppe A aus der ersten Typik vokale Ziele in den Vordergrund. Frau Derendinger (11) beschreibt im Interview eine ganze Reihe von Ursachen für den Widerstand ihrer Schülerinnen und Schüler, so z. B. Scham (11:231 ff.) oder Lustlosigkeit nach der Mutation bei den Jungen (11:157 ff.). Sie beschreibt auf einer deskriptiven Ebene die Herausforderungen und weist in der Begründung vielfach auf die Schulsituation hin und sieht keine Möglichkeiten für Besserung. So bemängelt sie die Stundendotation mit einer Lektion pro Woche (11:462 f.) und ist überzeugt, als Fachlehrerin keinen Einfluss auf Anstand und Respekt zu haben: „Weil man hat einfach keinen Einfluss.“ (11:634 f.) Zudem ist sie die jüngste Lehrkraft und hat mit 7 Jahren noch wenig Erfahrung. Frau von Rohr (06) und Herr Rothenburg (07) berichten ebenso wie Frau Derendinger (11), dass sie den Singunterricht abbrechen, wenn die Klassen nicht mitmachen. Herr Rothenburg (07) bringt zudem das ständig wechselnde Repertoire von Popsongs und die eigene Distanz zu den Wünschen der Lernenden

als Herausforderung ins Spiel: „Als ich begonnen habe, habe ich einfach meine Lieder gesungen. Welche ich cool finde, so das Repertoire von früher. Und das ist einfach immer gut gewesen.“ (07:419 ff.) Die berichteten disziplinarischen Schwierigkeiten lassen sich gut mit den Worten aus der obigen Explikation erklären. Scheinbar gelingt es den Lehrkräften nicht, das Desinteresse der Lernenden zu durchbrechen und ein kooperatives Verhalten zu erzeugen. Die Verantwortung wird bei den Schülerinnen und Schülern sowie extern im System verortet und Hinweise auf ein empathisches Verhalten oder eine Gestaltung der unbefriedigenden Situation können nicht gefunden werden. Vielmehr verharren die Lehrkräfte in Distanz zu den Schülerinnen und Schülern.

Spannend zeigt sich auch die Analyse der **Gruppe 2**. Die drei Lehrkräfte sind sich in verschiedener Hinsicht bewusst, dass im Singunterricht Schwierigkeiten auftauchen, und haben ihre individuellen Rezepte, damit umzugehen. Frau Basler (05) beschreibt zuerst ihre Enttäuschung: „Es ist der Frust, wenn ich das Gefühl habe, [...] ich erreich' die Kinder eigentlich nicht. Ich kann sie nicht abholen, ich kann sie nicht begeistern, sie lassen es über sich ergehen.“ (05:555 ff.) Sie beschreibt eine Distanz, die sie nicht überwinden kann und fragt sich, ob es an ihr oder den Lernenden liegt. Ihre Methode gegen das Desinteresse besteht darin, klassische Lieder von Löwe und Schubert zu singen, weil sie davon ausgeht, die emotionale Ebene dieser Lieder würde die Schülerinnen und Schüler innerlich treffen (05:565 f.). Das Zitat lässt allerdings vermuten, dass ihre Methode nicht zum Gelingen geführt hat. Auch Herr Buchser (02) äußert sich zu disziplinarischen Schwierigkeiten, insbesondere im tiefsten Leistungszug, und beklagt die Situation, dass „es zu viele Bremser, Störer, Verderber, Verhinderer hat“ (02:565 ff.). Er reagiert auf schwierige Situationen mit offenen Arbeitsaufträgen, schickt die Klasse z. B. für eine Viertelstunde nach draußen und verlangt anschließend einen Gesangsvortrag, den er auch erhält und positiv bewertet (02:151 f.). Obwohl er sich oft als Animateur gibt (z. B. 05:294 ff.), nimmt er sich mit den Gruppenarbeiten aus dem Fokus (05:149) und bricht mit einer neuen Sozialform bzw. einem Rhythmuswechsel die schwierige Situation auf. Es gelingt Herrn Buchser (02), neue Motivation zum Lernen zu erzeugen, indem er die Sozialformen variiert und den Schülerinnen und Schülern mehr Selbstverantwortung übergibt. Herr Neuendorf (08) pflegt einen vorausschauenden Umgang mit disziplinarischen Herausforderungen. So beschreibt er, dass er in jeder Klasse

die Jungen sucht, die möglicherweise Widerstand entwickeln, um diese dann „anzukicken, oder anzutippen, dass dort schon eine Aufmerksamkeit kommt“ (08:137 f.), worauf er zusätzlich spezifische Zugänge z. B. über die Wahl des Repertoires beschreibt. Zu Beginn des Abschnitts beschreibt Herr Neuendorf (08) diesen Vorgang als Basis schaffen bzw. Vertrauen schaffen. Insgesamt zeigen sich in dieser Gruppe eine Reflexionsebene und der Wunsch, den Widerstand zu überwinden. Frau Basler (05) scheint dies nicht wirklich zu gelingen, Herr Buchser (02) reagiert auf Schwierigkeiten mit dem Einsatz verschiedener Sozialformen und spezifischer Aufgabenstellungen und Herr Neuendorf sucht die Verbindung mit den anspruchsvollen Schülern, schafft Vertrauen und weist damit auf die Wichtigkeit der zwischenmenschlichen Beziehung hin (vgl. Bauer, 2008). Sowohl Herr Buchser (02) als auch Herr Neuendorf (08) arbeiten in ihren Klassen an der Entwicklung von Persönlichkeit und legen Wert auf die Beachtung von Grundbedürfnissen, wie das Erleben von Kompetenz, Verbundenheit und Autonomie (Ryan & Deci, 2000). Im Gegensatz dazu steht bei Frau Basler (05) die Entwicklung nicht im Zentrum. Sie hat die Überzeugung, dass die Lernenden während der adoleszenten Stimmentwicklung verstummen (05:79 ff.). In der Gruppe 2 zeigt sich in erster Linie eine reflektierte Haltung und der Wunsch, die herausfordernde Zeit mit Jugendlichen in der Sekundarstufe 1 zu meistern.

Auch in der **Gruppe 3** gibt es unterschiedliche Zugänge zu den Schülerinnen und Schülern, die Lehrkräfte berichten aber nur am Rand von Herausforderungen und können durchwegs beschreiben, wie sie mit den Klassen unterwegs sind. Herrn Brugger (01) gelingt während des Rituals des Einsingens zu Beginn jeder Stunde die Kontaktaufnahme und das Abtasten der Befindlichkeit einzelner Schülerinnen und Schüler (01:70 ff.) und die Gestaltung von längeren Singsequenzen mit mehreren aneinandergereihten Liedern erzeugt Energie (01:210 ff.). Dies wird in ähnlicher Form, aber weniger prägnant von Frau Fricker (03) berichtet. Herr Hilfiker (04) verwendet das Wort *Energie* sehr intensiv in der Beschreibung seiner Ziele: „dann probiere ich das Momentum zu erreichen, wo sie einfach daisitzen und völlig selbstvergessen vor ihrem Blatt, oder bestenfalls ohne Blatt und einfach singen und aufgehen.“ (04:379 ff.) Auch Herr Leimbacher (10) betont, dass Energie fließen muss, „durch Wertschätzung und durch Rückmeldungen“ ebenso wie durch „positiv verstärken“ (10:492 ff.). Herr Berner (09) betont, dass er trotz der einen Lektion pro Woche seine drei-

hundert Schülerinnen und Schüler kennt (09:100 ff.) und auf sofortiges Feedback sowie geschicktes Platzieren der Lerngruppe setzt. Die beschriebenen Situationen berichten von Beziehungsaufbau, Wertschätzung und Feedback sowie von Flow (Csikszentmihalyi, 2008), der durch das Singen von Liedern entsteht. Auch an dieser Stelle kann die Profilbildung des Samplings nicht als Begründung für die Zugehörigkeit zu dieser Gruppe hinzugezogen werden.

Insgesamt fallen in der Deskription und Analyse zwei Elemente auf. Einerseits gelingt es quer durch die Profile musikalisch und pädagogisch gebildeter Lehrkräfte, im Singunterricht qualitativ hochwertige bzw. energetische Momente zu erzeugen, die als Flow beschrieben werden können (Csikszentmihalyi, 2008; vgl. auch Evelein, 2010). Dies gelingt auf allen Leistungsniveaus, allerdings scheinen die tieferen Niveaus anspruchsvoller in der Erzeugung dieser Momente zu sein. Der Schlüssel zum Gelingen zeigt sich in der proaktiven Beziehungspflege, die zu Kooperation und Motivation führt. Nicht zu unterschätzen ist dabei die Kraft der Lieder selbst, da an einigen Stellen die Gestaltung von Singsequenzen mit mehreren Liedern betont wird. Dies kann am Singen als anthropologisches Merkmal (Sundberg, 2019) oder an den positiven Wirkungen gemeinsamen Singens (Gembris, 2015) festgemacht werden. In dieser Typik zeigt sich auch kein Zusammenhang mit dem Gebrauch von Sozialformen. Andererseits scheint eine Gruppe von Lehrkräften nur dann Erfolg zu haben, wenn sogenannte Zugpferde in der Klasse eine positive Spirale in Gang setzen. Diese Lehrkräfte scheinen ein eher distanziertes Verhältnis zu den Schülerinnen und Schülern zu haben, sprechen im Vergleich zur ersten Typik nicht von persönlichkeitsbildenden Momenten des Singens. Zudem schreiben sie Misserfolge externen Faktoren, wie zu knappes Zeitbudget, ihrer Rolle als Fachlehrkraft und dem ständig wechselnden Repertoire zu.

12.3 Erste Antworten auf Forschungsfragen

Im ersten Teil dieses Kapitels wurden Porträts erstellt, die wesentliche Elemente lokaler Unterrichtskulturen entlang der Perspektive der Lehrkräfte zusammenfassten und ein vielfältiges Bild von Individualkonzepten zeigten (Niessen, 2006). Ziel war das Verstehen und Würdigen dieser Kulturen (Strübing et al., 2018) sowie in einem ersten qualitativen Analyseschritt

das Ableiten von Typiken (Kelle & Kluge, 2010). Auf der Grundlage der Porträts wurden in der zweiten Hälfte des Kapitels zwei Typiken herausgearbeitet. Einerseits zeigten sich auseinanderdriftende Zielorientierungen im Singunterricht und andererseits unterschiedliche Beziehungsqualitäten zwischen Lehrkräften und Lernenden. Die beiden Typiken werden nun abschließend auf zwei Teilfragen der übergeordneten Forschungsfrage bezogen.

I. Welche Ziele verfolgen Lehrkräfte in ihrem Unterricht in Bezug auf das Singen und wie kann dies auf eine lokale Schulkultur bezogen werden?

Die erste Teilfrage bezieht sich direkt auf die Analyserichtung *Perspektive der Lehrkraft* der QIA. Durch den Vergleich der Porträts konnte herausgearbeitet werden, dass individuelle Zielorientierungen als Gegensatz zwischen dem Erreichen vokaler Kompetenzen und dem Singen als persönlichkeitsbildendes Moment beschrieben werden können. Im Gegensatz zum LP21, der hauptsächlich musikalische Kompetenzziele beschreibt und nur am Rand auf die Entwicklung einer persönlichen Beziehung zur Stimme hinweist, ist die Entwicklung der Stimme als Merkmal der Persönlichkeit für fast die Hälfte der Lehrkräfte ein wesentliches Ziel ihres Unterrichts. Die andere Hälfte orientiert sich mit vokalen Zielen hauptsächlich an den Zielen des Lehrplans. Dieser Gegensatz ist anschlussfähig an die Unterscheidung von Gesang als Produkt des Singens und Singen als Prozess (B. Jank, 2017). Spannend erweisen sich die Unterrichtsmerkmale, die sich bei den unterschiedlichen Zielvorstellungen zeigen.

- Lehrkräfte, die vokale Kompetenzen betonen, unterrichten in der Regel frontal. Einzig der Chorleiter (01) beschreibt aber seinen Unterricht diesbezüglich als erfolgreich, die anderen betonen die Notwendigkeit motivierter Zugpferde in den Klassen oder sind unzufrieden mit der Zielerreichung. Die Aussagen der Lehrkräfte weisen an der einen oder anderen Stelle auch auf Lücken in den Wissensdomänen CK oder PCK (Shulman, 1987) hin. Zudem scheint ein distanziertes Verhältnis zu den Lernenden das Potenzial zu besitzen, den Singunterricht negativ zu beeinflussen.

- Bei Lehrkräften, die persönlichkeitsbildende Momente des Singens betonen, zeigen sich vielfältige Lernarrangements in variablen Sozialformen. Insbesondere der Einsatz von differenziertem Feedback und eigenverantwortlichem Handeln deuten auf bereits etablierte formative Prozesse hin. Zudem scheint in den tieferen Leistungsniveaus eine Fokussierung auf Ziele der Persönlichkeitsentwicklung erfolgreich zu sein. Einem Teil der Lehrkräfte gelingt es, beide Ziele zu vereinen und sowohl den Prozess als auch das Produkt gelingend zu gestalten. Allerdings finden diese Beispiele meist im progymnasialen Unterricht, also auf höherem Leistungsniveau statt. Zudem sind diese Lehrkräfte sehr erfahren.
- An mehreren Stellen konnte gezeigt werden, dass sich ein Zusammenhang der Zielorientierungen mit den Ausbildungsprofilen nicht begründen lässt.

II. Welche Vorstellungen bezüglich individueller Förderung können bei Lehrkräften rekonstruiert werden und inwiefern werden dabei geschlechtsspezifische Verhaltensweisen im Unterricht wirksam?

Die zweite Teilfrage legt einen Fokus auf individuelle Förderung, die in Kapitel 5 als Ausgangspunkt für die Diskussion ganzheitlicher Assessment-Praktiken diente. Der Bezug zur Individualisierung lässt sich aus den erstellten Porträts nicht direkt ableiten, zeigt sich aber in beiden Typiken in seinen Auswirkungen. So wird in Typik 2 deutlich, dass Beziehungspflege und ein empathisches Verhalten durch die Lehrkraft zu einem kooperativen und motivierten Verhalten der Schülerinnen und Schüler führen oder zumindest das Desinteresse und den Widerstand zu überwinden vermögen. Im Gegensatz dazu gehen vor allem Lehrkräfte, die in ihrem Unterricht disziplinarischen Schwierigkeiten erleben, wenig auf Schülerinnen und Schüler ein, sehen externe Faktoren als Hinderungsgrund für das Gelingen. Sie gestalten ihren Unterricht üblicherweise frontal, brechen sogar den Unterricht ab, wenn die Schülerinnen und Schüler nicht kooperativ sind. Hinweise zur Individualisierung lassen sich in der Analyse der Porträts nur indirekt aufgrund der Beobachtung ableiten, dass Zeit für Beziehungspflege investiert wird und der Unterricht nicht nur in der frontalen Situation stattfindet. Heterogenität wird zwar wahrgenommen, aber eine

Individualisierung wird nicht umgesetzt (Linn, 2017). Interessanterweise sind dabei die Parameter der Ausbildung wiederum wenig aussagekräftig. Weder die musikalisch noch die pädagogisch Gebildeten beschreiben sich als erfolgreicher oder erleben weniger Widerstand. Ähnlich lässt sich auch ein Bezug zur Typik 1 herstellen. Die Betonung persönlichkeitsentwickelnder Momente des Singens basiert auf Vertrauen und Beziehung, die als Basis von Individualisierung gesehen werden kann.

Beide Typiken zeigen auf ihre Weise, dass gelingender Singunterricht ohne starken Bezug zwischen Lehrkraft sowie Schülerin und Schüler nicht auskommt. Zudem berichten Lehrkräfte oft von Erfolg, wenn sie Sozialformen vielfältig variieren und nicht nur vokale, sondern auch persönlichkeitsbildende Ziele verfolgen. Weder in der einen noch in der anderen Typik sind Bezüge zur adoleszenten Stimmentwicklung auszumachen.

13 Horizontale Achse: Perspektive Assessment

Die strukturierende Inhaltsanalyse bildet als Primärverfahren (Mayring, 2015) der vorliegenden Arbeit die methodische Grundlage für die inhaltliche Durchdringung des Materials und die Beantwortung der Forschungsfragen. Nach der quantitativen Analyse der Kodierhäufigkeiten (vgl. Kapitel 11) sowie der Erarbeitung individueller Porträts mit dem Anspruch, Zielorientierungen und Überzeugungen zu beschreiben (vgl. Kapitel 12), ist das Ziel dieses Kapitels eine inhaltliche Verdichtung der in den Kategorien gesammelten Informationen (Schreier, 2014, S. 14). Durch die thematische Sammlung und anschließende Analyse wird es möglich, spezifische Eigenheiten von Diagnose, Förderung und Beurteilung darzustellen, mit bestehenden Konzepten abzugleichen und auf die adoleszente Stimmentwicklung zu beziehen. Die Struktur des Kapitels folgt dabei einem Dreischritt. Im ersten Abschnitt wird auf geschlechtsspezifische Eigenheiten fokussiert, um mögliche Einflüsse auf formatives Assessment zu beschreiben. Im zweiten Abschnitt folgt eine Analyse entlang der drei Bereiche *assessment for*, *as* und *of learning* und im letzten Schritt rückt die Beantwortung der Forschungsfragen ins Zentrum.

13.1 Der Einfluss geschlechtsspezifischer Aspekte auf Assessment

Im ersten Abschnitt wird vornehmlich der Frage nachgegangen, inwiefern im Unterricht geschlechtsspezifische Phänomene ihre Wirkung auf ganzheitliches Assessment entfalten. Die Kategorie *FOKUS AUF DAS GESCHLECHT* dient dabei als Grundlage für die Analyse. Die Kodierhäufigkeiten dieser Kategorie zeigen, dass durchgehend Aussagen zu geschlechtsspezifischen Eigenheiten gemacht wurden (vgl. Tabelle 19). Eine vertiefende Untersuchung macht aber deutlich, dass die Kodierungen vor allem den Einfluss des Geschlechts der Schülerinnen und Schüler und die damit verbundene Stimmentwicklung betreffen und Aussagen zum Einfluss des Geschlechts der Lehrkräfte sowie zu Koedukation und Monoedukation spärlicher sind. Aus diesem Grund wird in spezifischen Bereichen die beschriebene

Methode der Explikation (Mayring, 2015, S. 90) verwendet, in der Aussagen der Interviews kontextualisiert und inhaltlich angereichert werden, um inhaltlich stimmige Begründungen zu generieren (vgl. 8.2). Der Aufbau dieses Abschnitts folgt der Logik des Kapitels 4 (vgl. Abbildung 4), in dem zuerst das Geschlecht der Lehrkräfte, anschließend das der Schülerinnen und Schüler sowie zum Schluss das Zusammenspiel in monoedukativen und koedukativen Settings diskutiert wird. In zwei eingeschobenen Exkursen werden spezielle Fälle untersucht, zuerst der *Stimmbruch* eines Mädchens und anschließend die individuelle Behandlung eines Brummers. Beide Beispiele zeigen spezifische Problemzonen der adoleszenten Stimmentwicklung und weisen gleichzeitig auf mögliche Formen formativen Assessments hin. Im letzten Teil des Abschnitts wird analysiert, ob die Beachtung der adoleszenten Stimmentwicklung und die Betonung geschlechtsspezifischer Phänomene sich als notwendige Perspektive für ganzheitliches Assessment erweisen.

Domänenspezifische Einflüsse des Geschlechts der Lehrkraft

In Kapitel 4.5 wurden Differenzen in der Notengebung zwischen Lehrerinnen und Lehrern im Fach Musik besprochen, eine allgemeine Tendenz für nachweisbare Effekte auf Schülerinnen und Schüler durch das Geschlecht der Lehrkraft konnte aber nicht nachgewiesen werden (Helbling, 2015). Der Autor weist allerdings auf die Möglichkeit domänenspezifischer Effekte hin. Freer (2015) betont denn auch für den Singunterricht die Wichtigkeit männlicher Rollenvorbilder für die Jungen, ebenso wie Kalmbach (2010), der zusätzlich Lehrerinnen als Vorbilder für Mädchen fordert. Als Begründung wird neben einer verbesserten vokalen Identitätsbildung der von Merrill (2020) beschriebene Timbral-Translationsfehler beim Nachsingen von Tönen aufgeführt. Dabei führen Schwierigkeiten beim Nachvollzug muskulärer Spannungsverhältnisse bei gegengeschlechtlichen Vorbildern zu Unsicherheit und falscher Oktavierung (vgl. 4.5).

Im Interviewleitfaden (siehe Kapitel 16.1) wurden geschlechtsspezifische Einflüsse der Lehrkräfte nicht explizit aufgeführt. Entsprechende Inhalte sind aber im Rahmen des unterschiedlichen Umgangs mit Mädchen und Jungen zur Sprache gekommen. Im Kategoriensystem wurde deshalb die Subkategorie *EINFLUSS DES GESCHLECHTS DER LEHRKRAFT* angelegt und auf einer weiteren Ebene mit zwei Subkategorien *TYPISCH LEHRERIN* und

TYPISCH LEHRER ergänzt¹⁵⁰. Kodiert wurden Textstellen, die einen Zusammenhang zum Geschlecht der Lehrkraft vermuten lassen oder diesen explizit ansprechen oder auch verneinen. Insgesamt wurden die beiden Subkategorien mit 23 Kodierungen aber wenig berücksichtigt. Als weitere Quelle für Hinweise wird aufgrund der domänenspezifischen Vorbildwirkung der Lehrkräfte auf die Subkategorie *VORBILD-FUNKTION* Bezug genommen, um die Überzeugungen zur eigenen Vorbildwirkung in der Argumentation mitzudenken. Aufgrund der Vorgaben zu Mehrfachkodierungen (vgl. Leitfadens Kodierung im Anhang) überschneiden sich die genannten Kategorien teilweise.

Nur in sechs von elf Interviews gibt es eine Aussage zur Wirkung des eigenen Geschlechts, verteilt auf drei Lehrerinnen und drei Lehrer. Die vorhandenen Kodierungen verweisen einerseits auf den Timbral-Translationsfehler und belegen andererseits indirekt im Rahmen methodischer Herausforderungen einen geschlechtsspezifischen Effekt. Frau Fricker (03) beschreibt spezifische Effekte bei adoleszenten Jungenstimmen. Sie schildert:

„Die, welche noch nicht so ganz, oder erst frisch im Stimmbruch sind, die singen noch das gleiche a. Aber die, welche schon eine Zeitlang unten angekommen sind, die singen es dann häufig eine Oktave tiefer. Aber es gibt viele, die haben den Stimmbruch gehabt und singen es dann so hoch wie ich. Logischerweise, wenn ich dann eine Oktave tiefer runtergehe, singen sie es dann zwei Oktaven tiefer, weil sie hören, das ist ein tiefer Ton, und dann singen whoaw (brummt) oder und das ist es dann auch nicht.“ (03:49 ff.)

Zuerst nimmt Frau Fricker (03) Bezug auf die Möglichkeit der Jungen, ihre Kinderstimme in der ersten Phase der Mutation weiterhin benutzen zu können (Peterson, 2019). Zudem scheinen die Jungen während der Mutation teilweise Schwierigkeiten zu haben, die akustische Vorlage der Lehrerin umzusetzen, und reagieren auf das dunklere Timbre der Oktavierung mit einem Doppeloktavsprung nach unten. Sie beschreibt damit den

¹⁵⁰ Ergänzend wird an dieser Stelle erwähnt, dass weder in den begleitenden Forschungsnotizen noch im Korpus der Interviews Hinweise gefunden werden konnten, eine Differenz zwischen biologischem Geschlecht und der Geschlechtsidentität der interviewten Lehrkräfte beleuchten zu müssen.

Timbral-Translationsfehler, betrachtet diesen aber hauptsächlich als Folge der adoleszenten Stimmentwicklung und den damit verbundenen Schwierigkeiten bei sensorischen und motorischen Anpassungsleistungen. Frau Basler (05) referiert das gleiche Phänomen (05:205 ff.), bezieht es aber nicht auf die Mutationsphase, sondern allgemein auf den Singunterricht mit Jungen. Als methodische Lösung nennt sie die Vorbildwirkung der anderen Jungen, ganz ähnlich wie Frau Fricker (03), welche die Lehrer in ihrem Schulchor als Vorbilder erwähnt. Frau Derendinger (11) nimmt nur auf Anfrage Bezug auf geschlechtsspezifische Phänomene beim Nachsingen:

„Könnte schon sein, dass es vielleicht einfacher wäre für einen Knaben, wenn es ein Mann singen würde, dass sie es genau in der Lautstärke hören, in welcher sie es dann machen müssen, dass das vielleicht ein wenig einfacher wäre, aber ich finde eigentlich grundsätzlich müsste es nicht unbedingt einen Unterschied machen.“ (11:171 ff.)

Sie zieht zwar eine gleichgeschlechtliche Vorbildwirkung in Erwägung, bezweifelt aber die Notwendigkeit entsprechender Maßnahmen für den Unterricht. Von insgesamt vier Lehrerinnen achten somit nur zwei explizit auf mögliche Schwierigkeiten von Jungen aufgrund ihres Geschlechts. In ihrem Unterricht setzen sie auf die Vorbildwirkung von Peers oder Lehrern (vgl. Freer, 2020), weitere Methoden werden aber nicht genannt. Insgesamt handelt es sich zwar um ein zu beachtendes Phänomen, das möglicherweise bei formativem Assessment sichtbar wird. Gewichtiger erscheint die Erkenntnis der beiden Lehrerinnen, Jungen seien auf Vorbilder angewiesen.

Auch Lehrer beziehen spezifische Herausforderungen im Singunterricht mit Mädchen auf ihre Vorbildrolle. Herr Buchser (02) spricht sein Unbehagen bezüglich der hohen Mädchenstimme folgendermaßen an: „Mir fällt immer wieder auf, [...] du vergisst die Mädchen, die singen zu tief.“ (02:327 f.) Er beobachtet, dass die Mädchen in seinem Unterricht das Randregister nicht benutzen. Er führt später aus, dass er diesem Umstand damit begegnet, dass er sie im Falsett anleitet und erwartet, sie damit aus der Bequemlichkeit des tiefen Singens herauszulocken, verweist aber gleichzeitig auf den mittelmäßigen Erfolg seiner Methode. Herr Berner (09) will sich nicht auf seine Vorbildrolle für Mädchen verlassen und hofft, dass „vor allem die Mädchen [...] in den Sologesang gehen“ (09:665 f.). Er spricht sich selbst damit die Fähigkeit ab, Mädchenstimmen adäquat zu

fördern. Herr Hilfiker (04) sagt auf Nachfrage zum Umgang mit dem Kopfreister der Mädchen, dass er nicht darauf eingeht: „Häufig haben sie [...] von alleine dann einfach getüpfelt. Oder, dann habe ich sie ein wenig sein lassen.“ (04:164 f.) Herr Hilfiker hofft, dass die Zeit für ihn arbeitet und die Mädchen allmählich die Töne automatisch treffen, er meint: „Mädchen die hoch singen, da bin ich noch nicht so wirklich darauf eingegangen, oder können [...]“. (04:162 f.) Er ist sich offenbar bewusst, keinen methodischen Zugang zum Randregister der Mädchenstimme zu haben. Bei diesen drei Lehrern zeigt sich ein Unbehagen im Umgang mit der Mädchenstimme, welches auf das Geschlecht der Lehrer bezogen werden kann. Zwei der Lehrer sind vokal nur minimal ausgebildet, bei einem (04) fehlt die vokale Ausbildung vollständig.

Im Gegensatz dazu nehmen drei weitere Lehrer keine Stellung zu ihrer Vorbildwirkung auf Mädchen, wie sich in der Subkategorie *FÖRDERUNG* zeigt. Herr Leimbacher (10) benutzt ebenso wie Herr Buchser (02) seine Falsettstimme als methodisches Werkzeug (10:70 ff.) und nennt mit dem Einbezug seines Trompetenspiels bei Sopranstimmen sogar ein weiteres: „ich spiele Trompete, ihr singt mit.“ (10:518) Er problematisiert den Umgang mit dem Randregister der Mädchen nicht und reiht sich mit Herrn Brugger (01) und Herrn Neuendorf (08) in eine Gruppe von vokal und schulmusikalisch hoch qualifizierten Lehrern ein, die das Randregister nicht als Herausforderung beschreiben. Allerdings kann nicht ausgeschlossen werden, dass der Timbral-Translationsfehler auch da eine gewisse Rolle spielt, denn die Qualität von Falsett ist nicht identisch mit der weiblichen Randstimme, ebenfalls muss der Klang der Trompete diesbezüglich infrage gestellt werden.

Zu bemerken ist an dieser Stelle allerdings, dass die Mädchen während ihrer Mutation das Phonieren mit der Randstimme in der Tendenz zu vermeiden versuchen (Gackle, 2019). So beschreibt auch eine Lehrerin den Verlust der Randstimme bei Mädchen im Verlauf der drei Jahre in der Sekundarstufe 1 (11:257 ff.). Sie ist ebenfalls vor die Herausforderung gestellt, Mädchen zu einer adäquaten Benutzung des Randregisters zu führen. So zeigt sich vor allem der methodische Umgang mit der Randstimme als Herausforderung, möglicherweise deutlicher bei Lehrern, wenn entsprechendes PCK fehlt.

Insgesamt zeigen die Beispiele, dass im Singunterricht mit dem Timbral-Translationsfehler (Merrill, 2020) ein geschlechtsspezifischer Einfluss zu

beobachten ist. Lehrerinnen scheinen während der Mutation der Jungen und Lehrer eher allgemein im Umgang mit der Randstimme der Mädchen Auffälligkeiten im Zusammenhang mit ihrer Vorbildrolle zu erleben. Abschließend stellt sich die Frage, inwiefern sich die beschriebenen Phänomene auf ganzheitliches Assessment auswirken. Der Timbral-Translationsfehler wird höchstens in formativen Prozessen (AFL) relevant, wenn durch Unkenntnis des Phänomens Unsicherheiten bei Schülerinnen und Schülern entstehen. Das Bewusstsein einer unzureichenden Vorbildwirkung scheint in gewissen Konstellationen, z. B. bei der Entwicklung der Randstimme, aber auch die vokale Förderung von Mädchen zu behindern, insbesondere bei männlichen Lehrkräften. Vokal hoch qualifizierte Lehrer zeigen damit aber keine Schwierigkeiten, interessanterweise aber eine Lehrerin (11), die als Sängerin ausgebildet ist. Die von Freer (2015) und Kalmbach (2010) geforderten gleichgeschlechtlichen vokalen Vorbilder sind aus dieser Optik zwar als sinnvoll zu erachten, lassen sich aber nicht als Notwendigkeit argumentieren, weil die Beispiele auch auf den Einfluss methodischer Defizite schließen lassen. Dabei darf aber nicht die Vorbildwirkung von Peers außer Acht gelassen werden (Freer, 2020).

Wahrnehmung von geschlechtsspezifischen Unterschieden bei Mädchen und Jungen

Die Diskussion der Wahrnehmung geschlechtsspezifischer Phänomene bei Mädchen und Jungen erfolgt unter zwei Gesichtspunkten. Einerseits wird untersucht, wie die Geschlechter allgemein durch die Lehrkräfte wahrgenommen werden und wie sich diese Aussagen mit dem Forschungsstand decken (CK; vgl. 4.3 & 4.4). Dabei werden auch die vorhandenen Vorstellungen mit den beschriebenen stereotypen Überzeugungen (vgl. 4.5) zu Fähigkeiten verglichen. Andererseits wird der methodische Umgang mit der Mutation beschrieben und analysiert (PCK). Ziel der Überlegungen ist wiederum die Evaluation der Einflüsse auf formatives Assessment im Unterricht.

Im Leitfaden für die Interviews ist die Thematik mit mehreren Fragestellungen vertreten. Gefragt wurde allgemein nach Unterschieden beim Singen mit Mädchen und Jungen, nach möglichen motivationalen Differenzen sowie dem Umgang mit der adoleszenten Stimmentwicklung. Dies führte im Kategoriensystem zu den Subkategorien *TYPISCH MÄDCHEN* und

TYPISCH JUNGEN in der Subkategorie *EINFLUSS DES GESCHLECHTS DER SCHÜLERINNEN UND SCHÜLER* sowie zu den Subkategorien *MUTATION MÄDCHEN* und *MUTATION KNABEN* in der Subkategorie *UMGANG MIT STIMMVERÄNDERUNG*. Diese deduktiv gesetzten Kategorien wurden ergänzt mit der Subkategorie *BRUMMER*, die sich während des Kodierens insbesondere als Abgrenzung zur Mutation bei den Jungen aufdrängte. Das Thema wird später in einem Exkurs aufgenommen.

Aussagen zur adoleszenten Mädchenstimme sind bei den Subkategorien *TYPISCH MÄDCHEN* (32 Kodierungen) und *MUTATION MÄDCHEN* (17 Kodierungen) mit einer Ausnahme bei allen Interviewpartnerinnen und -partnern zu finden. In der Analyse treten drei Themen hervor: Einmal der Umstand, dass die Arbeit mit Mädchen von einigen Lehrkräften als einfacher bezeichnet wird, weiter Beobachtungen zur adoleszenten Stimmentwicklung sowie abschließend Hinweise zu einem negativen Selbstkonzept.

Frau Derendinger (11) beurteilt die Arbeit mit den Mädchen folgendermaßen: „[...] aber grundsätzlich, wenn ich jetzt alle über einen Kamm drüber scheren würde, würde ich sagen, mit den Mädchen ist es ein wenig einfacher als mit den Knaben.“ (11:252 ff.) Sie ist mit dieser Aussage nicht allein, auch Herr Hilfiker (04) nennt die Mädchen „Selbstläufer“ (04:167). Frau Fricker (03) scheint froh zu sein, den Fokus nicht auf die Mädchen legen zu müssen: „aber ich habe wie ein wenig zu wenig ehm Zeit, mich auch noch um die Mädchen zu kümmern.“ (03:423 f.) Sie bietet allerdings zusätzlich zum regulären Unterricht ein Wahlfach mit Stimmbildung an, das meist nur von Mädchen besucht wird (03:257 ff.). Die drei Aussagen weisen in verschiedenen Qualitäten darauf hin, dass die Mädchen weniger Zeit und Aufmerksamkeit beanspruchen. Herr Buchser (02) beschreibt diesen einfacheren Umgang mit Mädchen allerdings als Gefahr: „Aber Mädchen, welche eigentlich gut singen, welche nicht stören, gehen in der Masse auch unter.“ (02:363 f.) Er legt den Finger auf den kritischen Punkt der reduzierten Aufmerksamkeit, denn dies verhindert eine adäquate Förderung und verunmöglicht formatives Assessment. Wenn Kalmbach (2010) vokale Gleichberechtigung fordert, gilt dies somit nicht nur für die Jungen, sondern aufgrund mangelnder Zuwendung spezifisch für die Mädchen.

Gleichzeitig nennen vereinzelt Lehrkräfte eine größere Leistungsfähigkeit der Mädchen (z. B. 03:421 ff.), wobei Herr Hilfiker differenziert, indem er auf interindividuelle Unterschiede hinweist (04:193 ff.). Einige Lehrkräfte nehmen Bezug auf die größere Singfreude der Mädchen, z. B.

Herr Berner (09:379 ff.): „Weil die Mädchen sind meistens einfach singiger [...]. Also, das ist jetzt, natürlich sehr plakativ, Mädchen Knaben. [...] Logisch ist es nicht nur so einfach, [...]“. Das Wort plakativ weist darauf hin, dass sich Herr Berner (09) seiner stereotypen Aussage bewusst ist. Zudem zeigen gerade seine monoedukativen Unterrichtssequenzen, die später zur Sprache kommen, wie wichtig ihm die Jungen sind. Im Gegensatz dazu nimmt Frau Derendinger (11) lediglich wahr, dass Mädchen lieber singen. Die von Green (1997) beobachtete stereotype Wahrnehmung von Lehrkräften bezüglich einer größeren Motivation und Leistungsfähigkeit bei Mädchen wird somit in den Interviews genannt, scheint aber nur ein nebensächliches Thema oder im Gegenteil Anlass für Interventionen zu sein.

Ein zweiter Bereich ist die adoleszente Stimmentwicklung der Mädchen. Herr Rothenburg (07) meint auf Nachfrage: „Nein, mich hat das noch nie beschäftigt.“ (07:354) Die Lehrkräfte nehmen zwar vereinzelt verschiedene von Gackle (2019) beschriebene Phänomene wahr, Maßnahmen im Umgang mit dem Stimmwechsel werden allerdings nicht genannt. Von Herrn Hilfiker (04) und Frau Basler (05) wird Schüchternheit wahrgenommen (04:186 f.; 05:81 f.) und Frau von Rohr (06) beschreibt Heiserkeit und leisen Gesang als Mutationsphänomen, ist sich aber nicht sicher, wie stark das Selbstbewusstsein dabei mitspielt (06:203 ff.). Nur Herr Berner (09) bemerkt bewusst, dass die Mädchenstimmen voller und tiefer werden. Er passt entsprechend auch die Tonarten von Liedern nach unten an (09:428 ff.). Die zu Beginn genannte Schüchternheit wird nur von einem Lehrer (10) als Ausdruck eines emotionalen Karussells beschrieben und führt später zu einem der Exkurse. Der Verweis von Gackle (2011) auf emotional schwierige Zeiten und die Beobachtung von Sweet (2018), dass auch vokal kompetente Mädchen große Unsicherheit erleben, wird somit kaum wahrgenommen. Frau Derendinger (11) nimmt allerdings implizit auf das vokale Selbstkonzept der Mädchen Bezug:

„Bei den Meitschi¹⁵¹ gibt es so ein Phänomen, dass ich wie das Gefühl habe, es gibt zwei Sorten von Mädchen. Die einen Mädchen, welche sagen, [...] das finde ich mega cool und Singen ist super und das kann ich gut. Und die andere Sorte, welche sagen, eigentlich ist es cool, aber ich kann es einfach nicht.“ (11:213 ff.)

¹⁵¹ Meitschi = Mädchen.

Das Phänomen der *non-singer identity* wird üblicherweise den Jungen zugeschrieben. So entwickelt sich teilweise bereits vor der Pubertät die Überzeugung, nicht singen zu können (Ashley, 2015). Die Beschreibung von Frau Derendinger (11) bezieht sich allerdings auf das negative vokale Selbstkonzept von Mädchen. Sie berichtet zudem davon, dass Mädchen oft vokale Vorbilder nachahmen wollen und „ihre eigene Stimme“ (11:218 f.) dabei verlieren. Bezieht man das Konzept der *vocal identity* nach Ashley (2015, S. 113) auf ihre Beschreibung, wird vor allem die Kluft zwischen dem Klang populärer Vorbilder und den gerade während der adoleszenten Stimmentwicklung beschränkten Möglichkeiten deutlich. Ähnlich wie Frau Derendinger (11) kennt Herr Brugger (01) Beispiele von Mädchen, die behaupten, nicht singen zu können: „Aber, wenn mir das jemand sagt, er könne nicht singen, dann gehe ich das aktiv an.“ (01:283 f.) Gemeinsam mit der Person sucht er auf der musikalischen Ebene nach Beweisen für das Gegenteil. Die beiden Lehrkräfte nehmen die Auswirkungen eines negativen Selbstkonzepts zwar wahr (Rieche et al., 2018), reagieren aber entweder nicht oder nur auf der musikalischen Ebene. Eine solche Intervention kann durchaus gelingen, wie der spätere Exkurs zum auffälligen Stimmwechsel eines Mädchens zeigt. Die Wahrnehmung von negativen Selbstkonzepten scheint aber nicht die Regel zu sein und auch nur vereinzelt zu Interventionen zu führen.

Daneben steht bei vielen Lehrkräften das Thema Randstimme im Fokus. Herr Brugger (01) weist auf den fehlenden Oktavsprung der Mädchen bei Popsongs hin (01:36 ff.)¹⁵², Herr Neuendorf (08) verweist auf die fehlende Lust, das Randregister zu bedienen (08:232 ff.), und Frau Derendinger (11) beschreibt ausführlich, wie sie versucht, die Mädchen an die Benutzung des Randregisters hinzuführen (11:506 ff.). Sie verweist auf einen langwierigen und „extrem schwierig[en]“ (11:520) Prozess, weil die Mädchen zuerst dazu gebracht werden müssten, das Randregister zu benutzen und dies danach in einer langen Konsolidierungsphase stabilisiert werden müsste. Der fehlende Zugang zum Randregister stimmt mit dem Phasenmodell von Gackle (2000) überein, wird aber von den Lehrkräften nicht als natürlicher

¹⁵² Der klassisch ausgebildete Chorleiter scheint hier nicht zu beachten, dass in Popsongs Männer oft hoch und Frauen oft tief singen und sich so aus klanglicher Perspektive nur eine kleine Differenz zwischen den Tonhöhen ergibt, sicher keine Oktave wie im klassischen Gesang.

Prozess der adoleszenten Stimmentwicklung eingeordnet. Als Maßnahme wird bei einer Mehrheit von Lehrkräften von Singübungen gesprochen, mit denen das Randregister trainiert werden soll. Dies deckt sich mit der Idee, mit adoleszenten Mädchen grundsätzlich einen Ambitus von zwei Oktaven zu trainieren (Gend et al., 2017) bzw. das Instrument als Ganzes zu beherrschen (Rutkowski, 2019).

Zusammenfassend wird deutlich, dass Mädchen als pflegeleichter wahrgenommen werden und teilweise sogar weniger Aufmerksamkeit erhalten. Gleichzeitig lassen sich zwei Bereiche eruieren, in welchen deswegen eine gezielte Förderung verhindert wird. Einerseits wird das Training der Randstimme nicht durchgehend, aber von einigen Lehrkräften als Herausforderung betrachtet. Dieser Punkt ist weiter oben bereits aufgetaucht und verweist auf eine fehlende Methodik (PCK). Andererseits werden Phänomene klanglicher Art sowie vereinzelt die Auswirkungen des vokalen Selbstkonzepts wahrgenommen, ein methodischer Umgang damit zeigt sich aber nicht. Insbesondere die Thematik der Randstimme, aber auch die Beachtung des vokalen Selbstkonzepts wird nicht mit der adoleszenten Stimmentwicklung der Mädchen in Zusammenhang gebracht, ließe sich dadurch aber bewusster vermitteln (Gackle, 2011).

Aussagen zur adoleszenten Jungenstimme sind in den Subkategorien *TYPISCH JUNGEN* (20 Kodierungen) und *MUTATION JUNGEN* (35 Kodierungen) zu finden. Diese Subkategorien wurden mit einer Ausnahme bei der Subkategorie *TYPISCH JUNGEN* durchwegs kodiert. Die Kodierhäufigkeiten weisen darauf hin, dass hauptsächlich die Mutation der Jungen im Zentrum steht. Trotzdem werden im Anschluss zuerst generelle Aussagen zur Arbeit mit Jungen untersucht, um dann zwei auffällige Beispiele des Umgangs mit der Mutation zu kontrastieren. Spezifisch erörtert wird ebenfalls der Umgang mit der Tessitura der Jungen.

Die Aussagen zu den Jungen sind vielfältig, haben aber in der Tendenz einen stereotypen Unterton. So wird von einer geringeren Leistungsfähigkeit ausgegangen (03:35 f.) oder angemerkt, dass Jungen mehr Aufmerksamkeit benötigen (04:188 f.; 10:279 ff.). Es wird sogar von Desinteresse (09:47 ff.) und Widerstand gesprochen. Frau Derendinger (11) legt ihren Jungen folgende Worte in den Mund: „ich kann das sowieso nicht und ich bin sowieso schlecht in dem und darum mache ich es gar nicht.“ (11:48 f.) Herr Neuendorf beugt diesem Umstand vor und schafft von Beginn weg eine Vertrauensebene (08:133 ff.) und Frau von Rohr (06) motiviert die

Jungen durch ein Mitspracherecht bei der Wahl von Liedern (06:135 ff.). Die Aussagen unterscheiden sich stark und wurden teilweise bereits im letzten Kapitel bezüglich der Notwendigkeit von Beziehungspflege als Typik 2 diskutiert (vgl. 12.2). Als zusätzlicher Aspekt wird von Frau Basler (05) und Frau Derendinger (11) beschrieben, wie ein Überhang an Jungen beim Singen zu Schwierigkeiten führen kann (05:154 ff.; 11:90 ff.). Diese Tatsache in Kombination mit einem oft höheren Anteil an Jungen in tieferen Leistungsniveaus (vgl. 6.3) untermauert zusätzlich die Wichtigkeit eines bewussten Umgangs mit der Klassendynamik als Grundlage von Förderprozessen.

Die Mehrheit der Kodierungen zu Jungen bezieht sich auf die Mutation selbst. Die Aussagen zweier Lehrer stehen in starkem Kontrast und werden deshalb fokussiert. Herr Brugger (01) reagiert bewusst nicht auf Jungen in der Mutation, um diese nicht bloßzustellen, Herr Buchser (02) dagegen stellt die Jungen in den Mittelpunkt. Herr Brugger (01) will auf keinen Fall Situationen provozieren, die eine bloßstellende Wirkung entfalten könnten. Er begründet dies mit dem Wissen um Personen, die durch Aussagen im Musikunterricht zur Überzeugung kamen, nicht singen zu können (01:270 ff.). Als Konsequenz verzichtet er im Unterricht generell auf individuelles Feedback, speziell auch bezüglich der Mutation (vgl. 12.1). Sein Ansatz bedeutet für die Jungen, in ihrer Stimmentwicklung allein gelassen zu werden. Herr Buchser stellt dagegen die Jungen zu Beginn der Mutation in den Mittelpunkt des Unterrichts:

„Immer wieder ein Thema, jeder Knabe, der in den Stimmbruch kommt, wird von mir gelobt, heißt, super, hey du kommst in den Stimmbruch. Gäll, jetzt kommen die tiefen Töne, mach mal ,na na na na na...‘ [...], dann kommen sie dort nach unten, dann sage ich whoaw [...]. Für uns ist es faszinierend. Machst du es bitte noch einmal? Und dann sind sie überrascht, dass ich dies alles positiv betone, machen das vor, und nachher beginnen wir uns zu überlegen, wieso ist das so.“ (02:371 ff.)

Auch wenn ein solches Vorgehen vor allem zum Schauspieler Buchser (02) passt und in keinem Artikel so beschrieben wird, können doch zwei spannende Eigenheiten eruiert werden. Erstens gelingt es ihm, dem Schüler jegliche Angst zu nehmen, mit seiner Stimme würde etwas nicht stimmen, und zweitens begibt er sich bewusst auf einen Weg, mit dem Schüler eine neue *vocal identity* zu finden. Der Unterschied der beiden Positionen zeigt

sich hauptsächlich in der bewussten Bearbeitung des vokalen Selbstkonzepts durch Herrn Buchser (02). Auf weniger theatralische Art und Weise begleitet auch Herr Leimbacher (10) die Jungen aktiv durch die Mutation, indem er sich regelmäßig über ihren Stimmzustand informiert und mit ihnen Übungen macht (10:182 ff.). Ein ähnliches Vorgehen wird später bei der Evaluation monoedukativer Unterrichtssequenzen bei weiteren Lehrern zu sehen sein. Der beschriebene Kontrast weist auf die psychologische Ebene hin, die von einigen Lehrkräften bewusst als Persönlichkeitsentwicklung (vgl. 12.2, Typik 1) adressiert wird. So wird von Herrn Buchser (02) bewusst das Selbstkonzept angesprochen, von Herrn Brugger (01) bewusst nicht. Ein ähnliches Bild zeigt sich auf der musikalischen Seite im Umgang mit der kontinuierlich tiefer werdenden Stimme.

Sowohl Harrison und Williams (2019) als auch Fuchs (2020) empfehlen während der Mutation ein schonendes Training in der tiefer werdenden, gut singbaren Tessitura. Ein Verzicht auf das Singen wird nur im Rahmen von professionellen Knabenchören diskutiert (Peterson, 2019). Es interessieren deshalb die Aussagen der Lehrkräfte zu adäquaten Tonlagen und einem möglichen Verzicht des Singens bei mutierenden Jungen. Die Analyse der Interviews zeigt zwei Vorgehensweisen, einerseits der Verzicht auf einen spezifischen Umgang mit der Mutation bis hin zum Verzicht auf das Singen selbst und andererseits die Verwendung einer adäquaten Tessitura. Die Mehrheit der Lehrkräfte lässt die Jungen in der oberen Oktave mitsingen, so lange es geht, und wartet dann ab, bis die Jungen in der unteren Oktave wieder einsteigen, die Zwischenphase wird aber unterschiedlich gehandhabt. Herr Rothenburg (07) lässt die Jungen lieber falsch singen als gar nicht (07:259 ff.). Herr Brugger (01) gibt allenfalls in der Bewertungssituation Tipps zur Mutation, hofft aber ansonsten, dass die Jungen schweigen. Er meint: „Und wenn sie clever sind, dann singen sie dann einfach nicht in der Hoffnung, ich würde es nicht sehen und machen den Mund auf und zu.“ (01:121 ff.) Weitere Lehrkräfte sind sich der tiefer werdenden Stimme nicht bewusst (04:118 ff.) oder fordern aktiv zum Verzicht auf das Singen auf (11:113 ff.; 05:201 ff.). Frau Fricker informiert ihre Jungen folgendermaßen: „Und ich sage ihnen einfach, es gibt eigentlich keine Lösung für das.“ (03:69 f.) Sie fasst damit das Nichteingehen auf die adoleszente Stimmentwicklung der Jungen zusammen. Im Gegensatz dazu erachten Herr Berner (09) und Herr Neuendorf (08) das Angebot eines adäquaten Tonumfangs als selbstverständlich. Herr Berner meint: „Ja unbedingt. Lo-

gisch. Ehm, ist noch witzig, ehm mit der Zeit hört man es einfach, wie jemand redet.“ (09:365 f.) Er eruiert geschickt die Sprechstimmlage, um die Tonart darauf anzupassen (Fuchs, 2020). Herr Neuendorf (08) transponiert im Musikunterricht ebenfalls, im Unterschied zum Knabenchor, in dem er die Jungen in die passenden Stimmlagen einteilt. Aus der Perspektive von Assessment ist der beschriebene Verzicht auf einen bewussten Umgang mit der Mutation gleichzeitig auch ein Verzicht auf Förderung. Singen in der Komfortzone ermöglicht die kontinuierliche muskuläre und sensorische Anpassungsleistung und ermöglicht so eine individuelle Förderung und Begleitung.

Beim Umgang allgemein mit Jungen und spezifisch mit der adoleszenten Stimmentwicklung zeigt sich ein disparates Bild. Einerseits haben verschiedene Lehrkräfte Methoden entwickelt, die mit dem Forschungsstand gut übereinstimmen, sei dies bezüglich adäquater Tonhöhen beim Singen oder der Entwicklung einer *vocal identity* durch die Förderung von Auftrittskompetenz in monoedukativen Settings. Es wird lokal generiertes Wissen gepflegt, intuitiv entwickelt auf der Basis einer langjährigen Erfahrung. Andererseits beobachtet eine größere Gruppe von Lehrkräften die Mutation der Jungen, ohne sich spezifisch um diese zu kümmern, in der Überzeugung es gebe keine Methoden im Umgang mit dem Stimmbruch. Die von verschiedenen Autorinnen und Autoren geforderte Betreuungsleistung (Freer, 2020; Fuchs, 2020; J. Williams et al., 2021) wird in diesem Fall nicht erbracht. An dieser Stelle muss noch einmal auf die fehlende Rezeption englischsprachiger Literatur und die eher spärlichen (A. Mohr, 2014) oder mehr deskriptiven (Schnitzer, 2008b) Angaben zum Umgang mit der adoleszenten Stimme als möglicher Grund für die disparate Situation verwiesen werden.

Aus diesem lokal gewachsenen Wissen um die Betreuung Jugendlicher während der Sekundarstufe 1 werden nun zwei herausstechende Beispiele explizit beleuchtet. Beide weisen auf formative Förderprozesse hin, die individuelle Personen betreffen, einmal die schwierig erlebte Mutation eines Mädchens und im zweiten Fall der mitten in der Klasse betreute Brummer bzw. *poor pitch singer*, wie diese Kategorie von Singenden später genannt wird.

Der auffällige Stimmwechsel bei einem Mädchen

Herr Leimbacher (10) erzählt von einer emotionalen Situation mit einem groß gewachsenen Mädchen, das nach einem massiven Stimmrutsch mit der eigenen Stimme nicht klarkommt. Die Situation führt zu einer Intervention des Lehrers. Im Sinne einer Explikation nach Mayring (2015) wird die Situation mehrperspektivisch beleuchtet und durch das Ergänzen von Informationen zu einer schlüssigen Erklärung gebracht. Nach der Darstellung der Episode wird zuerst das Phänomen und die Reaktion des Lehrers beleuchtet, um danach den Prozess als Ganzes bezüglich Assessment einzuordnen.

Die Jugendliche beginnt in der Musikstunde zu weinen, worauf sie vom Lehrer mit Begleitung auf die Mädchentoilette geschickt wird. Herr Leimbacher (10) beschreibt ausführlich, wie er aufgrund stundenplantechnischer Umstände auf einfache Weise eine Einzelbetreuung organisieren kann, und schildert anschließend seine Intervention:

„Und dann habe ich gemerkt, wie die hyperemotional reagiert, auf gewisse Töne, und sobald es ein wenig höher raufgeht, ist wirklich ein großes Mädchen, hat wirklich fast wie einen Stimmbruch gehabt, das Mädchen, ein massiver Stimmrutsch, von Stimmbruch sollte man ja nicht sprechen. Und dann haben wir da miteinander geübt. [...] Und plötzlich alleine hat sie gemerkt, aha, und dann habe ich ihr gesagt, hör, Körperstimme, Kopfstimme, probiere deine Kopfstimme jetzt ein wenig mehr zu lancieren, du bist jetzt wie fast wie ein wenig im Körper unten gefangen, oder kommst nicht mehr rauf. [...] Kommst nicht mehr rauf, komm mach jetzt noch einmal eine Acht, und wir summen noch einmal ein wenig. Ein wenig raufsummen, oder. [...] Heute strahlen wir uns gegenseitig an, weil wir durch die Krise durch sind, miteinander und zwar erfolgreich und wertschätzend, oder. Und das ist auch schön, oder. [...] Du hast nie ausgelernt.“ (10:241 ff.)

Herr Leimbacher (10) reagiert in der anberaumten Betreuungssequenz behutsam auf den emotionalen Zustand des Mädchens, interveniert zwar auf einer stimmtechnischen Ebene, verknüpft aber den positiven Ausgang der kritischen Situation hauptsächlich mit der gewonnenen Beziehungsqualität.

Aus der Beschreibung des Lehrers lässt sich ableiten, dass das Mädchen auf dem Höhepunkt der Mutation ist, mitten in der Phase 2B (vgl. Tabelle 2, S. 97) nach Gackle (2000): Die Stimmlage sinkt, es werden Registerbrüche hörbar und das Phonieren der Randstimme wird vermieden, weil es Unbehagen auslöst. Der Lehrer ordnet den Stimmwechsel dem ihm bekannten Phänomen der Mutation bei Jungen zu, macht dies einerseits an der hörbar tieferen Stimmlage fest und vermutet andererseits, dass auch die Größe des Mädchens einen Einfluss hat. Zudem beobachtet er, dass die Randstimmfunktion beim Mädchen zu Irritation führt, verweist aber nicht auf spezifisches Fachwissen. Herr Leimbacher (10) ist nicht weiter beunruhigt¹⁵³, sondern erklärt dem Mädchen die Funktion von Voll- und Randstimme, um Letztere anschließend mit verschiedenen Übungen wieder zu „lancieren“ (10:248). Durch den ausgeprägten Stimmwechsel hin zu einer tieferen Stimme sind die muskulären Anpassungsleistungen, die üblicherweise bei einem Mädchen in der Mutation nicht zu erwarten sind, relativ groß. Herr Leimbacher (10) hilft dem Mädchen insbesondere über diese Klippe hinweg. Die Intervention erfolgt auf einer rein musikalischen Basis: „Und dann haben wir da miteinander geübt.“ (10:245)

Der positive Effekt wird aber hauptsächlich auf der Beziehungsebene eingeordnet. Das Gelingen macht Herr Leimbacher (10) fest an der gegenseitigen Wertschätzung, die Vertrauen vonseiten der Schülerin voraussetzt. Es ist ihm offensichtlich gelungen, eine Unterrichtsatmosphäre aufzubauen, wo Krisen nicht zu Katastrophen führen, sondern bearbeitet werden. Es wurde bereits in Kapitel 12 deutlich, dass dieser Lehrer bewusst auf die Beziehungsebene achtet und sowohl musikalische als auch persönlichkeitsbildende Ziele im Singunterricht verfolgt. Die Krise ist außerordentlich, denn die Schülerin ist emotional überfordert und beginnt zu weinen. Als Sofortmaßnahme entscheidet sich der Lehrer, der Jugendlichen einen sicheren Hafen zu bieten, um die emotionale Situation zu deeskalieren. Sie darf sich mit einer Mitschülerin auf die Toilette zurückziehen,

¹⁵³ Im Gespräch mit einer Kinderärztin wurde deutlich, dass hier möglicherweise ein pathologisches Phänomen zu untersuchen wäre. Eine mögliche Überproduktion von Testoron wird aus medizinischer Sicht als Abweichung betrachtet und müsste entsprechend untersucht werden (Zwicky, 2016). Herr Neuendorf (08) macht die Beobachtung, dass Mädchen mit Migrationshintergrund tiefere Stimmen haben (08:232 ff.).

wird in dieser Situation nicht allein gelassen. Herr Leimbacher (10) scheint die heikle Situation souverän behandeln zu können. Es gibt keinen Grund anzunehmen, dass im Unterricht verletzende Handlungen vorausgegangen sind, denn aus dem Porträt über Herr Leimbacher (10) kann ein human-anerkennender Unterrichtsstil (Tellisch, 2015, S. 252) geschlossen werden (vgl. 12.1).

Neben der individuellen Ebene betrifft die Situation auch das soziale Gefüge der Klasse. Als Mädchen ist sie mit dem markanten Stimmwechsel in der Klasse und insbesondere in der Gruppe ihrer Peers ein Spezialfall. Sie kann die tiefe Stimmlage nicht einordnen und ist gezwungen, sich mit stereotypen Vorstellungen zur Tonhöhe der Frauenstimme auseinanderzusetzen. Sweet (2018) beschreibt ganz allgemein eine große Unsicherheit bei Mädchen während des Stimmwechsels, im obigen Beispiel wird diese durch die auffällige Tiefe der Stimme zusätzlich akzentuiert. Herr Leimbacher (10) hat intuitiv die Idee der Studie von Siple (1994) nachvollzogen, indem er in seiner Intervention Wissen zur Stimme mit adäquaten Stimmübungen kombiniert und damit einen Effekt im gestärkten Selbstkonzept auslöst. Das Mädchen konnte so vokal und emotional gestärkt und damit in die Klasse integriert werden.

In diesem speziellen Fall wurden auf der individuellen Ebene (Pubertät und Adoleszenz) und bezüglich der sozialen Komponente (Jugend, vgl. Abbildung 1) spezifische Auswirkungen der Stimmentwicklung beschrieben. Obwohl die Intervention nur auf der vokal Ebene erfolgt ist, zeigen sich Auswirkungen auf das gesamte System. Bezüglich Assessment gelingt es dem Lehrer, eine fehlertolerante Atmosphäre zu generieren, in der eine individuelle Förderung stattfinden kann. Es steht vor allem der Bereich AFL im Zentrum. Herr Leimbacher (10) ist gleichsam gezwungen, eine Diagnose vorzunehmen, um der Schülerin die notwendige Förderung angeeignet zulassen. Er wechselt dabei in die Rolle des Mentors, indem er ermutigt und fördert, und in die Rolle des Guides, indem er diagnostische Daten sammelt, um durch die Krise zu manövrieren (vgl. Tabelle 4). Als Kritik könnte einzig angebracht werden, dass das Phänomen nicht aufgrund von regulären Assessments, z. B. zur Entwicklung der Sprechstimmlage (vgl. Greuel & Horst, 2011), antizipiert wurde, sondern in der Krise wahrgenommen wurde.

PPS – poor pitch singer

In der vorliegenden Arbeit wurde das Thema *Brummer* bis jetzt nicht vertieft. Bei der Erstellung des Kategoriensystem zeigte sich dieses Phänomen aber als Nebenschauplatz der adoleszenten Stimmentwicklung. So wurde induktiv in die Subkategorie *UMGANG MIT STIMMVERÄNDERUNG* die Unterkategorie *BRUMMER* eingefügt, die allerdings nur in Abgrenzung zu Kodierungen von Mutationsphänomenen bei Mädchen und Jungen zum Einsatz kam. Mit Blick auf Assessment sind in dieser Arbeit Überzeugungen und Wissensbestände zum Umgang mit Brummern relevant, auf der einen Seite bezüglich der Diagnose und Wahrnehmung des Phänomens und auf der anderen Seite bezüglich der Förderung der betroffenen Schülerinnen und Schüler. Nachfolgend wird zuerst der Begriff Brummer klarer konturiert, um anschließend die Kodierungen zu analysieren und exemplarisch ein Beispiel zu besprechen.

Spital (2004) nennt zwei Kategorien von Singenden, die als Brummer eingestuft werden können, einerseits die Sprechsänger, welche gar nicht in den Modus des Singens kommen, und andererseits die Tiefsänger¹⁵⁴, die eine eigene Tonhöhe wählen und sich nicht in den Gesang der anderen einfügen (S. 9 f.). Auch in den verschiedenen besprochenen englischen Skalen (Rutkowski, 2019; Wise & Sloboda, 2008) wird Singen mit Sprechstimme bzw. wenig erkennbarer Kontur als tiefstes Niveau genannt (vgl. 5.2) und mit einer nicht bzw. noch nicht vollständig funktionierenden Singstimme in Verbindung gebracht. Einzig J. Hasselhorn (2015) verzichtet in seiner Skala (vgl. Tabelle 9) auf den Einbezug dieser Singqualität und ordnet diese als nicht sinnvoll bewertbar ein (S. 82). Merrill (2020) verweist in ihrer Arbeit zum Thema auf den englischen Begriff *poor pitch singer* (vgl. Turøy, 2018), der das Phänomen der ungenauen Intonation wörtlich beschreibt und dies zudem geschlechtsneutral formuliert. Der Begriff wird deshalb in dieser Arbeit mit der Abkürzung PPS übernommen. Zudem grenzt sie PPS von Menschen mit Amusie ab, die eine Störung beim „Erlernen, Abspeichern und Wiedererkennen von Melodien“ (S. 89)

¹⁵⁴ Spital verwendet das generische Maskulinum, ohne explizit auf Jungen zu verweisen.

haben¹⁵⁵. PPS werden meist im Rahmen von elementarer Musikpädagogik in der Primarschule mit einfachen Übungen trainiert, wobei mit Liftübungen auf den Registerwechsel fokussiert wird (Bruhn, 1991; A. Mohr, 2008; Brünger, 2011; Merrill, 2020). PPS haben somit meist untrainierte Singstimmen und bräuchten entsprechend individuelle Hilfestellung, um diesen Mangel auszugleichen¹⁵⁶.

In der QIA wurden 19 Kodierungen zu der Subkategorie *BRUMMER* vorgenommen. Eine Gruppe von vier Lehrkräften (04, 05, 06, 08) hat sich gar nicht zum Thema geäußert, ein Lehrer (01) nur implizit. Obwohl nur Mutmaßungen angestellt werden können, lohnt sich trotzdem der Blick auf die fünf spezifischen Situationen. Bei Herrn Hilfiker (04) lässt sich ein Rückschluss aufgrund seiner Überzeugung, Singen sei nicht lernbar, ziehen. So ist zu vermuten, dass er PPS aufgrund seiner Überzeugung nicht als untrainierte, sondern als untalentierte Sängerinnen und Sänger wahrnimmt. Frau von Rohr (06) macht kaum Aussagen zur Mutation von Jungen, weshalb vermutet werden kann, dass sie PPS nicht spezifisch wahrnimmt. Frau Basler (05) hingegen betreut die Kinder auch in der Primarstufe, was evtl. darauf hindeutet, dass sie auffallende PPS bereits dort behandelt, und Herr Neuendorf (08) beschreibt die Veränderungen der Mutation bei Mädchen und Knaben so detailliert, geht sogar auf ihre jeweilige Tessitura ein, dass davon ausgegangen werden darf, dass PPS bei ihm entsprechend als trainierbares Phänomen behandelt wird. Herr Brugger (01) geht nicht speziell auf Brummer ein, ist aber der Meinung, bei Jugendlichen sei noch viel zu retten (01:283 ff.).

Neben dieser Gruppe von Lehrkräften, die das Thema nicht oder nur implizit erwähnen, zeigen sich verschiedene Herangehensweisen. Einerseits nehmen Lehrkräfte wahr, dass PPS in der Klasse mehr oder weniger häufig

¹⁵⁵ Frau Derendinger (11) berichtet von einem Mädchen, das einerseits ungenau intoniert und andererseits ihre eigene Stimme auf dem Tonband nicht erkennt: „[...] das sei nicht sie. Das töne in ihrem Kopf ganz anders.“ (11:280) Dies ist ein starker Hinweis auf Amusie, die in erster Linie pathologische Störungen des Hirns beschreibt (Evers, 2012), aber als kongenitale Amusie auch angeboren sein kann (Schlesiger & Evers, 2005). Interessanterweise betrifft das Phänomen nicht Sprachfähigkeiten, sondern nur die musikalische Tonhöhe (Ayotte et al., 2002).

¹⁵⁶ Auf einen ausführlichen Literaturreview wird an dieser Stelle verzichtet, weil insbesondere die Betreuung im Vordergrund steht.

anzutreffen sind, ziehen aber nicht in Betracht diese zu fördern oder sind der Überzeugung, dass keine Besserung möglich ist. Auf der anderen Seite werden PPS trainiert und in den Klassenklang eingegliedert. In der ersten Gruppe drückt es Herr Rothenburg (07) so aus: „Nein, ich habe eigentlich die, die von Anfang an schlecht sind, die bleiben schlecht.“ (07:314 f.) Frau Derendinger (11) ist ebenfalls dieser Meinung, nimmt aber zusätzlich wahr, dass Jungen nach der Mutation nur noch Tiefsänger sind und nicht mehr wissen, „wie oder sie wollen einfach auch nicht“ (11:159 f.). In dieser Aussage wird das Phänomen PPS zusätzlich mit der Mutation der Jungen verknüpft, indem die Lehrerin beobachtet, dass die Mutation zu einer Einschränkung des Stimmgebrauchs führt, evtl. gekoppelt mit Desinteresse, sich vokal auszudrücken. Auch Herr Berner (09) stellt fest, dass es „erschreckend viele“ (08:99) gibt, welche die Töne nicht treffen, er nennt aber eine erste Methode, um den PPS zu begegnen. Er stellt diese zwischen korrekt singende Sängerinnen und Sänger, um ihnen vokale Vorbilder an die Seite zu stellen. Zusammen mit Frau Fricker (03) ist er zudem der Meinung, das Phänomen sei geschlechtsübergreifend anzutreffen, vermehrt aber bei Jungen (03:33 ff.). Frau Fricker (03) ist ebenfalls unsicher, ob viel dagegen gemacht werden kann, bringt aber eine weitere Maßnahme ins Spiel: So macht sie Sprechende darauf aufmerksam, zur Singstimme zu wechseln (03:386 ff.). Als Hindernis nennt sie allerdings die besonders großen Hemmungen während der Pubertät (03:400 ff.). Herr Leimbacher (10) und Herr Buchser (02) sind der Überzeugung, PPS trainieren zu können. Herr Buchser (02) weist zudem auf die geringe Anzahl von PPS hin¹⁵⁷, dies im Gegensatz zu Herr Berner (09), der von einer großen Anzahl spricht.

Einerseits beweisen hier Lehrkräfte, dass PPS auch in der Sekundarstufe 1 trainiert werden können und folgen damit den beschriebenen Prozeduren, die für die elementare Musikpädagogik entwickelt wurden (A. Mohr, 2008; Brünger, 2011), andererseits zeigt sich aber auch die Überzeugung, es könne wenig bis nichts gemacht werden, um PPS zu kurieren, was auf

¹⁵⁷ Die Einschätzung von Herrn Buchser (02) deckt sich mit der Erfahrung des Forschenden, der nur ganz wenige PPS bewusst wahrgenommen hat und diese meist trainieren konnte. Ausgenommen davon war ein Geschwisterpaar, das auch nach vier Jahren Singunterricht keine Melodiekontur korrekt wiedergeben konnten. Eventuell handelte es sich dabei um eine Form von angeborener Amusie (vgl. Fußnote 155).

Unwissenheit bezüglich „foundational knowledge about the physiological [...] changes“ (Freer, 2020, S. 3) hindeutet. Dies bedeutet wiederum, dass eine ganzheitliche und individuelle Betreuung nur möglich ist, wenn entsprechendes CK und PCK vorhanden ist. Herr Buchser (02) steuert als Highlight seines Unterrichts die Behandlung eines PPS bei, die anschließend exemplarisch besprochen wird.

Herr Buchser (02) ist bereits durch einen proaktiven und öffentlichen Umgang mit der Mutation von Jungen aufgefallen. Nun beschreibt er als Highlight seiner Tätigkeit als Musiklehrer, wie er einen PPS zum Singen gebracht hat:

„Also ein Highlight, ich habe einen Brummer gehabt. [...] Auf diese 1600 Schüler hat es einen Brummer gehabt. Ein ganz klarer Brummer. Also der hat gesungen, [...] und ich habe gemerkt, der singt einen Ton. Also einen kleinen tonalen Bereich. Und der hat mir so vertraut, und ich ihm auch, ich habe gesagt, ich lehre dich singen. [...] Und dann haben alle gehört, dass der jetzt singen kann. Dann hat die ganze Klasse applaudiert. Ich meine, das war Hühnerhaut.“ (02:723 ff.)

Ähnlich wie bei der beschriebenen Intervention zur auffällig tiefen Stimme eines Mädchens weiter oben, zeigt sich bei Herrn Buchser (02) der Unterricht als Ort des Vertrauens. Die Therapie des PPS findet direkt in der Klasse statt und wird am Ende gefeiert. Der Lehrer arbeitet auf musikalischer Basis an der Erweiterung der Tessitura. Er stellt den Schüler direkt neben sich, lässt ihn im Singkreis aber nach außen drehen, um die Konzentration zusätzlich zu verbessern und singt ihm direkt ins Ohr. Durch die Nähe haben beide die Möglichkeit der Erfolgskontrolle. Der Erfolg stellt sich ein und gipfelt im finalen Auftritt vor der Klasse. Zwar beschreibt Buchser (02) den musikalischen Erfolg, ebenso wichtig ist aber der Auftritt bzw. die damit verbundene Persönlichkeitsentwicklung, die sich bei Herrn Buchser (02) bereits mehrfach als Zielsetzung gezeigt hat.

Obwohl die Thematik PPS nur gestreift werden kann, zeigen sich doch bezüglich Assessment wichtige Hinweise. Bevor Lehrkräfte in der Lage sind, einer spezifischen Herausforderung zu begegnen, brauchen sie entsprechendes CK, um diese wahrnehmen zu können. Zudem ist PCK notwendig, um eine entsprechende Diagnose zu stellen und anschließend mit einer adäquaten Methode eine Förderung einzuleiten. Die Beispiele aus den Interviews zeigen nun, dass fehlende Kenntnisse zu einem Fehlkon-

zept bezüglich PPS führen und Schülerinnen und Schüler nicht gefördert werden. Defizite werden vielleicht beobachtet, aber nicht bearbeitet. Die Bearbeitung von individuellen Lernständen ist aber der zentrale Aspekt von Assessment.

Koedukative und monoedukative Settings

Bereits im autoethnografischen Teil der Arbeit wurden monoedukative Sequenzen als hilfreiches Setting beschrieben (siehe Kapitel 3.3). Kalmbach (2010) votiert sogar für Geschlechtertrennung beim Singen, um Gleichberechtigung herzustellen. Obwohl aktuelle Studienresultate zu Singleleistungen in monoedukativen Settings fehlen (vgl. Meier, 2008), vermutet Hess (2015) doch für performative Leistungen erweiterte Handlungsoptionen bei monoedukativen Settings (S. 330). Auch in Studien aus dem anglofonen Raum wird das Trennen der Geschlechter beim Singen als hilfreiches oder sogar notwendiges Setting beschrieben (Ashley, 2015; Orton & Pitts, 2019). In der vorliegenden Arbeit wurde der Anteil an monoedukativen Unterrichtssequenzen in der skalierenden Analyse der Sozialformen bereits untersucht (vgl. 11.3). An dieser Stelle interessieren nun die Beweggründe für temporäre Geschlechtertrennung, immer mit dem Blick auf Förderung und Beurteilung.

Die Frage nach monoedukativen und koedukativen Unterrichtssequenzen wurde im Interviewleitfaden vorgegeben und in den Gesprächen individuell vertieft. Entsprechend wurde im Kategoriensystem die Subkategorie *SOZIALFORMEN MIT FOKUS AUF DAS GESCHLECHT* mit den Subkategorien *MONOEDUKATIV* und *KOEDUKATIV* auf einer dritten Ebene ergänzt. Die Kodierungen dieser Subkategorien führten bereits im Rahmen der skalierenden Analyse der Sozialformen (vgl. 11.3) zu der Variable *GESCHLECHTERGETRENNT*, bei der fünf Lehrkräfte einen Unterrichtsanteil zwischen 10–40 % erreichten. Der Einbezug der Subkategorie *FORMEN DER BEWERTUNG* aus *ASSESSMENT OF LEARNING* zeigt, dass eine weitere Lehrkraft zumindest in Beurteilungssituationen auf eine Trennung achtet und zudem Situationen genannt werden, in denen innerhalb der ganzen Klasse Mädchen oder Jungen spezifisch angesprochen werden. In diesem Abschnitt wird deshalb auf drei Konstellationen fokussiert, zuerst auf Unterricht, in dem keine monoedukativen Sequenzen durchgeführt werden, anschließend auf monoedukativ

durchgeführte Sequenzen und ihre Begründungen und zuletzt auf koedukative Sequenzen mit direkter Ansprache an Mädchen und Jungen.

Eine knappe Mehrheit der Lehrkräfte erwähnt keine monoedukativen Settings. Vier Lehrkräfte sind bereits in der Skalierung der Sozialformen (vgl. 11.3) durch einen konsequent frontalen Unterricht aufgefallen (01, 05, 06, 07). Herr Brugger (01) und Frau Basler (05) gestalten ihren Unterricht bewusst als Chorsetting, beide sind auch entsprechend ausgebildet. Herr Brugger (01) begründet dieses Setting zusätzlich mit seiner Angst, möglicherweise jemanden bloßzustellen (01:258), bei Frau Basler (05) wird der Singunterricht in der Sekundarstufe 1 als Übergang zum Chor auf der Sekundarstufe 2 gestaltet¹⁵⁸. Frau von Rohr (06) und Herr Rothenburg (07) unterrichten ebenfalls konsequent frontal, ohne dies speziell zu begründen. Herr Rothenburg (07) gibt allerdings den Jugendlichen einen geschützten Rahmen für das Vorsingen, um deren Schamgefühle nicht zu verletzen (07:367 ff.). Zusätzlich sprechen zwei Lehrkräfte, die bewusst Gruppenarbeiten mit vokalen Aufgabenstellungen durchführen, nicht von monoedukativen Settings (02, 03). Ob die Gruppenbildung jeweils zu einer Geschlechtertrennung führt, geht aus dem Datenmaterial nicht hervor. In dieser Gruppe wird mehr oder weniger bewusst auf monoedukative Sequenzen verzichtet, eine Trennung findet höchstens in der Bewertungssituation statt, um Scham vorzubeugen. Als Begründung für das frontale Setting wird die Möglichkeit von bloßstellenden Situationen angeführt, es bleibt somit eine gewisse Anonymität bestehen. Auch Ashley (2015) fordert explizit für die Zeit der Mutation einen geschützten Rahmen im Klassenchorsetting, um sich mit der neuen Beschaffenheit der Stimme auseinandersetzen zu können (S. 115). Insbesondere im Chorsetting wird dabei aber nicht beachtet, dass im Rahmen des mehrstimmigen Singens beim Proben einzelner Stimmen durchaus bloßstellende Momente in Kauf genommen werden. Genau dies wird aber in der nächsten Gruppe in einem Fall als Begründung für eine partielle Trennung der Geschlechter angeführt.

Fünf Lehrkräfte trennen Jungen und Mädchen zugunsten monoedukativer Unterrichtssequenzen. Die konkreten Szenarien unterscheiden sich

¹⁵⁸ Zur Erinnerung: Frau Basler (05) unterrichtet an einer Waldorfschule und hat das Ziel, die Jugendlichen in der Sekundarstufe 1 im Wesentlichen zu begleiten (vgl. Kapitel 12.1/Die bewahrende Sängerin in der Waldorfschule).

aber deutlich¹⁵⁹. Herr Berner (09) führt ähnlich wie Herr Rothenburg (07) das Argument der Scham an und ermöglicht für Beurteilungssituationen eine Geschlechtertrennung. Bei ihm findet das Vorsingen aber mehrfach im Jahr statt, mit dem Ziel, die Auftrittskompetenz zu fördern (09:211 ff.). Zudem wird jeweils sofort ein Feedback gegeben, das laut Herrn Berner (09) gut akzeptiert wird (09:279 ff.). Er (09) schafft so regelmäßig Raum für die von Hess (2015) vorgeschlagenen erweiterten Handlungsoptionen. Herr Neuendorf (08) geht sogar noch einen Schritt weiter, indem er in monoedukativen Phasen geschlechtsspezifisches Repertoire einsetzt (08:110 ff.), wie dies Freer (2020) für die Jungen fordert. Gegen Ende der Schulzeit unternimmt er dann den Versuch, über das Singen und Gestalten von Musicalsongs die Geschlechter wieder bewusst zu vereinen (08:320 ff.). Beide Lehrer sehen das Singen spezifisch als Persönlichkeitsentwicklung (vgl. 12.2) und geben den Mädchen und Jungen die Möglichkeit, in einem geschützten Raum ihre *vocal identity* (Ashley, 2015) zu entwickeln.

Drei der fünf genannten Lehrkräfte führen monoedukative Sequenzen durch, sprechen aber Jungen und Mädchen auch in der ganzen Klasse direkt an. Herr Leimbacher (10) entscheidet situativ, ob eine Trennung im Sinne einer Registerprobe (10:279 ff.) oder aufgrund von Auffälligkeiten bezüglich der Mutation (10:175 ff.) notwendig ist, probt aber einzelne Stimmen auch in der ganzen Klasse (10:320 ff.). Der Unterricht von Frau Derendinger (11) findet in zwei verbundenen Räumen statt, was ihr die Möglichkeit gibt, regelmäßig und ohne organisatorischen Aufwand geschlechtergetrennt zu arbeiten (11:179 ff.). Als Begründung nennt sie die Konkurrenz zwischen Mädchen und Jungen (11:183 ff.). Sie spricht auch davon, dass Jungen in diesen Sequenzen ihre Ehre weniger verlieren, wenn es nicht so gut klingt (11:231 ff.). Die Konkurrenz nutzt sie allerdings im Anschluss im Klassensetting, um einen sozialen Druck aufzubauen

¹⁵⁹ Eine weitere Lehrerin aus der ersten Gruppe berichtet von einem Unterrichtsexperiment, das sie allerdings nicht dazu bewogen hat, monoedukative Sequenzen in ihren regulären Unterricht einzubauen. Während eines ganzen Schuljahres hat Frau Fricker (03) die Mädchen und Jungen zweier Klassen in einer der beiden Lektionen zusammengenommen, um einen geschlechtergetrennten Unterricht durchzuführen. Ihre Erfahrungen sind positiv, insbesondere erwähnt sie die Möglichkeit, mehrstimmig zu singen. Als Nachteil beschreibt sie das unterschiedliche Lerntempo von Jungen und Mädchen und die Anstrengung, während einer ganzen Lektion zu singen.

(11:186 ff.), indem abwechseln zwei Zeilen gesungen werden. Sie meint: „Das funktioniert auch noch recht gut. Weil da gibt es natürlich weniger, wo sie sich dann [...] quasi unten rausdrücken können und sagen, sie hat jetzt nicht gehört, dass ich nicht mitgemacht habe. Wenn dann nur noch fünf singen, dann hört man es dann schon.“ (11:189 ff.) Im Gegensatz zu den vorherigen Beispielen, in denen Scham als Begründungsmuster für monoedukative Sequenzen dient, befeuert Frau Derendinger (11) bewusst stereotype Vorstellungen eines Geschlechterkampfes. Der Einsatz von monoedukativen Unterrichtssequenzen scheint bei Herrn Hilfiger (04) eher zufälligen Charakter zu haben, z. B. wenn er die Knaben rausschickt, um den Mädchen mehr Selbstbewusstsein zu vermitteln (04:181 ff.). Aber auch er inszeniert in der frontalen Situation eine Konkurrenzsituation, beschreibt es aber lustvoller als Frau Derendinger (11). Er sagt: „manchmal mache ich [...] so ein Battle, Mädchen gegen Knaben, und dann müssen sie es gegeneinander vorsingen, und dann gibt es ein wenig Bewertung und dann dürfen die nächsten nachher einen Song auswählen.“ (04:230 ff.) Er beschreibt die Situation als Anreizsystem, um Energie zu gewinnen und zu zeigen, dass eine Leistungssteigerung möglich ist (11:392 ff.). Freer (2020) nennt Wettbewerb in der Klasse als gewünschtes Element von Jungen, ob damit eine Fokussierung auf die Geschlechterdifferenz gemeint ist, gibt er allerdings nicht an.

In der ersten beschriebenen Gruppe wird keine spezifische Förderung durch das Trennen von Mädchen und Jungen in Betracht gezogen, teilweise aufgrund der Angst, bloßstellende Momente zu generieren. In der zweiten Gruppe werden beide Geschlechter bewusst partiell getrennt, es sind aber zwei verschiedene Beweggründe auszumachen. Einerseits werden bewusst Räume für erweiterte Handlungsoptionen (Hess, 2015) geschaffen, um die Entwicklungen von *vocal identity* (Ashley, 2015) einzuleiten und Persönlichkeit aufzubauen, oder um musikalisches Lernen zu begünstigen. In diesem Fall ermöglicht ein *doing gender* einen Ausgleich von Differenzen, meist begründet mit der Beobachtung, so Scham vorbeugen zu können. Es zeigen sich sogar Momente von *undoing gender*, wenn nach monoedukativen Unterrichtsphasen bewusst gemischte Gruppen gebildet werden, oder Chorproben nur situativ aufgeteilt werden. Andererseits werden bewusst Differenzen betont und als Battle inszeniert, um Leistung einzufordern. *Doing gender* führt in diesem Fall nicht zu einem Ausgleich, sondern zu einer Tradierung stereotyper Rollenbilder. Insgesamt können in den Inter-

views im Zusammenhang mit monoedukativen Unterrichtssequenzen einige der von Freer (2020) gesammelten Wünsche von Jungen beobachtet werden, z. B. fähigkeitsbasierte musikalische Erfahrungen und Feedback sowie Wettbewerb und spezifisches Repertoire (S. 10).

Zusammenfassend zeigt sich, dass der Wunsch nach Förderung bei knapp der Hälfte der Interviewpartnerinnen und -partner zu monoedukativen Phasen führt, allerdings mit unterschiedlichen Beweggründen und im Rahmen des Genderdiskurses ambivalenter Bewertung. Neben dem formativen Assessment zeigt sich auch in Beurteilungssituationen fast durchgehend eine Geschlechtertrennung, wenn die Bewertung nicht sowieso in der Einzelsituation durchgeführt wird. Zuletzt bleibt an dieser Stelle das Gewicht dieser monoedukativen Sequenzen auf den gesamten Musikunterricht einzuordnen. Bei den fünf genannten Lehrkräften macht dieser 10–40 % ($M = 26\%$) aus, was im Schnitt deutlich weniger als die Hälfte des Unterrichts ist. Das gemeinsame Singen in der ganzen Klasse bleibt für alle zentral (vgl. 11.3).

Die Relevanz geschlechtsspezifischer Einflüsse auf Assessment

In diesem Abschnitt wurde untersucht, ob die im theoretischen Teil der Arbeit (vgl. Kapitel 4) diskutierten geschlechtsspezifischen Einflüsse auf Diagnose, Förderung und Beurteilung auch im vorliegenden Datenkorpus sichtbar werden oder diese sogar bestätigen. Trotz großer Heterogenität in den Aussagen zeigen sich einige Tendenzen, welche deren Relevanz direkt oder indirekt untermauern. Die Argumentation erfolgt nicht durch Repetition und Zusammenfassung der aufgeführten Belege, sondern durch drei Beispiele, die sich dem Thema unterschiedlich annähern.

In dieser Arbeit wurde mehrfach auf eine ganze Reihe von Erkenntnissen aus der englischsprachigen Musikpädagogik verwiesen, die im deutschsprachigen Raum nicht oder nur ungenügend rezipiert wurden. Aus diesem Grund war nicht zu erwarten, bei den interviewten Lehrkräften auf spezifisches Wissen (CK) und entsprechende Methoden (PCK) zu treffen. Trotzdem berichteten einige Lehrkräfte von Unterrichtssituationen, in denen spezifisch die adoleszente Stimmentwicklung im Zentrum steht. So wird in Einzelfällen bewusst die Tonhöhe von Liedern für mutierende Stimme angepasst und mehrere Lehrkräfte schaffen einen geschützten Raum der Entwicklung mit monoedukativen Unterrichtssequenzen. Die

Relevanz kann in diesen und weiteren Fällen direkt aus den positiven Auswirkungen der praktischen Beispiele geschlossen werden.

Auf indirekte Art zeigt sich die Relevanz an Merkmalen der adoleszenten Mädchenstimme. Das labile vokale Selbstkonzept und die Vermeidung der Randstimme wird in der Literatur beschrieben, von einigen Lehrkräften aber nur über die Beobachtung der Auswirkungen wahrgenommen, interveniert wird nicht. Andere versuchen den Phänomenen zu begegnen, sind sich aber über die Methode nicht im Klaren und nochmals andere scheinen mit einer profunden Chorleiterausbildung die Probleme zu umschiffen. Indirekt zeigt sich hier die Relevanz durch den Wunsch, auf ein Phänomen reagieren zu können. Dazu wäre aber eine treffende Verortung der Merkmale adoleszenter Stimmentwicklung und das Wissen um entsprechende Methoden hilfreich.

Nur im Rahmen eines Umkehrschlusses zeigt sich die Relevanz bei einigen Lehrkräften durch Fehlkonzepte. So fußt die Meinung, die Mutation von Jungen könne im Unterricht nicht adäquat behandelt werden, auf Unwissen, was auch als Verweigerung der Betreuungsleistung gelesen werden kann. Zudem werden in diesen Fällen stereotype Vorstellungen sichtbar, im Gegensatz zu den Beispielen, in denen Lehrkräfte bewusst mit der adoleszenten Stimmentwicklung umgehen.

Die Relevanz der adoleszenten Stimmentwicklung und geschlechtsspezifischer Eigenheiten ist aufgrund der Beispiele klar ausgewiesen und wird im nächsten Abschnitt thematisch mit den drei Bereichen von Assessment verknüpft. Es muss allerdings an dieser Stelle darauf hingewiesen werden, dass aufgrund des Forschungsdesigns Aussagen der Jugendlichen zur Relevanz fehlen (vgl. 15.1).

13.2 Ganzheitliches Assessment als Perspektive

Das Modell *assessment for learning, as learning* und *of learning* (Earl, 2013) wurde bereits ausführlich diskutiert. Es wurde dargelegt, dass Assessment den Lernprozess ganzheitlich von der Erfassung des Lernstands zu Beginn, bis hin zur abschließenden Beurteilung am Ende einer Lernphase begleiten soll. Terhart (2014, S. 884) konstatiert zwar eine allgemeine Tendenz weg von der summativen Bewertung hin zur formativen Förderung und Beurteilung in der Schule, Earl (2013) ist aber der Über-

zeugung, der überwiegende Anteil von Assessment im Unterricht sei summativ. Um diesem Umstand Rechnung zu tragen, ändert er in seiner Darstellung zu den unterschiedlichen Absichten, Referenzpunkten und Rollen des Assessments die konzeptionell vorgesehene Reihenfolge *for*, *as* und *of* auf die vermutete Gewichtung *of*, *for* und *as* (vgl. Tabelle 27). Damit folgt nicht summatives auf formatives Assessment, wie es die Logik formativer und summativer Evaluation ursprünglich vorsah (vgl. Scriven, 1967; Bloom et al., 1971) und im Rahmen der Betonung formativer Beurteilung auch aktuell (z. B. C. Schmidt, 2020) betont, sondern die abschließende Beurteilung AOL wird ins Zentrum gerückt. Diese Gewichtung wird auch von Lehmann-Wermser (2019) beobachtet. Er konstatiert zwar für den musikpädagogischen Bereich in Deutschland eine Tendenz hin zu formativen Elementen in der Notengebung, sieht diese aber auf den summativen Charakter von halbjährlichen Zeugnissen und somit auf summative Belege reduziert. Weiter wurde bereits dargelegt, dass auch in der musikpädagogischen Literatur zu Assessment das Interesse vornehmlich validen Bewertungssystemen musikalischer Leistungen gilt (vgl. J. Hasselhorn & Lehmann, 2015) und holistische Modelle zwar postuliert, aber nicht als konzeptuelle Unterrichtsstrategie ausformuliert wurden (vgl. Brophy, 2019c). Die genannten Argumente lenken den Blick auf eine erste wichtige Untersuchungsebene, nämlich auf die Frage, welche Aspekte von Assessment die Lehrkräfte im Singunterricht wie gewichten und in welcher

Ansatz	Zweck/Absicht	Referenzpunkt	Beurteilender
<i>assessment of learning</i>	Beurteilungen über Einstufung, Beförderung, Zeugnisse usw.	Lerngruppe Lehrplan externe Anforderungen	Lehrkraft
<i>assessment for learning</i>	Grundlage und Adaption der Planung	Unterrichtsziele Lehrplan	Lehrkraft
<i>assessment as learning</i>	Selbstevaluation Selbststeuerung	persönliche Ziele Lehrplan	Schülerin/Schüler

TABELLE 27: *Eigenschaften von assessment of, for und as learning nach Earl (2013, S. 43)*¹⁶⁰

¹⁶⁰ Übersetzung durch den Autor.

Reihenfolge sie diese einbringen. Brophy ordnet diese Fragestellung auch unter dem principle of valuing ein (S. 925), indem betont wird, wie Förderung und Bewertung stark davon abhängen, welche Bedeutung den Themen gegeben wird (vgl. Kapitel 12). Im ersten Abschnitt wird deshalb die Konzeption der Förderung und Leistungsbeurteilung untersucht.

Anschließend werden die drei Bereiche *assessment of*, *for* und *as learning* in der von Earl vorgeschlagenen Gewichtung dargestellt und analysiert. Im zweiten Abschnitt werden die Ergebnisse zu *assessment of learning* beschrieben, wobei Transparenz und Planbarkeit der Beurteilung im Zentrum stehen. Im dritten Abschnitt wird *assessment for learning* vertieft. Es interessieren dabei weniger Methoden frontaler Unterrichtssettings als bereits etablierte individuelle Diagnose- und Fördermethoden. Im vierten Abschnitt rücken Schülerinnen und Schüler ins Zentrum, wenn *assessment as learning* betrachtet wird. Untersucht werden Aussagen, welche die Lehrkräfte zur Selbststeuerung der Lernenden und zu einer lernförderlichen Atmosphäre machen. Zusammenfassend wird eingeschätzt, wo sich bezogen auf das Modell von Earl (2013) Assessment-Elemente in den Interviews zeigen und wie diese einzuordnen sind.

Analyse der Assessment-Konzeption

Zuerst wird nun untersucht, wie die Interviewpartnerinnen und -partner AOL und AFL aufeinander beziehen. Grundlage der Analyse sind hauptsächlich die Subkategorien *FORMEN DER BEWERTUNG* und *TRANSPARENZ DER INTERPRETATION* aus der Subkategorie *ASSESSMENT OF LEARNING*, die in allen Interviews kodiert wurden. Die Lehrkräfte nennen durchgehend das *Vorsingen*¹⁶¹ als summative Beurteilung, immer verknüpft mit einem mündlichen Feedback. Ein Teil des Samples weist auf formative Anteile des Feedbacks hin. Nur eine Lehrkraft beschreibt innerhalb der Beurteilungssituation eine Selbstbeurteilung, die als metakognitiver Anteil AAL zuzuordnen ist. Da AAL als Teilbereich von AFL postuliert werden kann und sich im Datenkorpus insgesamt wenig zeigt, wird dieser Bereich nicht in

¹⁶¹ Vorsingen: Liedvortrag allein, in Gruppen, vor der Lehrkraft oder der Klasse mit summativer Beurteilung.

die Untersuchung der Gewichtung und konzeptuellen Verknüpfung von Assessment-Elementen einbezogen.

Am Beispiel der Aussage von Herrn Brugger (01) wird nun gezeigt, wie die Mehrheit der interviewten Lehrkräfte die Leistungsbeurteilung des Singens konzeptuell gestaltet. Er beschreibt die Beurteilungssituation folgendermaßen: „Ja, [...] die Eins-zu-eins-Situation, wo ich dann [...] eine Standortbestimmung machen kann. Und schaue, wo ich ihnen helfen kann.“ (01:148 f.) Aus dem Zusammenhang des Interviews wird deutlich, dass diese Situation im Rahmen einer summativen Bewertung stattfindet. Herr Brugger (01) meint mit Standortbestimmung eine summative Beurteilung (AOL), in der er zusätzlich in Betracht zieht, ein förderorientiertes Feedback (AFL) zu geben. Die Aussage erscheint je nach Perspektive in einem anderen Licht. Bezogen auf den Wortsinn der aus dem Zusammenhang isolierten Aussage könnte von der Intention eines formativen Assessments ausgegangen werden. Der Lehrer macht in der Eins-zu-eins-Situation eine Standortbestimmung mit dem Ziel, ein formatives Feedback zu geben. Typischerweise würde dieser Vorgang im Rahmen von AFL zu Beginn einer längeren Unterrichtsphase im Sinne einer diagnostischen Einschätzung vor der Planung oder in der Mitte einer Lernphase als Mittel zur Adaption der Planung stattfinden. Das Setting rückt allerdings den summativen Charakter der Beurteilung ins Zentrum. Aus dieser Perspektive

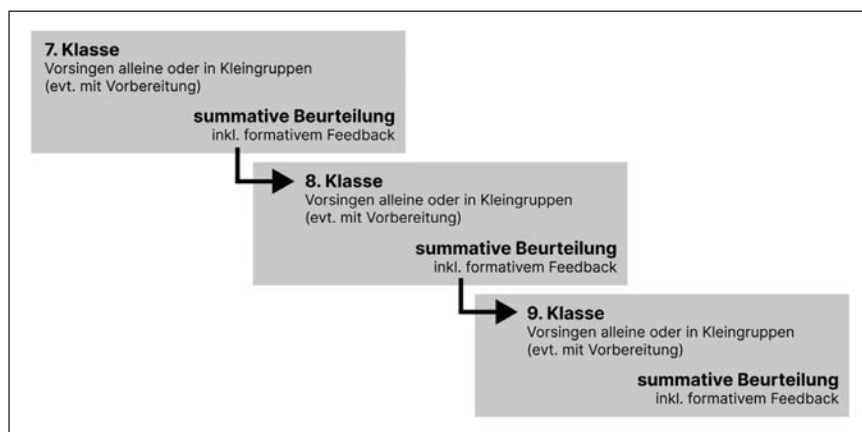


ABBILDUNG 12: Wirkung von formativem Feedback während dreier Schuljahre in der Sekundarstufe 1

erfolgt das Feedback hauptsächlich dazu, die Beurteilung transparent zu machen.

Fautley (2010) beschreibt diese Situation als „formative use of summative assessment“ (S. 198) und nimmt dabei die längerfristige Betreuung in den Blick. Er stellt fest, dass Feedback nach einer summativen Beurteilung dazu benutzt werden kann, um Ziele zu setzen und „for focusing pupil attention onto areas which need improvement“ (S. 198). Die summative Bewertung AOL wird zum Ausgangspunkt für ein Feedback, in dem eine Standortbestimmung durchgeführt wird und Ziele für eine folgende Phase gesetzt werden. Das Feedback als formatives Element AFL ist in diesem Fall der summativen Beurteilung nachgeschaltet. In der Abbildung 12 wird dieser zyklische Vorgang für den konkreten Fall der drei Schuljahre der Sekundarstufe 1 visualisiert. Schülerinnen und Schüler erhalten im Verlauf der drei Jahre von der 7.–9. Klasse jeweils gegen Ende des Schuljahres eine summative Beurteilung (AOL) mit einem formativen Feedback (AFL). Mit wenigen Ausnahmen zeigt sich dieser periodische Vorgang im ganzen Sample.

Es stellt sich aber die Frage, ob die von Fautley beschriebene Form von periodischem Assessment seine Wirkung tatsächlich entfalten kann. Mantie (2019) beschreibt Musikunterricht insgesamt als holistischen Förderprozess: „day-to-day instruction in music involves in situ informal assessments and evaluations.“ (S. 35) Unter dieser Prämisse macht es Sinn, im Singunterricht periodisch die Entwicklung der vokalen Kompetenzen zu überprüfen und individuelle Hinweise zur Entwicklung zu geben. Holcomb (2019) weist allerdings darauf hin, dass bei *Rubrics* mit einem holistischen Ansatz das Feedback zu unspezifisch ist, um wirksam werden zu können (vgl. Tabelle 8).

Im vorliegenden Fall könnte dieses Feedback jeweils in der 8. und 9. Klasse der Sekundarstufe 1 seine Wirkung entfalten (vgl. Abbildung 12). Aussagen zu einer längerfristigen Wirkung individueller Feedbacks werden aber nur von Herrn Brugger (01) und Frau Fricker (03) gemacht. Herr Brugger (01) benutzt im Klassensetting Handzeichen, um an Feedbacks zu erinnern: „durch das, [...] was im Vorsingen passiert ist, wo ich sie darauf hingewiesen habe, ist das irgendwie eine gute Basis, auch vor der Klasse, ich kann sie so ein wenig darauf hinweisen mit Zeichen.“ (01:63 ff.) Auch Frau Fricker (03) betont, dass Schülerinnen und Schüler sich an ihre Feedbacks aus der summativen Beurteilung erinnern: „Also dort gebe ich im-

mer, wirklich eine Rückmeldung und diese merken sie sich auch. Das habe ich schon ein paarmal festgestellt, das wissen sie dann noch. Das ist mir noch wichtig.“ (03:147 ff.) Das Feedback im Rahmen der summativen Beurteilung dient bei beiden Lehrkräften also nicht nur der Plausibilisierung der Singnote, sondern ist auch als formatives Element für die nächste Lernphase gedacht. Inwiefern Feedbacks der anderen Lehrkräfte ebenfalls wirksam werden, bzw. wie verbreitet die Wirkung in den Klassen ist, bleibt an dieser Stelle allerdings offen. Einige Lehrkräfte ermöglichen vor dem Vorsingen Übungsphasen, andere explizit nicht, Hinweise auf einen durch formative Assessments adaptierten Unterricht lassen sich aber nur an wenigen Stellen erahnen. Somit bestätigt sich im vorliegenden Sample die Dominanz von summativem Assessment, wie von Earl (2013) vermutet.

Ein Abgleich mit den *principles of assessment* (Brophy, 2019c) weist neben der Rochade von AFL und AOL auf eine weitere konzeptionelle Eigenheit des Vorsingens hin. Das Prinzip der *authenticity* verlangt nach authentischen Beurteilungssituationen. Schaumberger (2020a) startet in seinem Artikel zur Leistungsmessung mit der Beschreibung einer authentischen Beurteilungssituation beim Chorsingen. Der Chor ist aufgrund der Probenarbeit (AFL) in der Lage, das eigene Erleben und die vokale Leistung im Konzert selbst einzuschätzen (AAL) und wird belohnt mit der Einschätzung des Publikums, welches in Form von Applaus beurteilt (AOL). Alternativ wird der Chor im Wettbewerb von einer Fachjury beurteilt (S. 222), was deutlicher an AOL erinnert. Schaumberger betont zudem den emotionalen Aspekt des Auftritts. Im Jargon der Schule könnte gesagt werden, dass die Aufgabe gründlich geübt und anschließend geprüft wird. Der einzige Wechsel im Setting betrifft das Hinzukommen des Publikums.

In der beschriebenen Konzeption mit einem großen Anteil an gemeinsamem (Chor-)Singen und dem abrupten Wechsel in die solistische Situation beim Vorsingen zeigt sich ein Bruch im Setting der Aufgabenstellung, der die Authentizität infrage stellt. Insbesondere in Kombination mit einem rein frontalen Unterrichtssetting (01, 06, 07), in dem das Erarbeiten und Üben von Liedern in der ganzen Klasse stattfindet, ist der Wechsel in die Einzelsituation als starker Bruch zwischen Übungs- und Prüfungssituation einzuschätzen. Die in der skalierenden Analyse der Sozialformen (vgl. Abbildung 11) ausgewiesenen Anteile von 10 % Einzel- und Gruppenarbeit werden bei den meisten Lehrkräften entsprechend eingesetzt und verdeutlichen die Spezialsituation des Vorsingens.

Schülerinnen und Schüler sind in der Beurteilungssituation von einem Moment auf den anderen in vokaler und emotionaler Hinsicht auf sich gestellt. Möglicherweise könnten sie ein Lied als Mitsänger (Spital, 2004, S. 9 f.) in einer Gruppe richtig singen, auf sich alleine gestellt aber nicht. Zudem verändern der emotionale Stress und die Solo-Situation die Aufgabe grundsätzlich. Obwohl dieses Setting bewusst gewählt wird, um Bloßstellung zu verhindern (01:266 ff.), geht der Charakter des gemeinsamen Singens verloren und damit die Authentizität der Aufgabe. Einige Lehrkräfte haben diesbezüglich bereits Ideen entwickelt, indem einerseits das Vorsingen als Gruppenarbeit vorbereitet und vorgetragen und andererseits solistisches Singen im Klassensetting aufgebaut und beurteilt wird.

In diesem Abschnitt zeigte sich, dass summative Beurteilung teilweise dazu benutzt wird, förderorientiertes Feedback zu geben, wobei nur zwei Lehrkräfte davon berichten, später eine entsprechende Wirkung zu bemerken. Damit bestätigt sich das von Fautley (2010) beschriebene Modell formativer Nutzung summativer Beurteilung. Das vor allem im traditionell frontalen Unterricht zum Einsatz kommende Einzelsingen muss allerdings kritisch hinterfragt werden, da Training und Leistungsbeurteilung als in unterschiedlichen Settings stattfinden und damit das Prinzip der *authenticity* beschnitten wird.

Einige Lehrkräfte beschreiben nun ausgehend von Gruppenarbeiten und monoedukativen Settings auch Elemente individueller Förderung. Diese werden später genauer untersucht. Zuerst wird aber die summative Beurteilung der Lehrkräfte entlang der skizzierten Elemente von AOL (vgl. Tabelle 5) untersucht.

Assessment of Learning

Summative Beurteilungssituationen im Singunterricht werden in der Schweiz typischerweise *Vorsingen* genannt. Ausgehend von diesem *Vorsingen* entstehen Bewertungen, die meist relevant sind für die Schullaufbahn¹⁶². Damit stehen die Lehrkräfte in der Pflicht, ihre Entscheidungen

¹⁶² Die Relevanz der Zensuren ist kantonal unterschiedlich, sie ist ebenfalls in Bezug auf die anderen Domänen nicht zwingend gleichwertig, aber doch immer relevant für die sogenannte erreichte Promotion, wie die Versetzung in die nächste Schulstufe in der Schweiz genannt wird.

zur Beurteilung schriftlich zu belegen und damit nachvollziehbar zu machen. Dieser Anspruch des Schulsystems wird inhaltlich durch den Lehrplan 21 (D-EDK, 2015) als Maßstab der zu erreichenden Kompetenzziele gesteuert. Der Bezug zu AOL wird einerseits in der Forderung nach transparenten Beurteilungsschritten und andererseits mit dem Anspruch an die zu erreichenden Kompetenzen des Curriculums deutlich. Die Auswertung der Interviews soll zeigen, wie die Lehrkräfte mit den vom Bildungssystem an sie herangetragenen Forderungen umgehen und wie sich dies auf das vorgelegte Assessment-Modell nach Earl (2013) und die von Brophy (2019c) als *state of the art* bezeichneten Prinzipien von Assessment im Musikunterricht beziehen lassen.

Der summative Beurteilungsprozess unterliegt beim Singen allerdings nicht nur formalen und inhaltlichen Kriterien, sondern wird zusätzlich durch die adoleszente Stimmentwicklung beeinflusst, wobei nicht nur klangliche Einschränkungen, mögliche Blockaden durch eine verminderte *vocal identity* und die Gruppendynamik zwischen Mädchen und Jungen eine Rolle spielen (vgl. 13.1), sondern auch grundsätzliche Persönlichkeitsmerkmale wie Intro- und Extrovertierung, die je nach Setting der Leistungsbeurteilung zum Tragen kommen. Die verschiedenen Aspekte summativer Leistungsbeurteilung werden entlang der vier unter *ASSESSMENT FOR LEARNING* kodierten Subkategorien analysiert (vgl. Tabelle 23), wobei zusätzlich die nachgelieferten Beurteilungsunterlagen einbezogen werden.

Im Kategoriensystem der QIA wurden deduktiv zwei Subkategorien in Bezug auf AOL erstellt. Die Subkategorie *PLANUNG DER LERNSCHRITTE* bezieht sich auf den Vorbereitungsprozess der Leistungsbeurteilung. Kodiert wurden Aussagen, welche die Vorbereitung von AOL in Bezug auf Lernziele und Probesequenzen erwähnen. Bei einer Minderheit von 4 Lehrkräften wurden insgesamt 6 Textstellen kodiert. Die zweite Subkategorie wurde *TRANSPARENZ DER INTERPRETATION* genannt, um Textstellen zu sammeln, die beschreiben, wie die gesanglichen Leistungen bewertet werden, insbesondere unter dem Gesichtspunkt einer nachvollziehbaren Interpretation der Beurteilung. Dazu liegt aus der QIA mit 25 Kodierungen, verteilt auf 10 Lehrkräfte, eine ganze Reihe von Kodierungen vor. Weiter wurden im Verlauf der Entwicklung der QIA bei den Probanden zusätzlich individuelle Bewertungsraster eingefordert, worauf fünf Beispiele (01, 02, 06, 07, 10) zugesandt wurden.

Im Verlauf der Erstellung des Kategoriensystems zeigte sich der Bedarf, die Settings der Leistungsbeurteilungen und verschiedene Formen von Performance zu kodieren, weshalb induktiv zwei weitere Subkategorien ergänzt wurden. In der Subkategorie *FORMEN DER BEWERTUNG* wurden Aussagen zu Durchführungsszenarien kodiert, insbesondere um Umstände und verwendete Sozialformen zu eruieren. Auch in dieser Subkategorie kann von einem reichhaltigen Material ausgegangen werden, insgesamt sind es 45 Kodierungen verteilt auf alle interviewten Lehrkräfte. Die vierte Subkategorie erhielt den Namen *PERFORMANCE* und konnte an die Anforderungen des LP21 angeschlossen werden. Dieser setzt neu das Ziel¹⁶³, dass Schülerinnen und Schüler „ihre [...] stimmlichen Fähigkeiten vor Publikum oder auf der Bühne präsentieren“¹⁶⁴ (D-EDK, 2015) können. Kodiert wurden Textstellen, in denen Lehrkräfte von performativen Elementen berichteten. Auch diese Subkategorie wurde bei allen Interviews kodiert, insgesamt 39-mal. Die anschließenden Analysen folgen nun den vier beschriebenen Subkategorien.

Planung der Lernschritte

Die sechs verwertbaren Kodierungen der Subkategorie *PLANUNG DER LERN-SCHRITTE* sind auf vier Lehrkräfte verteilt, wobei drei davon auf Herrn Buchser (02) fallen. Er beschreibt mehrfach, wie die Planung von Lernschritten während einer summativen Beurteilung erfolgt, indem er konkretes Üben und eine Wiederholung der Leistungsbeurteilung vorschlägt (02:235 ff.; 259 ff.). Dieses Vorgehen wurde bereits als formative Benutzung summativer Bewertungen (Fautley, 2010) diskutiert, allerdings als zyklischer Vorgang über die drei Jahre der Sekundarstufe 1 hinweg. Das Angebot einer kurzfristigen Wiederholung lässt den summativen Messpunkt zu einer formativen Lernbegleitung werden und weist auf die Schwierigkeit einer konsequenten Trennung beider Bereiche hin. Direkt von einem planmäßigen Vorgehen kann allerdings nicht gesprochen werden.

Bezüglich der Planung einer Leistungsbeurteilung von Gesang stellt sich die Frage, in welcher Form Schülerinnen und Schüler auf die Vorsing-

¹⁶³ Diese Zielvorstellung ist neu im Lehrplan (siehe Huber, 2008).

¹⁶⁴ Siehe Vorlage LP21: <https://v-ef.lehrplan.ch/101XDKDtWn86CsyMNy-u647dmaxP5V9nnh> (abgerufen am 5.11.2021).

Momente am Tag X vorbereitet werden. Dies lässt sich aus einer Kodierung und implizit aus weiteren Aussagen ableiten. Herr Berner (09) beschreibt einen Ordner, in dem spezifische Lieder zum Vorsingen enthalten sind (09:195 ff.). Diese Lieder werden stetig mit der Klasse geübt und können am Tag X ohne spezifische Vorbereitung vorgetragen werden. Dieses Vorgehen scheint auch bei weiteren Lehrkräften Standard zu sein (z. B. 01:172 ff., 11:330 ff.). In diesem Fall gibt es keine Planung aufseiten der Lehrkräfte, sondern lediglich die Auswahl eines Liedes aus dem Repertoire durch die Schülerinnen und Schüler. Dabei vertrauen die Lehrkräfte auf das Einschätzungsvermögen der Lernenden (01:172 ff.; 03:214 ff.). Eine Zielorientierung bezogen auf konkrete Lernziele oder ein definierter Kompetenzerwerb kann aus diesem Setting nicht abgeleitet werden, auch wenn Herr Hilfiker (04:306 ff.) vorgängig Lernziele notiert und Herr Buchser (02:235 f.) Ziele bespricht. Im Wesentlichen werden in beiden Fällen lediglich die Beurteilungskriterien bekanntgegeben. An dieser Stelle kann eingewendet werden, dass durch stetiges Feedback beim Singen eine spezifische Planung obsolet ist, wie dies Mantie (2019) einbringt, wenn er den Musikunterricht per se als formativ erklärt. Was dabei unberücksichtigt bleibt, ist der individuelle Lernstand, der in einer solchen Konstellation nur sporadisch sichtbar wird¹⁶⁵.

Im Rahmen von Gruppenarbeiten geben einzelne Lehrkräfte zudem den Auftrag, auch Lieder eigener Wahl einzubringen, zu üben und vorzutragen (03, 08, 11). Wie in diesem Fall Planungsaspekte zu gewichten sind, zeigen Aussagen zum methodischen Vorgehen in Vorbereitungssequenzen. Frau Derendinger (11) gibt z. B. den Schülerinnen und Schülern die Möglichkeit, sich selbstständig auf das Vorsingen in kleinen Gruppen vorzubereiten, wobei sie als Ansprechperson Fragen beantwortet und Feedback gibt, allerdings nur, wenn sie gefragt wird. Sie meint: „dann kommen sie auch direkt zu mir und wollen ja wirklich, [...] wollen ja für sich als Person, als Gruppe ein Improvement erreichen.“ (11:400 ff.) In ihrer Arbeitsweise ist kein Bezug zu konkreten Lernschritten zu sehen, Feedbacks erfolgen nur,

¹⁶⁵ Diese Interpretation wird dem erwähnten Beispiel von Herrn Berner (09) nicht gerecht. Seine Schülerinnen und Schüler müssen als einzige mehrfach pro Jahr einen Song vorsingen, wobei hauptsächlich Wert auf Auftrittskompetenz gelegt wird. Summative Beurteilung, gleichzeitiges Feedback und die Erwartung stetiger Auftritte sind nicht zu vergleichen mit einem jährlichen Auftritt.

wenn aufgrund der intrinsischen Motivation der Schülerinnen und Schüler ein formatives Setting entsteht. Auch bei Frau Fricker (03) ist das Abholen von Feedback unverbindlich, Vorbereitung als Baustein zu guter Leistung scheint ihr aber wichtig zu sein (03:185 ff.). In beiden Aussagen bezieht sich somit Planung auf die Bekanntgabe eines Beurteilungstermins und dem Zur-Verfügung-Stellen von Übungszeit.

Eine Ausnahme bildet das Beispiel von Herrn Leimbacher (10), in dem, ausgehend von einem vorgelegten Beurteilungsraster (siehe Anhang) und Aussagen aus dem Interview (10:89 ff.) auf eine konkrete Planung rückgeschlossen werden kann. Er lässt im 6-stimmigen Stück *Evening Rise* alle Stimmen üben, um anschließend in der Beurteilung den Jugendlichen maximal je fünf Punkte für eine korrekt gesungene Melodie zu geben. Im Gegensatz zum Vorsingen eines Repertoire-Stücks wird hier auf ein konkretes Ziel hingearbeitet. Einerseits wird erwartet, dass die Stimmen nach Noten gesungen werden müssen, und andererseits wird neben sauberer Intonation und genauer Rhythmik die Benutzung des gesamten Ambitus (von G–d1/g–d2) angezielt. Auch wenn Herr Leimbacher (10) nicht konkret von Planung spricht, werden in diesem Beispiel klare Lernziele vorgegeben und überprüft. In diesem Beispiel werden zwei Forderungen an Lehrkräfte aus der Beschreibung des Assessment-Modells von Earl (2013, vgl. Tabelle 5) erfüllt: Einerseits bietet Herr Leimbacher (10) eindeutige und definierte Ziele des Lernens und andererseits eine Aufgabenstellung, die es den Lernenden ermöglicht, ihre Fähigkeiten anhand eines konkreten Beispiels zu entwickeln und nachzuweisen. Zudem folgt das Beispiel dem *principle of purpose* (Brophy 2019, S. 921), das eine Deklaration dessen einfordert, was gelernt werden soll und wie dies den musikalischen Lernprozess fördert.

Der Anspruch einer absichtsvollen Planung von Lernschritten und der anschließenden Beurteilung wird in dieser Deutlichkeit nur in einem Beispiel manifest. Im Gegensatz dazu steht der bereits diskutierte zyklische Ansatz. Die kleine Zahl von Kodierungen deutet somit weniger auf Planlosigkeit als auf einen übergeordneten Fokus auf allgemeine Ziele hin, die bereits ausführlich diskutiert wurden (vgl. 12.2). Dabei hat sich gezeigt, dass in der Zielvorstellung eine Polarität auszumachen ist zwischen musikalischer Förderung von Gesang oder einer Betonung von Persönlichkeitsentwicklung durch das Singen. Im Sinne der Orientierung an allgemeinen Zielen und der bereits beschriebenen zyklischen Beurteilung haben kon-

krete Ziele für eine summative Beurteilung weniger Bedeutung. Dies gilt in besonderem Maß, wenn die Auftrittskompetenz durch stetiges Auftreten Förderung erfährt (vgl. Fußnote 167).

Transparente Interpretation

Die von Earl (2013) formulierten Kriterien für die summative Leistungsbeurteilung (AOL) beinhalten unter anderem kommunizierte und belastbare Beurteilungskriterien sowie Klarheit darüber, wie diese interpretiert werden. In der vorliegenden Untersuchung ermöglichen die 25 Kodierungen zu *TRANSPARENTE INTERPRETATION* sowie die eingereichten Beurteilungsraster¹⁶⁶ eine Einschätzung zu beiden Perspektiven. Zuerst wird auf der inhaltlichen Ebene die Frage fokussiert, wie transparent die Beurteilungen sind und ob die Beurteilungskriterien eine valide Leistungsbeurteilung ermöglichen. Die geschilderte Anwendung dieser Raster und die Überzeugungen zur Notengebung dienen anschließend dazu, die Interpretation der Beurteilungskriterien einzuordnen und den summativen Beurteilungsprozess insgesamt auf das *principle of quality* (Brophy, 2019c, S. 920) zu beziehen.

Fast alle interviewten Lehrkräfte stützen sich bei der summativen Leistungsbeurteilung auf selbst entwickelte Beurteilungsraster, außer Frau Basler (05), die im Bereich Singen konsequent Stellung gegen eine Bewertung bezieht (05:431 ff.). Als Beurteilungskriterien werden durchgehend Elemente wie Melodie, Intonation, Rhythmus, Volumen, Textverständlichkeit und Einsatz genannt, allerdings in unterschiedlichen Kombinationen und uneinheitlichen Begrifflichkeiten. Einzelne Lehrkräfte eruieren zudem spezifische Klangqualitäten (06:Raster¹⁶⁷) oder formulieren ergänzende Kommentare (01, 11) als Stütze für das integrierte formative Feedback, das sich im Rahmen der zyklischen Leistungsbeurteilung etabliert hat. Bei zwei Lehrkräften (01, 10) sind die Beurteilungskriterien mit einer Ordinalskala versehen, wobei unklar bleibt, wie eine spezifische Punktzahl erreicht werden kann. In den anderen vorhandenen Beispielen sind zusätzlich zu

¹⁶⁶ Die Beurteilungsraster sind in der MAXQDA-Datei gespeichert (siehe Anhang).

¹⁶⁷ Raster bezieht sich jeweils auf das eingereichte Beurteilungsraster der entsprechenden Interviewnummer und ist im Anhang verfügbar.

den Beurteilungskriterien Ausprägungen formuliert, welche die Qualitätserwartungen beschreiben. Frau von Rohr (06) weist z. B. in ihrem Beurteilungsraster dem Kriterium *Melodie* die Werte „fehlerfrei“, „wenige falsche Töne“ sowie „mehrere falsche Töne“ (06:Raster) zu. Herr Rothenburg (07) beschreibt für das Kriterium *Volumen* fünf Stufen: „Flüstern – leise – mittel – Genug laut – mit Energie“ (07:Raster). In beiden Formen von Rastern führt die Vergabe von Punkten zu einer Note.

Der Anspruch an Transparenz wird mit diesen Rastern nur bedingt eingelöst. Zwar hat sich bei einer Mehrheit von Lehrkräften der Einsatz von Beurteilungsrastern etabliert und die Angabe von Beurteilungskriterien scheint zudem Schülerinnen und Schülern zu helfen, die Bewertung zu verstehen (09:235 ff.), die Vergabe von Punkten bietet aber großen Spielraum für Interpretation, insbesondere dort, wo keine Qualitätskriterien hinterlegt sind. Die beispielhaft genannten Ausprägungen der Kriterien müssten präzisiert werden, um auch von Außenstehenden angewendet werden zu können. Die individuell entwickelten Beurteilungsraster sind somit auf einen hohen Anteil an informellen Abmachungen angewiesen. Im Rahmen von mündlichen Feedbacks, die alle untersuchten Lehrkräfte als Abschluss der Beurteilung praktizieren, werden diese Abmachungen wirksam. Zwei Lehrkräfte (03, 04) verzichten auf das schriftliche Feedback, das Raster dient dann lediglich als Grundlage für das mündliche Feedback. Es darf also davon ausgegangen werden, dass auf einer informellen Ebene die summative Beurteilung als transparent erlebt wird. Inwiefern die Beurteilungskriterien eine valide Leistungsbeurteilung ermöglichen, kann aufgrund des hohen Anteils an informeller Information nicht beantwortet werden.

Im Vergleich zu bestehenden Skalen (z. B. Wise & Sloboda, 2008; Rutkowski, 2019) und dem vorgestellten *Rubric* von Holcomb (2019) gibt es inhaltliche Überschneidungen, wobei an dieser Stelle nur auf zwei Bereiche eingegangen wird. In zwei Rastern wird zwischen melodischer Genauigkeit und Intonation unterschieden, Frau von Rohr (06) nennt es *Melodie* und *Intonation*, Herr Rothenburg (07) *Melodie* und *Reinheit*. In den drei anderen Rastern wird nur ein Item aufgeführt, *Tonhöhe* (01), *Melodie* (02) und *Intonation* (Tonreinheit, 10). Der von Rutkowski vorgeschlagene Einbezug einer Einschätzung des Gebrauchs von Registern ist in keinem Raster zu finden, ebenso wenig wie Items zur adoleszenten Stimmentwicklung, die möglicherweise einen Einfluss auf die Tessitura haben. Die Unterscheidung von melodischer Genauigkeit und Intonation (06, 07) deckt

sich mit den oben genannten Skalen¹⁶⁸. Es zeigen sich allerdings auch Lücken. So wird das von Holcomb (2019) prominent gesetzte Kriterium *Atem* nur von Frau von Rohr (06) am Rand aufgeführt und die Kategorie der Sprechsänger findet sich nur im Raster von Herrn Rothenburg (07) beim Kriterium *Melodie* als „immer gleicher Ton“ (07:Raster). Zudem weisen in einigen Rastern Kriterien zu Äußerlichkeiten wie z. B. „Blick in die Noten“ (01:Raster) oder „Sitzen“ als Gegenpol von „Gute Körperspannung“ (07:Raster) auf eine gewisse Inkonsistenz der untersuchten Raster hin. Der Anspruch an eine valide Beurteilung muss so infrage gestellt werden, auch aufgrund der Tatsache, dass niemand auf etablierte und validierte Skalen und *Rubrics* zurückgreift. Es bestätigt sich somit auch die Beobachtung von Schaumberger (2020), der die Quellenlage auf individuell erarbeitete Beurteilungsraster reduziert sieht (S. 232).

Abgesehen von Frau Basler (05) benutzen die Lehrkräfte ihre Raster, ohne diese infrage zu stellen. Bei der Vergabe der Noten zeigt sich allerdings ein gewisses Unbehagen, das sich in der Vergabe von meist genügenden Noten ausdrückt. Frau Basler bringt das Unbehagen bezüglich der Notengebung als einzige Teilnehmende explizit auf den Punkt:

„[...] weil ich finde eine Note auf Stimme zu geben ist indiskutabel. Ich sag's so. [...] ich kann nicht sagen, du bist ein sehr guter Sänger. Das kann ich, gut, das kann ich noch einfügen [...] ich kann nicht sagen, du bist ein schlechter Sänger. [...] Das darf ich eigentlich auch nicht sagen. [...] das find ich, das geht nicht. Und die Stimme und der Mensch selber ist noch in der Entwicklung. Die Stimme ist ja dieses ganz Persönliche. [...] Wenn ich auf die Stimme eigentlich eine Note gebe, dann heißt das, ich bewerte den gesamten Menschen, und so kommt es auch an. Das mach ich nicht.“ (05:431 ff.)

¹⁶⁸ Ein Spezialfall stellt Herr Hilfiker (04) dar, der behauptet, die Bewertung der Intonation sei nicht zulässig: „ob sie die Töne treffen oder nicht, das ist nicht wirklich ein Bewertungskriterium bei mir. [...] Das ist einfach, keine Ahnung wie man das Lernen kann noch in diesem Alter, oder irgendwie es ist auch für uns schwierig, oder?“ (04:283 ff.). Mit dieser Argumentation steht Herr Hilfiker in Opposition zu allen anderen, welche die Tonhöhe selbstverständlich als Kriterium einbeziehen. Es wird allerdings verständlich, wenn ein biografischer Bezug hergestellt wird. Als Schlagzeuger und ohne schulmusikalische Ausbildung hat er kaum Bezug zur Stimme und sieht sich auch nicht als Vorbild.

Frau Basler (05) verbindet eine schlechte Beurteilung von Gesang mit einer schlechten Bewertung des singenden Menschen und entzieht mit dieser Argumentation der summativen Bewertung von Gesang jegliche Legitimation. Sie verweist zudem auf die adoleszente Stimmentwicklung als Grund, auf eine Bewertung in der Entwicklungsphase zu verzichten. Die Verknüpfung von Stimme und Person wurde an mehreren Stellen in dieser Arbeit begründet. Beachtet man zudem die bereits zitierte Vielzahl an Texten zum Singen, die einleitend über negative Jugenderlebnisse berichten (siehe z. B. Tellisch, 2015; Schaumberger, 2020a), macht es sicher Sinn, die traditionelle Weise des Vorsingens als Leistungsbeurteilung infrage zu stellen (vgl. 14.3). Als Lehrerin einer Waldorfschule hat Frau Basler (05) die Möglichkeit, ihre Überzeugung umzusetzen, im Gegensatz zu den Lehrkräften staatlicher Schulen, die verpflichtet sind, summative Beurteilungen durchzuführen.

Als Konsens zeigt sich bei den anderen Lehrkräften die gemeinsame Haltung, mit Ausnahmen nur genügende Noten zu geben. Frau Derendinger (11) beschreibt den Konsens so:

„In dem Moment, wo du hin stehst, in deiner Gruppe, und dein Möglichstes machst, und den Text, deinen Teil, ich sage jetzt eine Strophe ein Refrain, auswendig kannst, und dich bemühst, egal was dabei rauskommt, hast du einen Vierer.“ (11:438 ff.)

Die Note 4 entspricht der minimalen Leistung für eine genügende Bewertung¹⁶⁹. Frau Derendinger (11) weicht im Notfall, wenn der Gesang nicht bewertbar ist, auf außermusikalische Kriterien aus, indem sie motivationale und kognitive Kriterien hinzuzieht. Mit diesem Ausweichen folgt sie im Wesentlichen der Logik von Frau Basler (05) und stellt sicher, die Person nie als ungenügend zu bezeichnen. Frau von Rohr (06) geht sogar noch einen Schritt weiter und macht nur sehr gute Noten: „Die haben, ich mache sehr gute Vorsingnoten, weil ich sie motivieren möchte. [...] Ich will, nicht,

¹⁶⁹ Die Note 6 ist Ausdruck einer sehr guten Leistung, die Note 4 ist gerade noch genügend. Minimal wird im schweizerischen Beurteilungssystem die Note 1 vergeben. Wer am Ende des Schuljahres in allen Fächern einen Gesamtdurchschnitt unter 4 erreicht, muss je nach kantonaler Vorgabe die Klasse wiederholen oder in das nächsttiefere Leistungsniveau wechseln.

dass jemand aus der Bez.¹⁷⁰ rausgeht und sagt, ahh ich kann voll nicht singen. Ich bin immer huere schlächt gewesen.“ (06:348 ff.) Frau von Rohr (06) befürchtet, dass die Jugendlichen durch die Vergabe von Singnoten ein negatives vokales Selbstkonzept entwickeln und setzt ausschließlich sehr gute Noten. Sie betont zusätzlich, dass Jungen während der Mutation und Jugendliche mit Erkältungen einen Nachteilsausgleich erhalten (06:345). Frau Derendinger (11) und Frau von Rohr (06) verlassen die Logik der Beurteilungsraster mit den beschriebenen Beurteilungskriterien. Es stellt sich somit die Frage, ob die summative Beurteilung wirklich vokale Kompetenzen abzubilden vermag.

Herr Neuendorf (08) setzt auf einen Aushandlungsprozess. Er ist der einzige Lehrer im Sample, der eine Selbstbewertung verlangt. Er geht folgendermaßen vor:

„Und dann tue ich meine Bewertung, die ich mir vorstelle, [...] versuchen zu begründen, und nachher schauen wir, dass man den Mittelweg findet. Also wenn sie finden, sie hätten eine 5 verdient, und ich sage, ja ich habe das Gefühl eine 4.5 ist angemessen, dann gibt es halt eine 4.75, oder so.“ (08:371 ff.)

Der von Herrn Neuendorf (08) beschriebene Vorgang findet am Ende der Schulzeit im 9. Schuljahr statt. Es ist der Abschluss einer empathischen Reise (vgl. 12.1) und bezieht den geleisteten Aufbau von reflexiven Fähigkeiten, der später bei der Beurteilung von AAL nochmals thematisiert wird, mit ein. Obwohl an dieser Stelle diskutiert werden müsste, ob die Selbstbeurteilung eigenständig erfolgt oder das Urteil der Lehrkraft zu antizipieren versucht, eröffnet sich die Möglichkeit, Schülerinnen und Schüler nicht nur als Empfänger summativer Beurteilung zu betrachten, sondern in ihrer Urteilkraft ernst zu nehmen und insbesondere die Interpretation der Leistungsbeurteilung im Gespräch abzugleichen.

Insgesamt hat sich im untersuchten Sample eine transparente Beurteilungskultur etabliert, wobei nicht immer klar ist, welche Leistungsanforderungen mit den einzelnen Kriterien verbunden sind. Die Kriterien decken sich zum Teil mit den Kriterien etablierter Skalen und Raster, ob dies aber zu einer validen Beurteilung führt, kann nicht eruiert werden. Obwohl

¹⁷⁰ Bez = Bezirksschule (vgl. 6.3).

einige Lehrkräfte die summative Beurteilung mit beträchtlichem Aufwand betreiben, sogar ein Beispiel eines Aushandlungsprozesses beschrieben werden konnte, wird doch auch ein großes Unbehagen bezüglich der Bewertung von Singen sichtbar. Dies drückt sich dadurch aus, dass kaum ungenügende Noten oder gar nur Höchstnoten verteilt werden.

Formen der Bewertung und Performance

Mögliche Durchführungsszenarien der Leistungsmessung werden in der Literatur wenig thematisiert. Schaumberger (2020a) verweist auf den Chorauftritt als authentische Weise der Leistungsmessung und erinnert an die oft berichteten bloßstellenden Formen von Vorsingen (S. 223). Zudem zitiert er aus dem amerikanischen Kontext Einzel- und Gruppenprüfungen (S. 229). Holcomb (2019) skizziert individuelle Projekte, die zu einer abschließenden Bewertung von Audioaufnahmen führen (vgl. 5.2). Aus dem LP21 (D-EDK, 2015) lassen sich Szenarien höchstens aus den zu erreichenden Kompetenzen ablesen. So wird das gemeinsame chorische Singen als zu erreichendes Ziel für alle definiert, eine Solo-Performance wird nur als weiterführende Kompetenz in Betracht gezogen: „Die Schülerinnen und Schüler können ein begleitetes Lied solistisch oder in Gruppen vortragen.“ (S. 7)¹⁷¹ Gleichzeitig verlangt der LP21 aber auch das Training von Auftrittskompetenz, wobei thematisch auf gegenseitige Wertschätzung, Peerfeedback sowie Bühnenpräsentation fokussiert wird. Aus dem LP21 lassen sich somit Leistungsmessungen sowohl in Chor- oder Gruppenperformances als auch in solistischen Auftritten argumentieren. Zur logistischen Bewältigung (*principle of operability*; vgl. Abbildung 6) gibt der LP21 allerdings keine Auskunft.

Hinweise zu den Durchführungsszenarien von Bewertungen sowie zum Training von Auftrittskompetenz wurden in den Subkategorien *FORMEN DER BEWERTUNG* und *PERFORMANCE* kodiert. Insgesamt wurde *FORMEN DER BEWERTUNG* 45-mal kodiert, gleichmäßig auf die Interviews verteilt, minimal 2-mal und maximal 7-mal. Aufgrund der Kodierungen lassen sich einerseits verschiedene Settings beschreiben und andererseits gibt es bei

¹⁷¹ Siehe Vorlage LP21: <https://v-fe.lehrplan.ch/101JLDxVkUCUYdpsfsFau-HEdZCnTveKSr> (abgerufen am 20.01.2022).

einigen Lehrkräften Ansätze von aufbauenden Konzepten im Verlauf der drei Jahre der Sekundarstufe 1. Für die Subkategorie *PERFORMANCE* finden sich 39 Kodierungen, verteilt auf 9 Lehrkräfte, allerdings weniger gleichmäßig verteilt als bei *FORMEN DER BEWERTUNG*. Es zeigen sich hauptsächlich Unterschiede im Umgang mit der Förderung von Auftrittskompetenz. Einerseits wird auf die Möglichkeit des Schulchors als Auftrittsmöglichkeit verwiesen, andererseits explizit auf die individuelle Stärkung von Auftrittskompetenz im regulären Unterricht. Beide Subkategorien stehen in einer gewissen Abhängigkeit, weshalb sie in der Analyse gemeinsam betrachtet werden. So verändert sich z. B. der Grad an Privatheit der Beurteilungssituation mit dem Maß an verlangter Auftrittskompetenz. Der Start erfolgt mit der von Schaumberger (2020a) eingebrachten Feststellung, die natürlichste Art, Gesang von Gruppen zu bewerten, sei der Chorauftritt.

Keine einzige Lehrkraft zieht in Betracht, die Leistungsbewertung von vokalen Darbietungen in Form eines Chorauftritts durchzuführen. Zwar weisen sieben Lehrkräfte auf einen Schulchor und entsprechende performative Möglichkeiten hin, ziehen dies aber nicht in ihre Beurteilung mit ein. Die Vermutung liegt nahe, dies liege vor allem an der Unmöglichkeit einer individuellen Leistungsbeurteilung bei einer kollektiven Leistungserbringung. Drei Aspekte können als Begründung dienen, Chorauftritte als Leistungsmessung nicht a priori auszuschließen. Als erstes wird auf die Lehrkräfte verwiesen, die insgesamt nur sehr gute Noten mit einer sehr kleinen Streuung machen. Diese Lehrkräfte kämen auf eine übereinstimmende Bewertung, wenn sie ihre Klasse als Chor singen lassen und sich merken, wer sich mehr oder weniger stark engagiert. Je nach Engagement könnte eine gute oder sehr gute Note erteilt werden. Der Vorteil zum Vorsingen in Gruppen wäre neben einem äquivalenten Resultat eine authentische Prüfungssituation, da das Üben und die Bewertung kongruent sind (*principle of authenticity*, vgl. Abbildung 6). Nach wie vor problematisch wäre allerdings die Validität und Individualisierung der Notengebung, die üblicherweise bei einer summativen Leistungsmessung erwartet wird. Ein zweiter Aspekt ist die aus dem LP21 abgeleitete Anforderung, dass Schülerinnen und Schüler im Minimum grundlegende Kompetenzen im mehrstimmigen

Gesang erreichen sollen¹⁷². Obwohl es nicht die Aufgabe eines Lehrplans ist, Bewertungsmodi zu deklarieren, wird doch ganz selbstverständlich eine Kompetenzstufe deklariert, die nur schwerlich auf einfache Art und Weise geprüft werden kann. Ein Ansatz wäre diesbezüglich die Aufgabe von Herrn Leimbacher (10), der das Beherrschen mehrerer Stimmen eines Liedes verlangt. Bereits besprochen wurde die Einzelsituation dieser Leistungsbeurteilung, im Text erwähnt er allerdings auch die Möglichkeit eines zweistimmigen Vortrags (10:110 f.). Für die Umsetzung der Leistungsbeurteilung in einer Chorsituation müssten allerdings Wege entwickelt werden, entsprechende Kompetenzen authentisch, valide und auf individueller Basis messen zu können. Als dritter Aspekt wird noch einmal auf die genannten Schulchöre verwiesen. Wenn keine Lehrkraft eine Bewertung der Chorsituation in Betracht zieht, wäre auch zu überlegen, dies als Maßstab zu nehmen und auf eine individuelle Benotung gänzlich zu verzichten. Ziel eines ganzheitlichen Assessments ist nicht nur das summative Assessment, sondern auch die individuelle Stimmentwicklung, die zum Erreichen der Kompetenzziele beiträgt. Holcomb (2019) beschreibt einerseits *Rubrics* als adäquate Form der Leistungsbeurteilung im Chor, bringt aber auch individuell begleitete Aufgaben ins Spiel, die einzeln und in Gruppen geübt, aufgeführt und beurteilt werden. Die Lehrkräfte beschreiben nun ebenfalls verschiedene Wege, wie sie ihre Leistungsbeurteilungen organisieren.

Herr Buchser (02) hat eine ganze Palette von verschiedenen Settings entwickelt:

„Kann ich auch beim Vorsingen, wo es drei Kategorien gibt, nämlich vor mir, zu zweit oder zu dritt, oder vor der Klasse, aber zu mir hin singen und die Klasse ist im Hintergrund, und die dritte, ich sitze in die Klasse rein und sie singen zu der Klasse als Publikum, das ist am schwierigsten. Und dann machen wir verschiedene Kategorien, die meisten wählen immer das einfachste, also in Zweier-, Dreiergrüppchen vor mir vorsingen beim Computer und alle anderen sind draußen.“ (02:25 ff.)

¹⁷² Die Beschreibung der Kompetenz ist nicht eindeutig, gefragt ist ein „spezifischer Beitrag im mehrstimmigen chorischen Singen“ (<https://v-fe.lehrplan.ch/101PedBSEH5a5vtBnCMg7bz4t9TZTUyC5> – abgerufen am 05.02.2022).

Herr Buchser (02) lässt seine Schülerinnen und Schüler zwischen verschiedenen Herausforderungen wählen. Einerseits haben sie die Möglichkeit, ihren Gesang allein oder in einer kleinen Gruppe zu präsentieren, und zusätzlich können sie wählen, ob dies im Beisein der Klasse oder nur vor dem Lehrer stattfindet. Die Präsentation nur vor der Lehrperson findet in einem Schutzraum statt und ist nicht zwingend auf eine lernförderliche Atmosphäre angewiesen (vgl. die Diskussion zu Assessment as Learning, S. 335). Zudem ist der hohe Grad an Privatheit sicher von Vorteil für introvertierte Lernende. Die ultimativ private Variante, nur eine Audioaufnahme abliefern zu müssen (vgl. Holcomb, 2019), wird von keiner Lehrkraft genannt, wobei auch das Hören der eigenen aufgezeichneten Stimme von vielen als unangenehm empfunden wird und deshalb als Angebot erfolglos sein würde. Herr Brugger (01) verlangt von allen Lernenden ein Einzel-Vorsingen. Im Unterschied zu den anderen Lehrkräften gibt es keine Wahlmöglichkeit, was den Grad an Autonomie deutlich einschränkt. Herr Neuendorf (08) führt lediglich das letzte Vorsingen einzeln durch, begleitet von einer Selbsteinschätzung und einem Auswertungsgespräch und in Kombination mit einem Performance-Projekt mit großem Gestaltungsspielraum. Im Vergleich zeigt sich, dass bei Herrn Brugger (01) lediglich musikalische Ziele im Vordergrund stehen, im Gegensatz zu Herrn Neuendorf (08), der dieses Einzel-Vorsingen als Abschluss einer persönlichkeitsfördernden Entwicklung vor der Schulentlassung gestaltet. Herr Buchser (02) weist seine Lernenden neben diesen Durchführungsszenarien im geschützten Raum auch auf die Möglichkeit eines Vorsingens vor der Klasse hin, wobei dies, wie aus dem Zitat hervorgeht, bei ihm wenig gewählt wird.

Er beschreibt für das Vorsingen im Rahmen des Klassensettings zwei verschiedene Möglichkeiten mit unterschiedlichem Maß an Privatheit. In der einfacheren Version bleibt die Klasse im Hintergrund, was einerseits das Ablenkungspotenzial erniedrigt, da der Blickkontakt zu den Peers fehlt, andererseits aber auch wenig Verbindlichkeit durch die Zuhörenden voraussetzt oder generiert. Auch Herr Leimbacher (10) führt Einzelprüfungen durch, während denen die Klassen im Hintergrund eine selbstständige Arbeit erledigen (10:584 ff.). Beide Lehrkräfte bereiten so die Schülerinnen und Schüler für anspruchsvollere Auftrittssituation vor der Klasse vor. Herr Buchser (02) beschreibt denn auch das Vorsingen vor der Klasse als schwierigere Variante. Diese Form wird von vier Lehrkräften beschrieben,

wobei die Wahlfreiheit im Gegensatz zum Angebot von Herrn Buchser (02) teilweise eingeschränkt ist. Frau Derendinger (11) verlangt von allen Schülerinnen und Schülern ein Vorsingen in Gruppen vor der Klasse und setzt dabei sogar Mikrofone ein. Dabei sind die Lernenden zusätzlich zu einer erzwungenen Extrovertierung herausgefordert, ihre eigene Stimme aus dem Lautsprechen hören zu müssen. Als Zugeständnis erlaubt sie dann im 8. und 9. Schuljahr den Vortrag eines Raps, was aber oft lediglich deshalb gewählt wird, weil die Motivation zum Singen fehlt (11:314 ff.). Auch Herr Berner (09) verlangt das Vorsingen vor der Klasse, allerdings in monoedukativen Sequenzen. Durch Regelmäßigkeit des Vorsingens im Verlauf eines Schuljahrs erwartet er eine Zunahme der Auftrittskompetenz. Schülerinnen und Schüler benutzen zudem sein Angebot, allein vor der Klasse zu singen (09:471 ff.). Als dritte Lehrkraft verlangt Frau Fricker (03) das Vorsingen vor der Klasse in Gruppen, allerdings nur im letzten Schuljahr bei der Performance eines selbst komponierten Songs. Gleich wie bei Herrn Buchser (02) ist zumindest implizit der Aufbau von Auftrittskompetenz angelegt, weil im Verlauf der drei Schuljahre in der Sekundarstufe 1 der Auftritt vor der Klasse angestrebt wird. Im autoethnografischen Exkurs werden die zunehmenden Ansprüche an Auftrittskompetenz in ähnlicher Form beschrieben (vgl. 3.2). Die Beispiele mit einem kleinen Maß an Privatheit finden sich meist als Möglichkeit der Förderung von Auftrittskompetenz intendiert. Es fehlt allerdings auch in diesen Durchführungsszenarien teilweise die Wahlfreiheit.

Insgesamt zeigen sich verschiedene Tendenzen bei den Modi der Leistungsbeurteilung. Einige Lehrkräfte benutzen im Verlauf der drei Schuljahre in der Sekundarstufe 1 ein gleichbleibendes Setting, sei dies ein Einzelsingen vor der Lehrkraft oder das Vorsingen in Gruppen vor der Klasse. Dabei fällt auf, dass Schülerinnen und Schüler keine Wahlmöglichkeit haben und so ihre persönlichen Bedürfnisse nicht beachtet werden. Sowohl die Einzelsituation vor der Lehrkraft als auch der Auftritt mit Mikrofon kann einschüchternd oder von negativen Gefühlen begleitet sein. Andere Lehrkräfte ermöglichen unterschiedliche Settings und geben damit eine Wahlfreiheit, bauen sogar verschiedene Settings aufeinander auf, mit dem Ziel in verschiedenen Schwierigkeitsgraden vor Publikum aufzutreten und so Auftrittskompetenz zu entwickeln. Dies findet sowohl mit unterschiedlichen Graden an Beteiligung des Publikums als auch in monoedukativen Settings statt. Die Beurteilung von Chorsituationen innerhalb

der Klasse oder im Rahmen von Aufführungen in der Schule findet sich im Datenkorpus nicht. Entsprechende Überlegungen dazu könnten aber durchaus spannend sein, insbesondere unter Berücksichtigung von *Rubrics*.

Der summative Bereich AOL

Die Analyse zum Bereich *assessment of learning* zeigt eine reichhaltige Praxis der Leistungsbeurteilung. Eine konkrete Orientierung an Lernzielen kann nur implizit in einem Einzelfall gezeigt werden. Vielmehr steht bei der Mehrheit der Lehrkräfte eine zyklische Bewertung, orientiert an übergeordneten Zielen, im Vordergrund. So müssen Schülerinnen und Schüler beim jährlichen Vorsingen meist aus einem vorgegebenen Repertoire ein Lied auswählen und vortragen. Teilweise wird von einer höheren Kadenz der Beurteilungssituationen oder einer solistischen Performance gegen Ende der Schulzeit gesprochen. Auch der Anspruch an valide und transparente Beurteilungen wird nur teilweise eingelöst. Zwar sind vokale Beurteilungskriterien in den Rastern mit denen etablierter Raster und Skalen vergleichbar, aufgrund des verwendeten Punktesystems und der nur rudimentär beschriebenen Leistungsanforderungsniveaus kann aber nicht von einer validen Beurteilung gesprochen werden. Dies zeigt sich bei einigen Lehrkräften in einem Unbehagen zur Leistungsmessung und der Vergabe sehr guter Noten. Die Leistungsbeurteilung wird meist von einem Feedbackgespräch begleitet, das einerseits Transparenz entlang der Kriterien schafft, andererseits aber auch als formatives Assessment verstanden wird. Die Durchführung der Leistungsbeurteilung findet in unterschiedlichen Settings statt, die sich durch ein mehr oder weniger hohes Maß an Privatheit unterscheiden. Auf der einen Seite findet das Vorsingen nur vor der Lehrkraft statt, wobei dort kaum Wert auf die Entwicklung von Auftrittskompetenz gelegt wird. Auf der anderen Seite sind in Ansätzen aufbauende Settings sichtbar, in denen zunehmend Auftrittskompetenz vor der Klasse demonstriert werden kann.

Die ursprüngliche Konzeption formativer und summativer Evaluation (Scriven, 1967; Bloom et al., 1971) beruht auf der Idee, durch formative Prozesse die individuelle Leistungsfähigkeit zu verbessern. Die übergeordneten Ziele zum Singen und die zyklische Beurteilung lassen nun vermuten, AFL werde im Singunterricht wohl eher wenig gepflegt, insbesondere im Bereich der individuellen Förderung. Anschließend werden die

gefundenen Ansätze individueller Betreuung und Förderung beschrieben und mit dem Forschungsstand abgeglichen.

Assessment for Learning

Formative Praktiken zeigten sich bereits bei der Analyse der Porträts (vgl. 12.2). Insbesondere Lehrkräfte, welche die persönlichkeitsbildenden Momente des Singens akzentuieren, schaffen durch vielfältige Settings und die Betonung von Auftrittskompetenz ein Umfeld, in dem individuelles Feedback erfolgen kann. Zudem zeigte sich, dass Beziehungspflege und empathisches Verhalten bei einem Teil der Lehrkräfte bewusst gepflegt wird und sich daraus eine in Ansätzen individuelle Betreuung ergibt (vgl. 12.2). Formative Prozesse wurden auch als Appendix der summativen Beurteilung diskutiert, da zusätzlich zur Notengebung meist auch ein persönliches Feedback gegeben wird. Dieses Feedback hat zwar auch die Aufgabe, Transparenz in der summativen Beurteilung herzustellen, wird aber von den Lehrkräften als formatives Element im Rahmen ihrer Förderbemühungen verstanden, wobei nur zwei Lehrkräfte auf eine längerfristige Wirkung hinweisen. Bevor nun auf die konzeptionellen Elemente von AFL fokussiert wird, werden noch einmal deren Eckwerte in Erinnerung gerufen.

Der Bereich *Assessment for learning* schließt alle Aktivitäten von Lehrkräften ein, die während der Hinführung zur Leistungsmessung eine formative Komponente beinhalten: „Assessment for learning is designed to give teachers information that will allow them to modify the teaching and learning activities in which students are engaged to differentiate and understand how individual students approach their learning.“ (Earl, 2013, S. 13) Der Fokus von AFL liegt somit auf der Gewinnung von Informationen, die es den Lehrkräften ermöglicht, den Lernprozesse adaptiv zu gestalten. Entsprechende Praktiken für den Singunterricht finden in der musikpädagogischen Literatur noch wenig Beachtung (vgl. 5.3). Im deutschsprachigen Raum wird die explizite Betreuung individueller Bedürfnisse schon länger gefordert (Kalmbach, 2010) und aktuell mit Bezug auf Kompetenzentwicklung und Leistungsbeurteilung allgemein für das Fach Musik (Stern, 2020) und spezifisch für das Singen (Schaumberger, 2020a) diskutiert. Ansätze vokaler Diagnostik sind zwar vorhanden (Greuel & Horst, 2011), der Einbezug von Messinstrumenten als Basis formativer Prozesse

(Bojack-Weber, 2012) sowie anamnestische Stimmbeobachtungen (Fuchs, 2015) werden erwogen, die notwendigen diagnostischen Kompetenzen scheinen aber bei Lehrkräften noch wenig entwickelt zu sein (Brunner, 2014). Pezenburg und Dyllick (2018) beschreiben zudem erste Anzeichen einer Aufweichung monodidaktischer und frontaler Unterrichtsstrukturen beim Singen, was sich für AFL ebenfalls begünstigend auswirkt. Linn (2017) weist in seiner Untersuchung zur Wahrnehmung heterogener Unterrichtssituationen im Fach Musik allerdings darauf hin, dass gerade beim Singen aus Zeitgründen auf die Berücksichtigung individueller Bedürfnisse verzichtet wird. Zudem findet die Phase der adoleszenten Stimmentwicklung im deutschsprachigen Raum im Musikunterricht wenig Beachtung, wie der Vergleich zum anglophonen Raum (Gackle, 2019; Welch, 2019; Freer, 2020) und die Analyse zu den geschlechtsspezifischen Einflüssen nahelegt (vgl. 13.1). Im anglophonen Raum verweist Rutkowski (2019) auf die Möglichkeit, durch den Einbezug etablierter Skalen spezifische Lernstände direkt im Unterricht zu adressieren, und Holcomb (2019) zeigt, wie individualisierende Aufgaben im Bereich Singen mit formativen Prozessen unterstützt werden können. Das Datenmaterial wird nun bezüglich diagnostischer und differenzierender Unterrichtselemente untersucht, wobei auch Begründungen gegen eine Differenzierung beachtet werden.

Die Ausgestaltung der Subkategorie *ASSESSMENT FOR LEARNING* erfolgte gemischt deduktiv und induktiv. Für die zentralen Anliegen von AFL wurden zuerst im Kategoriensystem die Subkategorien *ZIELVORSTELLUNGEN*, *DIAGNOSE*, *DIFFERENZIERUNG*, *FÖRDERUNG* und *FEEDBACK* deduktiv erstellt. Es zeigte sich rasch, dass die Subkategorien *ZIELVORSTELLUNGEN* und *FÖRDERUNG* inhaltlich weiter ausdifferenziert werden müssen und die Inhalte sich meist auf die bereits erwähnten übergeordneten Ziele beziehen. Trotz dieser allgemeinen Ausrichtung wird noch einmal diskutiert, inwiefern trotz der bereits untersuchten allgemeinen Zielorientierung (vgl. 12.2) ein gewisses Maß an Individualisierung vorhanden ist. Methoden der vokalen Förderung in Gruppen, seien dies Strategien zum Einsingen oder die Wahl eines adäquaten Repertoires (vgl. Schnitzer, 2008b; Pachner & Barth, 2008), tragen nur bedingt zur Beantwortung der Forschungsfragen bei, weshalb insbesondere die Subkategorie *FÖRDERUNG* nicht vertieft analysiert wird. Anschließend wird die Subkategorie *FEEDBACK* untersucht, wobei hauptsächlich die individuelle Förderperspektive betrachtet wird. Abschließend werden die beiden Subkategorien *DIAGNOSE* und *DIFFERENZIERUNG*

und die induktiv ergänzte Subkategorie *DIFFERENZ* besprochen. Die quantitative Analyse der Kodierhäufigkeiten von AFL hat bereits gezeigt, dass diese spezifischen Elemente kaum Beachtung fanden (11.2).

Übergeordnete Ziele und Fördermethoden

Die Subkategorie *ZIELVORSTELLUNGEN* wurde 61-mal kodiert, durchgehend in allen Interviews. Die beiden in einer weiteren Ebene induktiv hinzugefügten Elemente *PERSÖNLICHKEIT* und *GESANG* führten bei der Analyse der Porträts zu der Typik 1 mit der Ausrichtung auf vokale oder persönlichkeitsbildende Ziele (vgl. 12.2). Entsprechende Tendenzen zur Individualisierung konnten indirekt festgemacht werden (vgl. 12.3). Speziell der Einsatz vielfältiger Unterrichtssettings (Pezenburg & Dyllick, 2018) und eine empathische Beziehungsgestaltung zeigen sich als Grundlage individualisierender Momente im Singunterricht. Die Subkategorie *FÖRDERUNG* wurde insgesamt 199-mal kodiert, was 52.4 % aller Kodierungen zu *ASSESSMENT FOR LEARNING* ausmacht. Um die Subkategorie fassbarer zu machen, wurde eine weitere Ebene mit folgenden sieben Subkategorien eingefügt: *GEHÖRBILDUNG/THEORIE, ARTIKULATION UND DYNAMIK, TONALE UND RHYTHMISCHE FÖRDERUNG, AUFTRITTSKOMPETENZ UND EMOTION, LIEDEINFÜHRUNG, STIMMBILDUNG (KÖRPER, ATMUNG, KLANG)* und *MOTIVATION*. Inhaltlich beziehen sich die beschriebenen Fördermethoden meist auf spezifisch stimmbildnerische Maßnahmen, die in verschiedenen Methodenbüchern entsprechend besprochen werden (Schnitzer, 2008b; Pachner & Barth, 2008; Bolender & Müller, 2012a) und nicht im Fokus der vorliegenden Arbeit stehen. Die Vielfalt an Aussagen zur Förderung schließt an den Gedanken von Mantie (2019) an, der Musikunterricht per se als Fördersetting beschreibt. In all diesen Aussagen finden sich allerdings keine Hinweise auf individuelle Assessments und darauf folgende spezifische Förderung, einzelne Lehrkräfte sprechen sich sogar gegen eine individuelle Förderung aus.

Auf der Suche nach den spezifischen Förderelementen von AFL besteht zudem die Gefahr, die Wirksamkeit allgemeiner Fördermethoden beim Singen zu unterschätzen oder in einer normativen Intention gar zu unterschlagen. Die skalierende Analyse zeigte, dass mindestens 50 % des Singens in einer frontalen Unterrichtssituation stattfindet (vgl. 11.3). Bereits mehrfach erwähnt wurde die Studie von Gembris (2015), die diesem Singen „pro-soziale Effekte gemeinsamen Musizierens auf der Basis von

Synchronisierungserfahrung“ (S. 12) beimisst. Zudem schlägt Ashley (2015) explizit vor, den Schülerinnen und Schülern einen anonymen Raum im Klassenchor zur Verfügung zu stellen, um experimentieren zu können. Beide Argumente weisen auf Lernfortschritte hin, die nicht auf einer individuellen Ebene evaluiert werden, aber trotzdem stattfinden, wie sich beispielsweise beim stark frontal ausgerichteten Unterricht von Herrn Brugger (01) in seiner eigenen Beschreibung zu einer gelungenen Kompetenzentwicklung zeigt. Weitere Lehrkräfte offenbaren allerdings eine gewisse Enttäuschung über die Resultate ihres frontal ausgerichteten Singunterrichts, wie die Analyse in der Typik 1 für die Gruppe der vokal Fördernden zeigt. Interessanterweise nehmen diese Lehrkräfte Stellung gegen eine individuelle Betreuung.

Die Argumente gegen Individualisierung beziehen sich einerseits auf fehlende Expertise und andererseits auf fehlende Zeit. Im Interview von Herrn Rothenburg (07) werden beide Argumente genannt: „sonst [...] ist immer [...] die ganze Klasse da. [...] Ich habe keine Zeit, nachher ist die Mathestunde, vorher ist die Turnlektion gewesen. Also das gibt es nicht. [...] Und ich bin auch nicht Sologesangslehrer.“ (07:286 ff.) Die fehlende Expertise wird auch in der Subkategorie *VORBILD-FUNKTION* unter *BERUFS-BIOGRAFIE* angesprochen, weshalb die Subkategorie entsprechend berücksichtigt wird. Drei Lehrer betonen, dass sie nicht Sologesangslehrer bzw. Sänger sind (01, 07, 09) und deshalb auf individuelles Coaching für die Entwicklung vokaler Ziele verzichten. Gleichzeitig zeigt sich in den Kodierungen, dass keine der professionell ausgebildeten Sängerinnen (05, 06, 11) in Betracht ziehen, vokale Herausforderungen in geplanten Unterrichtssequenzen auf einer individuellen gesangstechnischen Ebene anzugehen. Das zu Beginn erwähnte Argument, keine Gesangslehrkraft zu sein, weist deshalb hauptsächlich auf eine nicht vorhandene Gepflogenheit oder auf eine Zurückhaltung gegenüber einer individuellen Betreuung von Lernenden hin. Die Tendenz zur Zurückhaltung gegenüber individueller Betreuung hat sich bereits bei der Betreuung von PPS und beim Umgang mit der adoleszenten Stimmentwicklung in einer zumeist beobachtenden Haltung gezeigt (vgl. 13.1).

Das zweite Argument gegen Individualisierung und damit auch gegen die Umsetzung formativer Prozesse ist der Zeitfaktor, der auch von Linn (2017) ins Feld geführt wird. Herr Rothenburg (07) verweist auf den Umstand, dass er von einem Fach zum nächsten wechseln muss und die

organisatorischen Ansprüche individualisierende Unterrichtssequenzen nicht zulassen. Indirekt wird dieses Argument bei einzelnen Lehrkräften auch bei der Betreuung der Mädchen beschrieben (vgl. 13.1). So sind einige Lehrkräfte froh, dass die Mädchen als Selbstläufer wenig Betreuung benötigen. Die Herausforderung der zeitlichen Ressourcen wird zwar nur in Einzelfällen angeführt, erhält aber dort eine zusätzliche Brisanz, wo nur eine Wochenstunde für den gesamten Musikunterricht zur Verfügung steht. Gegen das Argument fehlender zeitlicher Ressourcen sprechen allerdings die bereits erwähnten Beispiele, die zeigen, wie Lehrkräfte mit vielfältigen Unterrichtssettings trotzdem zumindest auf implizite Weise individualisierend arbeiten.

Die Ausrichtung auf allgemeine Ziele und die vorhandenen Fördermethoden zeigen, dass der Wunsch nach vokaler Entwicklung zwar da ist (vgl. 12.2), meist aber auf einer allgemeinen Ebene und nur teilweise mit Bezug auf eine individuelle Entwicklung, wie sich nun gezeigt hat. Zudem werden zeitliche Ressourcen und fehlendes methodisches Knowhow als Argument gegen Individualisierung ins Feld geführt. Die Studie von Blanchard (2019) fügt dieser Erkenntnis zusätzlich Beispiele hinzu, in denen das Singen ohne Ansprüche an vokale Entwicklung unterrichtet wird. Obwohl dies in der vorliegenden Studie nicht bestätigt werden kann, wird doch am Rand auf ein entsprechendes Beispiel hingewiesen. Frau Basler (05) beschreibt dies wie folgt:

„Und ich hab mal bei einem anderen Kollegen, nicht an dieser Schule ein Negativ-Beispiel erlebt, find ich jetzt jedenfalls, der halt recht gut Klavier spielen konnte, haben sie sich rumgestellt und alle Popsongs runtergesungen, aber es war ein reines Runtersingen, es war kein Gestalten. Das ist nicht, find ich, der Auftrag für mich als Lehrerin an der Schule.“ (05:596 ff.)

Erfreulicherweise scheint dieser Modus bei den untersuchten Lehrkräften nicht die Regel zu sein. Insgesamt bleibt der Bezug zu den Anforderungen an AFL bei den beiden Subkategorien vage.

Feedback

Spätestens seit Hattie (2009) in seiner Metastudie für *feedback* eine Effektstärke von $d=0.70$ errechnet hat¹⁷³, wird in Schulen und im Musikunterricht (vgl. Biegholdt et al., 2020) vermehrt Wert auf Feedback gelegt. Dies zeigt sich in der vorliegenden Arbeit in der Anzahl der Kodierungen zu der Subkategorie *FEEDBACK*. Mit 64 Kodierungen bzw. 16.8 % aller Kodierungen von AFL scheint Feedback ein wichtiger Faktor zu sein. Im Interviewleitfaden wurde die Frage nach Feedback explizit auf verschiedene Unterrichtssituationen bezogen, von Einzelsituationen bis zum frontalen Unterricht. Kodiert wurden Textstellen, in denen von Feedback zwischen der Lehrkraft und den Lernenden gesprochen wird. Die meisten Aussagen betreffen allgemeines Feedback, in Form von Lob (z. B. 01:197 ff.), Appell (04:141 ff., 05:314) oder positiver Verstärkung (10:492, 11:393), und werden in frontalen Unterrichtssituationen im Rahmen des Singens erteilt. Ziel dieser allgemeinen Feedbacks ist meist auch eine lernförderliche Atmosphäre, die thematisch unter AAL eingeordnet ist und deshalb später besprochen wird.

An dieser Stelle werden zwei Beispiele besprochen, in denen Feedback über allgemeine Anweisungen hinausgeht und einen spezifischen Charakter von AFL hat. Herr Leimbacher (10) beschreibt mehrfach, wie er die frontale Unterrichtssituation aufbricht, sei es durch das situative Trennen der Klasse bei Bedarf (10:63 ff.) oder durch die Ansprache im frontalen Setting:

„Und dann kann es sein, dass ich mal schnell sage: Stopp, das ist nicht so gut gewesen, können wir da noch einmal schnell. Oder: Achtung, ihr steht falsch. Oder, den Jungs, T., du mit deiner schönen Männerstimme, du musst doch in die Mitte stehen. Links und rechts ist nicht ganz so sicher, komm, wechseln wir.“ (10:316 ff.)

Aus dem Kontext des Zitats geht hervor, dass die Klasse im Kreis um den Flügel steht und mehrstimmig singt. Er unterbricht, reorganisiert, um Unsicherheit zu begegnen, und baut gleichzeitig den Jungen auf, der bereits

¹⁷³ Die Effektstärke von $d=0.70$ übersteigt den Wert von 0.40 als Anzeige einer natürlichen Entwicklung deutlich. Daten aus 250+ INFLUENCES ON STUDENT ACHIEVEMENT (2017).

in der Phase 3 nach Cooksey (2000) angekommen ist. Im Anschluss beschreibt er sogar, wie einzelne Jungen in dieser Situation eine Zeile allein singen. Obwohl eine frontale Unterrichtssequenz stattfindet, behält der Lehrer die Bedürfnisse Einzelner im Auge und reagiert. Herr Leimbacher (10) adaptiert im vorliegenden Fall die Choraufstellung aufgrund einer situativen diagnostischen Überprüfung. Diese wird zwar nicht als spezifisches Assessment durchgeführt, sondern erfolgt direkt beim gemeinsamen Singen, belegt aber gerade für das Chorsetting eine Möglichkeit, AFL zu praktizieren, und verweist auf den fortwährenden formativen Prozess im Musikunterricht (Mantie, 2019). Herr Brugger (01) beschreibt ein ähnliches Beispiel im frontalen Unterrichtssetting, verzichtet allerdings auf eine direkte Ansprache der betroffenen Jungen, macht lediglich Handzeichen, um eine Bloßstellung zu verhindern (01:57 ff.). Er verzichtet dabei aber auch auf das wertschätzende mündliche Feedback, das die Unterrichtssequenz von Herrn Leimbacher (10) prägt.

Herr Berner (09) lässt seine Schülerinnen und Schüler in monoedukativen Settings mehrfach pro Jahr auftreten. Er beschreibt ausführlich die Akzeptanz und Wirkung seines Feedbacks:

„Ich sage das persönliche Feedback sofort, auch vor den anderen, das ist eigentlich nie das Problem, [...] ich kann jemandem mal sagen, schau mal, singen ist jetzt einfach mal nicht unbedingt dein Ding. Wissen wir, hören wir, aber ich finde es super, weißt du, wie du hin stehen kannst, und du machst es einfach so gut, es geht. [...] Und sie empfinden das nicht, [...] im Gegenteil, weil man muss sie einfach ermutigen, [...] und es passiert ja auch immer wieder, dass Leute, bei denen du vorher das Gefühl hattest, der oder die wird jetzt nie einen Ton treffen, und plötzlich, plötzlich geht es. Und nachher muss man das dann schon bei der Bewertung bei der mündlichen Bewertung, wahnsinnig groß rausheben, [...] sie sollen ja, es soll ja niemand geknickt rausgehen.“ (09: 279 ff.)

Herr Berner (09) scheint eine Form des Feedbacks gefunden zu haben, die unter Beachtung seiner Wortwahl vielleicht nicht den Ansprüchen an einen human-anerkennenden Musikunterricht (Tellisch, 2015) genügt, aber von den Lernenden akzeptiert wird. Die Betonung einer über das Vokale hinausgehenden Persönlichkeitsentwicklung durch das Singen wurde bereits in Typik 1 deutlich (12.2) und scheint im Feedback des Lehrers der Hebel zu sein, dass Lernende negative Kritik zu ihrem Gesang nicht *empfinden*.

Zudem scheint die Ermutigung während der wiederholten Auftrittssituationen sogar auf vokalem Gebiet Früchte zu tragen. Im Gegensatz zum jährlichen Vorsingen, das von den meisten Lehrkräften gepflegt wird, zeigt die von Herrn Berner (09) gepflegte Methode als formatives Assessment Wirkung¹⁷⁴. Treibende Kraft ist auch in diesem Beispiel wertschätzendes Feedback.

Bereits besprochen wurde das individuelle Feedback in summativen Bewertungssituationen. Dabei wurde hauptsächlich hervorgehoben, dass mit der zyklischen Überprüfung, die gegen Ende des Schuljahres stattfindet, wenig Gewähr besteht, dass die Hinweise in einem formativen Sinn wirksam werden können. Zudem scheinen die Beurteilungskategorien zu allgemein zu sein, um spezifische Kompetenzzuwächse zu generieren. Auf diese Problematik verweist auch Holcomb (2019), wenn er den holistischen *Rubrics* aufgrund der allgemeinen Beschreibung nur wenig Fördereffekte zutraut.

Insgesamt scheint die Wichtigkeit von Feedback den Lehrkräften bewusst zu sein, sei dies in frontalen Settings oder auch bei individuellen Feedbacks. Der bewusste Einsatz von Feedbacks, um im Sinne von AFL eine Wirkung zu erzeugen, zeigte sich nur in Einzelfällen oder bei der Besprechung formativer Elemente in summativen Beurteilungen.

Diagnose und der Unterschied zwischen Differenz und Differenzierung

Im Interviewleitfaden wurde mit der Frage nach Kenntnissen zu einzelnen Stimmen innerhalb einer Klasse das Thema Diagnose und Individualisierung implizit angestoßen. Die Kodierhäufigkeiten der drei Subkategorien *DIAGNOSE* (3.4 %), *DIFFERENZIERUNG* (5.5 %) und *DIFFERENZ* (5.8 %) sind allerdings marginal. Die Subkategorie *DIAGNOSE* vereint Aussagen, die direkt oder indirekt auf diagnostische Elemente verweisen. Sie wurde lediglich 13-mal kodiert, wobei sich nur bei knapp der Hälfte der Lehrkräfte Hinweise fanden. Auch die Subkategorie *DIFFERENZIERUNG* erhielt mit 21 Kodierungen, verteilt auf 7 Lehrkräfte, wenig Beachtung.

¹⁷⁴ Aus dem Interview wird klar, dass mehrfach in monoedukativen Settings vorgesungen werden muss und dies auch bewertet wird. Unklar bleibt, ob jedes Mal auch eine summative Beurteilung mit Note erfolgt. Das Vorgehen von Herrn Berner (09) hat aber einen stark formativen Charakter.

Zudem zeigte sich während der Erarbeitung der QIA die Notwendigkeit, zwischen wahrgenommener Differenz und einer differenzierenden Tätigkeit zu unterscheiden. So wurde induktiv die Subkategorie *DIFFERENZ* ins Kategoriensystem aufgenommen und 22-mal kodiert, verteilt auf 8 Lehrkräfte.

Bereits in der quantitativen Einschätzung der drei Subkategorien zeigte sich, dass diese für AFL spezifischen Elemente bei den Lehrkräften wenig Beachtung finden (vgl. 11.2). Zusätzlich wurden bereits bei der Analyse von Aussagen zum Umgang mit der adoleszenten Stimmentwicklung Elemente von Diagnose und Differenzierung in Einzelfällen sichtbar (vgl. 13.1). Diagnostische Elemente zeigten sich z. B. bei der Betrachtung des auffälligen Stimmwechsels eines Mädchens und bei der Betreuung von PPS. Im ersten Fall wird die Lehrkraft durch eine auftretende Schwierigkeit zu einer Reaktion herausgefordert und diagnostiziert zuerst, um dann eine Intervention durchzuführen. Bezüglich der Betreuung der PPS zeigt sich auch eine positive Einzelbehandlung, gleichzeitig werden bei anderen Lehrkräften auch Fehlkonzepte aufgedeckt, die zum Verzicht individueller Betreuung führen (vgl. 13.1). Die Kodierungen der Subkategorie *DIAGNOSE* nehmen ebenfalls mehrfach Bezug zur adoleszenten Stimmentwicklung (05:113 f., 06:213 ff., 08:200 ff., 11:100 ff.). Dies scheint eine Thematik zu sein, in der Lehrkräfte eine gewisse Achtsamkeit zeigen bzw. ein anamnestisches Gespür entwickelt haben. Die Analyse im Abschnitt zuvor spricht allerdings dagegen, denn die Lehrkräfte verzichten bei Mädchen meist und bei Jungen in der Regel auf spezifische Interventionen bei der Mutation, was wiederum auf fehlende Diagnose zurückverweist und die Analyse von Brunner (2014) zum Fehlen von Diagnosefähigkeit bestätigt. Neben diesen spezifisch auf die Entwicklung der adoleszenten Stimme ausgerichteten diagnostischen Elementen verweisen weitere Kodierungen auf Einzelbeispiele.

Zwei Lehrkräfte berichten von einer diagnostisch ausgerichteten Situation im frontalen Unterricht. In beiden Fällen werden einzelne Lernende aufgerufen, einen Ton oder eine Phrase allein zu singen. Frau Fricker (03) überprüft nach den Ferien den Stimmstatus (03:37 ff.) (vgl. Fuchs, 2015), Herr Brugger (01) nutzt die Kennenlernphase zu Beginn der Sekundarstufe 1, um einzelnen Stimmen eine Plattform zu geben. Er beschreibt wie folgt „Sing schnell vor, da hat es solche, die machen das – ist ja klar er hat es gesagt – und es ist noch keine Zeit gewesen, um Hemmungen zu haben.“

(01:253 f.) In beiden Situationen gibt es zwar Erkenntnisse für die Lehrkraft, es wird aber weder davon gesprochen, dass alle Lernenden in Form einer Reihenuntersuchung zum Zug kommen, wie das Greuel und Horst (2011) vorschlagen, noch wird eine Adaption des Unterrichts angestrebt. Weitere Lehrkräfte nehmen implizit Bezug auf diagnostisches Hören beim Einsingen, um zu sehen, wo die Klassen stehen (z. B. 02:63 ff.).

Spezifisch zum Thema Diagnostik nimmt Frau Basler (05) Stellung:

„Aber trotzdem wäre es mein Wunsch, eigentlich, und ich fände es absolut sinnvoll, am Anfang jedes Schuljahres, jeden Schüler einmal einzeln zu hören. [...] Einerseits [...] Diagnostik, aber auch, um [...] den Schülern eine gewisse Selbstsicherheit zu geben. Ich kann das, oder ich, da ist mein Schwerpunkt. Das krieg ich hin. Das krieg ich weniger hin. [...] Ich schaff es einfach nicht mit meiner Menge an Schülern, die ich hab.“ (05:278 ff.)

Frau Basler (05) betont, dass sie problematische Stimmverhältnisse im Einzelunterricht zu lösen versucht (05:219 ff., vgl. auch 01:283 ff.), ist aber auch stark an einer allgemeinen Diagnostik interessiert. Dabei vermutet sie durch das Generieren von diagnostischen Daten auch einen positiven Effekt im Bereich des vokalen Selbstkonzepts, weil die Schülerinnen und Schüler wissen, woran sie sind. Dies geht über die Erfassung des Stimmstatus hinaus (Greuel & Horst, 2011), deckt sich aber mit den Erkenntnissen von Bojack-Weber (2012), die diagnostischen Tools ebenfalls eine positive Wirkung auf die individuelle Entwicklung zuschreiben. Frau Basler (05) verweist auch an anderer Stelle auf ihre Hoffnung, methodische Werkzeuge zu erhalten (05:477 ff.), um im Sinne des *principles of operability* (Brophy, 2019c) Diagnosen im Unterricht einbauen zu können. Sie betont dann aber noch einmal den limitierenden Faktor der zeitlichen Ressourcen.

Die bereits mehrfach erwähnte Studie von Brunner (2014) zur fehlenden Diagnosekompetenz beim Singen bestätigt sich auch explizit in Beispielen der vorliegenden Studie. Frau Derendinger (11) beklagt z. B. den Misserfolg beim Singen in einer spezifischen Klasse: „[W]enn ich mit diesen singe, kriege ich fast eine Krise. Weil einfach gar nichts kommt. Und ich habe mit ihnen Sachen probiert, sie so ein wenig rauszukitzeln, und so ein wenig rauszufinden, was genau das Problem ist.“ (11:241 ff.) Die Lehrerin versucht mit verschiedenen Methoden, die Klasse zum Singen zu bringen, scheint aber nicht mehr weiter zu wissen. Ihr Vorgehen in

der Herausforderung erfolgt mit dem System *Fehler und Irrtum*, führt sie aber nicht zu neuen Erkenntnissen, sondern in die Krise. Obwohl in diesem Beispiel die zuweilen schwierige Situation im Singunterricht (Forge, 2016) nicht außer Acht gelassen werden darf, scheint doch an dieser Stelle das methodische Knowhow zu fehlen, die Klassendynamik zugunsten des Singens zu verbessern. Zu beachten wären insbesondere die individuellen vokalen Selbstkonzepte beim Singen (Ashley, 2015; Spychiger, 2017). Allerdings weisen gerade Rieche et al. (2018) auf die fehlende Wahrnehmungsfähigkeit von Selbstkonzepten hin, was an dieser Stelle als Schlüssel der krisenhaften Situation vermutet werden darf.

Neben der fehlenden Diagnosekompetenz sind einzelne Elemente der Subkategorien *DIFFERENZ* und *DIFFERENZIERUNG* zu beobachten. Insgesamt bestätigt sich die Analyse von Linn (2017) bezüglich der vorhandenen Wahrnehmung von Differenz bzw. Heterogenität: Unterschiede werden wahrgenommen, dies führt aber nicht zu adaptierten Unterrichtsprozessen. Zwei der von Hoene und Thurman (2011) aufgegriffenen Heterogenitätsmerkmale Lernstil, Lerntempo, Lerndisposition, Lernstand, Lerninteresse (S. 21) finden in den Kodierungen Beachtung. Bereits besprochen wurde der Bereich Lerndisposition, wobei insbesondere die Phänomene der adoleszenten Stimmentwicklung zu subsummieren sind. Einzig Frau Derendinger (11) bringt zusätzlich das Niveau der verschiedenen Leistungsstufen ins Spiel. Sie erwähnt, dass der Singunterricht in den tieferen Niveaus schwieriger sei (11:73 ff.). In dieser expliziten Form wird das Leistungsniveau bei den anderen Lehrkräften nicht diskutiert, wohl auch, weil sie die einzige Lehrkraft ist, die alle drei Niveaus unterrichtet. Einzig Herr Berner (09) erwähnt diese Unterschiede ebenfalls, obwohl er das höchste Niveau momentan nicht unterrichtet (09:124 ff.).

Die Lehrkräfte nehmen überdies hauptsächlich im Bereich Lerninteresse Differenzen wahr. Bereits besprochen wurden stereotype Äußerungen zu Mädchen und Jungen. Dabei zeigte sich ein stereotyper Unterton bezüglich einer geringeren Motivation bei Jungen, die sich in einer anspruchsvolleren Pflege der Klassendynamik niederschlägt (vgl. 13.1). Fehlende Motivation bemängelt z. B. Frau Derendinger (11), die bemerkt, dass einige Schülerinnen und Schüler kaum Freude am Singen zeigen (z. B. 11:52 ff.). Insgesamt fällt auf, dass die Wahrnehmung von Differenz nicht zur Adaption von Unterricht führt. Die Kodierungen zu *DIFFERENZIERUNG* zeigen allerdings auch nur basale Formen einer Unterrichtsadaption. So berichten

einige Lehrkräfte davon, dass bei der Wahl von Liedern zum Vorsingen eine gewisse Freiheit besteht, die durch die Breite des gesungenen Repertoires entsteht und den Schülerinnen und Schülern einfachere und schwierigere Lieder zur Auswahl überlässt (01:172 ff., 06:276 ff., 09:195 ff.). Insbesondere Herr Berner spricht von individuellen Anspruchsniveaus, die er damit ermöglicht (09:64 ff., 09:294 ff.). Daneben achten Lehrkräfte auf eine adäquate Gruppeneinteilung, die insbesondere die Tessitura innerhalb der Gruppe beachtet (03:199 ff.).

Die Ausbeute in den Subkategorien *DIAGNOSE*, *DIFFERENZ* und *DIFFERENZIERUNG* ist bescheiden. Im Gegensatz zu anderen Bereichen finden sich nur wenige konkrete Ansätze, die den Bereich AFL begünstigen.

Der formative Bereich AFL

Die Analyse zum Bereich *assessment for learning* bestätigt die Vermutung, dass der Bereich AFL bei den Interviewpartnerinnen und -partnern wenig Beachtung findet. Die Förderung findet aufgrund der übergeordneten Ziele und der zyklischen Beurteilung meist im Klassenverband durch allgemeine Stimmbildungsmaßnahmen statt. Formative Prozesse werden kaum explizit erwähnt, sondern erschließen sich aus dem breiten Repertoire an Fördermaßnahmen, wobei eine Zurückhaltung für individuelle Förderung auszumachen ist. Der Blick auf individuelle Bedürfnisse zeigt einige wenige Beispiele individueller Betreuung, so die bereits früher beschriebene Betreuung eines Stimmbruchmädchens oder die Aussagen weiterer Lehrkräfte, die individuelle Probleme angehen. Die Adaption des Unterrichts aufgrund von Diagnosen oder formativen Assessments findet sich im vorliegenden Datenmaterial aber nicht. Die zentrale Forderung von AFL, Informationen zu sammeln, um den Unterricht adaptiv zu gestalten, ist somit weder in einer allgemeinen Tendenz noch in Einzelfällen im Singunterricht zu sehen. Inwiefern Schülerinnen und Schüler trotzdem im Unterricht profitieren können, zeigt sich im Bereich AAL, in dem die lernförderliche Atmosphäre als Grundlage von metakognitiven Prozessen dargestellt wird.

Assessment as Learning

Assessment as learning (AAL) ist eine Teildimension formativen Assessments und somit eng mit Förderprozessen verknüpft (Earl, 2013, S. 39).

Das Herausstellen von AAL als eigenständigen Bereich ist vor allem der Absicht geschuldet, einen Fokus auf eigenverantwortliches Lernen von Schülerinnen und Schülern zu legen. Damit wird die kognitivistische Komponente von Assessment (C. Schmidt, 2020, S. 15) betont. In AAL werden metakognitive Elemente, wie die Fähigkeit zur Selbstevaluation sowie die Beobachtung, Reflexion und Steuerung der eigenen Entwicklung vereint, immer im Rahmen von Settings, in denen es für die Lernenden sicher ist, Risiken einzugehen (vgl. AAL in Tabelle 5).

Für die Subkategorie *ASSESSMENT AS LEARNING* wurden bei der Erstellung der QIA deduktiv sechs Subkategorien definiert, die diese Perspektive aufnehmen. In der quantitativen Analyse konnte bereits gezeigt werden, dass mit 13 % der Kodierungen innerhalb der Kategorie *ASSESSMENT* der Bereich AAL wenig Beachtung gefunden hat (vgl. 11.2). Aufgrund der geringen Kodierhäufigkeiten wird auch in diesem Bereich die Methode der Explikation (Mayring, 2015, S. 90) angewendet, um Einzelbeispiele aus den Interviews durch Beiziehung von weiteren Textstellen und Kontextwissen interpretieren zu können. Zudem wird noch einmal daran erinnert, dass Textstellen nicht nur kodiert wurden, wenn von einem absichtsvollen Handeln im Sinne von AAL berichtet wurde, sondern auch dann, wenn intuitiv entwickelte Elemente auf AAL hingewiesen haben. Vier der sechs Subkategorien beziehen sich direkt auf Aussagen zu den Lernenden: *SELBSTEINSCHÄTZUNG*, *REFLEXION*, *ZIELE SETZEN* und *ENTWICKLUNG VON EIGENVERANTWORTUNG*. Ergänzt wurden sie mit der Kategorie *UNTERSTÜTZENDE LERNUMGEBUNG*, in der Aussagen zu den Voraussetzungen und dem Nährboden für selbstevaluative Vorgänge kodiert wurden. Die Subkategorie *PEERFEEDBACK* ist in diesem Bereich zwar artfremd, wurde aber hier angegliedert, weil der Austausch auf der Ebene der Lernenden geschieht und nicht die formativen Prozesse von AFL tangieren. Gestartet wird mit der Subkategorie *UNTERSTÜTZENDE LERNUMGEBUNG* und der Diskussion der mehrfach genannten Thematik *BLOSSTELLEN* als zentrale Stellschraube einer guten Atmosphäre beim Singen.

Unterstützende Lernumgebung

Frau Fricker (03) beschreibt die Grundvoraussetzung für ihren Singunterricht wie folgt: „Erstens versuche ich, eine gute Laune zu verbreiten. Das finde ich ganz wichtig. Dass sie finden, das macht die gerne und die ist gern

da bei uns. Und sie hat uns gerne. Ich glaube, das ist wirklich die Grundvoraussetzung.“ (03:308 ff.) Frau Fricker (03) sieht ihr empathisches Verhalten und ihre Begeisterung für den Musikunterricht als Grundlage für das Verbreiten guter Laune. Ihre Aussage knüpft an die Typik 2 aus dem letzten Kapitel (vgl. 12.2) an. Wie sich dort gezeigt hat, können vokale oder persönlichkeitsbildende Ziele nur erreicht werden, wenn die Lehrkräfte ein empathisches Verhalten zeigen. Distanziertheit führt hingegen zu Widerstand im Klassenzimmer. Im Zusammenhang mit einer unterstützenden Lernumgebung verweist diese Erkenntnis in erster Linie auf die Wichtigkeit eines guten Classroom-Managements (Grohé, 2011), führt aber nicht zwingend zu den gewünschten metakognitiven Prozessen von AAL. Die 19 Kodierungen der Subkategorie *UNTERSTÜTZENDE LERNUMGEBUNG* zeigen über das Atmosphärische hinaus auf einen spezifischen Umgang mit Fehlern und Bloßstellung im Singunterricht, weshalb auch die Subkategorie *FEHLERKULTUR – BLOSSSTELLEN* in *KLASSENDYNAMISCHE PHÄNOMENE* in die Analyse integriert wird.

Oser und Spychiger (2005) benutzen den Begriff Fehlerkultur¹⁷⁵ im Rahmen ihres Vergleichs einer Fehlervermeidungs- mit Fehlerermutigungsdidaktik (S. 164). Sie vertreten darin die Überzeugung, das Richtige nur im Vergleich zum Falschen als dies erkennen zu können: „Man kann auf das Falsche nicht verzichten, will man das Richtige richtig erkennen. Man muss aber das Falsche bekämpfen, damit es ein Richtiges wird.“ (S. 13) Fehlervermeidungsdidaktik wird von Oser und Spychiger als Unterrichtsstrategie beschrieben, in der Fehler nicht nur vermieden werden, sondern der Unterricht möglichst so konzipiert ist, dass Schülerinnen und Schüler gar nicht dazu kommen, Fehler zu machen, und diese somit auch nicht sichtbar werden (S. 164). Dies steht im Gegensatz zu einem konstruktivistischen Lehr- und Lernverständnis, in dem Fehler nicht nur dazu gehören, sondern als Lernpotenzial verstanden werden, immer unter der Prämisse, dass diese in einem vertrauensvollen Umfeld aufgefangen werden können (S. 164). In einer Fehlerermutigungsdidaktik wird somit bewusst mit Fehlern gerechnet. Wie die Analyse der Kodierungen zeigt, befürchten einige Lehrkräfte negative Effekte, wenn beim Singen

¹⁷⁵ Oser und Spychiger (2005) prägten den Begriff Fehlerkultur im Rahmen der Entwicklung ihrer Theorie des Negativen Wissens.

Unzulänglichkeiten öffentlich werden, und leiten daraus didaktische Entscheidungen ab, die sich gut auf die Überlegungen von Oser und Spychiger beziehen lassen. Die Überzeugungen von zwei Lehrkräften werden deshalb exemplarisch als Beispiele abweichender Strategien im Umgang mit Bloßstellung auf die Fehlervermeidungs- und Fehlerermutigungsdidaktik von Oser und Spychiger (2005) bezogen und vertieft analysiert. Herr Neuendorf (08) betont die Notwendigkeit von Vertrauen, um sich vokal zu äußern, ohne bloßgestellt zu werden, Herr Brugger (01) verlässt nie das anonymisierende Setting des Klassenchors, um auf keinen Fall jemanden bloßzustellen.

Herrn Neuendorf (08) fasst seine Aufgabe im Singunterricht folgendermaßen zusammen:

„[M]an muss Stimmung erzeugen zuerst, man muss Basis schaffen, dass auch Vertrauen da ist, oder, weil Singen so etwas Persönliches ist. Ja, also das, das ist für mich eigentlich fast die zentrale Aufgabe gewesen in all den Jahren, eine Atmosphäre zu schaffen, dass sich das Kind nicht [...] gehemmt fühlt oder bloßgestellt wird, dass es der Umgebung Klasse vertrauen kann, sich zu äußern.“ (08:69 ff.)

Herr Neuendorf (08) verweist in seiner Aussage auf den starken Bezug zwischen Singen und Persönlichkeit, was in der vorliegenden Arbeit bereits mehrfach argumentiert wurde (vgl. 1.2, 4.1). Vokale Äußerungen im Klassensetting sind deshalb für ihn per se eine diffizile Angelegenheit und bedürfen einer Vertrauensbasis, nicht nur bezüglich der Beziehung zwischen Lernenden und Lehrkraft, sondern auch zwischen den Peers. Diese vertrauensvolle Atmosphäre ist dann die Grundlage für vokale Äußerungen und lässt auf der Ebene der Produzierenden den Abbau von Hemmungen zu, bedingt aber auf der Ebene der Rezipienten gleichzeitig einen wohlwollenden Umgang bzw. die Vermeidung von Bloßstellung. Den Ort der Bloßstellung verortet Herr Neuendorf (08) im Klassensetting, wo sich möglicherweise Peers gegenseitig bloßstellen. Aus seiner Sicht stellt nicht in erster Linie die Lehrkraft bloß, wie das in einigen Artikeln thematisiert wird (z. B. Weber, 2018) und in der Arbeit von Tellisch (2015) explizit untersucht wurde, sondern führen klassendynamische Phänomene zu bloßstellenden Momenten. Um seinen deklarierten Ansprüchen zu genügen, steuert er spezifisch die Klassendynamik, führt monoedukative Unterrichtssequenzen durch und berücksichtigt gezielt die Potenziale einzel-

ner Schülerinnen und Schüler (vgl. 12.1). Er sieht sich selbst in der Rolle eines Mentors und Guides (08:477 f.; vgl. Tabelle 4). Herr Neuendorf (08) spricht nicht explizit von Fehlern oder einer Fehlerkultur, seine Überzeugungen zeigen aber deutlich die Nähe zu einem konstruktivistischen Lehr- und Lernverständnis.

Im Gegensatz dazu beschreibt Herr Brugger (01) explizit seine Angst, Schülerinnen und Schüler bloßzustellen. Er begründet seine Haltung aufgrund biografischer Erlebnisse:

„Ehm, das ist vielleicht meine persönliche Optik, dass ich das gar nicht will, dass das [...] passiert, vielleicht im Hinterkopf ist diese Angst vom Bloßstellen, und darum also das provoziere ich auch nicht. Das ist mir so fremd. Ganz schlimm ist für mich, ich habe etwa zwei-drei Personen kennengelernt in meinem Leben, die gesagt haben, sie können nicht singen.“ (01:270 ff.)

Im weiteren Verlauf des Interviews beschreibt er, dass zum Teil die Eltern, aber auch Musiklehrkräfte Ursprung der negativen Selbstkonzepte gewesen seien. In der Konsequenz bleibt Herr Brugger (01) mit seinen Klassen konsequent im frontalen Setting und vermeidet jede Form individueller Ansprache oder Gruppenarbeiten, um jede Möglichkeit einer Bloßstellung zu vermeiden. Er ist sich bewusst, dass dies für Jugendliche mit einem negativen vokalen Selbstkonzept nicht hilfreich ist, bezeichnet sich selbst auch als feige (01:280 f.), hofft aber, dass die Jugendlichen „merken, wie ich den Umgang damit [mit dem Singen] pflege, sodass sie vielleicht aus dieser Situation rauskommen“ (01:282 f.). Er vertraut auf eine implizite Vorbildwirkung. Das von Herrn Brugger (01) beschriebene Chorsetting ermöglicht zwar den anonymisierten Umgang mit der eigenen Stimme im Chorklang, wie Ashley (2015) ihn vorschlägt, verunmöglicht aber gleichzeitig individuelle Betreuung und erschwert die für AAL vorgesehenen metakognitiven Prozesse. Im Sprachgebrauch von Oser und Spychiger (2005) kann dieser Ansatz als typisches Beispiel einer Fehlervermeidungs-didaktik bezeichnet werden.

Die beiden Lehrkräfte Neuendorf (08) und Brugger (01) stellen bezüglich der Erzeugung einer lernförderlichen Atmosphäre einen Kontrast dar. Beide sind sich bewusst, dass Singen zu heiklen Situationen, insbesondere zu Bloßstellung führen kann, sie reagieren darauf aber gänzlich verschieden. Herr Neuendorf (08) will in den Klassen einen Ort des Vertrauens

schaffen, einen Ort wo Jugendliche etwas riskieren können, was in seinem Unterricht zur Förderung individueller Potenziale und zu einer starken Betonung seiner Rolle als Mentor führt. Dies schafft die Grundlage für einen Unterricht, wo AAL eine Rolle spielen kann, wie sich später zumindest auf implizite Weise zeigt. Herr Brugger (01) dagegen vermeidet konsequent Situationen mit potenziellem Risiko zur Exposition Einzelner, bleibt konsequent in einem frontalen Unterrichtssetting und vermeidet damit Unterrichtselemente, wo Fehler als Lernchance (Oser & Spychiger, 2005) gesehen werden und Reflexion im Sinne AAL erfolgen kann.

Um den Anforderungen eines ganzheitlichen Assessments und spezifisch AAL zu genügen, ist eine vertrauensvolle Atmosphäre zentral. Dazu ist nicht nur die Beziehung zwischen Lehrkraft und Lernenden von Bedeutung, sondern auch die Beziehung zwischen den Peers. Über diese Grundlage hinaus wird im unterschiedlichen Umgang mit Bloßstellung deutlich, dass der Umgang mit individuellen Bedürfnissen und die Förderung individueller Potenziale mit einer Fehlerermutigungsdidaktik gekoppelt ist, die ihrerseits zu vielfältigen Unterrichtssettings führt. Diese grundlegenden Bedingungen für AAL sind nicht nur bei Herrn Neuendorf (08) sichtbar, sondern mehr oder weniger stark bei den Lehrkräften, die das Singen als Plattform für persönlichkeitsbildende Aspekte sehen (vgl. 12.2).

Das metakognitive Quartett

Auf dem Nährboden einer lernförderlichen Unterrichtskultur können nun die weiteren Elemente von AAL betrachtet werden. Die zugehörigen Subkategorien wurden unterschiedlich berücksichtigt: zu *SELBSTEINSCHÄTZUNG* wurden 15 Textstellen kodiert, zu *REFLEXION* 13, nur an einer Stelle wurde *ZIELE SETZEN* kodiert und bei der *ENTWICKLUNG VON EIGENVERANTWORTUNG* konnten 22 Textstellen ermittelt werden. In AAL werden die Jugendlichen zu Assessoren (*key assessor*) und die Absicht besteht darin, sie zu „Selbsteinschätzung und Selbstkorrektur oder Justierung“¹⁷⁶ (vgl. Tabelle 27) zu bewegen. Aufgrund der Kodierstrategie wurden Textstellen mit beabsichtigtem Handeln der Lehrkraft sowie Elemente, wo intuitiv günstige Situati-

¹⁷⁶ Eigene Übersetzung: Self-monitoring and self-correction or adjustment (Earl 2013, S. 42).

onen für Reflexion entstehen, kodiert. In der Analyse muss jeweils deutlich werden, ob Lehrkräfte mit Absicht die Jugendlichen zum Beobachten und aktivem Steuern des eigenen Lernens anleiten, oder ob sie lediglich von Beobachtungen solcher Elemente berichten, ohne eine ursprüngliche Absicht zu benennen. Allgemeine Tendenzen lassen sich aufgrund der kleinen Kodierhäufigkeiten kaum ablesen, spannend sind hauptsächlich einzelne Beispiele, in denen metakognitive Prozesse aufscheinen.

In der Subkategorie *SELBSTEINSCHÄTZUNG* gibt es zwar 15 Kodierungen, aber wenig konkrete Beschreibungen von beabsichtigtem Handeln. Herr Brugger erwartet von seinen Schülerinnen und Schülern, dass sie ihre Fähigkeiten bei der Wahl eines Liedes für das Vorsingen richtig einschätzen (01:172 ff.), und Herr Hilfiker (04) berichtet davon, dass er eine Selbstbeurteilung auf mündlicher Basis als Einstieg zum Beurteilungsgespräch einfordert (04:311 ff.). Herr Leimbacher (10) weist darauf hin, dass Lernende heute mutiger sind und sich mehr zutrauen, vor allem bezüglich solistischer Auftritte (10:414 ff.). Eine schriftliche Selbstbeurteilung führt Herr Neuendorf (08) im Rahmen der summativen Beurteilung durch. Die Jugendlichen füllen begleitend zum Vorsingen ein Raster aus und besprechen dieses anschließend mit dem Lehrer (08:352 ff.). Er äußert sich über die Qualität der Selbsteinschätzungen wie folgt:

„[S]ie wissen, dass sie sich selber bewerten müssen, und die meisten sagen, ich kann mich nicht bewerten, und wenn sie sich bewerten, dann bewerten sie sich sehr streng. Also ist meine Erfahrung. [...] Kommt auch keinem Knaben mehr in den Sinn, welche dort sagen, ja ich habe einen 6er verdient. Nur weil er das jetzt lustig findet.“ (08:365 ff.)

Er traut den Jugendlichen zu, sich selbst einzuschätzen, und beobachtet, dass sie sich meist zu streng bewerten. Die Formulierung „[k]ommt keinem [...] mehr in den Sinn“ (08:368 f.) lässt darauf schließen, dass sich zuweilen Lernende provokativ eine zu hohe Note gegeben haben, dies aber nicht mehr vorkommt. Herr Neuendorf (08) konnte offenkundig eine funktionierende Kultur der Selbsteinschätzung aufbauen. Dabei ist vor allem der bereits unter AOL diskutierte Aushandlungsprozess bei der Notengebung interessant. Inwiefern die Selbsteinschätzung im Sinne eines formativen Prozesses wirksam wurde, bleibt unklar. Die Mehrheit der untersuchten Lehrkräfte nutzt aber keine selbstevaluativen Verfahren.

In der deduktiv gesetzten Subkategorie *REFLEXION* können keine spezifischen Elemente gefunden werden, in denen Lernende explizit zum Nachdenken aufgefordert werden, um daraus Erkenntnisse für das eigene Singen zu gewinnen. Zwar stellen einige Lehrkräfte Fragen, in der Hoffnung Reflexion auszulösen, oder weisen darauf hin, dass sie aktives Zuhören gestalten, der Einbezug konkreter Konstrukte wie z. B. die Entwicklung positiver Selbstkonzepte mittels *possible selves*, wie sie Freer (2020) vorschlägt, konnte an keiner Stelle gefunden werden. Noch weniger ergiebig zeigt sich die Subkategorie *ZIELE SETZEN*, die nur einmal verwendet wurde. Bereits unter AOL konnte nur ein Beispiel einer auf Zielen beruhenden Bewertung gefunden werden, derselbe Lehrer (10) erwartet nun auch, dass seine Lernenden in spezifischen Unterrichtssequenzen die an sie herangetragenen Ziele eigenverantwortlich umsetzen, allerdings ohne Spielraum für ein reflektiertes Vorgehen mit eigenen Lernwegen. Das fast vollständige Fehlen von Reflexionsaufgaben und eigenverantwortlich gesteckten Zielen deckt sich mit den meist übergeordneten Zielen für das Singen, die zyklisch durch Vorsingen überprüft werden.

Die letzte deduktiv gesetzte Subkategorie aus dem metakognitiven Quartett heißt *ENTWICKLUNG VON EIGENVERANTWORTUNG*. Die Kodierungen betreffen hauptsächlich die Erwartung, eigenverantwortlich zum Klassenrepertoire beizutragen, oder äußern sich positiv, wenn Schülerinnen und Schüler Eigeninitiative zeigen. Herr Buchser (02) arbeitet als einziger mit Aufgabenstellungen, die an die Eigenverantwortung im Singunterricht appellieren. Ausgangspunkt sind disziplinarische Schwierigkeiten, die er dem tieferen Leistungsniveau zuschreibt. Seiner Beschreibung mangelt es nicht an einer gewissen Theatralik:

„Ich kann ein Lied machen, wo sie sagen, es tönt Scheiße, das tönt auch noch Scheiße und was ist heute, können wir heuten nicht singen? Ich hätte eigentlich dies und das vorgehabt, [...] soll ich euch rausschicken, ihr bereitet draußen vor und kommt rein, um vorzusingen, weil so läuft es nicht. Sie sagen, ja ist gut. [...], aber absolut erstaunlich, die kommen rein, [...] und singen alle gut vor, die Klasse kann mitmachen, und ich frage sie nachher, wieso ist es vorhin mit mir nicht gegangen? Sie sagen jeweils, wir wissen es nicht. Aber ich habe auch eine Lösung gefunden, wo ich mich aus dem Fokus nehmen kann, und sie singen dann trotzdem wieder.“ (02:139 ff.)

Herrn Buchser (02) startet mit einer frontalen Singsituation, bricht aber aus disziplinarischen Gründen ab. Drei weitere Lehrkräfte (06, 07, 11) berichten von derartigen Situationen, brechen den Singunterricht dann aber vollständig ab. Herr Buchser bricht nicht ab, sondern stellt die Aufgabe, in Gruppen ein Lied vorzubereiten und dann in der Klasse vorzutragen. Er gibt ihnen dadurch eine gewisse Autonomie (Ryan & Deci, 2000) und schafft es anschließend, die Gruppe wieder zum Funktionieren zu bringen. Aus dem Zusammenhang des Interviews geht hervor, dass dies nicht eine Einzelsituation ist, sondern das tägliche Brot einer städtischen Schule¹⁷⁷. Herr Buchser (02) berichtet von verschiedenen Gruppenarbeiten, die als Ziel immer eine Performance vor der Klasse haben, aber unterschiedlich viel Zeit beanspruchen. Auch bei Herrn Buchser zeigt sich ein konstruktivistisches Lehr- und Lernverständnis, geboren aus der Not disziplinarischer Schwierigkeiten.

Insgesamt zeigen sich bei den untersuchten Interviewpartnerinnen und -partnern nur wenige Beschreibungen von Unterrichtssituationen, die der metakognitiven Ebene zugeordnet werden können. Das Einzelbeispiel der Selbstbeurteilung von Herrn Neuendorf (08) und die Gruppenarbeiten von Herrn Buchser (02), die Eigenverantwortung aufbauen sollen, zeigen eine intendierte Form von AAL, die restlichen Beispiele können aber nur lose mit AAL in Verbindung gebracht werden. Noch eine weitere Lehrkraft (05) nimmt Bezug auf metakognitive Prozesse, die restlichen acht scheinen sich aber im Rahmen von Förderung und Beurteilung wenig Gedanken zu der kognitivistischen Ebene von Assessment zu machen. Frau Basler (05) formuliert als einzige Lehrkraft explizit das Anliegen, Schülerinnen und Schüler über ihre vokale Entwicklung reflektieren zu lassen:

„Also was ich mir von deiner Arbeit eigentlich erhoffe, ist auch so ein Maßstab, den die Schüler an sich selber legen können. [...] Dass sie sich selber einschätzen, und auch Ziele vornehmen, was kann ich erreichen, welchen Weg will ich eigentlich gehen, weil da habe ich eigentlich den

¹⁷⁷ Bei der Definition des Samples wurde darauf verzichtet, städtische und ländliche Schulen als Merkmal der Profilbildung in die Diskussion einzubeziehen. Tatsächlich ist aber die Schule, an der Herr Buchser (02) arbeitet, ein Brennpunkt mit tieferen Leistungsniveaus und einer anspruchsvollen Mischung von Lernenden.

Eindruck noch bei mir im Unterricht, [...] das machen sie alles mit, aber dass sie ein Ziel haben, dass sie mit der Stimme erreichen wollen, mh [...] das fehlt irgendwie noch.“ (05:477 ff.)

Frau Basler (05) zeigt sich an dieser Stelle sehr selbstkritisch¹⁷⁸ und bemängelt in ihrem Unterricht die fehlenden Reflexionsmöglichkeiten. Ihre Klassen machen zwar mit, sind aber angepasst bzw. ziellos bezüglich ihrer vokalen Entwicklung. Ihre Aussage lässt sich zwar auch als Reaktion auf die thematische Vorgabe im Leitfaden auslegen, insgesamt wird aber aus dem Zusammenhang ihres Interviews klar, dass sie dem *Verstummen* der Lernenden (05:80 f.) gerne entgegenwirken möchte. Nimmt man ihr Anliegen ernst, wird deutlich, wie wichtig die Überlegungen zu Methoden eines ganzheitlichen Assessment sein könnten. Die insgesamt marginale Ausbeute bezüglich der metakognitiven Elemente im Singunterricht zeigt aber auch, wie wenig etabliert dieses Anliegen im vorliegenden Sample ist. Ein ähnliches Bild zeigt sich bezüglich der Subkategorie *PEERFEEDBACK*, die abschließend besprochen wird.

Peerfeedback

Der Anspruch, im Musikunterricht Peerfeedback einzubauen, ergibt sich aus dem Bereich Gestaltungsprozesse¹⁷⁹ des LP21 (D-EDK, 2015), in dem Wert auf einen wertschätzenden Umgang mit eigenen und fremden Vorführungen gefordert wird. Die Verknüpfung von AAL und Peerfeedback ist nur indirekt gegeben und ergibt sich hauptsächlich aus der Idee, die Subkategorie in AAL auf der Ebene der Lernenden einzuordnen, in Abgrenzung zu Ansprüchen an Lehrkräfte, die unter AFL eingeordnet sind. Zusätzlich ist die Einteilung auch der Vermutung geschuldet, Peerfeedback werde im Unterricht dort eingesetzt, wo auch auf einer Reflexionsebene gearbeitet wird.

¹⁷⁸ Frau Basler (05) ist die einzige interviewte Person, die sich in dieser Weise selbstkritisch zeigt. Meist ist eine große Überzeugung da, der eigene Unterricht sei in Ordnung. Auftretende Schwierigkeiten werden dem Schulsystem attribuiert (11:460 ff.) oder es kommt eine gewisse Enttäuschung für das Misslingen des Singunterrichts zum Ausdruck (07:324 ff.).

¹⁷⁹ <https://v-fe.lehrplan.ch/101kkyW7RzBHMNAyyKEkkqHgFPH6DSDsB> (abgerufen am 28.01.2022).

Die Subkategorie *PEERFEEDBACK* wurde nur in drei Interviews kodiert (03, 08, 10). Das bedeutet in erster Linie, dass eine Mehrheit der Lehrkräfte Peerfeedback als Methode nicht einsetzt. Ein Abgleich mit der Subkategorie *SCHÜLER ZIEHEN SCHÜLER MIT* aus *KLASSENDYNAMISCHE PHÄNOMENE* bestätigt das Fehlen von angeleiteten Peerfeedbacks. In den kodierten Textstellen wird sowohl von positiven (z. B. 01:368 ff.) als auch negativen (z. B. 02:556 ff.) Einflüssen innerhalb der Klasse berichtet, weitere Hinweise zu Methoden mit Anklang an Peerfeedback können allerdings auch dort nicht gefunden werden. In den Kodierungen zu Peerfeedback berichten zwei Lehrkräfte von positiven Effekten in Gruppenarbeiten. Frau Fricker (03) berichtet von positiven Wirkungen, wenn sich Lernende in Gruppen gegenseitig ansingen (03:123 ff.), Herr Neuendorf (08) weist auf gegenseitige Komplimente in seinem Unterricht hin. Herr Leimbacher beschreibt eine Situation, in der sich Schülerinnen und Schüler gegenseitig motivieren, mit Mikrofon zu singen, indem sie sich „fast wie trösten, und sagen, du kannst doch, komm probier“, komm wir machen es doch zusammen und so“ (10:441 f.). In diesem Beispiel motivieren sich die Lernenden innerhalb der Klasse, höhere Ziele zu stecken und das Wagnis einzugehen, sich zu exponieren. In der positiven Atmosphäre des Unterrichts von Herrn Leimbacher (10) wird so Peerfeedback wirksam. Bereits im autoethnografischen Exkurs wurde das Phänomen einer positiven Klassendynamik durch Peerfeedback beim Mikrofonsingen beschrieben (siehe Kapitel 3.4). In keinem der genannten Beispiele können aber angeleitete Beispiele von Peerfeedback oder die Intention, einen methodischen Gewinn aus diesem zu generieren, ausgewiesen werden.

Die metakognitive Ebene AAL

Die von Earl (2013) definierten Ansprüche an *assessment as learning* treffen in der vorliegenden Untersuchung auf einen Singunterricht, der sich weitgehend an übergeordneten vokalen und persönlichkeitsentwickelnden Zielen orientiert. Die geringen Kodierhäufigkeiten weisen bereits in quantitativer Hinsicht auf eine kaum etablierte metakognitive Ebene des Singunterrichts hin (vgl. 11.2). In der qualitativen Auswertung zeigt sich nun, dass bei einigen Lehrkräften großes Gewicht auf eine lernförderliche Atmosphäre gelegt wird und in Einzelbeispielen auch Elemente von Selbsteinschätzung und Eigenverantwortung in den Unterricht einbezogen

werden. Spezifische Aufgabenstellungen oder Unterrichtssequenzen, die ausgehend von Reflexion oder Selbsteinschätzung eigenverantwortliches Lernen zum Ziel haben, konnten an keiner Stelle gefunden werden. Insgesamt zeigt sich somit zwar eine Basis für metakognitive Prozesse, aber keine konkreten und intendierten Maßnahmen zugunsten von AAL. Die Basis kann bei einigen Lehrkräften in der lernförderlichen Atmosphäre festgemacht werden. Im Einzelbeispiel zeigt sich aber, dass zusätzlich Maßnahmen, die Bloßstellung im Unterricht verhindern, bedacht werden müssen. So werden Unterrichtssequenzen durchgeführt, in denen risikolos mit der Stimme ausprobiert und gelernt werden kann. Diese für AAL grundlegende Maßnahme schließt an den konstruktivistisch ausgerichteten Begriff Fehlerermutigungsdidaktik von Oser und Spychiger (2005) an, der sich in Abgrenzung zu einer Fehlervermeidungsdidaktik bei Lehrkräften findet, die aus Angst vor Bloßstellung ihren Unterricht so durchführen, dass möglichst keine Fehler auftreten. In dieser Differenz wird gleichzeitig sichtbar, wie sich die Rolle der Lehrkraft hin zur Mentorin und zum Guide verschiebt (vgl. Tabelle 4).

Die Relevanz von Assessment im Singunterricht

In diesem Abschnitt wurde ganzheitliches Assessment in der von Earl (2013) vorgeschlagenen Einteilung *assessment for, as* und *of learning* auf das Datenmaterial der elf Interviews projiziert. Die qualitative Auswertung erfolgte in vier Schritten. Zuerst wurde die Gewichtung von Förderung und Beurteilung im Licht des *principle of valuing* (Brophy, 2019c) untersucht, um das Vorsingen als auffällige Komponente der summativen Beurteilung in der Assessment-Konzeption nach Earl (2013) zu verorten. In der anschließenden Analyse der drei Assessmentbereiche AOL, ALF und AAL wurden eine Vielzahl von individuellen Zugängen zu Teilaspekten der Assessment-Konzeption von Earl (2013) nachgewiesen, allgemeine Tendenzen, die den normativen Setzungen von Earl entsprechen, zeigten sich aber nicht. Bevor im nächsten Abschnitt im Rahmen der Beantwortung der Forschungsfragen die Erkenntnisse zu geschlechtsspezifischen Einflüssen (vgl. 13.1) in die Diskussion einbezogen werden, folgt ein Gesamtblick auf die vier Teilschritte dieses Abschnitts.

Die Vermutung Earls (2013), AOL habe üblicherweise im Unterricht das größte Gewicht, führte zuerst zur Überprüfung der Assessment-

Konzeption der elf interviewten Lehrkräfte. Durchgehend zu beobachten ist eine summative Leistungsbeurteilung mit integriertem Feedback, also eine Verbindung von AOL und AFL (Fautley, 2010), in der übergeordnete Ziele zyklisch überprüft werden. Bei der Förderung dieser Ziele zeigen sich zwei Stoßrichtungen, die bereits zuvor als Typik 1 gefasst wurden (vgl. 12.2). Bei der Gruppe von Lehrkräften, die vokale Ziele in den Vordergrund stellen, wird abgesehen von Stimmübungsritualen keine spezifische Förderung betrieben, vielmehr wird auf die natürliche Stimmentwicklung und auf die positive Wirkung von Vorbildern in der Klasse gehofft. Bei Lehrkräften, die den Fokus auf persönlichkeitsbildende Ziele setzen, können dagegen eine größere Vielfalt an Unterrichtsformen sowie ein stärkerer Fokus auf individuelle Bedürfnisse ausgemacht werden. Die Vermutung Earls ist insofern richtig, als sich die summative Beurteilung und damit AOL im Datenkorpus am deutlichsten konturiert. Unter Beachtung der *principles of assessment* (Brophy, 2019c) wurde allerdings das Kernstück dieser Konzeption, das Vorsingen, als nicht authentische Form der Leistungsüberprüfung infrage gestellt, da eine Unstimmigkeit zwischen Übungsanlage und Leistungsbeurteilung auszumachen ist.

Die Analyse des Bereichs AOL zeigte zwar, dass Transparenz als eines der Hauptkriterien von AOL den Lehrkräften wichtig erscheint und durchgehend mit mündlichen Feedbacks und teilweise mit Rastern hergestellt wird, die Leistungsbeurteilungen selbst aber kaum als valide einzuschätzen sind. So stimmt die Auswahl der Kriterien zu großen Teilen mit validierten Rastern überein (z. B. J. Hasselhorn, 2015), die konkreten Leistungsanforderungen bleiben aber meist unklar. Weiter wird als Kriterium eine summative Beurteilung aufgrund von Lernzielen verlangt, was nur in einem Beispiel explizit beschrieben werden konnte. Der Fokus liegt, wie bereits oben erwähnt, auf übergeordneten gesanglichen und persönlichkeitsbildenden Zielen, die nicht auf Teilziele heruntergebrochen und überprüft werden. Der Charakter summativer Bewertung zwecks Einteilung in Leistungsniveaus wird zudem durch eine bewusst kleine Streuung der Noten bzw. durch die Vermeidung ungenügender Noten verwässert oder die summative Bewertung wird per se abgelehnt. Als Begründung zeigt sich das Bewusstsein, dass Singen stark mit der Persönlichkeit verknüpft ist und mit ungenügenden Noten Schaden angerichtet werden könnte. Unterschiede zeigen sich zwischen Lehrpersonen, die hauptsächlich frontal unterrichten und denjenigen, die auch andere Sozialformen einbeziehen. Bei

Ersteren wird kaum auf Auftrittskompetenz geachtet und die Leistungsbeurteilung findet in Einzelsettings oder kleinen Gruppen statt. Bei Letzteren wird der Unterricht offener gestaltet und Aspekte des Auftretens werden gefördert, auch unter Einbezug der Klasse als Publikum.

Formative Prozesse, die zum Bereich von AFL gehören, finden sich in verschiedenen Bereichen im Datenkorpus. Das der summativen Leistungsbeurteilung nachgelagerte Feedback wird von den Lehrkräften neben der Plausibilisierung der Beurteilung als förderorientierte Maßnahme betrachtet, bleibt aber in seiner Wirkung ungewiss. Die Lehrkräfte weisen hauptsächlich auf verschiedene Formen von allgemeinem Feedback in der frontalen Unterrichtssituation hin. Damit kommt zum Ausdruck, dass die individualisierenden Aspekte von AFL wenig gebräuchlich sind. Einzelne Lehrkräfte sind zudem skeptisch bezüglich ihrer eigenen zur Individualisierung notwendigen Wissensbestände. Zwei Beispiele weisen allerdings darauf hin, dass individualisierendes Feedback nicht a priori utopisch ist, sondern auf der Basis lokaler Unterrichtskulturen funktionieren kann. So gibt ein Lehrer regelmäßig Individual-Feedback vor der Klasse, was aus seiner Perspektive gut akzeptiert wird, und ein anderer Lehrer adaptiert seinen Unterricht während gemeinsamer Singsequenzen begleitet von individuellen Feedbacks. Grundlegende Elemente von AFL können in der QIA aber kaum gefunden werden. So konnte weder eine Datengewinnung zwecks Adaption des Unterrichts noch der bewusste Einsatz von Assessments für diagnostische und anamnestiche Reihenuntersuchungen (Greuel & Horst, 2011; Fuchs, 2015) festgestellt werden. Die vorhandenen Ansätze in Einzelbeispielen belegen lediglich intuitiv entwickelte Lokalkulturen impliziter Individualförderung.

Die Analyse von AAL zeigt eine Diskrepanz in der Kultur von Singunterricht. Auf der einen Seite wird von Lehrkräften eine Vertrauenskultur aufgebaut, womit vielfältige Settings und vokale und persönlichkeitsentwickelnde Momente möglich werden. Auf der anderen Seite bleibt der Unterricht aus Angst vor „Bloßstellung“ in einem frontalen Setting und die Förderung im Singunterricht bleibt auf einer allgemeinen Ebene. Dieser Unterschied zeigt sich auch in der Typik 2 (vgl. 12.2), in der eine Polarität festgestellt wurde zwischen Lehrkräften, die bewusst Beziehungen und einen empathischen Umgang zu Schülerinnen und Schülern pflegen, und Lehrkräften, die eher distanziert sind und Widerstand vonseiten der Lernenden erfahren. Zudem kann in dieser Polarität an das Konzept der

Fehlerkultur (Oser & Spychiger, 2005) angeknüpft werden. Es konnte dargelegt werden, dass die Angst vor Bloßstellung zu einer Fehlervermeidungsdidaktik führt und sich von einer Fehleraufsuchsdidaktik unterscheidet, in der durch Vertrauen zwischen Lehrkräften und Lernenden eine vokale und persönliche Entwicklung ermöglicht wird. Metakognitive Phasen sind im Sample dünn gesät, denn weder reflexive noch selbstevaluative Unterrichtselemente konnten ausgemacht werden. Ein einziges Beispiel von Selbstbeurteilung wies exemplarisch auf die Möglichkeiten von AAL hin und eine Lehrerin wünscht sich Werkzeuge für diese metakognitive Ebene.

Mit der Perspektive des Assessment-Konzepts wurde ein gerichteter Blick auf die Aussagen zu Diagnose, Förderung und Beurteilung ermöglicht. Die starke Ausrichtung auf übergeordnete Ziele beim Singen lässt sich nur bedingt mit den Bereichen AOL, AFL und AAL in Verbindung bringen. Trotzdem konnten einige Tendenzen und vielfältige individuelle Konzepte analysiert werden, die eine gewisse Nähe zu den drei Bereichen zeigen. Im nächsten und letzten Abschnitt des Kapitels werden nun die Forschungsergebnisse des Bereichs D zusammengeführt.

13.3 Verknüpfung der verschiedenen Perspektiven

Im abschließenden Abschnitt des Bereichs D werden die Resultate der verschiedenen Analyseschritte aufeinander bezogen. Damit wird ein letzter Zyklus des iterativen Forschungsprozesses (Strübing et al., 2018) durchgeführt. In kurzen Rückblenden werden zentrale Bausteine der Forschung, von der Entwicklung der Fragestellung über die Durchführung der strukturierenden Inhaltsanalyse nach Mayring (2015) bis zur mehrperspektivischen Darstellung und Analyse des Forschungsmaterials nachvollzogen. Anschließend werden die Forschungsfragen beantwortet und in einem Ausblick wird überlegt, welche Desiderate sich für die Praxis und die Forschung aus der vorliegenden Arbeit ergeben.

Entwicklung der Fragestellung und inhaltliche Fokussierung

Singen hat im Musikunterricht in einigen europäischen Ländern einen zentralen Stellenwert (Schaumberger et al., 2020), so auch in der Schweiz.

Obwohl im regionalen Kontext der Studie Singen eine ungebrochene Tradition aufweist (Marty, 2021; Huber et al., 2021), werden daselbst auch Zweifel an der Qualität der vokalen Förderung laut. So bestätigt Blanchard (2019) für den Unterricht der Sekundarstufe 1 den Bereich Singen zwar als „unhintergehbare Setzung des schulischen Musikunterrichts“ (S. 287), in seinem Sample wird aber gleichzeitig deutlich, dass die Gesangsqualität nicht als Kompetenzziel angestrebt wird. Zudem werden in vielen Studien Beispiele erniedrigender Momente im Singunterricht beschrieben (z. B. Tellisch, 2015), wobei immer das solistische Vorsingen vor der Klasse in der Kritik steht. Dieses *Vorsingen* wird auch von Schaumberger (2020a) in seiner Analyse zur Leistungsbeurteilung kritisch beurteilt, erfolgt aber gerade in der Schweiz flächendeckend in den letzten Jahren der Volksschule. Mit der Fokussierung der Forschungsarbeit auf die Sekundarstufe 1 akzentuiert sich die problematische Seite der Leistungsbeurteilung durch die adoleszente Stimmentwicklung. Für die 7.–9. Klassenstufe konnten zu dieser Thematik für den deutschsprachigen Raum keine empirischen Arbeiten gefunden werden, obwohl das Singen in Teilen Deutschlands eine Renaissance erlebt (Hoos de Jokisch, 2014; Oberschmidt, 2016). Die musikpädagogische Forschung zum Singen fokussiert hauptsächlich auf den Bereich der Primarstufe (Forge & Gembris, 2012; Brunner, 2014; Henning, 2014), den Kontext der Chorarbeit (z. B. Fuchs, 2019a; Peterson, 2019; Schaumberger, 2020b) sowie auf Chorklassen (Gütay, 2012). Die Förderung und Beurteilung des Singens stellt sich somit für die Sekundarstufe 1 als Forschungsdesiderat dar, das in der vorliegenden Arbeit untersucht wird.

Die Fokussierung auf das Singen bedingt einerseits eine Klärung des Lerngegenstandes und andererseits die Betrachtung von Förder- und Beurteilungskonzepten. Bei beiden Perspektiven wurde aufgrund der Litteraturrecherche bewusst eine internationale Perspektive eingenommen, um im deutschsprachigen Raum kaum rezipierte Studien aus dem anglofonen Raum zu berücksichtigen. Bei der Untersuchung des Lerngegenstandes zeigte sich hauptsächlich die Bedeutsamkeit der Mutation als prägendes Element der Sekundarstufe 1. Sowohl für die Jungen (Freer, 2020; J. Williams et al., 2021) als auch für die Mädchen (Sweet, 2018; Gackle, 2019) wurden Studien rezipiert, welche die Mutation als biologischen Prozess beschreiben, aber auch auf die damit verbundenen entwicklungspsychologischen Veränderungen hinweisen. Neben der Beschreibung der etablierten Phasenmodelle der adoleszenten Stimmentwicklung (Cooksey,

2000; Gackle, 2000; Welch, 2019) zeigen sich insbesondere die Entwicklung einer *vocal identity* (Ashley, 2015) sowie die Beachtung der Auswirkungen der Mutation, bei Mädchen auf die emotionale Befindlichkeit (Sweet, 2018) und bei Knaben auf die Identitätsbildung (Freer, 2015), als wichtige Elemente, die bei der Förderung von Jugendlichen berücksichtigt werden müssen. Zudem wurde deutlich, dass monoedukative Unterrichtssequenzen verschiedentlich empfohlen (Kalmbach, 2010; Orton & Pitts, 2019; Schaumberger, 2020b) und auch aus der Perspektive des Genderdiskurses als adäquates Mittel angesehen werden (Hess, 2018).

Bezüglich Förderung und Beurteilung wurde einerseits das Modell *assessment for, as and of learning* (Earl, 2013; Winter, 2015) als ganzheitliches Instrument formativer und summativer Prozesse dargestellt und theoretisch an konstruktivistisch ausgerichtete Lehr- und Lernkonzepte angeschlossen (C. Schmidt, 2020). Andererseits wurde nach bestehenden Förder- und Beurteilungsinstrumenten im Bereich des Singens gesucht. Es zeigte sich, dass einige validierte Skalen (Wise & Sloboda, 2008; J. Hasselhorn, 2015; Rutkowski, 2019) zur Beurteilung von Stimmentwicklung vorhanden sind, aber vor allem der diagnostische und formative Bereiche noch wenig entwickelt wurde (Greuel & Horst, 2011; Brunner, 2014; Holcomb, 2019).

Aufgrund des ermittelten Forschungsstandes wurde abschließend eine übergeordnete Forschungsfrage formuliert, welche die Beschreibung unterschiedlicher Singkulturen und eine deskriptive Situationsanalyse ermöglicht sowie einen Bezug zu den Erkenntnissen zur adoleszenten Stimmentwicklung und zum Assessment-Konzept nach Earl (2013) als Perspektiven der Analyse erlaubt. Der Hinweis auf die Perspektive der Lehrkräfte ist dem Umstand geschuldet, für die Datengenerierung auf leitfadengestützte Interviews zu fokussieren. Dies ist mit der Hoffnung verbunden, neben vielfältigen Unterrichtsszenarien in der Analyse auch Haltungen und Überzeugungen berücksichtigen zu können (Niessen, 2017).

Etappen des Forschungsprozesses

Im Rahmen des Forschungsdesigns (vgl. 8.1) wurden für das Sampling ausgehend vom Modell professioneller Handlungskompetenz (Baumert & Kunter, 2006) unterschiedliche Profile von Musiklehrpersonen definiert, um möglichst diverse Aussagen zum Singunterricht zu erhalten. Dies

führte zu einem Sample von elf Musiklehrkräften, die aufgrund der Informationen aus den Interviews nochmals in drei neue Profile eingeteilt wurden (vgl. 9.3). Neben hauptsächlich musikalisch gebildeten Personen, z. B. eine Sängerin und ein Schlagzeuger, zeigten sich eine Gruppe von generalistisch ausgebildeten Lehrkräften sowie eine Reihe von Lehrkräften mit einem musikalisch-pädagogisch breiten Profil (vgl. Tabelle 15). Insgesamt kann das Sampling als gelungen betrachtet werden, da die inhaltliche Palette von Rückmeldungen vielfältig und divergent war. Die Profile rückten im Forschungsprozess aber zunehmend in den Hintergrund, da die Zielvorstellungen und die Beschreibung von Methoden und Unterrichtssequenzen sich nicht schlüssig mit Elementen der Ausbildung verbinden ließen. Einzig die teilweise jahrzehntelange Erfahrung scheint individuell entwickelte Umgangsweisen hervorgebracht zu haben (vgl. Individualkonzepte nach Niessen, 2006), die sich auf den aktuellen Forschungsstand, sowohl für geschlechtsspezifische Phänomene als auch für formative Förderprozesse beziehen lassen.

Die Durchführung der QIA (Mayring, 2015; Mayring & Fenzl, 2019) zeigte sich als unproblematisch. Die Entwicklung des Kategoriensystems erfolgte sowohl deduktiv als auch induktiv und wurde im Interrater-Reliabilitätstest mit überzeugenden Werten abgeschlossen. Spannend war die kritische Betrachtung des Wertes Kappa κ_n , in der aufgrund anfänglich bescheidener Werte eine Analyse der Qualitätsbeschreibungen (vgl. Tabelle 12) und der Berechnungsmethode mit MAXQDA durchgeführt wurde, um festzustellen, dass der Wert Kappa κ_n ohne Angabe einer detaillierten Methodenbeschreibung mit Vorsicht zu genießen ist. In der anschließenden Analysearbeit wurden verschiedene Werkzeuge der Inhaltsanalyse (Schreier, 2014) eingesetzt, um im Sinne eines integrierten Methodenverständnisses nach Mayring (2015, S. 53) zuerst quantitative und anschließend hauptsächlich qualitative Aspekte der Forschungsarbeit zu ergründen.

In der quantitativen Analyse konnte gezeigt werden, dass die Thematik Förderung und Beurteilung von den Interviewpartnerinnen und -partnern insgesamt als relevant erachtet wird, was sich in einer guten Verteilung der Kodierhäufigkeiten zeigte. Für die vorliegende Arbeit zeigte sich aber hauptsächlich die skalierende Strukturierung der Sozialformen (vgl. 11.3) als wichtiger Erkenntnissschritt, konnte doch später mehrfach darauf referiert werden, um inhaltliche Elemente der qualitativen Analyse zu verdeut-

lichen. So zeigte sich beispielsweise, dass gemeinsames Singen nach wie vor einen zentralen Stellenwert genießt, aber doch bei mehr als der Hälfte der Lehrkräfte mit monoedukativen Unterrichtssequenzen und Gruppenarbeiten ergänzt wird und diese Vielfalt als Gradmesser für Individualisierung und empathische Beziehungen zwischen Lehrkräften und Lernenden betrachtet werden kann.

Die Entwicklung der elf Porträts ermöglichte durch die Beschreibung von Überzeugungen und Haltungen eine individuelle Situationsanalyse und zusammenfassend die Formulierung von zwei Typiken (Kelle & Kluge, 2010), die sich als wichtige Perspektiven in der nachfolgenden thematischen Analyse entlang des Kategoriensystems zeigten. Insbesondere die Orientierung an übergeordneten Zielen erwies sich als wertvoller Baustein für das Verständnis der herausgearbeiteten Assessment-Konzeption sowie der Orientierung an wertschätzender Beziehungsarbeit für den Singunterricht.

Nach der quantitativen (vgl. Kapitel 11) und qualitativen (Kapitel 12) Annäherung an die Thematik Förderung und Beurteilung erfolgte mit der horizontalen Analyse entlang der Themen des Kategoriensystems der spezifische Blick auf die adoleszente Stimmentwicklung (vgl. 13.1) und die Elemente von AFL, AAL und AOL (vgl. 13.2). Die Erkenntnisse der beiden Abschnitte werden nun in der Beantwortung der Forschungsfragen zusammengeführt.

Beantwortung der Forschungsfragen

Die übergeordnete Forschungsfrage und die einzelnen Teilfragen geben nun die Struktur vor für einen ganzheitlichen Blick auf die Resultate dieser Arbeit¹⁸⁰.

Welche Unterrichtspraxen zeigen sich in Bezug auf Diagnose, Förderung und Beurteilung von Singunterricht aus der Perspektive der Lehrkräfte in der Sekundarstufe 1, insbesondere unter Berücksichtigung

¹⁸⁰ In der hier gewählten konzentrierten Form der Darstellung ist es unumgänglich, die in den vorangegangenen Kapiteln erarbeiteten Feinheiten in der Argumentation zugunsten klar fassbarer Phänomene komprimiert darzulegen.

individueller Bedürfnisse von Schülerinnen und Schülern und in Bezug auf geschlechtsspezifische Phänomene?

Die übergeordnete Forschungsfrage wird an dieser Stelle hauptsächlich bezüglich der ersten zwei Worte der Fragestellung besprochen. Die im weiteren Verlauf der Formulierung genannten spezifischen Themen rücken bei den nachfolgenden Teilfragen ins Zentrum. Mit der Einstiegsformulierung wird spezifisch danach gefragt, welche Situationen im Singunterricht angetroffen werden. In der QIA wurden dazu fünf verschiedene Sozialformen kodiert und in der skalierenden Strukturierung deren vier quantifiziert (vgl. 11.3). Nicht in die Skalierung einbezogen wurden die ebenfalls kodierten Situationen von Einzelbetreuung im Setting des frontalen Unterrichts. Die Resultate bilden aufgrund der gewählten Datenerhebungsmethode die Perspektive der Lehrkräfte ab und sind aufgrund des Analysevorgangs nicht als exakte Messungen, sondern als Richtwerte zu verstehen (siehe Abbildung 11, S. 245). Nachfolgend werden die fünf Situationen skizziert, um die Vermutung von Pezenburg und Dyllick (2018), vermehrt individualisierendes und eigenverantwortliches Lernen beim Singen zu beobachten, auf einer empirischen Basis zu bestätigen und inhaltlich auszu-differenzieren.

- Bei den interviewten Lehrkräften steht mit im Schnitt mehr als 70 % des Unterrichts nach wie vor das gemeinsame Singen im Zentrum ($M = 71.8\%$). Vier Lehrkräfte bleiben während 90 % des Unterrichts in der frontalen Unterrichtssituation, gestalten also abgesehen von der Phase der summativen Überprüfung den ganzen Unterricht als Chorsetting. Fünf Lehrkräfte arbeiten während 60–70 % in der ganzen Klasse und zwei Lehrkräfte während 50 % der Zeit. Die zentrale Stellung bestätigt sich auch in den hohen Kodierhäufigkeiten der Fördermethoden, die sich zu einem überwiegenden Teil direkt auf dieses Unterrichtssetting beziehen und deren Wirkung nicht unterschlagen werden darf, wie sich in der Analyse des Bereichs AFL gezeigt hat. Zudem gibt es auch einzelne Argumente gegen das Aufbrechen dieser Unterrichtsform. Zum einen wird aus Angst vor dem Bloßstellen von Schülerinnen und Schülern im Chorsetting verharret und zum anderen werden fehlendes Wissen (CK, PCK) und fehlende zeitliche Ressourcen als Hinderungsgründe für individualisierende Formen genannt. Insgesamt hat somit der monodidaktische Zugang zum Singen nach wie vor einen großen

Stellenwert, wird aber, wie die Prozentangaben zeigen, bei einigen Lehrkräften im Sinne von Pezenburg und Dyllick (2018) aufgebrochen.

- Die ebenfalls bei allen Lehrkräften anzutreffende Form von Einzel- und Gruppenbetreuung im Rahmen von 10 % des Unterrichts ist in der Regel für die summative Leistungsbeurteilung reserviert. Obwohl Feedback in diesen Unterrichtssequenzen ein wesentlicher Bestandteil ist und somit von Individualisierung ausgegangen werden könnte, wurde gezeigt, dass der formative Anteil dieses Unterrichtsbereichs aufgrund der langwierigen jährlichen Zyklen wenig Wirkung zeigt und deshalb nicht als Aufbrechen der frontalen Unterrichtssituation gewertet werden darf. Vielmehr weisen diese Unterrichtsanteile hauptsächlich auf die schweizerische Gepflogenheit des Vorsingens hin, das in den meisten Texten zum Singen in dieser Form negativ bewertet oder zumindest infrage gestellt wird (Tellisch, 2015; Schaumberger, 2020a).
- Bei fast der Hälfte der Lehrkräfte haben sich monoedukative Unterrichtsanteile etabliert. Meist werden Jungen und Mädchen separiert, um Raum für Entwicklung zu geben, oder um deren Handlungsoptionen zu erweitern (Hess, 2015). Die meisten Lehrkräfte arbeiten auch mit Aufgabenformaten in Gruppen, meist um das Vorsingen vorzubereiten, aber auch um eigenständige Abschlussarbeiten wie z. B. das Schreiben und Vortragen eigener Songs vor der Klasse zu ermöglichen. Vermehrt rückt dabei auch die Auftrittskompetenz ins Zentrum. Zudem ist bei einer Lehrkraft spezifisch die Intention für selbstverantwortliches Lernen zu sehen. Wie sich bisher gezeigt hat und später bei der Beantwortung der Teilfragen nochmals bestätigen wird, sind beide Unterrichtsformen grundlegende Bausteine für individualisierende Prozesse.
- Die nicht quantifizierte Form von Einzelbetreuung während der Arbeit mit der ganzen Klasse wird von einigen Lehrkräften sporadisch eingesetzt, um spontan den Stimmstatus einzelner Schülerinnen und Schüler zu überprüfen oder in monoedukativen Settings Einzelfeeds zu geben. Diese Unterrichtsmethode wird aber nur von einem Teil der Lehrkräfte erwähnt und kann nicht als etablierte Methode bezeichnet werden.

Die Untersuchung der auftretenden Sozialformen im Singunterricht weist zwar auf eine nach wie vor große Dominanz des gemeinsamen Singens hin, bei einer Mehrheit der untersuchten Lehrkräfte zeigen sich aber Tendenzen für Unterrichtsformate mit Gruppenarbeiten und monoedukativen

Settings. Dabei erweisen sich diese Formate als Auslöser für Prozesse, welche die in dieser Arbeit gesuchten ganzheitlichen Assessmentprozesse und die Beachtung geschlechtsspezifischer Phänomene begünstigen. Wie sich dies ausgestaltet, zeigt sich in der Beantwortung der Teilfragen.

I. Welche Ziele verfolgen Lehrkräfte in ihrem Unterricht in Bezug auf das Singen und wie kann dies auf eine lokale Schulkultur bezogen werden?

Bereits im letzten Kapitel wurde diese Teilfrage entlang der vertikalen Perspektive beantwortet (vgl. 12.3). An dieser Stelle werden diese Erkenntnisse mit der thematischen Analyse der vertikalen Achse abgeglichen und deshalb noch einmal aufgegriffen. In der Analyse der Porträts zeigten sich hauptsächlich auseinanderdriftende Zielvorstellungen, die an die Unterscheidung von Gesang und Singen nach B. Jank (2017) angeschlossen werden können. Einerseits sind vokale Ziele mit dem Produkt Gesang verbunden und andererseits wird Singen als Prozess verstanden, der als Motor die Persönlichkeitsentwicklung begünstigen soll. Das Interesse gilt nun den Bezügen dieser Polarität auf die Perspektive Assessment und die geschlechtstypischen Phänomene.

- In der Ausrichtung auf vokale Ziele ist eine Gruppe von Lehrkräften vereint, die sowohl in der Analyse von AFL als auch bei der Suche nach metakognitiven Elementen von AAL kaum Berücksichtigung fand. An den vokalen Zielen wird im Rahmen einer allgemeinen Förderung gearbeitet, individualisierende Prozesse mit formativer Unterstützung finden nicht statt und es wird auch kaum an Auftrittskompetenzen gearbeitet. Zudem zeigte sich in der Analyse der geschlechtstypischen Phänomene, dass mutierende Stimmen in dieser Gruppe zwar wahrgenommen werden, aber wenig Betreuung erhalten. Die von Kalmbach (2010) geforderte Betreuungsleistung für alle, bzw. spezifisch angepasst für Mädchen und Jungen, wird somit nicht geleistet. In dieser Gruppe ist zudem zu bemerken, dass die vokalen Ziele zwar deklariert werden, die Lehrkräfte sich selbst aber tendenziell nicht als wirksam in der Erreichung dieser Ziele beschreiben. Der Unterricht findet bei den Lehrkräften dieser Gruppe überwiegend im Chorsetting statt. Zudem wird wenig Wert auf Beziehungspflege gelegt oder sogar von Wi-

derstand berichtet. Schwierigkeiten werden auch auf die Stellung des Fachs oder die knappen zeitlichen Ressourcen projiziert. Abgesehen von der Tatsache, dass in dieser Gruppe alle ausgebildeten Sängerinnen versammelt sind, kann die Ausrichtung auf vokale Ziele nicht mit Ausbildungsprofilen in Verbindung gebracht werden. Warum gerade die Sängerinnen Mühe bekunden, vokale Ziele zu erreichen, bleibt an dieser Stelle unklar. Als Begründung könnte höchstens die von einer Sängerin eingebrachte Differenz zwischen dem eigenen professionellen Gesang und den Stimmen der Schülerinnen und Schüler dienen (11:57). Einzig der Schulmusiker mit Chorleiterausbildung (01), der eine lokale Chorkultur im Klassenunterricht pflegt, beschreibt diese Art von Unterricht als erfolgreich.

- Eine zweite Gruppe von Lehrkräften betont die mit dem Singen verbundene Persönlichkeitsentwicklung. Teilweise wird explizit auf die Verbundenheit von Person und Stimme hingewiesen (09, 10). Ihre Überzeugungen lassen sich gut an die Literatur aus dem Forschungsstand anschließen. So wird die Stimme als anthropologisches Merkmal (Neus, 2017; Sundberg, 2019), als das natürlichste Instrument (Hammar, 2016) sowie als Kommunikationsinstrument von Persönlichkeitsmerkmalen (Fuchs, 2010) gesehen. Die Zielsetzung nimmt die Ansprüche an Auftrittskompetenzen im LP21 auf, geht aber mit der ganzheitlichen Sichtweise und dem Wunsch, die Stimme als Persönlichkeitsmerkmal zu bilden, deutlich darüber hinaus. Dies zeigt sich im Unterricht in einer Fülle von methodischen Herangehensweisen, immer auf der Basis eines lernförderlichen Klimas (AAL). Ausgangspunkt ist Beziehungspflege, Vertrauen und eine Fehleraufsuchsdidaktik (Oser & Spychiger, 2005). Im Unterricht werden mit monoedukativen Arbeitsformen und Gruppenarbeiten Räume geschaffen, in denen sich Schülerinnen und Schüler entwickeln können (Freer, 2015; Hess, 2018). Insbesondere in tieferen Leistungsniveaus beschreiben die Lehrkräfte die Betonung der Persönlichkeitsentwicklung als erfolgreiche Strategie. In höheren Leistungsstufen kann über diese Betonung hinaus auch an anspruchsvolleren vokalen Zielen gearbeitet werden. Weiter ist zu beobachten, dass individuelle Stimmprobleme in dieser Gruppe bearbeitet werden, ebenfalls wird Wert auf die Beachtung von Mutationsphänomenen gelegt, z. B. wird mit mutierenden Jungen in der gerade aktuellen Tonlage gesungen. Auch in dieser Gruppe lassen sich die Ausbildungsprofile nicht

direkt mit der Zielvorstellung verbinden, interessant sind allerdings zwei Faktoren. Einerseits sind in dieser Gruppe Lehrkräfte mit einer jahrzehntelangen Erfahrung zu finden und andererseits sind spezifisch zwei Lehrkräfte dabei, die eine Schauspielausbildung absolviert haben. Langjährige Erfahrung und eine Zusatzausbildung bezüglich Auftrittskompetenz scheinen wichtige Faktoren für eine gelingende und ganzheitliche Methode im Singunterricht zu sein.

In der nochmaligen Betrachtung der ersten Teilfrage konnte aufgrund der horizontalen Analyse die These verdichtet werden, dass eine alleinige Ausrichtung auf vokale Ziele im Singunterricht meist nicht zum Erfolg führt. Vielmehr bestätigt sich die Wichtigkeit, die Stimme als Persönlichkeitsmerkmal zu betrachten, das ganzheitlich gefördert werden muss. Dies deckt sich mit aktueller Forschung für Jungen (Freer, 2020; J. Williams et al., 2021) und Mädchen (Sweet, 2018; Gackle, 2019). Dabei zeigt sich, dass vokale Ziele erreicht werden können, aber vor allem in tieferen Niveaus nachgelagert zu einer Stärkung der Person. Ins Zentrum rückt dann vor allem die Bildung von *vocal identity* (Ashley, 2015) bzw. die Berücksichtigung einer konzeptuellen Ebene, z. B. durch den Einbezug des Konzepts der *possible selves* (Markus & Nurius, 1986; Freer, 2015).

II. Welche Vorstellungen bezüglich individueller Förderung können bei Lehrkräften rekonstruiert werden und inwiefern werden dabei geschlechtsspezifische Verhaltensweisen im Unterricht wirksam?

Auch für die zweite Teilfrage wurde im letzten Kapitel eine erste Auslegeordnung vorgenommen (12.3). Insbesondere empathisches Verhalten und Beziehungspflege konnte als notwendige Grundlage für individualisierende Prozesse festgemacht werden. Distanziertheit von Lehrkräften und als Gegenstück Widerstand bei Lernenden zeigte sich vornehmlich im frontal ausgerichteten Unterricht. Bei dieser Teilfrage ergeben sich über diese atmosphärisch geprägten Elemente hinaus aus dem Kapitel 1 interessante Bezüge zu den darin untersuchten Perspektiven.

- Individualisierung ist im vorliegenden Datenkorpus kein zentrales Thema und scheint nicht als pädagogischer Konsens (M. Hasselhorn et al., 2019) im Singunterricht angekommen zu sein. Neben Argumenten gegen Individualisierung wird auch deutlich, dass deren Fehlen nicht

als Mangel angesehen wird, sondern gemeinsames Singen im Chorsetting als akzeptierter Standard gilt. Damit folgen diese Lehrkräfte allerdings nicht den Ansprüchen des LP21 (D-EDK, 2015), der neben dem gemeinsamen Singen auch solistische Kompetenzen vorsieht und insbesondere die Förderung von Auftrittskompetenz anstrebt. Auch die Lehrkräfte, die implizit individuelle Förderung praktizieren oder sogar über individuelle Ansprüche sprechen, bezeichnen dies nicht als Individualisierung.

- Das Fehlen von Individualisierung ist direkt gekoppelt mit der mehrfach beschriebenen Ausrichtung auf übergeordnete Ziele. So fehlen formative Prozesse und adaptive Unterrichtssequenzen, ausgerichtet auf individuelle Lernprozesse, fast vollständig, wie sich in der nächsten Teilfrage zeigt. Auch individuell ausgerichtete summative Beurteilungen sind nur in Einzelfällen zu sehen. Letztlich zeigen sich im Sample nur in Einzelfällen konstruktivistisch ausgerichtete Überzeugungen. Außer bei einem Lehrer, der die Selbstbeurteilung des eigenen Gesangs gegen Ende der Schulzeit verlangt, sind auch kaum kognitivistische Elemente zu finden (C. Schmidt, 2020).
- In dieser Arbeit wurde mehrfach auf die Bedürfnisse von Lernenden während der adoleszenten Stimmentwicklung verwiesen und es wurden Raster diskutiert, die eine individuelle Förderung erlauben (z. B. Rutkowski, 2019). Im vorliegenden Sample zeigen sich zwei Strategien im Umgang mit der Stimmentwicklung. Einige Lehrkräfte nehmen zwar die Phänomene der Mutation bei Mädchen und Jungen wahr, bearbeiten dies aber nicht. Sie sind auch betreffend PPS der Meinung, eine Behandlung sei nicht möglich. Zudem konnte gezeigt werden, dass einige Lehrkräfte auch Mühe bekunden, Selbstkonzepte wahrzunehmen und zu adressieren (Rieche et al., 2018). Im Gegensatz dazu reagieren einige Lehrkräfte selbstverständlich auf individuelle Stimmprobleme im Zusammenhang mit der Mutation, zelebrieren die Entwicklung sogar oder ermöglichen den Schülern wenigstens ein adäquates Training in der verfügbaren Tessitura. Explizit zur besseren Betreuung von Mädchen und Jungen werden an einigen Orten zudem monoedukative Unterrichtssequenzen durchgeführt, wie verschiedentlich vorgeschlagen (z. B. Orton & Pitts, 2019). In dieser zweiten Gruppe wird Individualisierung zwar nicht explizit genannt, aber doch auf implizite Weise in den Unterricht eingebaut.

Spezifisch bezüglich dieser Teilfrage zeigen sich die Individualkonzepte der Lehrkräfte (Niessen, 2006). Einige Lehrkräfte sehen keine Notwendigkeit, in ihrem Unterricht individualisierende Lehr- und Lernformen einzusetzen, andere individualisieren auf implizite Weise durch das Training von Auftrittskompetenz oder durch die spezifische Beachtung von geschlechtsspezifischen Phänomenen. Explizite Vorstellungen bezüglich individueller Lernvoraussetzungen können im untersuchten Sample nicht festgemacht werden, sondern zeigen sich nur implizit über den ganzheitlichen Zugang zu Stimme und Person. Dieser Bereich zeigt damit eine Divergenz zum Forschungsstand. So spielen individuelle Lernvoraussetzungen wie motivationale Dispositionen und Selbstkonzepte sowie Volition und lernbegleitende Emotionen (M. Hasselhorn & Gold, 2017, S. 66 f.) im Musikunterricht und speziell im Singunterricht aufgrund der starken Verknüpfung von Stimme und Person eine besondere Rolle. Fiedler und J. Hasselhorn (2020) fordern deshalb mit Bezug auf Spychiger:

„Inhaltlich schließen wir uns deshalb Spychiger (2015) an, die fordert, ‚dass die Förderung der musikalischen Selbstkonzepte der Schülerinnen und Schüler als zentrale Lehrkompetenz (an-)erkannt werden muss‘ (S. 69; vgl. auch Spychiger, 2017).“ (S. 25)

Das Aufbrechen von Unterrichtsformen, Beziehungspflege und das Beachten der adoleszenten Stimmentwicklung führen im untersuchten Sample zumindest auf implizite Weise zu individualisierenden Formen.

III. Von welchen diagnostischen Mitteln berichten Lehrkräfte und welchen Einfluss haben diese in der Vorbereitung und Durchführung von Unterricht und im Umgang mit Schülerinnen und Schülern?

Diese Teilfrage nimmt zuerst spezifisch Bezug auf diagnostische Elemente, deckt aber insgesamt die Bereiche AFL und AAL ab. Die Analysen des Datenkorpus in diesen beiden Bereichen haben sich als wenig ergiebig erwiesen und mussten meist auf Einzelbeispiele bezogen werden. Obwohl spezifische Diagnoseverfahren und eine darauffolgende Adaption des Unterrichts nicht zu finden waren, betonten einige Lehrkräfte die Wichtigkeit einer lernförderlichen Atmosphäre. An dieser Stelle werden die gefundenen formativen Prozesse noch einmal verortet.

- Das spezifische von AFL ist die Gewinnung von Daten zugunsten einer Optimierung der Lernwege. Dieses spezifische Element zeigte sich im vorliegenden Datenkorpus nicht. Im Rahmen einer allgemeinen Förderung wird zwar im frontalen Unterricht adaptiv reagiert, meist aber auf einer allgemeinen Ebene. Dies kann als Spezifikum des Singunterrichts im Rahmen frontaler Unterrichtssituationen bzw. als typisch und adäquat für die als Chorarbeit angelegten Phasen bezeichnet und durch eine Vielzahl an beschriebenen Fördermethoden bestätigt werden. Diese allgemeine Förderung ist im untersuchten Sample der *Courant normal* des Musikunterrichts (Mantie, 2019).
- Die Lehrkräfte führen zwar nicht spezifische formative Assessments durch, orientieren sich aber doch in den zyklisch durchgeführten Leistungsbeurteilungen über den aktuellen Entwicklungsstand ihrer Schülerinnen und Schüler. Für die summative Leistungsbewertung werden Raster verwendet, die allerdings den besprochenen diagnostischen Instrumenten nicht genügen (Fuchs, 2015; Greuel & Horst, 2011; Rutkowski, 2019). Zudem hat die Situation der Leistungsmessung wenig gemeinsam mit einer bewusst herbeigeführten Unterrichtssequenz mit Diagnose und Anamnese. Trotzdem beschreiben zumindest zwei Lehrkräfte eine formative Wirkung der in dieser Situation gegebenen Feedbacks. Mehrere Lehrkräfte führen zudem deutlich mehr Beurteilungsanlässe durch und benutzen diese Situationen, um Rückmeldungen zu geben. Diese etablierten Formen entsprechen zwar nicht den intendierten Abläufen nach Earl (2013), sind aber zumindest für das gemeinsame Singen und einen Unterricht im Chorsetting mitzudenken. Von Vorteil wäre allerdings der Einbezug von holistischen und spezifischen Rubrics (Krippner, 2016; Holcomb, 2019).
- Im Sample zeigen sich vereinzelt Momente der Diagnose, wenn vokale Probleme oder negative Selbstkonzepte auftreten. Wie die besprochenen Einzelbeispiele zeigen, handelt es sich meist um geschlechtsspezifische Phänomene der Mutation. Einige Lehrkräfte bearbeiten diese Herausforderungen, allerdings meist außerhalb des Unterrichts in Pausen und auf individueller Basis. Dabei wird allerdings intuitiv gehandelt (Brunner, 2014), der Einbezug von Skalen (z. B. J. Hasselhorn, 2015), die Entwicklungsstufen zeigen, oder Rubrics (z. B. Holcomb, 2019), die Kompetenzstufen ausweisen, wird nicht erwähnt.

- Reflexive Phasen und Selbsteinschätzungen im Sinne des Einbezugs einer metakognitiven Ebene werden im Allgemeinen nicht thematisiert. Allerdings können die Grundlagen für eine lernförderliche Atmosphäre bei einigen Lehrkräften gefunden und an die Fehlerermutigungsdidaktik von Oser und Spychiger (2005) angeschlossen werden. So entsteht der Eindruck, dass einige Lehrkräfte mit intuitiv entwickelten Methoden in ihrer lokalen Singkultur die Grundlagen für Elemente von AAL und auch AFL gelegt haben und für nächste Schritte bereit wären. Es scheint allerdings entsprechendes Knowhow (PCK) zu fehlen.

Die Assessment-Perspektive weist für den formativen Bereich, sei dies spezifisch aus der Sicht der Lehrkräfte mit adaptiven Unterrichtselementen oder aus der Sicht von Schülerinnen und Schülern bezüglich metakognitiver Elemente, eine Leerstelle bzw. einen möglichen Entwicklungsschwerpunkt auf. Einzig das Element von AAL bringt bei einer Gruppe von Lehrkräften einen Unterricht hervor, der die Grundlagen für formatives Assessment, spezifisch metakognitive Prozesse beinhaltet. Als Grund für die Leerstelle erweist sich die Bearbeitung übergeordneter Ziele und deren zyklische Überprüfung.

IV. Wie werden Leistungen von Schülerinnen und Schülern gemessen und wie wird darauf vorbereitet?

Die letzte Teilfrage dieser Arbeit bezieht sich auf den Bereich des summativen Assessments (AOL). Dieser Bereich ließ sich im Datenkorpus am deutlichsten konturieren, da alle Lehrkräfte im Musikunterricht eine Leistungsmessung durchführen müssen und dies im Bereich Singen auch regelmäßig tun. Unter Beachtung der verschiedenen Perspektiven gibt es einige Diskussionspunkte.

- Bei allen interviewten Lehrkräften findet mindestens einmal im Jahr eine Leistungsbeurteilung in Form eines Vorsingens statt. In der Analyse der Assessment-Konzeption der Lehrkräfte zeigte sich eine Differenz zu der ursprünglichen Idee eines summativen Assessments, das auf der Basis formativer Förderprozesse erfolgen soll (Scriven, 1967; Bloom et al., 1971; Earl, 2013). In der Regel wird nach dem Vorsingen auf der Grundlage der vergebenen Punkte eine Note gesetzt, gefolgt von einem mündlichen, in einigen Fällen sogar schriftlichen Feed-

back. Das mündliche Feedback dient als Erklärung und Legitimation der Beurteilung, ist dabei sowohl Mittel zur Herstellung von Transparenz (AOL) als auch Hinweis für die vokale Weiterentwicklung (AFL). Die Lernenden erhalten somit kein förderorientiertes Feedback vor der Leistungsbeurteilung, sondern erst danach. Fautley (2010) beschreibt die formative Nutzung summativer Beurteilungen zwar als taugliches Mittel, die konkreten Umstände im vorliegenden Sample weisen aber darauf hin, dass die formativen Aspekte kaum Wirkung entfalten, da der Zyklus zu lang dauert und die Rückmeldungen zu wenig spezifisch sind (Holcomb, 2019).

- In dieser zyklischen Leistungsbeurteilung findet abgesehen von einer allgemeinen Förderung im Rahmen frontaler Unterrichtssequenzen keine zielorientierte Hinführung auf die Leistungsbeurteilung statt. Meist wird ohne Vorbereitung ein Lied aus dem Repertoire vorgesungen. In wenigen Fällen wird unter Beachtung der Mutation bei Jungen Einfluss auf die Gruppenkonstellation beim Vorsingen genommen und Zeit für eine individuelle Vorbereitung gegeben, Unterrichtssequenzen mit bewusst gestalteten Trainings gibt es aber nicht. Dies entspricht nicht der Intention von AOL, das eine Leistungsbeurteilung aufgrund von Zielvorgaben vorsieht. In einem Beispiel wird die Möglichkeit skizziert, durch das Erlernen mehrerer Stimmen in einem Chorstück Punkte zu sammeln (10). Dies deckt sich mit den Wünschen von Jungen, die gerne Fertigkeiten trainieren und zeigen (Freer, 2020). Die fehlenden Zielvorgaben bzw. die Orientierung an allgemeinen Zielen ist aus diesem Blickwinkel problematisch.
- Die Anforderung an Transparenz ist durch das mündliche Feedback gegeben, die Leistungsbeurteilungen können aber nicht als valide eingestuft werden (*principle of quality*). Zwar werden die in der Literatur beschriebenen Kriterien angewandt, die Leistungsanforderungen werden aber nicht expliziert. Zudem löst die Verknüpfung von Stimme und Person bei den Lehrkräften Unbehagen aus und die Noten werden meist nur in einem eingeschränkten genügenden Bereich erteilt, teilweise sogar mit minimaler Streuung. Diese Vorgehensweise nimmt zwar die Befürchtung auf, ein Urteil über eine Person zu fällen (Lau, 2016), verhindert aber ebenso wie die fehlenden Leistungsanforderungen eine valide Benotung. Es stellt sich hier die Frage, wie die Lehrkräfte grundsätzlich zu summativer Beurteilung stehen (*principle of valuing*).

Trotzdem darf an dieser Stelle nicht vergessen werden, dass eine summativ Bewertung ein nicht zu unterschätzender extrinsischer Faktor für die Leistungsmotivation ist (Lau, 2016) und gerade beim in der Schweiz etablierten Vorsingen akzeptiert wird.

- Eine weitere Diskrepanz wird an dieser Stelle noch einmal aufgegriffen. Der Wechsel von der Chorsituation beim gemeinsamen Üben zum Vorsingen allein vor der Lehrkraft ergibt einen Bruch in der Aufgabenstellung, denn die Bedingungen in beiden Situationen lassen sich nicht vergleichen (*principle of authenticity*). Hier lohnt sich aber ein differenzierterer Blick. In wenigen Fällen ist die Kadenz dieser zyklischen Beurteilung höher und die Aufgabenstellungen sind konkreter. So gibt es Lehrkräfte, die im Verlauf der drei Schuljahre die Leistungsanforderungen steigern oder den Schülerinnen und Schülern Wahlmöglichkeiten anbieten, was zu einer höheren Authentizität der Aufgabenstellungen führt. Von Leistungsmessungen in Form von Castings, wie sie z. B. in Musiklehrmitteln vorgeschlagen werden (Detterbeck, 2011), wird nicht berichtet, vielmehr wird meist auf die Möglichkeit der Teilnahme im Schulchor verwiesen.
- Die Berücksichtigung der adoleszenten Stimmentwicklung ist marginal. Weder die Mutation der Jungen noch die Behauchtheit der Mädchen wird während der adoleszenten Stimmentwicklung in den Äußerungen zur Leistungsmessung einbezogen. Allenfalls wird im formativen Teil der Rückmeldung darauf Bezug genommen, entsprechende Ängste aufgrund der stimmlichen Herausforderungen werden aber kaum thematisiert. Obwohl einige Lehrkräfte die biologischen Merkmale spezifisch bearbeiten, wird doch die entwicklungspsychologische Seite nicht beachtet und muss als Leerstelle deklariert werden.

Insgesamt ergibt sich aus den genannten Gründen Evidenz, die Praxis der Leistungsbeurteilung zu überdenken, diese vermehrt an die Ansprüche von Lehrplänen und zu entwickelnde Kompetenzerwartungen anzugleichen und insbesondere das Training von Auftrittskompetenz einzubeziehen. Dabei sind vor allem die Leistungsanforderungen zu klären, um zu validen Leistungsmessungen zu kommen. Entsprechende Möglichkeiten werden im Bereich D skizziert.

Ausblick

Elf Lehrkräfte berichteten in der vorliegenden Interviewstudie über Methoden und Zielsetzungen bei der Begleitung Jugendlicher im Singunterricht. Die Varianzoptimierung (Patton, 2015) führte zu einem diversen Sample und ermöglichte eine umfassende Situationsanalyse bezüglich Diagnose, Förderung und Beurteilung. Dabei zeigte sich durchgehend eine Assessment-Konzeption mit einer zyklischen Bewertung, wie sie Fautley (2010) als „formative use of summative assessment“ (S. 198) beschreibt. Die ursprüngliche Konzeption von formativen und darauf folgendem summativen Assessment (Bloom et al., 1971; Scriven, 1967) und die auf metakognitive Aspekte erweiterte Perspektive von Earl (2013) zeigten sich lediglich in Einzelbeispielen und schwachen Tendenzen, wie z. B. dem Einsatz vielfältiger Sozialformen und einem empathischen Umgang mit Schülerinnen und Schülern. Obwohl ein großer Reichtum an Fördermethoden sichtbar wurde, konturierte sich im Datenkorpus hauptsächlich die summative Leistungsbeurteilung. Auch die Perspektive der adoleszenten Stimmentwicklung zeigte kein einheitliches Bild der Betreuung. So wird bei einigen Lehrkräften die Mutation bei Jungen im Unterricht berücksichtigt, bei anderen wird aus Angst vor Bloßstellung jegliche persönliche Ansprache vermieden.

Formatives Assessment müsste aus der Perspektive von Earl (2013) einen deutlich höheren Stellenwert haben. Dieser normative Anspruch bedingt allerdings ein Repertoire an förderorientierten Methoden sowie Kompetenzbeschreibungen des Singens (vgl. Krippner, 2016), die im deutschsprachigen Raum kaum diskutiert werden (Schaumberger, 2020a). Positive Einzelbeispiele und intuitiv entwickelte Methoden, wie sie Brunner (2014) ebenfalls beschrieben hat, finden sich auch in der vorliegenden Arbeit und können als Ausgangspunkt für Überlegungen zur Unterrichtsentwicklung dienen. Zudem gilt es die aktuelle Fokussierung auf formative Prozesse kritisch zu reflektieren, da auch summative Beurteilungen für Schülerinnen und Schüler eine gewisse Wertigkeit haben (Lau, 2016). Neben Fragen der Wertigkeit formativer und summativer Prozesse sind in der vorliegenden Arbeit Hinweise auf positive Wirkungen formativer Prozesse gefunden worden, diese bedürfen allerdings einer empirischen Überprüfung.

In den beiden folgenden Kapiteln werden abschließend Möglichkeiten skizziert, einerseits in der Praxis den Singunterricht weiterzuentwickeln,

indem auf konstruktivistische Lehr- und Lernmethoden ausgerichtete Konzepte postuliert werden (C. Schmidt, 2020), und andererseits in der Forschung die Relevanz der adoleszenten Stimmentwicklung zu bestätigen sowie Kompetenzbeschreibungen und entsprechende Diagnose- und Fördermethoden zu entwickeln. Dabei steht auch das in dieser Arbeit immer wieder im Fokus stehende *Vorsingen* zur Diskussion. Ziel wäre eine authentische Leistungsbeurteilung, die den *principles of assessment* (Brophy, 2019c) genügt, insbesondere die Frage der Validität von Leistungsbeurteilung und die verschiedenen Aspekte von AFL, AAL und APL aufnimmt sowie einen human-anerkennenden Singunterricht (Tellisch, 2015) ermöglicht.

E AUSBLICK

Die vorliegende Arbeit startete mit der Einordnung des Forschungsvorhabens in relevante Kontexte (vgl. Kapitel 1) und einem autoethnografischen Exkurs (vgl. Kapitel 3). Im Ausblick wird zu beiden Ausgangspunkten ein Bogen gespannt, zuerst zu drei Thesen für das Assessment im Singunterricht (vgl. Kapitel 14) und nachfolgend zu möglichen Anschlussfragen und Forschungsdesideraten (vgl. Kapitel 15). Die Trennung in zwei Kapitel erscheint zunächst künstlich, schließt aber an den schulmusikalischen Diskurs der Schweiz an. So beschreibt Huber (2016) die „Gegenüberstellung von positiv eingeschätzter Praxis und negativ bewerteter Wissenschaft“ (S. 52) als ein häufiges Argumentationsmuster im Diskurs der schweizerischen Schulmusik zwischen 1970–2010. Die Thesen stellen in diesem Sinn die Bedeutsamkeit dieser Forschungsarbeit für die Praxis heraus, legitimiert sowohl durch die Berufsbiografie des Forschenden als auch aus einer schweizerischen schulmusikalischen Optik. Das letzte Kapitel verweist abschließend auf die Wichtigkeit, weitere wissenschaftliche Erkenntnisse für den Singunterricht zu gewinnen, sei dies innerhalb der Fachdidaktik, wenn Wege gesucht werden, Diagnosefähigkeit bei Musiklehrkräften und geeignete Fördermethoden zu entwickeln, oder im Rahmen möglicher Interventionsstudien zur Überprüfung der entwickelten Thesen. So ist zu hoffen, dass sich mit dem Einbezug des erziehungswissenschaftlich ausgerichteten Assessment-Konzepts von Earl (2013) die „Ausgestaltung [des] Pädagogischen“ in dieser Arbeit gegenüber „eine[r] genuin musikalische[n] Pädagogik und Didaktik“ (Huber, 2016, S. 53) im schweizerischen schulmusikalischen Kontext behaupten kann.

14 Thesen: Assessment im Singunterricht

In diesem Kapitel werden die Erkenntnisse aus dem Forschungsstand und der QIA zu drei Thesen verdichtet, die als Empfehlungen für den Singunterricht dienen sollen. Damit wird der erste Bogen ausgehend vom autoethnografischen Exkurs zurück in die Praxis gespannt. Die erste These orientiert sich an den Erkenntnissen zur adoleszenten Stimmentwicklung sowie den Einsichten zu geschlechtsspezifischen Einflüssen auf den Singunterricht und diskutiert, was im Unterricht diesbezüglich vermehrt zu beachten wäre. Die zweite These geht vom Assessment-Modell Earls (2013) aus und revidiert das Modell insofern, als auf eine Priorisierung formativer Prozesse verzichtet wird. Die Überlegungen fassen Maßnahmen zusammen, die einen gewinnbringenden Einsatz dieses neutralen Modells versprechen. In der dritten und letzten These wird über diese Erkenntnisse hinaus die Idee eines Portfolios skizziert, um einerseits eine Alternative für das in der Schweiz gebräuchliche Vorsingen zu lancieren, aber auch um ein Modell in den Diskurs zur Leistungsbeurteilung des Singens einzubringen.

14.1 Wissen zur adoleszenten Stimmentwicklung ist ein wichtiger Faktor in der Begleitung von Jugendlichen.

Die erste These nimmt die ungenügende Rezeption anglofonen Forschungsergebnisse in der deutschsprachigen Musikpädagogik (vgl. 4.3 & 4.4) sowie die Situationsanalyse zu den geschlechtsspezifischen Einflüssen im Singunterricht (vgl. 13.1) zum Anlass, auf zentrale Herausforderungen und mögliche Methoden in der Begleitung von Jugendlichen während der adoleszenten Stimmentwicklung hinzuweisen. Die These schließt an die von Kalmbach (2010) geforderte spezifische Begleitung von Jungen und Mädchen in Chorklassen an, erweitert aber das Blickfeld zugunsten des regulären Musikunterrichts. An dieser Stelle wird nochmals auf die Analyse von Freer (2020) verwiesen, der auch aus einer internationalen Perspektive auf fehlendes Wissen (CK) speziell zu diesem Thema hinweist: „[...] others lack foundational knowledge about the physiological, psychological and sociological changes that affect adolescent boys and their singing voices.“

(S. 3) Seine Analyse gilt es auch auf Mädchen zu beziehen, da ihre Stimm-entwicklung oft unbeachtet bleibt. Die Analyse der Interviews weist auf einige Übereinstimmungen mit dem anglofonen Forschungsstand hin, entweder bestätigend durch Beobachtungen der Lehrkräfte oder in der Beschreibungen intuitiv gewachsener Methoden (Brunner, 2014). Auf dieser Grundlage erfolgen nun, fokussiert auf die erste These, einige Empfehlungen zu einer wissensbasierten Begleitung der Stimmentwicklung Jugendlicher:

- Die adoleszente Stimmentwicklung ist eine prägende Zeit für Schülerinnen und Schüler, bestimmt durch biologische, psychologische und soziale Phänomene. Es gilt ein Bewusstsein für die Bedeutsamkeit dieser Entwicklung aufzubauen und die notwendige Betreuungsleistung (Kalmbach, 2010) aktiv zu gestalten. Die in Artikeln häufig anzutreffenden Hinweise auf schlechte Erfahrungen im schulischen Musikunterricht während der Mutation – oft betreffen solche Beispiele Jungen oder Männer – weisen zusätzlich auf die Notwendigkeit einer proaktiven Begleitung durch einen human-aner kennenden Musikunterrichts hin (Tellisch, 2015).
- Beachtenswert sind im untersuchten Unterricht positive Effekte in der Betreuung dort, wo frontale Unterrichtssituationen zugunsten von Gruppenarbeiten oder monoedukativen Unterrichtssequenzen aufgebrochen werden. Einerseits sind erweiterte Handlungsoptionen (Hess, 2015) für beide Geschlechter in monoedukativen Unterrichtssequenzen zu beobachten und andererseits ermöglicht erst ein erweitertes Set an Sozialformen individuelle Zugänge, um die Stimmen adäquat diagnostizieren, fördern und beurteilen zu können (Greuel & Horst, 2011; Fuchs, 2015).
- Die Mädchen stehen in Gefahr übersehen zu werden, da ihre Stimm-entwicklung durch Behauchtheit und Schüchternheit verschleiert wird und ihre Leistungen wenig Intervention auslösen. Trotzdem sind laut dem Phasenmodell von Gackle (2000) individuell deutliche Phänomene auszumachen. Verschiedene Autoren (Sipley, 1994; Gend et al., 2017; Sweet, 2018) und auch das Beispiel mit der erstaunlich tiefen Mädchenstimme (vgl. 13.1) weisen darauf hin, wie wichtig die Vermittlung von Wissen sowie der Aufbau eines Repertoires an vokalen Übungen sind.
- Spezielle Beachtung sollte das Training des Kopfreisters erhalten. Gend et al. (2017) empfehlen, bei Mädchen in dieser Zeit trotz

Unbehagen beim Phonieren des Randregisters einen Ambitus von zwei Oktaven zu trainieren. Dies wird von einigen Lehrkräften erfolgreich gemacht, andere erleben dabei Misserfolge. Gackle (2019) empfiehlt zudem, die Mädchen nicht vorschnell in Alt und Sopran einzuteilen, sondern sie im Alter von 9–16 als *treble voice* (vgl. Fußnote 59) zu behandeln und abwechselnd in unterschiedlichen Stimmlagen singen zu lassen.

- Betreffend der adoleszenten Stimmentwicklung der Mädchen scheint es zudem wichtig, eine psychologische Betrachtungsweise einzunehmen und besonders die Labilität des Selbstkonzepts der Mädchen in Betracht zu ziehen. Diese benötigen speziell in dieser Zeit ein soziales Umfeld, in welchem sie mit ihrer Unsicherheit umgehen können. Dies betrifft auch die nach außen sicher wirkenden Sängerinnen, wie Sweet (2015) in ihrer Studie feststellt. Speziell zu beachten ist an dieser Stelle die Vorbildwirkung aktueller Stimmen aus der Popmusik, die möglicherweise aufgrund der Nichterreichbarkeit bei einigen Mädchen zu einer *non-singer identity* führen (vgl. auch Ashley, 2015).
- Bei der Begleitung von adoleszenten Jungenstimmen gibt es eine große Disparität in den Aussagen der Interviewpartnerinnen und -partner, von der Nichtbeachtung der Mutation aus Angst vor Bloßstellung bis zum Feiern des hörbaren Eintritts in die Mutation. Zwar wird bei einigen Lehrkräften der Versuch unternommen, die physiologischen Vorgänge zu erklären, es bleibt aber im besten Fall bei einer gesanglichen Betreuung. Auch bei den Jungen macht es Sinn, die Phasen nach Cooksey (2000) als Grundlage für ein bewusstes Erleben der Mutation zu verwenden. Dabei kann das Monitoring der entsprechenden stimmlichen Parameter helfen, um einen sicheren Umgang während der vokalen Einschränkungen zu ermöglichen (J. Williams et al., 2021).
- Als wohl wesentlichste Empfehlung für den Umgang mit Jungen zeigt sich die von mehreren Autoren eingebrachte Idee, die Jungen während der mittleren Phasen der Mutation in einer komfortablen Tessitura ihre sensorischen und motorischen Veränderungen konsolidieren zu lassen (Fuchs, 2020; J. Williams et al., 2021). Dies wird erfolgreich von zwei der untersuchten Lehrkräfte angewendet, benötigt aber entsprechende Fertigkeiten. Die Idee, mit dem Singen zu pausieren, kann aus der aktuellen Literatur nicht argumentiert werden, insbesondere nicht für den regulären Musikunterricht.

- Zudem ist speziell bei den Jungen die Entwicklung einer *vocal identity* (Ashley, 2015) zu fördern. Dazu soll noch einmal Freer (2020) zu Wort kommen: „boys want to be treated as the young men they are becoming.“ (S. 11) Die Betonung von Persönlichkeitsentwicklung durch das Training von Auftrittskompetenz weist auf eine mögliche Methode hin. Zudem wird auf die Vorbildwirkung von Peers verwiesen. Dies deckt sich mit der Idee Freers, das Konstrukt der *possible selves* nach Markus und Nurius (1986) als Basis für die Entwicklung eines positiven Selbstkonzepts bezüglich Singen zu verwenden. Entsprechende Methoden zur Implementation in den Unterricht müssten allerdings noch weiterentwickelt werden.
- Stimmliche Defizite werden von einigen Lehrkräften im Sample spezifisch behandelt, andere beschreiben lediglich die Defizite der *poor pitch singer* (PPS). Rutkowski (2019) beschreibt als Bedingung für korrekte Intonation das Beherrschen der Stimmregister und Dalla Valle (2007) zeigte, dass die meisten Menschen in der Lage sind, korrekt zu singen. Unter dieser Prämisse darf davon ausgegangen werden, dass die meisten stimmlichen Probleme, die unter dem Stichwort *Brummer* angesiedelt werden, mit Training gelöst werden können. Das erfolgreiche Training von PPS bedingt allerdings eine Unterrichtsgestaltung, welche die individuelle Betreuung begünstigt.
- Nicht zuletzt ist die Vorbildwirkung von Lehrkräften zu beachten. Insbesondere der Timbral-Translationsfehler (Merrill, 2020) löst möglicherweise in gegengeschlechtlichen Individualsituationen Unsicherheit aus, ist aber in seiner Wirkung nicht dramatisch. Er weist allerdings auf die Wichtigkeit von vokalen Vorbildern insgesamt hin, seien dies bei Lehrerinnen ältere Jungen in der Klasse oder dem Schulchor (vgl. Freer, 2020) oder bereits entwickelte junge Frauenstimmen bei Lehrern. Dies wiederum bedingt Kenntnisse zu individuellen Stimmen als Grundlage für Hinweise auf Vorbilder.

Harrison und Williams (2019) sind überzeugt, dass *content knowledge* zu der Stimmentwicklung bei Jungen ein wirkungsvolles Instrument für die Betreuung ist (S. 547 ff.). Dies gilt selbstverständlich ebenso für Mädchen und muss mit dem Knowhow entsprechender Methoden gekoppelt sein (PCK). Nicht vergessen werden darf allerdings das *principle of valuing* (Brophy, 2019c), das die Wichtigkeit der ersten Empfehlung nochmals aufnimmt. Solange geschlechtsspezifische Einflüsse und die adoleszente

Stimmentwicklung nicht als relevant erachtet wird, bleiben Schülerinnen und Schüler in ihrer Entwicklung auf sich selbst gestellt. In der nächsten These wird deshalb argumentiert, wie das Konzept eines ganzheitlichen Assessments (Earl, 2013) individuelle Förderprozesse begünstigen kann.

14.2 Mit der Umsetzung eines ganzheitlichen Assessment-Modells gelingt individuelle Förderung im Singunterricht.

In der zweiten These wird ein Assessment-Modell skizziert, das sowohl die Ansprüche eines ganzheitlich fördernden Singunterrichts als auch die Erfordernisse summativer Beurteilung abdeckt. Zuerst wird die Konzeption des Modells erklärt sowie der Verzicht auf die Priorisierung des formativen Assessments begründet, um nachher ausgehend von den Erkenntnissen der QIA Empfehlungen für die Bereiche AFL, AAL und AOL zu diskutieren.

Im Modell wird einerseits formatives und summatives Assessment als zusammengehöriges Paar mit unterschiedlichen Aufgaben definiert und andererseits auf die Begriffe Fremd- und Selbsteinschätzung fokussiert. Die erste Ebene ist als Rückgriff auf die ursprüngliche Idee von Scriven

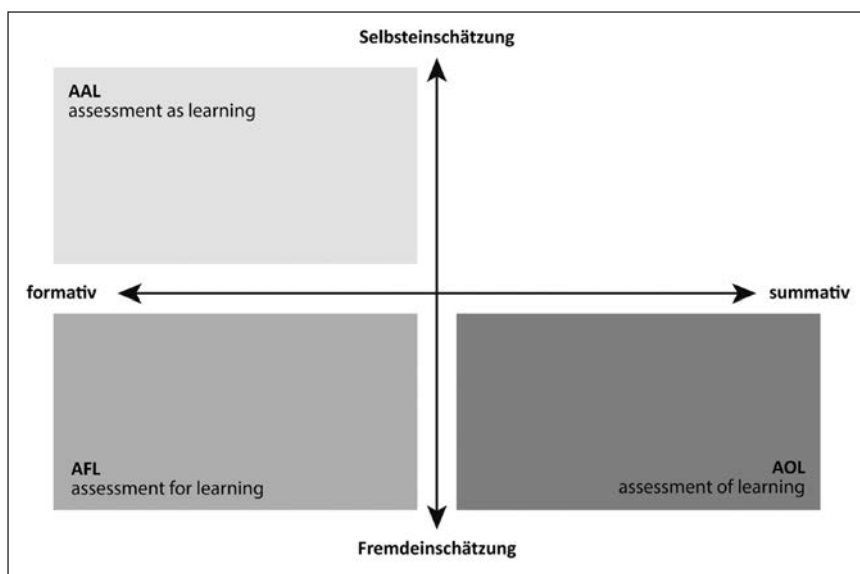


ABBILDUNG 13: Revidiertes Assessment-Modell

(1967) und Bloom et al. (1971) zu verstehen und die zweite berücksichtigt gleichermaßen die vielfältigen Rollen der Lehrkräfte (Wilson, 1996) sowie konstruktivistische und insbesondere kognitivistische Lerntheorien (C. Schmidt, 2020). Das neue Modell lässt entlang zweier Achsen eines Koordinatensystems vier Quadranten entstehen. Dazu werden auf der x-Achse die Aspekte formativ und summativ links und rechts platziert, auf der y-Achse Fremd- und Selbsteinschätzung unten und oben (vgl. Abbildung 13). Gemäß der Achsendefinitionen werden in die Quadranten die drei Assessment-Bereiche *assessment for*, *as* und *of learning* verteilt, in den beiden linken formatives Assessment, wobei AFL unten und AAL oben eingefügt wird, sowie AOL als summatives Assessment in den Quadranten rechts unten unter Fremdeinschätzung.

Im Gegensatz zu der Darstellung in Pyramidenform, in der Earl (2013) auf normative Weise einen gewünschten Zustand beschreibt oder auf den vermuteten Ist-Zustand hinweist (S. 43 f., vgl. Abbildung 5), wird im neuen Modell auf eine Priorisierung verzichtet, visualisiert durch die übereinstimmende Größe der Quadranten. Die von Black und Wiliam (1998) eingeführte und seither tradierte Begünstigung von formativem Assessment (Lau, 2016) wird so zugunsten einer Fokussierung auf die doppelte Aufgabe der Lehrkräfte aufgegeben. Weder ihre Rolle als individuelle Förderer und Lernbegleiter noch ihre Pflicht als Entscheidungsträger schulischer Laufbahnentscheide wird so durch eine normative Setzung überhöht oder marginalisiert. Auch der Bereich AAL wird im vorgeschlagenen Modell gleichwertig in einen Quadranten eingeordnet und erhält kein übergroßes Gewicht wie im Modell von Earl (2013). So werden Lehrende in der Steuerung von Lernprozessen (AFL) und Lernende in ihrem individuellen Lernprozess als ebenbürtige Partnerinnen und Partner definiert und gleichzeitig im Lehr-Lern-Prozess aufeinander bezogen.

Das vorgeschlagene Modell lässt im rechten oberen Quadranten eine Lücke entstehen, die summative Beurteilung und Selbsteinschätzung vereint. Dieser Bereich lädt ein, darüber nachzudenken, wie summative Beurteilung unter Einbezug der Selbsteinschätzung aussehen könnte. Die Idee, mit Portfolios zu arbeiten, liegt dabei auf der Hand, sei dies, um den Entwicklungsprozess abzubilden, Leistungen zu dokumentieren, über die eigene Leistung zu reflektieren oder im Rahmen digitaler Erweiterungen des Singunterrichts (*self-evaluating tools*) Fremdeinschätzung einzubeziehen. Diese Idee wird im letzten Abschnitt dieses Kapitels

(vgl. 14.3) weitergesponnen. Die Darstellung der Empfehlungen folgt der von Earl (2013) geprägten Reihenfolge *assessment for, as und of learning*.

Assessment for Learning

In der vorliegenden Arbeit wurden große Unterschiede in der Zielorientierung und der damit verbundenen Förderung sichtbar. Obwohl in der QIA die Subkategorie *ASSESSMENT FOR LEARNING* sehr oft kodiert wurde, zeigten sich bezüglich der Kernkomponenten von AFL wenig konkrete Beispiele. Insbesondere sind differenzierte Lehrstrategien und Lernangebote, um einzelne Schülerinnen und Schüler in ihrem Lernen zu unterstützen, nur in Einzelfällen zu beobachten und individuelles Feedback wird meist nicht zugunsten einer Verbesserung vor einer summativen Leistung gegeben, sondern im Anschluss an summative Beurteilung. Im Vordergrund stehen übergeordnete Ziele, die bei einigen Lehrkräften erkennbar zu individuellem Lernzuwachs bei Schülerinnen und Schülern führen, bei anderen aber lediglich zu einem gemeinsamen Singen ohne vokale Kompetenzentwicklung. Insgesamt scheint im untersuchten Sample weder die ursprüngliche Idee von Assessment (Scriven, 1967; Bloom et al., 1971) noch die in der musikpädagogischen Literatur für den Singunterricht vorgeschlagene individuelle Betreuung (Rutkowski, 2019; Fuchs, 2020) von Bedeutung zu sein. Dieses Desiderat führt zu drei Empfehlungen für die Entwicklung von AFL im Singunterricht.

- An mehreren Stellen dieser Arbeit wurde der positive Einfluss eines diversen Einsatzes von Sozialformen deutlich. So steht das gemeinsame Singen zwar bei allen Lehrkräften im Zentrum (mind. 50 % des Unterrichts, vgl. 11.3), wird aber bei einigen Lehrkräften durch Gruppenarbeiten und monoedukative Sequenzen ergänzt. Das Aufbrechen der frontalen Unterrichtssituation zeigt sich in den vorgelegten Beispielen als Türöffner für individuelle Zugänge zu Schülerinnen und Schülern. Dies deckt sich mit den von Hess (2015, 2018) spezifisch für den performanceorientierten Musikunterricht vermuteten verbesserten Handlungsoptionen in monoedukativen Settings, aber auch mit den Überlegungen zum Umgang mit Heterogenität (Budde et al., 2016, S. 37; Linn, 2017). In diesem Sinn wird empfohlen, Aufgabenstellungen zu entwickeln, die einen vielfältigen Einsatz von Sozialformen ermöglichen, seien dies Gruppenperformances oder Karaoke-Shows, reflexive

Partnerarbeiten oder auch Trainingseinheiten für mutierende Jungenstimmen.

- Die Interviewpartnerinnen und -partner berichteten hauptsächlich von der Erfassung individueller Lernstände im Rahmen formativer Feedbacks im Anschluss an summative Beurteilungen. Diagnostische Assessments mit einer nachfolgenden Adaption des Unterrichts (vgl. Tabelle 5) werden zwar in der Literatur gefordert (Greuel & Horst, 2011; Fuchs, 2015), konnten aber in den Interviews nicht eruiert werden. Rieche et al. (2018) weisen zudem auf die fehlende Wahrnehmungsfähigkeit von Selbstkonzepten durch die Lehrkräfte hin. Unabdingbar scheint somit die Erweiterung des Repertoires an diagnostischen Methoden (Brunner, 2014). Einige der interviewten Lehrkräfte stehen allerdings individualisierenden Formen kritisch gegenüber, z. B. aus Angst vor bloßstellenden Momenten im Unterricht. Diese Fehlervermeidungsdidaktik (Oser & Spychiger, 2005) weist darauf hin, dass die Revision biografisch geprägter Selbsttheorien und Überzeugungen in Betracht gezogen werden müsste (Wahl, 2013). Zudem sind geeignete Instrumente zur Diagnose oder Beurteilung (z. B. Holcomb, 2019) im deutschsprachigen Raum noch zu entwickeln (Schaumberger, 2020a).
- Der Anspruch an Individualisierung, der im ersten Teil der Arbeit als allgemeiner *pädagogischer Konsens* herausgearbeitet wurde (vgl. 5.1), steht in einem Konflikt um zeitliche Ressourcen, was Linn (2017) mit seiner Untersuchung zur Wahrnehmung von Heterogenität herausgearbeitet hat. Er zeigt, dass Heterogenität zwar durchgehend wahrgenommen wird, den Lehrkräften aber keine Zeit bleibt, darauf zu reagieren. Es wird lediglich der Hoffnung Ausdruck gegeben, beim gemeinsamen Singen würden sich durch die natürliche Entwicklung die Probleme von selbst lösen. Es sei an dieser Stelle an die Lehrkraft erinnert, die froh ist, keine Zeit für die Mädchen aufwenden zu müssen (vgl. 13.1). Der Konflikt um zeitliche Ressourcen wird weiter zugespitzt, wenn die beim gemeinsamen Singen zu beobachtenden positiven prosozialen und emotionsregulierenden Effekte (Gembris, 2015) bedacht werden. Glücklicherweise zeigen verschiedene Beispiele aus dem Sample (08, 09, 10), dass eine Synthese aus individualisierenden und gemeinschaftlichen Elementen möglich ist. Wichtig erscheint bei dieser Empfehlung, fehlende zeitliche Ressourcen nicht als Argument gegen den Anspruch

einer geforderten Betreuungsleistung (Kalmbach, 2010) auszuspielen, sondern nach kreativen Lösungen zu suchen.

Der Bereich AFL benötigt somit als Grundlage eine Unterrichtsstruktur mit vielfältigen Sozialformen, um Gestaltungsspielraum für individuelle Förderung zu erhalten. Damit dies gelingt, braucht es Assessment-Methoden, die einfach und verständlich in den Unterricht integriert werden können, um dem *principle of operability* (Brophy, 2019c, S. 922) zu genügen. Ansätze dazu gibt es bereits in einzelnen Lehrmitteln (Hoene & Thurman, 2011; Bolender & Müller, 2012a), ein ganzheitliches Setting, das auch die adoleszente Stimmentwicklung einbezieht, müsste noch entwickelt werden (vgl. 15.2).

Assessment as Learning

Der zweite formative Bereich betrifft die metakognitiven Prozesse von AAL. In der Analyse in Kapitel 13.2 wurde vor allem die Wichtigkeit einer lernförderlichen Atmosphäre als Grundlage von Selbstevaluation und Reflexion herausgearbeitet und mit einigen gelingenden Beispielen aus dem Datenkorpus belegt. Im Gegensatz dazu konnte nur in Einzelfällen auf konkrete Beispiele metakognitiver Praxen zurückgegriffen werden. In der Auswertung der Porträts wurde zudem deutlich, wie wichtig Beziehungsgestaltung für die Vertrauensbildung im Singunterricht ist (vgl. 12.2, Typik 2). Zusätzlich spielen im Singunterricht aufgrund der starken Verknüpfung von Stimme und Person individuelle Lernvoraussetzungen wie motivationale Dispositionen und Selbstkonzepte sowie Volition und lernbegleitende Emotionen eine besondere Rolle (M. Hasselhorn & Gold, 2017, S. 66 f.). Sichtbar wird dies vor allem im vokalen Selbstkonzept bzw. der *vocal identity*, die während der adoleszenten Stimmentwicklung zusätzlichen Turbulenzen ausgesetzt ist (Freer, 2020; Sweet, 2018). Die lernförderliche Atmosphäre ist deshalb nur ein Pfeiler für den Bereich AAL. Als zweiter Pfeiler zeigt sich der Umgang mit Selbstkonzepten, der nicht selbstverständlich gelingt, wie Rieche et al. (2018) gezeigt haben. Die Voraussetzungen, Schülerinnen und Schüler zu einem selbstbestimmten Beschreiten von Lernwegen, zu Selbstevaluation und Reflexion zu führen, sind vielfältig und in der Implementierung anspruchsvoll.

- Die in Kapitel 12 erstellten Porträts und deren Analyse haben gezeigt, wie ein auf persönlichkeitsentwickelnde Ziele ausgerichteter Singun-

terricht insgesamt und speziell in tieferen Niveaustufen erfolgsversprechend ist. *Eine Stimme zu haben* bleibt dann nicht eine musikalische Kompetenz, sondern wird zum Zeichen von Persönlichkeit. Da eine direkte Steuerung von AAL durch die Lehrkräfte anspruchsvoll ist, kann als erste Empfehlung der Hinweis erfolgen, das Singen nicht nur auf sein musikalisches Produkt, also den Gesang zu reduzieren, sondern multiperspektivische Zugänge zu wählen. Speziell der beschriebene Zugang zur Stimme als Persönlichkeitsmerkmal scheint besonders begünstigend zu sein.

- In der Studie von Fiedler und J. Hasselhorn (2020) wird vorgeschlagen, die Förderung von Selbstkonzepten als Lehrinhalt in die Ausbildung für das Fach Musik einzubauen. Dass Lehrkräfte Mühe bekunden, Selbstkonzepte wahrzunehmen, bestätigen Rieche et al. (2018). Gerade für die metakognitive Ebene scheint es zentral, negative wie positive Selbstkonzepte wahrnehmen zu können. Empfohlen wird deshalb, Urteile über einzelne Lernende nicht vorschnell und aufgrund vordergründiger Merkmale zu bilden, sondern Befragungsinstrumente zu entwickeln, die Selbstkonzepte aufdecken. Allenfalls lohnt sich dabei eine Adaption von bestehenden Forschungsinstrumenten (Harnischmacher & Hörtzsch, 2012; Spychiger, 2013; Carmichael & Harnischmacher, 2015; Fiedler, 2016; Fiedler & J. Hasselhorn, 2020). Deren Resultate erleichtern dann Feedback zwischen Peers oder Lehrkräften und Lernenden.
- Earl und Katz (2006) fordern die Lehrkräfte auf, durch regelmäßiges Üben die Schülerinnen und Schülern zu befähigen, selbstbewusste, kompetente Selbsteinschätzer zu werden (vgl. Tabelle 5), ähnlich wie die Forschungsgruppe um Hattie, die aufbauend auf das Konzept von *visible learning* (Hattie, 2009) ein neues Konzept empfiehlt, das Lernende zu „assessment-capable visible learners“ (N. Frey et al., 2018, S. 7) macht. Earl (2013) verweist diesbezüglich auch auf die Vorbildwirkung der Lehrkräfte als Modellierende metakognitiver Prozesse. So erwartet er von Lehrkräften, die Fähigkeit zur Selbsteinschätzung vorzuleben und zu vermitteln (vgl. Tabelle 5). Im Bereich der Musikpädagogik verweist z. B. Holcomb (2019) bei den „Model Cornerstone Assessments (MCAs)“ (S. 385) auf eine Selbstevaluation. In der vorliegenden Arbeit berichtet Herr Neuendorf (08) von einem Aushandlungsprozess, der auf eine Selbstbeurteilung der Lernenden aufbaut.

Herr Buchser (02) setzt auf Aufgabenformate, die neben vokalen Anteilen Wert auf die Entwicklung von Reflexionsfähigkeit legt. Für den Bereich AAL wird deshalb empfohlen, Schülerinnen und Schülern darin zu unterstützen, ihre eigene Entwicklung kritisch zu reflektieren und auch ihren Peers wertschätzendes Feedback zu geben.

Der Aufbau metakognitiver Fähigkeiten beruht einerseits auf einem guten Unterrichtsklima sowie einer guten Beziehung zwischen Lernenden und Lehrkräften, ist andererseits aber auch auf deren Fähigkeit, Selbstkonzepte einzuschätzen, angewiesen. Erst auf dieser Basis können Schülerinnen und Schüler zu assessment-fähigen Lernern werden.

Assessment of Learning

Für den Bereich des summativen Assessments ergeben sich im Anschluss an den Forschungsstand und die Forschungsergebnisse drei Empfehlungen. Im Fokus steht dabei die Validität von Leistungsbeurteilungen, die Konzeption der Leistungsbeurteilung und die Authentizität der Aufgabenstellungen. Ausgangspunkt der Empfehlungen ist die Diskrepanz zwischen den beschriebenen validierten Tests (J. Hasselhorn, 2015; Rutkowski, 2019) und der intuitiv entwickelten Raster der Lehrkräfte, die zwar die Beurteilung für die Lernenden transparent machen, aber kaum Auskunft geben über den Leistungsstand, da meist mit einem Punktesystem gearbeitet wird.

- In der ersten Empfehlung wird auf die fehlende Validität der Leistungsbeurteilung fokussiert. Brophy (2019c) weist auf ein allgemeines Desiderat bezüglich Validität und Reliabilität von Assessments hin (S. 914). Obwohl es einige validierte Beurteilungsskalen für das Singen gibt (vgl. 5.2, 5.3) und auch im deutschsprachigen Raum eine Skala validiert wurde (J. Hasselhorn, 2015), werden in der Praxis meist selbstentwickelte Instrumente benutzt (Schaumberger, 2020a), deren Validität insgesamt infrage gestellt werden muss. Der Vorschlag lautet deshalb, auf Noten für das Vorsingen zu verzichten, oder zumindest *Rubrics* zu entwickeln, die einen multiperspektivischen Einblick in den Entwicklungsstand gewähren, analog zu den *progress cards* englischer *choristers* (Krippner, 2016). Insbesondere in der zyklischen Art der Leistungsbeurteilung, wie sie die interviewten Lehrkräfte durchführen, wäre dies eine bedenkenswerte Alternative.

- Im Assessment-Modell von Earl (2013) wird für AOL eine Zielorientierung mit klaren Bewertungskriterien eingefordert. Den Schülerinnen und Schülern soll im Unterricht klar sein, auf was sie hinarbeiten und welche Inhalte bewertet werden. In der Untersuchung zeigte sich nur ein Beispiel einer klar definierten Aufgabenstellung, die im Unterricht bearbeitet und anschließend summativ bewertet wurde. Es wird deshalb empfohlen, vermehrt konkrete Aufgabenstellungen zu entwickeln, die im Anschluss zu einer Bewertung führen. Als durchaus gewünschter Nebeneffekt wird auch die Frage der Validität entschärft, da die Beurteilungskriterien sich auf natürliche Weise konkretisieren. Dies käme auch der Arbeit mit Jungen entgegen, da diese selbst konkretes Erlernen von Fertigkeiten weniger zielgerichtetem gemeinsamem Singen vorziehen (Freer, 2020).
- Lau (2016) weist in seiner Untersuchung zum Zusammenspiel von formativem und summativem Assessment auf die extrinsisch treibende und nicht zu unterschätzende Kraft summativer Beurteilung hin. In der Annahme, eine valide Leistungsbeurteilung (Holcomb, 2019) in einem human-aner kennenden Unterricht (Tellisch, 2015) sei möglich, kann ein Gleichgewicht zwischen intrinsischer und extrinsischer Motivation gesucht werden. Dazu braucht es im Unterricht allerdings einen Spannungsbogen vom Üben bis zur Bewertung mit einem Fokus auf authentischen Lerngelegenheiten. Als letzte Empfehlung in der zweiten These wird deshalb gefordert, ein aufbauendes Set von Aufgaben zu entwickeln, das die verschiedenen Aspekte von Performance abdeckt und von Schülerinnen und Schülern in einem Spannungsbogen erlebt werden können. Als Beispiel diene wiederum die von Holcomb (2019) beschriebenen „Model Cornerstone Assessments (MCAs)“ (S. 385).

Die Empfehlungen für ein ganzheitliches Assessment in den Bereichen AFL, AAL und AOL sind divers. Neben möglichen Türöffnern wurden auch normative Setzungen besprochen, die das Überdenken von Haltungen und Überzeugungen notwendig machen. Zudem haben diese Empfehlungen teilweise einen hypothetischen Charakter und sind auf zusätzliche Forschung angewiesen, einerseits um ein Set von Assessment-Methoden zu entwickeln und diese konzeptionell zu verbinden (vgl. 15.2), und andererseits, um deren Wirksamkeit kritisch zu überprüfen (vgl. 15.3). In der nächsten These wird ein holistisches Beurteilungskonzept mit Fokus auf die Selbst- und Fremdbeurteilung vorgelegt.

14.3 Formative und summative Förder- und Beurteilungsprozesse können mit einem Stimmportfolio adäquat abgebildet werden.

In der dritten These steht das Stimmportfolio als verbindendes Element der beschriebenen Assessment-Prozesse im Zentrum. Eingesetzt im vierten Quadranten tanzt der Begriff *Portfolio* aus der Reihe, denn logisch wäre eine Erweiterung der Begrifflichkeiten AFL, AAL und AOL mit einer vierten Begrifflichkeit (vgl. Abbildung 14). Es findet sich allerdings keine adäquate Präposition, die nicht schon besetzt ist (vgl. Fußnote 2) oder das Spezifische der Kombination summativer Leistungsbeurteilung und Selbsteinschätzung auszudrücken vermag. Es wird deshalb in dieser These zuerst der Einbezug von Portfolios in das Modell argumentiert, zusätzlich auf die Durchlässigkeit der vier Quadranten verwiesen, um anschließend spezifisch für den Singunterricht zu skizzieren, wie ein Stimmportfolio konkret aussehen könnte.

Der Einbezug von Portfolios zur Abbildung von Lernprozessen wird seit Längerem diskutiert. Wurde zuerst vor allem nach alternativen Möglichkeiten für die Beurteilung gesucht (Häcker, 2005, 2011), steht aktuell

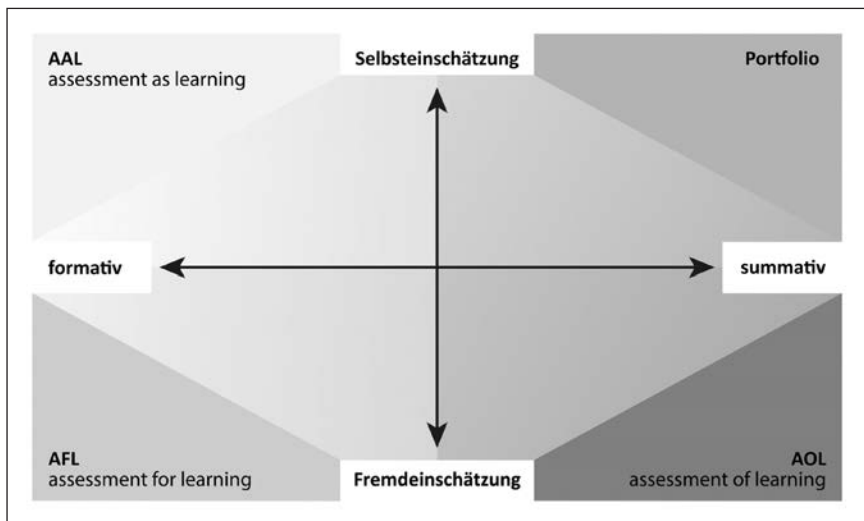


ABBILDUNG 14: Durchlässiges Assessment-Modell als Basis eines Prozess- und Produktportfolios

die Optimierung von Lehr-, Lern- und Entwicklungsprozessen im Zentrum. Ziegelbauer und Gläser-Zikuda (2016) schreiben: „Auf der Unterrichts- und Schulebene wird mit dem Einsatz von Portfolios vor allem das Ziel der Optimierung individueller Lernentwicklungen, aber auch die Unterstützung von Unterrichts- und Schulentwicklungsprozessen verfolgt.“ (S. 9) Diese zweifache Zielsetzung ist auch Grundlage der dritten These. Einerseits soll die Arbeit mit Portfolios Schülerinnen und Schülern die Möglichkeit geben, die eigene Stimmentwicklung durch die Dokumentation im Portfolio bewusster zu erleben und andererseits wird im normativen Sinn eine Unterrichtsentwicklung angestoßen, welche die problematische Seite des Vorsingens überwinden soll. Zur Umsetzung von Portfolios gibt es mittlerweile reichhaltiges Material, das die verschiedenen Formen von Portfolios und die Methoden zur Umsetzung darlegen (Winter, 2015; Keller & König, 2017). Zudem wurde auch für den vokalen Bereich die Idee von Portfolioarbeit bereits erprobt (Morrison, 2005) und findet aktuell Berücksichtigung bei Dyllick (2019).

Die Einordnung von Portfolios in den rechten oberen Quadranten (vgl. Abbildung 14) verortet diese nicht nur in den Bereich der Selbstbewertung, sondern zieht gleichzeitig in Betracht, sie als summatives Bewertungsinstrument gelten zu lassen. Im ersten Moment scheint eine summative Beurteilung in der Verantwortung von Schülerinnen und Schülern undenkbar, sind es doch externe Maßstäbe, denen die summative Beurteilung genügen soll. Zwei Argumente weisen aber auf einen Spielraum hin, der ausgenutzt werden sollte. Das erste Argument bezieht sich auf die Qualitätsansprüche der Leistungsbeurteilung und die unterschiedlichen Vorgehensweisen der Lehrkräfte. Das zweite Argument fokussiert auf die Interpretationsmöglichkeiten der gesetzlich vorgegebenen Rahmenbedingungen.

In der vorliegenden Untersuchung plädieren die Lehrkräfte meist für Noten im genügenden Bereich, geben gar nur Höchstnoten oder sind dezidiert dagegen, für Leistungen im Singunterricht eine Note zu setzen. Diesbezüglich stellte sich in der Untersuchung die Frage, inwiefern von einer validen Beurteilung gesprochen werden kann. Solange Lehrkräfte nur einen sehr eingeschränkten Bereich der Notenskala bespielen, wird die Leistungsbeurteilung per se infrage gestellt und verlieren die Leistungsunterschiede zwischen Schülerinnen und Schülern an Bedeutung. Wird nun der Anspruch an Validität zugunsten einer Selbstbeurteilung aufgegeben, bleibt nur zu klären, ob die Differenz zwischen den von Lehrkräften

gesetzten Noten und eigenverantwortlichen Selbstbeurteilungen sich als relevant erweist. Das Beispiel von Herrn Neuendorf (08) zeigt, dass sogar mit einer eher strengen Selbstbeurteilung durch die Lernenden zu rechnen ist. Unter der Voraussetzung einer etablierten Selbstbeurteilungskultur wäre der Verzicht auf eine kaum differenzierende Fremdbeurteilung zu verkraften.

Ein weitaus entscheidender Grund für eine Verschiebung der Beurteilungshoheit zugunsten der Schülerinnen und Schülern zeigt sich in der Synopse von Beurteilungsvorgaben im Bildungsraum Nordwestschweiz (Institut Weiterbildung und Beratung PH FHNW, Arbeitsgruppe Schwerpunkt Unterrichtsentwicklung, 2016). Die Autoren zeigen auf, wie groß der Spielraum in der Gestaltung zeugnisrelevanter Noten ist. In keinem der vier Kantone wird auf einen exakten Schnitt Wert gelegt, vielmehr liegt es im Ermessen der Lehrkräfte, aus den Beurteilungsbelegen eine belastbare Zeugnisnote herzustellen. Dieser Spielraum zeigt sich auch in dem von Lehmann-Wermser (2019) eingebrachten Beispiel eines Semesterberichts in Form eines Kompetenzrasters, das auf eine Beurteilung in Zahlen verzichtet, stattdessen Fertigkeiten in Worte fasst und Kommentare ermöglicht (S. 241). In dieser These wird deshalb vorgeschlagen, Kompetenzraster oder *Rubrics* sowie eine Sammlung von Belegen des eigenen Lernwegs in einem Stimmportfolio zu vereinigen und wenn gefordert, als Grundlage für eine zeugnisrelevante Note zu verwenden.

Als weitere konzeptionelle Idee des Modells wird vorgeschlagen, die ausdrückliche Unterscheidung von formativen und summativen Assessments aufzugeben. Im *principle of purpose* (Brophy 2019, S. 921) wird zwar gefordert, den Zweck eines Assessments klar zu fokussieren, Earl (2013) verweist aber auch auf die sich teilweise ergänzenden oder auch widersprechenden Rollen von Assessment (S. 36). In der vorliegenden Arbeit zeigt sich die Durchmischung in der Assessment-Konzeption der Interviewpartnerinnen und -partner durch die Kombination der summativen Leistungsbeurteilung mit einem integrierten formativen Feedback. Durch das Feedback wird einerseits Transparenz für die summative Beurteilung hergestellt, aber durch einen formativen Teil auch die zukünftige Entwicklung angeregt. Dieses auch von Fautley (2010) beschriebene Verfahren weist auf die Möglichkeit einer mehrfachen Absicht von Assessments hin, die nicht ausgeschlossen werden sollte. Im vorgeschlagenen Stimmportfolio könnten Belege einer ursprünglich formativen Einschätzung im

Rahmen einer holistischen Betrachtung und unter Berücksichtigung der Reflexion der Lernenden zu einer summativen Beurteilung führen. Dies bedingt allerdings die aktive Gestaltung des Bereichs AAL und das Hinführen der Lernenden zu Selbst-Assessment (N. Frey et al., 2018). Im Modell selbst wird die Idee der Durchlässigkeit der vier Quadranten durch Übergänge zwischen Graustufen dargestellt (vgl. Abbildung 14).

Für die konkrete Umsetzung eines Stimmportfolios werden nun Themen und Inhalte skizziert, die ein selbstbestimmtes Handeln von Schülerinnen und Schülern ermöglichen und es insgesamt zulassen, aus den gesammelten Belegen eine Zeugnisnote zu bilden. Die Sammlung erhebt keinen Anspruch auf Vollständigkeit, sondern vereint einige aus Sicht des Forschenden zentrale Bausteine.

- Holcomb (2019) skizziert ein holistisches Rubric, verweist aber darauf, dass vertieftes Feedback und spezifischer Entwicklungsbedarf mit detaillierteren Rubrics ausgewiesen werden müssten. Krippner (2016) bringt dazu sogenannte *progress cards* ins Spiel, die englische *choristers* in ihrer Entwicklung begleiten. An dieser Stelle wird auch noch einmal auf das Lehrmittel von Bolender und Müller (2012b) verwiesen, das in Teilbereichen die Visualisierung der Kompetenzentwicklung vorschlägt. Die Entwicklung einer holistischen Entwicklungsmatrix und entsprechender Materialien, bezogen auf die ganze Breite der adoleszenten Stimmentwicklung und ausgerichtet auf den Musikunterricht in der Sekundarstufe 1, zeigt sich als erster und grundlegender Schritt und wird auch von Schaumberger (2020a) angeregt.
- Neben einer inhaltlichen Strukturierung gilt es auch die Perspektive von Schülerinnen und Schülern einzunehmen und zu überlegen, welche Elemente sie gerne in einem Portfolio ablegen möchten. Möglich wäre ähnlich einer Skizzenmappe im Bildnerischen Gestalten die Sammlung von selbst gesungenen Songs und Liedern, seien dies Gruppenarbeiten oder Soloauftritte. Freer (2020) verweist z. B. auf den Wunsch von Jungen, Fertigkeiten zu trainieren, was auf die Möglichkeit hinweist, auf konkrete Ziele hinarbeiten und Zertifikate zu erhalten. Dabei sollte auch der Einsatz digitaler Tools in die Überlegungen einbezogen werden, einerseits um die Lehrkräfte zu entlasten, andererseits aber auch um die Vorteile von individuellem und sofortigem Feedback einzubeziehen (vgl. Imthurn, 2020).

- Bereits in der ersten These wurde das Wissen um die adoleszente Stimmentwicklung betont. Für das Portfolio wird deshalb vorgeschlagen, Aufgaben zu entwickeln, die einerseits Wissen zur Mutation und andererseits Instrumente zur Selbstbeobachtung beinhalten. Mädchen könnten z. B. ihre Tessitura regelmäßig überprüfen und mit dem Phasenmodell nach Gackle (2000) abgleichen, um den eigenen Stimmklang und die auftauchenden Registerbrüche zu verstehen. Jungen könnten ihre Sprechstimmlage beobachten, um selbstständig zu wissen, wie sich ihre Tessitura nach unten verschiebt, um die adäquate Tonlage zum Singen zu kennen. Die Beobachtungen der adoleszenten Stimmentwicklungen können helfen, die sensorischen und motorischen Anpassungsleistungen und die emotionalen Begleitumstände besser zu verstehen.
- Die letzte Empfehlung nimmt noch einmal Bezug auf die ambivalente Aufgabe der Lehrkräfte. Sie sollen Förderer sein (AFL), haben aber auch Entscheidungsgewalt für die Zukunft der Lernenden (AOL). Im Wechsel von formativer und summativer Beurteilung liegt eine nicht aufzulösende Spannung, die zusätzlich durch die Verknüpfung vokaler und emotionaler Aspekte aufgeladen wird. Aus der Perspektive dieser mehrfachen Spannungen wird das Vorsingen als akzeptierte Tradition der summativen Beurteilung infrage gestellt. Die Empfehlung lautet deshalb, nicht punktuell und nur aus Sicht der Lehrkräfte Entscheidungen zu fällen, sondern das Vorsingen zugunsten eines Stimmportfolios aufzugeben und konsensual aufgrund der entwickelten Potenziale zu einer Leistungsbeurteilung zu gelangen.

Das skizzierte Stimmportfolio ist als alternative Form der Leistungsmessung konzipiert, kann sehr unterschiedliche Formen von Belegen enthalten, sollte aber nicht nur die musikalische Seite des Gesangs berücksichtigen, sondern im Sinne eines holistischen Ansatzes auch entwicklungspsychologische Elemente beinhalten und insbesondere die adoleszente Stimmentwicklung im Auge behalten. Mögliche Elemente zur Bildung eines Portfolios wurden vom Forschenden als Basischeck Stimme konzipiert (Imthurn, 2020). Dabei wird ein Fokus auf die Entwicklung von digitalen Instrumenten gelegt, die einerseits den Schülerinnen und Schülern direktes Feedback und größere Autonomie gewähren (Ryan & Deci, 2000) und andererseits den Lehrkräften die Grundlage für Diagnose und Förderung

liefern¹⁸¹. Das Stimmportfolio kann so Basis eines ganzheitlichen Assessment-Prozesses werden, der formative, summative und metakognitive Aspekte vereint. Ob dies zielführend ist, gilt es zu überprüfen (vgl. 15.3).

Den Ideen und Gedanken der vorgestellten Thesen folgen im nächsten Kapitel drei Forschungsdesiderate, die darin münden, die Wirksamkeit ganzheitlichen Assessments im Rahmen von Unterrichtsforschung zu überprüfen.

¹⁸¹ Eine weitere Ressource wurde ebenfalls vom Forschenden entwickelt und beinhaltet bereits einige zum skizzierten Portfolio passende Aufgabenstellungen, z. B. das Aufgabenset (vgl. Luthiger et al., 2018) JEDE STIMME IST EINZIGARTIG (<https://www.lernumgebungen.ch/lu-singen/stimme-entwickeln/jede-stimme-ist-einzigartig> – abgerufen am 15.01.2021), in dem Wissen zu Stimmparametern entwickelt wird, das abschließend in einer Partnerarbeit auf die eigene Stimme angewandt wird. In den verschiedenen Aufgabenstellungen wird mehrperspektivisch auf die Einzigartigkeit der menschlichen Stimme verwiesen.

15 Weiterführende Fragestellungen

Die vorliegende Studie zeigt am Beispiel eines kleinen Samples ein heterogenes Bild des Singunterrichts in der Nordwestschweiz. Mit der Betrachtung durch die Assessment-Linse wurde eine fokussierte Situationsanalyse bezogen auf Diagnose, Förderung und Beurteilung ermöglicht. Während der Analyse zeigten sich verschiedene Desiderate, die nun abschließend als weiterführende Fragestellungen skizziert werden. Das bereits in der Einleitung der Arbeit zitierte didaktische Dreieck (Bönsch, 2006) zeigt sich

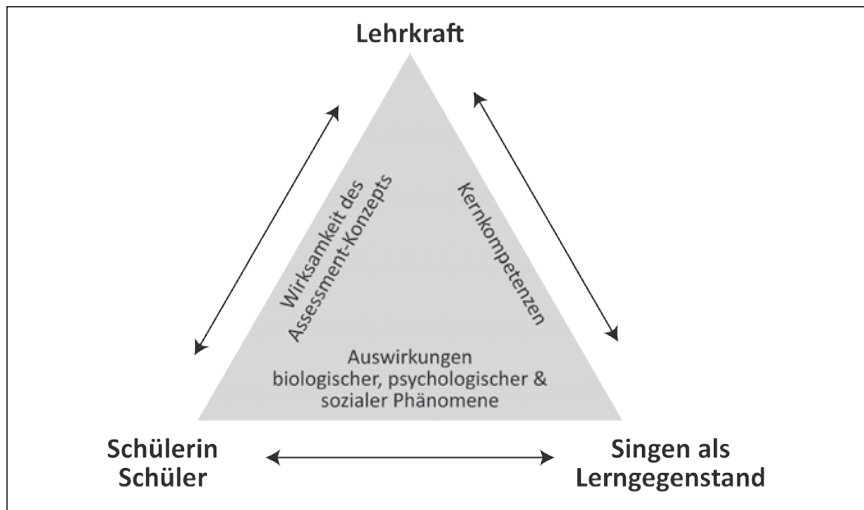


ABBILDUNG 15: Forschungsdesiderate im didaktischen Dreieck (adaptiert nach W. Jank und Meyer, 2014, S. 55)

dabei nochmals als hilfreiche Folie für die Verortung dieser Desiderate, allerdings in einer adaptierten Form. So wird von W. Jank und Meyer (2014, S. 55) das ursprünglich unidirektional auf Vorbereiten, Lehren und Lernen ausgerichtete didaktische Modell erweitert. Statt der unidirektionalen Ausrichtung betonen sie neu die Wichtigkeit der Wechselwirkungen zwischen den drei Komponenten Lehrkraft, Schülerin/Schüler und Lerngegenstand (vgl. Abbildung 15). Diese Wechselwirkungen wurden in der vorliegenden Arbeit zwischen allen Komponenten des Modells sichtbar,

bedürfen aber einer vertieften Untersuchung, unter anderem auch aus Perspektiven, die in der vorliegenden Studie nicht möglich waren.

Zwischen den Komponenten Schülerin/Schüler und dem Lerngegenstand Singen muss die Relevanz der biologischen, psychologischen und sozialen Auswirkungen der adoleszenten Stimmentwicklung auf das Unterrichtsgeschehen geklärt werden. In den Interviews konnten zwar einige Phänomene der adoleszenten Stimmentwicklung aufgrund des Forschungsstandes eingeordnet werden, die tatsächliche Relevanz für Schülerinnen und Schüler bleibt aber aufgrund des Forschungszugangs als Postulat stehen. Auch bezüglich der Wechselwirkungen zwischen Lerngegenstand und Lehrkraft präsentiert sich ein vielfältiges Bild von Zugängen und Methoden. So stellt sich die Frage, inwiefern diese und die Ansprüche für ein ganzheitliches Assessment im Singunterricht als *core practices* (Millican & Forrester, 2019) verdichtet werden können. Und nicht zuletzt findet der Singunterricht als Miteinander von Schülerinnen und Schülern sowie Lehrkräften statt, wobei die Wirksamkeit von Assessment-Methoden überprüft werden müsste. Mit diesem Kapitel wird ein Bogen ausgehend von relevanten Forschungslücken zu Beginn der Arbeit hin zu sich öffnenden Forschungsfeldern am Ende der Arbeit gespannt.

15.1 Relevanz der adoleszenten Stimmentwicklung im Singunterricht

Der Forschungsstand zur adoleszenten Stimmentwicklung wurde in der vorliegenden Arbeit auf biologische, entwicklungspsychologische und soziale Phänomene hin untersucht und in der QIA mit den empirischen Daten abgeglichen. In der Studie konnten so die drei genannten Perspektiven exemplarisch herausgearbeitet werden (vgl. 13.1). Die aufgrund des Forschungsstandes postulierte Relevanz dieser Entwicklung für den Musikunterricht zeigt sich bei einigen Lehrkräften, bei anderen wird aber die Mutation im Unterricht kaum thematisiert, teilweise aus Angst, Jugendliche bloßzustellen. Zudem kann die in den Interviews eingenommene Perspektive die Relevanz der adoleszenten Stimmentwicklung für die Jugendlichen selbst nicht belegen, einmal aufgrund des Forschungszugangs, aber auch weil Selbstkonzepte meist nicht oder nur aufgrund ihrer Auswirkungen erkannt werden (Rieche et al., 2018). Nach der Rekapitulation der

drei Blickwinkel wird eine geeignete Untersuchungsstrategie vorgestellt, welche die Perspektive der Schülerinnen und Schüler mitdenkt.

Die in Kapitel 4 vorgestellten Studien zur adoleszenten Stimmentwicklung (Cooksey, 2000; Gackle, 2019; Freer, 2020) stammen aus dem anglofonen Raum. Die verwendeten Daten sind zwischen 20 und 50 Jahre alt. Aktuellere Daten zur Stimmentwicklung hat Fuchs (2019b) im deutschsprachigen Raum zwar vorgelegt, der Ansatz folgt allerdings nicht den vorgestellten Modellen und der Beschreibung individueller Verläufe, sondern der aus der Medizin bekannten Darstellung von Perzentilen, die Normwerte für die zu erwartende Entwicklung aufzeigen. Dieser Ansatz ist dann hilfreich, wenn auffällige Werte beobachtet werden, um diese einer medizinischen Diagnostik zuzuführen. Ein solcher Fall wäre z. B. die von Herrn Leimbacher (10) beschriebene auffällig tiefe Mädchenstimme, die sich unter Beachtung der Perzentile möglicherweise als auffällig herausstellen würde. Obwohl die Modelle zur Entwicklung von Mädchen- und Jungenstimme aktuell immer noch rezipiert werden (Welch, 2019), müsste deren Validität und der praktische Einsatz auch im deutschsprachigen Raum überprüft werden. Neben diesem biologischen Blickwinkel bestätigen aktuelle Studien aus der psychologischen Perspektive einen für den Musikunterricht wirksamen Zusammenhang zwischen dem musikalischen Selbstkonzept und der Motivation im Unterricht (Fiedler & J. Hasselhorn, 2020). Erst in Ansätzen wurden dagegen vokale Selbstkonzepte bei Mädchen (Sweet, 2018) und Jungen (Ashley, 2015; Freer, 2020) untersucht. Obwohl einige Forschende die Effekte der Persönlichkeitsentwicklung während der adoleszenten Stimmentwicklung in ihren Forschungen einbeziehen (vgl. auch McBride & Palkki, 2020), gibt es nur im Rahmen des Genderdiskurses Untersuchungen, wie sich diese im sozialen Umfeld der Schule auswirken (Hess, 2018). Insgesamt fehlen somit Studien, welche die biologische, entwicklungspsychologische und soziale Perspektive verbinden und für den Schulbereich fruchtbar machen.

Aufgrund der Komplexität der adoleszenten Stimmentwicklung wird vorgeschlagen, im Rahmen einer Triangulations- oder Mixed-Methods-Studie durch verschiedene Blickwinkel „multiperspektivische Forschungsergebnisse“ (Lüdemann & Otto, 2019, S. 3) anzustreben. Dabei wäre es möglich, eine Stimmfeldmessung (Fuchs et al., 2006; Riedmüller et al., 2010) als biologisch geprägte Perspektive mit entwicklungspsychologischen Aspekten des Selbstkonzepts (Fiedler & J. Hasselhorn, 2020; Freer, 2020)

sowie sozialen Phänomenen im Klassenverband (Hess, 2018) abzugleichen, um unterschiedliche „Theorien, Perspektiven, Methoden, Daten und/oder Forschenden innerhalb eines Forschungsprojektes“ (Gabriel, 2019, S. 14) zu beschreiben und zu verbinden. Als Ziel wäre anzustreben, positive und negative Faktoren der adoleszenten Stimmentwicklung für den allgemeinen Musikunterricht zu benennen und evtl. auch zu quantifizieren, um über Postulate und skizzierte Handlungsoptionen hinaus (vgl. Kap. 14) erprobte Modelle zu entwickeln.

15.2 Kernpraktiken für ganzheitliches Assessment

In der vorliegenden Arbeit zeigen sich interessante Phänomene rund um Assessment. So führt fast die Hälfte der Lehrkräfte neben dem gemeinsamen Gesang Gruppenarbeiten und monoedukative Unterrichtssequenzen durch oder fördert bewusst die Entwicklung der Persönlichkeit durch das Singen. Fast durchgehend werden die vokalen Leistungen summativ mit einem *Rubric* beurteilt. Ausnahmslos sind aber bewusst gestaltete diagnostische und formative Assessments nicht auszumachen. Die individuellen Methoden und Zugänge werden z. T. begründet, scheinen aber mehrheitlich intuitiv gewachsen zu sein (Brunner, 2014). Zudem fällt auf, dass in weiten Teilen der Arbeit die erstellten Ausbildungsprofile nicht stringent mit typischen Zielvorstellungen oder Handlungsaspekten von Assessment in Verbindung gebracht werden können. Auf der Suche nach belastbaren Begründungen für die intuitiv entwickelten Assessment-Methoden zeigt sich somit höchstens die langjährige Erfahrung als wichtiger Prädiktor. Insgesamt wird deutlich, dass nicht nur die Ausbildungsbiografien eine große Heterogenität aufweisen, sondern auch die individuell entwickelten Singkulturen. Die lokal und intuitiv entwickelten Praktiken lassen sich zwar an theoretische Überlegungen und Studienresultate anschließen, bauen aber nicht direkt darauf auf. Dies umschreibt Huber (2023, in Vorbereitung) für Schweizer Verhältnisse als „Hang fürs Praktische“ (S. 1) und stellt sogar eine Abgrenzung zu einer wissenschaftsbasierten Musikpädagogik fest.

Diese Unbestimmtheit weist einerseits auf die Problematik der Definition eines einheitlichen Sets von Methoden und Praktiken hin, macht andererseits aber deutlich, dass die Verankerung derselben in der Ausbildung nicht zwingend gelingt. Es gilt also die Herausforderung zu meistern, den

Umgang mit Schülerinnen und Schülern im Singunterricht sowie ganzheitliche Assessment-Methoden und gelingende Praktiken systematisch zu beschreiben, aber auch zu überlegen, wie diese den Lehrkräften verfügbar gemacht werden können.

Die Forschung von Puffer und Hofmann (2017, 2021) zum Aufbau fachdidaktischer Kompetenzen (PCK) und insbesondere von Diagnosekompetenz für den Singunterricht wurde bereits erwähnt. Wie sich in der Studie zeigte, konnten Studierende zwar durch den Einsatz von Videovignetten ihre diagnostischen Fähigkeiten steigern, im Vergleich zu den Fähigkeiten erfahrener Lehrkräfte zeigt sich aber nur ein bescheidener Fortschritt. Ein Katalog von Kompetenzen für Kinder- und Jugendchorleiter hat Schaumberger (2020b) mittels einer Interviewstudie evaluiert. Als Resultat schlägt er ein Bündel von Kompetenzen vor, die sich gut an die Modelle von Shulman (1987) sowie Baumert und Kunter (2006) anschließen lassen. Insbesondere die *Personale Kompetenz* scheint sich in seinem Sample als sehr wichtig erwiesen zu haben, was sich auch in der vorliegenden Studie bestätigt. So steht in der zweiten Typik die Pflege der Beziehung zu Schülerinnen und Schülern im Zentrum (vgl. 12.2). Die beiden Arbeiten zeigen, dass die Inventarisierung von PCK (oder auch CK oder PK) bereits aufgenommen wurde, die Relation zur Praxis aber nicht unbeachtet bleiben darf.

Als möglicher Weg zeigt sich das Konzept der *core practices*, das als Brücke zwischen Theorie und Praxis beschrieben wird (Grossman, 2018). Grossman et al. (2009) weisen in ihrem initialen Artikel auf Lücken hin, die durch die Betonung von Wissensbeständen entstehen: „This emphasis on knowledge has obscured the importance of other aspects of teaching, including the need for skill in orchestrating instructional activities, and the relational work involved in creating classroom communities.“ (S. 273) Kernpraktiken gehen somit über das Modell professioneller Handlungskompetenz nach Baumert und Kunter (2006) hinaus, indem sie das praktische Handeln im komplexen Unterrichtsgeschehen ins Zentrum rücken. Fraefel (2018) beschreibt vier Perspektiven zur Verdeutlichung der Relevanz von Kernpraktiken: Erstens erweist sich ihre Relevanz aus „den praktischen beruflichen Herausforderungen und Notwendigkeiten“ (S. 6), zweitens orientieren sie sich an wissenschaftlichen Theorien, drittens haben sie kein Praxisdefizit, weil die „*Habitualisierung* wirkungsvollen Handelns“ (S. 6, Hervorhebung im Original) im Zentrum steht, und viertens haben sie

kein Wirkungsdefizit, „weil deren Qualität sich an den potenziell günstigen Wirkungen auf die Schülerinnen und Schüler misst“ (S. 6). Er fasst die Vorteile für die Ausbildung folgendermaßen zusammen: „In der kontinuierlichen Beschäftigung mit Kernpraktiken bilden Studierende nach und nach einen Cluster von Impulsen, Wissensbausteinen, Erfahrungen und Experimenten, die sich zu einem stabilen und wirkungsvollen Handlungswissen verdichten.“ (S. 6) Für den deutschsprachigen Raum bezieht sich Fraefel zudem auf die Arbeit von Aebli (1961, 2011), für den englischsprachigen Raum auf die bereits genannten Autorinnen.

Erste Ansätze zu einem Diskurs von Kernpraktiken in der Musikpädagogik beschreiben Millican und Forrester (2019) auf der Grundlage von zwei vorausgehenden Arbeiten (Millican & Forrester, 2017, 2018). Im Rahmen einer Delphi-Studie wurde ein vorläufiges Set von Kernpraktiken definiert und bewertet. Interessanterweise wurde wiederum die Wichtigkeit der Beziehung zu den Schülerinnen und Schülern betont, aber auch auf die Wichtigkeit diagnostischer Zugänge verwiesen (S. 88). Damit zeigt sich eine Verbindung zu formativem Assessment und den vielfältigen Rollen und Aufgaben, die ganzheitliches Assessment ausmachen. Die Definition eines Sets von Kernpraktiken spezifisch für den Singunterricht mit Bezug auf die adoleszente Stimmentwicklung und ganzheitliches Assessment könnte so über die reine Auflistung von zu erreichenden Kompetenzen einen praxisbezogenen Beitrag in der Aus- und Weiterbildung leisten.

15.3 Gestaltung von Assessment im Singunterricht

Die vorliegende Arbeit untersuchte bereits bestehende meist intuitiv gewachsene Assessment-Praktiken und entwickelt auf der Basis des Forschungsstandes und der Resultate der QIA drei Thesen für mögliche Gelingensbedingungen von ganzheitlichem Assessment (vgl. Kapitel 14). Deren Wirksamkeit gilt es aber in der Praxis zu evaluieren. So müsste z. B. deutlich werden, dass die von einigen Lehrkräften vor allem auf einer beobachtenden Basis festgestellte natürliche Entwicklung der adoleszenten Stimmen durch geeignete Interventionen, wie z. B. der Erstellung eines Portfolios (vgl. 14.3), positiv beeinflusst werden kann. Die Untersuchung hat aber ebenso deutlich gemacht, dass Singunterricht unter unterschiedlichen Bedingungen, sei dies eine lokale Schulkultur, das Niveau der Schülerinnen

und Schüler oder die Biografie von Lehrerinnen und Lehrern, diverse Ausprägungen aufweist. Diese lokalen Singkulturen erschweren die Überprüfung der Thesen durch ein uniformes Interventionskonzept. Gesucht wird somit eine Evaluationsmethode, die lokale Praxen und theoretische Ansätze verbinden kann. Konrad (2019) beschreibt diese Herausforderung als geeignete Ausgangslage für Design-Based-Research (DBR): „Besteht für eine Studie im Bereich der Lehr-Lern-Forschung ein Desiderat von praktischer und theoretischer Relevanz, bildet der Design-Based-Research-Ansatz einen geeigneten methodologischen Rahmen.“ (S. 110) DBR wird speziell bezüglich einer doppelten Zielsetzung hervorgehoben. Lehmann-Wermser und Konrad (2016) beschreiben einerseits den Fokus auf die sich verbessernde Wirksamkeit eines Design-Produktes durch iterative Design-Zyklen. Die formulierte Idee eines Sing-Portfolios könnte so angepasst auf lokale Gegebenheiten mehrfach überarbeitet werden, um eine optimale Wirksamkeit zu entfalten. Andererseits kann im DBR-Prozess eine lokale Theoriebildung ermöglicht werden (S. 268). Geplante Interventionen sind dann nicht als Rezepte zu verstehen, sondern als Designprodukte, die kurzfristig lokale Theorien und längerfristig zum Theoriediskurs einer ganzheitlichen Leistungsbeurteilung für das Singen beitragen.

F LITERATURVERZEICHNIS

- 250+ *Influences on Student Achievement*. (2017). Visible Learning plus.
- Adorno, T. W. (1991). *Dissonanzen: Musik in der verwalteten Welt* (7. Aufl.). Vandenhoeck und Ruprecht.
- Aebli, H. (1961). *Grundformen des Lernens. Ein Beitrag zur psychologischen Grundlegung der Unterrichtsmethode*. Klett.
- Aebli, H. (2011). *Zwölf Grundformen des Lernens. Eine Allgemeine Didaktik auf psychologischer Grundlage* (14. Auflage). Klett.
- Altman, D. G. (1991). *Practical Statistics for Medical Research*. Chapman & Hall.
- American Educational Research Association, American Psychological Association & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Anderson, L. (2006). Analytic autoethnography. *Journal of contemporary Ethnography*, 35(4), 373–395. <https://doi.org/10.1177/0891241605280449>
- Anderson, L. & Glass-Coffin, B. (2016). I learn by going. In S. H. Jones, T. E. Adams & C. Ellis (Hrsg.), *Handbook of autoethnography* (S. 57–83). Routledge.
- Antwerpen, S. (2014). *Singen in der Schule: Ästhetische Bildungspotenziale des Singens und Gesangs*. Waxmann.
- Ashley, M. (2006). You sing like a girl? An exploration of ‘boyiness’ through the treble voice. *Sex Education*, 6(2), 193–205. <https://doi.org/10.1080/14681810600578792>
- Ashley, M. (2009). *How high should boys sing? Gender, authenticity and credibility in the young male voice*. Routledge.
- Ashley, M. (2015). *Singing in the lower secondary school*. Oxford University Press.

- Ayotte, J., Peretz, I. & Hyde, K. (2002). Congenital amusia: A group study of adults afflicted with a music-specific disorder. *Brain*, 125(2), 238–251. <https://doi.org/10.1093/brain/awf028>
- Bauer, J. (2008). *Lob der Schule: Sieben Perspektiven für Schüler, Lehrer und Eltern* (2. Aufl.). Heyne.
- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469–520. <https://doi.org/10.1007/s11618-006-0165-2>
- Baumert, J. & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert, W. Blum, U. Klusman, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (S. 29–53). Waxmann.
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 18(1), 5–25. <https://doi.org/10.1080/0969594X.2010.513678>
- Berelson, B. (1952). *Content analysis in communication research*. Free Press.
- Bergen, H. von (1995). *Unsere Stimme: Ihre Funktion und Pflege* (2. Aufl.). Müller & Schade.
- Berger Kofmel, N. (2015). *Diversity an der FHNW: Ein Wegweiser für die Hochschulpraxis* (FHNW, Hrsg.).
- Biegholdt, G., Witt, K.-U. & Lutz, J. (2020). *Wie hängen Selbsteinschätzung, Feedback, Beurteilung und Bewertung zusammen?* Friedrich Verlag.
- Birrer, R. (15. August 2017). Lehrermangel wird sich zuspitzen. *Tages-Anzeiger*. <https://www.tagesanzeiger.ch/schweiz/standard/fuer-eine-entwarnung-ist-es-zu-frueh/story/27218234>
- BKS Aargau. (2019). *Evaluationsbericht: Lehrmittel Musik Oberstufe* [Evaluationsbericht]. Departement Bildung, Kultur und Sport BKS.
- Black, P. & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *The Phi Delta Kappan*, 80(2), 139–148.

- Black, P. & Wiliam, D. (2010). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 92(1), 81–90. <https://doi.org/10.1177/003172171009200119>
- Blanchard, O. (2019). *Hegemonie im Musikunterricht: Die Befremdung der eigenen Kultur als Bedingung für den verständigen Umgang mit kultureller Diversität*. Waxmann.
- Blanchard, O. (2021). Schweizer Schulmusik: Schule oder Musik? In J. Huber, M.-A. Camp, O. Blanchard, S. Chatelain, F. Joliat, R. Steiner, & J. Zurmühle (Hrsg.), *Kulturen der Schulmusik in der Schweiz. Les cultures d l'enseignement musical à l'école en Suisse* (S. 63–78). Chronos.
- Blömeke, S. (2011). Kompetenzen von Lehramtsstudierenden in gering strukturierten Domänen. Erste Ergebnisse aus TEDS-LT. In S. Blömeke, A. Bremerich-Vos, H. Haudeck, G. Kaiser, G. Nold & K. Schwippert (Hrsg.), *Kompetenzen von Lehramtsstudierenden in gering strukturierten Domänen: Erste Ergebnisse aus TEDS-LT* (S. 8–25). Waxmann.
- Bloom, B. S., Hastings, J. T., Madaus, George F. & Madaus, G. F. (1971). *Handbook on formative and summative evaluation of student learning*. McGraw-Hill.
- Bönsch, M. (2006). *Allgemeine Didaktik: Ein Handbuch zur Wissenschaft vom Unterricht*. Kohlhammer.
- Bojack-Weber, R. (2012). *Singen in der Grundschule: Eine Untersuchung zur Singfähigkeit und zum Singverhalten von Grundschulkindern*. Wißner.
- Bolender, R. & Müller, G. (2012a). *Gesangsklasse: Leitfaden*. Helbling.
- Bolender, R. & Müller, G. (2012b). *Gesangsklasse: Schülerheft*. Helbling.
- Bolli, F., Caviezel, A. & Vonarburg, P. (1999). *Musik Sekundarstufe 1*. Verlag Schweizer Singbuch.
- Bottoms, J. F. (1995). *Investigation of William Vennard's light mechanism principle in the adolescent female voice*. University of Colorado.
- Braun, J. & Ehrenspeck-Kolasa, Y. (2019). Die qualitative Inhaltsanalyse als ‚Sehangebot‘ und als Medium der Rekonstruktion professionellen Handelns und erziehungswissenschaftlicher Theoriebildung. In

- S. Gottuck, I. Grünheid, P. Mecheril & J. Wolter (Hrsg.), *Sehen lernen und verlernen: Perspektiven pädagogischer Professionalisierung* (S. 175–195). Springer Fachmedien. https://doi.org/10.1007/978-3-658-19496-3_8
- Brennan, R. L. & Prediger, D. J. (1981). Coefficient Kappa: Some uses, misuses, and alternatives. *Educational and Psychological Measurement*, 41(3), 687–699. <https://doi.org/10.1177/001316448104100307>
- Bromme, R. (1992). *Der Lehrer als Experte: Zur Psychologie des professionellen Wissens*. Huber.
- Brophy, T. S. (Hrsg.). (2019a). *The Oxford handbook of assessment policy and practice in music education* (Bd. 1). Oxford University Press.
- Brophy, T. S. (Hrsg.). (2019b). *The Oxford handbook of assessment policy and practice in music education* (Bd. 2). Oxford University Press.
- Brophy, T. S. (2019c). Assessment in music education: The state of the art. In T. S. Brophy (Hrsg.), *The Oxford handbook of assessment policy and practice in music education* (Bd. 2, S. 903–932). Oxford University Press.
- Brophy, T. S. & Fautley, M. (20. April 2017). *International principles for assessment in music education*. Presented at Context matters: The 6th International Symposium on Assessment in Music Education. Birmingham City University, Birmingham, England. Unpublished manuscript.
- Brophy, T. S. & Haning, M. (Hrsg.). (2020). *Advancing music education through assessment: Selected papers from the seventh international symposium on assessment in music education*. GIA.
- Brünger, P. (2011). Sing doch mal richtig! Wie soll man im Musikunterricht mit Brummern umgehen? *Grundschule*, 9, 38–39.
- Bruhn, H. (1991). Hör auf zu singen – Du bist ja ein Brummer! In R.-D. Kraemer (Hrsg.), *Musikpädagogik Unterricht – Forschung – Ausbildung* (S. 49–63). Schott.
- Brunner, G. (2014). Diagnose- und Förderstrategien von Lehrkräften im Musikunterricht der Grundschule zur Qualitätsverbesserung des Sin-

- gens. In M. Fröhlich & G. Brunner (Hrsg.), *Impulse der Musikdidaktik. Festschrift für Mechtild Fuchs* (S. 231–270). Helbling Verlag.
- Buchborn, T. & Malmberg, I. (2013). Forschung aus der Perspektive musikpädagogischer Praxis. *Diskussion Musikpädagogik*, 57, 4–13.
- Budde, J., Kansteiner, K. & Bossen, A. (2016). *Zwischen Differenz und Differenzierung: Erziehungswissenschaftliche Forschung zu Mono- und Koedukation*. Springer VS.
- Busch, T. & Kranefeld, U. (2012). Sind Mädchen die besseren Optimisten im Umgang mit Musik? Prädiktoren des musikalischen Selbstkonzepts in der Klassenstufe 2. *Beiträge empirischer Musikpädagogik*, 3(1), 1–28.
- Carmichael, M. (2014). *Zum Kompetenzerleben von Schülern im Musikunterricht – Eine empirische Studie mit Schülern der Jahrgangsstufe 7–12* [MA, Hochschule der Künste]. http://www.fem-berlin.de/files/Masterarbeit_Carmichael-fem.pdf
- Carmichael, M. & Harnischmacher, C. (2015). Ich weiß, was ich kann! Eine empirische Studie zum Einfluss des musikbezogenen Kompetenzerlebens und der Motivation von Schülerinnen und Schülern auf deren Einstellung zum Musikunterricht. In A. Niessen & J. Knigge (Hrsg.), *Theoretische Rahmung und Theoriebildung in der musikpädagogischen Forschung* (S. 177–198). Waxmann.
- Chang, H. (2008). *Autoethnography as method*. Routledge.
- Ciorba, C. R. & Smith, N. Y. (2009). Measurement of instrumental and vocal undergraduate performance juries using a multidimensional assessment rubric. *Journal of Research in Music Education*, 57(1), 5–15.
- Clifford, J. & Marcus, G. E. (1986). *Writing culture: The poetics and politics of ethnography: A school of American research advanced seminar*. University of California Press.
- Cohen, A. J. (2015). The AIRS test battery of singing skills: Rationale, item types, and lifespan scope. *Musicae Scientiae*, 19(3), 238–264. <https://doi.org/10.1177/1029864915599599>

- Cooksey, J. M. (1977a). The development of a contemporary, eclectic theory for the training and cultivation of the junior high school male changing voice: Part I Existing theories. *The Choral Journal*, 18(2), 5–14.
- Cooksey, J. M. (1977b). The development of a contemporary, eclectic theory for the training and cultivation of the junior high school male changing voice: Part II Scientific and empirical findings: Some tentative solutions. *The Choral Journal*, 18(3), 5–16.
- Cooksey, J. M. (1977c). The development of a contemporary, eclectic theory for the training and cultivation of the junior high school male changing voice: Part III Developing an integrated approach to the care and training of the junior high school male changing voice. *The Choral Journal*, 18(4), 5–15.
- Cooksey, J. M. (1978). The development of a contemporary, eclectic theory for the training and cultivation of the junior high school male changing voice: Part IV Selecting music for the junior high school male changing voice. *The Choral Journal*, 18(5), 5–18.
- Cooksey, J. M. (2000). Voice transformation in male adolescents. In L. Thurman & G. Welch (Hrsg.), *Bodymind & voice: Foundations of voice education* (S. 718–738). VoiceCare Network; National Center for Voice & Speech; Voice Center of Fairview.
- Cooksey, J. M. & Welch, G. F. (1998). Adolescence, singing development and national curricula design. *British Journal of Music Education*, 15(1), 99–119. <https://doi.org/10.1017/S026505170000379X>
- Crawford-Ferre, H. G. & Wiest, L. R. (2013). Single-sex education in public school settings. *The Educational Forum*, 77(3), 300–314. <https://doi.org/10.1080/00131725.2013.792906>
- Criblez, L. & Lehmann, L. (2007). Lehrerinnen- und Lehrerbildung in der Schweiz: Ausgewählte Ergebnisse aus der Bildungsberichterstattung. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 25(3), 377–391.
- Csikszentmihalyi, M. (2008). *Flow: The Psychology of Optimal Experience*. Harper Perennial Modern Classics.

- Cslovjcek, M. (2009). ‚Alle müssen wollen‘ – Musikalische Aktivität im Klassenunterricht als Herausforderung und Chance. In H.-U. Grunder & A. Gut (Hrsg.), *Zum Umgang mit Heterogenität in der Schule* (Bd. 1, S. 82–92). Schneider Hohengehren.
- Dalla Bella, S., Giguère, J.-F. & Peretz, I. (2007). Singing proficiency in the general population. *The Journal of the Acoustical Society of America*, 121(2), 1182–1189. <https://doi.org/10.1121/1.2427111>
- Dartsch, M., Knigge, J., Niessen, A., Platz, F. & Stöger, C. (Hrsg.). (2018). *Handbuch Musikpädagogik: Grundlagen – Forschung – Diskurse*. Waxmann. <https://elibrary.utb.de/doi/book/10.36198/9783838550404>
- D-EDK. (2014). *Grundlagen*. Deutschschweizer Erziehungsdirektoren-Konferenz. www.lehrplan.ch
- D-EDK (Hrsg.). (2015). *Lehrplan 21 – Schulfach Musik*. <http://vorlage.lehrplan.ch/>
- Demorest, S. M., Kelley, J. & Pfordresher, P. Q. (2017). Singing ability, musical self-concept, and future music participation. *Journal of Research in Music Education*, 64(4), 405–420. <https://doi.org/10.1177/0022429416680096>
- Demorest, S. M. & Pfordresher, P. Q. (2015). *Seattle singing accuracy protocol – SSAP* [Measurement Instrument]. Seattle Singing Accuracy Protocol – SSAP. <https://ssap.music.northwestern.edu>
- Demorest, S. M., Pfordresher, P. Q., Bella, S. D., Hutchins, S., Loui, P., Rutkowski, J. & Welch, G. F. (2015). Methodological perspectives on singing accuracy. *Music Perception*, 32(3), 266–271. <https://doi.org/10.1525/mp.2015.32.3.266>
- Denzin, N. K. (2006). Analytic autoethnography, or déjà vu all over again. *Journal of Contemporary Ethnography*, 35(4), 419–428. <https://doi.org/10.1177/0891241606286985>
- Denzin, N. K. & Lincoln, Y. (Hrsg.). (2000). *Handbook of qualitative research* (2. Aufl.). SAGE Publications.
- Detterbeck, M. (2011). *Musix: Das Kursbuch Musik*. Helbling.

- Dixson, D. D. & Worrell, F. C. (2016). Formative and summative assessment in the classroom. *Theory Into Practice*, 55(2), 153–159. <https://doi.org/10.1080/00405841.2016.1148989>
- Douglas, K. & Carless, D. (2016). A history of autoethnographic inquiry. In S. Holman Jones, T. E. Adams & C. S. Ellis (Hrsg.), *Handbook of autoethnography* (S. 84–106). Routledge.
- Driscoll, M. J. (1998). *Learning about assessment, learning through assessment*. National Academy Press.
- Dyllick, N. (2011). Vokalpraxis in der Schule: Durch eine systemisch-konstruktivistische Perspektive zu einem veränderten fachdidaktischen Verständnis? In B. Clausen (Hrsg.), *Vergleich in der musikpädagogischen Forschung* (S. 59–84). Die blaue Eule.
- Dyllick, N. (2019). *Vokalpraxis in der Schule: Aus der Perspektive einer systemisch-konstruktivistischen Pädagogik*. Verlag Dohr.
- Earl, L. M. (2003). *Assessment as learning: Using classroom assessment to maximize student learning*. Corwin.
- Earl, L. M. (2013). *Assessment as learning: Using classroom assessment to maximize student learning* (2. Aufl.). Corwin.
- Earl, L. M. & Katz, M. S. (2006). *Rethinking classroom assessment with purpose in mind: Assessment for learning, assessment as learning, assessment of learning*. Manitoba Education, Citizenship & Youth.
- Eckhardt, D. (2019). *Musik machen inmitten Anderer: Reflexionen über das Singen im Chor*. LIT.
- EDK. (2015). *Harmonisierung der verfassungsmässigen Eckwerte (Art. 62 Abs. 4 BV) für den Bereich der obligatorischen Schule [Bilanz 2015]*. Schweizerische Konferenz der kantonalen Erziehungsdirektoren EDK.
- Elliott, D. J., Silverman, M. & McPherson, G. E. (2019). *The Oxford handbook of philosophical and qualitative assessment in music education*. Oxford University Press.
- Ellis, C. S. (2004). *The ethnographic I: A methodological novel about autoethnography*. AltaMira Press.

- Ellis, C. S., Adams, T. E. & Bochner, A. P. (2010). Autoethnografie. In G. Mey & K. Mruck (Hrsg.), *Handbuch Qualitative Forschung in der Psychologie* (S. 345–357). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92052-8_24
- Ellis, C. S. & Bochner, A. P. (2000). Autoethnography, personal narrative, reflexivity: Researcher as subject. In N. K. Denzin (Hrsg.), *Handbook of qualitative research* (2. Aufl., S. 733–769). SAGE Publications.
- Ellis, C. S. & Bochner, A. P. (2006). Analyzing analytic autoethnography: An autopsy. *Journal of Contemporary Ethnography*, 35(4), 429–449. <https://doi.org/10.1177/0891241606286979>
- Evelein, F. (2010). Interaktion, Kreativität und Autonomie. Herausforderungen für eine moderne Musikdidaktik. *mip-Journal*, 28, 6–11.
- Evers, S. (2012). Amusie nach Schlaganfall. *Der Neurologe & Psychiater*, 13(3), 27–27. <https://doi.org/10.1007/s15202-012-0107-1>
- Fautley, M. (2010). *Assessment in music education*. Oxford University Press.
- Fiedler, D. (2016). Verification and Validation of the Musical Self-Concept Inquiry (MUSCI) to Measure ‚Musical Self-Concept‘ of German Students at Secondary Education Schools. In B. Burger, J. Bamford & E. Carlson (Hrsg.), *The 9th International Conference of Students of Systematic Musicology (SysMus16), Jyväskylä yliopisto, June 8–10 2016: Programme, abstracts & proceedings*. Department of Music, University of Jyväskylä & Finnish Centre for Interdisciplinary Music Research.
- Fiedler, D. & Hasselhorn, J. (2020). Zum Zusammenhang von musikalischem Selbstkonzept und Motivation im Musikunterricht. *Beiträge empirischer Musikpädagogik*, 11, 1–34.
- Fleiss, J. L., Levin, B. & Paik, M. C. (2003). *Statistical methods for rates and proportions* (3. Aufl.). J. Wiley. <https://doi.org/10.1002/0471445428>
- Flick, U. (2019). Gütekriterien qualitativer Sozialforschung. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S. 473–488). Springer Fachmedien. https://doi.org/10.1007/978-3-658-21308-4_33

- Fluri, F., Fricker, R., Mattle, S., Nösberger, D. & Weber, B. (2017). *Musik aktiv: Klasse 7–9 (3. Zyklus)*. Verlag Schweizer Singbuch.
- Forge, S. (2016). *Die Nachhaltigkeit von Lehrerbildung am Beispiel des Singförderprojekts „SMS – Singen macht Sinn“* [Dissertation]. Universität Paderborn.
- Forge, S. & Gembris, H. (2012). *Singförderung in der Grundschule. Evaluation des Projekts ‚Singen macht Sinn‘*. LIT.
- Fraefel, U. (2018). *Portfolio: Ein Arbeitsinstrument für den Aufbau wirkungsvoller Praktiken*. Pädagogische Hochschule FHNW.
- Frank, F. & Sparber, M. (1970a). Die Prämutationsstimme, die Mutationsstimme und die Postmutationsstimme im Sonagramm. *Folia Phoniatica et Logopaedica*, 22(6), 425–433. <https://doi.org/10.1159/000263422>
- Frank, F. & Sparber, M. (1970b). Stimmumfänge bei Kindern aus neuer Sicht. *Folia Phoniatica et Logopaedica*, 22(6), 397–402. <https://doi.org/10.1159/000263419>
- Freer, P. K. (2015). Perspectives of European boys about their voice change and school choral singing: Developing the possible selves of adolescent male singers. *British Journal of Music Education*, 32(1), 87–106. <https://doi.org/10.1017/S026505171400031X>
- Freer, P. K. (2020). The adolescent boy’s changing voice. In J. van der Sandt (Hrsg.), *Singing with children: International perspectives* (S. 3–18). Libreria Musicale Italiana.
- Freer, P. K. & Detterbeck, M. (2019). Wenn die Stimme ‚bricht‘. *mip-Journal*, 55, 6–11.
- Frey, N., Hattie, J. & Fisher, D. (2018). *Developing assessment-capable visible learners, grades K-12: Maximizing skill, will, and thrill*. Corwin.
- Frey, O. (2012). Städtische Milieus. In F. Eckardt (Hrsg.), *Handbuch Stadtsoziologie* (S. 503–525). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-94112-7_22

- Frierson-Campbell, C. & Newman, T. U. (2008). Towards a rationale for a scholarly practicum for jazz performance majors: A collaborative auto-ethnography. *Proceedings 28th ISME World Conference. Music at All Ages*, 97–101.
- Frühholz, S. & Belin, P. (2019). The Oxford handbook of voice perception. In *The Oxford Handbook of Voice Perception*. Oxford University Press. <https://doi.org/10.1093/oxfordhnb/9780198743187.001.0001>
- Fuchs, M. (2010). Das Kind ist kein verkleinerter Erwachsener! Gemeinsamkeiten und Unterschiede zwischen kindlichen, jugendlichen und erwachsenen Stimmen. In M. Fuchs (Hrsg.), *Wechselwirkungen zwischen Erwachsenen- und Kinderstimme* (S. 6–16). Logos Verlag.
- Fuchs, M. (2015). Die heisere Kinder- und Jugendstimme. *mip-Journal*, 43, 6–10.
- Fuchs, M. (2018). Bewusstheit – Entfaltung – Wirkung: Wie entwickelt sich bei Kindern und Jugendlichen die Beziehung zur eigenen Stimme? In M. Fuchs (Hrsg.), *Beziehungssystem Stimme* (S. 11–18). Logos Verlag.
- Fuchs, M. (2019a). *Gesangsstile – Stimmtechniken – Stimmgesundheit*. Logos Verlag.
- Fuchs, M. (2019b). Phoniatriische Betreuung kritischer Entwicklungsphasen der Kinder- und Jugendstimme anhand von Normwertbereichen. In M. Fuchs (Hrsg.), *Gesangsstile – Stimmtechniken – Stimmgesundheit* (S. 41–54). Logos Verlag.
- Fuchs, M. (2020). Voice training during mutation. In J. van der Sandt (Hrsg.), *Singing with children: International perspectives* (S. 19–32). Libreria Musicale Italiana.
- Fuchs, M., Heide, S., Hentschel, B., Gelbrich, G., Makuch, A., Thiel, S., Täschner, R. & Dietz, A. (2006). Stimmleistungsparameter bei Kindern und Jugendlichen. *HNO*, 54(12), 971–980. <https://doi.org/10.1007/s00106-005-1307-1>
- Gabriel, S. (2019). Triangulation als theoretisierte Verhältnisfrage zwischen Gegenstandskonstruktionen in qualitativen Forschungsprojekten. In J. Lüdemann & A. Otto (Hrsg.), *Triangulation und Mixed-Methods*:

Reflexionen theoretischer und forschungspraktischer Herausforderungen (S. 13–38). Springer VS. <http://link.springer.com/10.1007/978-3-658-24225-1>

Gackle, L. (1991). The adolescent female voice: Characteristics of change and stages of development. *The Choral Journal*, 31(8), 17–25.

Gackle, L. (2000). Understanding voice transformation in female adolescents. In L. Thurman & G. F. Welch (Hrsg.), *Bodymind and voice: Foundations of voice education* (S. 739–744). VoiceCare Network; National Center for Voice & Speech; Voice Center of Fairview.

Gackle, L. (2006). Finding Ophelia's voice: The female voice during adolescence. *The Choral Journal*, 47(5), 28–37.

Gackle, L. (2011). *Finding Ophelia's voice, opening Ophelia's heart: Nurturing the adolescent female voice*. Heritage Music Press.

Gackle, L. (2019). Adolescent girls' singing development. In G. F. Welch, D. M. Howard & J. Nix (Hrsg.), *The Oxford handbook of singing* (S. 551–569). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199660773.013.22>

Gamoran Sherin, M. & Es, E. A. van (2009). Effects of video club participation on teachers' professional vision. *Journal of Teacher Education*, 60(1), 20–37. <https://doi.org/10.1177/0022487108328155>

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. Basic Books.

Geimer, A. (2011). Performance Ethnography und Autoethnography: Trend, Turn oder Schisma in der qualitativen Forschung? *Zeitschrift für Qualitative Forschung*, 12(2), 299–230.

Gembris, H. (2015). *Transfer-Effekte und Wirkungen musikalischer Aktivitäten auf ausgewählte Bereiche der Persönlichkeitsentwicklung: Ein Überblick über den aktuellen Stand der Forschung*. Bertelsmann Stiftung.

Gembris, H., Kraemer, R.-D. & Maas, G. (2001). *Macht Musik wirklich klüger? Musikalisches Lernen und Transfereffekte*. Wißner.

- Gend, M. van, Matthias, P., Allan, C. & Scott, N. (2017). Adolescent girls' changing voices: Re-thinking classification. *Australian Voice*, 18, 1–13.
- Gläser, J. & Laudel, G. (2010). *Experteninterviews und qualitative Inhaltsanalyse: Als Instrumente rekonstruierender Untersuchungen* (4. Aufl.). VS Verlag für Sozialwissenschaften.
- Glaser, B. G., Paul, A. T. & Strauss, A. L. (1998). *Grounded theory: Strategien qualitativer Forschung*. Huber.
- Gohl, W. & Hux, A. (1975). *Musik auf der Oberstufe*. Verlag Schweizer Singbuch Oberstufe.
- Gohl, W. & Löhner, R. (1997). *Musik auf der Oberstufe: Lieder, Tänze, Musikkunde*. Schweizer Singbuch Oberstufe.
- Gouzouasis, P., Bakan, D., Ryu, J. Y., Ballam, H., Murphy, D., Ihnatovych, D., Virag, Z. & Yanko, M. (2014). Where do teachers and learners stand in music education research? A multi-voiced call for a new ethos of music education research. *International Journal of Education & the Arts*, 15(15), 1–23.
- Graefe-Hessler, D. & Jank, W. (2015). Primacanta – Ein Erfahrungsbericht nach fünf Projektjahren. In S. Schmid (Hrsg.), *Musikunterricht(en) im 21. Jahrhundert: Begegnungen – Einblick – Visionen* (S. 219–226). Wißner.
- Graff, U. (2013). Sonderangebot oder Notlösung? Zum Status geschlechts-homogener Pädagogik in der ‚reflexiven Koedukation‘. In D.-T. Chwalek, M. Diaz, S. Fegter & U. Graff (Hrsg.), *Jungen – Pädagogik: Praxis und Theorie von Genderpädagogik* (S. 91–99). Springer Fachmedien. https://doi.org/10.1007/978-3-531-94290-2_7
- Green, L. (1993). Music, gender and education: A report on some exploratory research. *British Journal of Music Education*, 10(3), 219–253. <https://doi.org/10.1017/S0265051700001789>
- Green, L. (1997). *Music, gender, education*. Cambridge University Press.
- Greuel, T. (2006). Musikpädagogische Stimmdiagnostik in der allgemein bildenden Schule. *Afs-Magazin*, 22, 14–18.

- Greuel, T., Heyer, T. & Horst, U. (2007). Musikpädagogische Stimmdiagnostik in der allgemein bildenden Schule. In T. Greuel (Hrsg.), *In Möglichkeiten denken – Qualität verbessern: Auf dem Weg zu einer musikpädagogischen Diagnostik* (S. 89–96). Bosse.
- Greuel, T. & Horst, U. (2011). Singen zwischen Wunsch und Wirklichkeit: Musikpädagogische Stimmdiagnose in der allgemein bildenden Schule. In T. Greuel, U. Kranefeld & E. Szczepaniak (Hrsg.), *Singen und Lernen – Perspektiven auf schulische und außerschulische Vokalpraxis* (S. 69–86). Shaker.
- Grohé, M. (2011). *Der Musiklehrer-Coach. Professionelles Handeln in konflikthafter Unterrichtssituationen*. Helbling.
- Grossman, P. (Hrsg.). (2018). *Teaching Core Practices in Teacher Education*. Harvard Education Press.
- Grossman, P., Hammerness, K. & McDonald, M. (2009). Redefining teaching, re-imagining teacher education. *Teachers and Teaching*, 15(2), 273–289. <https://doi.org/10.1080/13540600902875340>
- Gütay, W. (2012). *Singen in Chorklassen. Eine Längsschnittstudie zu stimmphysikalischen, kognitiven und psychosozialen Auswirkungen*. IBIT – Universitätsbibliothek.
- Gütay, W. & Kreutz, G. (2010). Entwicklung der Singstimme von Kindern und Chorklassen: Eine medizinisch-physikalische Längsschnittuntersuchung von Stimmleistungsparametern. In B. Schwarz, P. Nenniger & R. S. Jäger (Hrsg.), *Erziehungswissenschaftliche Forschung – Nachhaltige Bildung. Beiträge zur 5. DGfE-Sektionstagung ‚Empirische Bildungsforschung‘ – AEPF-KBBB im Frühjahr 2009* (S. 461–468). Verlag Empirische Pädagogik.
- Gwet, K. L. (2014). *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters* (4. Aufl.). Advanced Analytics, LLC.
- Häcker, T. (2005). Portfolio als Instrument der Kompetenzdarstellung und reflexiven Lernprozesssteuerung. *bwp@*, 8, 1–11.
- Häcker, T. (2011). Portfolio revisited – über Grenzen und Möglichkeiten eines viel versprechenden Konzepts. In T. Meyer, K. Mayrberger,

- S. Münte-Goussar & C. Schwalbe (Hrsg.), *Kontrolle und Selbstkontrolle: Zur Ambivalenz von E-Portfolios in Bildungsprozessen* (S. 161–183). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92722-0_17
- Hammar, J. (2016). *Gesang lehren und lernen im Spannungsfeld zwischen Instinkt und Wissenschaft: Besonderheiten der Gesangsausbildung unter Berücksichtigung neuer Feedbackmöglichkeiten* (3. Aufl.). Wißner.
- Hansen, H., Kübler, M. & Seher, A. (Hrsg.). (2018). *Clever Aufgaben: Bausteine der Unterrichtsentwicklung*. hep der bildungsverlag.
- Harnischmacher, C. & Hörtzsch, U. (2012). Motivation und Musikunterricht. Eine empirische Studie zum Vorhersagewert des Motivationsmodells Musikalischen Handelns auf die Einstellung zum Musikunterricht aus Schülersicht. In *Musikpädagogisches Handeln. Begriffe, Erscheinungsformen, politische Dimensionen: Music Education: Concepts, Practices, and Political Dimensions* (S. 70–85). Die Blaue Eule.
- Harrison, S. D. (2007). A perennial problem in gendered participation in music: What's happening to the boys? *British Journal of Music Education*, 24(3), 267–280. <https://doi.org/10.1017/S0265051707007577>
- Harrison, S. D. & Williams, J. (2019). Boys' singing voice change in adolescence. In G. F. Welch, D. M. Howard & J. Nix (Hrsg.), *The Oxford handbook of singing* (S. 533–549). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199660773.013.22>
- Hasselhorn, J. (2015). *Messbarkeit musikpraktischer Kompetenzen von Schülerinnen und Schülern: Entwicklung und empirische Validierung eines Kompetenzmodells*. Waxmann.
- Hasselhorn, J. & Lehmann, A. C. (2015). Leistungsheterogenität im Musikunterricht. Eine empirische Untersuchung zu Leistungsunterschieden im Bereich der Musikpraxis in Jahrgangsstufe 9. In A. Niessen & J. Knigge (Hrsg.), *Theoretische Rahmung und Theoriebildung in der musikpädagogischen Forschung* (S. 163–176). Waxmann.
- Hasselhorn, M., Decristan, J. & Klieme, E. (2019). Individuelle Förderung. In O. Köller, M. Hasselhorn, F. W. Hesse, K. Maaz & J. Schrader

- (Hrsg.), *Das Bildungswesen in Deutschland. Bestand und Potenziale* (S. 375–401). Verlag Julius Klinkhardt.
- Hasselhorn, M. & Gold, A. (2017). *Pädagogische Psychologie: Erfolgreiches Lernen und Lehren*. Kohlhammer Verlag. <http://ebookcentral.proquest.com/lib/fhnw/detail.action?docID=4902108>
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.
- Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. Routledge.
- Heinzer, M. & Hangartner, J. (2016). Politische Zuständigkeiten für die Volksschule in der Schweiz. In J. Hangartner & M. Heinzer (Hrsg.), *Gemeinden in der Schul-Governance der Schweiz: Steuerungskultur im Umbruch* (S. 31–53). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-13092-3_2
- Helbling, M. (2015). Brauchen Mädchen und Jungen gleichgeschlechtliche Lehrkräfte? Eine Überblicksstudie zum Zusammenhang des Lehrergeschlechts mit dem Bildungserfolg von Jungen und Mädchen. In *Enzyklopädie Erziehungswissenschaft Online* (S. 1–45). Beltz Juventa. <https://content-select.com/de/portal/media/view/55e564e5-2900-4c2d-b8a3-46c9b0dd2d03>
- Helfferich, C. (2019). Leitfaden- und Experteninterviews. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S. 669–686). Springer Fachmedien. https://doi.org/10.1007/978-3-658-21308-4_44
- Henning, H. (2013). Künstler in die Schulen? *Üben und Musizieren, 1*, 40–43.
- Henning, H. (2014). *Qualität in der vokalpädagogischen Praxis mit Kindern im Grundschulalter*. [Dissertation]. Hochschule für Musik Würzburg

- Herzog, M. (2018). Gemeinsames Musizieren fördert Inklusion!? Überlegungen zu Bedeutungskonstruktionen am Beispiel von inklusivem Klassenmusizieren. In Arbeitskreis Musikpädagogische Forschung (Germany), B. Clausen & S. Dressler (Hrsg.), *Soziale Aspekte des Musiklernens (Social aspects of music learning)* (S. 77–89). Waxmann.
- Hess, F. (2015). Das ist doch nichts für echte Kerle! In J. Wedl & A. Bartsch (Hrsg.), *Teaching Gender? Zum reflektierten Umgang mit Geschlecht im Schulunterricht und in der Lehramtsausbildung*. transcript Verlag. <https://doi.org/10.14361/9783839428221>
- Hess, F. (2018). *Gendersensibler Musikunterricht: Empirische Studien und didaktische Konsequenzen*. Springer Fachmedien.
- Hlawatsch, A. & Krickl, T. (2019). Einstellungen zu Befragungen. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S. 357–364). Springer Fachmedien. https://doi.org/10.1007/978-3-658-21308-4_24
- Hoene, S. & Thurman, B. (2011). *Umgang mit Heterogenität im Musikunterricht. Band 1: Grundlagen*. Institut für Qualitätsentwicklung an Schulen Schleswig-Holstein (IQSH).
- Holcomb, A. D. (2019). Assessment practices of American choral music educators. In T. S. Brophy (Hrsg.), *The Oxford handbook of assessment policy and practice in music education* (Bd. 2, S. 379–399). Oxford University Press.
- Holman Jones, S. (2005). Autoethnography: Making the personal political. In N. K. Denzin & Y. S. Lincoln (Hrsg.), *The SAGE handbook of qualitative research* (S. 763–791). SAGE Publications.
- Holman Jones, S., Adams, T. E. & Ellis, C. (2016a). Coming to know autoethnography as more than a method. In S. Holman Jones, T. E. Adams & C. S. Ellis (Hrsg.), *Handbook of autoethnography* (S. 17–48). Routledge.
- Holman Jones, S., Adams, T. E. & Ellis, C. S. (Hrsg.). (2016b). *Handbook of autoethnography*. Routledge.

- Hoos de Jokisch, B. (2014). Singen im Aufwind: Über die Wiederentdeckung des Singens im 21. Jahrhundert. *Üben & Musizieren*, 31(2), 6–10.
- Hornbach, C. M. & Taggart, C. C. (2005). The relationship between developmental tonal aptitude and singing achievement among Kindergarten, first-, second-, and third-grade students. *Journal of Research in Music Education*, 53(4), 322–331. <https://doi.org/10.1177/002242940505300404>
- Huber, J. (2008). *Standards im Musikunterricht: Pilotstudie* (Forschungsbericht Nr. 9; Forschungsbericht der Pädagogischen Hochschule Zentralschweiz, Hochschule Luzern). Institut für Lehren und Lernen (ILeL) der Pädagogischen Hochschule Zentralschweiz Hochschule Luzern, Institut Forschung & Entwicklung (F&E) der Hochschule Luzern – Musik.
- Huber, J. (2016). „Tasten- und Saitenhandwerker“ vs. „Forschergilde“: Der Diskurs um die „richtige“ Musiklehrer_innenbildung in der Deutschschweiz. In *Musikpädagogik und Erziehungswissenschaft* (Bd. 37, S. 45–58). Waxmann.
- Huber, J. (2023, in Vorbereitung). Des Schweizerers ‚Hang fürs Praktische‘ oder: Wie man sich Autorität im Diskurs um die Musikpädagogik erwirbt. In M.-A. Camp, B. Brabec de Mori & D. Klebe (Hrsg.), *Autoritätsbildungen in der Musik*. Chronos.
- Huber, J., Camp, M.-A., Blanchard, O., Chatelain, S., Joliat, F., Steiner, R., & Zurmühle, J. (Hrsg.) (2021). *Kulturen der Schulmusik in der Schweiz. Les cultures d l’enseignement musical à l’école en Suisse*. Chronos.
- Huff-Gackle, L. (1985). The young adolescent female voice (ages 11–15): Classification, placement, and development of tone. *The Choral Journal*, 25(8), 15–18.
- Hunziker, M. (2020). ‚Kinder müssen keinen Geschlechtstvorstellungen gerecht werden‘. Interview mit Diego Valsecchi von der Beratungsstelle Transgender Network Switzerland. *Das Heft*, 5, 31.
- Imthurn, G. (2020). Basic voice check: A diagnostic tool for the singing class. In T. S. Brophy & M. Haning (Hrsg.), *Advancing music edu-*

cation through assessment: Selected papers from the seventh international symposium on assessment in music education. GIA.

Institut Weiterbildung und Beratung PH FHNW, Arbeitsgruppe Schwerpunkt Unterrichtsentwicklung (2016). *Lernförderliches Beurteilen und Bewerten im Kontext des Lehrplans 21* [Grundlagenpapier. Unveröffentlichtes Papier].

Jank, B. (2013). Singen als Dialog zwischen menschlicher Kommunikation und musikalischer Bildung: Ausgangspunkte und Begründungen für das Musikprojekt ‚Belcantare Brandenburg‘. In B. Jank & J. Buschmann (Hrsg.), *Belcantare Brandenburg: Jedes Kind kann singen! Ein Singprojekt zur Fortbildung für musikunterrichtende Grundschullehrerinnen und Grundschullehrer im Land Brandenburg. Repräsentative Dokumentation* (S. 57–66). Universitätsverlag Potsdam.

Jank, B. (2017). Lieddidaktische Überlegungen zum Singen in der Grundschule. In J. A. Rodríguez-Quiles (Hrsg.), *Internationale Perspektiven zur Musik(lehrer)ausbildung in Europa* (S. 103–116). Universitätsverlag Potsdam.

Jank, B. & Bossen, A. (Hrsg.). (2017). *Musikarbeit im Kontext von Inklusion und Integration*. Universitätsverlag Potsdam.

Jank, B. & Buschmann, J. (Hrsg.). (2013). *Belcantare Brandenburg: Jedes Kind kann singen! Ein Singprojekt zur Fortbildung für musikunterrichtende Grundschullehrerinnen und Grundschullehrer im Land Brandenburg. Repräsentative Dokumentation*. Universitätsverlag Potsdam.

Jank, W. (Hrsg.). (2013). *Musik Didaktik: Praxishandbuch für die Sekundarstufe I und II* (5. Aufl.). Cornelsen Scriptor.

Jank, W. & Meyer, H. (2014). *Didaktische Modelle* (11. Aufl.). Cornelsen Scriptor.

Jaschke, A. C., Eggermont, L. H. P., Honing, H. & Scherder, E. J. A. (2013). Music education and its effect on intellectual abilities in children: A systematic review. *Reviews in the Neurosciences*, 24(6), 665–675. <https://doi.org/10.1515/revneuro-2013-0023>

Jedinger, A. & Michael, T. (2019). Interviewereffekte. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung*

- (S. 365–376). Springer Fachmedien. https://doi.org/10.1007/978-3-658-21308-4_25
- Johnson, D. (2020). The decline of singing in compulsory Swedish music education. In J. van der Sandt (Hrsg.), *Singing with children: International perspectives* (S. 227–240). Libreria Musicale Italiana.
- Jones, W. M. (1998). Investigation of William Vennard's light mechanism principle in the adolescent female voice [Review of *Investigation of William Vennard's light mechanism principle in the adolescent female voice*, von J. F. Bottoms]. *Bulletin of the Council for Research in Music Education*, 136, 60–62.
- Jordan, A.-K. (2012). *Entwicklung und Validierung eines Kompetenzmodells im Fach Musik: Wahrnehmen und Kontextualisieren von Musik*. Beltz.
- Kälin, P. (1976). Zur Situation der Musikerziehung in der Schweiz. *Schweizer Schule*, 63(12), 458–474. <https://doi.org/10.5169/seals-530071>
- Kalmbach, H.-J. (2010). Für die vokale Gleichberechtigung – Duale Sing-erziehung: Wie man die Singverweigerer in die Gemeinschaft zurückholt. In M. Fuchs (Hrsg.), *Wechselwirkungen zwischen Erwachsenen- und Kinderstimme* (S. 129–142). Logos Verlag.
- Kelle, U. & Kluge, S. (2010). *Vom Einzelfall zum Typus: Fallvergleich und Fallkontrastierung in der qualitativen Sozialforschung* (2. Aufl.). VS Verlag für Sozialwissenschaften.
- Keller, S. & König, F. (2017). *Kompetenzorientierter Unterricht mit Portfolio*. hep verlag.
- King, V. (2013). Jugend- und Adoleszenzforschung. In V. King & H.-C. Koller (Hrsg.), *Die Entstehung des Neuen in der Adoleszenz: Individuation, Generativität und Geschlecht in modernisierten Gesellschaften* (2. Auflage, S. 29–55). Springer VS. https://doi.org/10.1007/978-3-658-01350-9_2
- Kleinen, G. (2006). Komparative Studie zur Musikpädagogik in der Schweiz. In N. Knolle (Hrsg.), *Lehr- und Lernforschung in der Musikpädagogik* (S. 299–324). Die blaue Eule.

- Klieme, E. & Warwas, J. (2011). Konzepte der individuellen Förderung. *Zeitschrift für Pädagogik*, 57(6), 805–818.
- Konrad, U. (2019). „Flexibilisierung der Lehrperson“ – Entwicklung eines Designprinzips zwischen starren Strukturen und individuellen Bedürfnissen im Bandklassenunterricht. In A. Bikner-Ahsbahs & M. Peters (Hrsg.), *Unterrichtsentwicklung macht Schule* (S. 109–127). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-20487-7_7
- Kracauer, S. (1952). The challenge of qualitative content analysis. *Public Opinion Quarterly*, 16(4, Special Issue on International Communications Research), 632–642. <https://doi.org/10.1086/266427>
- Kraemer, R.-D. (2007). *Musikpädagogik – Eine Einführung in das Studium* (2. Aufl.). Wißner.
- Krauss, S., Lindl, A., Schilcher, A., Fricke, M., Göhring, A., Hofmann, B., Kirchhoff, P. & Mulder, R. H. (Hrsg.). (2017). *FALKO: Fachspezifische Lehrerkompetenzen. Dokumentation der Testinstrumente zur Erhebung des fachdidaktischen Wissens und des Fachwissens in allen FALKO-Disziplinen Deutsch, Englisch, Latein, Physik, Musik und Evangelische Religion sowie der Vignetten zur Erfassung pädagogischer Kompetenz*. Waxmann.
- Krebs, D. & Menold, N. (2019). Gütekriterien quantitativer Sozialforschung. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S. 489–504). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-21308-4_34
- Krippendorff, K. (1980). *Content analysis: An introduction to its methodology*. SAGE Publications.
- Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2. Aufl.). SAGE Publications.
- Krippendorff, K. (2019). *Content analysis: An introduction to its methodology* (4. Aufl.). SAGE Publications.
- Krippner, R. (2016). In quires and places where they sing. Ausbildungs- und Arbeitsstrukturen in der englischen Knabenchortradition. In M. Fuchs (Hrsg.), *Stimme – Leistung – Gesellschaft* (Bd. 10, S. 103–117). Logos Verlag.

- Krüger, D. & Riemeier, T. (2014). Die qualitative Inhaltsanalyse – eine Methode zur Auswertung von Interviews. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 133–145). Springer. https://doi.org/10.1007/978-3-642-37827-0_11
- Kuckartz, U. (2016). Typenbildung und typenbildende Inhaltsanalyse in der empirischen Sozialforschung. In M. W. Schnell, C. Schulz, U. Kuckartz & C. Dunger (Hrsg.), *Junge Menschen sprechen mit sterbenden Menschen: Eine Typologie* (S. 31–53). Springer Fachmedien. https://doi.org/10.1007/978-3-658-12317-8_2
- Kuckartz, U. (2018). *Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung* (4. Aufl.). Beltz Verlagsgruppe.
- Küstners, I. (2009). *Narrative Interviews. Grundlagen und Anwendungen* (2. Aufl.). VS Verlag für Sozialwissenschaften.
- Kunz, Y. (23. Mai 2015). BILDUNG: Werden Lehrer ungenügend vorbereitet? *Luzerner Zeitung*. <https://www.luzernerzeitung.ch/zentralschweiz/luzern/bildung-werden-lehrer-ungenuegend-vorbereitet-ld.100761>
- Lagerberg, D. (1975). *Kontext och funktion: Bidrag till innehållsanalysens teori och metod*. Fyris-Tr.
- Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. <https://doi.org/10.2307/2529310>
- Largo, R. H. (2010). *Kinderjahre: Die Individualität des Kindes als erzieherische Herausforderung* (19. Aufl.). Piper.
- Largo, R. H. (2017a). *Babyjahre: Entwicklung und Erziehung in den ersten vier Jahren* (Vollständig überarbeitete Neuauflage [18. Aufl.]). Piper.
- Largo, R. H. (2017b). *Das passende Leben: Was unsere Individualität ausmacht und wie wir sie leben können*. S. Fischer.
- Largo, R. H. & Czernin, M. (2017). *Jugendjahre: Kinder durch die Pubertät begleiten*. Piper.

- Lau, A. M. S. (2016). 'Formative good, summative bad?' – A review of the dichotomy in assessment literature. *Journal of Further and Higher Education*, 40(4), 509–525. <https://doi.org/10.1080/0309877X.2014.984600>
- Legrand, R. (2018). *Geschmack im Musikunterricht. Das strukturelle Selbstverständnis der Musikpädagogik zwischen Wertevermittlung und Wissenschaftlichkeit*. Springer Fachmedien.
- Lehmann, L. (2007). *Harmonisierung der Stufen- und Fächerprofile in der Lehrerinnen- und Lehrerbildung* (S. 27) [Expertenbericht]. Pädagogische Hochschule der Fachhochschule Nordwestschweiz.
- Lehmann-Wermser, A. (2002). Vom Verschwinden der Jungen aus der Musikdidaktik. *Zeitschrift für Kritische Musikpädagogik*, 1–21.
- Lehmann-Wermser, A. (2008). Singen im Unterricht zeitenübergreifend. In A. Niessen & A. Lehmann-Wermser (Hrsg.), *Aspekte des Singens: Ein Studienbuch* (S. 78–92). Wißner.
- Lehmann-Wermser, A. (2019). Assessment in German music education. In T. S. Brophy (Hrsg.), *The Oxford handbook of assessment policy and practice in music education* (Bd. 1, S. 233–252). Oxford University Press.
- Lehmann-Wermser, A. & Konrad, U. (2016). Design-Based Research als eine der Praxis verpflichtete, theoretisch fundierte Methode der Unterrichtsforschung und -entwicklung. Methodologische Grundlagen, dargestellt am Beispiel eines Forschungsprojektes im Bandklassen-Unterricht. In J. Knigge & A. Niessen (Hrsg.), *Musikpädagogik und Erziehungswissenschaft*. Waxmann.
- Leiprecht, S., Mattle, S., Oberhänsli, W. & Raas, A. (2017). *Aufbaureihen Heft 1: Ein praktischer Reiseführer in die Welt der Musik – 7. Klasse (3. Zyklus)*. Verlag Schweizer Singbuch.
- Linn, F. (2017). *Überzeugungen von Musiklehrenden zum Umgang mit Heterogenität im Musikunterricht*. universi – Universitätsverlag Siegen.
- Lohmeier, C. (2018). Zwischen „gone native“ und „eine von uns“: Reflexionen zu etischer und emischer Positionierung zum Forschungsfeld. In A. M. Scheu (Hrsg.), *Auswertung qualitativer Daten* (S. 29–39).

Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-18405-6>

Ludwig, P. H. (2003). Partielle Geschlechtertrennung – Enttäuschte Hoffnungen? Monoedukative Lernumgebungen zum Chancenausgleich im Unterricht auf dem Prüfstand. *Zeitschrift für Pädagogik*, 49(5), 640–656.

Lüdemann, J. & Otto, A. (Hrsg.) (2019). *Triangulation und Mixed-Methods: Reflexionen theoretischer und forschungspraktischer Herausforderungen*. Springer VS.

Lupatsch, J. & Hadjar, A. (2011). Determinanten des Geschlechterunterschieds im Schulerfolg: Ergebnisse einer quantitativen Studie aus Bern. In A. Hadjar (Hrsg.), *Geschlechtsspezifische Bildungsungleichheiten* (S. 177–202). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92779-4_8

Luthiger, H., Wilhelm, M., Wespi, C. & Wildhirt, S. (2018). *Kompetenzförderung mit Aufgabensets: Theorie – Konzept – Praxis*. hep.

Mahlert, U. (2017). Glück im Musizieren: Erscheinungsformen – Pädagogische Überlegungen. In M. Fuchs (Hrsg.), *Die Stimme im pädagogischen Alltag* (S. 11–22). Logos Verlag.

Malmberg, I. (2012). *Projektmethode und Musikunterricht*. LIT.

Manovski, M. P. (2014). *Arts-based research, autoethnography, and music education: Singing through a culture of marginalization*. Brill Sense.

Mantie, R. (2019). The philosophy of assessment in music education. In T. S. Brophy (Hrsg.), *The Oxford handbook of assessment policy and practice in music education* (Bd. 1, S. 33–55). Oxford University Press.

Markus, H. & Nurius, P. (1986). Possible selves. *The American Psychologist*, 41(9), 954–969. <https://doi.org/10.1037/0003-066X.41.9.954>

Marty, C. (2021). Singen oder Musik: Beweggründe zur Änderung einer Fachbezeichnung. In J. Huber, M.-A. Camp, O. Blanchard, S. Chatelain, F. Joliat, R. Steiner, & J. Zurmühle (Hrsg.), *Kulturen der Schulmusik in der Schweiz. Les cultures de l'enseignement musical à l'école en Suisse*. Chronos.

- Massoth, D. (2019). Technical issues related to computerized music performance assessment. In T. S. Brophy (Hrsg.), *The Oxford handbook of assessment policy and practice in music education* (Bd. 2, S. 855–876). Oxford University Press.
- Mattle, S. & Oberhäsli, W. (2018). *Aufbaureihen Heft 2: Ein praktischer Reiseführer in die Welt der Musik – 8.–9. Klasse* (3. Zyklus). Verlag Schweizer Singbuch.
- Mayring, P. (1982). *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. Deutscher Studienverlag.
- Mayring, P. (2015). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (12. Aufl.). Beltz.
- Mayring, P. & Brunner, E. (2006). Qualitative Textanalyse – Qualitative Inhaltsanalyse. In V. Flaker & T. Schmid (Hrsg.), *Von der Idee zur Forschungsarbeit: Forschen in Sozialarbeit und Sozialwissenschaft* (S. 453–462). Böhlau.
- Mayring, P. & Fenzl, T. (2019). Qualitative Inhaltsanalyse. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S. 633–648). Springer Fachmedien. https://doi.org/10.1007/978-3-658-21308-4_42
- McBride, N. R. & Palkki, J. (2020). Big boys don't cry (or sing) ... still? A modern exploration of gender, misogyny, and homophobia in college choral methods texts. *Music Education Research*, 22(4), 408–420. <https://doi.org/10.1080/14613808.2020.1784862>
- McGinnis, P. (2018). Engaging students in learning through assessment. *Science Scope*, 41(5), 1–1.
- Mecke, A.-C. (2007). *Mutantenstadt: Der Stimmwechsel und die deutsche Chorpraxis im 18. und 19. Jahrhundert*. Wissenschaftlicher Verlag Berlin.

- Mehr, S. A., Singh, M., Knox, D., Ketter, D. M., Pickens-Jones, D., Atwood, S., Lucas, C., Jacoby, N., Egner, A. A., Hopkins, E. J., Howard, R. M., Hartshorne, J. K., Jennings, M. V., Simson, J., Bainbridge, C. M., Pinker, S., O'Donnell, T. J., Krasnow, M. M. & Glowacki, L. (2019). Universality and diversity in human song. *Science*, 366(64–68). <https://doi.org/10.1126/science.aax0868>
- Meier, M. D. (2008). *Musikunterricht als Koedukation? Eine dreijährige Longitudinalstudie an einer ländlichen hessischen Gesamtschule*. Cuvillier Verlag.
- Merrill, J. (2020). Das Phänomen ‚Brummer‘ aus neurowissenschaftlicher Sicht. In M. Fuchs (Hrsg.), *Stimmen Hören – Potentiale Entwickeln – Störungen Behandeln* (S. 83–94). Logos Verlag.
- Millican, J. S. & Forrester, S. H. (2017). *Developing core teaching practices in music education* [Presentation]. 2017 Symposium on music teacher education, Minneapolis.
- Millican, J. S. & Forrester, S. H. (2018). Core practices in music teaching: A Delphi expert panel survey. *Journal of Music Teacher Education*, 27(3), 51–64. <https://doi.org/10.1177/1057083717736243>
- Millican, J. S. & Forrester, S. H. (2019). Music teacher rankings of selected core teaching practices. *Journal of Music Teacher Education*, 29(1), 86–99. <https://doi.org/10.1177/1057083719867682>
- Möller, S. (2012). Inhaltsanalyse. In B. Schäffer & O. Dörner (Hrsg.), *Handbuch Qualitative Erwachsenen- und Weiterbildungsforschung* (S. 381–394). Verlag Barbara Budrich.
- Mohr, A. (1997). *Handbuch der Kinderstimmbildung*. Schott.
- Mohr, A. (2008). Der Brummer. Das Phänomen. Hilfen. In *Lieder, Spiele, Kanons. Stimmbildung in Kindergarten und Grundschule*. Schott.
- Mohr, A. (2014). *Handbuch der Kinderstimmbildung*. Schott Music. <http://ebookcentral.proquest.com/lib/fhnw/detail.action?docID=4571977>
- Mohr, K. & Schäfer-Lembeck, H.-U. (Hrsg.). (2008). *Leistung im Musikunterricht: Beiträge der Münchner Tagung 2008*. Allitera Verlag.

- Morrison, S. (2005). Sure, I'm a singer! The empowering effect of vocal portfolios at the secondary school level. In K. Adams (Hrsg.), *The Phenomenon of Singing International Symposium V* (Bd. 5, S. 240–246).
- Muschg, A. (17. Dezember 2020). *Gredig direkt – Mit Schriftsteller Adolf Muschg* (U. Gredig) [Interview]. <https://www.srf.ch/play/tv/gredig-direkt/video/mit-schriftsteller-adolf-muschg?urn=urn:srf:video:aec13217-db8f-470f-b532-631ba44ae29b>
- Naidr, J., Zbořil, M. & Ševčík, K. (1965). Die pubertalen Veränderungen der Stimme bei Jungen im Verlauf von 5 Jahren. *Folia Phoniatrica et Logopaedica*, 17(1), 1–18. <https://doi.org/10.1159/000263019>
- Nethsinghe, R. (2012). The Influence of Informal Music Education in Teacher Formation: An Autoethnography. *Qualitative Report*, 17, 1–16.
- Neus, I. (2017). *Singen. Zentrale Begriffe, psychosoziale Wirkfunktionen und musikpädagogische Handlungsfelder: Eine interdisziplinäre Untersuchung*. LIT.
- Newlin, G. A. (2004). *The effects of part-work instruction on first grade part-singing acquisition and achievement* [Dissertation]. University of Hartford.
- Niessen, A. (2006). *Individuale Konzepte von Musiklehrern*. LIT.
- Niessen, A. (2008). Anmerkungen zum Singen aus sozialisationstheoretischer Perspektive. In A. Lehmann-Wermser & A. Niessen (Hrsg.), *Aspekte des Singens: Ein Studienbuch* (S. 145–156). Wißner.
- Niessen, A. (2013). Evaluation und Leistungsmessung. In W. Jank (Hrsg.), *Musik Didaktik: Praxishandbuch für die Sekundarstufe I und II* (5 Aufl., S. 141–144). Cornelsen Scriptor.
- Niessen, A. (2014). Individuale Konzepte von Musiklehrenden – eine qualitative Studie als Beitrag zur musikpädagogischen Grundlagenforschung. *Art Education Research*, 9, 1–9.
- Niessen, A. (2016). Zum Verhältnis von Individualkonzepten und (musik-) didaktischen Konzeptionen. In A. Lehmann-Wermser (Hrsg.), *Musikdidaktische Konzeptionen: Ein Studienbuch* (S. 145–156). Wißner.

- Niessen, A. (2017). Möglichkeiten von Interviews in musikpädagogischer Forschung. In M. L. Schulten & K. S. Lothwesen (Hrsg.), *Methoden empirischer Forschung in der Musikpädagogik: Eine anwendungsbezogene Einführung* (S. 103–120). Waxmann.
- Niessen, A., Knigge, J. & Vogt, J. (2014). „Forschung aus der Perspektive musikpädagogischer Praxis“ aus der Perspektive musikpädagogischer Forschung. *Zeitschrift für Kritische Musikpädagogik (ZfKM)*, 68–80.
- Nurjannah, I. & Siwi, S. M. (2017). Guidelines for analysis on measuring interrater reliability of nursing outcome classification. *International Journal of Research in Medical Sciences*, 5(4), 1169–1175. <https://doi.org/10.18203/2320-6012.ijrms20171220>
- Oberschmidt, J. (2016). Es wird wieder gesungen!? *Musikunterricht aktuell*, 4, 8–16.
- Oebelsberger, M. (2009). Singen ist Mädchensache. In T. Schweiger & T. Hascher (Hrsg.), *Geschlecht, Bildung und Kunst* (S. 79–90). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-91332-2_5
- Orgass, S. (2017). Ein System als Teil einer Theorie. Überlegungen zum Abgrenzungskriterium einer reflexionslogischen Wissenschaftstheorie der Musikpädagogik. In A. J. Cvetko & C. Rolle (Hrsg.), *Musikpädagogik und Kulturwissenschaft* (S. 119–136). Waxmann.
- Orton, A. & Pitts, S. E. (2019). Adolescent perceptions of singing: Exploring gendered differences in musical confidence, identity and ambition. *Music Education Research*, 21(1), 40–51. <https://doi.org/10.1080/14613808.2018.1507019>
- Oser, F. & Spychiger, M. (2005). *Lernen ist schmerzhaft: Zur Theorie des Negativen Wissens und zur Praxis der Fehlerkultur*. Beltz.
- Pachner, R. & Barth, V. (2008). *Vokalpädagogik: Theorie und Praxis des Singens mit Kindern und Jugendlichen* (3. Aufl.). Bosse.
- Panadero, E. & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, 9, 129–144. <https://doi.org/10.1016/j.edurev.2013.01.002>

- Patton, M. Q. (2015). *Qualitative research & evaluation methods integrating theory and practice* (Fourth Edition). SAGE.
- Pellegrino, K., Conway, C. M. & Russell, J. A. (2015). Assessment in performance-based secondary music classes. *Music Educators Journal*, 102(1), 48–55.
- Peterson, N. O. (2019). Ohne Sorge in den Stimmwechsel: Gesangspädagogischer Umgang mit der männlichen Stimme in der Entwicklung. In M. Fuchs (Hrsg.), *Gesangsstile – Stimmtechniken – Stimmgesundheit* (S. 25–40). Logos Verlag.
- Pezenburg, M. & Dyllick, N. (2018). Singen als didaktisches Handlungsfeld. In M. Dartsch, J. Knigge, A. Niessen, F. Platz & C. Stöger (Hrsg.), *Handbuch Musikpädagogik. Grundlagen – Forschung – Diskurse* (S. 294–302). Waxmann.
- Pfordresher, P. Q. & Demorest, M. (2020). Construction and validation of the Seattle singing accuracy protocol (SSAP): An automated online measure of singing accuracy. In F. A. Russo, B. Ilari & A. J. Cohen (Hrsg.), *The Routledge Companion to Interdisciplinary Studies in Singing, Volume I: Development* (S. 322–333). Routledge.
- Phillips, K. H. (1995). The changing voice: An albatross? *The Choral Journal*, 35(10), 25–27.
- Plenarversammlung der EDK. (2007). *Interkantonale Vereinbarung über die Harmonisierung der obligatorischen Schule (HarmoS-Konkordat)* (649.11). EDK.
- Ptok, M., Schwemmle, C., Iven, C., Jessen, M. & Nawka, T. (2006). Zur auditiven Bewertung der Stimmqualität. *HNO*, 54(10), 793–802. <https://doi.org/10.1007/s00106-005-1310-6>
- Puffer, G. & Hofmann, B. (2016). FALKO-M: Zur Konzeptualisierung des Professionswissens von Musiklehrkräften. In A. Niessen & J. Knigge (Hrsg.), *Musikpädagogik und Erziehungswissenschaft* (S. 107–120). Waxmann.
- Puffer, G. & Hofmann, B. (2017). FALKO-M: Entwicklung und Validierung eines Testinstruments zum domänenspezifischen Professionswissen von Musiklehrkräften. In S. Krauss, A. Lindl, A. Schilcher,

- M. Fricke, A. Göhring, B. Hofmann, P. Kirchhoff & R. H. Mulder (Hrsg.), *FALKO: Fachspezifische Lehrerkompetenzen. Dokumentation der Testinstrumente zur Erhebung des fachdidaktischen Wissens und des Fachwissens in allen FALKO-Disziplinen Deutsch, Englisch, Latein, Physik, Musik und Evangelische Religion sowie der Vignetten zur Erfassung pädagogischer Kompetenz*. Waxmann.
- Puffer, G. & Hofmann, B. (2021). LeHet Musik: Förderung professioneller Kompetenzen im Umgang mit Heterogenität in Singe-Situationen. In A. Hartinger, E. Matthes, Nett, U. Nett, E. Dresel & K. Peuschel (Hrsg.), *Lehrprofessionalität im Umgang mit Heterogenität: Theoretische Konzepte, Förderansätze, empirische Befunde*. Waxmann.
- Rädiker, S. & Kuckartz, U. (2019). *Analyse qualitativer Daten mit MAXQDA: Text, Audio und Video*. Springer VS.
- Rat für deutsche Rechtschreibung. (2019). *Bericht und Vorschläge der AG „Geschlechtergerechte Schreibung“ zur Sitzung des Rats für deutsche Rechtschreibung am 16.11.2018: Revidierte Fassung aufgrund des Beschlusses des Rats vom 16.11.2018*. Institut für Deutsche Sprache. https://www.rechtschreibrat.com/DOX/rfdr_2018-11-28_anlage_3_bericht_ag_geschlechterger_schreibung.pdf
- Reichertz, J. (2003). *Die Abduktion in der qualitativen Sozialforschung: Über die Entdeckung des Neuen*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-531-93163-0>
- Reichertz, J. (2013). *Die Abduktion in der qualitativen Sozialforschung: Über die Entdeckung des Neuen* (2. Aufl.). Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-531-93163-0>
- Richter, B. (2013). *Die Stimme: Grundlagen, künstlerische Praxis, Gesunderhaltung*. Henschel.
- Rieche, H., Fischer, A. K., Geißler, C., Eitel, A., Brunner, G. & Renkl, A. (2018). Wenn Schülerinnen und Schüler glauben, unmusikalisch zu sein. *Beiträge empirischer Musikpädagogik*, 9, 1–23.
- Riedmüller, S., Decoster, V. & Brockmann, M. (2010). Relevanz des Stimmfeldes für Diagnostik und Therapieevaluation. *Forum Logopädie*, 24(4), 11–15.

- Robinson, R. L. & Gesler, S. (2007). Middle School Singers: Turning Their Energy Into Wonderful Choirs! *The Choral Journal*, 48(3), 41–48.
- Röhner, J. & Schütz, A. (2016). *Psychologie der Kommunikation*. Springer Fachmedien. <https://doi.org/10.1007/978-3-658-10024-7>
- Rohrbach, K. (2019). *Im.puls. 2: Das Schweizer Musikbuch – Schülerbuch*. Helbling.
- Rost, D. H. (2008). Multiple Intelligenzen, multiple Irritationen. *Zeitschrift für Pädagogische Psychologie*, 22(2), 97–112. <https://doi.org/10.1024/1010-0652.22.2.97>
- Roulston, K. (2006). Mapping the possibilities of qualitative research in music education: A primer. *Music Education Research*, 8(2), 153–173. <https://doi.org/10.1080/14613800600779592>
- Rüttgers, P. (2016). *Von Rock 'n' Roll bis Hip-Hop: Geschlecht und Sexualität in Jugendkulturen*. Springer VS.
- Rutkowski, J. (1990). The measurement and evaluation of children's singing voice development. *The Quarterly*, 1(1–2), 81–95.
- Rutkowski, J. (2010). The measurement and evaluation of children's singing voice development. *Visions of Research in Music Education*, 16(1), 81–95.
- Rutkowski, J. (2019). Assessing singing voice development. In T. S. Brophy (Hrsg.), *The Oxford handbook of assessment policy and practice in music education* (Bd. 2, S. 629–651). Oxford University Press.
- Ruzicka, B. (2019). *Singstimme und Angst: Annäherungen aus stimmbildnerischer und psychologischer Perspektive* [Diplomarbeit]. Linz.
- Ryan, R. M. & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. <https://doi.org/10.1037/110003-066X.55.1.68>
- Sadolin, C. (2013). *Complete vocal technique* (Deutsche Ausgabe; 2. Version). Bosworth Music.

- Sammler, D. (2017). Neurophysiologische Aspekte des Singens. In M. Fuchs (Hrsg.), *Die Stimme im pädagogischen Alltag* (S. 187–196). Logos Verlag.
- Sandt, J. van der (Hrsg.). (2020). *Singing with children: International perspectives*. Libreria Musicale Italiana.
- Schaumberger, H. (2020a). Leistungsbeurteilung und Singen im Musikunterricht – eine Annäherung. In M. Schwarzbauer & K. Steinhäuser (Hrsg.), *„Nur“ Geschmackssache? Der Umgang mit kreativen Leistungen im Musik- und Kunstunterricht* (S. 222–247). LIT.
- Schaumberger, H. (2020b). *Professionalisierung von Kinder- und Jugendchorleitern*. LIT.
- Schaumberger, H., Wolf, M. & Sandt, J. van der (2020). The status of singing in schools: An overview of national curricula in Austria, England, Italy and South Tyrol (Italy). In J. van der Sandt (Hrsg.), *Singing with children: International perspectives* (S. 199–226). Libreria Musicale Italiana.
- Scheidegger, J. (2008). *Musiklehrausbildung in der Zentralschweiz 1960–2005*. Hochschule Luzern – Musik.
- Schlesiger, C. & Evers, S. (2005). Amusien – Eine Einführung in das Phänomen angeborener und erworbener cerebraler Störungen der Musikalität. *Musikphysiologie und Musikmedizin*, 12(1), 7–15.
- Schlömerkemper, J. (2010). *Konzepte pädagogischer Forschung: Eine Einführung in Hermeneutik und Empirie* (Bd. 3273). Klinkhardt.
- Schmidt, C. (2020). Formatives Assessment: Begriffsannäherung und theoretische Einordnung. In C. Schmidt, *Formatives Assessment in der Grundschule* (S. 7–23). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-26921-0_2
- Schmidt, H.-C. (1979). Versungen und vertan? Das Lied als problematischer Gegenstand der Musikpädagogik. *Zeitschrift für Literaturwissenschaft und Linguistik*, 34, 112–129.
- Schneider, M. C., McDonel, J. S. & DePasquale, C. A. (2019). Performance assessment and rubric design. In T. S. Brophy (Hrsg.), *The*

Oxford handbook of assessment policy and practice in music education (Bd. 1, S. 629–650). Oxford University Press.

Schnitzer, R. (2008a). *Singen ist klasse*. Schott.

Schnitzer, R. (2008b). *Singen ist klasse: Lehrerband mit DVD*. Schott.

Schreier, M. (2014). Varianten qualitativer Inhaltsanalyse: Ein Wegweiser im Dickicht der Begrifflichkeiten. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 15(1), 1–27. <https://doi.org/10.17169/fqs-15.1.2043>

Schulten, M. L. & Lothwesen, K. S. (2017). *Methoden empirischer Forschung in der Musikpädagogik. Eine anwendungsbezogene Einführung*. Waxmann.

Schumacher, R. (Hrsg.). (2009). *Pauken mit Trompeten: Lassen sich Lernstrategien, Lernmotivation und soziale Kompetenzen durch Musikunterricht fördern?* Bundesministerium für Bildung und Forschung (BMBF).

Scriven, M. (1967). *The methodology of evaluation*. Social Science Education Consortium.

Scriven, M. (1972). Die Methodologie der Evaluation. In C. Wulf (Hrsg.), *Evaluation: Beschreibung und Bewertung von Unterricht, Curricula und Schulversuchen*. (S. 60–91). Piper.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14. <https://doi.org/10.3102/0013189X015002004>

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1–22.

Siedenburg, I. (27. April 2015). *Populäre Musik, Gender und Musikpädagogik: Wirkungsweisen der Kategorie Geschlecht und Perspektiven für die Forschung* [Vorlesung]. Antrittsvorlesung an der Westfälischen Wilhelms-Universität, Münster.

Sipley, K. (1994). Improving vocal self-image and tone quality in adolescent girls: A study. *The Choral Journal*, 35(3), 35–38.

- SKBF. (2014). *Bildungsbericht Schweiz 2014*. Schweizerische Koordinationsstelle für Bildungsforschung.
- SKBF (2018). *Bildungsbericht Schweiz 2018*. Schweizerische Koordinationsstelle für Bildungsforschung.
- Snyder, M. (2015). Leaning into autoethnography: A review of Heewon Chang's autoethnography as method. *The Qualitative Report*, 20(2), 93–96.
- Sparber, M. & Frank, F. (1980). *Die Mutationsstimme im Gesangsunterricht*. Musikhaus Pan AG.
- Spital, H. (2004). *Stimmstörungen im Kindesalter: Ursachen, Diagnose, Therapiemöglichkeiten*. Georg Thieme Verlag.
- Spry, T. (2001). Performing autoethnography: An embodied methodological praxis. *Qualitative Inquiry*, 7(6), 706–732. <https://doi.org/10.1177/107780040100700605>
- Spychiger, M. (1997). Howard Gardners Konzept der musikalischen Intelligenz im Rahmen der Entwicklung der Intelligenzforschung. In J. Scheidegger (Hrsg.), *Persönlichkeitsentfaltung durch Musikerziehung* (S. 71–78). Musikedition Nepomuk.
- Spychiger, M. (2013). Das musikalische Selbstkonzept. Wer ich bin und was ich kann in der Musik. *üben & musizieren*, 6, 18–21.
- Spychiger, M. (2015). Lernpsychologische Perspektiven für eine grundschulspezifische Musikdidaktik. In *Musikdidaktik Grundschule: Theoretische Grundlagen und Praxisvorschläge* (S. 50–71). Helbling Verlag.
- Spychiger, M. (2017). Teaching towards the promotion of students' musical self-concept. In R. Girdzijauskiene & M. Stakelum (Hrsg.), *Creativity and innovation* (S. 133–146). Helbling.
- Spychiger, M., Olbertz, F. & Gruber, L. (2010). *Das musikalische Selbstkonzept. Konzeption des Konstrukts als mehrdimensionale Domäne und Entwicklung eines Messverfahrens* (Wissenschaftlicher Schlussbericht an den SNF 100013–116208).

- Stahlke Wall, S. (2016). Toward a moderate autoethnography. *International Journal of Qualitative Methods*, 15(1), 1–9. <https://doi.org/10.1177/1609406916674966>
- Stamann, C., Janssen, M. & Schreier, M. (2016). Qualitative Inhaltsanalyse – Versuch einer Begriffsbestimmung und Systematisierung. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 17(3), 1–17.
- Stern, T. (2020). Leistungsbeurteilung: Wie behindert oder fördert sie das Lernen? In M. Schwarzbauer & K. Steinhauser (Hrsg.), *„Nur“ Geschmackssache? Der Umgang mit kreativen Leistungen im Musik- und Kunstunterricht* (S. 11–24). LIT.
- Stevens, D. D. & Levi, A. J. (2013). *Introduction to rubrics: An assessment tool to save grading time, convey effective feedback, and promote student learning*. Stylus Publishing, LLC.
- Strübing, J. (2013). *Qualitative Sozialforschung. Eine komprimierte Einführung für Studierende*. Oldenbourg Wissenschaftsverlag.
- Strübing, J. (2017). Theoretischer Konservatismus und hegemonialer Gestus: Über ungute professionspolitische Spaltungen. Ein Kommentar auf Ronald Hitzlers „Zentrale Merkmale und periphere Irritationen...“. *Zeitschrift für Qualitative Forschung*, 18(1), 91–99. <https://doi.org/10.3224/zqf.v18i1.07>
- Strübing, J., Hirschauer, S., Ayaß, R., Krähnke, U. & Scheffer, T. (2018). Gütekriterien qualitativer Sozialforschung. Ein Diskussionsanstoß. *Zeitschrift für Soziologie*, 47(2), 83–100. <https://doi.org/10.1515/zfsoz-2018-1006>
- Sundberg, J. (2019). The singing voice. In S. Frühholz & P. Belin (Hrsg.), *The Oxford handbook of voice perception* (S. 117–142). Oxford University Press.
- Sweet, B. (2015). The adolescent female changing voice: A phenomenological investigation. *Journal of Research in Music Education*, 63(1), 70–88. <https://doi.org/10.1177/0022429415570755>
- Sweet, B. (2016). ‚Choral Journal‘ and the adolescent female changing voice. *The Choral Journal*, 56(9), 53–64.

- Sweet, B. (2018). Voice change and singing experiences of adolescent females. *Journal of Research in Music Education*, 66(2), 133–149. <https://doi.org/10.1177/0022429418763790>
- Tellisch, C. (2015). *Lehrer-Schüler-Interaktionen im Musikunterricht als Beitrag zur Menschenrechtsbildung*. Verlag Barbara Budrich.
- Terhart, E. (2014). Die Beurteilung von Schülern als Aufgabe des Lehrers: Forschungslinien und Forschungsergebnisse. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2. Aufl., S. 883–904). Waxmann.
- Terhart, E., Bennewitz, H. & Rothland, M. (Hrsg.) (2014). *Handbuch der Forschung zum Lehrerberuf* (2. Aufl.). Waxmann.
- Thompson, J. D. (2015). Towards cultural responsiveness in music instruction with black detained youth: An analytic autoethnography. *Music Education Research*, 17(4), 421–436. <https://doi.org/10.1080/14613808.2014.930117>
- Tosi, P. F. & Agricola, J. F. (1757). *Anleitung zur singkunst. Aus dem italiänischen des herrn Peter Franz Tosi mit erläuterungen und zusätzen von Johann Friedrich Agricola*. G. L. Winter.
- Trehub, S. E., Becker, J. & Morley, I. (2015). Cross-cultural perspectives on music and musicality. *Philosophical Transactions B Biological Sciences*, 370(1664), 1–9. <https://doi.org/10.1098/rstb.2014.0096>
- Turøy, A. K. W. (2018). Once a poor pitch singer, always a poor pitch singer? A bottom up study of factors that may support singing development. *British Journal of Music Education*, 35(1), 91–103. <https://doi.org/10.1017/S026505171700016X>
- Verbis. (2020). *MAXQDA | Die #1 Software für Qualitative & Mixed-Methods-Forschung*. MAXQDA. <https://www.maxqda.de/>
- Verordnung über die Laufbahnentscheide an der Volksschule (2009). SAR 421.352. https://gesetzessammlungen.ag.ch/app/de/texts_of_law/421.352

- Vries, P. de (2012). Autoethnography. In S. Delamont & A. Jones (Hrsg.), *Handbook of qualitative research in education* (S. 354–363). Edward Elgar.
- Wahl, D. (2013). *Lernumgebungen erfolgreich gestalten. Vom trägen Wissen zum kompetenten Handeln* (3. Aufl., mit Methodensammlung). Klinkhardt.
- Wallbaum, C. (2008). Zur ästhetisch-kulturellen Bildung mit Stimme. In A. Lehmann-Wermser & A. Niessen (Hrsg.), *Aspekte des Singens. Ein Studienbuch* (S. 93–110). Wißner.
- Weber, A. (2018). Stimme sein. In M. Fuchs (Hrsg.), *Beziehungssystem Stimme* (Bd. 12, S. 31–54). Logos Verlag.
- Weinert, F. E. (2001). Concept of competence: A conceptual clarification. In D. S. Rychen & L. Hersch Salganic (Hrsg.), *Defining and selecting key competencies* (S. 45–65). Hogrefe & Huber Publishers.
- Welch, G. F. (2019). Singing and vocal development. In G. F. Welch, D. M. Howard & J. Nix (Hrsg.), *The Oxford handbook of singing* (S. 509–532). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199660773.013.22>
- Welch, G. F., Howard, D. M. & Nix, J. (Hrsg.) (2019). *The Oxford handbook of singing*. Oxford University Press.
- Whitcomb, R. (1999). Writing rubrics for the music classroom. *Music Educators Journal*, 85(6), 26–32. <https://doi.org/10.2307/3399518>
- Wienhausen, S. (2016). Belt, Scream, Distortion – Parameter zur Leistungsbeurteilung von Kinder- und Jugendstimmen im Pop- und Musicalbereich? Über die Schwierigkeit der Beurteilung von Gesangsqualitäten im nicht-klassischen Bereich. In M. Fuchs (Hrsg.), *Stimme – Leistung – Gesellschaft* (S. 119–124). Logos Verlag.
- Williams, B. B. (1990). *An investigation of selected female singing- and speaking-voice characteristics through comparison of a group of pre-menarcheal girls to a group of post-menarcheal girls* [Dissertation]. University of North Texas.

- Williams, J., Welch, G. F. & Howard, D. M. (2021). Which sung pitch range is best for boys during voice change? *Journal of Voice*, 35(4), 581–588. <https://doi.org/10.1016/j.jvoice.2019.12.002>
- Willis, E. C. & Kenny, D. T. (2011). Voice training and changing weight: Are they reflected in speaking fundamental frequency, voice range, and pitch breaks of 13-year-old girls? A longitudinal study. *Journal of Voice*, 25(5), e233–e243. <https://doi.org/10.1016/j.jvoice.2010.06.004>
- Wilson, R. J. (1996). *Assessing students in classrooms and schools*. Allyn & Bacon Canada.
- Winter, F. (2015). *Lerndialog statt Noten: Neue Formen der Leistungsbeurteilung*. Beltz.
- Wirtz, M. & Kutschmann, M. (2007). Analyse der Beurteilerübereinstimmung für kategoriale Daten mittels Cohens Kappa und alternativer Maße. *Die Rehabilitation*, 46(6), 370–377. <https://doi.org/10.1055/s-2007-976535>
- Wise, K. J. & Sloboda, J. A. (2008). Establishing an empirical profile of self-defined “tone deafness”: Perception, singing performance and self-assessment. *Musicae Scientiae*, 12(1), 3–26. <https://doi.org/10.1177/102986490801200102>
- Xie, Q. (2013). *Agree or disagree? A demonstration of an alternative statistic to Cohen’s Kappa for measuring the extent and reliability of agreement between observers*. MacroSys, LLC. https://nces.ed.gov/FCSM/pdf/J4_Xie_2013FCSM.pdf
- Ziegelbauer, S. & Gläser-Zikuda, M. (Hrsg.). (2016). *Das Portfolio als Innovation in Schule, Hochschule und LehrerInnenbildung: Perspektiven aus Sicht von Praxis, Forschung und Lehre*. Verlag Julius Klinkhardt.
- Zinnecker, J. (2000). Pädagogische Ethnographie. *Zeitschrift für Erziehungswissenschaft*, 3(3), 381–400. <https://doi.org/10.1007/s11618-000-0039-y>
- Zwicky, S. (23. Februar 2016). Stimmentwicklung und Hormone. *Schweizer Musikzeitung*. <https://www.musikzeitung.ch/de/basis/smm/2016/03/Stimmentwicklungund-Hormone.html>

G ANHANG

Der Weblink und der QR-Kode führen zu einem öffentlich zugänglichen Ordner. <https://drive.switch.ch/index.php/s/rhFS0afwomVzpcb>

Passwort: *imthurn-23*

Folgende Dokumente sind zu finden:

- Leitfaden Interview
- Leitfaden Transkription
- Leitfaden Kodierung
- Detaillierter Ablauf der QIA
- Kodierhandbuch
- Eingereichte Dissertation
- Das Beurteilungsraster von Herrn Leimbacher (10)



Das Datenmaterial dieser Dissertation ist als MAXQDA-Datei verfügbar. Die Audiodateien sind nicht anonymisiert und deshalb nicht öffentlich verfügbar. Folgende Inhalte sind in der MAXQDA-Datei inkludiert:

- 11 Transkriptionen der Interviews
- Einzelne Beurteilungsraster
- QIA
- Forschungsnotizen

Überblick Bildungssystem der Schweiz

Der Überblick ist auf Seite 2 unter dieser Web-Adresse einzusehen:

https://www.skbf-csre.ch/fileadmin/files/pdf/bildungsberichte/2018/Bildungsbericht_Schweiz_2018.pdf

Kontakt:

Gabriel Imthurn

gabriel.imthurn@fhnw.ch



Empirische Forschung zur Musikpädagogik
hrsg. von Prof. Dr. Andreas Lehmann-Wermser (Universität Bremen)

Kristin Sander-Steinert

Anbahnung ästhetischer Erfahrung im Musikunterricht

Möglichkeiten und Grenzen beim Einsatz des iPads

Bd. 11, 2022, 306 S., 34,90 €, br., ISBN 978-3-643-15271-8

Joana Grow

Komponieren im Musikunterricht der Grundschule

Bd. 10, 2018, 854 S., 94,90 €, br., ISBN 978-3-643-13936-8

Marc Godau

Gemeinsam allein: Klassenmusizieren mit Populärer Musik

Eine systemisch-konstruktivistische Grounded Theory über Prozesse selbstständigen Lernens von Gruppen mit informellen Lernmethoden im schulischen Musikunterricht

Bd. 9, 2017, 480 S., 64,90 €, br., ISBN 978-3-643-13777-7

Inka Neus

Singen

Zentrale Begriffe, psychosoziale Wirkfunktionen und musikpädagogische Handlungsfelder. Eine interdisziplinäre Untersuchung

Bd. 8, 2017, 248 S., 34,90 €, br., ISBN 978-3-643-13625-1

Elias Zill

Den eignen Ohren folgen

Musikalisch-ästhetische Erfahrungen im Kontext produktionsorientierter Schulprojekte

Bd. 7, 2016, 342 S., 39,90 €, br., ISBN 978-3-643-13026-6

Heike Gebauer

Kognitive Aktivierung im Musikunterricht

Eine qualitative Videostudie

Bd. 6, 2016, 456 S., 49,90 €, br., ISBN 978-3-643-12622-1

Natalia Ardila-Mantilla

Musiklernwelten erkennen und gestalten

Eine qualitative Studie über Musikschularbeit in Österreich

Bd. 5, 2016, 558 S., 34,90 €, br., ISBN 978-3-643-50555-2

Thomas Busch

„Was, glaubst Du, kannst Du in Musik?“

Musikalische Selbstwirksamkeitserwartungen und ihre Entwicklung zu Beginn der Sekundarstufe I

Bd. 4, 2013, 352 S., 34,90 €, br., ISBN 978-3-643-11787-8

Susanne Naacke

Eine Schule auf dem Weg – Gelingende Schulentwicklung mit Chor- und Bläserklassen

Eine qualitative Fallstudie

Bd. 3, 2011, 296 S., 29,90 €, br., ISBN 978-3-643-11237-8

LIT Verlag Berlin – Münster – Wien – Zürich – London

Auslieferung Deutschland / Österreich / Schweiz: siehe Impressumseite

Untersucht wird die Förder- und Beurteilungspraxis von Schweizer Musiklehrkräften. Unter Berücksichtigung internationaler Studien zur adoleszenten Stimmentwicklung dient als Heuristik für die Analyse das Konzept eines ganzheitlichen Assessments, verstanden als Leistungsmessung im Interesse des Lernens. Sichtbar werden dabei sowohl intuitiv entwickelte Förderstrategien als auch Defizite im Umgang mit der Leistungsmessung. Die Studie schließt mit Impulsen für zukünftige Entwicklungen im Singunterricht.

Dr. Gabriel Imthurn ist seit 2022 Professor für Musikpädagogik an der Pädagogischen Hochschule der Nordwestschweiz. Sein Arbeitsschwerpunkt liegt im Bereich Singen in der Sekundarstufe.

