

Weber, Jannis

Mathematische Modellbildung und Videoanalyse zum Lernen der Newtonschen Dynamik im Vergleich

Berlin : Logos Verlag 2022, 381 S. - (Studien zum Physik- und Chemielernen; 345) - (Dissertation, Universität Frankfurt, 2022)



Quellenangabe/ Reference:

Weber, Jannis: Mathematische Modellbildung und Videoanalyse zum Lernen der Newtonschen Dynamik im Vergleich. Berlin : Logos Verlag 2022, 381 S. - (Studien zum Physik- und Chemielernen; 345) - (Dissertation, Universität Frankfurt, 2022) - URN: urn:nbn:de:0111-pedocs-334366 - DOI: 10.25656/01:33436; 10.30819/5566

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-334366>

<https://doi.org/10.25656/01:33436>

in Kooperation mit / in cooperation with:



<https://www.logos-verlag.de>

Nutzungsbedingungen

Dieses Dokument steht unter folgender Creative Commons-Lizenz: <http://creativecommons.org/licenses/by-nc-nd/4.0/deed.de> - Sie dürfen das Werk bzw. den Inhalt unter folgenden Bedingungen vervielfältigen, verbreiten und öffentlich zugänglich machen: Sie müssen den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen. Dieses Werk bzw. dieser Inhalt darf nicht für kommerzielle Zwecke verwendet werden und es darf nicht bearbeitet, abgewandelt oder in anderer Weise verändert werden.

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-nc-nd/4.0/deed.en> - You may copy, distribute and transmit, adapt or exhibit the work in the public as long as you attribute the work in the manner specified by the author or licensor. You are not allowed to make commercial use of the work or its contents. You are not allowed to alter, transform, or change this work in any other way.

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

Studien zum Physik- und Chemielernen

M. Hopf, H. Niedderer, M. Ropohl, E. Sumfleth [Hrsg.]

345

Jannis Weber

Mathematische Modellbildung und Videoanalyse zum Lernen der Newtonschen Dynamik im Vergleich



λογος

Studien zum Physik- und Chemielernen

Herausgegeben von Martin Hopf, Hans Niedderer, Mathias Ropohl und Elke Sumfleth

Diese Reihe im Logos Verlag Berlin lädt Forscherinnen und Forscher ein, ihre neuen wissenschaftlichen Studien zum Physik- und Chemielernen im Kontext einer Vielzahl von bereits erschienenen Arbeiten zu quantitativen und qualitativen empirischen Untersuchungen sowie evaluativ begleiteten Konzeptionsentwicklungen zu veröffentlichen. Die in den bisherigen Studien erfassten Themen und Inhalte spiegeln das breite Spektrum der Einflussfaktoren wider, die in den Lehr- und Lernprozessen in Schule und Hochschule wirksam sind.

Die Herausgeber hoffen, mit der Förderung von Publikationen, die sich mit dem Physik- und Chemielernen befassen, einen Beitrag zur weiteren Stabilisierung der physik- und chemiedidaktischen Forschung und zur Verbesserung eines an den Ergebnissen fachdidaktischer Forschung orientierten Unterrichts in den beiden Fächern zu leisten.

Martin Hopf, Hans Niedderer, Mathias Ropohl und Elke Sumfleth

Studien zum Physik- und Chemielernen

Band 345

Jannis Weber

**Mathematische Modellbildung und
Videoanalyse zum Lernen der
Newtonschen Dynamik im Vergleich**

Logos Verlag Berlin



Studien zum Physik- und Chemielernen

Martin Hopf, Hans Niedderer, Mathias Ropohl und Elke Sumfleth [Hrsg.]

Bibliografische Information der Deutschen Nationalbibliothek

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



© Copyright Logos Verlag Berlin GmbH 2022

Alle Rechte vorbehalten.

ISBN 978-3-8325-5566-5

ISSN 1614-8967

Logos Verlag Berlin GmbH
Georg-Knorr-Str. 4, Geb. 10
D-12681 Berlin

Tel.: +49 (0)30 / 42 85 10 90

Fax: +49 (0)30 / 42 85 10 92

<https://www.logos-verlag.de>

Mathematische Modellbildung und Videoanalyse zum Lernen der Newtonschen Dynamik im Vergleich

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich Physik
der Johann Wolfgang Goethe-Universität
in Frankfurt am Main

von
Jannis Valentin Weber
aus Wiesbaden

Frankfurt 2022

vom Fachbereich Physik der
Johann Wolfgang Goethe-Universität als Dissertation angenommen.

Dekan: Prof. Dr. Harald Appelshäuser

Gutachter: Prof. Dr. Thomas Wilhelm
Prof. Dr. Jochen Kuhn
Prof. Dr. Volkhard Nordmeier

Datum der Disputation: 30.06.2022

Für Jarno

Kurzzusammenfassung

Die zentralen Aussagen der Newtonschen Mechanik sind für Lernende erfahrungsgemäß nur schwierig zu erlernen. Hartnäckige, mit der physikalischen Theorie unvereinbare Schülervorstellungen erschweren Verbesserungen im Konzeptverständnis. Ein Grund dafür könnte der Fokus auf Idealsituationen im Unterricht sein. Lernende haben Schwierigkeiten, sich in diese Idealsituationen hineinzusetzen und es besteht die Gefahr einer Kluft zwischen dem Unterricht und den real beobachtbaren Bewegungen im Alltag. Komplexe Alltagsbewegungen sind vor allem mathematisch kompliziert. Helfen kann hier Computereinsatz im Unterricht.

Die Studie im Prä-Post-Design mit $N = 274$ Schüler*innen der elften Jahrgangsstufe (nach G9) untersucht zwei unterschiedliche Arten, wie komplexe Bewegungen mit Reibung mithilfe des Computers (mathematische Modellbildung und Videoanalyse) diskutiert werden können. Dabei wird die Wirksamkeit eines Unterrichts mit Computereinsatz untersucht, mit dem die Newtonsche Dynamik nach dem schulischen Unterricht vertieft wird. Dazu wurden äquivalente Interventionen gestaltet, die sich zwischen den Gruppen nur in der Art des Computereinsatzes unterscheiden. In vier verschiedenen Aufgaben arbeiten die Schüler*innen nach einer Durchführung des jeweiligen Experiments in Zweiergruppen. Der eingesetzte Test basiert auf bekannten Tests und eigenen Items und wurde für diese Studie erstellt und pilotiert. Zudem wurden während einer Arbeitsphase von $N = 45$ Schüler*innen Bildschirmvideos mit Tonaufnahmen der Gespräche aufgenommen, um die Arbeitsweise analysieren zu können.

Im Konzeptverständnis der Dynamik zeigt sich ein signifikanter Unterschied zwischen Vor- und Nachtest mit großer Effektstärke in beiden Gruppen. Es gibt keinen Unterschied zwischen den beiden Varianten im gesamten Lernzuwachs. In Items zum ersten Newtonschen Gesetz mit Kräftekompensation verbessert sich die Gruppe der Modellbildung allerdings stärker. Dies könnte mit der stärkeren Reduktion der Schülervorstellung in der Gruppe der Modellbildung zusammenhängen, dass eine Kraft in Bewegungsrichtung wirken muss. Gleichzeitig zeigt sich in den angefertigten Aufnahmen der Schülergespräche, dass die Gruppe der Modellbildung diese Vorstellung in der Arbeitsphase häufiger aktiviert. Das Modellieren führt zu einem Explizieren der Schülervorstellung, welche dann durch den Vergleich von Modell und Messdaten durch die Software widerlegt wird. Dies reduziert die Schülervorstellung stärker. Die Gruppe der Modellbildung verbessert ihr Modellverständnis durch das Modellieren. Weitere Unterschiede (z. B. im Cognitive Load und affektiven Merkmalen) werden diskutiert.

Eine Mehrebenenanalyse zeigt, dass der Vortest, das Interesse an theoretischen Zusammenhängen in der Physik, die kognitive Belastung und die Physiknote das Nachtestergebnis beeinflussen. Zudem zeigt ein Vergleich von erfolgreichen und weniger erfolgreichen Gruppen, dass die Phase der Ergebnissicherung und ein Zurückgreifen auf Modell bzw. Messdaten bei der Formulierung von Ergebnissen wichtig für den Lernerfolg sein könnten.

Abstract

Newtons laws of motion are known to be difficult to understand for students. Many persistent preconceptions exist, that are incompatible to the physical theory. Those preconceptions hinder the progress in conceptual understanding of Newtons laws. One reason for the persistence of those conceptions is the focus on idealized situation in physics class. Students have difficulties with those situations and a gap between physics class and the real world can arise. Complex motions from everyday life are too mathematically difficult for physics lessons in school. The computer as a tool can help to bridge that gap.

This pre-post-design study with $N = 274$ students from 11th grade of German (Hessian) schools investigates two different ways of using the computer to discuss complex motions with friction regarding their efficacy in improving the conceptual understanding of Newtons first two laws after the school lessons. For that, two equivalent interventions were designed that only differ in the way the computer is used. In each intervention, four different experiments are discussed. The students always work in groups of two after the respective experiment is shown in the plenum. The test, which was used to measure conceptual understanding and other variables, is partly based on known tests. Other parts of the test were created and piloted for this study. Additionally, screen recordings in combination with the conversations of $N = 45$ students were used to investigate, among other things, the way students use the software.

A significant difference with large effect size between conceptual understanding in the pre-test and post-test was found in both cohorts. A comparison showed no difference between computational modelling and video motion analysis in the overall conceptual understanding. However, a significant difference in items regarding Newtons first law was found. The computational modelling cohort improved significantly more. It was also found that the common preconception, that a force in the direction of motion is necessary, was reduced further in the computational modelling cohort. An analysis of the screen recordings reveals that this preconception is expressed more often in that group as well. This leads to the conclusion that modelling forces students to activate the preconception which is then disproved by the comparison of the model and the data in the software. This cognitive conflict leads to a bigger reduction of the preconception which in turn leads to a larger improvement in the items regarding Newtons first law. The computational modelling cohort also improves in model understanding. More differences (e.g. cognitive load and different affective variables) were found and are being discussed.

A linear hierarchical model shows that the pre-test score, the interest in theoretical relations in physics, the cognitive load and the physics grade influence the post-test score. Additionally, a comparison between more successful and less successful pairs shows that the phase of discussing the results could be an important factor for the learning gain. The more successful pairs also used the data or model more often to derive end results.

Inhaltsverzeichnis

KURZZUSAMMENFASSUNG	5
ABSTRACT	7
INHALTSVERZEICHNIS	9
1 EINLEITUNG	15
<u>I. THEORETISCHER HINTERGRUND</u>	<u>17</u>
2 FACHLICHE GRUNDLAGEN DER KLASSISCHEN MECHANIK.....	19
2.1 <i>Kinematik</i>	19
2.2 <i>Dynamik</i>	20
2.3 <i>Elementarisierungsentscheidungen in der Schule</i>	22
3 SCHÜLERVORSTELLUNGEN	27
3.1 <i>Generelle Aspekte von Schülervorstellungen</i>	27
3.2 <i>Schülervorstellungen zur Mechanik</i>	28
3.3 <i>Kontextabhängigkeit von Schülervorstellungen</i>	34
3.4 <i>Conceptual Change</i>	35
3.5 <i>Umgang mit Schülervorstellungen</i>	37
4 LERNEN MIT NEUEN MEDIEN	41
4.1 <i>Cognitive Load Theory</i>	42
4.2 <i>Cognitive Theory of Multimedia Learning</i>	44
4.3 <i>Design-Empfehlungen für digitale Lernumgebungen</i>	44
5 MODELLE UND MODELLIEREN	47
5.1 <i>Modelle in der Physik</i>	47
5.2 <i>Modellbildung</i>	49
5.3 <i>Zusammenspiel von Modell und Experiment</i>	50
5.4 <i>Mathematische Modellbildung</i>	51
5.4.1 <i>Definition</i>	51
5.4.2 <i>Numerische Verfahren</i>	51
5.4.3 <i>Arten der Mathematischen Modellbildung</i>	55
5.4.4 <i>Gründe für den Einsatz von mathematischer Modellbildung</i>	64

5.4.5	Forschungsergebnisse.....	66
5.4.6	Heutige Nutzung von mathematischer Modellbildung in Schulen.....	72
5.4.7	Zusammenfassung.....	72
6	DATENERHEBUNG UND -AUSWERTUNG IM PHYSIKUNTERRICHT	75
6.1	<i>Allgemeines zur Datenerhebung im Physikunterricht.....</i>	75
6.2	<i>Videoanalyse von Bewegungen</i>	76
6.2.1	Definition	76
6.2.2	Entwicklung der Videoanalysesoftware	76
6.2.3	Gründe für den Einsatz von Videoanalyseprogrammen im Physikunterricht	78
6.2.4	Forschungsergebnisse.....	79
6.2.5	Heutige Nutzung von Videoanalysesoftware in Schulen	85
6.2.6	Zusammenfassung.....	86
II.	<u>FORSCHUNGSANSATZ.....</u>	89
7	EINORDNUNG DER ARBEIT UND FORSCHUNGSFRAGEN	91
7.1	<i>Einordnung der Arbeit in den aktuellen Forschungsstand.....</i>	91
7.2	<i>Forschungsfragen.....</i>	91
7.2.1	Quantitativer Teil	91
7.2.2	Qualitativer Teil	93
8	ANLAGE UND DURCHFÜHRUNG DER EMPIRISCHEN STUDIE.....	95
8.1	<i>Vorstudie zur Software im Bereich der mathematischen Modellbildung und Videoanalyse</i>	95
8.1.1	„Newton-II“.....	95
8.1.2	„measure dynamics“.....	97
8.2	<i>Allgemeine Design-Entscheidungen auf Basis der bekannten Theorie.....</i>	98
8.3	<i>Konkrete Ausgestaltung der Studie</i>	98
8.3.1	Stichprobe	99
8.3.2	Interventionen	100
8.4	<i>Testinstrument</i>	109
8.4.1	Vorhandene Tests zur Mechanik.....	109
8.4.2	Entwicklung des Testinstruments.....	110
8.4.3	Pilotierung des Testinstruments	113

8.4.4	Gütekriterien des Testinstruments.....	130
8.5	<i>Lehrerfragebogen</i>	136
8.6	<i>Bildschirmvideos</i>	137
8.6.1	Auswahl des Versuchs für die Aufnahme der Bildschirmvideos.....	137
8.6.2	Qualitative Inhaltsanalyse.....	137
III.	EVALUATION DER STUDIE	147
9	ERGEBNISSE ZUM KONZEPTVERSTÄNDNIS.....	149
9.1	<i>Power Analysis</i>	150
9.2	<i>Wirksamkeit</i>	151
9.2.1	Mathematische Modellbildung.....	152
9.2.2	Videoanalyse.....	167
9.3	<i>Vergleich der Interventionen</i>	177
9.3.1	Kontrollvariablen.....	177
9.3.2	Nachtestergebnisse.....	180
9.3.3	Vortestergebnisse.....	181
9.3.4	Absoluter Lernzuwachs.....	183
9.3.5	Lernzuwachs nach Terzilen.....	186
9.3.6	Kontrolle der Vortestergebnisse.....	188
9.3.7	Unterscheidung zwischen Geschlechtern.....	195
9.3.8	Hierarchisches lineares Modell.....	199
9.3.9	Unterschiede in den einzelnen Inhaltsbereichen.....	225
10	ERGEBNISSE ZU WEITEREN QUANTITATIVEN TESTS.....	229
10.1	<i>Modellverständnis</i>	229
10.2	<i>Einschätzung der Lernenden</i>	231
10.3	<i>Cognitive Load</i>	233
10.4	<i>Theoretisierende vs. Experimentierende</i>	235
10.5	<i>Schülervorstellungen</i>	240
10.6	<i>Auswertung des Lehrerfragebogens</i>	244
10.6.1	Computereinsatz im Unterricht.....	244
10.6.2	Bewertung der Items.....	246

10.6.3	Konzeptverständnis.....	252
10.7	<i>Zusammenfassung zum quantitativen Teil.....</i>	254
11	ERGEBNISSE DES QUALITATIVEN TEILS DER STUDIE	257
11.1	<i>Schwierigkeiten</i>	257
11.1.1	Inhaltliche Schwierigkeiten	257
11.1.2	Schwierigkeiten mit der Software.....	261
11.2	<i>Argumentationsrichtung.....</i>	266
11.3	<i>Schülervorstellungen.....</i>	270
11.4	<i>Vorgehensweise.....</i>	272
11.4.1	Gruppe Modellbildung.....	272
11.4.2	Gruppe Videoanalyse.....	275
11.5	<i>Zusammenhänge mit dem Lernerfolg.....</i>	276
11.5.1	Gruppe Modellbildung.....	277
11.5.2	Gruppe Videoanalyse.....	280
11.5.3	Gesamtbetrachtung	284
11.6	<i>Zusammenfassung zum qualitativen Teil.....</i>	285
12	KOMBINATION DER ERGEBNISSE.....	287
12.1	<i>Verbesserung der Ergebnisse im ersten Newtonschen Gesetz</i>	287
12.2	<i>Prädiktoren für einen erfolgreichen Einsatz der genutzten Software</i>	288
12.3	<i>Argumentationsrichtungen.....</i>	288
IV.	<u>ZUSAMMENFASSUNG</u>	291
13	ZUSAMMENFASSUNG UND DISKUSSION	293
13.1	<i>Wirksamkeit der durchgeführten Interventionen.....</i>	294
13.2	<i>Vergleich der durchgeführten Interventionen</i>	295
13.3	<i>Einflussfaktoren auf die Änderung des Konzeptverständnisses</i>	296
13.4	<i>Weitere Beobachtungen aus den aufgezeichneten Schüleraktivitäten.....</i>	297
13.5	<i>Diskussion der Ergebnisse</i>	299
13.6	<i>Fazit.....</i>	300
14	SUMMARY AND DISCUSSION	303
14.1	<i>Effect of the interventions.....</i>	304

14.2	<i>Comparison of the two interventions.....</i>	305
14.3	<i>Predictors of learning gain in conceptual understanding.....</i>	306
14.4	<i>Further findings from the screen recordings and conversations</i>	307
14.5	<i>Discussion</i>	308
14.6	<i>Conclusion.....</i>	309
15	DANKSAGUNG	311
16	LITERATURVERZEICHNIS.....	313
17	ANHANG	345
17.1	<i>Vortest</i>	345
17.2	<i>Nachtest.....</i>	355
17.3	<i>Kodiermanual.....</i>	365

1 Einleitung

Newtons Principia (Newton, 1687) wird von vielen Wissenschaftler*innen als eines der wichtigsten Werke in der Geschichte der Wissenschaft angesehen (Guicciardini, 2003) und bildet den Grundstein der klassischen Physik. In diesem Werk beschreibt Newton u. a. drei Gesetze, die unser tägliches Leben steuern und Grundlage für viele physikalische, auf der klassischen Mechanik aufbauende Theorien sind. Bewegungen unseres Alltags bis hin zu Bewegungen von Planeten waren mit Newtons Bewegungsgesetzen erstmals erklärbar (Feynman et al., 2011). Obwohl die klassische Mechanik in manchen Bereichen, wie hohen Geschwindigkeiten oder kleinen Größenordnungen, durch andere Theorien abgelöst wurde, ist sie nach wie vor für die Beschreibung von alltäglichen Bewegungen geeignet. Newtons Aussagen zur Mechanik lassen sich auf wenige Gesetze (oder Axiome) reduzieren, mit denen jede Art von Bewegung beschrieben, erklärt und vorausgesagt werden kann. Im Schulunterricht spielt die Newtonsche Mechanik dementsprechend eine berechtigterweise große Rolle und wird normalerweise in die Bereiche Kinematik und Dynamik unterteilt.

Ernüchternd sind wiederum die Ergebnisse, die der Schulunterricht in diesem Bereich erzielt. Die vermeintlich einfachen Aussagen der Newtonschen Gesetze sind für Schüler*innen nur sehr schwer zu erlernen. Schülervorstellungen sind im Bereich der Mechanik gut erforscht und besonders hartnäckig (Schecker et al., 2018). Diese Hartnäckigkeit ist u. a. darin begründet, dass Schüler*innen ihr gesamtes Leben Erfahrungen mit Bewegungen und Kräften machen und sich so gewisse Vorstellungen festigen können, die physikalisch nicht tragfähig sind. Diese Diskrepanz zwischen dem Ziel des Physikunterrichts und den bei Schüler*innen beobachteten Leistungen motiviert zu Forschungen über das Lernen von Mechanik.

Wenn Lernende selbst nach der Aufgabe der Mechanik gefragt werden, antworten sie häufig mit der Beschreibung von realen Abläufen (Schecker et al., 2018). Die Realität im Klassenraum sieht aber meist so aus, dass aufgrund mathematischer Komplexität gerade auf diese realen Abläufe verzichtet wird und damit eine Überbetonung von Idealsituationen zustande kommt. Lernende haben jedoch Schwierigkeiten, sich in diese Idealsituationen hineinzudenken (Schecker, 1985). Das kann dazu führen, dass die bestehenden Präkonzepte auf Schülerseite durch den Unterricht nicht beeinflusst oder abgeschwächt werden, was zu einem auch nach dem Unterricht bestehenden, mangelnden Kraftverständnis führt.

Der Computereinsatz kann ein Mittel sein, den Konflikt zwischen der Erwartung der Lernenden an den Physikunterricht und dem tatsächlichen Unterricht aufzulösen, da mithilfe des Computers auch alltagsnahe und damit komplexe Bewegungen analysiert und modelliert werden können. Die Möglichkeiten, den Computer im Physikunterricht einzusetzen, sind vielfältig und ändern sich im Laufe der Zeit durch Weiterentwicklung der Software. Im Bereich der Mechanik ist die Videoanalyse beliebt, die ein einfaches Messverfahren für zweidimensionale Bewegungen darstellt. Forschungen zur Videoanalyse zeigen ihren Erfolg bisher im Bereich der Kinematik. Zur Dynamik liegen kaum empirische Ergebnisse vor. Außerdem kann der Computer zum Modellieren von Bewegungen genutzt werden, wobei den Nutzer*innen von entsprechender Software die nötige Mathematik abgenommen wird. Forschungen zur mathematischen Modellbildung zeigen, dass diese je nach Art, Software und Altersgruppe der Lernenden einige Schwierigkeiten mit sich bringt. Neuere Software versucht jedoch, diese Probleme zu lösen. Beide Arten des Computereinsatzes eignen sich theoretisch, um authentische und alltagsnahe Bewegungen im Schulunterricht zu thematisieren, unterscheiden sich in der Art der Vorgehensweise aber stark voneinander.

Inhalt dieser Arbeit ist die Entwicklung und Evaluation einer Intervention, die durch intensiven Computereinsatz in Zweiergruppen komplexe Bewegungen mit Reibungseinflüssen thematisiert und so die beschriebene Kluft zwischen Physikunterricht und Realität schließen soll. Dabei wird untersucht, inwiefern eine solche Intervention erfolgreich für das Lernen der Newtonschen Mechanik ist und wie diese die vorhandenen Schülervorstellungen beeinflusst. Außerdem werden die zwei Herangehensweisen – die mathematische Modellbildung und die Videoanalyse – miteinander verglichen.

Dazu wird zunächst berichtet, welche Schwierigkeiten im Mechanikunterricht vorliegen und welche Möglichkeiten es gibt, diese zu umgehen. Es werden klassische Schülervorstellungen im Bereich der Mechanik vorgestellt und erklärt, welche Konzeptwechselstrategien es in der didaktischen Forschung gibt. Danach wird zusammengefasst, was bereits über den Erfolg von Computereinsatz im Physikunterricht bekannt ist und wie sich die beiden Methoden mathematische Modellbildung und Videoanalyse einsetzen lassen. Auf Basis dessen wird eine Intervention vorgestellt, die die Lernwirksamkeit der beiden Methoden untersucht. Deren Erfolg wird mit einem Mixed-Methods-Ansatz evaluiert, der aus einer Kombination von Daten aus einem eigens erstellten Fragebogen und Bildschirmaufnahmen besteht.

I. Theoretischer Hintergrund

2 Fachliche Grundlagen der klassischen Mechanik

Die klassische Mechanik umfasst jegliche Beschreibungen von Bewegungen, die sich nicht in der Nähe der Lichtgeschwindigkeit abspielen und in Größenordnungen oberhalb der Quantenmechanik anzusiedeln sind (van Dongen, 2021). Dabei lässt sich die klassische Mechanik auf wenige Grundaussagen reduzieren, die, sobald der Zustand eines Systems zu einem Zeitpunkt bekannt ist, die Zukunft dieses Systems vorhersagen können. Häufig wird die klassische Mechanik in der Physik dabei in die Kinematik und die Dynamik unterteilt. In den Ingenieurwissenschaften ist die Dynamik mit der Statik und der Festigkeitslehre ein Gebiet der technischen Mechanik und setzt sich wiederum aus Kinematik und Kinetik zusammen (Wilhelm et al., 2021). Aufgrund des Schulbezugs dieser Arbeit wird die Einteilung der Mechanik in Kinematik und Dynamik beibehalten.

2.1 Kinematik

Die Kinematik beschreibt den Teil der Mechanik, der sich mit der Beschreibung von Bewegungen befasst. Es werden Bewegungen analysiert und anhand von verschiedenen Größen charakterisiert. Damit der Ort eines Objekts beschrieben und gemessen werden kann, muss zunächst ein Bezugssystem festgelegt werden. Dieses Bezugssystem ist willkürlich. Es ist ein Inertialsystem, wenn es in Ruhe ist oder sich mit einer konstanten Geschwindigkeit bewegt. Im Schulkontext wird das Bezugssystem normalerweise auf der Erdoberfläche verankert. Da diese aber stets beschleunigt wird, kommen Scheinkräfte wie die Corioliskraft zustande. Scheinkräfte werden eingeführt, um die Newtonsche Mechanik auch in beschleunigten Bezugssystemen anwenden zu können, auch wenn diese Scheinkräfte die Voraussetzungen einer Kraft im Newtonschen Sinne nicht erfüllen. Die Corioliskraft ist in den Größenordnungen, in denen in der Schule gerechnet wird, aber meist zu vernachlässigen. Die Wahl des Bezugssystems wird meist nicht näher diskutiert, da es häufig auch eine „natürliche“ Wahl zu geben scheint. Gerechnet wird mit verschiedenen Koordinaten, wobei kartesische Koordinaten den Standard in der Schule bilden. Der Ort eines Objekts kann dann durch die Angabe des Abstands zum Ursprung in der jeweiligen Koordinatenrichtung angegeben werden. $\vec{r}(t)$ ist dementsprechend der n-dimensionale (in der Schule in der Regel ein- bis dreidimensionale) Ortsvektor des Punkts, an dem sich das Objekt zur Zeit t befindet. Unterschieden werden muss dieser von dem Weg oder der Strecke, die das Objekt zurücklegt. Der Weg ist definiert als die Strecke, die ein Objekt zwischen zwei Zeitpunkten zurücklegt und ist damit eine skalare Größe.

$$s = \int_a^b |\dot{\vec{r}}(t)| dt$$

Die Geschwindigkeit ist die Ableitung des Ortes nach der Zeit: $\vec{v} = \lim_{\Delta t \rightarrow 0} \frac{\Delta \vec{r}}{\Delta t} = \frac{d\vec{r}}{dt}$, und damit auch eine vektorielle Größe. Man kann die Geschwindigkeit zerlegen in ihren Betrag und ihre Richtung, die im Sprachgebrauch Tempo oder Schnelligkeit und Bewegungsrichtung genannt werden. Die Beschleunigung ist die Ableitung der Geschwindigkeit nach der Zeit ($\vec{a} = \lim_{\Delta t \rightarrow 0} \frac{\Delta \vec{v}}{\Delta t} = \frac{d\vec{v}}{dt}$) und damit ebenfalls eine vektorielle Größe. Auch hier kann zwischen dem Betrag der Beschleunigung und der Richtung der Beschleunigung unterschieden werden. Die genannten Größen dienen zur Beschreibung einer Bewegung. Die Gesetze, die bestimmen, wie sich ein Objekt bewegt, werden im folgenden Kapitel beschrieben.

2.2 Dynamik

Unter Dynamik wird der Teil der Mechanik verstanden, der sich mit wirkenden Kräften beschäftigt. Über die reine Beschreibung von Bewegungen hinaus ist hier die Ursache einer Bewegung von Interesse. Die Grundlage der Dynamik bilden die drei Newtonschen Gesetze:

- (1) „Jeder Körper beharrt in seinem Zustande der Ruhe oder der gleichförmigen geradlinigen Bewegung, wenn er nicht durch einwirkende Kräfte gezwungen wird, seinen Zustand zu ändern.“ (Newton, 1963)
- (2) „Die Aenderung der Bewegung ist der Einwirkung der bewegenden Kraft proportional und geschieht nach der Richtung derjenigen geraden Linie, nach welcher jene Kraft einwirkt.“ (Newton, 1963)
- (3) Die Wirkung ist stets der Gegenwirkung gleich, oder die Wirkungen zweier Körper auf einander sind stets gleich und von entgegengesetzter Richtung.“ (Newton, 1963)

Diese deutsche Übersetzung aus dem lateinischen Original stellt den Wortlaut der Gesetze der Bewegung von Newton dar. Die genaue Auslegung der Newtonschen Gesetze (oder auch Axiome) hat sich im Laufe der Zeit teilweise verändert. Insbesondere die ersten beiden Gesetze bilden Grundlage zur Diskussion über die genaue Auslegung des Wortlauts. In der Schule wird manchmal so damit umgegangen, dass das zweite Gesetz die Grundlage der Mechanik bildet und aussagt, dass eine Impulsänderung von einer Kraft hervorgerufen wird. Dabei ist diese Impulsänderung proportional zur Einwirkungsstärke und Einwirkungsdauer und passiert in Einwirkungsrichtung.

In Gleichungsform und heutiger Schreibweise sieht das folgendermaßen aus:

$$\vec{F} = \frac{d\vec{p}}{dt} = \frac{d}{dt}(m \cdot \vec{v})$$

Dabei ist $\vec{p} = m \cdot \vec{v}$ der Impuls. Wenn man davon ausgeht, dass die Masse konstant bleibt, was in der Schule häufig vorausgesetzt wird, kann die Gleichung nun wie folgt geschrieben werden:

$$\vec{F} = m \cdot \frac{d\vec{v}}{dt} = m \cdot \vec{a}$$

Diese Schreibweise wird in der Schule am häufigsten verwendet und ist Schüler*innen in der Regel nach dem Unterricht bekannt. In der Schule wird das erste Gesetz dann als Spezialfall des zweiten Gesetzes verstanden. Wenn keine Kraft auf das Objekt wirkt, folgt aus der Gleichung, dass die Änderung der Geschwindigkeit Null ist und diese damit konstant bleibt. Das gilt sowohl für Ruhe als auch für eine gleichförmige Bewegung. Manchmal wird aber auch das erste Gesetz zuerst unterrichtet.

Das dritte Newtonsche Gesetz beschreibt die Wechselwirkung zwischen zwei Körpern. Die auf die beiden Körper wirkenden Kräfte sind immer gleich groß und zeigen in entgegengesetzte Richtungen. Dieses Gesetz vervollständigt die anderen beiden in der Hinsicht, dass eine Kraft immer alle drei Gesetze erfüllen muss. Sie bilden gemeinsam also die Definition einer Kraft. Die bereits genannte Corioliskraft erfüllt beispielsweise nicht das dritte Newtonsche Gesetz und ist damit keine Kraft (auch Scheinkraft genannt).

Mit den oben beschriebenen Zusammenhängen ist die Geschwindigkeit eines Objekts mit konstanter Masse dann abhängig von den wirkenden Kräften:

$$\vec{v} = \frac{1}{m} \int \vec{F} dt = \int \frac{\vec{F}}{m} dt$$

$\vec{F}$ ist dabei die Summe aller wirkenden Kräfte: $\vec{F} = \vec{F}_1 + \dots + \vec{F}_n$. Daraus folgt der Ort durch erneute Integration:

$$\vec{r} = \int \left(\int \frac{\vec{F}}{m} dt \right) dt$$

Die wirkenden Kräfte können orts-, zeit- oder auch geschwindigkeitsabhängig sein. Es sind deshalb nicht für alle entstehenden Differentialgleichungen analytische Lösungen möglich. Die entstehenden Integrationskonstanten müssen über die Anfangsbedingungen

(oder den Zustand des Systems zu einem beliebigen anderen Zeitpunkt) der Geschwindigkeit und des Ortes ermittelt werden. Zur Mechanik gehören weitere Größen und Erhaltungssätze (Energie, Impuls etc.), auf die aufgrund der Schwerpunktsetzung der Studie hier nicht weiter eingegangen wird.

2.3 Elementarisierungsentscheidungen in der Schule

Je nach Bundesland und Schulform wird die Mechanik meist in der Mittelstufe (z. B. achte Jahrgangsstufe) und nochmals in der Oberstufe unterrichtet. In beiden Fällen muss elementarisiert werden, weil einerseits das mathematische Handwerkszeug fehlt, andererseits auch die physikalischen Konzepte teilweise zu komplex für die Zielgruppe sind.

Beispielhaft kann der Ort eines schräg durch die Luft geworfenen Balls nach einer bestimmten Zeit t_1 mit der obigen Gleichung bestimmt werden. Da es sich um eine zweidimensionale Bewegung handelt, reicht es aus, ein zweidimensionales, kartesisches Koordinatensystem zu definieren. Die x-Achse verläuft, wie üblich, parallel zur Erdoberfläche mit positiver Richtung nach rechts und die y-Achse senkrecht zur Erdoberfläche mit positiver Richtung nach oben. Die wirkenden Kräfte sind eine senkrecht nach unten gerichtete Gravitationskraft $\vec{F}_G = m \cdot \vec{g}$ mit $\vec{g} = -9,81 \cdot \vec{e}_y$ und eine Luftreibungskraft, die der Bewegungsrichtung entgegenwirkt und von der Querschnittsfläche des Objekts, der Luftdichte, der Form des Objekts und dem Tempo abhängig ist. Es gilt $\vec{F}_R = -\frac{1}{2}A\rho c_W |\vec{v}| \cdot \vec{v}$, wenn von einer turbulenten Luftströmung ausgegangen wird. Also:

$$\vec{r} = \int \left(\int \frac{-m \cdot 9,81 \frac{m}{s^2} \cdot \vec{e}_y - \frac{1}{2}A\rho c_W v \vec{v}}{m} dt \right) dt$$

Neben der sichtbaren Komplexität der Gleichung ist diese Differentialgleichung analytisch mit Elementaroperationen nur sehr schwierig zu lösen und eine geschlossene Lösung liegt gar nicht vor (Ray & Fröhlich, 2015), weshalb auf numerische Methoden zurückgegriffen werden müsste, was händisch ebenfalls nahezu unmöglich bzw. sehr mühsam und aufwändig ist.

Vektorielle Größen werden in der Mathematik (z. B. in Hessen) zudem erst nach Behandlung der Mechanik in Physik eingeführt und deren Ableitung und Integral ist meist nicht Inhalt der Schulmathematik (Hessisches Kultusministerium, 2016). In den folgenden Abschnitten sollen einige mögliche Elementarisierungsentscheidungen in der Mechanik

vorgestellt werden, die versuchen, das oben dargestellte Problem zu lösen. Diese sind für den Mechanikunterricht der Oberstufe typisch.

Reduktion der Dimensionen

In hessischen Gymnasien lernen Schüler*innen im Mathematikunterricht in der Einführungsphase in die gymnasiale Oberstufe (E-Phase, 11. Jahrgangsstufe in G9 bzw. 10. Jahrgangsstufe in G8)) den Umgang mit Ableitungen und Integralen und im zweiten Halbjahr der Qualifikationsphase (Q2, 12. Jahrgangsstufe in G9) mit Vektoren (Hessisches Kultusministerium, 2016). Da die Mechanik im Physikunterricht in der Oberstufe in der E-Phase unterrichtet wird, können Physiklehrkräfte also nicht auf diese Kenntnisse aus dem Mathematikunterricht bauen. Während die Lernenden Ableitungen und Integrale parallel im selben Schuljahr kennenlernen, ist die Einführung von Vektoren für die Physik zu spät. Aufgrund dessen entscheiden sich viele Lehrkräfte gegen das Verwenden von vektoriellen Größen. Am einfachsten gelingt dies, wenn auf eindimensionale Bewegungen zurückgegriffen wird, bei denen nur noch das Vorzeichen von der Richtung der Größe übrigbleibt.

Für das in Kapitel 2.3 dargestellte Beispiel würde das bedeuten, dass nur noch ein senkrechter Wurf bzw. Fallbewegung betrachtet wird. Dann wird aus der Bewegungsgleichung:

$$r = \int \left(\int \frac{-m \cdot 9,81 \frac{m}{s^2} - \frac{1}{2} A \rho c_W v^2}{m} dt \right) dt$$

Dabei ist v die Geschwindigkeit und ändert je nach Bewegungsrichtung das Vorzeichen.

Im zweidimensionalen Fall werden die Koordinatenrichtungen voneinander getrennt und separat als skalare Gleichungen ausgewertet. Dies ist aber nur für den Fall möglich, dass keine Luftreibung wirkt, da sonst die Koordinatenrichtungen nicht mehr unabhängig voneinander sind. Mit Luftreibung ergeben sich folgende Bewegungsgleichungen:

$$\vec{x} = \int \left(\int \frac{-\frac{1}{2} A \rho c_W |\vec{v}| \vec{v} \vec{e}_x}{m} dt \right) dt$$

und

$$\vec{y} = \int \left(\int \frac{-m \cdot 9,81 - \frac{1}{2} A \rho c_w |\vec{v}| \vec{v} \vec{e}_y}{m} dt \right) dt$$

Da in diesem Fall $v = \sqrt{v_x^2 + v_y^2}$ ist, sind die beiden Koordinatenrichtungen nicht voneinander unabhängig und die Gleichungen sind nach wie vor nicht lösbar. Es muss in diesem Fall also auf Reibung verzichtet werden.

Alternativ können zweidimensionale Bewegungen auf eindimensionale zurückgeführt werden (Wilhelm et al., 2021). In diesem Fall wird Abstand s entlang der Bahnkurve von einem Startpunkt betrachtet. In diesem Fall ändert sich der Einheitsvektor in Abhängigkeit der Position des Objekts. Auch wenn man damit die kinematischen und dynamischen Zusammenhänge erklären und zeigen kann, ist es u. a. unmöglich daraus abzuleiten, dass eine Kraft zu einer Richtungsänderung führen kann und es sich bei Kraft, Beschleunigung und Geschwindigkeit um vektorielle Größen handelt (Wilhelm, 2018a; Wilhelm et al., 2021).

Simplifizierungen von wirkenden Kräften und Betrachtung von Idealsituationen

Da in den Beispielen zuvor die Luftreibungskraft das Hauptproblem darstellt und diese vermeintlich nicht essenziell für das Verständnis der Bewegungsgesetze ist, wird sie häufig weggelassen. Es handelt sich dann um eine Idealsituation, die es in der Realität nur gibt, wenn sich ein Objekt im Vakuum bewegt.

Die Bewegungsgleichungen vereinfachen sich damit zu

$$\vec{x} = \int \left(\int 0 dt \right) dt$$

und

$$\vec{y} = \int \left(\int -9,81 \frac{m}{s^2} dt \right) dt.$$

Damit liegen keine Differentialgleichungen mehr vor und das Integral kann herkömmlich gelöst werden. Da auch die Integralrechnung allerdings nach der Mechanik unterrichtet wird, können Lernende dies zu dem Zeitpunkt mathematisch ebenfalls noch nicht. Eine weitere Vereinfachung, die im Physikunterricht dann häufig vorgenommen wird, ist die

Bewegungsgleichung zu lösen und die Lösungen als spezielle Formeln herauszuheben, mit denen man beispielsweise den schiefen Wurf und analoge Bewegungen berechnen kann.

Problematisch an der konsequenten Vernachlässigung von Reibung ist es, dass sich Alltagsbeobachtungen der Schüler*innen dadurch nicht mit den Vorhersagen der Newtonschen Mechanik decken. Im Alltag kommen bewegte Gegenstände nach einer bestimmten Zeit immer zur Ruhe oder müssen aktiv angetrieben werden. Daher liegt es nahe, dass Lernende davon ausgehen, dass eine Kraft in Bewegungsrichtung wirken muss, damit sich etwas überhaupt bewegen kann (siehe auch Kapitel 3). Eine solche Kraft wird dann erfunden und z. B. „Bewegungskraft“ genannt. Nach verschiedenen Conceptual Change-Theorien muss die Physik den Lernenden eine zu ihren Präkonzepten alternative Theorie liefern, die plausibel genug ist, dass Schüler*innen eine Notwendigkeit darin sehen, zu ihr überzugehen oder sie zumindest in Anwendungssituationen zu verwenden (z. B. Potvin, 2017). Das häufige Vorkommen und die Hartnäckigkeit der genannten Vorstellung könnte auf einen Mangel an Möglichkeiten zurückzuführen sein, Reibung in der Schule aktiv zu besprechen.

Betonung von Spezialfällen

Um die (mathematische) Komplexität verschiedener Bewegungen zu umgehen, werden in der Schule häufig Spezialfälle betrachtet, in denen bspw. Reibungsfreiheit herrscht und nur eine Kraft auf den Körper wirkt. Ziel ist es dann, Berechnungsformeln zu erarbeiten, mit denen gewisse Größen ausgerechnet werden können. Beispielsweise gilt für den schiefen Wurf dann

$$x = v_{0x} \cdot t$$

und

$$y = \frac{1}{2} g t^2 + v_{0y} \cdot t.$$

Dies gilt wiederum nur, wenn das Objekt sich im Zeitnullpunkt genau im Koordinatenursprung befindet. Sobald eine der genannten Bedingungen nicht mehr zutrifft, kann die Gleichung nicht mehr benutzt werden. In vielen Fällen nehmen solche Gleichungen bei Schüler*innen einen hohen Stellenwert ein, da sie eine Möglichkeit bieten, Größen auszurechnen (Schecker, 1998a). Dies kann zusätzlich durch die Art der Klausuren bedingt sein, in denen häufig bei gewissen gegebenen Größen andere ausgerechnet werden müssen. Dazu bieten sich dann die genannten Berechnungsformeln an. Zur Nutzung einer

solchen Formel bedarf es dabei allerdings keinerlei Konzeptverständnis. Problematisch wird es dann, wenn Lernende sich nicht bewusst sind, in welchen Fällen eine solche Gleichung eingesetzt werden kann und sie keinen qualitativen Unterschied zur Grundgleichung der Mechanik $\vec{a} = \frac{\vec{F}}{m}$ sehen. Zudem bleibt hinter einer solchen Formel verborgen, warum sich das Objekt überhaupt in dieser Weise bewegt. Es sollte im Unterricht also mehr Fokus auf grundlegenden Zusammenhängen (*power tools*, (MacDonald et al., 1988)) liegen als auf den genannten Spezialfällen. Dies wird häufig aus rein mathematischen Gründen nicht getan.

Statt die Situation aufgrund ihrer mathematischen Komplexität immer mehr zu vereinfachen, gibt es auch die Möglichkeit, Hilfsmittel zu nutzen, die es den Schüler*innen ermöglichen, sich mit der Physik einer solchen Situation zu beschäftigen, ohne mit der komplexen Mathematik konfrontiert zu sein. Ein Beispiel für solche Hilfsmittel sind Computer, die eine Vielzahl an Software anbieten, mit denen beschriebene Bewegungen betrachtet werden können (siehe Kapitel 5 und 6).

3 Schülervorstellungen

3.1 Generelle Aspekte von Schülervorstellungen

Schülervorstellungen sind Konzepte und Vorstellungen, mit denen Schüler*innen in den Unterricht kommen. Solche alltagsnahen Vorstellungen werden teilweise auch Präkonzepte, Alltagstheorien oder auch Fehlvorstellungen genannt. Da Letzteres vor allem auf die negativen Aspekte der Vorstellung hindeutet und die positiven Aspekte unterschlägt, ist dieser Begriff ungünstig. Auch wenn die hier diskutierten Vorstellungen nicht nur bei Schüler*innen auftreten wird dennoch der gängige Begriff Schülervorstellung genutzt.

Schülervorstellungen bezeichnen im engeren Sinn Vorstellungen über physikalische Größen und Begriffe, deren Beziehungen untereinander und über physikalische Gesetzmäßigkeiten (Wilhelm, 2005b). Mittlerweile werden auch allgemeinere Dinge, wie Denkrahmen über Erkenntnisgewinnung in der Physik (Schecker, 1983) oder den eigenen Lernprozess als Schülervorstellungen bezeichnet. Diese Vorstellungen sind vor allem durch alltägliche Erfahrungen und Formulierungen der Alltagssprache geprägt und haben sich dementsprechend im Laufe des Lebens gebildet und gefestigt. Sie erweisen sich in der Umwelt der Schüler*innen als sinnvoll und hilfreich, da sie in einem gewissen Maß korrekte Vorhersagen über Abläufe erlauben (Wilhelm, 2005b), entsprechen aber meist nicht der physikalischen Vorstellung. Sie beeinflussen den Lernerfolg der Lernenden maßgeblich. Als Lehrkraft ist es wichtig, typische Schülervorstellungen zu kennen und zu identifizieren, um den Unterricht daran anpassen zu können, da deren Muster bei den meisten Schüler*innen ähnlich sind. Insbesondere unter einer konstruktivistischen Perspektive des Lernens sind die Voraussetzungen der Lernenden für die Konstruktion neuen Wissens zu berücksichtigen. Lernimpulse werden von Seiten der Schüler*innen danach immer im Kontext des vorhandenen Wissens und der vorhandenen Präkonzepte aufgenommen und verarbeitet, wodurch die Lernergebnisse erheblich beeinflusst werden können. In jedem Teilgebiet der Physik müssen also bei der didaktischen Rekonstruktion der Thematik die Schülervorstellungen berücksichtigt werden. Übergeordnet stellt sich zudem die Frage nach der grundlegenden Strategie im Umgang mit solchen Präkonzepten. Es ist einerseits möglich, ausbaufähige Präkonzepte aufzugreifen und aufzubauen, während unerwünschte nicht thematisiert werden. Andererseits kann intensiv mit kognitiven Konflikten und Konfrontationen gearbeitet werden, um einen Anreiz zum Wechsel auf die physikalisch erwünschten Konzepte zu schaffen. Unter anderem mit diesen Strategien befasst sich die Forschung im Bereich des Conceptual Change (Kapitel 3.4).

Lernende können ihre Vorstellungen meist nicht präzise artikulieren. Die heute bekannten Schülervorstellungen konnten in Forschungsarbeiten herausgearbeitet werden, da es in allen Bereichen der Physik gewisse wiederkehrende Muster gibt, nach denen Lernende antworten. Diese Antworten lassen sich dann so interpretieren, als ob die entsprechende Person die formulierte Schülervorstellung hielt. Niedderer und Schecker (1992) unterscheiden hier zwischen Tiefen- und Oberflächenstruktur. In der Tiefenstruktur sind die konkreten Vorstellungen enthalten, welche sich in der Oberflächenstruktur äußern. Lehrkräfte und Forscher*innen können durch diese Oberflächenstruktur nur durch Interpretation auf die Tiefenstruktur schließen. Obwohl Schülervorstellungen also nicht direkt beobachtbar sind, stellen die Forschungsmethoden und die zahlreichen Replikationen der entsprechenden Studien sicher, dass gewisse Schülervorstellungen bei der Beantwortung von Fragen oder der Konstruktion von Wissen eine Rolle spielen. Weniger Einigkeit besteht darin, ob die Schülervorstellungen als übergreifende Annahmen (*Frameworks*) existieren (Vosniadou, 1994) oder eine Sammlung einzelner, unzusammenhängender Schülerideen (*p-prims*) sind (diSessa, 1993). Dies wird in Kapitel 3.4 diskutiert.

3.2 Schülervorstellungen zur Mechanik

Speziell in der Mechanik gibt es viele hartnäckige Schülervorstellungen, die sehr gut erforscht sind. Teilweise werden diese Schülervorstellungen durch gewisse Elementarisierungsentscheidungen (siehe dazu Kapitel 2.3) hervorgerufen, verstärkt oder zumindest nicht ausreichend herausgefordert. Im Folgenden¹ werden die Schülervorstellungen näher beleuchtet, die relevant für die vorliegende Arbeit sind. Die Einteilung der Schülervorstellungen ist nicht eindeutig. Es können sich mehrere Schülervorstellungen aus einer anderen übergeordneten Schülervorstellung ableiten, andere wiederum bieten Überschneidungen und hängen eng zusammen. Die folgende Auflistung der Schülervorstellungen ist also weder vollständig, noch in ihrer Einteilung eindeutig. Nahezu alle der aufgelisteten Schülervorstellungen lassen sich grundsätzlich auf ein mangelndes Verständnis des Zusammenhangs zwischen Kräften und Bewegung (also der ersten beiden Newtonschen Gesetze) zurückführen. Hier werden im Hinblick auf die spätere Auswertung nur Schülervorstellungen vorgestellt, die in dem in der Intervention genutzten Fragebogen oder der Diskussion der Schüler*innen während der Arbeitsphasen enthalten sind. Dabei handelt

¹ Die hier dargestellten Schülervorstellungen basieren hauptsächlich auf dem Lehrbuchkapitel von Schecker und Wilhelm (2018). Das Lehrbuch fasst die zahlreichen Ergebnisse verschiedener Forschungsstudien zu dem Thema zusammen. Ein gute Übersicht über diese Studien gibt Duit (2009). Teilweise werden zusätzlich entsprechende Primärquellen zu den jeweiligen Vorstellungen angegeben.

es sich durch den Fokus der Studie nahezu ausschließlich um Schülervorstellungen im Bereich der Dynamik. Eine einzelne Vorstellung aus dem Bereich der Kinematik ist ebenfalls inkludiert.

Kraft ist ein Sammelbegriff²

Diese Schülervorstellung meint, dass der Begriff der Kraft als Sammelbegriff für andere Größen genutzt wird (Energie, Impuls etc.). Es findet ähnlich wie zwischen Geschwindigkeit und Beschleunigung (s. u.) eine mangelnde Unterscheidung zwischen verschiedenen Größen statt. Dies kann durch generell mangelndes Verständnis der verschiedenen Größen entstehen. Andererseits kann auch diese Schülervorstellung wieder mit anderen zusammenhängen. Zum Beispiel kann bei der Suche nach einer Kraft, die in Bewegungsrichtung wirkt, statt einer Kraft die Energie genannt werden. Daraus wird z. B. gefolgert, dass eine Energie in Bewegungsrichtung wirkt. Hierunter kann auch gefasst werden, dass Schüler*innen Kraft als Fähigkeit zum Wechselwirken auffassen und nicht als die Wechselwirkung selbst (siehe auch „Ein bewegter Körper hat Kraft“).

Zur Aufrechterhaltung einer Bewegung bedarf es einer Kraft in Bewegungsrichtung³

Sehr viele Schüler*innen gehen von einer Proportionalität zwischen Kraft und Geschwindigkeit statt Beschleunigung aus. Diese Vorstellung ist damit verbunden, dass es einer Kraft bedarf, damit eine Bewegung aufrechterhalten werden kann. Dies kann durch Alltagssituationen, wie z. B. dem Fahrradfahren, entstehen. Eine Person, die sich mit einer konstanten Geschwindigkeit auf dem Fahrrad fortbewegt, erfährt verschiedene Kräfte (Gravitationskraft, Normalkraft der Straße, Antriebskraft, Reibungskräfte etc.). Alle diese Kräfte müssen sich aus Sicht des Experten bzw. der Expertin gegenseitig aufheben, so dass die Summe aller Kräfte Null ist, damit sich das Objekt mit konstanter Geschwindigkeit bewegen kann. Für Lernende ist es aber ebenfalls schlüssig, Reibungskräfte zu ignorieren und durch die Beobachtung, dass man schneller fährt, wenn man stärker in die Pedale tritt, darauf zu schließen, dass diese Kraft die einzige ist und in Bewegungsrichtung wirkt. Die Proportionalität zwischen dieser Kraft und dem Tempo ist dann schnell hergeleitet.

² Jung (1981)

³ Schecker (1985)

In Aufgaben äußert sich diese Vorstellung so, dass Kräfte eingeführt werden, die nicht vorhanden sind, um zu erklären, dass sich etwas bewegen kann („Abschlagskraft“, „Bewegungskraft“ etc.). Ohne eine aktive Thematisierung von Reibungskräften erscheint es schwierig, die Rahmenbedingungen zu schaffen, dass Schüler*innen von dieser Vorstellung zur physikalisch gewünschten Vorstellung übergehen. Diese Schülervorstellung ist eine der häufigsten und nahezu bei jedem Lernenden zu finden. Das Konzeptverständnis von Kräften ist stark davon abhängig, ob die Schülervorstellung überwunden werden kann.

Ein bewegter Körper hat Kraft

Nach dieser Vorstellung hat ein sich bewegender Körper Kraft (Impetus). Diese wird teilweise mit Bewegungskraft oder auch Abschlags- oder Abschusskraft bezeichnet. Es handelt sich dabei in der Regel nicht um eine Verwechselung der Begriffe Kraft und Energie (oder Impuls), da eine solche als Bewegungskraft bezeichnete Kraft auch mit anderen Kräften addiert und generell als Kraft behandelt wird. Diese Schülervorstellung hängt auch wiederum eng mit der Vorstellung zusammen, dass eine Kraft in Bewegungsrichtung wirken muss. Diese Notwendigkeit kann nämlich dazu führen, dass nach einer Kraft gesucht wird, die für die Bewegung verantwortlich ist. Wenn keine äußere Kraft als Ursache gefunden wird, wird die im Körper gespeicherte Kraft als Ursache gesehen.

Alle Körper kommen natürlicherweise zum Stillstand

Diese Vorstellung wurde als Teil anderer Vorstellungen im Vorfeld bereits diskutiert, wird hier aber nochmal gesondert vorgestellt. Der Alltag liefert die direkte Beobachtung, dass jedes Objekt, das nicht aktiv versucht, den Bewegungszustand aufrecht zu erhalten, nach einer Zeit zur Ruhe kommt. Naheliegend ist, dass dies der natürliche Verlauf ist und nicht durch äußere Einflüsse (Reibung) verursacht wird. Reibung ist also auch hier der Schlüssel zu einem Konzeptwechsel zur gewünschten und physikalisch korrekten Vorstellung.

Bei Kreisbewegungen wirkt die Zentrifugalkraft⁴

Bei der Besprechung von Kreisbewegungen wird von Lernenden meist die Zentrifugalkraft als erstes genannt. Die Zentrifugalkraft ist aus dem Alltag bekannt und sorgt dafür,

⁴ Jung (1981)

dass man bei Kreisbewegungen nach außen gedrückt wird (Karussell, Kurvenfahrt etc.). Da man diese Kraft vermeintlich spürt, nimmt sie bei Lernenden einen hohen Stellenwert ein.

Da man sich im Karussell in einem beschleunigten Bezugssystem befindet, gelten die Newtonschen Gesetze in diesem Bezugssystem nicht. Um die Bewegung dennoch erklären zu können, wird die Scheinkraft Zentrifugalkraft eingeführt. Wenn eine solche Bewegung von außen in einem stationären Bezugssystem analysiert wird, wird aber klar, dass der Körper nach innen beschleunigt, also auch die Kraft nach innen zeigen muss (Zentripetalkraft). Diese Schülervorstellung der Zentrifugalkraft ist häufig anzutreffen, betrifft allerdings im Gegensatz zur vorigen Schülervorstellung eine spezielle Bewegung und nicht das zweite Newtonsche Gesetz selbst.

Bei einer gleichförmigen Kreisbewegung wirken keine Kräfte

Eine verbreitete Vorstellung zur gleichförmigen Kreisbewegung ist, dass bei dieser keine Kräfte wirken. Dies kann unterschiedliche Ursachen haben. Die Vorstellung könnte damit begründet sein, dass sich das Objekt „natürlicherweise“ im Kreis bewegt und daher keine Kräfte notwendig sind. Andererseits kann man auch zum dem Schluss kommen, dass keine Kraft wirkt, wenn man die Geschwindigkeit als skalare Größe auffasst. Diese würde sich in dem Beispiel dann nicht ändern, wodurch der Schluss gezogen werden könnte, dass keine Kraft wirkt. Die Vorstellung kann also entweder in einem mangelnden Verständnis der Newtonschen Gesetze begründet liegen oder in einer unzureichenden Vorstellung über die Geschwindigkeit bei gleichzeitig korrekter Anwendung der Grundgleichung der Mechanik.

Die Bewegungsform prägt sich dem bewegenden Objekt ein

Bei einer Kreisbewegung, die durch Beendigung der Zentripetalkraft (bspw. durch Durchschneiden eines Fadens) unterbrochen wird, vermuten Schüler*innen häufig, dass sich der Körper in einem Bogen weiterbewegt, da sich die vorige Bewegungsform dem Körper eingeprägt hat. Dies ist mit der Grundgleichung der Mechanik nicht vereinbar. Das Wegfallen der Kraft bedeutet, dass diese auch keine Beschleunigung mehr verursachen kann, was zu einer konstanten Geschwindigkeit nach Betrag und Richtung führt.

Diese Schülervorstellung passt zu den sogenannten Impetustheorien und der oben bereits vorgestellten Schülervorstellung, die besagen, dass ein sich bewegendes Objekt einen Impetus besitzt, der eine Art gespeicherte Kraft darstellt. Die Ursache zu glauben, dass ein solcher Impetus vorliegen muss, ist ebenfalls darin begründet, dass ein Objekt von etwas

angetrieben werden muss. Wenn keine äußere Kraft vorhanden ist, muss dies eine im Körper gespeicherte Kraft sein. In diesem Beispiel ist dann die Kreisbewegung in dem Körper gespeichert, die nach dem Wegfallen der Zentripetalkraft erst nach und nach verbraucht wird, sodass sich das Objekt in einem Bogen weiterbewegt.

Ein Objekt kann sich nur mit konstanter Geschwindigkeit bewegen, wenn keine Kräfte wirken: Das Beharrungsprinzip gilt nur für Spezialfälle⁵

Eine Schülervorstellung, die speziell das erste Newtonsche Gesetz betrifft, ist die, dass sich ein Körper nur dann mit konstanter Geschwindigkeit bewegen kann, wenn keine Kräfte wirken. Fachlich korrekt ist, dass die resultierende Gesamtkraft, die auf den Körper wirkt, null sein muss, damit sich ein Objekt mit konstanter Geschwindigkeit bewegt. Der erste Fall ohne Kräfte ist im Alltag nicht anzutreffen, da immer mehrere Kräfte wirken. Daher ist es umso wichtiger, dass internalisiert wird, dass auch mehrere wirkende Kräfte zu einer konstanten Geschwindigkeit führen können, wenn diese sich ausgleichen.

Die Kräfte, die zwischen zwei Körpern wirken, sind abhängig von deren Masse: „Je schwerer, desto stärker“

Diese Schülervorstellung betrifft das dritte Newtonsche Gesetz (Wechselwirkungsprinzip). Wenn ein PKW und ein LKW zusammenstoßen, ist zu beobachten, dass dieser Zusammenstoß den PKW in seiner Geschwindigkeit stärker beeinflusst als den LKW. Daraus wird geschlossen, dass die Kraft, die der LKW auf den PKW ausübt, größer ist als die umgekehrte Kraft. Beobachtet werden die Beschleunigungen der beiden Fahrzeuge, die sich stark unterscheiden. Dass die Kräfte, die wirken, betragsmäßig dennoch gleich groß sind, ist für viele Schüler*innen schwierig zu verinnerlichen. Es kann hier argumentiert werden, dass die Schülervorstellung zwar das dritte Newtonsche Gesetz betrifft, aber ebenso durch ein Missverstehen des zweiten Gesetzes entsteht, da von der Wirkung (also Beschleunigung des PKW) ohne Beachtung der Masse auf die Kraft geschlossen wird.

Nur aktive Körper können Kräfte ausüben, passive leisten Widerstand⁶

Häufig werden Kräften eine Wertigkeit zugesprochen. In einer Wechselwirkung ist dann ein Körper aktiv (z. B. ein Ball) und der andere passiv (z. B. die Wand). Der Ball übt bei

⁵ Schecker (1985)

⁶ Brown (1989)

einem Stoß gegen die Wand dann eine Kraft aus, während die Wand lediglich Widerstand leistet. Die Wechselwirkung zwischen beiden Objekten wird als asymmetrisch angesehen. Diese Vorstellung kommt aus der alltäglichen Erfahrung, dass häufig ein Wechselwirkungspartner lebend und/oder in Bewegung ist, während der andere Partner unbelebt und/oder stillstehend ist (Hand und Wand, Ball und Boden etc.). Daher liegt es nahe, dass Objekte in aktiv und passiv unterteilt werden und ihnen dementsprechend eine unterschiedliche Wertigkeit zugeschrieben wird. Physikalisch gibt es aber keinen Unterschied zwischen den Kräften, die die beiden Körper aufeinander ausüben.

Eine Kraft braucht eine Zeit, bis sie sich durchsetzen kann: Die stärkere Kraft gewinnt

In manchen Fällen gehen Schüler*innen davon aus, dass eine Kraft zwar vorhanden ist, sich ihre Wirkung aber erst nach einiger Zeit zeigt, wenn sie sich gegen eine andere Kraft durchsetzt. Kombiniert ist dies oft mit der Vorstellung, dass ein Körper eine Kraft hat („Bewegungskraft“, siehe Impetus) und dass die Kraft immer in Bewegungsrichtung zeigen muss. So wird die Bewegung eines senkrecht nach oben geworfenen Objekts bspw. so analysiert, dass die Bewegungskraft des Objekts zunächst für die Bewegung nach oben verantwortlich ist. Nach einiger Zeit kann sich aber die Erdanziehungskraft gegen diese Kraft durchsetzen und sorgt dann dafür, dass der Körper sich wieder nach unten bewegt. Es kann also argumentiert werden, dass diese Vorstellung durch das Vorhandensein anderer Schülervorstellungen entsteht, die wiederum ein Mangel an Verständnis des zweiten Gesetzes zeigen. Sie wird dennoch einzeln betrachtet.

Undifferenziertheit von Geschwindigkeit und Beschleunigung

Diese Schülervorstellung ist aus dem Bereich der Kinematik und entsteht durch den allgemeinen Sprachgebrauch und einer damit einhergehenden mangelhaften Differenzierung zwischen den beiden Begriffen. Es hat sich gezeigt, dass die Beschleunigung als zweite Ableitung des Ortes nach der Zeit allgemein schwierig ist für Schüler*innen. Hohe Beschleunigungen werden im Alltag zusätzlich meistens mit hohen Geschwindigkeiten assoziiert. Ein Auto mit einer hohen Beschleunigung fährt schnell schnell. Dass wiederum ein Ball, der auf den Boden fällt, deutlich stärker beschleunigt, findet sich in der Alltagssprache nicht wieder. Viele Schüler*innen sehen die Beschleunigung dementsprechend häufig als etwas Ähnliches wie die Geschwindigkeit und sind nicht in der Lage diese Größen zu unterscheiden.

Zusammenfassung

Wie eingangs bereits argumentiert, hängen viele Schülervorstellungen von anderen Vorstellungen ab. Vielen der genannten Vorstellung passen entweder zur Impetustheorie oder es liegen ihr die Vorstellungen zugrunde, dass eine Kraft in Bewegungsrichtung wirken muss. Diese beiden Vorstellungen sind ebenfalls eng miteinander verwandt und deuten auf ein mangelndes Verständnis des zweiten Newtonschen Gesetzes hin. Diese Vorstellungen sind historisch bei vielen prominenten Denkern vorhanden gewesen und auch heute noch in den Köpfen vieler Menschen verankert. Begründet werden kann dies durch Alltagserfahrungen, in denen scheinbar alle Objekte natürlicherweise zur Ruhe kommen. Dieses „natürlicherweise“ oder „von allein“ kann bzw. muss aber durch bewusstes Thematisieren der vielfältigen, in unserer Umwelt vorhandenen Reibungskräften aufgelöst werden. Es kann argumentiert werden, dass der Konzeptwechsel von der aristotelischen Sichtweise bzw. Impetustheorie zur Newtonschen Sichtweise die größte Herausforderung und wichtigste Aufgabe im Dynamikunterricht in der Schule darstellt und den größten potenziellen Fortschritt erzielt, was die Schülervorstellungen im Bereich der Dynamik betrifft. Eine Thematisierung von Reibung scheint demnach unumgänglich, gestaltet sich aber schwierig. Wie ein solcher Konzeptwechsel vonstattengehen kann, wird im folgenden Kapitel betrachtet.

3.3 Kontextabhängigkeit von Schülervorstellungen

Es ist bekannt, dass die Aktivierung von Schülervorstellung stark vom Kontext abhängig ist (Gunstone & White, 1981; z. B. Muis & Gierus, 2014). Dementsprechend ist nicht nur die Kenntnis über das Vorkommen gewisser Schülervorstellungen hilfreich für die Unterrichtsplanung, sondern auch deren Abhängigkeit vom besprochenen Kontext. Gezeigt werden konnte bereits, dass bei Veränderung der Konzepte und Kontexte Schülervorstellungen in der Mechanik in unterschiedlicher Weise aktiviert werden. Einfluss darauf haben können u. a. die Art (Palmer, 1997) und Richtung der Bewegung (horizontal/vertikal) (Ferreira et al., 2019; Lemmer, 2013) und bei 2D-Bewegungen die Ebene, in der die Bewegung stattfindet (horizontal/vertikal) (Lemmer, 2013) etc.

Im Bereich der Dynamik gibt es Erhebungen, die zeigen, dass es für Schüler*innen einen Unterschied macht, ob sie argumentativ von den wirkenden Kräften auf die Eigenschaften der kinematischen Größen der Bewegung schließen sollen oder umgekehrt. Speziell im Bereich des ersten Newtonschen Gesetzes mit einem sich mit konstanter Geschwindigkeit bewegendem Objekt haben Testitems, die von Kräften auf die Bewegung schließen, eine andere Lösungswahrscheinlichkeit als die analogen Testitems, die sich nur in ihrer

Argumentationsrichtung unterscheiden (Just et al., 2021). Erklärt wird dies von den Autor*innen durch die unterschiedliche Aktivierung von Schülervorstellungen. In diesem konkreten Beispiel fällt es Schüler*innen leichter von den wirkenden Kräften auf die Bewegung zu schließen. Dies könnte deshalb der Fall sein, da die Vorstellung, dass ein Objekt eine Kraft erfahren muss, damit es sich bewegen kann, unterschiedlich aktiviert wird. Die Autor*innen sehen die Begründung in der Alltagserfahrung, dass sich Objekte im Alltag häufig mit einer konstanten Geschwindigkeit bewegen, wenn man selbst eine konstante Kraft ausübt. Dies führt zu dem Schluss, dass auf ein Objekt, welches sich mit konstanter Geschwindigkeit bewegt, eine Kraft wirken muss. In der entgegengesetzten, erfolgreicheren Argumentationsrichtung könnte eher die Vorstellung aktiviert werden, dass sich nichts ändert, wenn keine Kraft wirkt (Alonzo & Aufschnaiter, 2018). Diese Vorstellungen führen in dem Beispiel zu einer richtigen Antwort, was die unterschiedlichen Lösungswahrscheinlichkeiten der Argumentationsrichtungen erklären könnte. Es ist allerdings nicht bekannt, welche Argumentationsrichtung Schüler*innen selbst bevorzugt wählen und ob sich die Ergebnisse auch auf Gespräche über physikalische Ergebnisse übertragen lassen.

3.4 Conceptual Change

Physikunterricht hat das Ziel, dass Schüler*innen im „Laufe der Schulzeit eine Struktur von möglichst konsistenten und vernetzten Wissenselementen erwerben“ (Hopf & Wilhelm, 2018, S. 26). Dazu ist es notwendig, dass sie neben ihren bestehenden Präkonzepten (situationsbedingt) die wissenschaftlich erwünschten Vorstellungen annehmen. Dieser Prozess wird als Conceptual Change oder Konzeptwechsel bezeichnet. Alltagsvorstellung und physikalische Vorstellung können parallel existieren, was meist sehr lang anhält. Aufgrund der Wichtigkeit der dort ablaufenden Prozesse stellt Conceptual Change ein eigenes Forschungsfeld dar, welches seit den 1980er-Jahren existiert (Potvin, Nenciovici et al., 2020). Konzeptwechselmodelle stellen einen Versuch dar, die Mechanismen und Faktoren zu berücksichtigen, die den Prozess einschränken oder beeinflussen, wie sich die anfänglichen Ideen von Schüler*innen zu wissenschaftlicheren Konzepten entwickeln (Potvin, Nenciovici et al., 2020; Schnotz, 1999). Da es auch in der vorgestellten Studie um das Verändern von hartnäckigen Präkonzepten bzw. Schülervorstellungen geht, wird hier kurz auf den aktuellen Stand im Bereich des Conceptual Change eingegangen.

Insgesamt können mittlerweile 86 verschiedene Conceptual Change-Modelle unterschieden werden, auf die eine Vielzahl von vorigen Arbeiten (z. B. von Piaget) Einfluss hatten (Potvin, Nenciovici et al., 2020). Hier wird nur auf die prominentesten Modelle

eingegangen. Diese wurden in einer Meta-Studie von Potvin, Nenciovici et al. (2020) auf Basis von positiven und negativen Erwähnungen und empirischen Überprüfungen in peer-reviewten Journals identifiziert und in der Reihenfolge nach der Häufigkeit ihrer Nennung in den Fachjournals vorgestellt. Da dies ein eigenes großes Forschungsfeld darstellt, können die einzelnen Modelle nur kurz umrissen werden. Abgeleitet werden sollen aus diesen Darstellungen die Design-Entscheidungen für die Interventionen.

Das am häufigsten referenzierte Konzeptwechselmodell ist von Posner et al. (1982). Konzeptwechsel findet demnach statt, indem von einer Menge an Konzepten auf eine andere Menge gewechselt wird, welche miteinander nicht kompatibel sind. Posner nennt hierfür folgende Bedingungen: 1. Lernende müssen mit den bisherigen Konzepten unzufrieden sein. 2. Die neuen Konzepte müssen plausibel, (3.) verständlich und (4.) erfolgreich erscheinen.

Nach Vosniadou (1994) ist ein Konzeptwechsel die graduelle Anpassung von mentalen Modellen der physischen Welt. Dies kann durch Anreicherung oder Überarbeitung der bestehenden mentalen Modelle passieren. Die Anreicherung stellt dabei eine Hinzunahme von neuen Informationen zu den bestehenden Strukturen dar, während die Überarbeitung mit einer Veränderung in den eigenen Vorstellungen oder in den Relationen in einer Struktur einer bestehenden Theorie einhergeht. Chi et al. (1994) unterscheiden alles in den drei ontologischen Kategorien Materie, Prozesse und mentale Zustände. Ein Konzeptwechsel bedeutet in diesem Modell, dass ein Konzept einer neuen Kategorie zugeordnet wird.

Hewson und Hewson (1983) nutzen eine ähnliche Terminologie wie Posner et al. (1982). Ein neues Konzept wird zu den bestehenden Konzepten hinzugefügt, wenn es plausibel, verständlich und erfolgsversprechend ist. Zentral ist hier aber die Aussage, dass ein neues Konzept vom Lernenden geblockt wird, wenn es den bereits vorhandenen Konzepten widerspricht. Damit ein Konzeptwechsel stattfinden kann, muss also der Status einer Vorstellung, die das neue Konzept blockiert, sinken, bevor der Status der neuen Vorstellung steigen kann. Hewson und Hewson (1983) nennen diesen Vorgang Konzeptaustausch. Die Hauptaussage von Pintrich et al. (1993) ist, dass ein Konzeptwechsel von der Interaktion von drei unterschiedlichen Arten von Faktoren bestimmt wird. Diese sind motivationale Faktoren, wie persönliche Interessen, kontextuelle Faktoren, wie z. B. Aufgabenstruktur, und kognitive Faktoren, wie die Aktivierung von Vorwissen. Zentraler Bestandteil des Konzeptwechselmodells (*Knowledge in Pieces*) von diSessa (1993) sind die sogenannten *p-prims*. *p-prims* sind fragmentartige Kausalzusammenhänge, die keine zusammenhängende Theorie ergeben und die Intuition im Bereich der physikalischen Welt darstellen. Ein Konzeptwechsel äußert sich in diesem Modell dadurch, dass diese *p-prims*

umorganisiert und in ein höheres kognitives System eingebettet werden. Das Konzeptwechselmodell von diSessa (1993) unterscheidet sich womöglich am meisten von den anderen und wird häufig auch als Gegenpol zum Konzeptwechselmodell von Vosniadou (1994) gesehen.

Alle die genannten Modelle haben Befürworter und werden teils durch empirische Evidenz gestützt. Teilweise gibt es Überschneidungen, sodass die Modelle als unterschiedliche Varianten eines gemeinsamen Modells gedeutet werden können. Teilweise sind die Unterschiede aber auch so groß, dass es vermeintlich nicht möglich sein kann, beide Modelle als eine angemessene Beschreibung der Realität zu sehen. Wie bei anderen Modellen passen auch die Konzeptwechselmodelle nicht in jeder Situation. Vielmehr muss je nach Lernprozess (z. B. einzelner Begriff vs. umfängliches Inhaltsgebiet) ein anderes Modell zur Beschreibung und Erklärung herangezogen werden (Hopf & Wilhelm, 2018). Zwei in der Physikdidaktik wichtige und relativ unterschiedliche Modelle für Konzeptwechsel sind das bereits oben kurz beschriebene synthetische Modell der Konzeptentwicklung nach Vosniadou (1994) und die Theorie der *Knowledge in Pieces* nach diSessa (1993). In der näheren Vergangenheit haben sich die Modelle grundsätzlich eher einander angenähert (Amin et al., 2014).

Diese unterscheiden sich vor allem darin, dass Vosniadou (1994) von konsistenten Präkonzepten in den Köpfen der Lernenden ausgeht, während diSessa (1993) einzelne unzusammenhängende Wissens Elemente (*p-prims*) annimmt, die eine einzelne Schlussfolgerung darstellen und in unterschiedlichen Situationen aktiviert werden können (z. B. „Von nichts kommt nichts.“). Auch hier hat im Laufe der Zeit aber eine Annäherung stattgefunden, da Vosniadou zunehmend eher von synthetischen Modellen spricht, die situationsbedingt unterschiedliche Vorstellungen aktivieren – im Gegensatz zu den ursprünglich postulierten starren und konsistenten Präkonzepten. Die Auswirkungen der beiden unterschiedlichen Modelle auf Lernkonzepte wird im folgenden Kapitel besprochen. Die Vorstellung dieser Konzeptwechselmodelle zielte vor allem darauf ab, daraus Empfehlungen für das Design von Lernumgebungen und Interventionen abzuleiten. Ein wichtiger Punkt hierbei ist u. a. der Umgang mit kognitiven Konflikten.

3.5 Umgang mit Schülervorstellungen

Wie diskutiert, ist es nicht das Ziel des Physikunterrichts, die Alltagsvorstellungen zu eliminieren, da dies ohnehin nicht möglich zu sein scheint (Kelemen & Rosset, 2009; Lombrozo et al., 2007). Ein bewusstes und reflektiertes Nebeneinander der Vorstellungen ist aber ein angestrebter Zustand. Ein Bewusstsein der typischen Alltagsvorstellungen

ermöglicht das Vermeiden gewisser Formulierungen auf der einen Seite und das aktive Ansprechen und Diskutieren gewisser Dinge auf der anderen Seite. Grundlegender kann aber die Frage nach der besten Art des Lernens und dem Umgang mit kognitiven Konflikten gestellt werden. Zu unterscheiden sind dabei kontinuierliche und diskontinuierliche Lernwege (Wilhelm & Schecker, 2018). Für einen diskontinuierlichen Lernweg typisch ist das Erzeugen von kognitiven Konflikten. Dies soll Unzufriedenheit der Lernenden mit ihren bestehenden Konzepten erzeugen. Der Konflikt wird dann durch einen Konzeptwechsel aufgelöst. Kontinuierliche Lernwege zeichnen sich wiederum durch eine schrittweise Entwicklung der bestehenden Schülervorstellungen in Richtung der gewünschten Vorstellungen aus. Es wird also nicht das Gefühl erzeugt, dass die Vorstellungen falsch sind, sondern an die ausbaufähigen Teilkonzepte angeknüpft und diese weiterentwickelt. Für welchen Lernweg man sich entscheidet hängt stark von dem akzeptierten Konzeptwechselmodell ab. Wenn Lernende, wie von Vosniadou (1994) propagiert, konsistente Konzepte besitzen, ist es schwierig die Notwendigkeit der Annahme der gewünschten Konzepte zu vermitteln, ohne dass die bestehenden Konzepte als unzureichend oder falsch angesehen werden. Es müssen also Fehler bzw. Unstimmigkeiten in den vorhandenen Konzepten aufgedeckt werden, was durch einen kognitiven Konflikt hervorgerufen werden kann. Dazu werden zu Beginn einer Unterrichtseinheit ausgewählte Versuche gemacht oder Inhalte diskutiert, die gerade im Widerspruch zu den Alltagsvorstellungen stehen, wonach die gewünschten Konzepte eingeführt werden. Dabei muss der kognitive Konflikt nicht nur objektiv vorliegen, sondern auch subjektiv wahrgenommen werden (Wilhelm & Schecker, 2018). Eine objektive Möglichkeit, Vorhersagen mit den tatsächlichen Ergebnissen zu vergleichen, stellt der Computer dar. Egal, ob dies mit Messwerterfassungssystemen, Videoanalyse oder auch mathematischer Modellbildung erfolgt, kann somit ein Experiment durchgeführt und Vorhersagen bezüglich gewisser Größen getroffen werden. Schüler*innen können dann das Experiment selbstständig durchführen und die Ergebnisse mit ihren Vorhersagen vergleichen. Dies hat den Vorteil, dass die Lehrkraft nicht in bewertender Funktion tätig sein muss. Bei dem Vorgehen bleibt die Frage, an welchem Punkt im Laufe der Unterrichtseinheit und in welcher Art und Weise mit einem kognitiven Konflikt gearbeitet wird, da das Vorgehen auch Nachteile haben kann. Einerseits kann eine aktive Äußerung der ungewollten Vorstellung diese erst manifestieren. Dies ist vor allem dann ein Problem, wenn es sich um eine Vorstellung handelt, die nicht so verbreitet ist, dass sie in der Auseinandersetzung mit der Thematik ohnehin aktiviert worden wäre. Auch die falschen Vorhersagen anderer Schüler*innen können bei Lernenden ähnlich in Erinnerung bleiben, wie die gewünschte Vorstellung nach der Besprechung des Experiments. Deutlich muss auch gemacht werden, dass es einen Unterschied zwischen Vorhersage und Ausgang des Experiments gibt, da die

Wahrnehmung von Abläufen stark von der Erwartung der Lernenden beeinflusst sein kann. Bekannt ist, dass bei Erkennen einer Diskrepanz möglicherweise eine Erklärung gesucht wird, warum die Vorhersage in dieser Situation nicht zutreffend war, ohne dass die eigene Vorstellung falsch ist. Dadurch wird diese dann nicht geändert (Duit, 1993; Wilhelm & Schecker, 2018). Einerseits muss viel Zeit für diesen Vorgang eingeplant werden, da ein kognitiver Konflikt in mehreren Situationen verdeutlicht werden muss, andererseits muss den Lernenden dann das Handwerkszeug für eine alternative Theorie gegeben werden.

Ein anderer Ansatz, der eher mit dem *Knowledge in Pieces*-Modell von diSessa (1993) vereinbar ist, ist an ausbaufähigen Stellen anzusetzen ohne direkt auf Fehler in den bestehenden Vorstellungen hinzuweisen⁷. Das ergibt dann Sinn, wenn man davon ausgeht, dass Lernende gewisse einfache Schlussregeln (*p-prims*) besitzen, die in unterschiedlichen Kontexten unterschiedlich aktiviert werden, es sich also nicht um konsistente Vorstellungen bzw. Konzepte handelt. Ziel ist dann, mit der richtigen Strategie die richtigen *p-prims* in der richtigen Situation zu aktivieren (Wiesner et al., 2010). Das heißt auch, dass es vermieden wird, ungeeignete *p-prims* zu aktivieren, was wiederum der kognitiven Konfliktstrategie widerspricht. Alternativ können diese aber auch umgedeutet, statt vermieden werden. Diese Strategie kann vor allem bei Themen sinnvoll sein, in denen keine weit verbreiteten und den Lernprozess stark hindernden Schülervorstellungen bekannt sind.

Nach Potvin (2017) sollte ohnehin eher von konzeptueller Prävalenz (Conceptual Prevalence) statt Konzeptwechsel (Conceptual Change) gesprochen werden, da Alltagsvorstellungen lebenslang erhalten bleiben. Forschungen zeigen, dass auch Expert*innen lediglich die Fähigkeit entwickeln, diese in den gewünschten Situationen zu unterdrücken (Potvin, Malenfant-Robichaud et al., 2020). Für den Unterricht heißt das, die beiden Konzeptionen stehen stets in einem Wettstreit (Potvin et al., 2015). Bei der Schwächung eines Konzepts ist es wichtig, dass ein alternatives Konzept bereits zur Verfügung steht. Potvin (2017) plädiert daher auch für einen veränderten Umgang mit kognitiven Konflikten. Demnach sollten kognitive Konflikte kontextbasiert genutzt werden, nicht um den Status der Alltagstheorie vollständig zu mindern, sondern um zu verdeutlichen, dass die Alltagstheorie in diesem Kontext keine sinnvolle Beschreibung der Realität erlaubt. Ziel ist, dass sich die gewünschten Konzepte langfristig durch eine Überlegenheit in der kognitiven Nützlichkeit gegenüber den Alltagsvorstellungen in den entsprechenden Situationen

⁷ Wie ein Unterrichtskonzept aussehen kann, welches auf dieser Strategie nach diSessa (1993) aufbaut, wird dargestellt in: Burde und Wilhelm (2020).

durchsetzen. Hier kann also von einer Kombination der beiden Strategien gesprochen werden.

Zusammenfassend kann man sagen, dass sich kognitive Konflikte vor allem dann anbieten, wenn es um Schülervorstellungen geht, die sehr weit verbreitet sind. Zudem sollten kognitive Konflikte nicht direkt am Anfang stehen, wenn die Lernenden noch nicht das gewünschte Konzept zur Erklärung des Phänomens zur Verfügung haben. Beachtet werden muss dabei, dass sich kognitive Konfliktstrategien gerade bei schwächeren Schüler*innen u. a. negativ auf das Selbstbild auswirken können (Dreyfus et al., 1990). Eine Verbindung kontinuierlicher und diskontinuierlicher Strategien erscheint erfolgversprechend. Man verfolgt zunächst die Aufbaustrategie und arbeitet nach dem Aufbau des neuen Wissens gezielt mit kognitiven Konflikten in einem engen Rahmen.

Es bedarf noch mehr Forschung, um herauszufinden, in welcher Situation welche Lernstrategie sinnvoll und zielführend ist. Dies ist von dem zugrundeliegenden Konzeptwechselmodell, aber auch von der Lernbedingungen, der Lerngruppe, dem Lerngegenstand etc. abhängig. Die Forschungsergebnisse können jedoch so interpretiert werden, dass eine Strategie, die darauf aufbaut, anschlussfähige Wissens Elemente herauszugreifen, diese zu aktivieren und zu verstärken, vor allem bei Themen gelingt, in denen keine stark verbreiteten lernhinderlichen Vorstellungen vorliegen. Für die Variante eines kognitiven Konfliktes sprechen Lerninhalte, in denen ein Großteil der Lernenden eine bestimmte Vorstellung hat, die physikalisch nicht erwünscht ist. Diese kann dann durch einen kognitiven Konflikt geschwächt werden. Gleichzeitig sollte aber eine Alternative für die Lernenden vorliegen, sodass nicht mit einem großen kognitiven Konflikt begonnen werden sollte. Vorteilhaft könnte es auch sein, wenn die Schüler*innen in einer solchen Situation ihre Vorhersagen wertungsfrei äußern können und wertungsfrei eine Rückmeldung über die Richtigkeit der Vorhersage erhalten. Dies kann durch den Computereinsatz allgemein und speziell durch die eigenständige Arbeit mit dem Computer gelingen.

4 Lernen mit neuen Medien

Neben den spezifischen fachdidaktischen Vorteilen gewisser Arten des Computereinsatzes, auf die in Kapitel 5.4 und 6.2 eingegangen wird, hat das Lernen mit neuen Medien gewisse allgemeine Charakteristika. In diesem Kapitel wird beleuchtet, was über den Einsatz von neuen Medien im mathematisch-naturwissenschaftlichen Unterricht bekannt ist und wie sich dies auf Interesse und Motivation der Lernenden auswirkt. Nach der Vorstellung von zwei Lerntheorien mit Bezug zu digitalen Medien werden Design-Prinzipien abgeleitet, nach denen multimediale Lernumgebungen gestaltet werden sollten, um einen möglichst großen Lernerfolg zu erzielen. Diese Design-Prinzipien wurden bei der Erstellung der im Zuge dieser Studie erforschten Lernumgebungen ebenso berücksichtigt wie die Schülervorstellungen (siehe Kapitel 3.2) und die Conceptual Change-Modelle (siehe Kapitel 3.4).

Die Nutzung von digitalen Medien hat einen mittleren bis starken positiven Einfluss auf den Lernerfolg. Diese Erkenntnis basiert auf einer Meta-Studie von Hillmayr et al. (2017), in der 79 Primärstudien untersucht wurden. Festgestellt wurde bei dieser Analyse, dass im Bereich Mathematik, Physik, Biologie und Chemie insgesamt ein Vorteil von dem Einsatz digitaler Medien gegenüber der herkömmlichen Unterrichtsweise zu beobachten war. Sung et al. (2016) beziffern den positiven Effekt von mobilen digitalen Medien auf den Lernerfolg gegenüber herkömmlichen Methoden auf $g = 0,52$. Zusätzlich wurde eine Effektstärke von $g = 0,45$ für die kollaborative Nutzung des Computers auf den Lernzuwachs beobachtet (Chen et al., 2018). Hedges g^8 stellt dabei ein Maß für den Lernzuwachs bezogen auf die Streuung innerhalb der Stichprobe dar (Hedges, 1981) und ähnelt der Effektstärke d nach Cohen (1988). Hillmayr et al. (2020) nennen in einer neueren Meta-Studie sogar insgesamt einen positiven Effekt von $g = 0,65$ für die Nutzung von digitalen Medien in den mathematisch-naturwissenschaftlichen Fächern. Hier wurden aber viele unterschiedliche Arten des Computereinsatzes in unterschiedlichen Fächern und Kontexten mit unterschiedlichen Zielen zusammengefasst.

Hillmayr et al. (2017) haben zusätzlich einige Gelingensbedingungen für den Einsatz neuer Medien formuliert. Es scheint vorteilhaft zu sein, wenn ihr Einsatz ergänzend statt ersetzend zu herkömmlichen Unterrichtsmethoden erfolgt. Auch die Dauer des Einsatzes

⁸ $(\mu_1 - \mu_2) \cdot \left(\frac{(n_1-1)\sigma_1^2 + (n_2-1)\sigma_2^2}{n_1+n_2-2} \right)^{-1/2}$ mit μ für die Mittelwerte und σ für die Standardabweichungen zu den zwei Zeitpunkten.

spielt eine Rolle. Je länger der Einsatz der digitalen Medien andauert, desto geringer ist der beobachtete Effekt. Dies lässt sich mit dem Neuheitseffekt solcher Software erklären. Auch wenn bei längerer Nutzung die Auswirkungen geringer ist, so gibt es dennoch weiterhin einen positiven Effekt, auch wenn der Neuheitseffekt verflogen ist. Ein positiver Effekt zeigt sich zusätzlich, wenn die Lehrkraft vor dem Einsatz der digitalen Medien eine Schulung dazu besucht hat. In Gruppen (vor allem zu zweit an einem Gerät) lernen Schüler*innen mit neuen Medien besser als allein (Carnes et al., 1987; Hillmayr et al., 2017), was die Autor*innen im Austausch zwischen den Schüler*innen während der Arbeitsphasen begründet sehen. Nichtsdestotrotz ist eine gute Steuerung bzw. Unterstützung durch die Lehrkraft (oder auch durch Mitschüler*innen) wichtig (Hillmayr et al., 2017). Der Einsatz von digitalen Medien kann die Motivation für das jeweilige Unterrichtsfach erhöhen, wenngleich dies teilweise ebenfalls durch den Neuheitseffekt verursacht wird. Auch wenn wenig Studien dazu vorliegen, ergab eine Expertenbefragung, dass sich digitale Medien auch sehr gut zur Differenzierung innerhalb einer Lerngruppe eignen (Hillmayr et al., 2017). Die größten beobachteten Lernzuwächse ergaben sich im Bereich der dynamischen mathematischen Werkzeuge, zu denen die Autor*innen u. a. die Software GeoGebra zählen (Hillmayr et al., 2020). Zusätzlich wurden im Fach Physik und in den höheren Jahrgangsstufen (gymnasiale Oberstufe, Jahrgangsstufen 11-13) die größten Effektstärken beobachtet (Hillmayr et al., 2020).

In der Literatur ergibt sich grundsätzlich ein sehr positives Bild, was den Einsatz von neuen Medien im (Physik-)Unterricht angeht. Der Einsatz von digitalen Medien allein ist aber kein Garant für Erfolg. Vielmehr muss sich der Einsatz der Medien an dem Unterrichtsgegenstand und der Lerngruppe orientieren. Verschiedene Lerntheorien treffen darüber hinaus auch im Bereich des Einsatzes von digitalen Medien konkrete Aussagen. Mit diesen Lerntheorien kann auch der Erfolg von neuen Medien in gewissen Situationen erklärt werden. In der Folge werden zwei Lerntheorien vorgestellt, die eng mit dem Einsatz digitaler Medien verknüpft sind, woraufhin Design-Prinzipien abgeleitet werden, die in entsprechenden Lernumgebungen beachtet werden sollten.

4.1 Cognitive Load Theory

Die grundlegende Annahme der Cognitive Load Theory (CLT) ist, dass Menschen eine limitierte Verarbeitungskapazität ihres Arbeitsgedächtnisses haben und neues Wissen in Schemata abspeichern (Sweller et al., 2011). Die Kapazität des Arbeitsgedächtnisses muss bei dem Design von Lernumgebungen berücksichtigt werden, sodass dieses nicht überlastet wird. Konkret werden dabei drei Arten der kognitiven Belastung unterschieden. Die intrinsische kognitive Belastung (Intrinsic Cognitive Load, ICL) bezeichnet die dem

Lerngegenstand inhärente kognitive Belastung, welche von dessen Schwierigkeit und dem Vorwissen des Lernenden abhängt. Die ICL ist stark davon abhängig, wie vernetzt ein Lerngegenstand ist (Sweller, 2003). Vokabelartige Lernelemente, die nur schwach vernetzt sind, verursachen eine geringe kognitive Belastung, während stärker vernetzte Lernelemente, wie sie beim Verständnislernen vorkommen, eine höhere kognitive Belastung mit sich bringen. Ein weiterer relevanter Faktor ist das Vorwissen der lernenden Person. Je höher das Vorwissen ist, desto geringer kann die intrinsische kognitive Belastung bei gleichem Lerngegenstand werden. Die extrinsische oder lernirrelevante kognitive Belastung (Extraneous Cognitive Load, ECL) wird von der Lernumgebung, vor allem von ihrer Darstellung und Gestaltung beeinflusst. Überflüssige Informationen führen zu einer Erhöhung der extrinsischen kognitiven Belastung. Diese Belastung ist vor allem auch bei der Nutzung von Multimediaumgebungen zu beachten, da diese viele Möglichkeiten der Darstellungen von Lerninhalten ermöglichen. Die dritte Art der kognitiven Belastung ist die lernrelevante kognitive Belastung (Germane Cognitive Load, GCL), welche für die Übertragung von Wissen in das Langzeitgedächtnis verantwortlich ist. Diese muss dementsprechend gefördert werden, u. a. dadurch, dass vor allem der ECL geringgehalten wird. Dies kann durch die Beachtung verschiedener Effekte gelingen.

Bedeutsam für die Gestaltung von Lernumgebungen ist der Split-Attention-Effekt, zu dem es bei der Trennung von verschiedenen Informationsquellen kommt. Wenn das der Fall ist, müssen mentale Integrationsprozesse stattfinden, die den ECL erhöhen und damit den Lernerfolg mindern (Sweller & Chandler, 1994). Der Lösungsbeispieleffekt (*worked example effect*) ermöglicht insbesondere zu Beginn einer Lerneinheit eine Reduktion des ECL durch Nutzung von Beispiellösungen, die weniger mentale Integrationsprozesse erfordern als Problemlösungsaufgaben (Hollender et al., 2010). Weitere Effekte, die aus der CLT folgen, aber speziell in multimedialen Settings relevant sind, werden im folgenden Kapitel besprochen.

Auch wenn die CLT keine spezielle Theorie für die Nutzung von digitalen Medien ist, haben ihre Grundannahmen einen großen Einfluss auf den Erfolg beim Einsatz von digitalen Medien. Programme sollten entsprechend gestaltet sein und die Lehrkraft sollte die Grundannahmen bei der Gestaltung des Unterrichts berücksichtigen. Aus der CLT lässt sich für den Einsatz digitaler Medien also vor allem ableiten, dass die Lernumgebung so gestaltet sein sollte, dass der ECL möglichst gering ist, was in der Theorie einen hohen GCL und damit einen hohen Lernzuwachs ermöglicht. Der GCL kann aber auch spezifisch gefördert werden (Paas & van Merriënboer, 1994).

4.2 Cognitive Theory of Multimedia Learning

Die Cognitive Theory of Multimedia Learning (kognitive Theorie multimedialen Lernens, CTML) (Mayer, 1999, 2003, 2009) ist eine Theorie, die sich mit dem Lernen in multimedialen Settings beschäftigt. Sie baut auf der These der Cognitive Load Theory auf, dass das Arbeitsgedächtnis eine begrenzte Kapazität besitzt, wodurch die kognitive Belastung eine zentrale Rolle spielt (Mayer & Moreno, 2003). Zusätzlich wird das Arbeitsgedächtnis in zwei Kanäle unterteilt, die jeweils ebenfalls begrenzt sind. Diese Kanäle unterscheiden sich darin, welche Informationen in ihnen verarbeitet werden. Während ein Kanal für die Verarbeitung von verbalisierten Informationen verantwortlich ist, verarbeitet der andere Kanal visuell-bildhafte Informationen. Nach der CTML sollte eine Lernumgebung beide Kanäle während eines Lernprozesses ansprechen (Mayer, 2014).

Den Split-Attention-Effekt versucht die CTML durch das Kontiguitätsprinzip zu vermeiden (Mayer, 2014, 2017). Dies bedeutet, dass zusammenhängende Informationen, die verarbeitet werden müssen, räumlich (räumliches Kontiguitätsprinzip) und zeitlich (zeitliches Kontiguitätsprinzip) nicht voneinander getrennt sein sollten (Moreno & Mayer, 1999). Während eine räumliche Trennung durch Suchen der relevanten Informationen zu einer Erhöhung der lernirrelevanten (ECL) kognitiven Belastung führt, verursacht eine zeitliche Trennung die Notwendigkeit, dass eine mentale Repräsentation länger aufrechterhalten werden muss, was ebenfalls zu einer höheren kognitiven Belastung führt. Eine weitere Annahme der CTML ist das Interaktivitätsprinzip. Dieses besagt, dass es lernförderlich ist, wenn Lernende den Ablauf einer Lernumgebung selbst bestimmen können, da sie so eine kognitive Überlastung vermeiden können (Mayer & Chandler, 2001).

4.3 Design-Empfehlungen für digitale Lernumgebungen

Digitale Lernumgebungen sollten so gestaltet sein, dass sie den beiden Kontiguitätsprinzipien nachkommen, den Nutzer*innen eine eigene Steuerung ermöglichen und damit den ECL reduzieren. Die Informationen sollten so präsentiert sein, dass die Nutzer*innen nicht dazu gezwungen sind, Informationen im Arbeitsgedächtnis zu speichern, die für die Erklärung eines später relevanten Sachverhalts benötigt werden. Auch ein räumliches Trennen von zusammengehörenden Informationen sollte durch die Gestaltung der Software vermieden werden. Nach Möglichkeit ist die Lernumgebung selbst steuerbar (also interaktiv), was z. B. bei Lernvideos nur bedingt der Fall ist.

Darüber hinaus folgen aus der CTML weitere Design-Prinzipien. Textuelle Informationen sollten stets durch bildliche Informationen ergänzt werden, da so beide Kanäle angesprochen werden und damit am besten gelernt werden kann. Das Kohärenzprinzip besagt,

dass bildliche Informationen, welche für die angestrebten Lernziele irrelevant sind, die Lernleistung reduzieren. Empfohlen wird außerdem, textuelle Informationen als eingesprochene auditive Informationen zu übermitteln (Clark & Mayer, 2016), damit beide Kanäle optimal genutzt werden können. Redundanzen, die sich durch die gleichzeitige Darstellung von Text als auditive und visuelle Informationen ergeben, sollten vermieden werden (Clark & Mayer, 2016). Diese Design-Grundsätze sind vor allem für schwächere Schüler*innen relevant, da stärkere Schüler*innen ungünstige Lernumgebungen eher durch Vorwissen ausgleichen können. Die Auswahl der Software für die Nutzung in dieser Studie und die Gestaltung der Lernmaterialien wurden auf Basis der Grundsätze dieser Theorien getroffen

.

5 Modelle und Modellieren

5.1 Modelle in der Physik

Modelle spielen im Physikunterricht eine große Rolle. In jedem Teilgebiet und in jeder Klassenstufe werden Modelle im Physikunterricht genutzt, um Sachverhalte darzustellen, zu veranschaulichen, Vorhersagen zu treffen, den Ausgang eines Experimentes zu erklären etc. Dabei folgt der Physikunterricht in dieser Hinsicht der Fachdisziplin Physik. Auch hier sind Modelle für die Erkenntnisgewinnung und damit den Fortschritt der Wissenschaft essenziell. In der Schule geht es somit einerseits um das „Lernen mit Modellen“, was bedeutet, dass Modelle von der Lehrkraft gezielt genutzt werden, um den Schüler*innen das Lernen von Fachinhalten zu erleichtern und um das „Lernen von Modellen“. Andererseits ist ein wichtiger Teil, der von der Physikdidaktik für den Unterricht gefordert wird, auch das „Lernen über Modelle“, also das Diskutieren, was ein Modell überhaupt ist, welche Eigenschaften und Grenzen es hat und welche Idealisierungen für das Modell aus welchem Grund vorgenommen wurden. Dies kann anhand einzelner Modelle exemplarisch diskutiert werden, führt dann aber auf eine Thematisierung von epistemologischen Aspekten im Allgemeinen hinaus.

Ein besseres Modellverständnis führt zu einem besseren Wissenschaftsverständnis (Leisner-Bodenthin, 2006). Ein gute Modellkompetenz könnte also einerseits den Lernerfolg steigern, ist andererseits bereits für sich genommen ein Lernziel des Physikunterrichts, da Modellverständnis notwendig ist für Wissenschaftsverständnis (Leisner-Bodenthin, 2006). Zu definieren, was ein Modell ist und welche Rolle es im Erkenntnisgewinnungsprozess spielt, ist dabei keineswegs einfach. Kircher (2007) stellt dabei die große Spannweite des Modellbegriffs dar. Modelle im engeren Sinn sind demnach Modelle, die eine Veranschaulichung für einen bestimmten Aspekt darstellen und auch als solche bezeichnet werden, wie etwa das Teilchenmodell. Umfassender kann aber noch mehr als Modell bezeichnet werden, wenn man berücksichtigt, dass „unser Wissen über die Realität prinzipiell hypothetisch ist“, was ebenfalls eine Eigenschaft von Modellen ist (Kircher, 2007, S. 680). In diesem Fall lassen sich dann Hypothesen und sogar zuverlässige Theorien unter den Modellbegriff fassen. Um eine konkrete Definition des Modellbegriffs liefern zu können, muss zunächst eine Lerntheorie zugrunde gelegt werden, da das Wesen und der Zweck von Modellen davon abhängen, wie die Existenz von Phänomenen und der damit zusammenhängende Lernprozess bewertet wird. Kircher, Girwidz und Häussler (2007) sprechen sich dabei für den indirekten Realismus aus, nach dem Objekte in der realen Welt zwar unabhängig vom Beobachter oder der Beobachterin existieren, diese aber nicht vollständig wahrgenommen werden können. Alternative

Lerntheorien sind der naive Realismus und der Konstruktivismus, die andere Grundaussagen zum Verhältnis von Subjekt und existierender Realität haben. Kircher, Girwidz und Häussler (2007, S. 683) kommen unter Voraussetzung des indirekten Realismus zu der Definition, dass ein Modell M „ein von einem Subjekt S für bestimmte Zwecke und für eine bestimmte Zeit benutzter bzw. geschaffener Gegenstand oder theoretisches Konstrukt derart [ist], dass zu bestimmten Elementen von M Analogien zu Elementen des Objekts O bestehen“. Die Art der Analogie kann dabei ebenso umfassend diskutiert werden. Eine mögliche Annahme ist ein partieller Isomorphismus. Dieser verdeutlicht, dass gewisse Merkmale des Modells auf Merkmale des Objekts bzw. Phänomens abgebildet werden können und umgekehrt, es aber sowohl Merkmale des Modells gibt, die das Objekt nicht hat, als auch andersherum (Vollmer, 2008). Dabei lässt sich diese Relation nur im naiven Realismus, nicht aber im indirekten Realismus als Abb. im mathematischen Sinne sehen (Bunge, 1973). Im letzteren ist die Relation nur im metaphorischen Sinne homomorph (Kircher, Girwidz & Häussler, 2007). Dabei sollen Modelle anschaulich, vertraut, einfach und transparent sein (Kircher, Girwidz & Häussler, 2007). Sie erfüllen je nach Ziel unterschiedliche Funktionen. Entweder man kennt bereits ein Modell und möchte damit das Verhalten eines Objekts vorhersagen oder die Merkmale eines Objekts sollen mithilfe des Modells erklärt werden. Zudem haben Modelle in der Physikdidaktik auch lernökonomische Funktionen (Kircher, 2007). Mikelskis-Seifert et al. (2005) klassifizieren Modelle nach den Zielen, die mit ihnen verfolgt werden. So gibt es einerseits Modelle, die als Mittel zur Veranschaulichung genutzt werden. Weiter gibt es Modelle, die als Mittel zur Vereinfachung (Idealisierung) genutzt werden, und Modelle, die zur Erklärung beobachteter Phänomene, zur Erforschung komplexer Zusammenhänge und zur Generierung neuer Erkenntnisse dienen. Auf eine umfassendere Diskussion von Modelldefinitionen und dem Verhältnis von Subjekt, Modell und Objekt je nach zugrundeliegender Lerntheorie wird an dieser Stelle verzichtet. Es ist zu beachten, dass sowohl in der Alltagssprache als auch in der Wissenschaft der Modellbegriff (ebenso Modellbildung und Modellierung) teilweise sehr unterschiedlich verwendet wird (Mikelskis-Seifert et al., 2005). Schecker (1998a) unterscheidet drei unterschiedliche Arten von Modellen: Gegenständliche Modelle, Gedankenmodelle und mathematische Modelle. Vor allem letztere Modelle, welche nach dem Autor ein Satz von Differentialgleichungen und funktionalen Beziehungen sind, der Vorhersagen über den Ablauf von physikalischen Prozessen ermöglicht (Schecker, 1998a), sind für den Inhalt dieser Arbeit relevant.

Modellkompetenz besteht nach Leisner-Bodenthin (2006) aus deklarativem Wissen, prozeduralem Wissen und dem Selbstständigkeitsgrad. Dem deklarativen Wissen ist neben spezifischen Modelleigenschaften das allgemeine Modellverständnis zugeordnet. Unter Modellverständnis versteht die Autorin das Wissen, dass Modelle von Menschen gemacht

sind, wenn die Grenzen der Wahrnehmbarkeit erreicht sind oder nicht wahrnehmbare Mechanismen erklärt, vorhergesagt oder veranschaulicht werden sollen. Modelle dienen also der Beschreibung und Veranschaulichung bekannten Wissens, aber auch der Erschließung neuen Wissens. Zudem sind Idealisierungen notwendig. Modelle werden außerdem für einen bestimmten Zweck erschaffen, sind vorläufig und müssen sich in der Community durchsetzen (Leisner-Bodenthin, 2006). Das prozedurale Wissen stellt die jeweilige Anwendung (eines konkreten Modells und des Modellverständnisses) dar. Dabei kann noch zwischen domänenspezifischer und domänenübergreifender Modellkompetenz unterschieden werden, da das Vorhandensein von Modellkompetenz in Bezug auf ein bestimmtes Modell nicht zwangsläufig eine generelle Modellkompetenz nach sich zieht (Mikelskis-Seifert & Leisner, 2005). Studien zeigen, dass ein besseres Modellverständnis und eine gesteigerte Modellkompetenz durch entsprechenden Unterricht bzw. Interventionen erzielt werden können. Ebenso ist bekannt, dass Schüler*innen und auch Lehrkräfte häufig ein unzureichendes Bild von der Bedeutung von Modellen in der Erkenntnisgewinnung haben (Crawford & Cullin, 2004; Crawford & Cullin, 2005; Meisert, 2008; Mikelskis-Seifert et al., 2005; Trier & Upmeier zu Belzen, 2009). Insbesondere dieser Aspekt sollte also stärker betont werden.

5.2 Modellbildung

Da Modelle so eine umfassende Bedeutung besitzen, ist nach Kircher, Girwidz und Häussler (2007) „Physik lernen“ immer „physikalische Modellbildung“. Modellbildung kann als Prozess der Erstellung oder Konstruktion von Modellen gesehen werden oder auch deren Nutzung und Anpassung beinhalten (Gilbert & Justi, 2016). Upmeier zu Belzen und Krüger (2010) ordnen die Modellbildung als Teil der Modellkompetenz ein und sehen folgende Teilkompetenzen als Teil der Modellbildung: Zweck von Modellen, Testen von Modellen und Ändern von Modellen (Vgl. Crawford & Cullin, 2005; Grosslight et al., 1991; Justi & Gilbert, 2003). Damit besteht die Modellkompetenz insgesamt aus dem Wissen über Modelle und der Modellbildung. Modellbildung stellt den Umgang mit Modellen im Modellbildungsprozess dar und ist folglich Teil des prozeduralen Wissens (Justi & Gilbert, 2002). Es liegt hier also ebenfalls eine Trennung zwischen deklarativem und prozeduralem Wissen vor. Im Modellbildungsprozess ist es wichtig, dass Hypothesen formuliert, Prognosen mit dem Modell erstellt und diese mit der Wirklichkeit verglichen werden. Zudem sollten die Gültigkeit eines Modells hinterfragt, die Grenzen des Modells aufgezeigt und die eigene Vorgehensweise reflektiert werden (Mikelskis-Seifert et al., 2005). Hilfreich kann dabei nicht zuletzt die Nutzung von Computern sein (Mikelskis-Seifert et al., 2005).

Auch wenn Modellbildung im Physikunterricht ähnliche Ziele verfolgt, wie die Modellbildung bzw. Theoriebildung in der Fachwissenschaft, und teilweise von Physikdidaktiker*innen gefordert wurde, dass der Physikunterricht die Methodologie und die Wissensbestände der physikalischen Fachwissenschaft erarbeitet, scheint dies aus verschiedenen Gründen nicht möglich zu sein. So spielen vor allem individuelle Bedingungen (z. B. Alltagsvorstellungen) der Schüler*innen im Unterricht eine große Rolle, während die allgemeinen Bedingungen (Voraussetzungen, Hintergründe, Anlässe der Modellbildung) weniger Bedeutung haben (Kircher, Girwidz & Häussler, 2007). Zudem ist die wissenschaftliche Methode in der Realität nicht algorithmusartig und damit nicht gut planbar (Kircher, Girwidz & Häussler, 2007). Physikalische Theorie- bzw. Modellbildung ist also nicht unmittelbar für den Physikunterricht geeignet, beeinflusst aber dessen Lernziele. Modellbildung selbst ist ein wichtiger Bestandteil des Physikunterrichts. Diese orientiert sich an der Modellbildung aus der Fachwissenschaft, bildet sie aber nicht vollständig ab. Insgesamt ist es wichtig, im Unterricht die Bedeutung von Modellen und Modellbildung für den naturwissenschaftlichen Erkenntnisprozess zu diskutieren.

5.3 Zusammenspiel von Modell und Experiment

Modelle spielen im Erkenntnisgewinnungsprozess der Naturwissenschaften eine zentrale Rolle. Um dies zu verdeutlichen, ist insbesondere das Verhältnis von Modell bzw. Theorie und Experiment relevant. Der Erkenntnisgewinnungsprozess stellt dabei klassisch einen iterativen Prozess zwischen Theoretisieren und Experimentieren dar (Mikelskis-Seifert et al., 2005). Experimente sind dabei selbst Lernobjekte, dienen zusätzlich aber auch als Lernhilfe (Kircher, Girwidz & Häussler, 2007). Der Einsatz von quantitativen Experimenten im Physikunterricht verdeutlicht dabei das Wechselspiel zwischen Theorie und Experiment, die Bedeutung quantitativer Experimente für die Entwicklung der Physik, das Ringen um sinnvolle Daten und deren Interpretation und das genaue Beobachten und Experimentieren (Kircher, Girwidz & Häussler, 2007). Im Unterricht wird allerdings häufig der Eindruck vermittelt, dass ein Experiment ein zuvor entwickeltes Modell unmittelbar beweist oder die Daten eines Experiments eindeutig auf ein bestimmtes Modell hinweisen, was nicht der Fall ist, sich aus Gründen des Zeitmangels im Unterricht aber teilweise auch nicht anders darstellen lässt.

Mittlerweile ergänzt der Computer in der Fachwissenschaft immer häufiger dieses Wechselspiel zwischen Theorie und Experiment (Teodoro & Neves, 2011) und ersetzt letzteres teilweise komplett. Diese Entwicklung in der Fachwissenschaft (Caballero & Merner, 2018) sollte sich auch im Physikunterricht wiederfinden. Wenngleich im Unterricht selten Themen behandelt werden, in denen ein ComputermodeLL das Experiment ersetzen muss

(z. B. Entstehung des Universums), bietet sich ein ergänzender Einsatz des Computers zur Veranschaulichung, als Hilfsmittel etc. häufig an. Eine Möglichkeit, den Computer zum Modellieren im Wechselspiel zwischen Theorie und Experiment einzusetzen, stellt die mathematische Modellbildung mit entsprechender Modellbildungssoftware dar.

5.4 Mathematische Modellbildung

5.4.1 Definition

Mathematische Modellbildung meint in diesem Kontext die „Konstruktion eines Netzwerks physikalischer Begriffe und Beziehungen, mit denen das Verhalten eines physikalischen Systems beschrieben und vorhergesagt werden kann“ (Schecker, 1998a, S. 23). Mathematische Modellbildungssysteme sind Computerprogramme, mit denen die modellierten Zusammenhänge dieses physikalischen Systems numerisch berechnet werden können. Damit werden bei der mathematischen Modellbildung Modelle genutzt, um die Realität zu beschreiben. Diese Modelle werden von den Nutzer*innen (und damit von den Schüler*innen) aktiv konstruiert. Dieser Punkt unterscheidet die mathematische Modellbildung von der Nutzung von Simulationen, da im letzteren Fall die der Simulation zugrundeliegenden Modelle nicht selbst erstellt werden, worin der primäre Bildungswert der Modellbildung liegt (Hucke, 2000; Schecker, 1993b). Dadurch verfolgen diese beiden Arten des Computereinsatzes unterschiedliche Ziele. In der Praxis kommen verschiedene Arten von Modellbildungsprogrammen in Frage, von denen manche eher die qualitativen Zusammenhänge fokussieren, während andere unmittelbar mit gleichungsbasierten Zusammenhängen arbeiten (Kapitel 5.4.3). Nachfolgend werden außerdem die didaktischen Gründe für den Einsatz von mathematischer Modellbildungssoftware beleuchtet (Kapitel 5.4.4). Dabei wird der Einsatz von Modellbildungssoftware im Kontext der Mechanik beleuchtet, obgleich sich auch andere Teilgebiete der Physik für den Einsatz von Modellbildungssoftware eignen. Außerdem werden die Forschungsergebnisse im Bereich der mathematischen Modellbildung dargestellt (Kapitel 5.4.5).

5.4.2 Numerische Verfahren

Numerische Methoden sind für die Benutzung von mathematischer Modellbildungssoftware essenziell. Die beim Modellieren entstehenden Differentialgleichungen haben durch ihre Form oft keine geschlossene Lösung. Weg- und geschwindigkeitsabhängige Kräfte führen beispielsweise dazu, dass die Differentialgleichung nicht mehr klassisch gelöst werden kann. In einem solchen Fall können durch numerische Lösungsverfahren Näherungslösungen berechnet werden. Die dabei möglichen Berechnungsverfahren sind

vielfältig. Für die gängige Modellbildungssoftware spielen vor allem das Euler-Verfahren und das Runge-Kutta-Verfahren eine wichtige Rolle (Hermann, 2010).

Das Euler-Verfahren

Aus mathematischer Sicht hat das Euler-Verfahren eher historische Bedeutung, da es in einigen Fällen zu unerwünschten Problemen führt und in der Praxis daher bessere Verfahren eingesetzt werden. Für den Schulunterricht ist die Bedeutung des Euler-Verfahrens aber nach wie vor sehr aktuell. Das liegt vor allem daran, dass es leicht zu erklären und mit einer Schulklasse durchzuführen ist. Sobald der Lösungsalgorithmus mit Programmiersprachen oder der Nutzung von Tabellenkalkulationsprogrammen selbst geschrieben wird, ist die Nutzung des Euler-Verfahrens nahezu alternativlos, da andere Verfahren zu komplex sind.

Das Euler-Verfahren geht davon aus, dass die für die Bewegung relevanten Größen in kleinen Zeitintervallen konstant sind. Die Beschleunigung ergibt sich durch die wirkenden Kräfte und die Masse des Objekts: $\vec{a} = \frac{\vec{F}_{ges}}{m}$

Die Kräfte werden in dem Zeitintervall Δt als konstant angesehen, wodurch die Beschleunigung ebenfalls konstant ist. Daher ergibt sich die Geschwindigkeit mit:

$$\Delta \vec{v} = \vec{a} \cdot \Delta t$$

Also:

$$\vec{v}_{neu} - \vec{v}_{alt} = \vec{a} \cdot \Delta t$$

Und damit:

$$\vec{v}_{neu} = \vec{v}_{alt} + \vec{a} \cdot \Delta t$$

Für den Ort gilt:

$$\Delta \vec{x} = \vec{v} \cdot \Delta t$$

Ein analoges Verfahren liefert dann:

$$\vec{x}_{neu} = \vec{x}_{alt} + \vec{v} \cdot \Delta t$$

In den Berechnungsschritten für $\vec{v}_{neu}$ und $\vec{x}_{neu}$ greift das klassische Euler-Verfahren auf die Werte des vorigen Zeitschritts zurück: $\vec{v}_{neu} = \vec{v}_{alt} + \vec{a}_{alt} \cdot \Delta t$ und $\vec{x}_{neu} = \vec{x}_{alt} + \vec{v}_{alt} \cdot \Delta t$ (Wilhelm & Schecker, 2021). Insbesondere anfällig ist das oben dargestellte

Euler-Verfahren gegenüber Schwingungsprozessen. Hier kommt es trotz einer realistischen Modellierung zu einem unrealistischen Aufschaukeln der Schwingungen im Modell. In einer verbesserten Variante des Euler-Verfahrens wird bei den Berechnungen einmal auf einen Wert aus dem neuen Zeitintervall genutzt (Wilhelm, 2020): Entweder $\vec{v}_{neu} = \vec{v}_{alt} + \vec{a}_{neu} \cdot \Delta t$ und $\vec{x}_{neu} = \vec{x}_{alt} + \vec{v}_{alt} \cdot \Delta t$ oder stattdessen $\vec{v}_{neu} = \vec{v}_{alt} + \vec{a}_{alt} \cdot \Delta t$ und $\vec{x}_{neu} = \vec{x}_{alt} + \vec{v}_{neu} \cdot \Delta t$.

Als Beispiel für die Lösung einer Differentialgleichung mit dem Euler-Verfahren soll ein Fallkegel mit Luftreibung dienen (Wilhelm, 2000). Die Bewegungsgleichung lautet

$$m \cdot a(t) = m \cdot g + c \cdot v(t)^2$$

mit $g = -9,81 \text{ ms}^{-2}$, $v_0 = 0 \text{ ms}^{-1}$ und $x_0 = 0$. Beispielhaft wird $\frac{c}{m} = 2,14 \frac{1}{\text{m}}$ angenommen (diese realistischen Daten stammen aus einem Realexperiment). Die Beschleunigung im ersten Schritt wird berechnet mit

$$a_{alt} = g + \frac{c}{m} v_{alt}^2 = g.$$

Die Geschwindigkeit ergibt sich dann durch

$$v_{neu} = v_{alt} + a_{alt} \cdot \Delta t = g \cdot \Delta t$$

Bei einer Wahl von $\Delta t = 0,1 \text{ s}$ ist $v_{neu} = -0,981 \frac{\text{m}}{\text{s}}$. Die Berechnung vom Ort x erfolgt nach dem verbesserten Euler-Verfahren dann mit der neuen Geschwindigkeit.

$$x_{neu} = x_{alt} + v_{neu} \cdot \Delta t = 0 - 0,981 \frac{\text{m}}{\text{s}} \cdot 0,1 \text{ s} = -0,0981 \text{ m}$$

Im nächsten Zeitschritt ist

$$a_{alt} = g + \frac{c}{m} v_{alt}^2 = g + \frac{c}{m} \left(-0,981 \frac{\text{m}}{\text{s}}\right)^2 = g + 0,962 \frac{\text{m}^2}{\text{s}^2} \cdot \frac{c}{m} = -7,75 \frac{\text{m}}{\text{s}^2}$$

und damit

$$v_{neu} = v_{alt} + a_{alt} \cdot \Delta t = -0,981 \frac{\text{m}}{\text{s}} + \left(-7,75 \frac{\text{m}}{\text{s}^2}\right) \cdot 0,1 \text{ s} = -1,75 \frac{\text{m}}{\text{s}}$$

und

$$x_{neu} = x_{alt} + v_{neu} \cdot \Delta t = -0,0981 \text{ m} + \left(-1,75 \frac{\text{m}}{\text{s}}\right) \cdot 0,1 \text{ s} = -0,273 \text{ m}.$$

Das Runge-Kutta-Verfahren

Das Runge-Kutta-Verfahren ist in den meisten Programmen für die mathematische Modellbildung ebenfalls implementiert und liefert bessere Resultate (Schecker, 1998a). Das Verfahren ist so komplex, dass es in der Schule nicht besprochen werden kann. Die Grundideen sollen hier dennoch kurz vorgestellt werden.

Das klassische Runge-Kutta-Verfahren ist vierstufig und wird oft mit Runge-Kutta 4 (RK4) abgekürzt. Auch beim RK4 werden Differentialquotienten mit abschnittsweise konstanten Differenzenquotienten approximiert. Dabei wird die Funktion in jedem Schritt, also zu jedem Zeitpunkt, vier Mal ausgewertet und so vier Hilfssteigungen berechnet, die mit unterschiedlichen Gewichtungen in die Lösung mit einfließen. Der gewichtete Durchschnitt dieser vier Hilfssteigungen führt zu der berechneten Steigung. Als Beispiel kann wieder die Lösung der Differentialgleichung des Fallkegels mit Luftreibung beschrieben werden:

$$m \cdot v'(t) = m \cdot g + c \cdot v(t)^2$$

Um nun die Geschwindigkeit im ersten Berechnungszeitpunkt zu erhalten, müssen vier Hilfsfunktionen berechnet werden:

$$k_1 = g + \frac{c}{m} \cdot v_0 = g$$

$$k_2 = g + \frac{c}{m} \cdot \left(v_0 + \frac{dt}{2} \cdot k_1 \right)^2 = g + \frac{c}{m} \cdot \left(\frac{dt}{2} \cdot g \right)^2$$

$$k_3 = g + \frac{c}{m} \cdot \left(v_0 + \frac{dt}{2} \cdot k_2 \right)^2 = g + \frac{c}{m} \cdot \left(\frac{dt}{2} \cdot \left(g + \frac{c}{m} \cdot \left(\frac{dt}{2} \cdot g \right)^2 \right) \right)^2$$

$$\begin{aligned} k_4 &= g + \frac{c}{m} \cdot (v_0 + dt \cdot k_3)^2 \\ &= g + \frac{c}{m} \cdot \left(dt \cdot \left(g + \frac{c}{m} \cdot \left(\frac{dt}{2} \cdot \left(g + \frac{c}{m} \cdot \left(\frac{dt}{2} \cdot g \right)^2 \right) \right)^2 \right) \right)^2 \end{aligned}$$

Mit den Daten von oben erhält man als Näherungslösung dann:

$$\begin{aligned}v(t_1) = v_1 &\approx v_0 + dt \cdot \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \\&= 0,1 \cdot \frac{1}{6}(-9,81 + 2 \cdot (-9,30) + 2 \cdot (-9,35) + (-7,94)) \frac{m}{s} \\&= -0.9175 \frac{m}{s}\end{aligned}$$

Die Auswertung der Differentialgleichung an den folgenden Zeitschritten wird dabei noch komplexer und somit weder per Hand durchführbar noch im Unterricht erklärbar.

Das Verfahren hat die Konsistenzordnung 4 und ist dem Euler-Verfahren mit Konsistenzordnung 1 überlegen (Hermann, 2010). Die Konsistenzordnung gibt Auskunft über den möglichen lokalen Diskretisierungsfehler (Stoer & Bulirsch, 2005) und ist neben Kondition und Stabilität eine der drei Fehlerbewertungen für numerische Verfahren.

Wenn also durch die genutzte Software die Möglichkeit zur Nutzung des RK4 oder eines Runge-Kutte-Verfahrens höherer Ordnung besteht, sollte dies genutzt werden, um von der Realität abweichende Resultate zu verhindern. Sobald das numerische Verfahren explizit thematisiert und verstanden werden soll, muss allerdings die verbesserte Variante des Euler-Verfahren verwendet werden.

5.4.3 Arten der Mathematischen Modellbildung⁹

5.4.3.1 Programmiersprachen

Programmiersprachen können genutzt werden, um mathematische Modellbildung in der Schule zu nutzen. Die Idee bei der Nutzung von Programmiersprachen für die Modellierung der Realität ist, dass man dem Computer durch den Quellcode mitteilt, was in welcher Reihenfolge berechnet werden soll. Das heißt, dass die Nutzer*innen Kenntnisse über die physikalischen Zusammenhänge, über die nötige Mathematik und über die Syntax und Logik der Programmiersprache haben müssen. In der Regel wird bei der Benutzung von Programmiersprachen das Eulersche Polygonzugverfahren (Euler-Verfahren) verwendet, welches einfach zu erklären ist (siehe Kapitel 5.4.2).

⁹ Teile dieses Kapitels sind an Weber und Wilhelm (2020a) angelehnt.

Für die Implementierung des Euler-Verfahrens benötigt man eine Schleife, in der die folgenden Zeilen enthalten sind:

$$v = v + \frac{F_{ges}}{m} \cdot \Delta t$$

$$y = y + v \cdot \Delta t$$

$$t = t + \Delta t$$

Zudem muss die Gesamtkraft aus den angreifenden Kräften berechnet werden. Möglich ist es außerdem, die Berechnung der Beschleunigung in einer eigenen Zeile vorzunehmen und diesen Wert für die Berechnung der neuen Geschwindigkeit zu nehmen. Ein Beispiel für eine Schleife für einen Fallkegel mit Luftreibung in der Programmiersprache „Python“ ist in Abb. 1 dargestellt.

```
3 while y>0:
4     F_Gewicht=m*g
5     F_Reibung=k*v**2
6     F_Gesamt=F_Gewicht+F_Reibung
7     a=F_Gesamt/m
8     v=v + a*dt
9     y=y + v*dt
10    f1.plot(pos=(t,y))
11    t=t+dt
```

Abb. 1: Schleife unter Nutzung des verbesserten Euler-Verfahrens in Python.

Durch die Programmierung der nötigen Schleifen wird der iterative Charakter von numerischen Berechnungen deutlich. Die Zusammenhänge zwischen den kinematischen Größen können so ebenfalls vertieft und gefestigt werden. Außerhalb der Schleife müssen die Anfangsbedingungen und Konstanten definiert werden. Wenn außerdem ein Diagramm gezeichnet werden soll, muss definiert werden, welche Größen gegeneinander aufgetragen werden sollen (Abb. 2).

Die Anfänge der mathematischen Modellbildung mit Programmiersprachen liegt bereits in den 1970er Jahren (Bork, 1980; Peters, 1980), wobei meist die Programmiersprachen „Fortran“ und „Basic“ verwendet wurden. Vereinzelt wurde auch „Grafit“ eingesetzt (Kane & Sherwood, 1980). Der erste größere Einsatz von Programmiersprachen zur mathematischen Modellbildung wurde von dem „Maryland University Project in Physics and Educational Technology“ (M.U.P.P.E.T.) umgesetzt. Dabei wurde die Programmiersprache „Pascal“ eingesetzt, um die sogenannten „power tools“ der Physik zu vermitteln (MacDonald et al., 1988).

```

3 #Diagramm
4 f1=gcurve(color=color.red)
5
6 #Konstanten
7 dt=0.001
8 m=0.00103
9 g=-9.81
10 k=0.0022
11
12 #Anfangsbedingungen
13 t=0
14 y=1.045
15 v=0
16
17 #Schleife
18 while y>0:
19     F_Gewicht=m*g
20     F_Reibung=k*v**2
21     F_Gesamt=F_Gewicht+F_Reibung
22     a=F_Gesamt/m
23     v=v + a*dt
24     y=y + v*dt
25     f1.plot(pos=(t,y))
26     t=t+dt

```

Abb. 2: Modellierung mit Erstellung eines Diagramms in Python.

In Schulen wurde als Erstes der Commodore 64 eingesetzt, um mit imperativen Programmiersprachen, wie „Basic“, „Comal“ und „Pascal“ zu modellieren. Später wurden andere Programme, wie „Dynamos“ für MS-DOS entwickelt und eingesetzt. Ein für die Schule ebenfalls denkbare Vorgehen ist die Nutzung von programmierbaren Taschenrechnern, mit denen die Differentialgleichungen von physikalischen Systemen gelöst werden können (Eisberg & Eisberg, 1978).

Als Ausgabe der fertigen Berechnung können alle Größen in verschiedenen Diagrammen dargestellt werden. Dies war über eine lange Zeit die einzige Darstellungsform, was für Schüler*innen problematisch sein kann, da sie sich mit dem Interpretieren von Graphen erfahrungsgemäß schwertun (Beichner, 1994; McDermott et al., 1987; Thornton & Sokoloff, 1998). Mittlerweile ist es mit der Programmierumgebung „VPython“ auch möglich, mit ein paar zusätzlichen Zeilen Code, 3D-Animationen der modellierten Bewegung zu erstellen, was beim Bewerten des Modells helfen kann (Scherer et al., 2000). Dazu sind zusätzliche Zeilen im Code und Kenntnisse des Erstellens der 3D-Objekte nötig (Abb. 3). Dieser Code erzeugt eine 3D-Animation mit Orts-Zeit-Diagramm, welches in Abb. 4 zu sehen ist.


```
3 #3D-Objekte und Geschwindigkeitspfeil
4 display(width=638,height=500,center=vector(0,0.5,0),
5         background=color.white)
6 cone = cone(axis=vector(0,-0.2,0), radius=0.2, color=color.yellow)
7 vcone = arrow(pos=cone.pos, axis=vector(0,0,0), shaftwidth=0.015,
8              color=color.red)
9 floor = box(pos=vector(0,-0.2,0), length=2, height=0.01, width=1,
10            color=color.gray(0.9))
11
12 #Diagramm
13 f1 = gcurve(color=color.red)
14
15 #Konstanten
16 dt=0.001
17 m=0.00103
18 g=-9.81
19 k=0.0022
20
21 #Anfangsbedingungen
22 t=0
23 cone.pos=vector(0,1.045,0)
24 cone.velocity=vector(0,0,0)
25
26 #Schleife
27 while cone.pos.y>0:
28     rate(100)
29     F_Gewicht=m*g
30     F_Reibung=k*cone.velocity.y**2
31     F_Gesamt=F_Gewicht+F_Reibung
32     cone.acceleration=vector(0,F_Gesamt/m,0)
33     cone.velocity.y=cone.velocity.y + cone.acceleration.y*dt
34     cone.pos=cone.pos + cone.velocity*dt
35     vcone.pos=cone.pos
36     vcone.axis=cone.velocity/5
37     f1.plot(pos=(t,cone.pos.y))
38     t=t+dt
```

Abb. 3: Modellierung in VPython mit 3D-Animation.

So kann, wenn entsprechende Programmierkenntnisse vorliegen oder dies von der Lehrkraft gut vorbereitet ist, eine Animation erstellt werden, die synchron zu dem Graphen darunter abläuft. Heute werden Programmiersprachen zum Modellieren im Physikunterricht nur noch selten eingesetzt, wohl auch aufgrund der beschriebenen Hürden, die mit dem Erlernen von Programmiersprachen einhergehen.

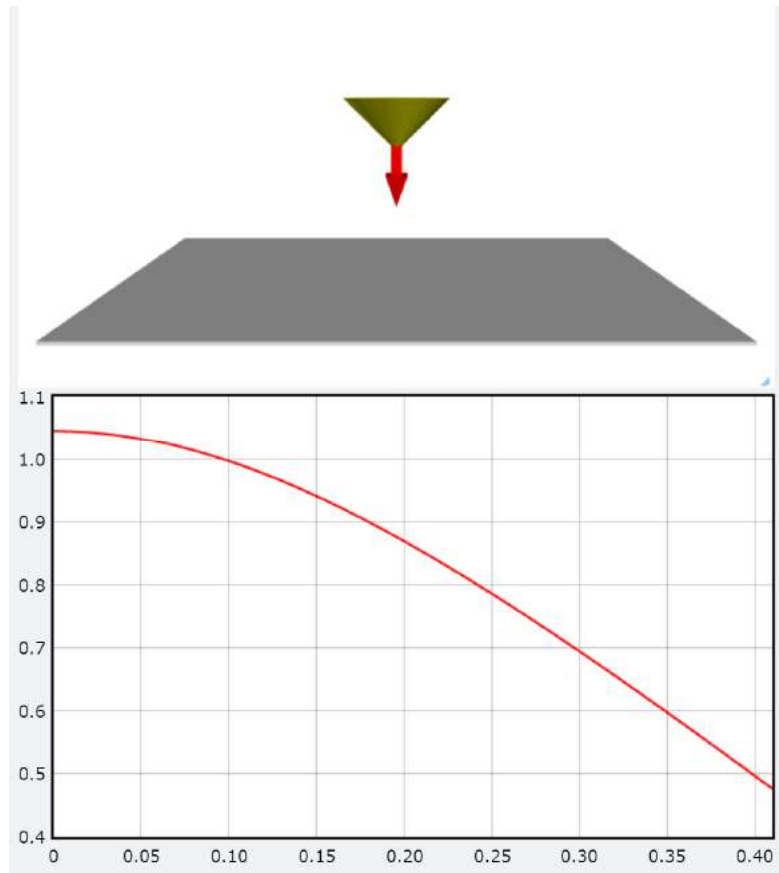


Abb. 4: Ausgabe vom VPython-Code: Animation mit Geschwindigkeitspfeil oben, Orts-Zeit-Diagramm unten.

5.4.3.2 Tabellenkalkulationsprogramme

Eine andere Art mathematische Modellbildung im Physikunterricht einzusetzen, ist die Nutzung von Tabellenkalkulationsprogrammen. Die Grundideen bei der Berechnung der physikalischen Größen aus anderen Größen sind identisch zur Nutzung von Programmiersprachen. Ein Hauptunterschied und Vorteil der Tabellenkalkulationsprogramme gegenüber Programmiersprachen ist, dass die Nutzer*innen keine Programmierkenntnisse benötigen (Hsiao, 1985). Damit soll bezweckt werden, dass der Fokus auf dem Lehren von Physik mit dem Computer liegen kann und nicht Computerkenntnisse mit Hilfe von Physik beigebracht werden (Laws, 1991a). Dafür sind hier aber gute Kenntnisse der verwendeten Software nötig, um erfolgreich zu modellieren. Bei der Berechnung des Modells entsteht eine Tabelle voller Zahlenwerte, wodurch es schwierig ist, Fehler zu finden und die modellierten Gleichungen zu modifizieren (Wilhelm, 2018b).

In der Regel wird so gearbeitet, dass jede Zeile in der Tabelle einem Zeitschritt entspricht und in jede Spalte eine physikalische Größe eingegeben wird. Über Zellenbezüge können

die Nutzer*innen dann die Gleichungen des Euler-Verfahrens implementieren, wodurch sich die Zellen automatisch mit den berechneten Werten füllen. Je nachdem wie lang die Modellierung sein soll, werden neue Zeilen erzeugt, bis man am Ende der Modellierung angelangt ist (Abb. 5).

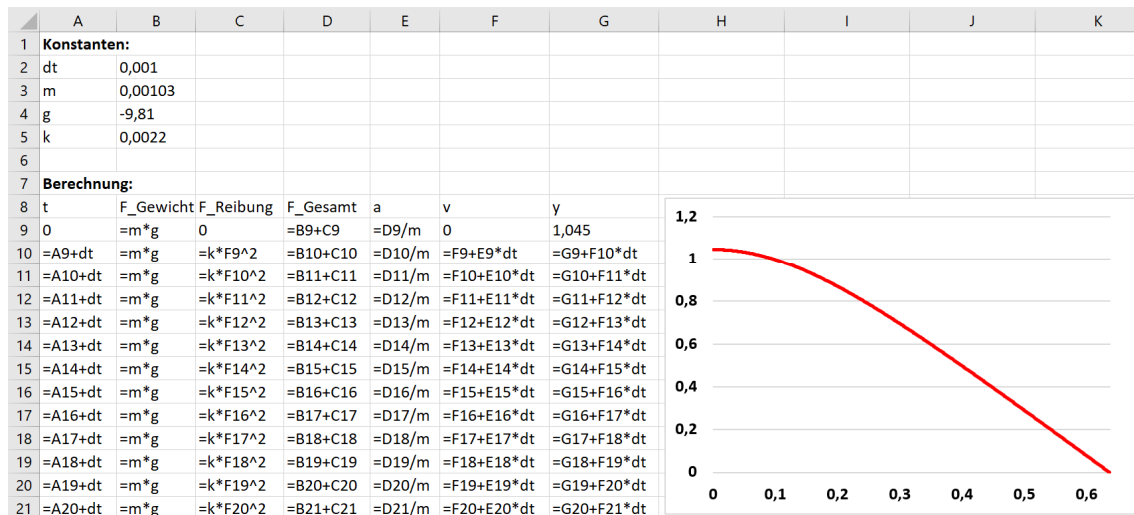


Abb. 5: Modellierung mit Excel.

Tabellenkalkulationsprogramme zum Modellieren werden im Physikunterricht bereits ähnlich lang eingesetzt wie imperative Programmiersprachen. 1979 wurde VisiCalc für den Apple II entwickelt (Baker & Sugden, 2003), was dann bereits 1984 für den Unterricht genutzt wurde (Arganbright, 1984). Auch aktuell wird Excel häufig im Physikunterricht eingesetzt, da das Programm den meisten Lehrkräften geläufig ist und für verschiedene Dinge genutzt werden kann (Herman, 2009). Die Ausgabe bleibt bei Tabellenkalkulationsprogrammen auf Diagramme beschränkt.

5.4.3.3 Graphische Modellbildungsprogramme

Ein wichtiges Ziel der mathematischen Modellbildung im Physikunterricht ist die Verdeutlichung von physikalischen Zusammenhängen. Diese Zusammenhänge werden vom Nutzer bzw. der Nutzerin selbst modelliert und verdeutlichen die unveränderliche Grundstruktur der sogenannten Newton-Maschine.

Um diese Struktur im Vergleich zu der Nutzung von Programmiersprachen und Tabellenkalkulationsprogrammen stärker zu verdeutlichen und vor allem visuell darzustellen, wurden graphische Modellbildungsprogramme speziell für den Unterricht entwickelt. Diese Programme ermöglichen es den Nutzer*innen, auf einer graphischen Oberfläche zunächst qualitativ die Struktur zu erstellen (Abb. 6), wonach die Zusammenhänge erst

im zweiten Schritt quantifiziert werden. Damit kann die Bewegung dann berechnet werden.

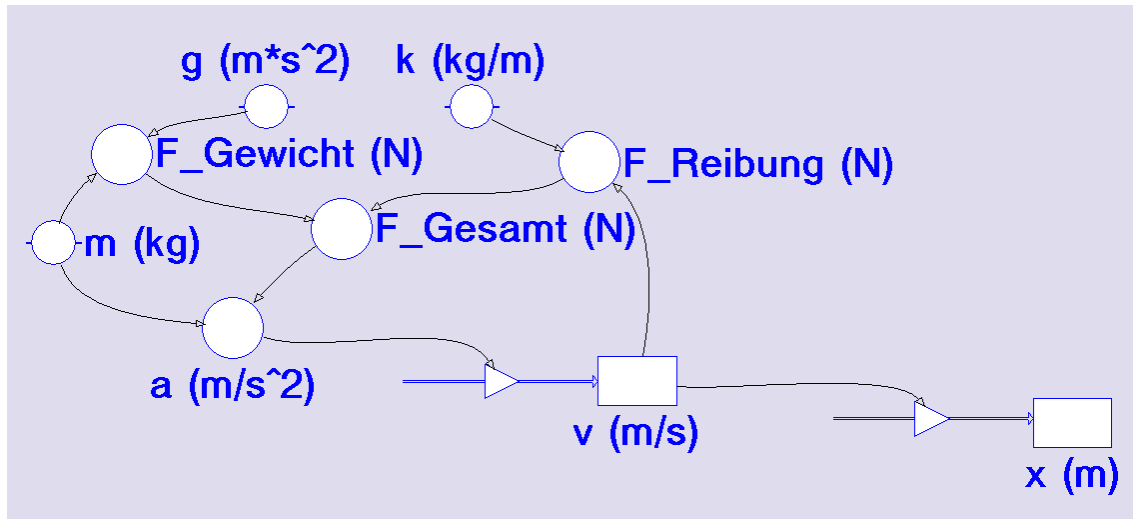


Abb. 6: Modellierung mit graphischer Oberfläche in Coach 6.

Dabei gibt es verschiedene Schaltflächen, die auf die Logik der Systemdynamik nach Forrester (1968) zurückgehen und ursprünglich für betriebswirtschaftliche Systeme entwickelt wurden (Bossel, 1992; Vgl. Wilhelm, 2005a):

- (1) Vorgabegrößen: Größen, die von außen auf das System einwirken, aber nicht von diesem beeinflusst werden (hier z. B. g).
- (2) Zustandsgrößen: Größen, in denen sich der Zustand eines Systems als Ausdruck seiner Geschichte darstellt (hier z. B. v).
- (3) Zwischengrößen: Größen, die aus Vorgabe- und Zustandsgrößen berechnet werden (hier z. B. F_{Gesamt}).

Graphische Modellbildungsprogramme können die auftretenden Differentialgleichungen mit verschiedenen numerischen Verfahren lösen. Die dabei nötigen Rechenschritte rücken aber in den Hintergrund und der Nutzer wird nicht nur von der Berechnung entlastet, sondern muss sich auch über die Berechnungsverfahren selbst keine Gedanken mehr machen.

Ein in den 1990er-Jahren häufig eingesetztes graphisches Modellbildungsprogramm ist „STELLA“. Es wurde ursprünglich für die Wirtschaftswissenschaften entwickelt. Speziell für den Schulunterricht entwickelte Programme sind „Dynasys“, „Modus“ und „Coach“, die teilweise auch noch im Unterricht eingesetzt werden.

Auch bei graphischen Modellbildungsprogrammen gibt es je nach Software die Möglichkeit, die berechnete Bewegung in einer Animation ausgeben zu lassen, was bei der Bewertung des Modells helfen kann. Dazu wurden in den 1990er und 2000er Jahren spezielle Programme wie „VisEdit/PAKMA“ und „JPAKMA“ entwickelt. Die Software „Modellus“ bzw. „ModellusX“ ist auch in der Lage, Animationen auszugeben (Teodoro, 2004).

5.4.3.4 Gleichungsbasierte Modellbildungsprogramme

Alle bisher vorgestellten Programmtypen haben unterschiedliche Vorteile und verfolgen unterschiedliche Ziele. Gemeinsam ist ihnen aber ein erheblicher Aufwand zur Einarbeitung in die jeweilige Software. Diese Zeit kann im Schulunterricht oft fehlen, sodass sich ein Einsatz der bisher vorgestellten Software nur bei konsequenter und regelmäßiger Nutzung lohnt. Deswegen gibt es mittlerweile auch Programme, die komplexitätsreduziert und damit einfacher zu bedienen sind. Um bei Programmen wie „Newton-II“ ein mathematisches Modell zu erstellen, müssen die Nutzer*innen lediglich die an einem Körper wirkenden Kräfte, dessen Masse und die Anfangsbedingungen der Bewegung in Gleichungsform eingeben und erhalten daraufhin die berechnete Bewegung (Abb. 7). Die kinematischen Zusammenhänge sind bereits von der Software vorgegeben, wodurch die Software nur für den Einsatz in der Mechanik geeignet ist.

Das führt dazu, dass weder die „Methode der kleinen Schritte“ bzw. die Euler-Methode noch die Zusammenhänge zwischen den kinematischen Größen thematisiert werden. Die Konzentration liegt außerdem bei den Zusammenhängen zwischen Kräften und der Bewegung und damit der zentralen Gleichung der Mechanik und nicht bei den Zusammenhängen zwischen den kinematischen Größen. Diese Einschränkung verfolgt das Ziel, dass die Software mit weniger Einarbeitungszeit verbunden ist und daher schneller im Unterricht eingesetzt werden kann.

Eine Bewertung des Modells erfolgt hier einerseits über die entstehenden Diagramme. Andererseits kann diese über den Vergleich mit Messdaten, welche direkt in demselben Diagramm angezeigt werden können, erfolgen.

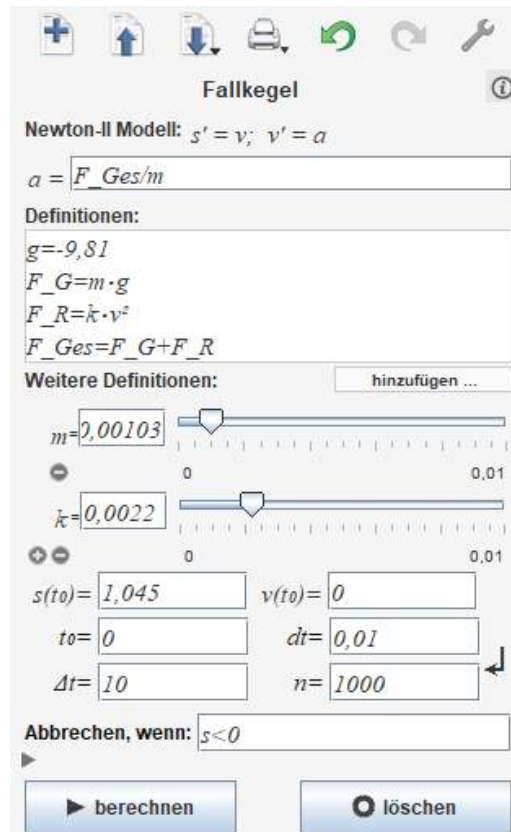


Abb. 7: Gleichungsbasierte Modellierung in „Newton-II“.

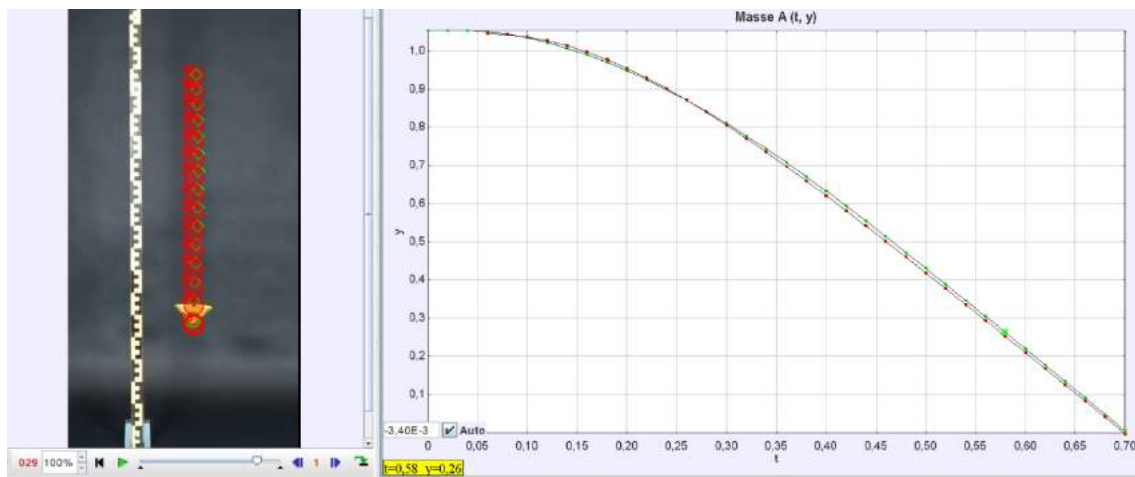


Abb. 8: Vergleich von Modell und Realität in Video und Diagramm in „Tracker“.

Auch möglich ist z. B. bei dem Programm Tracker eine Kombination mit einer Videoanalyse, wodurch das erzeugte Modell direkt im Video angezeigt wird und damit ein Vergleich zwischen Objekt und Modell erfolgen kann (Abb. 8). Dazu im Menüpunkt

„Dynamisches Modell“ neben den Anfangsbedingungen lediglich eine Gleichung für die wirkende Gesamtkraft eingegeben werden.

5.4.4 Gründe für den Einsatz von mathematischer Modellbildung¹⁰

Das computerbasierte Modellieren von Zusammenhängen physikalischer Systeme (*Computational Physics*) ist ein zentrales Element der physikalischen Forschung, wird teilweise als dritte Säule der Physik (neben theoretischer Physik und Experimentalphysik) angeführt (Obsniuk et al., 2015) und als genauso wichtig wie ebenjene erachtet. Moderne Wissenschaft wäre ohne den Einsatz von Computern zur Modellierung der Realität nicht möglich. Das liegt nicht zuletzt daran, dass die Modelle zur Beschreibung der Realität immer mächtiger und damit auch komplexer werden. Die Beschreibung der Realität in der Schulphysik ist mathematisch ebenfalls so komplex, dass sich der klassische Physikunterricht meist vor allem in idealen Laborsituationen bewegt. Dies stellt ein Problem dar, da Lernende Schwierigkeiten haben, sich in Idealsituationen hineinzudenken (Schecker, 1985). Die Reduktion führt auch dazu, dass sich die Anzahl an möglichen Experimenten stark reduziert und Reibungskräfte in der Schule oft keine Rolle spielen. Das kann einerseits dazu führen, dass Schüler*innen der Meinung sind, Physik habe nichts mit der Realität zu tun (Schecker, 1985). Andererseits kann gerade das Fehlen von Reibungskräften im Unterricht zu einer aristotelischen Sichtweise auf die Mechanik führen, da im Alltag alle Bewegungen zum Stillstand kommen. Bei fehlender Betrachtung und Betonung von Reibungskräften liegt der Schluss nahe, dass dies eine inhärente Eigenschaft von sich bewegenden Körpern ist.

Ebenso sollte im Unterricht generell ein Fokus auf Alltagserfahrungen der Schüler*innen gelegt werden. Speziell in der Mechanik besteht für Schüler*innen die Aufgabe der Physik in der Beschreibung realer Abläufe (Schecker et al., 2018). Die Lernenden kommen mit einer Fülle an Alltagsvorstellungen über die Mechanik in den Unterricht. Um reale Bewegungsabläufe in der Schule thematisieren zu können, bedarf es aber Hilfsmittel, da die Mathematik zur Beschreibung dieser Bewegungen für den Schulunterricht zu schwierig ist. Das liegt daran, dass in der Regel mehrere Kräfte gleichzeitig wirken und dass Reibungskräfte auf den Körper wirken, welche häufig abhängig von der Geschwindigkeit des Körpers sind. Als geschwindigkeitsabhängige Kraft tritt meist die Luftreibung auf, die bei Fall- und Fahr-, aber auch Flugbewegungen eine große Rolle spielt. Nicht vom Tempo, aber von der Richtung der Geschwindigkeit, abhängig ist die Coulomb-Reibung.

¹⁰ Teile dieses Kapitels sind an Weber und Wilhelm (2020a) angelehnt.

Schwingungen entstehen durch ortsabhängige Kräfte, während auch Beispiele für zeitabhängige Kräfte gefunden werden können. Diese Tatsache führt zwangsläufig zu sehr schwierig oder gar nicht zu lösenden Differentialgleichungen. Das ist ein Hauptgrund dafür, dass in der Schule nahezu ausschließlich idealisierte Phänomene thematisiert werden. Abhilfe kann hier mathematische Modellbildungssoftware schaffen, bei der die Nutzer*innen die physikalischen Zusammenhänge modellieren und quantifizieren, woraufhin die Software die entstehenden Differentialgleichungen numerisch löst. Die modellierten Daten können dann auf verschiedene Weise dargestellt werden, sodass das erstellte Modell optimiert und daraufhin angewendet werden kann, um Vorhersagen zu treffen.

Die Nutzung von mathematischer Modellbildungssoftware führt zu einer Betonung der Zusammenhänge physikalischer Größen, was eine grundsätzliche Aufgabe von Physikunterricht ist. Gerade in der Mechanik bietet es sich an, diese Kette von Abhängigkeiten zu thematisieren, welche grundlegend für das Verständnis der Newtonschen Mechanik ist. Die „Resultierende“ oder „Summe aller Kräfte“ bewirkt, dass ein Körper beschleunigt wird, während die Masse des Körpers diese Beschleunigung mitbestimmt. Die Beschleunigung, die Geschwindigkeit und der Ort sind über die sogenannte kinematische Kette verbunden. Die Beschleunigung ist die Geschwindigkeitsänderung pro Zeit, während die Geschwindigkeit die Ortsänderung pro Zeit ist. Diese auch „Newton-Maschine“ oder Standardmodell (Schecker, 1993a) genannte Kette von Abhängigkeiten ist für jede Bewegung identisch und kann damit als Grundstruktur der Newtonschen Mechanik bezeichnet werden. Dass die Grundstruktur für alle Bewegungen gleichbleibt, wird durch eigene Modellierungen deutlich. Das steht im Kontrast dazu, dass i. d. R. im Unterricht ein großer Fokus auf spezielle Gleichungen wie $x(t) = \frac{1}{2}a \cdot t^2$ gelegt wird, die zwar mit einfach mathematischen Methoden gelöst werden können, aber nur unter sehr speziellen Bedingungen gelten (hier: $x(0) = 0$, $v(0) = 0$, $a = \text{konst.}$). Modellbildungsprogramme arbeiten hingegen mit den allgemein gültigen Definitionen wie $\vec{v} = \frac{\Delta \vec{x}}{\Delta t}$ und der Grundgleichung der Newtonschen Mechanik $\vec{a} = \frac{\Sigma \vec{F}}{m}$. Dabei ist die Idee, eine Vielzahl an Phänomenen mit einer geringen Anzahl an Gesetzen und Definitionen zu erklären (Bethge, 1992). Diese Gesetze und Definition werden auch *Power Tools* genannt (MacDonald et al., 1988).

Eine Modellbildungssoftware gibt zudem entscheidendes Feedback über die Hypothesen der Lernenden. Das Modellieren entspricht dem Explizieren des Gedankenmodells, welches ein Lernender von einer bestimmten Situation hat (Wilhelm, 2005a). Ein Explizieren dieses Gedankenmodells führt zu einer aktiven Auseinandersetzung mit den Vorstellungen, die über eine bestimmte Bewegung vorherrschen. Die Modellierung selbst ist eine

aktive Formulierung von Hypothesen über diese Bewegung. Die Software gibt den Nutzer*innen unmittelbares Feedback über diese Hypothesen, sodass sie bei Bedarf korrigiert und verbessert werden können (Lück & Wilhelm, 2011). Es bietet sich also an, den Grad der Idealisierung des Modells immer weiter zu reduzieren und im Laufe des Prozesses der Modellierung weitere Einflüsse in das Modell aufzunehmen. Das Modell wird während des Modellierungsprozesses fortlaufend evaluiert. Die Lernenden entscheiden zudem selbstständig, welche Einflüsse in das Modell aufgenommen und welche ignoriert werden.

Daher lassen sich mit mathematischer Modellbildung ebenso Aspekte von Nature of Science im Unterricht unterbringen und diskutieren (Büffler et al., 2008). Es birgt für Lernende die Möglichkeit, die Fähigkeiten des Schätzens und Approximierens zu erlernen (Redish & Wilson, 1993). Modelle werden in einem dynamischen iterativen Prozess erstellt (Mikelskis-Seifert & Kasper, 2011) und der wissenschaftliche Erkenntnisprozess und das Zusammenspiel von Theorie, Experiment und Computer können beleuchtet werden (Teodoro & Neves, 2011). Mit mathematischer Modellbildung lässt sich authentischer, schülerzentrierter und projektbasierter Unterricht realisieren, der sowohl inhaltlich relevante Aspekte verdeutlicht als auch wissenschaftliche Arbeitsweisen vermitteln kann.

5.4.5 Forschungsergebnisse¹¹

Die Gründe, warum mathematische Modellbildung im Unterricht eingesetzt wird, sind vielfältig. Dementsprechend ist die Forschung, um herauszufinden, ob der Einsatz im Unterricht erfolgreich ist, ebenso vielfältig. Neben dem Konzeptverständnis der Newtonschen Mechanik wurden auch allgemeinere Fähigkeiten, wie systemisches Denken, und Aspekte der Natur der Naturwissenschaften erforscht.

Bereits in den 1980er Jahren gab es zwei große Umstrukturierungen der Physik für Erstsemester an US-amerikanischen Universitäten, die mathematische Modellbildung beinhalteten. Das M.U.P.P.E.T.-Projekt nutzte Computer, um reale physikalische Systeme zu analysieren (Redish & Wilson, 1993). Die Studierenden mussten dazu in Pascal programmieren. Erforscht wurde, ob diese Umstrukturierungen es ermöglichen, dass wissenschaftliche Fähigkeiten, die für das Durchführen von wissenschaftlichen Projekten notwendig sind, früher im Studium erlernt werden können. Das Ergebnis war ein signifikanter Anstieg an guten Forschungsprojekten, die von den Studierenden durchgeführt wurden. Einen anderen Ansatz wählten die Begründer des Workshop Physics (Laws et al.,

¹¹ Teile dieses Kapitels sind an Weber und Wilhelm (2020a) angelehnt

2004), bei dem allerdings keine Programmiersprache integriert wurde. Diese Entscheidung basierte darauf, dass der Workshop Physics sich eher auf die Physik als auf die Berechnungen konzentrieren sollte. Der Computer wurde hier in vielfältiger Weise genutzt. Unter anderem wurde mathematische Modellbildung mit Tabellenkalkulationsprogrammen durchgeführt. Der Workshop Physics wurde in verschiedener Weise evaluiert. Dabei hat sich unter anderem gezeigt, dass das Konzeptverständnis von Mechanik bei den Teilnehmenden höher war als nach dem traditionellen Vorgehen (Laws et al., 2015) und dass das Interesse der Studierenden im Workshop Physics ebenfalls größer war, während sich die Leistung beim Lösen von herkömmlichen Textbuch-Aufgaben im Vergleich zum klassischen Unterricht nicht verschlechterte (Laws, 1991b). Da nicht speziell der Einsatz der mathematischen Modellbildung erforscht wurde und viele Änderungen im Vergleich zum herkömmlichen Vorgehen gemacht wurden, ist unklar, welcher Anteil der beobachteten Verbesserungen auf die mathematische Modellbildung zurückzuführen ist.

Die Studien, welche in Deutschland in den 1990er-Jahren durchgeführt wurden, fokussierten eher allgemeinere Fähigkeiten wie etwa systemisches Denken (Hassel, 1987; Ossimitz, 1996; Schecker et al., 1999). Folglich wurden die Inhalte so ausgewählt, dass ein Verständnis des Systemverhaltens gefördert werden sollte (Zumann & Weaver, 1988). Laborstudien konnten dabei zum Teil zeigen, dass die Software zum Lernen von systemischem Denken geeignet war. Feldstudien in verschiedenen fachlichen Kontexten (Klieme & Maichle, 1994) konnten diese Effekte aber nicht nachweisen (Schecker et al., 1999). Castro Kurtz dos Santos (1992) konnte mit dem Einsatz der Programme „IQON“ und „STELLA“ zeigen, dass die Studierenden ein verbessertes Verständnis von Kausalität in der Mechanik hatten. Sie konnten sich die Welt außerdem besser in Form von Variablen vorstellen und verbesserten ihr Verständnis von Differentialgleichungen in iterativen Schleifen. Dies galt allerdings nur für Studierende mit einem Physik-Hintergrund (Castro Kurtz dos Santos, 1992).

Seither haben die meisten Studien zu mathematischer Modellbildung das Konzeptverständnis der zugrundeliegenden Physik im Fokus. Guzdial (1994) kreierte bspw. eine computerbasierte Herangehensweise an das Modellieren von Physik in einer Programmierungsumgebung auf dem Apple Macintosh (Guzdial, 1994). Mit Interviews konnte er zeigen, dass alle seiner fünf Teilnehmenden sich in der Konzeptualisierung von Geschwindigkeit, Beschleunigung und der Wurfbewegung verbesserten. Durch einen fehlenden Vergleich ist aber auch hier unklar, welche Veränderung dabei auf die Variable der Modellbildung zurückzuführen ist.

Schecker et al. führten in verschiedenen Physikkursen der elften bis 13. Jahrgangsstufe in Bremen Testungen durch (Schecker, 1992, 1998a; Schecker & Niedderer, 1991). Ein

großes Projekt „Physiklernen mit Modellbildungssystemen“ wurde von der Deutschen Forschungsgemeinschaft finanziert und von 1996 bis 1999 durchgeführt (Schecker, 1998b; Schecker et al., 1999). In zwei Leistungskursen der elften Klassenstufe wurde „STELLA“ in Kleingruppen eingesetzt, während die Kontrollgruppe keinen Computereinsatz hatte. Der Computereinsatz in der Treatment-Gruppe betrug dabei etwa ein Fünftel der gesamten Unterrichtszeit. Von der Studie konnte nicht bestätigt werden, dass die Gruppe mit dem Computereinsatz ein höheres Level an Konzeptverständnis erreichte (Schecker et al., 1999; Schecker & Gerdes, 1999a). Weiterhin wurde untersucht, ob sich durch den Einsatz von Modellbildungssoftware in der Mechanik eine Verbesserung der Problemlösestrategie in neuen Situationen zeigt, die selbst keine Modellbildung erfordern. Es konnte bestätigt werden, dass Proband*innen aus der Treatment-Gruppe signifikant häufiger die physikalisch korrekte Argumentation bezogen auf die mechanischen Kräfte anwendeten (Schecker et al., 1999; Schecker & Gerdes, 1998). Bei nicht-mechanischen Kräften konnten keine Unterschiede zwischen den beiden Gruppen festgestellt werden (Schecker et al., 1999; Schecker & Gerdes, 1998). Bekannte Strukturen der Mechanik, wie die „Newton-Maschine“, wurden von der Gruppe der mathematischen Modellbildung häufiger für das Modellieren neuer Situationen angewandt (Schecker et al., 1999). Insgesamt hat sich kein Unterschied im systemischen Denken zwischen den beiden Gruppen ergeben. Als Fazit wird in der Studie berichtet, dass die Entwicklung von Newtonschen Ideen durch mathematische Modellbildung nicht in dem erhofften Maß verbessert werden konnte. Mathematische Modellbildung eignet sich demnach lediglich für eine Verbesserung des physikalischen Verständnisses, nicht jedoch für die Entwicklung übergreifender Kompetenzen (Schecker et al., 1999).

Darüber hinaus wurde „STELLA“ mit angehenden Lehrkräften im Physik-Praktikum getestet (Sander, 2000). Dabei stand vor allem die Wechselwirkung zwischen Modell und Experiment im Mittelpunkt der Erhebung. Sander (2000) fand dabei heraus, dass durch die Nutzung der mathematischen Modellbildung mehr über die dahinterliegende Theorie gesprochen wurde als beim reinen Experimentieren. Neues Wissen konnte durch die Nutzung nicht erarbeitet werden (Sander et al., 2002). Der Einsatz von mathematischer Modellbildung eignete sich aber zum Vertiefen des bereits bekannten Wissens. Die Interaktion zwischen Modellierung und Experiment konnte kaum beobachtet werden, was der Autor aber damit erklärt, dass der Vergleich von Modell und Messdaten auf zwei unterschiedlichen Bildschirmen stattfinden musste (Sander, 2000; Tinker, 1993). Huckle und Fischer (1999) konnten zeigen, dass sechs Studierende nach dem Einsatz von „STELLA“ beim Modellieren mehr auf die physikalischen Beziehungen zwischen den Größen geachtet haben. Deren erstellte Concept Maps zeigten aber kaum Änderungen (Huckle & Fischer, 1999).

Tinker (1993) fand außerdem heraus, dass „STELLA“ selbst problematisch für Schüler*innen sein kann. Er setzte das Programm im Mathematikunterricht ein und beobachtete, dass die Lernenden Probleme mit den Begriffen Zustand und Änderungsrate und der Interpretation der entstehenden Graphen hatten. Er folgerte, dass „STELLA“ nicht optimal geeignet ist, um diese Dinge zu lernen.

In einer Studie mit Elftklässler*innen in bayerischen Gymnasien untersuchte Wilhelm den Einsatz der Modellbildungssoftware „VisEdit“ in Kombination mit „PAKMA“ (Wilhelm, 2005a). Dabei wurde die Software in lehrerzentriertem Unterricht eingesetzt. Es konnte gezeigt werden, dass Schüler*innen zu fixiert auf spezielle Lösungsfunktionen waren und die grundlegenden Zusammenhänge in der Mechanik nicht kennen. Ein Umgang mit mehreren Kräften war den Lernenden oft nicht möglich. Der Einsatz der Software wurde von Schülerseite sehr gelobt und die Lerngruppe hatte nach dem Unterricht eher das Gefühl, dass Physik etwas mit dem Alltag zu tun hätte. Untersucht wurden außerdem die Veränderung in den Concept Maps, die die Proband*innen vor und nach der Unterrichtseinheit anfertigten. Darin sollten die Zusammenhänge und Beziehungen in der Mechanik dargestellt werden. Es zeigte sich, dass zu Beginn die Geschwindigkeit für viele die zentrale Größe war und häufig von einer Einwirkung von Masse und Beschleunigung auf die Kraft ausgegangen wurde. Nach der Unterrichtseinheit hatte die „Summe aller Kräfte“ eine viel größere Bedeutung und die Beschleunigung stellte die zentrale Größe dar. Auch der Zusammenhang zwischen Kraft, Masse und Beschleunigung wurde häufiger in der gewünschten Richtung angegeben (Wilhelm, 2005a).

Aktuellere Forschungsergebnisse kommen vor allem aus den USA. Der Modeling Instruction-Lehrplan ist ein Reformprogramm für US-amerikanische High-Schools. Die Erstellung und Umsetzung wurde dabei von der National Science Foundation der Vereinigten Staaten unterstützt und setzt auf schülerzentrierten Unterricht mit intensivem Einsatz des Computers als wissenschaftliches Medium. Dabei geht es unter anderem darum, die Grenzen von physikalischen Modellen wissenschaftlich zu testen und über Erkenntnisgewinnung in den Naturwissenschaften zu sprechen (Jackson et al., 2008). Über den Erfolg des Modeling Instruction-Lehrplans liegen Daten von 20.000 Studierenden vor. Dabei konnte herausgefunden werden, dass der Ansatz zu einem besseren physikalischen Verständnis als traditioneller Unterricht führt (Jackson et al., 2008), die Mitarbeit im Unterricht verstärkt (Brewer et al., 2010), die Ansichten über die Natur der Naturwissenschaften verbessert (Brewer et al., 2009) und eine höhere Selbstwirksamkeitserwartung verursacht (Sawtelle et al., 2010). Mathematische Modellbildung spielte in der ursprünglichen Version jedoch keine Rolle. Die Schülerzentrierung und Fokussierung auf Modelle und den wissenschaftlichen Erkenntnisprozess schaffen aber eine Umgebung, in der

mathematische Modellbildung als Ergänzung eingebaut werden kann. Dies wurde von Caballero et al. (2014) erfolgreich durchgeführt und untersucht. Dabei wurde „VPython“ im Unterricht mit Neuntklässler*innen eingesetzt. Sie fanden heraus, dass Schüler*innen in der Lage sind, „computational thinking“ anzuwenden und numerische Berechnungen durchzuführen. Allerdings wurde auch berichtet, dass die Lernenden eine erhebliche Zeit damit verbrachten, sich in das Programm einzuarbeiten. Auch nach dem ersten Einsatz der Software wurde viel Zeit benötigt, um sich erneut in das Programm einzufinden. Nur ein Drittel der Schüler*innen war nach dem Unterricht in der Lage, ein Modell für ein neues physikalisches System zu erstellen, wobei der Erfolg stark davon abhing, wie gut sie Physik mit „computational knowledge“ verknüpfen konnten (Aiken et al., 2012). Es bleibt also die Frage, ob Programmierumgebungen mit mehr Hilfestellung zu einem besseren Ergebnis führen kann oder ob Programmiersprachen wie Python für den Schulunterricht ungeeignet sind.

Auch in anderen Studien wurde der Einsatz von „VPython“ untersucht, wobei der Erfolg in Universitätskursen höher ist als in der Schule. Mit 1357 Studierenden wurde der Einsatz der „VPython“-Umgebung in der Einführungsveranstaltung in die universitäre Physik untersucht, wobei in elf von 13 Laborversuchen mathematische Modellbildung enthalten war und zusätzlich 13 Probleme als Hausaufgabe mit mathematischer Modellbildung gelöst werden mussten. Anders als in der Schule konnten hier nach dem Kurs nahezu zwei Drittel der Studierenden ein neues Problem erfolgreich modellieren (Caballero et al., 2012). Es konnte außerdem gezeigt werden, dass die größten Probleme der Studierenden damit zusammenhingen, die Gesamtkraft zu berechnen, die auf den Körper wirkt.

In einer anderen Studie an einer Universität wurde „Mathematica“ für die mathematische Modellbildung genutzt. Die Studierenden betonten, dass der Wechsel von Physik auf dem Papier zu Physik am Computer wichtig für das Verständnis der Materie war. Es konnte eine positive Änderung des Selbstbewusstseins und der Motivation beobachtet werden (Caballero & Pollock, 2014). Ausgehend von diesen Resultaten entwickelten die Autor*innen eine aktivitätsorientierte Lernumgebung P³ (Irving et al., 2017), bei der unter anderem mathematische Modellbildung im Zentrum steht und die auf dem Curriculum „Matter and Interactions“ (M and I) basiert (Chabay & Sherwood, 2015). P³ versucht moderne Problemlösestrategien einzuführen und allgemeine Grundprinzipien statt spezielle Lösungsformeln zu fokussieren, die im Unterricht häufig eine viel zu große Rolle spielen (Chabay & Sherwood, 2008). In sieben von 30 realistischen Problemen wird dabei mathematische Modellbildung eingesetzt. Erste empirische Ergebnisse zeigten einen durchschnittlichen Lernzuwachs von 60 % im FMCE-Test (Thornton & Sokoloff, 1998), während die traditionelle Einführung in die Physik einen durchschnittlichen Lernzuwachs

von 10-35 % erreicht. Damit verbunden war auch eine leicht positive Veränderung im Colorado Learning Attitudes about Science Survey (CLASS) (Adams et al., 2006), welcher die Einstellung der Studierenden zur Physik und zur Nature of Science misst. In klassischen Einführungsveranstaltungen geht das durchschnittliche Ergebnis in diesem Test hingegen zurück (Irving et al., 2017). Auch hier war die mathematische Modellbildung aber nur eine von vielen Änderungen und die Ergebnisse sind nicht allein darauf zurückzuführen.

Andere Studien konnten auch Erfolge von verschiedenen Modellbildungsprogrammen nachweisen. „Modellus“ (Neves et al., 2010) wurde erfolgreich eingesetzt und führte zu einer Verbesserung der Interpretation kinematischer Graphen (Araujo et al., 2008). Die Nutzung von „Mathematica“ und „Python“ in einer anderen Studie wurden von Studierenden ebenfalls gelobt und führte zu einer drastischen Verbesserung der damit durchgeführten Projekte (Araujo et al., 2008). Die Nutzung von „VPython“ half Studierenden außerdem nach eigener Einschätzung dabei, drei verschiedene Probleme zu modellieren und die Physik dahinter besser zu verstehen (Hawkins et al., 2017). Benacka (2015b) untersuchte in verschiedenen Studien die motivierenden Eigenschaften der mathematischen Modellbildung mit Tabellenkalkulationsprogrammen in Oberstufenschulen (Benacka, 2015a, 2015b) und kam zu dem Ergebnis, dass 97 % der Schüler*innen die Einheit, in der ein neuartiges Problem mit dem Euler-Verfahren modelliert wurde, sehr interessant fanden (Benacka, 2016).

Viele bisher getätigte Studien untersuchten nur einzelne Aspekte der mathematischen Modellbildung oder änderten im Vergleich zum herkömmlichen Vorgehen mehrere Aspekte, sodass Veränderungen nicht nur auf die mathematische Modellbildung zurückzuführen sind. Es scheint sich aber anzudeuten, dass in Universitäten die Umstrukturierung der Physik unter Einbezug des Computers und mathematischer Modellbildung zu großen Erfolgen führen kann. Diese Ergebnisse sind aber nicht ohne Weiteres auf die Schule übertragbar. In der Schule ergeben sich meist andere Probleme. Der Einsatz von dem graphischen Modellbildungsprogramm „STELLA“ produzierte einige erfreuliche Ergebnisse, konnte aber andere Erwartungen nicht erfüllen, was auch an Schwierigkeiten mit der Software gelegen haben könnte. Programmiersprachen mit Ausgabe von Animationen wie „VPython“ bringen in der Schule auch Probleme mit sich. Zum Einsatz von Tabellenkalkulationsprogrammen gibt es positive Ergebnisse, was die Motivation der Schüler*innen angeht, allerdings wurden dort keine empirischen Daten zum Lernerfolg erhoben. Keine empirischen Ergebnisse gibt es bislang zu gleichungsbasierten Modellbildungsprogrammen wie „Newton-II“. Naheliegend ist es also zu untersuchen, ob

Programme, die mit zentralen Gleichungen arbeiten, in der Schule die Einarbeitungszeit reduzieren und gleichzeitig zu einem erfolgreichen Lernprozess beitragen können.

5.4.6 Heutige Nutzung von mathematischer Modellbildung in Schulen

Von heutiger Software für den Physikunterricht wird neben den Kontiguitätsprinzipien (Kapitel 4.2) erwartet, dass sie intuitiv zu bedienen ist und stabil läuft. Die Nutzung von Software für die mathematische Modellbildung sollte nicht dazu führen, dass man anstelle von mathematischen Schwierigkeiten technische Schwierigkeiten hat. Bei verschiedenen Programmen, die bereits in der Schule eingesetzt wurden, wurde von Schwierigkeiten in der Bedienung berichtet. Das kann auch eine Erklärung dafür sein, dass fertige Animationen häufiger im Unterricht eingesetzt werden als die mathematische Modellbildung (Wenzel, 2018). Ziel sollte es sein, dass eine Software so einfach zu bedienen ist, dass sie ohne lange Einarbeitungszeit von Schüler*innen selbstständig genutzt werden kann. In der Theorie verspricht dies moderne Software zur mathematischen Modellbildung.

Aktuelle Zahlen zur Computernutzung an deutschen Schulen stammen von einer Erhebung, die von 2014 bis 2015 in drei deutschen Regionen stattgefunden hat. Es wurden Daten von insgesamt $N = 163$ Lehrkräften erhoben, die sich auf die drei Regionen verteilten (Wenzel, 2018). Dabei ergaben sich Unterschiede in der Häufigkeit der Nutzung zwischen den Regionen, die mit unterschiedlichen Lehrplänen und auch der Herkunft der verschiedenen Software zu erklären sind. Insgesamt werden insbesondere neuere Programme sehr wenig eingesetzt, da sie wenig bekannt sind. Am häufigsten wurde angegeben, dass mathematische Modellbildung mit Microsoft Excel gemacht würde. Dabei ist fraglich, ob das Programm tatsächlich zum Modellieren oder eher für die Datenauswertung, -aufbereitung und -darstellung genutzt wird.

5.4.7 Zusammenfassung

Mathematische Modellbildung bezeichnet die Konstruktion von Zusammenhängen von physikalischen Größen. Sie unterscheidet sich dabei grundlegend von der Nutzung fertiger Simulationen, in denen nur einzelne Parameter verändert werden können. Es gibt verschiedene Arten von Modellbildungssoftware für den Physikunterricht, die je nach Zielsetzung eingesetzt werden. Unter anderem durch die Entwicklung neuer Programme hat sich der Einsatz mathematischer Modellbildung in den letzten Jahrzehnten verändert.

In der jüngeren Vergangenheit wurden nahezu ausschließlich Studien zum Einsatz mathematischer Modellbildung in (US-amerikanischen) Universitäten durchgeführt. Hier scheint es einige Hinweise darauf zu geben, dass eine Umstrukturierung der

Einführungsveranstaltungen in die Physik unter einem konstruktivistischen Gesichtspunkt, in die die Nutzung von mathematischer Modellbildung eingebaut wird, positive Resultate erzielen kann. Dabei wird meist die Software „VPython“ genutzt. Dies ergibt insbesondere Sinn, da Studierende der Physik und der Ingenieursstudiengänge ohnehin grundlegende Programmierkenntnisse besitzen sollten und sich so ein Synergieeffekt ergibt. „VPython“ wurde auch bereits in Schulen eingesetzt. Schüler*innen der Sekundarstufe haben mit der Programmiersprache „Python“ offenbar aber größere Probleme, sodass es einer intensiven Einarbeitung bedarf. Da diese Zeit im Physikunterricht oft nicht zur Verfügung steht, bietet es sich an, Programme zu nutzen, die speziell für den Physikunterricht gemacht und damit einfacher zu bedienen sind. Weitere aktuelle Studien zum Einsatz mathematischer Modellbildung in Sekundarstufen fehlen. In den 1990er- und 2000er-Jahren wurden vor allem graphische Modellbildungsprogramme untersucht, wobei die weitreichenden Erwartungen nicht gänzlich erfüllt werden konnten. Es ist nötig, dass weitere Studien zum Einsatz von mathematischer Modellbildung in Schulen gemacht werden. Die Datenlage bei gleichungsbasierenden Modellbildungsprogrammen wie „Newton-II“ ist schlechter.

Außerdem ist unklar, wann ein*e Nutzer*in erfolgreich mit dem Computer Physik lernt. Nötig sind je nach Ausführung physikalische, mathematische und informationstechnologische Kenntnisse. Es ist wichtig herauszufinden, welche Voraussetzungen erfolgreiche Lernende schon mitbringen und wie sich das Lernen am Computer bei ebenjenen gestaltet. Es ist auch fraglich, ob das Modellieren am Computer sich an bekannten Modellierungsstrukturen orientiert und ob dies ein Kriterium für Erfolg ist. Zudem ist nicht erforscht, ob und inwiefern sich das Modellverständnis der Lernenden durch die Nutzung von mathematischer Modellbildungssoftware im Mechanikunterricht verändert.

6 Datenerhebung und -auswertung im Physikunterricht

6.1 Allgemeines zur Datenerhebung im Physikunterricht

Für das wissenschaftliche Arbeiten in der Physik und auch im Physikunterricht sind Experimente essenziell. Sie dienen neben vielen anderen Funktionen u. a. dazu, Daten zu generieren, die bspw. zur Überprüfung theoretisch hergeleiteter Gesetzmäßigkeiten dienen können (Kircher, Girwidz & Häussler, 2007). Die Art, wie Daten erhoben und ausgewertet werden, ist stark von der Art des Experiments abhängig. In der Mechanik haben historisch insbesondere Längenmessungen mit Lineal, Maßband etc. und Zeitmessungen mit einer Stoppuhr einen hohen Stellenwert. Diese Art der Messung dient in der Regel bei gleichförmigen Bewegungen zur Bestimmung der Geschwindigkeit oder bei beschleunigten Bewegungen zur Bestimmung einer durchschnittlichen Geschwindigkeit. Bei komplexeren Bewegungen kommt diese Messmethode allerdings an ihre Grenzen. Hinzu kommt, dass die Datenauswertung von Hand auf Papier erfolgen muss. Andere Bewegungen, die vielversprechend erscheinen, was das Erlernen der kinematischen Größen und auch der Zusammenhänge zwischen Kräften und Bewegung angeht, lassen sich mit diesen einfachen Methoden kaum mehr untersuchen.

Es gibt mittlerweile eine Vielzahl an Sensoren und Messwerterfassungssystemen, die seit einiger Zeit einen festen Platz in den Physikräumen in deutschen Schulen haben. Im Bereich der Mechanik sind unter anderem Kraft-, Beschleunigungs-, Bewegungs-, und Rotationssensoren zu nennen, mit denen man eine Vielzahl an Bewegungen ausmessen kann. Die Hersteller solcher Sensoren vertreiben ebenfalls Software, mit der man die erhobenen Daten automatisiert auswerten und in verschiedener Weise darstellen kann. Aber nicht jede Bewegung lässt sich auf diese Weise optimal vermessen. Häufig ist es im Bereich der Mechanik einfacher, ein Video einer Bewegung anzufertigen und dieses mit Hilfe von Videoanalyse zu analysieren (Suleder, 2020). Neben der geringen Komplexität des Aufbaus sind auch die Kosten der Ausrüstung ein Vorteil der Videoanalyse. Schüler*innen besitzen mobile Geräte mit guter Kamera, auf deren Betriebssystemen die Videoanalyse betrieben werden kann. Auch wenn man die Videoanalyse mit gesonderten Videokameras durchführt und mit Software am PC kombiniert, ist dies meist einfacher und günstiger als einen geeigneten Aufbau mit Sensoren vorzunehmen.

Da es auch in der Forschung einige Ergebnisse gibt, die der Videoanalyse im Vergleich zu klassischen Unterrichtsmethoden bzw. Messwerterfassungen gesteigertes Interesse und Motivation und einen höheren Lernzuwachs zuschreiben (z. B. Becker et al., 2020a),

wird dieses Instrument hier näher beleuchtet und die historische Entwicklung und Forschungsergebnisse dargestellt.

6.2 Videoanalyse von Bewegungen

6.2.1 Definition

Videoanalyse im Allgemeinen meint die systematische Analyse von Videos jeglicher Art. Mit Videoanalyse können Bewegungsabläufe im Sport zur Leistungsoptimierung oder auch soziale Prozesse im Schulunterricht untersucht werden. Die Videoanalyse von Bewegungen, auf die sich hier ausschließlich bezogen wird, ist eine kontaktlose Messmethode, die sich für den Physikunterricht vor allem im Bereich der Mechanik anbietet. Es ist mit ihr möglich, ohne weitere Sensoren zweidimensionale Bewegungen zu analysieren und die Messdaten auf vielfältige Weise darzustellen. Dazu wird der Ort eines Objekts, welches sich maximal in zwei Dimensionen bewegen darf (sofern nur eine Kamera genutzt wird) in regelmäßigen Zeitabständen über die Pixel in einem Video gemessen. Dies wiederum basiert darauf, dass ein Video aus Einzelbildern besteht, die in einer gewissen Rate aneinandergereiht sind. Aus der Veränderung des Ortes können dann durch Differenzenquotienten andere Größen (Geschwindigkeit und Beschleunigung) abgeleitet werden. Für die Berechnung weiterer Größen, wie Kraft oder Energie, werden dann zusätzliche Informationen benötigt (z. B. die Masse) (Suleder, 2020). Neben der Messung und Berechnung erfüllt die Videoanalysesoftware auch den Zweck der Datenauswertung, -aufbereitung und -darstellung. Die Videoanalyse ist in den Bereich der Messverfahren einzuordnen. Sie hat im Vergleich zu Messwerterfassungssystemen einige Vorteile, die im Folgenden genauer beleuchtet werden, nachdem die historische Entwicklung der Videoanalysesoftware und ihres Einsatzes dargestellt wird.

6.2.2 Entwicklung der Videoanalysesoftware

Videoanalyse wird für das Lehren von Physik seit den 1980er-Jahren eingesetzt (Laws, 1991a). Zu Beginn bestand die Videoanalyse daraus, manuell an einem Bildschirm die Position eines Objektes zu bestimmen und die Bewegung zu analysieren (z. B. Ducas, 1993). In der Folge wurden Programme entwickelt, mit denen eine Analyse durch mehrmaliges Klicken auf das Objekt mit der Maus gelang, woraufhin auch Programme mit einer automatisierten Analysefunktion und der Option für eine Echtzeit-Videoanalyse auf den Markt kamen (Weber & Wilhelm, 2020b). Neben der Qualität und der Art der Aufnahme der Messdaten verbesserte sich analog auch deren Darstellung, sodass immer mehr Programme auch didaktischen Design-Prinzipien gerecht wurden (siehe Kapitel 4.3). Der

Verlauf der Entwicklung des Einsatzes von Videoanalysesoftware wird in der Folge dargestellt.

Bevor es spezielle Videoanalyseprogramme gab, konnte man diese Analysetechnik nutzen, indem man ein Video einer Bewegung auf einem beliebigen Bildschirm abspielte und die Position des Objekts in regelmäßigen Zeitabständen festhielt (bspw. durch Abpausen). Diese Messpunkte mussten dann durch ein Koordinatensystem skaliert und entweder per Hand ausgewertet oder in eine andere Software eingetragen werden. Da dieses Verfahren sehr aufwendig ist, war der Bedarf für spezielle Videoanalysesoftware vorhanden, die den Nutzer*innen diese Arbeit abnimmt. Die ersten Programme, die speziell für die Videoanalyse gemacht wurden, entstanden in den 1990er-Jahren. Der Funktionsumfang dieser Programme war allerdings sehr limitiert und beschränkte sich fast ausschließlich auf die Darstellung von Messdaten in Tabellenform. Vereinzelt konnten auch Kurven zu den Messdaten gefittet werden. Didaktisch waren diese Programme allerdings nicht optimiert und hatten mit einigen technischen Hürden zu kämpfen, die erst in den folgenden Generationen an Videoanalyseprogrammen überwunden wurden (Suleder, 2020). Von der vorigen Vorgehensweise der Videoanalyse unterschieden sich die neueren Programme vor allem darin, dass man die Position des Objekts per Mausklick ermitteln konnte, wodurch die Software die Position automatisch übernahm und dass die Messdaten innerhalb der Software bereits ausgewertet werden konnten.

Eine weitere große Verbesserung der Programme in der folgenden Dekade war die automatische Analysefunktion, die es dem Nutzer ermöglicht, mit einem einzigen Klick das Objekt im gesamten Video zu verfolgen. Dies ist vor allem bei langen Videos mit vielen Einzelbildern eine große Erleichterung gegenüber der manuellen Analyse. Für ein Gelingen der automatischen Videoanalyse ist allerdings die Qualität des Videos von größerer Bedeutung, da sich zwar auch die Analysealgorithmen im Laufe der Zeit verbessert haben, die automatische Analyse bei ungünstigen Lichtverhältnissen aber bspw. immer noch fehlschlagen kann. Auch im Bereich der Datenauswertung und -darstellung entwickelte sich die Software weiter. Neben der Darstellung der Messdaten in Diagrammform gibt es mittlerweile viele weitere didaktisch sinnvolle Darstellungen in Videoanalyseprogrammen. Beispielsweise bietet die Software „measure dynamics“ neben einer Wertetabelle verschiedene, parallel zum Video entstehende Diagramme an. Außerdem können Pfeile (z. B. Geschwindigkeit und Beschleunigung) direkt im Video an das Objekt gezeichnet werden. Weitere Optionen der Darstellung der Bewegung sind Stroboskopbilder und -videos. Neben „measure dynamics“ wird international vor allem die Open-Source Software „Tracker“ eingesetzt, die es neben der Videoanalyse auch erlaubt, eine

mathematische Modellbildung vorzunehmen, die ebenfalls mit dem Video verknüpft und in diesem angezeigt wird (Weber, 2017).

Die nächste Stufe der Entwicklung für Videoanalyseprogramme war der Sprung zu mobilen Endgeräten, die eine Kamera bereits integriert haben und damit keine weitere Ausrüstung außer dem Gerät mehr nötig ist. Der Schritt des Überspielens des Videos fällt damit weg. Zudem sind Schüler*innen heutzutage mit Mobilgeräten in der Regel vertrauter als mit Laptops. Als Beispiel ist hier das Programm „Viana“ zu nennen, das auf Tablet-PCs und Smartphones funktioniert. Einen Vergleich verschiedener Videoanalyse-Apps auf Tablets gibt Suri (2021).

Neuerdings gibt es auch Programme, wie „eXperilyser“, die eine Echtzeit-Videoanalyse ermöglichen (Weber & Wilhelm, 2020b), die dadurch eine hohe zeitliche Kontiguität (siehe Kapitel 4) schafft und den Sprung vom Realexperiment zum entsprechenden Bewegungsgraphen weiter erleichtern soll. Hier wird ein weiterer Schritt eingespart. Es muss nun nicht mehr das Video aufgezeichnet und danach die Videoanalyse gestartet werden. Es werden gewisse Größen, wie z. B. Geschwindigkeit-Zeit-Diagramme, direkt im Bildschirm angezeigt, während sich das Objekt bewegt. Dies ist vor allem für längere Bewegungen (wie Schwingungen) geeignet, bringt bei kürzeren, schnelleren Bewegungen allerdings keinen erheblichen Mehrwert, insbesondere da der Funktionsumfang des Programms im Vergleich zu anderen Programmen sonst eingeschränkter ist. Die genannten Entwicklungsschritte erfolgten dabei nacheinander, allerdings wird heutzutage sowohl Software für PCs und Laptops als auch Software für mobile Endgeräte weiterentwickelt, beides wird erforscht und im Schulunterricht eingesetzt.

6.2.3 Gründe für den Einsatz von Videoanalyseprogrammen im Physikunterricht

Videoanalyse ist eine einfache Messmethode für Bewegungen jeglicher Art. Heutzutage hat jede*r Schüler*in ein mobiles Endgerät, welches alle für die Videoanalyse nötigen Funktionen enthält. Es ist in vielen Fällen einfacher, eine Bewegung aufzunehmen und mit einer Videoanalysesoftware zu analysieren als einen Aufbau mit externen Sensoren zu entwerfen, welcher die Daten an einen PC schickt, auf dem diese mit einem entsprechenden Programm ausgewertet werden können. In manchen Fällen ist eine Messwerterfassung über Sensoren sogar nahezu unmöglich, während dies mit einer Videokamera und einem Videoanalyseprogramm leicht gelingt (Suleder, 2010). Viele Messmethoden über Sensoren funktionieren zudem nur bei eindimensionalen Bewegungen und die beeinflussen die Bewegungen teilweise selbst (bspw. bei Messung über einen Faden). Gleichzeitig ermöglichen Smartphones auch, dass Lernende einzeln, in Partnerarbeit oder in Gruppen

an eigenen Projekten arbeiten können, da keine weitere Ausstattung benötigt wird. Das macht einen schülerzentrierten und projektbasierten Unterricht deutlich einfacher. Gleichzeitig kann die Kluft zwischen Physikunterricht und alltäglichem Leben reduziert werden, da reale Bewegungen aus dem Alltag der Schüler*innen Anlass zur Untersuchung der physikalischen Gesetzmäßigkeiten geben können. Videoanalyse zeichnet sich gegenüber anderen Messmethoden aber nicht nur durch seine Einfachheit und Praktikabilität aus, sondern weist einige Funktionen auf, die für den Lernprozess förderlich sein können.

Im Mechanikunterricht spielen Diagramme eine bedeutende Rolle, um Zusammenhänge zwischen verschiedenen Größen darzustellen (Suleder, 2010). Solche Diagramme sind auf die wesentlichen Informationen reduziert, die nötig sind, um eine Bewegung zu verstehen. Diese Reduktion auf das Wesentliche geht aber mit einer erheblichen Abstraktion von der eigentlichen Bewegung einher. Die Videoanalyse kann einen Ansatz bieten, mit Zwischenstufen zwischen der eigentlichen Bewegung und den Diagrammen das Verständnis dieser zu erhöhen (Beichner, 1996). Zudem kann die Verknüpfung zwischen Graph und Bewegung verbessert werden (Boyd & Rubin, 1996). In ähnlicher Form kann die Videoanalyse auch bei der Übersetzung der realen Bewegung in andere mathematische Repräsentationen, wie z. B. Tabellen, als Zwischenstufe agieren und den Transfer damit vereinfachen (Pappas et al., 2002). Multicodierung bezeichnet dabei die Verwendung verschiedener Symbolsysteme zur Verdeutlichung eines Sachverhalts (Weidenmann, 2002). Eine komplexe Situation kann dadurch auf unterschiedlichen Abstraktionsniveaus dargestellt werden (z. B. Video, Pfeile und Diagramme). Nach Weidenmann (2002) fördert dies u. a. Interesse, die Entwicklung adäquater mentaler Modelle und flexibles Denken.

Videoanalysesoftware bietet eine Vielzahl an Darstellungsvarianten, die sich gegenseitig ergänzen. Die Arbeit mit solchen multiplen Repräsentationen (Becker et al., 2020a) kann sich positiv auf den Lernerfolg auswirken. Auch im Bereich des Cognitive Load (siehe Kapitel 4.1) scheint die Videoanalyse vielversprechend zu sein. So könnte durch ein hohes Maß an räumlicher und zeitlicher Kontiguität die kognitive Belastung reduziert werden. Inwiefern diese hypothetischen Vorteile der Videoanalyse bereits durch Forschungsstudien belegt wurden, wird im folgenden Kapitel diskutiert.

6.2.4 Forschungsergebnisse

Es existieren bereits einige Studien, die den Erfolg von Videoanalyse in unterschiedlichen Kontexten zeigen. Dabei ist einerseits der Bereich der fachlichen Auswirkungen zu

nennen, der sich auf verschiedene Bereiche aufteilt (Kinematik, Dynamik). Andererseits wurde auch der Einfluss der Videoanalyse auf viele affektive Merkmale wie Interesse und Motivation untersucht.

Erste Untersuchungen zur Übersetzung einer Bewegung in den zugehörigen Bewegungsgraphen mit der Hilfe von Videos machte Beichner (1990), an dessen Studie $N = 237$ Studierende und Schüler*innen teilnahmen. Zwar nutzte er dazu keine klassische Videoanalysesoftware, lies den Bewegungsgraphen aber simultan zu einer Animation der Bewegung entstehen, die wiederum aus echten Fotos der Bewegung bestand. Dabei zeigt sich, dass Proband*innen, die simultan zum entstehenden Bewegungsgraphen ein Video sehen, kein größeres Verständnis für den Graphen zeigen, obwohl die Parallelität zwischen Realexperiment und Graphenentstehung in anderen Studien bereits als lernförderlich eingestuft wurde. Es könnte also einen Unterschied machen, ob der Graph parallel zum realen Experiment oder zu einem aufgezeichneten Video erscheint. Die Kontrollgruppe nutzte Stroboskopbilder zur Analyse der Bewegung. Beide Funktionen sind heutzutage in Videoanalyseprogrammen standardmäßig enthalten, sodass auch die Nutzung von Stroboskopbildern mittlerweile in den Bereich der Nutzung von Videoanalyse gezählt werden kann. Es konnte außerdem kein Unterschied zwischen den Gruppen beobachtet werden, die einerseits das Experiment zusätzlich real zu sehen bekamen, und denjenigen, die nicht diese Möglichkeit hatten.

Brungardt und Zollman (1995) führten eine Studie mit 31 Oberstufenschüler*innen durch. Die Lernenden analysierten verschiedene Bewegungen aus dem Bereich des Sports, indem sie ein Papier über den Bildschirm legten und die Position des Objekts zu verschiedenen Zeitpunkten festhielten. Dieses Papier wurde für die Skalierung auf ein Koordinatensystem aufgebracht und die so gewonnenen Daten wiederum in eine Software eingepflegt, die dann Geschwindigkeit und Beschleunigung berechnete und die Graphen anzeigen konnte. Die unabhängige Variable der Untersuchung war die Zeit zwischen der Erzeugung des Graphen und dem Abspielen des Videos, die abhängige Variable das Verständnis von Bewegungsgraphen. Bei einer Gruppe erfolgte das Abspielen simultan, in der anderen lag einige Zeit zwischen dem Ablaufen des Videos und der Erzeugung des Graphen. Es zeigte sich kein Unterschied zwischen beiden Gruppen, allerdings wurden die Prätestergebnisse nicht kontrolliert. Es konnte aber gezeigt werden, dass die Gruppe mit simultanem Ablauf motivierter war. Bryan (2010) argumentiert allerdings, dass solche Studien aufgrund der neueren Entwicklungen im Bereich der Videoanalysesoftware repliziert werden müssen, um Hinweise darauf zu erhalten, ob die gefundenen Ergebnisse unter heutigen Bedingungen immer noch gültig sind.

Eine weitere Studie von Beichner (1996) konnte eine positive Auswirkung von Videoanalyse auf das Interpretationsvermögen von Bewegungsgraphen nachweisen. Dazu nahmen 386 Schüler*innen und Student*innen an der Studie teil, in der die Software „Videograph“ genutzt wurde. Beichner (1996) konnte zeigen, dass der Lernerfolg umso größer war, je intensiver die Videoanalyse eingesetzt wurde. Dabei war es wichtig, dass die Lernenden selbstständig mit der Software arbeiten konnten. Ein alleiniger Einsatz durch den Lehrer brachte keine positiven Effekte. In demselben Jahr konnten Escalada et al. (1996) beim Einsatz von Videoanalyse für reale zweidimensionale Bewegungen zeigen, dass dies zu einer Verbesserung bei Lernenden im Bereich der Visualisierung, Untersuchung, Analyse und Verständnis von Bewegungen führen kann. Sie nutzten die Programme „VideoAnalyzer“ und „Visual Space-Time“ für fünf verschiedene Aktivitäten, die mit Studierenden durchgeführt wurden. Die Autor*innen berichten von wenig Schwierigkeiten bei der Nutzung der Programme und postulieren, dass sie sehr effektiv beim Lernen der Konzepte im Bereich der Koordinatensysteme waren (Escalada et al., 1996).

Boyd und Rubin (1996) untersuchten die Wirksamkeit von Stroboskopbildern für das Konzeptverständnis von Bewegungen. Dabei ging es ihnen vor allem darum, durch die Stroboskopbilder eine Brücke zwischen der Bewegung und der nötigen mathematischen Repräsentationen zu schlagen. Die Verbindung zwischen alltäglichen Wahrnehmungen und den mathematischen Repräsentationen stellt das Video dar. Dabei erstellten die Lernenden die Bewegungsgraphen selbst, was sonst häufig von der Software erledigt wird. Die Autoren sehen diese Unterrichtsmethode als geeignet an, wobei die Ergebnisse auf Interviews mit lediglich zwei Sechstklässlern basieren.

Laws und Pfister (1998) setzten Videoanalyse als zentrales Werkzeug ein, um einen auf Experimenten basierenden und studentenzentrierten Einführungskurs in die Mechanik an der Universität zu kreieren. Dazu arbeiteten die Studierenden projektbasiert mit „Video point“ und „VideoGraph“ und meldeten zurück, dass sie durch diese Projekte mehr lerneten als bei anderen zum Kurs gehörenden Aktivitäten. Eine Quantifizierung dieser Aussagen oder eine Messung des Lerneffekts wurde von den Autoren allerdings nicht durchgeführt. Andere Untersuchungen zum viel eingesetzten „Workshop Physics“ zeigten aber einige positive Ergebnisse. Zu nennen sind hier vor allem ein gesteigertes Konzeptverständnis in Mechanik im Vergleich zum traditionellen Vorgehen (Laws et al., 2015), ein größeres Interesse und kein Unterschied im Lösen von klassischen Textbuch-Aufgaben (Laws, 1991b). Diese Ergebnisse betreffen den gesamten „Workshop Physics“, in dem im Vergleich zu herkömmlichen Kursen einiges verändert wurde, und können damit nicht auf die Videoanalyse zurückgeführt werden.

Rodrigues et al. (2001) untersuchten Videoanalyse und Messwerterfassung mit externen Sensoren im Praktikum für Physikstudierenden. Während 47 Studierende eine Intervention mit Videoanalyse absolvierten, führten 160 Studierende den herkömmlichen Verlauf mit Messwerterfassung durch. Die Studierenden lobten den Kurs generell, Unterschiede zwischen den Vorgehensweisen wurden aber nicht gefunden. Angemerkt wird von den Autor*innen, dass Studierende möglicherweise nur ihre bestehenden Präkonzepte untermauern und ihre Ideen nicht infrage stellen. Dies deutet darauf hin, dass es nötig ist, die Ergebnisse mit Lernenden zu diskutieren und während der Arbeitsphasen zum Hinterfragen der bestehenden Vorstellungen anzuregen.

Pappas et al. (2002) hatten das Ziel, die Videoanalyse als Brücke zwischen Realsituationen und mathematischen Repräsentationen einzusetzen. Sie konzentrierten sich in ihrer Studie dabei auf den Bereich der Tabellen. Die Studie wurde mit Lehramtsstudierenden durchgeführt, wobei auch hier nur eine kleine Anzahl an Proband*innen teilnahm ($N = 5$). Deswegen sind die Ergebnisse kaum generalisierbar. Genutzt wurde die Software „Videopoint“, um den Proband*innen ein Verständnis von Koordinatentabellen als Repräsentation einer realen Bewegung und die Bedeutung des Referenzkoordinatensystems zu vermitteln. Es ist allerdings nicht sicher, inwiefern die beobachteten Verbesserungen (Interpolation von Kurven, Schnittpunktbestimmung, Vorzeichen der Geschwindigkeit, Referenzsysteme) auf die Videoanalyse zurückzuführen sind, da keine Vergleichs- oder Kontrollgruppe vorlag.

Ohne Videoanalyse selbst einzusetzen nutzte Wilhelm (2005a) mit großem Erfolg dynamisch ikonische Repräsentationen, die in heutigen Videoanalyseprogrammen vorhanden sind, um Schüler*innen zweidimensionale Bewegungen näher zu bringen. Die Elftklässler*innen, die u. a. dynamisch ikonische Repräsentationen nutzten, verbesserten sich im Vergleich zum herkömmlichen Unterricht stärker – insbesondere in der zweidimensionalen Kinematik und der Dynamik. Dieser Unterschied kann allerdings nicht einer spezifischen Variable zugeordnet werden (Wilhelm, 2005a). Auf den Ergebnissen aufbauend wurde ein Unterrichtskonzept zur Einführung in die Mechanik in der Sekundarstufe 1 entwickelt, welches von zweidimensionalen Bewegungen ausgeht und dynamisch ikonische Repräsentationen nutzt. Videoanalyse kann hier sinnvoll eingebunden werden (Waltner et al., 2010).

Palazzo und Schools (2008) setzten „LoggerPro“ mit Studierenden ein und kamen zu dem Schluss, dass die Videoanalyse bei einem Großteil der Studierenden dazu führt, eine Verbindung zwischen der Physik im Klassenraum und der realen Welt herzustellen. Der eigene Lernzuwachs im Hinblick auf den Einsatz von Videoanalyse wurde von den

Studierenden aber eher negativ bewertet. Die Lehrkräfte waren unterschiedlicher Meinung, was die Sinnhaftigkeit eines Einsatzes von Videoanalyse im Mechanik-Praktikum angeht.

Eine andere Art der Videoanalyse wurde von Kanim und Subero (2010) eingesetzt. Sie untersuchten den Einsatz von digitaler Fotografie mit langer Belichtungszeit in der Mechanik von Universitätskursen. An den sich bewegenden Objekten waren regelmäßig blinkende Leuchtdioden befestigt. Die Fotos wurden genutzt, um Studierenden ein besseres Verständnis der vektoriellen Natur von kinematischen Größen zu vermitteln und die Beschleunigung mit der Gesamtkraft zu verknüpfen. Diese Technik stellte sich allerdings als nur bedingt erfolgreich heraus. Die meisten Studierenden hatten auch nach der Intervention Schwierigkeiten mit der Beschleunigung und ihrer Verbindung zu den wirkenden Kräften (Kanim & Subero, 2010).

Struck und Yerrick (2010) konnten, wie bereits andere Autor*innen vor ihnen, nachweisen, dass die Videoanalyse beim Verständnis und dem Erstellen von Geschwindigkeits-Zeit-Graphen und Beschleunigungs-Zeit-Graphen hilft. Dazu wurden 39 Schüler*innen einer US-amerikanischen Oberstufenschule untersucht und der Einsatz von Sensoren mit dem der Videoanalyse mit „LoggerPro“ in einem Kreuzdesign verglichen. Die erbrachten Leistungen bei den genannten Graphenarten waren nach der Intervention höher als nach der Nutzung eines MBL (Microcomputer-based Laboratory). Bei Ort-Zeit-Graphen war kein Unterschied zwischen den Methoden zu erkennen.

Wee et al. (2012) nutzten „Tracker“ (Brown & Cox, 2009), um den Flug eines Projektils zu untersuchen. Die positiven Rückmeldungen der Studierenden zeigen die hohe Akzeptanz der Nutzung einer solchen Software. Die Wahrnehmung der Unterrichtssequenz war durchweg positiv, wurde allerdings nicht quantifiziert. Ebenfalls „Tracker“ nutzten Rodrigues und Simeão Carvalho (2013), um in Mittel- und Oberstufe Mechanik mithilfe des Spiels „Angry Birds“ zu unterrichten. Der Lernerfolg wurde hier allerdings ebenfalls nicht analysiert.

Hockicko et al. (2014) nutzten die Videoanalysesoftware „Tracker“, um das Verständnis für den Bremsprozess eines Automobils zu unterrichten. Sie fanden heraus, dass die Lernenden beim Einschätzen des Bremswegs einen systematischen Fehler machten, der durch die Intervention mit Videoanalyse korrigiert werden konnte. Das Konzeptverständnis verbesserte sich im Vergleich zu herkömmlich unterrichteten Studierenden stärker. Außerdem meldeten die Lernenden zurück, dass es sich für sie um eine interessante Unterrichtsmethode handelte, bei der ihnen die Gruppenarbeit zusagte.

Ambrosis et al. (2015) setzten Videoanalyse für das Lehren der Rollbewegung mit 20 höhersemestrigen Lehramtsstudierenden für das Gymnasium ein. Dabei wurde beobachtet, dass die Proband*innen durch die Sequenz einige Lernschwierigkeiten überwinden konnten und sich ihr Konzeptverständnis der Rollbewegung verbesserte. Allerdings gab es hier keine Kontrollgruppe und es ist nicht zu ermitteln, welchen Anteil daran die Videoanalyse hatte.

Wee et al. (2015) erforschten den Einsatz der Videoanalysesoftware „Tracker“ beim Erlernen eines senkrechten Wurfs und der freien Fallbewegung mit $N = 123$ 15-jährigen Oberstufenschüler*innen. Die Ergebnisse der Interventionsgruppe in Prä- und Posttest unterschieden sich mit großer Effektstärke. Es lag allerdings keine Kontrollgruppe vor. Der relative Lernzugewinn wurde anhand von Hakes g^{12} (hier $g = 0,42$) mit dem für den FCI-Test (Hestenes et al., 1992) typischen relativen Lernzugewinn von g von 0,23 verglichen, der aber nur auf dem Durchschnitt der Metaanalyse von Hake zum FCI-Test beruht (Hake, 1998), sodass keine Aussage darüber getroffen werden kann, wieviel die Proband*innen in einer äquivalenten Unterrichtseinheit ohne Videoanalyse gelernt hätten.

Videoanalysesoftware kann auch genutzt werden, um Animationen statt Videos von Bewegungen zu analysieren (Claessens, 2017; Filsinger & Wilhelm, 2020). Die Autoren berichten aber nicht über den Erfolg dieser Herangehensweise.

Während „Tracker“ meist an Laptops eingesetzt wird, gibt es einige neuere Studien, die sich speziell mit dem Einsatz von mobilen Endgeräten für die Videoanalyse beschäftigen. Klein et al. (2018) untersuchten den Einfluss von Videoanalyse auf die Repräsentationskompetenz und den Experimentbezug in der Mechanik mit $N = 76$ Erstsemestern. Dabei zeigte sich, dass Studierende, die im Mechanik-Praktikum mit Videoanalyse lernen, eine mit mittlerer bis großer Effektstärke höhere Repräsentationskompetenz im Vergleich zu den herkömmlich unterrichteten Studierenden haben. Die Repräsentationskompetenz umfasst dabei Interpretationen im Umgang mit Diagrammen, Formeln, Vektoren und Bildern (Klein et al., 2018).

Becker et al. (2019a) setzten Videoanalyse auf Tablet-PCs in den Bereichen „Gleichförmige Bewegung“ und „Beschleunigte Bewegung“ in der Sekundarstufe II ein. An der Studie nahmen im Bereich der gleichförmigen Bewegung $N = 109$ und im Bereich der „Beschleunigten Bewegung“ $N = 70$ Proband*innen teil. Sie konnten in beiden

¹² $g = \frac{\text{tatsächlicher Zugewinn}}{\text{maximal möglicher Lerngewinn}}$

Bereichen ein mit großer Effektstärke höheres Konzeptverständnis bei den Lernenden feststellen, die mit der Videoanalyse arbeiteten als bei denjenigen, die herkömmlich unterrichtet wurden. Beim anspruchsvolleren Thema der beschleunigten Bewegungen war der Unterschied größer als bei der gleichförmigen Bewegung (Becker et al., 2019a). Die Autoren konnten außerdem zeigen, dass die kognitive Belastung (siehe Kapitel 4.1) durch den Einsatz von Videoanalyse reduziert werden konnte (Becker et al., 2019a, 2020a). Gleichzeitig wurden die in der vorigen Studie berichteten Ergebnisse mit größerer Probandenzahl repliziert und bestätigt. Außerdem konnte eine Reduktion von „negativ-deaktivierenden“ Emotionen bei Nutzung der Videoanalyse nachgewiesen werden, während kein Unterschied bei den „positiv-aktivierenden“ Emotionen beobachtet wurde. Diese negativen Emotionen wirken sich nachweislich auf den Extranous Cognitive Load (ECL) aus und führen damit auch zu einem geringeren Lernzuwachs (Becker et al., 2020a).

Hochberg et al. (2020) testeten eine Lernumgebung zu Schwingungen mit mobiler Videoanalyse mit 52 Schüler*innen von Oberstufengymnasien. Es konnte ein signifikanter Unterschied mit mittlerer Effektstärke festgestellt werden, was den Lernerfolg der Proband*innen zugunsten der Gruppe der Videoanalyse betrifft. Bei Interesse und Neugierde ergab sich hier allerdings kein signifikanter Unterschied zur Kontrollgruppe, die mit Lineal und Stoppuhr arbeitete. Mögliche Gründe sind laut Autor*innen ein Unterschied zwischen Smartphones, bei denen Studien bereits positive Effekte auf Interesse und Neugierde nachgewiesen haben, und Tablets, die hier verwendet wurden, und die Stichprobe. In dieser waren einerseits nur fünf Kurse, andererseits hatten sich diese bereits aktiv für Physik entschieden und wiesen bereits zu Beginn ein hohes Interesse auf, was den Einfluss der Videoanalyse auf das Interesse reduzieren könnte.

6.2.5 Heutige Nutzung von Videoanalysesoftware in Schulen

Ähnlich wie bei der mathematischen Modellbildung wird von heutiger Software für den Physikunterricht erwartet, dass sie intuitiv zu bedienen ist, stabil läuft und die Kontiguitätsprinzipien berücksichtigt. Die Nutzung von Videoanalysesoftware im Unterricht ist immer mit einem gewissen Zusatzaufwand verbunden, der durch das Einarbeiten in die Software entsteht. Dieser darf nicht so groß sein, dass er zu viel Unterrichtszeit kostet und die Vorteile der Videoanalyse zunichtemacht. Anders als im Bereich der mathematischen Modellbildung berichten nahezu alle Studien von geringen Schwierigkeiten beim Einsatz von Videoanalysesoftware. Grund dafür könnte auch sein, dass Schüler*innen insbesondere im Bereich der mobilen Endgeräte bereits Routine in der Bedienung haben.

Aktuelle Zahlen zur Computernutzung im Physikunterricht an deutschen Schulen stammen auch im Bereich der Videoanalyse von einer Erhebung, die von 2014 bis 2015 in drei deutschen Regionen stattgefunden hat. Es wurden Daten von insgesamt $N = 163$ Lehrer*innen erhoben, die sich auf die drei Regionen verteilten (Wenzel, 2018). Auch hier ergaben sich Unterschiede in der Häufigkeit der Nutzung zwischen den Regionen, die mit unterschiedlichen Lehrplänen und auch der Herkunft der verschiedenen Software zu erklären sind. Insgesamt werden Messwerterfassungssysteme mit externen Sensoren häufiger eingesetzt als die Videoanalyse, welche nur von 5 % der befragten Lehrkräfte häufiger als alle paar Monate genutzt wird (im Vergleich zu 27 % bei Messwerterfassung). 37 % der Befragten gaben an, Videoanalyse noch nie genutzt zu haben (im Vergleich zu 22 % bei Messwerterfassung) (Wenzel, 2018).

6.2.6 Zusammenfassung

Einige der vorgestellten Forschungsstudien im Bereich der Videoanalyse leisteten vor allem Entwicklungsarbeit, evaluierten diese allerdings nicht. Weitere Studien setzten die Videoanalyse zwar ein und analysierten den Erfolg dieses Einsatzes, änderten außer dem Einsatz der Videoanalyse allerdings weitere Variablen, sodass wenig über den Effekt der Videoanalyse selbst ausgesagt werden kann. Aus anderen Studien lassen sich Handlungsempfehlungen für den Einsatz von Videoanalyse im Unterricht und auch in weiteren Studien ableiten. Dabei sind neuere Ergebnisse durch die zwischenzeitliche Weiterentwicklung der Software höher zu gewichten. Außerdem muss bei der Gewichtung und Einordnung der Forschungsergebnisse beachtet werden, wie kontrolliert die Studien durchgeführt wurden.

In den Ergebnissen der vorgestellten Forschungsstudien lassen sich einige Gemeinsamkeiten erkennen. Die Ergebnisse bezüglich des Interesses und der Motivation von Lernenden sind eindeutig. Der Einsatz von Videoanalyse wirkt motivierend und Lernende haben ein großes Interesse an diesem Hilfsmittel. Zusätzlich lässt sich durch den Einsatz ein schülerzentrierter und projektorientierter Unterricht realisieren, der sich abermals positiv auf Interesse und Motivation auswirkt (Hake, 1998). Inwiefern diese Steigerung der Motivation auf einen Neuheitseffekt (siehe Kapitel 4) zurückzuführen ist, ist nicht abschließend geklärt. Für das fachliche Lernen scheint es aber wichtig zu sein, dass Schüler*innen selbstständig mit der Software arbeiten. Dass dies einfach möglich ist, ist gerade einer der Vorteile der Videoanalyse. Insbesondere im Bereich der Kinematik wurden einige Zusammenhänge zwischen Videoanalyse und einem positiven Effekt auf den Lernzuwachs im Bereich von Diagrammen, Tabellen, Vektoren und verschiedenen Bewegungen gefunden.

Besonders hervorzuheben sind die aktuellen Studien von Becker et al. (2019a, 2019b, 2020a, 2020b), in denen durch ein kontrolliertes Studiendesign ein positiver Effekt der Videoanalyse auf die Repräsentationskompetenz, das Konzeptverständnis von gleichförmigen und beschleunigten Bewegungen, der kognitiven Belastung und positiven emotionalen Aspekten gezeigt werden konnte. Diese positiven Beobachtungen lassen sich durch die jeweiligen Forschungsdesigns eindeutig dem Faktor Videoanalyse zuordnen.

Insgesamt etabliert sich die Videoanalyse immer mehr im Physikunterricht, was auch durch die physikdidaktische Forschung gestützt wird. Die Nutzung von Videoanalyse kann im Vergleich zu herkömmlichen Herangehensweisen zu einem erfolgreicherem Unterricht der Kinematik führen. Weniger bekannt ist, ob sich die Videoanalyse auch zum Erlernen der Grundgleichung der Mechanik (also für den Bereich der Dynamik) und damit zu den Zusammenhängen zwischen Bewegung und wirkenden Kräften eignet. Auch wenn die Videoanalyse also als etabliertes und erfolgreiches Unterrichtsmedium angesehen werden kann, besteht weiterhin ein Forschungsinteresse – vor allem im Bereich der Dynamik.

II. Forschungsansatz

7 Einordnung der Arbeit und Forschungsfragen

7.1 Einordnung der Arbeit in den aktuellen Forschungsstand

Die neueren Ergebnisse im Bereich der Videoanalyse zeigen, dass diese den klassischen Unterrichtsmethoden mit einer replizierten mittleren Effektstärke überlegen ist, wenngleich die Ergebnisse im Bereich der Dynamik rar sind. Auch wenn in Schulen ähnlich positive Ergebnisse der mathematischen Modellbildung nicht so überzeugend gezeigt werden konnten, wird aufgrund der in Kapitel 5.4.4 dargestellten Argumente hypothetisiert, dass die mathematische Modellbildung für das Erlernen der Dynamik gut geeignet ist. Speziell geeignet für das Verbessern des Konzeptverständnisses scheint eine Software, die einfach zu bedienen und speziell für die Dynamik erstellt worden ist. Die Schüler*innen können sich somit auf den Zusammenhang zwischen Kraft und Bewegung konzentrieren. Es bietet sich ein Vergleich zwischen der mathematischen Modellbildung und einem Vorgehen mit Videoanalyse an, da Letztere dem klassischen Unterricht in weiten Teilen der Mechanik überlegen zu sein scheint.

7.2 Forschungsfragen

Die aus der Theorie abgeleiteten Forschungsfragen, die mit dieser Arbeit beantwortet werden sollen, teilen sich in zwei unterschiedliche Bereiche auf: Fragen, die eher mit einer quantitativen Erhebung beantwortet werden können, und Fragen, die eher mit einer qualitativen Erhebung beantwortet werden können. Durch diese Teilung wird die Studie einen Mixed-Methods-Ansatz verfolgen und zur Beantwortung der jeweiligen Forschungsfragen auch Daten aus dem anderen Teil heranziehen.

7.2.1 Quantitativer Teil

Bevor ein Vergleich der beiden Vorgehensweisen (mathematische Modellbildung und Videoanalyse) gezogen wird, ist zunächst von Interesse, inwiefern die beiden Methoden für das Erlernen der Dynamik überhaupt wirksam und geeignet sind. Daher lautet die erste Forschungsfrage:

- Führt eine computergestützte Intervention mit mathematischer Modellbildung oder Videoanalyse nach dem schulischen Unterricht der Newtonschen Gesetze zu einem messbaren Zuwachs im Konzeptverständnis?

Diese Frage ist in der bisherigen Forschung ungeklärt, da im Bereich der mathematischen Modellbildung aktuelle Studien vor allem mit Studierenden naturwissenschaftlicher Studiengänge und nicht mit Schüler*innen durchgeführt wurden. In der Videoanalyse ist der Erfolg innerhalb der Kinematik in vielen Fällen belegt, es fehlen jedoch Studien im Bereich der Dynamik. Weiterhin sollen diese beiden Vorgehensweisen miteinander verglichen werden:

- Welche Unterschiede lassen sich zwischen einem modellierenden Vorgehen mit mathematischer Modellbildung und einem messenden Vorgehen mit Videoanalyse beim Erlernen der Newtonschen Mechanik erkennen?

Die formulierte Hypothese ist, dass die mathematische Modellbildung aufgrund der intensiveren Beschäftigung mit Kräften und dem direkteren Feedback der Software besser für das Erlernen der Dynamik geeignet ist. Insbesondere werden dabei neben dem Konzeptverständnis auch andere Parameter wie Interesse und kognitive Belastung (siehe Kapitel 4.1) untersucht, da diese einen Einfluss auf den Lernerfolg haben können (Plass & Kaplan, 2016). Neben der reinen Gegenüberstellung dieser Variablen ist auch von Interesse, welche weiteren Variablen einen Einfluss auf den Lernerfolg haben.

Die beiden Methoden unterscheiden sich nicht nur in der Software. Bei der mathematischen Modellbildung wird von der Theorie ausgehend eine Bewegung modelliert und deren kinematische Größen dann mit der Realität verglichen, während bei der Videoanalyse ausgehend vom Ort des Objekts im Realexperiment verschiedene kinematische Größen berechnet werden, sodass dann auf die wirkenden Kräfte geschlossen werden kann. Es handelt sich somit um unterschiedliche Argumentationsrichtungen. Eine diesbezügliche Forschungsfrage, die der vorigen unterzuordnen ist, lautet:

- Wirken sich die unterschiedlichen Argumentationsrichtungen der beiden Methoden auf die Lösungswahrscheinlichkeiten von Items aus, die diese oder die entgegengesetzte Argumentationsrichtung verlangen? Also: Gibt es Unterschiede im Lernzuwachs der beiden Gruppen in den unterschiedlichen Argumentationsrichtungen zwischen Kräften und Bewegung?

Es kann außerdem argumentiert werden, dass sich die beiden Vorgehensweisen in ihrem Bezug zu Theorie bzw. Experiment unterscheiden. Fraglich ist, ob es Schülertypen gibt, die sich eher als Theoretiker*innen bzw. Experimentierende sehen, und ob dies einen Einfluss auf den Lernzuwachs in den jeweiligen Gruppen hat:

- Lassen sich anhand des Interesses an Theorie bzw. Experiment unterschiedliche Typen bei den Lernenden unterscheiden und ist es für den Lernzuwachs in einer jeweiligen Intervention entscheidend, welchem Typ sie angehören?

In Kapitel 5.1 wurde außerdem dargelegt, dass das Modellieren in den Naturwissenschaften essenziell ist und Schüler*innen oft eine naive Vorstellung von Modellen haben. Dies könnte u. a. daran liegen, dass in der Schule kaum Eigenschaften von Modellen thematisiert werden und wenig selbst modelliert wird. Ein Argument für den Einsatz von mathematischer Modellbildung wäre, wenn sich die Ansichten zu Modellen in der Physik durch das selbstständige Modellieren in positive Richtung verändern würden:

- Kann durch das aktive Modellieren eine Veränderung der Einstellungen gegenüber Modellen erreicht werden, ohne dass die Eigenschaften von Modellen aktiv thematisiert werden?

Da der Lernerfolg in der Mechanik, wie bereits dargelegt, stark von den vorhandenen Schülervorstellungen beeinflusst wird, betrifft die abschließende Forschungsfrage die Aktivierung von Schülervorstellungen:

- Gibt es bzgl. der Aktivierung von Schülervorstellungen Unterschiede zwischen den Methoden? In welcher Art beeinflusst diese Aktivierung das Konzeptverständnis?

Diese Forschungsfrage ist sowohl dem qualitativen als auch dem quantitativen Teil zugehörig.

7.2.2 Qualitativer Teil

Da noch sehr wenig darüber bekannt ist, wie Schüler*innen mit mathematischer Modellbildungssoftware und Videoanalysesoftware arbeiten, soll auch dies ein Fokus der vorliegenden Forschungsarbeit sein. Dabei soll es einerseits darum gehen, wie mit der Software gearbeitet wird und andererseits welche Prädiktoren den Lernerfolg beeinflussen. Zudem soll analysiert werden, welche Schwierigkeiten bei der Nutzung der beiden Programme entstehen. Die erste Forschungsfrage beschäftigt sich mit dem generellen Arbeitsverhalten der Schüler*innen:

- Wie gehen Schüler*innen beim Arbeiten mit dem Modellbildungsprogramm und dem Videoanalyseprogramm vor?

Eine bereits im vorigen Kapitel formulierte Hypothese ist, dass die Software und das damit verbundene Vorgehen im Bezug zu Theorie und Experiment unterschiedlich sind. Beiden unterliegt eine unterschiedliche Argumentationsrichtung. Fraglich ist, ob sich das auch in den Gesprächen der Proband*innen während der Arbeitsphase niederschlägt:

- Überträgt sich die Argumentationsrichtung der jeweiligen Methode auf die Arbeitsweise der Lernenden?

Unabhängig von der Argumentationsrichtung ist die übergeordnete Frage, mit welcher Methode man den grundlegenden Zusammenhang der Mechanik zwischen Beschleunigung, Masse und wirkenden Kräften besser lernt. Daher ist von Interesse, bei welchem Vorgehen mehr über diesen Zusammenhang gesprochen und diskutiert wird:

- Welche Variante regt mehr zu Gesprächen zum Zusammenhang zwischen Kräften und Bewegung an?

Da insbesondere im Bereich der mathematischen Modellbildung bereits von einigen Schwierigkeiten im Umgang mit der Software berichtet wurde, ist von Interesse, ob die heutige Software ebenfalls zu größeren Schwierigkeiten im Umgang führt. Zusätzlich soll beobachtet werden, welche Schülervorstellungen auftreten und ob sich diese in den Gruppen unterscheiden. Das würde dafürsprechen, dass gewisse Vorstellungen durch eine der beiden Varianten eher angeregt werden:

- Welche typischen Schwierigkeiten und Schülervorstellungen sind identifizierbar? Unterscheiden sich diese zwischen den beiden Vorgehensweisen?

Abschließend gibt es in Analogie zu der Forschungsfrage des quantitativen Teils die Frage, welche Prädiktoren grundsätzlich für den Lernerfolg entscheidend sind:

- Lassen sich beim Arbeiten Faktoren erkennen, die den Lernerfolg der Lernenden beeinflussen?

Dabei soll untersucht werden, wie zielgerichtet die Lernenden vorgehen, welche Unterschiede zwischen erfolgreichen und weniger erfolgreichen Schüler*innen bestehen und ob daraus Handlungsempfehlungen für den Einsatz einer solchen Software im Physikunterricht abgeleitet werden können.

8 Anlage und Durchführung der empirischen Studie

8.1 Vorstudie zur Software im Bereich der mathematischen Modellbildung und Videoanalyse

Als Vorstudie zu dieser Arbeit wurden im Rahmen einer wissenschaftlichen Hausarbeit zunächst Programme für den Einsatz in der Schule gesichtet und bewertet (Weber, 2017; Weber & Wilhelm, 2018). Die Vorstudie untersuchte die Kombination von mathematischer Modellbildung und Videoanalyse im Physikunterricht. Es wurden sowohl Modellbildungsprogramme als auch Videoanalyseprogramme analysiert und bewertet, mit denen Daten Modellbildung und Videoanalyse verglichen werden können. Von den Programmen, die in die Vorstudie einbezogen wurden („Tracker“, „Coach 6 Studio MV“, „measure dynamics“, „Newton-II“), sind manche in der Lage, sowohl mathematische Modellbildung als auch Videoanalyse durchzuführen, während andere nur eines von beiden beherrschen aber die jeweils anderen Daten importieren können. Bei der kriteriengeleiteten Bewertung schnitten die Programme „Newton-II“ für die Modellbildung und „measure dynamics“ für die Videoanalyse am besten ab. Bewertet wurden die Programme nach Bedienung, Genauigkeit, Schnelligkeit, Darstellungsmöglichkeiten, Einstellungs-möglichkeiten und Preis und Verfügbarkeit mit unterschiedlicher Gewichtung.

Wichtig war neben den Kriterien, die mit der Benutzung und Handhabung der Software zusammenhängen, auch lerntheoretische Ansätze, wie die Kontiguitätsprinzipien (siehe Kapitel 4.2). Die beiden für die hier vorgestellte empirische Studie relevanten Programme werden in der Folge im Hinblick auf verschiedene lerntheoretische Aspekte genauer vorgestellt.

8.1.1 „Newton-II“

„Newton-II“ ist ein gleichungsbasiertes Modellbildungsprogramm, welches für den Einsatz im Bereich der Mechanik gemacht wurde. Die Zusammenhänge zwischen den kinematischen Größen sind im Programm bereits implementiert und müssen nicht selbst eingegeben werden. Modelliert werden muss der Zusammenhang zwischen Kräften und der Beschleunigung. Außerdem können in dem Modellierungsfenster neben den Kräften beliebige weitere Größen modelliert und definiert werden. Die Anfangsbedingungen der Bewegung und die Masse des Objekts können in speziell dafür vorgesehenen Feldern eingetragen werden. Außerdem können die Schrittweite, die Modellierungsdauer, die Anzahl an Modellierungszeitpunkten und die Startzeit festgelegt werden. Die Software bietet weitere Optionen wie bedingte Variablen oder Schieberegler für beliebige Größen. Die

Darstellung der Modellierung erfolgt in Form verschiedener Diagramme, die manuell eingestellt werden können. Alles wichtige der Software befindet sich in einem Fenster, womit den zuvor beschriebenen Kontiguitätsprinzipien nachgekommen werden soll (Abb. 9).

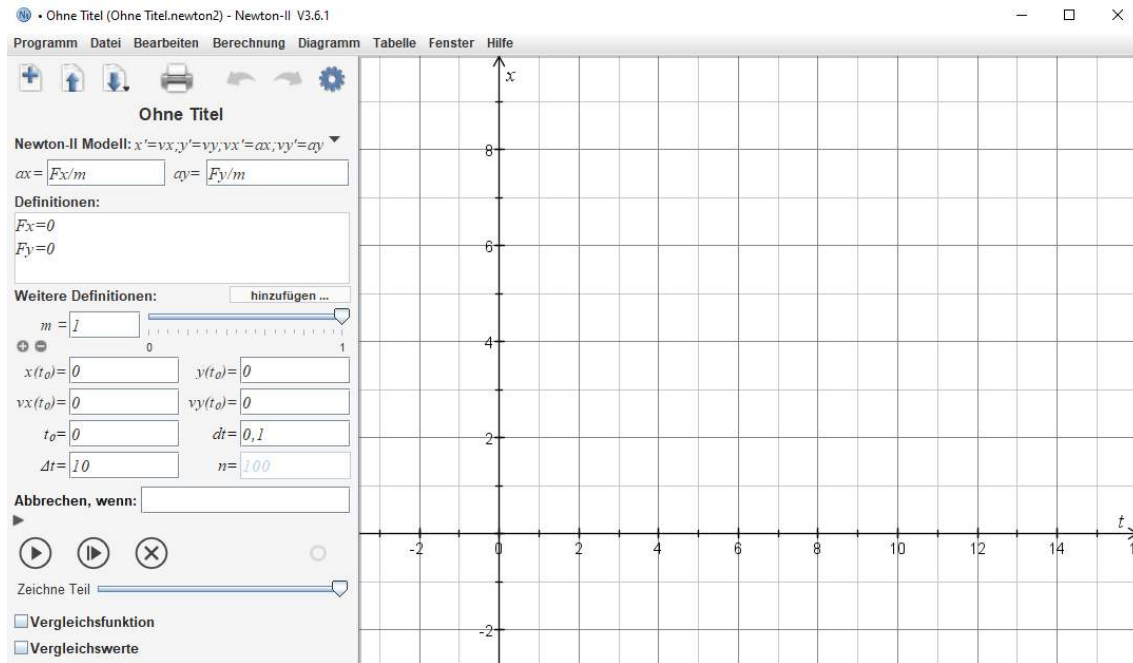


Abb. 9: Programmoberfläche "Newton-II" in der Version 3.6.1 in der Form für die Modellierung von zweidimensionalen Bewegungen.

Das Programm kann ein- bis dreidimensionale Modelle anfertigen. Je nach Anzahl der Dimensionen sind die Größen unterschiedlich benannt. Es kann zudem eine Abbruchbedingung festgelegt werden, die die Modellierung beendet (z. B. wenn das Objekt den Boden erreicht). Außerdem ist das Programm in der Lage, Messdaten direkt in das Diagramm einzufügen und neben der Modellierung anzuzeigen. In der Software ist das Standard-Runge-Kutta-Verfahren (vgl. Kapitel 5.4.2) voreingestellt. Alternativ kann auch das Euler-Verfahren, das Heun-Verfahren und das Runge-Kutta-Verfahren mit Schrittanpassung ausgewählt werden. „Newton-II“ ist leicht zu bedienen, läuft stabil und ist kostenlos für alle gängigen Betriebssysteme verfügbar. Dementsprechend wird es als geeignetes Programm für den Einsatz mathematischer Modellbildung in der Mechanik im Physikunterricht und für diese Studie gesehen.

8.1.2 „measure dynamics“

„measure dynamics“ ist ein Videoanalyseprogramm, welches auf Windows-Betriebssystemen läuft. Die Software benötigt ein Video, welches in unterschiedlichen Dateiformaten vorliegen kann, und ermittelt daraus mithilfe der automatischen Videoanalysefunktion Daten, die in vielfältiger Weise dargestellt werden können. Zunächst muss dazu eine Skalierung erfolgen. Außerdem kann das Koordinatensystem festgelegt werden. Danach kann die Videoanalyse erfolgen, die Algorithmen für Bewegungserkennung, Mustererkennung und Farberkennung anbietet. Eine Voruntersuchung garantiert in den meisten Fällen die Wahl des geeigneten Algorithmus, sodass die automatische Analyse gut funktioniert. Falls dies nicht der Fall ist, können die Einstellungen manuell geändert oder eine manuelle Analyse durchgeführt werden.

Für die Darstellung der Daten bietet die Software beliebige Diagramme, Stroboskopbilder und -videos und dynamisch ikonische Repräsentationen wie Geschwindigkeits- oder Beschleunigungspfeile. Die Software kommt den Design- und Kontiguitätsprinzipien nach (Abb. 10) und ist weitestgehend einfach und intuitiv zu bedienen, solange man nicht spezielle Wünsche an Einblendungen und dynamisch ikonische Repräsentationen hat. Hier können die vielen Einstellungsmöglichkeiten kompliziert werden.

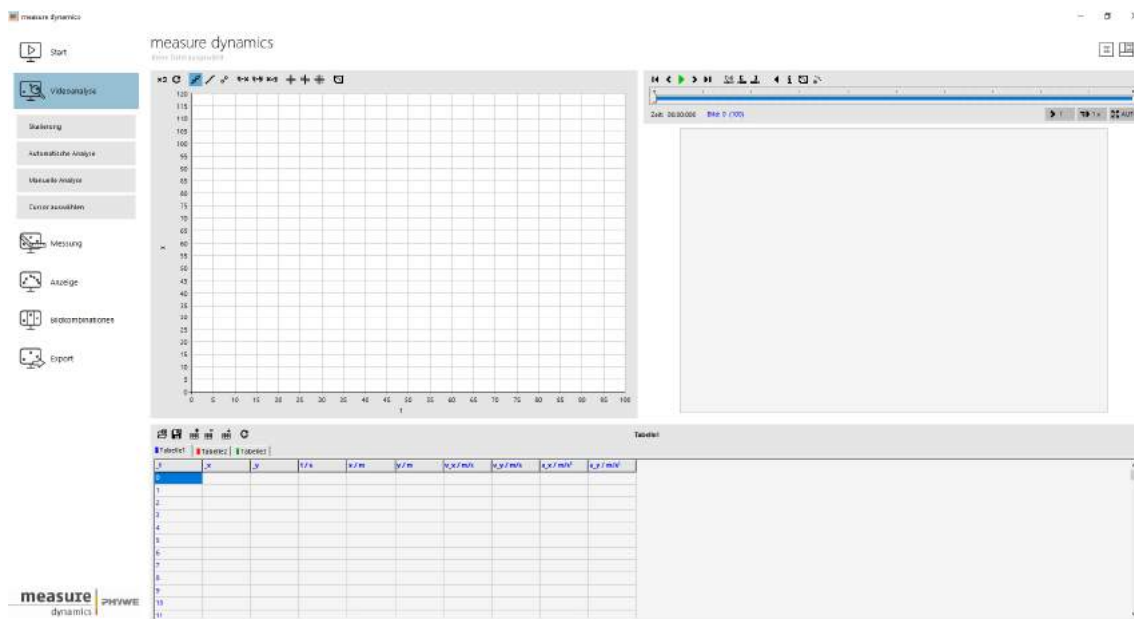


Abb. 10: Programmoberfläche von "measure dynamics" in der Version 1.5.

Auf der Benutzeroberfläche werden im Normalfall gleichzeitig das Video, das eingestellte Diagramm und die Messdaten angezeigt. Die Diagramme entstehen parallel zum ablaufenden Video, was didaktisch sinnvoll ist. In der Vorstudie wurden die genannten

Programme von den untersuchten am besten bewertet (Weber, 2017). Aufgrund dessen werden sie für die Interventionen dieser Studie genutzt.

8.2 Allgemeine Design-Entscheidungen auf Basis der bekannten Theorie

Im bisherigen Teil der Arbeit wurden relevante Aspekte der Theorie in den Bereichen Schülervorstellungen (Kapitel 3), Conceptual Change (Kapitel 3.4), Lernen mit neuen Medien (Kapitel 4), Mathematische Modellbildung (Kapitel 5.4) und Videoanalyse (Kapitel 6.2) und deren Software (Kapitel 8.1) vorgestellt. Diese bilden die Grundlage für die Entscheidungen, die für die Gestaltung der Interventionen getroffen wurden.

Dabei sind die Interventionen innerhalb der Studie auf die bekannten Schülervorstellungen ausgelegt und sollen vor allem häufig auftretende Schülervorstellungen wie „Es muss eine Kraft in Bewegungsrichtung wirken“ verringern. Auf Basis der in Kapitel 3.4 vorgestellten Konzeptwechselstrategien wird der Computer genutzt, um den Lernenden eine Rückmeldung zu geben und damit Unterschiede zwischen der physikalischen und der Alltagsvorstellung offen zu legen. Es wird darauf Wert gelegt, dass die wissenschaftlich erwünschte Theorie (Gesetze der Newtonschen Mechanik) bereits aus dem Schulunterricht bekannt ist, sodass es nicht zu früh im Lernprozess zu kognitiven Konflikten kommt. Im Bereich der mathematischen Modellbildung ist der Vergleich zwischen eigener Vorstellung und Realität möglicherweise intensiver, da die eigene Vorstellung hier in Form einer Gleichung expliziert werden muss. Auch bei der Videoanalyse wird aber durch das Arbeitsheft eine Formulierung von Hypothesen verlangt, bevor die entsprechende Auswertung der Daten erfolgt. Die Theorie über das Lernen mit neuen Medien führte zu der Entscheidung, dass die Lernenden selbstständig und in Zweiergruppen mit der Software arbeiten. Wichtig sind aber auch die regelmäßigen Besprechungsphasen im Plenum. Zudem wurde die Software und die anleitenden Materialien nach der Cognitive Load Theory und der Cognitive Theory of Multimedia Learning ausgesucht, sodass die kognitive Belastung durch Einhaltung der Kontiguitätsprinzipien möglichst gering sein soll. Im Folgenden wird die konkrete Ausgestaltung der Studie vorgestellt, die aus den allgemeinen und aus der Theorie abgeleiteten Design-Entscheidungen folgt.

8.3 Konkrete Ausgestaltung der Studie

Um die Forschungsfragen beantworten zu können, wurde eine quasi-experimentelle Studie im Prä-Post-Design durchgeführt. Diese untersuchte den Einfluss einer Intervention mit mathematischer Modellbildung bzw. Videoanalyse auf das Konzeptverständnis der Newtonschen Mechanik, aber auch auf affektive Merkmale der Proband*innen.

Dazu wurden Klassen der Einführungsphase der gymnasialen Oberstufe an hessischen Gymnasien in das Goethe-Schülerlabor Physik der Goethe-Universität Frankfurt eingeladen, die dann randomisiert klassenweise einer der beiden Interventionen zugeordnet wurden. Die klassenweise Zuordnung ist für eine Gewährleistung der internen Validität ungünstig, lies sich durch die organisatorischen Rahmenbedingungen aber nicht vermeiden. Bei einer experimentellen Studie und einer damit einhergehenden randomisierten Zuordnung der Studienteilnehmenden geht man davon aus, dass sich die Störvariablen gleichmäßig in den Gruppen darstellen. Dies ist bei der hier gewählten Zuteilung nicht a priori zu gewährleisten. Da eine vollständig randomisierte Zuteilung der Schüler*innen durch die gegebenen Randbedingungen nicht möglich war, wurden verschiedene Kontrollvariablen erhoben, die die Gleichheit der beiden Gruppen kontrollieren sollten (siehe Kapitel 9.3.1).

Die bei Gruppenzuordnung von ganzen Klassen zu erwartenden Lehrer- und Klasseneffekte sind in diesem Setting als gering einzuschätzen. Die Lehrkraft ist nicht an der Durchführung der Intervention beteiligt. Der Großteil der Intervention findet in Partnerarbeit statt, sodass klassenweite Sozialeffekte gering sind. Dennoch ist es unter diesen Umständen wichtig, die Intervention nach Möglichkeit mit einer großen Anzahl an Klassen durchzuführen.

Die Lernenden haben die Newtonschen Gesetze vor der Intervention bereits in der Schule kennengelernt. Diese wurden dann in der Intervention vertieft. Um den Lernzuwachs messen zu können, wurden ein Vor- und ein Nachtest konzipiert, wobei der Vortest in der letzten Physikstunde vor dem Besuch des Schülerlabors in der Schule durchgeführt wurde und der Nachtest sich direkt an die Intervention anschloss. Es wäre nicht ausreichend, das Konzeptverständnis der Proband*innen lediglich nach der Intervention zu überprüfen, da so nicht ermittelt werden kann, welcher Anteil des Nachtestergebnisses durch das Treatment zu erklären ist.

8.3.1 Stichprobe

Die geplante Stichprobe bestand aus ca. 630 Schüler*innen aus 33 Schulklassen, mit denen ein Termin vereinbart war. Durch die Schulschließung im Zuge der Corona-Pandemie im Erhebungszeitraum konnten dann leider nur insgesamt 312 Schüler*innen aus 21 Kursen an der Intervention teilnehmen. Von diesen lagen für $N = 274$ Proband*innen vollständige Datensätze mit nur wenigen fehlenden Werten vor. Die Schüler*innen stammten aus dem ersten Schuljahr der Oberstufe an hessischen Gymnasien (zehnte Klassenstufen nach G8 bzw. elfte Klassenstufen nach G9). Ausnahme bildeten dabei zwei Kurse, die zu

Beginn der Qualifikationsphase an den Interventionen teilnehmen, um das Gelernte aus dem letzten Schuljahr zu wiederholen und zu vertiefen. In Hessen ist es in der sogenannten Einführungsphase in die Oberstufe (E-Phase) verpflichtend, Physik zu belegen. Am Ende des Schuljahrs kann das Fach abgewählt werden. Es befinden sich somit einige Schüler*innen in den Kursen, die sich bereits gegen Physik entschieden haben und bei denen das Schuljahr ihr letztes in Physik sein wird, was sich negativ auf die Motivation auswirken kann. Gleichzeitig waren von den 21 Kursen sieben Kurse sogenannte Leistungsvorkurse oder auch Orientierungskurse, in denen sich die Lernenden bewusst für Physik entschieden haben und voraussichtlich bis zum Abitur einen Leistungskurs in Physik belegen. Die Stichprobe ist also, was das Vorwissen, das Interesse an Physik und andere affektive Merkmale angeht, als repräsentativ für hessische Gymnasialkurse anzusehen.

8.3.2 Interventionen

Die durchgeführten Interventionen waren darauf ausgelegt, den Kern der Newtonschen Dynamik in verschiedenen Situationen zu verdeutlichen und dabei die in den vergangenen Kapiteln hergeleiteten Design-Prinzipien zu berücksichtigen. Dabei wurde immer mit einem realen Experiment gearbeitet, was den Ausgangspunkt für eine genauere Betrachtung der Bewegung mit dem Computer darstellte. Die Experimente wurden so ausgewählt, dass verschiedene Arten von Bewegungen besprochen wurden, die außerdem typische Schülerfehlvorstellungen adressieren.

Die Interventionen waren auf 3,5 Stunden ausgelegt und es wurden vier Experimente besprochen. Zu Beginn wurden den Schüler*innen zunächst der allgemeine Ablauf erklärt, wonach kurz die wichtigsten Größen (Kraft, Beschleunigung, Geschwindigkeit, Ort) und deren Zusammenhänge wiederholt wurden, was den Lernenden durch den Unterricht zuvor aber bereits bekannt war.

Es wurde keine theoretische Einführung in das Programm gegeben. Der erste Versuch (die Fallbewegung eines Fallkegels mit Luftreibung) wurde gemeinsam mit den Schüler*innen in einem lehrergelenkten Unterricht modelliert bzw. analysiert, wobei dies immer dieselbe Lehrperson war. Diese hat dabei die Funktionsweisen des Programms über einen Beamer gezeigt, während die Lerngruppe das Gezeigte in Partnerarbeit am Laptop nachmachen konnte. Dabei wurde ein lehrergeleitetes Unterrichtsgespräch geführt, was die physikalischen Zusammenhänge verdeutlichen sollte. In den weiteren Versuchen, die durchgeführt wurden, konnten die Schüler*innen dann selbstständig in Zweiergruppen mit dem Programm arbeiten. Es wurde dabei immer gemeinsam im Plenum das

Experiment durchgeführt und beschrieben, woran sich die Arbeitsphase in Gruppen anschloss. Am Ende der Arbeitsphase wurde eine Besprechung des Herausgefundenen im Plenum durchgeführt. Als Anleitung während der Arbeitsphase diente ein Arbeitsheft, in dem als Zusammenfassung für jedes Experiment immer Tabellen enthalten waren, in denen bei den relevanten Größen angekreuzt werden musste, ob diese sich während der Bewegung ändern, dauerhaft null oder konstant und von null verschieden sind. Dies sollte verdeutlichen, dass die Zusammenhänge zwischen Kräften und den kinematischen Größen (Grundgleichung der Mechanik und kinematische Kette) unabhängig vom Versuch immer gleich sind und betonen, dass die Beschleunigung eines Objekts proportional zu der Summe aller wirkenden Kräfte ist. Die Tabellen waren dabei analog zur Argumentationsrichtung der zugrundeliegenden Methode sortiert: Bei der mathematischen Modellbildung wurde zuerst nach speziellen Kräften, dann nach der Summe aller Kräfte, nach der Beschleunigung und der Geschwindigkeit gefragt, während dies bei der Videoanalyse in umgekehrter Richtung bearbeitet wurde. Während der Arbeitsphasen wurde von der Lehrperson und einem Betreuer bzw. einer Betreuerin bei Schwierigkeiten Hilfe geleistet.

Die vier durchgeführten Experimente waren (1) eine geradlinige Fallbewegung mit Luftreibung, (2) eine geradlinige, abschnittsweise beschleunigte Bewegung, (3) ein schiefer Wurf mit und ohne relevanten Einfluss der Luftreibung und (4) eine Kreisbewegung mit Ausschalten der Zentripetalkraft. Die beiden Interventionen unterscheiden sich ausschließlich in der Nutzung der Software und sind sonst identisch aufgebaut.

Versuch 1: Fallbewegung mit Luftreibung

Die Fallbewegung eines Fallkegels aus Papier wurde als Einstieg gewählt, bei dem sich die Proband*innen gleichzeitig in das Programm einarbeiten. Der Fallkegel fällt aus der Ruhe in einer bestimmten Höhe zu Boden (Abb. 11).

Während der Bewegung wirken eine konstante Gravitationskraft nach unten und eine im Betrag zunehmende Luftreibungskraft nach oben auf den Kegel, die sich am Ende der Bewegung ausgleichen und zu einer konstanten Geschwindigkeit führen. Dabei wird betont, dass ein Kräftegleichgewicht zu einer konstanten Geschwindigkeit führt und die Gesamtkraft nicht in Bewegungsrichtung zeigen muss, damit sich ein Objekt bewegen kann. Es soll gezeigt werden, dass die Summe aller Kräfte proportional zur Beschleunigung und nicht zur Geschwindigkeit des Objekts ist. Bei dem Versuch soll zunächst die Bewegung des Kegels von den Schüler*innen beschrieben werden, woraufhin Begriffe wie „geradlinig“ fallen. Dabei kann sowohl geäußert werden, dass der Kegel beschleunigt, als auch, dass er sich mit einer konstanten Geschwindigkeit bewegt. Die Anmerkungen werden

dabei gesammelt und zunächst nicht kommentiert. Anders als bei den folgenden Versuchen wird im Plenum auch bereits darüber gesprochen, welche Kräfte die Bewegung verursachen könnten und in welche Richtung diese zeigen. Nach der Besprechung starten die Schüler*innen ihre Laptops und es wird gemeinsam modelliert bzw. analysiert.

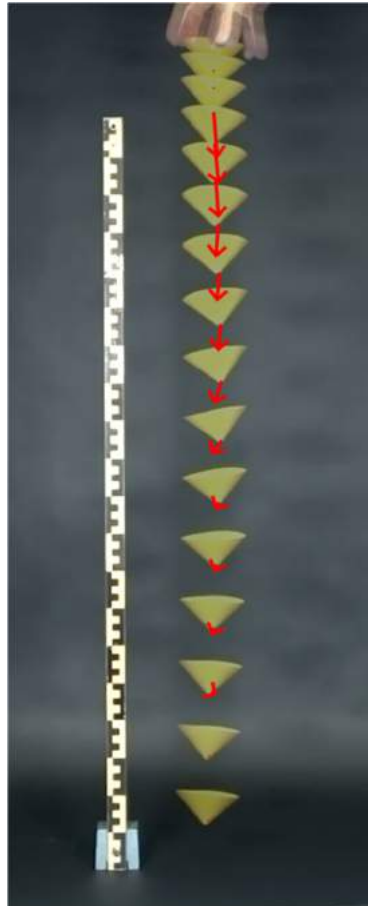


Abb. 11: Stroboskopbild der Fallbewegung mit Beschleunigungspfeilen in Rot.

Beim Modellieren wird dabei auf die Äußerungen aus der Lerngruppe reagiert und somit nur das modelliert, was von den Schüler*innen vorgeschlagen wird. Es ist dabei wichtig, dass alle Hypothesen ausprobiert werden können und nicht von der Lehrkraft bewertet werden. Es wird herausgearbeitet, dass es sinnvoll ist, zunächst ohne Reibungseinflüsse zu modellieren und daraufhin eine erste Prüfung des Modells anhand des entstehenden Diagramms zu tätigen. Sobald der dort sichtbare Verlauf plausibel erscheint, werden die Messdaten des Experiments zum Vergleich herangezogen. Dann kann die Modellierung der Reibung erfolgen.

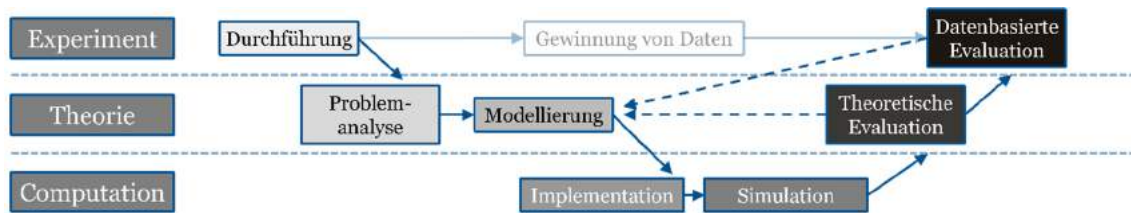


Abb. 12: Ablauf der mathematischen Modellbildung im Wechselspiel zwischen Experiment, Theorie und Computation. Die Schattierung gibt die Reihenfolge wieder (von hell nach dunkel).

Der in Abb. 12 dargestellte Verlauf ist ein idealtypischer Ablauf für die Modellierung in der Intervention. Zunächst wird das Experiment durchgeführt, worauf eine Problemanalyse folgt. Nach der Problemanalyse erfolgt die Modellierung. Wenn die Hypothesen mathematisiert wurden, werden sie in die Software implementiert und das Modell daraufhin simuliert. Anschließend sollte zunächst eine erste theoretische bzw. diagrammbasierte Evaluation stattfinden, in der überprüft wird, ob der berechnete Verlauf mit dem erwarteten Verlauf übereinstimmt. Falls dies nicht der Fall ist, wird die Modellierung überarbeitet. Falls das Modell der ersten Evaluation standhält, werden die bereits vorbereiteten Messdaten hinzugezogen und das Modell mit diesen Daten verglichen. Es wird so oft die Modellierung verbessert und erneut evaluiert, bis die Proband*innen mit ihrem Modell zufrieden sind. Diese hier eingeübte Reihenfolge sollte in den kommenden Stationen von den Schüler*innen übernommen werden.

Da die meisten Lerngruppen keine Erfahrungen mit der Beschreibung von Luftreibung haben, wird hier gemeinsam diskutiert und ausprobiert, bis man einen sinnvollen mathematischen Ausdruck für die Luftreibung gefunden hat. Die Einflussgrößen Luftdichte, Querschnittsfläche und Luftwiderstandsbeiwert werden nicht explizit modelliert, sondern zu einer Konstante zusammengefasst (Abb. 13). Wichtig ist vielmehr die Abhängigkeit von der Geschwindigkeit. Nach dem gemeinsamen Modellieren haben die Schüler*innen Zeit, das Besprochene in ihrem Arbeitsheft zu vermerken, in dem die dort enthaltenen Fragen beantwortet werden. Nach dieser Arbeitsphase in Partnerarbeit wird eine Abschlussbesprechung im Plenum gemacht.

Die Videoanalyse läuft so ab, dass ebenfalls gemeinsam das Video analysiert wird. Dabei werden die Funktionen des Programms erklärt und die verschiedenen Darstellungen (Pfeile, Diagramme) interpretiert und besprochen. Der Ablauf in der Videoanalyse unterscheidet sich von der Modellbildung, da die Modellierung, Implementation und Simulation entfallen, dafür die Daten aber selbst von den Schüler*innen generiert werden (Abb. 14).

Abb. 13: Beispielmmodellierung des Fallkegels in „Newton-II“.

Nach der anfänglichen Durchführung des Experiments schließt sich eine Problemanalyse an. Danach werden Hypothesen über die Bewegung und bestimmte kinematische Größen aufgestellt. Parallel dazu werden mithilfe der Software die Daten erhoben. Anschließend findet durch eine Auswertung der Daten eine Evaluation der Hypothesen statt. Die Hypothesen können daraufhin verbessert und erneut mit den Daten verglichen werden. Dieser Prozess wird für verschiedene Darstellungen durchgeführt. Die Hypothesen werden hier allerdings nicht quantifiziert, sodass ein Abweichen der Hypothesen von der Realität möglicherweise weniger wahrgenommen wird.

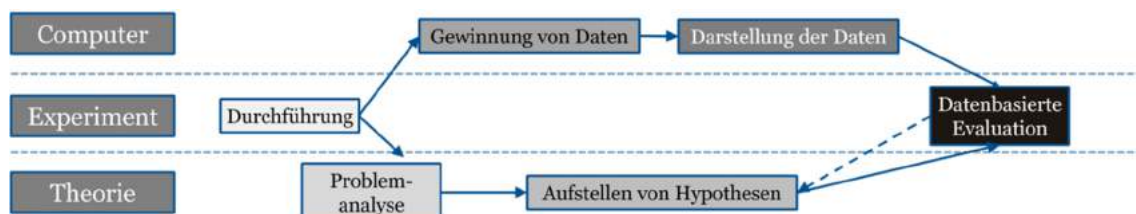


Abb. 14: Ablauf der Videoanalyse im Wechselspiel zwischen Experiment und Theorie. Die Schattierung gibt die Reihenfolge wieder.

Nachdem die Lerngruppe alle wichtigen Funktionen des Programms kennengelernt hat, schließt sich auch hier eine Arbeitsphase an, in der die Proband*innen alles Gelernte selbst noch einmal ausprobieren können, das Besprochene im Arbeitsheft vermerken und die Fragen im Heft beantworten. Danach schließt sich ebenfalls eine Abschlussbesprechung an. Im Anschluss an die Besprechung schließt sich unmittelbar der nächste Versuch an.

Versuch 2: Abschnittsweise beschleunigter Wagen

Der abschnittsweise beschleunigte Wagen (Abb. 15) hat einige Ähnlichkeiten zum vorigen Versuch und bietet sich somit für den ersten Versuch an, den die Schüler*innen selbstständig bearbeiten. Es handelt sich ebenfalls um eine geradlinige Bewegung, die zu Beginn beschleunigt ist und am Ende eine (nahezu) konstante Geschwindigkeit besitzt. Der Übergang zwischen diesen beiden Phasen ist aber nicht mehr stetig, sondern sprunghaft. Ein Gewicht, welches in der Luft hängt, ist über ein Seil, welches über eine Umlenkrolle geleitet wird, mit einem Wagen verbunden. Der Wagen wird an den Anfang der Fahrbahn gezogen und losgelassen. Dabei wird er zunächst von der Gravitationskraft, die auf das Gewicht wirkt und über das Seil auf den Wagen einwirkt, beschleunigt. Das Gewicht kommt nach einer bestimmten Distanz auf dem Boden auf, wonach der Wagen aber noch etwa einen Meter weiterrollt. Im letzten Abschnitt bewegt er sich mit einer (nahezu) konstanten Geschwindigkeit, da er (nahezu) kräftefrei ist. In Realität sieht man, dass sich die Geschwindigkeit aufgrund der wirkenden Reibungskräfte leicht ändert, welche hier diskutiert werden. Es soll einerseits deutlich werden, dass es bei Kräftefreiheit zu einer konstanten Geschwindigkeit kommt und andererseits, dass es in Realität nahezu nie zu einer solchen Kräftefreiheit kommt.

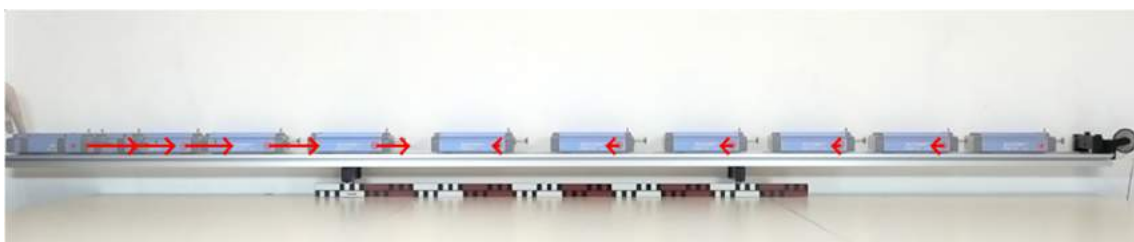


Abb. 15: Stroboskopbild des Wagens mit Beschleunigungspeilen in Rot.

Nach Beschreibung der Bewegung haben die Schüler*innen nun selbstständig die Möglichkeit, mit der Software die Bewegung zu modellieren bzw. zu analysieren und die Fragen im Arbeitsheft zu beantworten. Am Ende findet wieder eine Abschlussbesprechung statt.

Versuch 3: Schiefer Wurf

Das dritte Experiment ist ein schiefer Wurf (Abb. 16). Der schiefe Wurf war lange ein klassisches Experiment an hessischen Gymnasien, wird mittlerweile aber häufig nur noch in Leistungskursen besprochen, in diesem Fall häufig ohne Reibung.

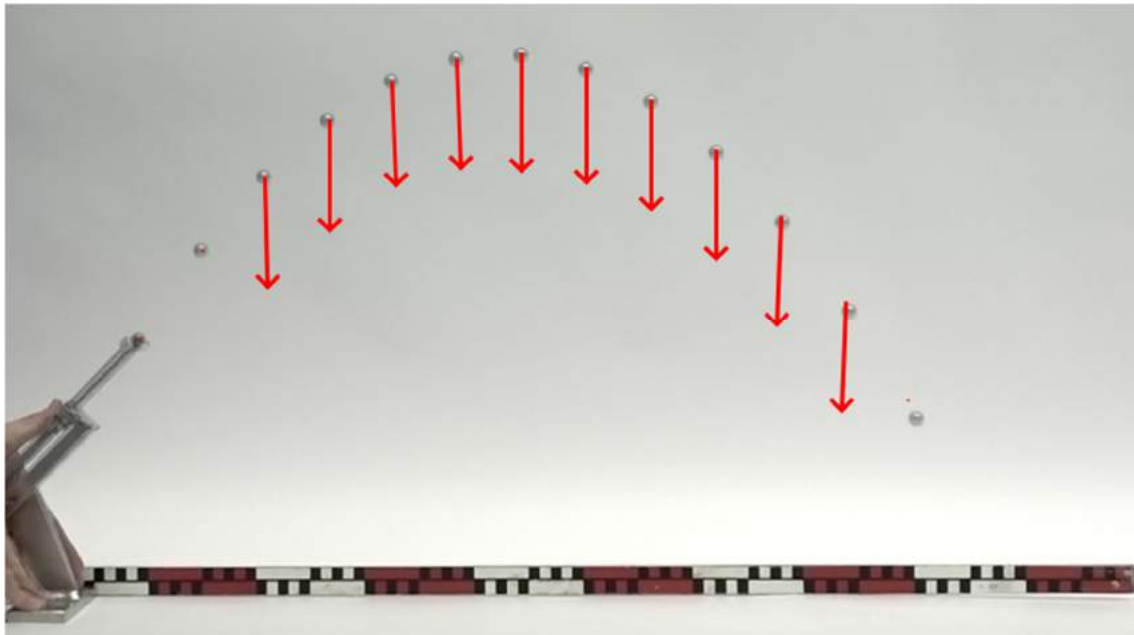


Abb. 16: Stroboskopbild der Wurfbewegung einer Stahlkugel mit Beschleunigungsvektoren in Rot.

In den hier beschriebenen Interventionen wird nicht die Form der Bewegung (Parabel) in den Vordergrund gestellt, sondern die Zusammenhänge zwischen wirkenden Kräften und den kinematischen Größen, die dazu führen, dass sich die Kugel in einem Bogen bewegt. Auch wird in der Schule Reibung meist vernachlässigt, ohne dass die Schüler*innen die Möglichkeit haben zu überprüfen, ob diese Vernachlässigung gerechtfertigt ist. Deswegen werden hier mit einem Katapult eine Stahlkugel und eine Styroporkugel abgeschossen. Der Vergleich mit Messdaten zeigt, dass die Stahlkugel nicht merklich vom Luftwiderstand beeinflusst wird, während dies bei der Styroporkugel sehr wohl der Fall ist (Abb. 17).

Bei der Stahlkugel spielen Reibungskräfte keine Rolle, sodass die Kugel konstant nach unten beschleunigt wird. Die Geschwindigkeit in horizontale Richtung ist daher konstant, was das Unabhängigkeitsprinzip verdeutlicht. Bei der Styroporkugel kommt eine Kraft entgegen der Bewegungsrichtung hinzu, die dazu führt, dass sich die Geschwindigkeit in beide Koordinatenrichtungen ändert. Da die Modellierung der Reibung hier kompliziert

ist, wurde sie als Zusatzaufgabe für schnellere und leistungstärkere Schüler*innen konzipiert und dient somit auch zur Binnendifferenzierung.

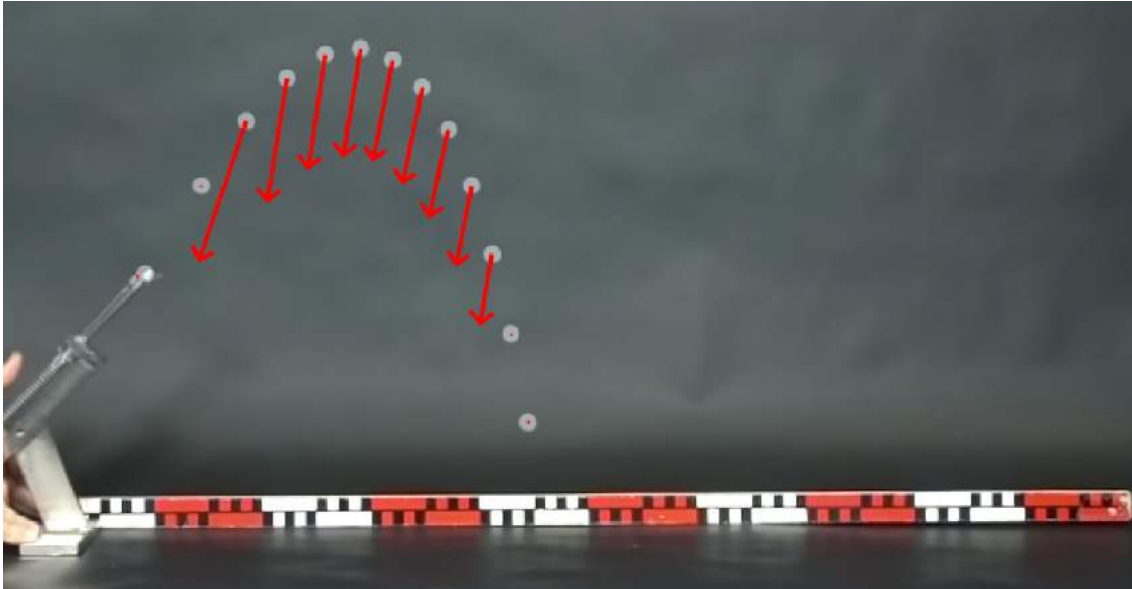


Abb. 17: Stroboskopbild der Wurfbewegung einer Styroporkugel mit Beschleunigungspfeilen in Rot.

Versuch 4: Kreisbewegung

Die Besprechung der Kreisbewegung (Abb. 18) ist in hessischen Lehr- bzw. Rahmenplänen vorgesehen. Studien zeigen, dass die Kreisbewegung mit konstantem Tempo bei vielen Lernenden Fehlvorstellungen hervorruft. Ein Großteil geht bei einer solchen Bewegung davon aus, dass die Beschleunigung null ist¹³. Das liegt daran, dass die Beschleunigung als Schneller- oder Langsamerwerden konzeptualisiert wird (siehe Kapitel 3). Es muss betont werden, dass die Geschwindigkeit aus Tempo und Richtung besteht, wodurch eine Richtungsänderung auch eine Geschwindigkeitsänderung und damit eine Beschleunigung darstellt. Eine weitere Herausforderung stellt die Zentrifugalkraft dar. Erfahrungsgemäß sind die Fehlvorstellungen nach dem Thematisieren der Kreisbewegung weiterhin vorhanden.

¹³ In einer Studie an hessischen Gymnasien konnten Wilhelm und Gemici (2017) zeigen, dass fast niemand der 346 Schüler*innen am Ende der E-Phase die Beschleunigung bei der Kreisbewegung korrekt angeben konnte.

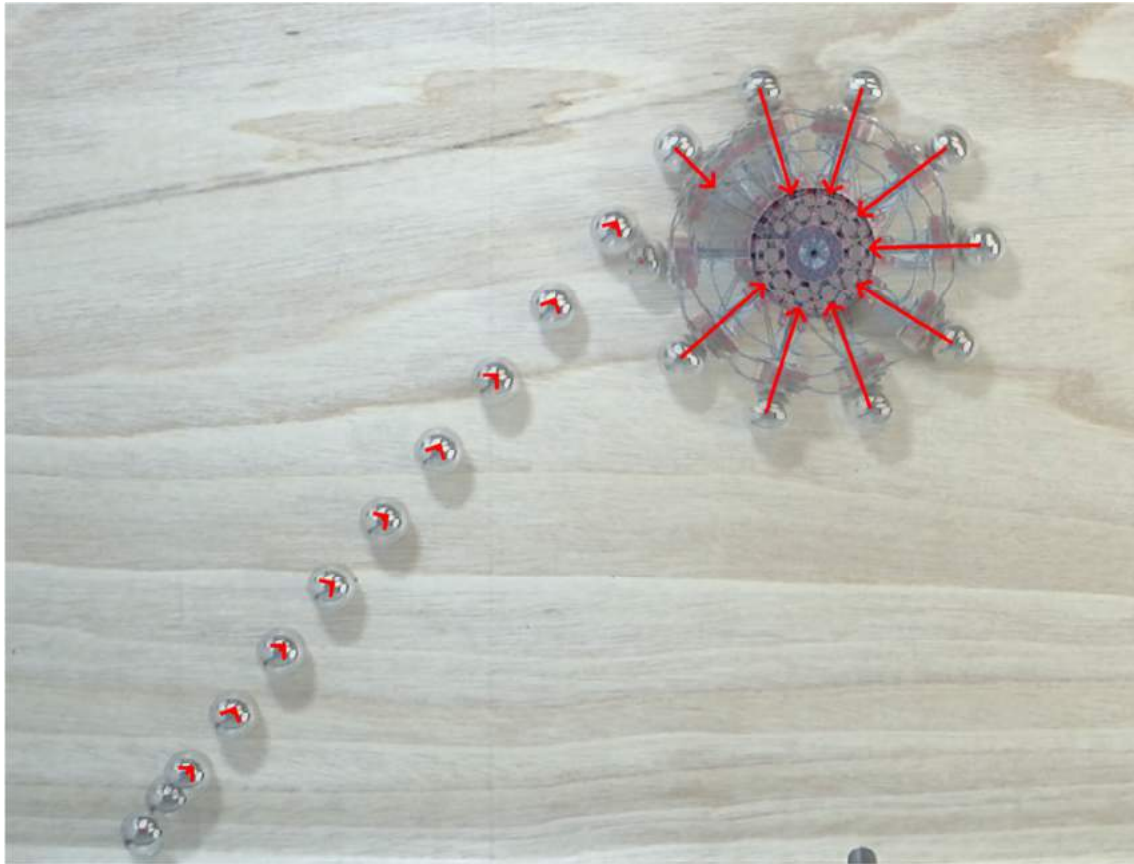


Abb. 18: Stroboskopbild der Kreisbewegung mit Beschleunigungspfeilen in Rot.

In der hier thematisierten Variante wird eine Stahlkugel von einem Elektromagneten angezogen, der mit einem Elektromotor verbunden ist und eine Kreisbewegung ausführen kann. Der Elektromagnet übt die Zentripetalkraft aus. Nach ein paar Umdrehungen wird der Magnet ausgeschaltet. Davor werden Vermutungen über das Weiterrollen der Kugel eingeholt, wobei viele Schüler*innen von einer bogenartigen Bewegung ausgehen. Durch das Wegfallen der Zentripetalkraft muss die Kugel gradeausrollen, da außer Reibung keine weiteren Kräfte auf die Kugel wirken. Auch hier schließt sich nach der Durchführung des Experiments eine Arbeitsphase an. Die Modellierung der Kreisbewegung ist schwierig, wozu einige Hilfen angeboten werden. Für diejenigen, die auch mit den Hilfen nicht zu einer erfolgreichen Modellierung gelangen, wird die fertige Modellierung angeboten, die dann genutzt werden kann, um die Fragen zu beantworten.

8.4 Testinstrument

8.4.1 Vorhandene Tests zur Mechanik

Es existieren verschiedene, überregional bekannte Tests zur Messung des Konzeptverständnisses der Newtonschen Mechanik. Hervorzuheben ist der Force Concept Inventory-Test (FCI-Test) von Hestenes et al. (1992), der das Konzeptverständnis von Kraft überprüft und in vielen Studien zur Mechanik bereits eingesetzt wurde (Hake, 1998). Die Autoren sind der Meinung, dass die restlichen Inhalte der Mechanik ohne dieses Konzeptverständnis von Kräften „nutzlos“ sind und es bedenklich sein sollte, wenn Lernende, die im FCI-Test schlecht abschließen, gute Punktzahlen in herkömmlichen Tests erreichen und gute Noten erhalten (Hestenes et al., 1992). Der Ankreuztest besteht aus Items zu verschiedenen Teilgebieten der Mechanik. Dabei sind die Antwortmöglichkeiten so konzipiert, dass die Disktraktoren auf typischen Alltagsvorstellungen basieren und dadurch für Schüler*innen besonders attraktiv sind. Das führt dazu, dass man in vielen Fällen die hinter einer Antwort stehende Vorstellung identifizieren kann, das macht den Test für Schüler*innen aber auch sehr schwierig. Die Fragen können ohne Kenntnisse von Gleichungen oder anderen mathematischen Fähigkeiten beantwortet werden, was den Test vielseitig einsetzbar macht. Die Bekanntheit des FCI-Tests macht ihn für den Einsatz in Forschungsstudien attraktiv, da eine hohe Vergleichbarkeit mit bereits getätigten Studien besteht. Trotz seiner weiten Verbreitung weist der FCI-Test aber einige Schwächen auf. Zunächst zeigt sich in verschiedenen Studien, dass Schüler*innen in dem Test auch nach dem Unterricht oft sehr schlechte Ergebnisse erzielen (Wilhelm, 2005b), wodurch keine optimale Passung zu der hier untersuchten Stichprobe besteht. Das schlechte Abschneiden ist dabei nicht verwunderlich, da der Test ursprünglich für die Universität entwickelt wurde und auch dort eher niedrige Punktzahlen erreicht wurden (z. B. Hake, 1998). Das liegt, wie bereits diskutiert, vor allem an den hartnäckigen Fehlvorstellungen in der Mechanik und der Attraktivität der Distraktoren im Test. Die Reliabilität des FCI-Tests wurde in vielen Studien untersucht und erwies sich als gut. Der Test misst reliabel und valide das gesamte Konzeptverständnis der Newtonschen Mechanik (Savinainen & Scott, 2002). Obwohl die Autoren des Tests aber verschiedene Alltagsvorstellungen aus verschiedenen Bereichen der Newtonschen Mechanik unterscheiden („Kinematik“, „Impetus“, „Aktive Kraft“, „Actio/Reactio-Paare“, „Verknüpfung von Einflüssen“, „Andere Einflüsse auf die Bewegung“) und die Items in sechs Inhaltsbereiche (Kinematik, 1. Newtonsches Gesetz, 2. Newtonsches Gesetz, 3. Newtonsches Gesetz, Superposition und Kraftarten) einteilen (Hestenes et al., 1992), ist es nicht gelungen, inhaltlich interpretierbare und reliable Subskalen empirisch nachzuweisen. Nach Huffman und Heller (1995) misst der FCI-Test einzelne Wissensteile von Studierenden, die nicht zwangsläufig ein

kohärentes Kraftkonzept ergeben. Dies erschwert eine differenzierte Betrachtung. Hestenes und Halloun (1995) geben an, dass die Gesamtpunktzahl im Test ein valides Maß für das Verständnis des Newtonschen Kraftbegriffs ist. Die fehlende Clusterung in die Inhaltsbereiche besteht laut den Autoren durch falsch-positive Antworten und würde sich bei Newtonschen Denkern verbessern (Hestenes & Halloun, 1995). Diese falsch positiven Ergebnisse (Hestenes et al., 1992) könnten aber z. B. durch ein zweistufiges Fragensdesign verbessert werden.

Weiterhin bekannt ist auch der FMCE-Test (Force and Motion Concept Evaluation), der ebenfalls ein Standardtest aus den USA ist (Thornton & Sokoloff, 1998). Die Ergebnisse des FMCE korrelieren stark mit denen des FCI (Thornton et al., 2009). Der FMCE testet aber nur eindimensionale Kräfte und Bewegungen und hat damit einen engeren Fokus als der FCI, welcher eher für das allgemeine Verständnis der Newtonschen Mechanik herangezogen wird.

8.4.2 Entwicklung des Testinstruments

Testteil zur Mechanik

Die genannten Schwächen der obigen Tests führten dazu, dass in dieser Studie ein eigener Test erstellt und pilotiert wurde. Ausgangspunkt für die Entwicklung des Testinstruments waren bekannte Tests, wie der FCI- (Hestenes et al., 1992) oder der FMCE-Test (Thornton & Sokoloff, 1998). Die vorhandenen Fragen wurden dabei zum Teil direkt übernommen und zum Teil an die Zielsetzungen dieser Studie angepasst. Bei den Anpassungen handelte es sich um Änderungen der Formulierung, der Antwortmöglichkeiten oder einem Hinzufügen einer Begründung, wodurch die hinter einer Antwort stehende Schülervorstellung besser identifiziert werden kann. Außerdem wurden eigene Items entwickelt, die als leicht zu charakterisieren sind, um zu ermöglichen, dass auch die Fähigkeiten schwacher Schüler*innen mit dem Test aufzulösen sind und es nicht zu starken Bodeneffekten kommt.

Durch die gezielte Auswahl und Erstellung von Items sollten verschiedene Bereiche der Dynamik abgedeckt werden, die im Zentrum der Intervention stehen. Zusätzlich wurden auch rein kinematische Items verwendet, um auch in diesem Bereich einen Vergleich zwischen den beiden Interventionen zu haben. Dabei wurde sich auf Beschleunigungsdiagramme beschränkt, da diese innerhalb der Kinematik für Schüler*innen erfahrungsgemäß besonders herausfordernd sind (Struck & Yerrick, 2010; Wilhelm, 2005a). Aufgrund

der begrenzten Testzeit und geistigen Kapazitäten der Proband*innen musste auf die Testung weiterer Bereiche aus der Kinematik verzichtet werden.

Schlussendlich wurden für den endgültigen Test, der 19 Items in dem Testteil zur Mechanik hat, neun aus dem FCI-Test (Hestenes et al., 1992) in der Übersetzung von Gerdes und Schecker (1999) mit Bebilderung von Wilhelm (2005b) übernommen. Davon wurden drei Items unverändert gelassen, vier Items wurden geändert und zwei durch eine zweite Stufe (eine Begründung der Antwort) erweitert. Weiterhin sind vier Items aus dem FMCE-Test (Thornton & Sokoloff, 1998) in der Version von Wilhelm (2005a). Drei Items sind komplett selbst erstellt und zwei von Wilhelm (2007) nach Flores et al. (2004). Ein weiteres ist aus Warren (1979) übernommen. Der gesamte Test befindet sich im Anhang.

Testteil zu Vorstellungen über Modelle in der Physik

Eine weitere Forschungsfrage der Studie ist, ob eine Intervention mit mathematischer Modellbildung zu einer geänderten Ansicht über die Bedeutung von Modellen in der Physik führen kann, ohne dass die Eigenschaften von Modellen explizit thematisiert werden. Dazu wurde ein eigenes Set an Items entwickelt, das auf den von der Physikdidaktik postulierten Eigenschaften von Modellen in der Physik basiert (siehe Kapitel 5). Ziel war es zunächst, einen multidimensionalen Fragebogen zu erhalten, in dem verschiedene Eigenschaften von Modellen, wie „Modelle als Darstellung bekannten Wissens“, „Erkenntnisgewinn durch Modelle“, „Exaktheit von Modellen“ und „Zusammenspiel von Modell und Experiment“ enthalten sind. Durch die Erkenntnisse der Pilotierung (siehe Kapitel 8.4.3) wurde dieser Teil des Fragebogens zu einem eindimensionalen Fragebogen abgeändert, der allgemein Fragen zu den Eigenschaften von Modellen erhält. Dadurch lässt sich nicht feststellen, ob sich nur die Ansichten zu einer bestimmten Eigenschaft von Modellen geändert haben, aber ob die Einstellungen zu Modellen durch die Intervention insgesamt beeinflusst wurden. Die Items wurden alle mit einer fünfstufigen Likert-Skala konzipiert.

Testteil zu affektiven Merkmalen

Die in der Studie erhobenen affektiven Merkmale dienen einerseits zur Kontrolle, dass sich die beiden Gruppen in keiner entscheidenden Variable systematisch unterscheiden. Andererseits wurde auch das Interesse am Experimentieren und das Interesse an theoretischen Überlegungen erhoben, um damit den Einfluss dieser Variablen auf den

Lernzuwachs in den unterschiedlichen Gruppen messen zu können¹⁴. Als Kontrollvariablen wurden das fachspezifische Selbstkonzept, das Fachinteresse an der Physik und das Interesse an Computern erhoben. Dabei wurden die Items z. T. aus bereits bestehenden und etablierten Fragebögen übernommen und nochmals pilotiert und teilweise selbst erstellt und pilotiert.

Im zweiten Teil geht es darum, mit verschiedenen Items herauszufinden, ob es bei Schüler*innen unterschiedliche Typen hinsichtlich Interesse an Theorie und Experiment gibt. Die Physik unterteilt sich klassisch in Theoretische Physik und Experimentalphysik, wobei diese Unterscheidung im Physikunterricht im Normalfall keine Rolle spielt. Eine Übersicht über die in diesem Testteil enthaltenen Skalen und deren Quelle gibt Tab. 1.

Tab. 1: Skalen im Testteil zu affektiven Merkmalen.

Skalenbezeichnung	Items	Test	Quelle
Fachspezifisches Selbstkonzept	4	Vortest	Helmke (1992)
Fachinteresse	4	Vortest	Kunter et al. (2002)
Interesse an theoretischen Zusammenhängen	5	Vortest	Eigene Erstellung
Interesse am Experimentieren	4	Vortest	Eigene Erstellung
Interesse an Computern/Computeraffinität	4	Vortest	Brell (2008)
Bewertung der Software	5	Nachtest	Eigene Erstellung
Relevanz des behandelten Themas	3	Nachtest	Laukenmann et al. (2000)
Spaß an der Intervention	4	Nachtest	Laukenmann et al. (2000)

In allen Skalen wurden fünfstufige Likert-Skalen verwendet. Zusätzlich wurden die Mathematik- und die Physiknote in Notenpunkten erhoben.

¹⁴ Das Interesse an theoretischen Überlegungen könnte dabei Ähnlichkeiten zum *Brain Type* des Systematisierers aufweisen (Baron-Cohen et al. (2005); Skorsetz et al. (2021)), der zusammen mit dem Empathisierer in der Empathisierer-Systematisierer-Theorie die Wahl eines naturwissenschaftlichen Studiengangs besser vorhersagt als das Geschlecht (Billington et al. (2007); Skorsetz et al. (2021)).

Testteil zum Cognitive Load

Wie zuvor diskutiert, kann die kognitive Belastung einen Einfluss auf den Lernerfolg haben. Darüber hinaus ist bereits bekannt, dass sich mit einer Lernumgebung, die Videoanalyse nutzt, die kognitive Belastung im Vergleich zu herkömmlichem Unterricht der gleichen Thematik senken lässt. Ein Vergleich der empfundenen kognitiven Belastung der beiden Methoden ist deshalb sinnvoll. Dazu wurde ein Fragebogen von Becker et al. (2019b) übernommen, der aus dem Original von Leppink et al. (2013) übersetzt wurde. Der Fragebogen besteht aus zehn Items, die zwischen den unterschiedlichen Arten der kognitiven Belastung unterscheiden: Intrinsische kognitive Belastung, lernirrelevante kognitive Belastung und lernrelevante kognitive Belastung. Genutzt wird dabei eine sechsstufige Likert-Skala. Dieser Testteil wird nur im Nachtest eingesetzt.

8.4.3 Pilotierung des Testinstruments¹⁵

Der Großteil des oben beschriebenen Tests wurde vor dem Einsatz in der Studie mit insgesamt $N = 85$ Schüler*innen aus fünf Klassen der E-Phase aus hessischen Gymnasien pilotiert. Teile wurden nach vorgenommenen Änderungen mit $N = 116$ Schüler*innen ein weiteres Mal pilotiert. Wiederum weitere Teile wurden lediglich übernommen und die statistischen Kennwerte der Pilotierung durch die ursprünglichen Autoren lediglich berichtet. Eine der Klassen für die eigene Pilotierung war ein Vorleistungskurs, die anderen Grundkurse. Die Pilotierung hat unter den gleichen Bedingungen stattgefunden wie die darauffolgende Hauptstudie. Die Klassen haben den Vortest in der Schule nach dem Unterricht der Newtonschen Gesetze durchgeführt, sind daraufhin in das Physik-Schülerlabor der Goethe-Universität Frankfurt gekommen und haben dort eine der beiden Interventionen durchgeführt. Dabei waren drei Klassen in der Gruppe der mathematischen Modellbildung und zwei Klassen in der Gruppe der Videoanalyse, wobei hier nicht auf die Unterschiede zwischen Gruppen eingegangen wird. Nach der Intervention schloss sich der Nachtest an. Ziel der Pilotierung war es, inhaltliche Skalen zu finden, die Reliabilität des Tests zu bestimmen und die Items auf ihre Schwierigkeit zu überprüfen.

¹⁵ Der Test wird aufgrund der eher geringen Stichprobengröße in der Pilotierung und der Mehrdimensionalität und damit einer geringen Anzahl an Items pro Dimension nach der klassischen Testtheorie (z. B. Gulliksen (1950)) und nicht nach der Item-Response-Theory bzw. mit einem Rasch-Modell ausgewertet.

Testteil zur Mechanik

In der Pilotierung des Testteils zum Konzeptverständnis wurde eine Testversion eingesetzt, die aus 31 Items aus verschiedenen Tests und eigener Kreation bestanden (siehe Kapitel 8.4.2). Durch die Pilotierung sollte der Umfang an Items reduziert und die am besten geeigneten Items ausgewählt werden. Dazu wurden Itemanalysen durchgeführt und die verbliebenen Items auf Basis von Itemschwierigkeit, Trennschärfe und Reliabilität ausgewählt. Die folgenden Kennzahlen beziehen sich auf die 19 Items, welche im Test verblieben sind und damit in der Hauptstudie eingesetzt wurden. Dabei fehlen in der Liste die Items 4, 11, 13a und 13b. Dies sind Items zum dritten Newtonschen Gesetz, die nicht in der Pilotierung enthalten waren und nur als Kontrollskala in den Test der Hauptstudie aufgenommen wurden. Diese vier Items wurden aus dem FCI-Test (Hestenes et al., 1992) übernommen. In der gesamten Analyse wurden im fachlichen Teil zweistufige Items nur als richtig gewertet, wenn Antwort und Begründung richtig waren. Nicht angekreuzte Items wurden als falsch gewertet, da die Proband*innen beliebig viel Zeit zur Verfügung hatten und ein Nichtbeantworten damit einem Nichtwissen (mit Ausnahme von einzelnen Unachtsamkeiten) gleichzusetzen ist. Die Items wurden dichotom ausgewertet. Zunächst können die inhaltlichen Items auf ihren Schwierigkeitsindex in Vor- und Nachtest überprüft werden. Der Schwierigkeitsindex (oder Itemschwierigkeit) sagt aus, von welchem Anteil der Proband*innen der Gesamtstichprobe ein Item richtig beantwortet wurde und wird folgendermaßen berechnet (Lienert & Raatz, 1998):

$$P = \frac{N_R}{N}$$

Dabei ist N_R die Anzahl an Proband*innen, die das Item richtig beantwortet haben und N die Gesamtanzahl an Proband*innen. Einen Überblick gibt Tab. 2.

Die Itemschwierigkeiten liegen im Vortest also zwischen 8,89 % und 51,9 % und im Nachtest zwischen 26,8 % und 76,8 %. Da es sich um einen Niveautest und nicht um einen Schnelligkeitstest handelt, ist zu begrüßen, dass die Itemschwierigkeiten stark um 50 % herum streuen (Lienert & Raatz, 1998). Neun von 19 Items konnten im Nachtest von mehr als 50 % der Probanden gelöst werden (lediglich zwei im Vortest). Abb. 19 zeigt die Itemschwierigkeiten aller Items in Vor- und Nachtest.

Tab. 2: Itemschwierigkeiten der verbliebenen Items in Vor- und Nachtest.

Item	$P_{Vortest}$	$P_{Nachtest}$
1	16,5 %	40,2 %
2	8,89 %	59,8 %
3	24,1 %	57,3 %
5a	51,9 %	56,1 %
5b	27,8 %	39,0 %
6	22,8 %	51,2 %
7	44,3 %	76,8 %
8	32,9 %	41,5 %
9	43,0 %	52,4 %
10a	39,2 %	37,8 %

Item	$P_{Vortest}$	$P_{Nachtest}$
10b	50,6 %	59,8 %
12	20,3 %	37,8 %
14	20,3 %	57,3 %
15a	27,8 %	46,3 %
15b	27,8 %	36,6 %
15c	17,7 %	26,8 %
15d	26,6 %	39,0 %
16a	25,3 %	53,7 %
16b	22,8 %	48,8 %

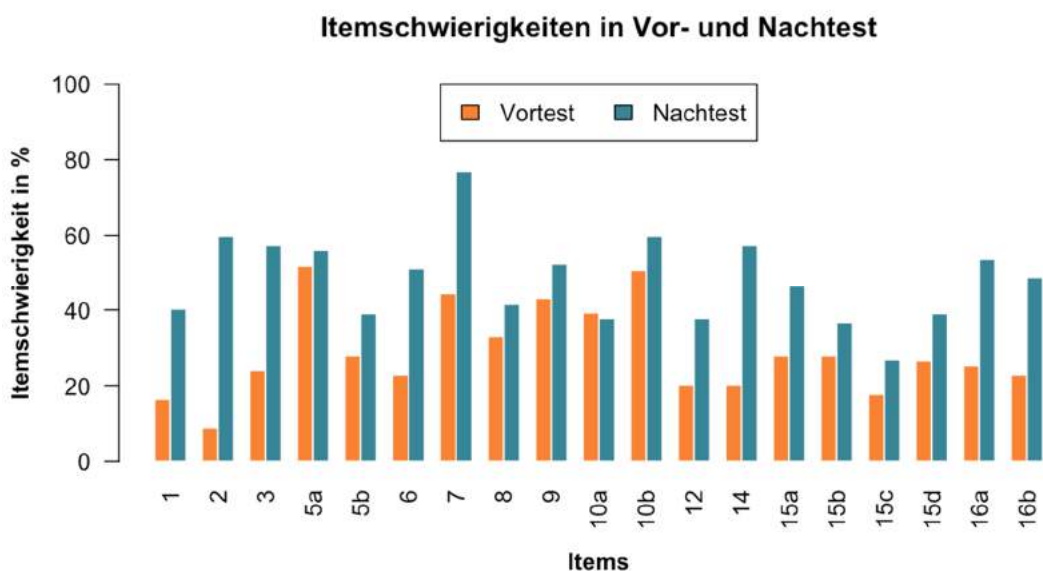


Abb. 19: Itemschwierigkeiten aller fachlichen Items in Vor- und Nachtest.

Außerdem muss die Trennschärfe der Items betrachtet werden. Die Trennschärfe gibt an, wie stark ein Item mit dem Gesamt-Testergebnis korreliert (Lienert & Raatz, 1998). Dadurch kann etwas über die Aussagekraft des Items ausgesagt werden. Der Wert der Trennschärfe kann theoretisch zwischen -1 und 1 liegen, wobei ein negativer Trennschärfekoeffizient bedeutet, dass ein Item von guten Proband*innen eher falsch beantwortet wird als von schlechten oder in der Auswertung falsch gepolt wurde. Dies kam bei dem Test allerdings nicht vor (Tab. 3).

Tab. 3: Trennschärfe der verbliebenen fachlichen Items in Vor- und Nachtest (<0,3 rot, >0,5 grün nach Weise).

Item	Trennschärfe Vortest	Trennschärfe Nachtest	Item	Trennschärfe Vortest	Trennschärfe Nachtest
1	0,322	0,342	10b	0,350	0,493
2	0,591	0,388	12	0,325	0,488
3	0,275	0,113	14	0,443	0,498
5a	0,334	0,282	15a	0,615	0,502
5b	0,512	0,485	15b	0,645	0,485
6	0,431	0,475	15c	0,699	0,446
7	0,312	0,289	15d	0,632	0,491
8	0,548	0,535	16a	0,342	0,502
9	0,336	0,502	16b	0,370	0,503
10a	0,440	0,488			

Dabei sollte die Trennschärfe nach Weise (1975) über 0,30 liegen, wobei Werte über 0,50 als gut gelten. Dies ist bei Items 3, 5a und 7 im Nachtest nicht erfüllt. Insbesondere bei Item 3 ist die Trennschärfe bedenklich gering und auch im Vortest unter der Schwelle von 0,30. Bei der obigen Berechnung der Trennschärfe wird die Korrelation zwischen Items und Gesamtergebnis aller fachlichen Items berechnet, was für einen mehrdimensionalen Test allerdings nicht sinnvoll ist. Wenn durch den Test mehrere inhaltliche Dimensionen getestet werden, ist vielmehr die Korrelation zwischen Item und Ergebnis der jeweiligen Skala entscheidend als die Korrelation zwischen Item und Gesamtergebnis (dies wird später diskutiert und in Tab. 6 dargestellt). Grundsätzlich ist außerdem sowohl

bei leichten, als auch bei schwierigen Items die Trennschärfe tendenziell geringer (Bortz & Döring, 2006), was bei einem Niveautest aber gewünscht sind.

Um herauszufinden, ob der Test, wie geplant, unterschiedliche inhaltliche Dimensionen enthält, wurde eine explorative Faktorenanalyse durchgeführt. Eine explorative Faktorenanalyse dient dazu, über Korrelationen Zusammenhänge zwischen einzelnen Items in einem größeren Set an Items zu finden (Backhaus et al., 2016). Die Gruppen an Variablen, die hoch miteinander korrelieren, nennt man Faktoren. Sie lassen sich je nach Test inhaltlich interpretieren und damit zu gewissen Skalen zusammenfassen. Diese Skalen sollten dann inhaltlich interpretierbar sein und dienen letztlich der Datenreduktion, da nicht mehr auf Itemebene, sondern auf Skalenebene gerechnet und argumentiert wird. Außerdem weist ein gewisses Ergebnis in einer Menge an korrelierenden Items eine höhere Reliabilität auf als einzelne, richtig oder falsch beantwortete Items. Um eine explorative Faktorenanalyse durchführen zu können, sollten zunächst einige Bedingungen überprüft werden. Dazu gehört das Überprüfen der Eignung der Daten für eine Faktorenanalyse durch das Kaiser-Meyer-Olkin-Kriterium (Kaiser, 1970; Kaiser & Rice, 1974). Die „Measure of Sampling Adequacy“ (MSA) gibt an, wieviel Varianz in den Variablen durch zugrundeliegende Faktoren verursacht sein könnte. Dabei geben unterschiedliche Autoren unterschiedliche Grenzwerte an, die erfüllt sein sollten. Nach Field (2003) sollte der MSA über 0,50 liegen, Pallant (2016) empfiehlt Werte über 0,60. Nach Hair (2010) sind Werte über 0,50 zu akzeptieren, Werte zwischen 0,50 und 0,70 mittelmäßig und Werte zwischen 0,70 und 0,80 gut. Nach Kaiser und Rice (1974) sind Werte unter 0,50 nicht zu akzeptieren („unacceptable“), Werte zwischen 0,50 und 0,60 miserabel („miserable“), Werte zwischen 0,60 und 0,70 mittelmäßig („mediocre“), Werte zwischen 0,70 und 0,80 mittelmäßig („middling“), Werte zwischen 0,80 und 0,90 verdienstvoll („meritorious“) und über 0,90 großartig („marvelous“). Dabei kann dieser Wert für jedes Item berechnet werden (Tab. 4).

Aussagekräftiger ist der Gesamtwert aller Items. Mit einem Wert von insgesamt 0,71 eignen sich die hier vorliegenden Daten gut für das Durchführen einer explorativen Faktorenanalyse. Weiterhin kann mit dem Bartlett-Test auf Sphärizität überprüft werden, ob sich die zugrundeliegende Datenmatrix signifikant von einer Einheitsmatrix unterscheidet (Bartlett, 1937). Da eine Einheitsmatrix bedeuten würde, dass keine Korrelationen zwischen den Items bestünden, sollte der Test also ein signifikantes Ergebnis liefern, damit eine Faktoranalyse durchgeführt werden kann. Signifikanz in diesem Test bedeutet also die Ablehnung der Nullhypothese, dass die Korrelationsmatrix eine Einheitsmatrix ist. Die hier vorliegende Datenmatrix unterscheidet sich signifikant von der Einheitsmatrix ($p < 5 \cdot 10^{-67}$).

Tab. 4: Measure of Sampling Adequacy (MSA) für alle fachlichen Items (<0,5: rot, 0,5-0,6: orange, 0,6-0,8: schwarz, 0,8-0,9: hellgrün, >0,9 grün nach Kaiser).

Item	MSA	Item	MSA
1	0,64	10b	0,79
2	0,67	12	0,66
3	0,52	14	0,68
5a	0,61	15a	0,81
5b	0,72	15b	0,73
6	0,69	15c	0,72
7	0,59	15d	0,84
8	0,79	16a	0,61
9	0,79	16b	0,61
10a	0,62		

Für die weitere Analyse sind die Eigenwerte dieser Korrelationsmatrix wichtig. Als Extraktionsmethode wurde die Hauptkomponentenanalyse gewählt, die ein paar Unterschiede zur Faktorenanalyse im engeren Sinne (Hauptachsenanalyse) aufweist. Die Hauptkomponentenanalyse versucht, die Datenstruktur durch möglichst wenige Faktoren zu erklären (Backhaus et al., 2016). Man geht hier davon aus, dass die gesamte Varianz durch die Faktoren erklärt werden kann. Nach der Prüfung der Eignung der Daten muss eine Entscheidung über die Anzahl an Faktoren getroffen werden, wonach dann die Items zu den Faktoren zugeordnet werden müssen, was ein Optimierungsproblem darstellt.

Um die Anzahl an Faktoren bestimmen zu können, gibt es verschiedene Kriterien. Bei allen Kriterien werden die Eigenwerte der Korrelationsmatrix absteigend sortiert und den Faktoren zugeordnet. Ein häufig verwendetes Kriterium ist das Kaiser-Kriterium, wonach alle Faktoren, deren Eigenwert größer als 1 ist, verwendet werden (Kaiser, 1960). Dieses Kriterium ist zwar eindeutig, jedoch häufig zu pauschal. Es überschätzt meist die Dimensionalität und sollte deshalb nicht als einziges Kriterium für die Anzahl an zu extrahierenden Faktoren herhalten (Lance & Vandenberg, 2009). Andererseits wird häufig der Scree-Test bzw. das Ellenbogenkriterium genutzt, welches den Scree-Plot nutzt (Cattell,

1966). Beim Scree-Test wird im Scree-Plot nach einer Knickstelle gesucht, nach der sich der Graph asymptotisch an die Abszisse annähert. Der Scree-Plot zeigt dabei die Eigenwerte der Hauptkomponenten bzw. Faktoren. Bei mehreren Knicken ist in der Regel der Rechte von beiden zu nutzen. Die Anzahl an zu extrahierenden Faktoren ist dann die Anzahl der Faktoren links neben diesem Knick.

Eine Weiterentwicklung des Scree-Tests ist die sogenannte Parallelanalyse nach Horn (1965), die auch auf dem Plot der Eigenwerte der Korrelationsmatrix basiert, aber zusätzlich noch einen weiteren Scree-Plot von Eigenwerten einer zufällig generierten Matrix einzeichnet. Die Anzahl an Faktoren, die einen höheren Eigenwert haben als die Zufallseigenwerte gibt die Anzahl an zu extrahierenden Faktoren an (Abb. 20).

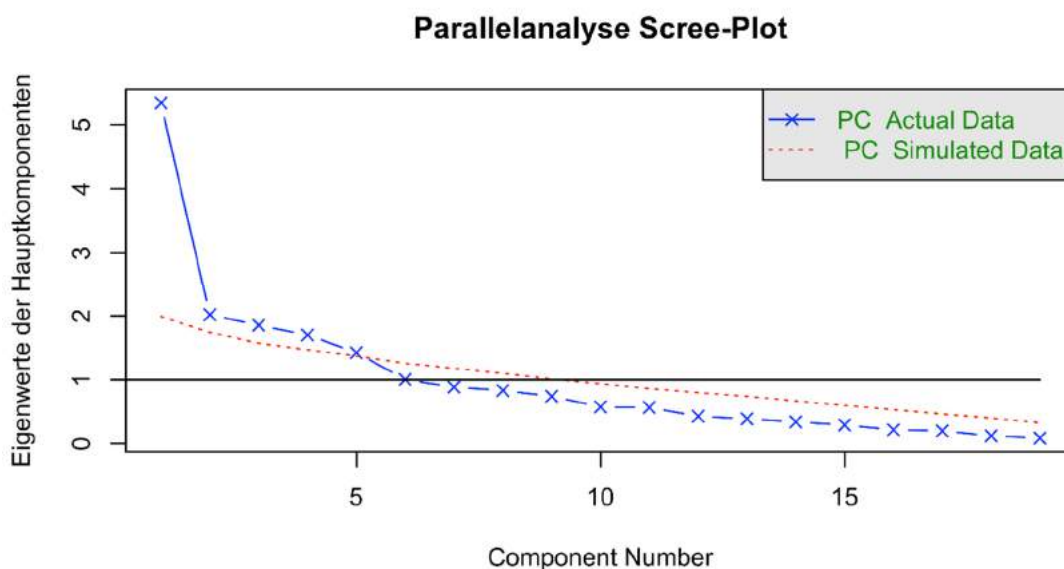


Abb. 20: Scree-Plot der Daten (x) und der simulierten Daten (.) mit Grenze für das Kaiser-Kriterium bei 1 im fachlichen Teil des Tests.

Das Kaiser-Kriterium liefert sechs Faktoren, wobei der Eigenwert des sechsten Faktors lediglich 1,005 ist und damit nur unwesentlich über der Grenze von 1 liegt. Der Scree-Test und die Parallelanalyse liefern fünf Faktoren. Bei inhaltlicher Betrachtung dieser Faktoren fällt auf, dass die Items, die auf zwei entstehende Faktoren laden, stark miteinander korrelieren und inhaltlich ebenfalls zu einer Skala passen, weswegen die beiden Skalen zusammengelegt wurden und mit vier Faktoren gerechnet wird. Die Faktorladungen, die inhaltliche Benennung der sich ergebenden Skalen und die Werte für Cronbachs Alpha dieser Skalen sind in Tab. 5 sichtbar.

Tab. 5: Faktorladungen der explorativen Faktorenanalyse der inhaltlichen Items (Faktorladungen unter 0,35 unterdrückt, fett gedruckte Ladungen geben die endgültige Zuordnung an).

Items	Komponente				Name	α_c
	1	2	3	4		
15b	0,892				Beschleunigungsdiagramme	0,91
15a	0,874					
15d	0,828					
15c	0,782					
16b		0,875			Von bekannter Bewegung auf Kräfte schließen	0,73
16a		0,786				
1		0,622				
9	0,359	0,514				
14		0,475				
12			0,815		Erstes Newtonsches Gesetz mit Kräftekompensation	0,75
6			0,776			
8	0,353		0,739			
3			0,585			
7				0,712	Von bekannten Kräften auf Bewegung schließen	0,70
5a				0,699		
2				0,600		
10a				0,596		
5b	0,404			0,522		
10b	0,414			0,370		

Die Hauptkomponenten wurden mit einer Varimax-Rotation mit Kaiser-Normalisierung gefunden. Die Varimax-Rotation (Kaiser, 1958) entspricht einer wiederholten Rotation der Koordinatenachsen im Raum. Die Achsen bleiben dabei orthogonal zueinander und es wird so lange rotiert, bis die Varianz der quadrierten Faktorladungen maximiert wird. Die Kaiser-Normalisierung steht dafür, dass die Faktorladungen vor der Rotation normalisiert werden.

Die Faktorladungen unter 0,35 wurden unterdrückt. Der fettgedruckte Wert und die Position in der Tabelle geben die endgültige Zuordnung an. Das Item 10b wurde dabei aus inhaltlichen Gründen zu Skala 4 zugeordnet, obwohl die Faktorladung auf Faktor 1 geringfügig höher ist.

Diese Faktoren lassen sich inhaltlich interpretieren. Auf Faktor 1 laden nur Items, die Beschleunigungsdiagramme betreffen, auf Faktor 2 laden nur Items, bei denen die Proband*innen von einer gegebenen Bewegung auf die wirkenden Kräfte schließen sollen, auf Faktor 3 laden Items, die das erste Newtonsche Gesetz mit Kräftekompensation betreffen und auf Faktor 4 laden Items, bei denen die Proband*innen bei bekannter Kraft die Bewegung bestimmen sollen. Die Werte für Cronbachs Alpha (α_c) sind ein Maß für die Reliabilität einer Skala. Sie sollten nach Nunnally und Bernstein (2008) über 0,70 liegen, damit eine ausreichende Reliabilität gegeben ist. Die genauere Bedeutung wird in Kapitel 8.4.4 diskutiert. Da nun vier Skalen gefunden wurden, kann eine erneute Betrachtung der Trennschärfen erfolgen (Tab. 6).

Tab. 6: Trennschärfen der Items im fachlichen Teil in den verschiedenen Skalen (<0,3 rot, >0,5 grün nach Weise).

Items Skala 1	Trenn- schärfe	Items Skala 2	Trenn- schärfe	Items Skala 3	Trenn- schärfe	Items Skala 4	Trenn- schärfe
15a	0,767	1	0,406	3	0,330	2	0,398
15b	0,849	9	0,419	6	0,623	5a	0,401
15c	0,772	14	0,371	8	0,620	5b	0,477
15d	0,776	16a	0,571	12	0,607	7	0,437
		16b	0,726			10a	0,452
						10b	0,398

Die Trennschärfen liegen nun alle über dem Grenzwert von 0,30, womit die Items für den Test geeignet sind. Dies bestätigt zudem erneut die Multidimensionalität des Tests.

Testteil zu Modellen und Theorie und Experiment in der Physik

Nach einer ersten Pilotierung der Items zu den Vorstellungen über Modelle in der Physik wurden auf Basis der Resultate einige Änderungen vorgenommen und die neuen Items erneut pilotiert. Anders als bei der ersten Pilotierung wurden diese Items nicht mit dem gesamten Fragebogen im Schülerlabor Dynamik, sondern mit ausgewählten Klassen der E-Phase von gymnasialen Oberstufen im Unterricht pilotiert. 122 Schüler*innen füllten den Bogen aus. Davon waren $N = 116$ Bögen vollständig. In dem Fragebogen waren einerseits Fragen über das Wesen von Modellen enthalten (deklaratives Modellverständnis) und andererseits wurde erfragt, wo die Interessen der Schüler*innen in Bezug auf Theorie oder Modellieren und Experimentieren liegen.

Eine explorative Faktorenanalyse sollte zeigen, ob sich diese Dreiteilung der Items in den Daten zeigt. Um die Eignung der Daten für eine explorative Faktorenanalyse zu untersuchen, wurde das Mayer-Olkin-Kriterium überprüft und der Bartlett-Test auf Sphärizität durchgeführt (Tab. 7).

Tab. 7: Measure of Sampling Adequacy (MSA) für Items zu Modellen und Theorie/Experiment in der Physik (<0,5: rot, 0,5-0,6: orange, 0,6-0,8: schwarz, 0,8-0,9: hellgrün, >0,9 grün nach Kaiser).

Item	MSA
F1a	0,77
F1b	0,67
F1c	0,71
F2a	0,69
F2b	0,64
F2d	0,72
F3a	0,72
F8a	0,68

Item	MSA
F8b	0,65
F8c	0,73
F8d (umkodiert)	0,68
F9a	0,81
F9b	0,76
F9c	0,73
F9d	0,68

Die vorliegenden Daten haben ein MSA von 0,71 und der Bartlett-Test ist signifikant ($p < 5 \cdot 10^{-47}$). Damit eignen sie sich für eine explorative Faktorenanalyse. Der Scree-Plot mit Parallelanalyse sieht wie folgt aus (Abb. 21):

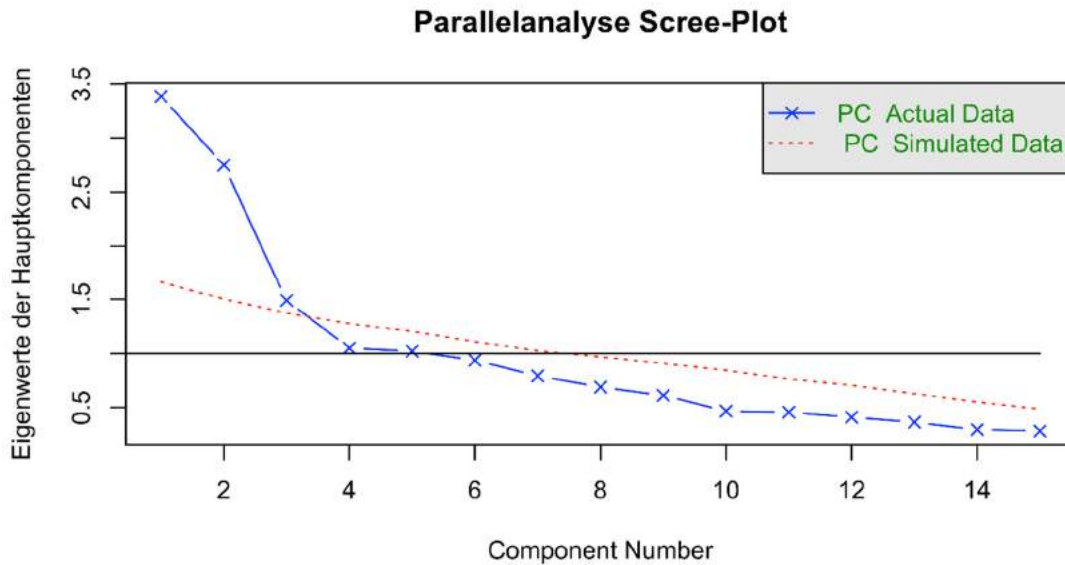


Abb. 21: Scree-Plot der Daten (x) und der simulierten Daten (.) mit Grenze für das Kaiser-Kriterium bei 1 im Testteil zu Modellen und Theorie/Experiment in der Physik.

Das Kaiser-Kriterium liefert fünf Faktoren, während der Scree-Test und die Parallelanalyse drei Faktoren vorschlagen. Bei drei Faktoren ergeben sich folgende Faktorladungen (Tab. 8).

Die Hauptkomponenten wurden mit einer Varimax-Rotation mit Kaiser-Normalisierung gefunden. Die Faktorladungen unter 0,35 wurden unterdrückt. Der fettgedruckte Wert und die Position in der Tabelle geben die endgültige Zuordnung an. Die Faktorladungen lassen sich inhaltlich gut deuten und sind außerdem reliabel mit $\alpha_C > 0,70$.

Tab. 8: Faktorladungen für die Items im Testteil zu Modellen und Theorie/Experiment in der Physik (Faktorladungen unter 0,35 unterdrückt, fett gedruckte Ladungen geben die endgültige Zuordnung an).

Item	Komponente			Name	α_c
	1	2	3		
F2a	0,717			Vorstellungen über Modelle	0,72
F3a	0,686				
F1b	0,602				
F1a	0,589				
F2b	0,563				
F1c	0,557				
F2d	0,510		0,367		
F8c		0,798		Interesse am Experimentieren	0,72
F8b		0,788			
F8a		0,751			
F9b			0,779	Interesse an theoretischen Zusammenhängen	0,77
F9a			0,758		
F9c			0,733		
F9d			0,677		
F8d (umkodiert)			0,583		

Zudem können wieder die Itemschwierigkeiten betrachtet werden, die der durchschnittlichen anteiligen Punktzahl in einem von 0 bis 4 kodierten Item entsprechen (Tab. 9). Bei Items, in denen affektive Merkmale auf Likert-Skalen abgefragt werden, wird in der Regel eine Itemschwierigkeit von 50 % angestrebt und Items, deren Itemschwierigkeiten unter 20 % oder über 80 % liegen, ausgeschlossen (Mummendey & Grau, 2014). Die Schwierigkeiten sind hier als angemessen anzusehen.

Tab. 9: Itemschwierigkeiten für die Items im Testteil zu Modellen und Theorie/Experiment in der Physik.

Item	P	Item	P	Item	P
F2a	60,9 %	F8c	68,2 %	F9b	36,1 %
F3a	50,6 %	F8b	63,4 %	F9a	41,4 %
F1b	54,0 %	F8a	58,2 %	F9c	37,5 %
F1a	56,3 %			F9d	34,9 %
F2b	54,2 %			F8d (umkodiert)	33,8 %
F1c	52,2 %				
F2d	65,7 %				

Außerdem ergeben sich folgende Trennschärfen (Tab. 10). Wie im vorigen Kapitel diskutiert, sollte die Korrelation zwischen einem Item und dem Ergebnis in der Skala über einem Wert von 0,30 liegen. Dies ist bei allen Items der Fall.

Tab. 10: Trennschärfen der Items Testteil zu Modellen und Theorie/Experiment in der Physik in den verschiedenen Skalen (<0,3 rot, >0,5 grün nach Weise).

Items Skala 1	Trenn- schärfe	Items Skala 2	Trenn- schärfe	Items Skala 3	Trenn- schärfe
F2a	0,546	F8c	0,616	F9b	0,613
F3a	0,499	F8b	0,572	F9a	0,633
F1b	0,389	F8a	0,454	F9c	0,500
F1a	0,442			F9d	0,551
F2b	0,431			F8d (umkodiert)	0,435
F1c	0,338				
F2d	0,360				

Testteil zu affektiven Merkmalen

Die Fragen zu affektiven Merkmalen in der Physik wurden mit $N = 79$ Proband*innen in der Pilotierung vollständig beantwortet. Es wurden Items aus Kunter et al. (2002) zum Fachinteresse und aus Helmke (1992) zum fachspezifischen Selbstkonzept übernommen. Um den Umfang des Fragebogens zu reduzieren, wurde eine Itemanalyse durchgeführt und die Reliabilitäten der Skalen berechnet, um einige Items aus den verwendeten Fragebögen zu streichen. Bei den im Test verbliebenen Items ergeben sich folgende Itemschwierigkeiten (Tab. 11):

Tab. 11: Itemschwierigkeiten der Items zu affektiven Merkmalen.

Item	<i>P</i>	Item	<i>P</i>
Ia	48,1 %	Ib	58,5 %
If (umkodiert)	55,9 %	Id (umkodiert)	53,2 %
Ii (umkodiert)	55,7 %	Ig	54,7 %
Ij (umkodiert)	58,5 %	Ik	57,0 %

Auf die Bestimmung der Anzahl an Hauptkomponenten wurde hier aufgrund der Bekanntheit der enthaltenen Skalen verzichtet, da die Items aus bekannten Fragebögen zum „Fachinteresse“ und „Fachspezifischen Selbstkonzept“ übernommen wurden und somit von zwei Hauptkomponenten auszugehen ist. Die Faktorladungen können dennoch als Bestätigung der Zuordnung zu den beiden Skalen dienen (Tab. 12).

Die Hauptkomponenten wurden ebenfalls mit einer Varimax-Rotation mit Kaiser-Normalisierung gefunden. Die Faktorladungen unter 0,35 wurden unterdrückt. Der fettgedruckte Wert und die Position in der Tabelle geben die endgültige Zuordnung an. Dabei ist es nicht verwunderlich, dass Items zum „Fachinteresse“ z. T. auch mit dem Faktor „Fachspezifisches Selbstkonzept“ korrelieren. Die beste Zuordnung bestätigt aber die vorher vermutete Struktur. Die Reliabilitäten der beiden Skalen sind ebenfalls als sehr gut anzusehen.

Tab. 12: Faktorladungen der Items zu affektiven Merkmalen (Faktorladungen unter 0,35 unterdrückt, fett gedruckte Ladungen geben die endgültige Zuordnung an).

Item	Komponente		Name	α_c
	1	2		
If (umkodiert)	0,882		Fachspezifisches Selbstkonzept	0,93
Ii (umkodiert)	0,873			
Ij (umkodiert)	0,807	0,431		
Ia	0,782	0,372		
Id (umkodiert)		0,817	Fachinteresse	0,87
Ik	0,448	0,795		
Ig		0,782		
Ib	0,575	0,651		

In den beiden Skalen können ebenfalls die Trennschärfen der Items betrachtet werden (Tab. 13). Die Trennschärfen der pilotierten Items sind durchweg als gut einzuschätzen, da sie alle über der Schwelle von 0,50 liegen.

Tab. 13: Trennschärfe der Items zu affektiven Merkmalen (<0,3 rot, >0,5 grün nach Weise).

Items Skala 1	Trennschärfe	Items Skala 2	Trennschärfe
Ia	0,922	Ib	0,752
If (umkodiert)	0,896	Id (umkodiert)	0,648
Ii (umkodiert)	0,893	Ig	0,685
Ij (umkodiert)	0,896	Ik	0,844

Testteil zur Bewertung des Schülerlabors

Auch der Testteil zur Bewertung des Schülerlabors wurde mit dem gesamten Test zusammen mit $N = 85$ Schüler*innen pilotiert. Da auch hier bei der Erstellung des Fragebogens eine Mehrdimensionalität angestrebt und vermutet wurde, sollen zunächst die

Bedingungen für eine explorative Faktorenanalyse überprüft werden. Das Kaiser-Mayer-Olkin-Kriterium ergibt für die Korrelationsmatrix der zwölf Items einen Wert von $MSA = 0,69$ (Tab. 14):

Tab. 14: Measure of Sampling Adequacy (MSA) für Items zu affektiven Merkmalen ($<0,5$: rot, $0,5-0,6$: orange, $0,6-0,8$: schwarz, $0,8-0,9$: hellgrün, $>0,9$ grün nach Kaiser).

Item	MSA
IIj	0,66
IIk	0,73
IIm	0,79
IIn	0,69
IIo (umkodiert)	0,67
IIb	0,63

Item	MSA
IIe	0,67
IIi	0,74
IIa	0,76
IIc	0,69
IIId	0,82
IIg	0,88

Die Korrelationsmatrix unterscheidet sich zudem signifikant ($p < 5 \cdot 10^{-47}$) von einer Einheitsmatrix, wodurch die Daten für eine explorative Faktorenanalyse geeignet sind. Dafür wird zunächst der Scree-Plot betrachtet (Abb. 22):

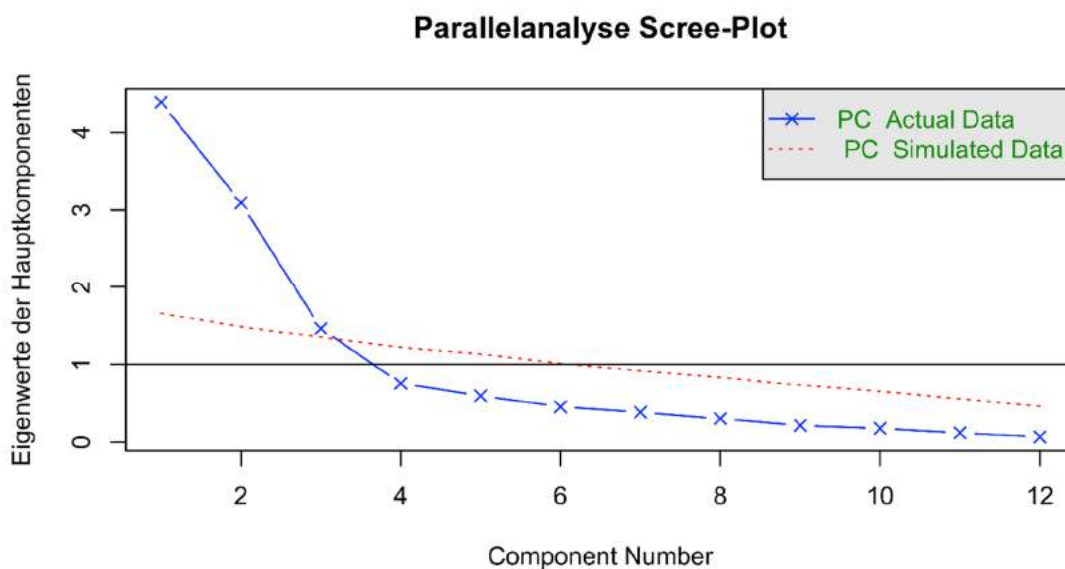


Abb. 22: Scree-Plot der Daten (x) und der simulierten Daten (.) mit Grenze für das Kaiser-Kriterium bei 1 im Testteil zur Bewertung des Schülerlabors.

Der Screeplot zeigt, dass Kaiser-Kriterium, Scree-Test und Parallelanalyse hier alle drei Faktoren vorschlagen. Die Faktorladungen sehen mit drei Faktoren wie in Tab. 15 dargestellt aus.

Die Hauptkomponenten wurden mit einer Varimax-Rotation mit Kaiser-Normalisierung gefunden. Die Faktorladungen unter 0,35 wurden unterdrückt. Der fettgedruckte Wert und die Position in der Tabelle geben die endgültige Zuordnung an. Die Faktorladungen lassen sich inhaltlich gut deuten und die gefunden Skalen sind außerdem reliabel mit $\alpha_C > 0,7$.

Im Test außerdem enthalten war ein Testteil zum Cognitive Load (siehe Kapitel 8.4.2). Dieser wurde von Becker et al. (2020b) pilotiert (siehe auch Kapitel 8.4.4), weswegen eine eigene Pilotierung hier nicht durchgeführt wurde.

Tab. 15: Faktorladungen der Items zur Bewertung des Schülerlabors (Faktorladungen unter 0,35 unterdrückt, fett gedruckte Ladungen geben die endgültige Zuordnung an).

Item	Komponente			Name	α_C
	1	2	3		
II _n	0,869			Bewertung der verwendeten Software	0,85
II _m	0,813				
II _o (umkodiert)	0,805				
II _k	0,741				
II _j	0,664				
II _b		0,943		Relevanz des behandelten Themas	0,78
II _i		0,882			
II _e		0,871			
II _c			0,890	Spaß an der Intervention	0,85
II _a			0,865		
II _d			0,814		
II _g		0,485	0,666		

8.4.4 Gütekriterien des Testinstruments

Allgemein lassen sich drei Gütekriterien eines Tests unterscheiden, die erfüllt werden sollten (z. B. Przyborski & Wohlrab-Sahr, 2014). Diese drei Gütekriterien sind Objektivität, Reliabilität und Validität. Sie sind jeweils Oberbegriffe für verschiedene Ausprägungen. Das erstellte Testinstrument soll in den folgenden Kapiteln im Hinblick auf die Erfüllung der Gütekriterien diskutiert werden.

Objektivität

Objektivität von Testinstrumenten bedeutet, dass die Durchführung und die Ergebnisse des Tests unabhängig von der den Test durchführenden Person sein sollten (Bortz & Döring, 2006). Bei der Objektivität können drei Unterformen unterschieden werden.

Die Durchführungsobjektivität sagt aus, dass eine Probandin bzw. ein Proband bei unterschiedlichen Testanwender*innen dasselbe Resultat erzielt. Dies kann verletzt sein, wenn durch den Testanwendenden zusätzlich Instruktionen nötig sind, damit die Proband*innen den Test verstehen. Verhindert wird eine Verletzung der Durchführungsobjektivität meist durch standardisierte Testinstruktionen (Bortz & Döring, 2006).

Die Auswertungsobjektivität bezieht sich auf die Person des Auswertenden. Das heißt, dass verschiedene Auswerter*innen zu exakt demselben Ergebnis kommen sollten. Dies kann ebenfalls durch genaue Instruktionen erreicht werden. Wenn es eine unterschiedliche Bepunktung von Aufgaben gibt, muss also klar sein, welche Antwort wie viele Punkte wert ist. Einfach zu erreichen ist dies durch Antworten im Richtig-Falsch-Format oder bei Single- oder Multiple-Choice Aufgaben (Bortz & Döring, 2006).

Die Interpretationsobjektivität bezieht sich auf die Interpretation des Testergebnisses selbst. Dabei dürfen keine individuellen Deutungen in die Bewertung des Testergebnisses einfließen. Vielmehr ist ein Vergleich mit standardisierten Normen vorzuziehen (Bortz & Döring, 2006). Standardisierte Normen ermöglichen den Vergleich der Testergebnisse und eine angemessene Einordnung. Die Ergebnisse der Interpretation des Tests sollten also nicht von der Person abhängen.

Die drei Objektivitätskriterien sind für den erstellten Fragebogen als erfüllt anzusehen. Die Durchführungsobjektivität ist dadurch gewährleistet, dass die Instruktionen des Testinstruments in schriftlicher Form auf dem Test vorliegen. Die Auswertungsobjektivität ist bei Multiple-Choice-Tests generell in hohem Maße erfüllt. Die Interpretationsobjektivität ist bedingt gewährleistet. In den unterschiedlichen Testteilen kann zu einem unterschiedlichen Maß auf Normen zurückgegriffen werden. Beispielsweise kann bei der

Interpretation der Ergebnisse des Mechanik-Teils auf Ergebnisse des FCI-Tests zurückgegriffen werden. In anderen Teilen ist dies nicht immer möglich. Da der Test aber vor allem zu einem Vergleich zwischen den beiden Gruppen herangezogen wird und die Ergebnisse nicht alleinstehend bewertet werden, stellt die Interpretationsobjektivität ein geringeres Problem dar.

Reliabilität

Die Reliabilität ist ein Genauigkeitsmaß. Es gibt an, wie präzise ein untersuchtes Merkmal mit dem Test gemessen werden kann. In der Praxis sind Messwerte immer fehlerbehaftet. Dies kann durch Unkonzentriertheit, Müdigkeit oder auch das Raten von Proband*innen hervorgerufen werden (Bortz & Döring, 2006). Die Reliabilität misst den Anteil der wahren Varianz an der beobachteten Varianz und kann daher Werte von 0 bis 1 annehmen (Bortz & Döring, 2006). Es gibt verschiedene Möglichkeiten, die Reliabilität zu bestimmen. Diese Verfahren dienen dazu, die wahre Varianz abzuschätzen und damit die Reliabilität berechnen zu können.

Die Idee der sogenannten Retestreliabilität ist es zu bestimmen, wie ähnlich sich zwei Testergebnisse sind, die zu unterschiedlichen Zeitpunkten stattgefunden haben. Je höher die Korrelation zwischen beiden Tests ist, desto eher kann man davon ausgehen, dass der Test das Konstrukt genau misst. Da es hier um eine zeitliche Stabilität der Testergebnisse geht, wird dieses Kriterium auch Stabilität genannt (Bortz & Döring, 2006). Wichtig ist dabei, dass der Test von derselben Stichprobe erneut durchgeführt wird und dass in der Zwischenzeit keine Intervention stattgefunden hat, die das Testergebnis beeinflussen könnte. Wenn also Konzeptverständnis eines Unterrichtsinhalts in der Physik getestet werden soll, sollte in der Zwischenzeit kein Physikunterricht zu diesem Thema stattgefunden haben. Generell sollte der Zeitraum zwischen den Testungen nicht zu groß sein, insbesondere, wenn instabile Merkmale gemessen werden sollen. Die Berechnung der Retestreliabilität kann bei instabilen Merkmalen aber ohnehin problematisch sein. Andererseits muss ausreichend Zeit dazwischenliegen, damit die Proband*innen sich nicht mehr an ihre Antworten aus dem Test zuvor erinnern können und so dasselbe nochmal ankreuzen, da sonst die Reliabilität überschätzt werden würde. Üblich sind mehrere Wochen (Bortz & Döring, 2006), wobei dies auch von der Art des Tests abhängen kann. Die Korrelation der beiden Testergebnisse gibt dann an, welcher Anteil der Unterschiede zwischen Testergebnissen durch wahre Unterschiede in den gemessenen Merkmalen zu erklären sind. Eine (unrealistische) perfekte Korrelation von 1 zwischen den beiden Testzeitpunkten würde bedeuten, dass die Unterschiede der Testergebnisse zwischen zwei Probanden nur auf wahre Unterschiede in der Merkmalsausprägung zurückzuführen sind.

Die Überprüfung der Retestreliabilität bietet sich in dieser Studie nicht an, da eine zweifache Durchführung des Tests mit vielen Schüler*innen einerseits einen großen organisatorischen Aufwand mit sich bringen würde und viele Lehrkräfte dafür im Unterricht keine Zeit bereitstellen können. Andererseits darf, wie bereits beschrieben, zwischen den Testungen keine merkmalsverändernde Intervention stattfinden. Da der Physikunterricht aber gerade eine solche Intervention ist, ist diese Methode ungeeignet.

Für die Paralleltestreliabilität müssen zwei Testversionen erstellt werden, die derselben Population kurz hintereinander gegeben werden. Die Ähnlichkeit der beiden Testergebnisse ist ein Maß dafür, wie genau das zu messende Merkmal gemessen werden kann. Das Resultat ist dann ein Kennwert beider Testvarianten, sodass entweder beide reliabel oder beide nicht reliabel sind. Die Erstellung von zwei Testvarianten ist mit einigem Aufwand verbunden. Insbesondere da die Items zunächst von einer Stichprobe durchgeführt werden müssen, damit sie nach Schwierigkeit sortiert werden können und dann, nach der Erstellung von zwei äquivalenten Testvarianten, von einer zweiten Stichprobe bearbeitet werden. Zudem sind mindestens doppelt so viele Items nötig, wie im fertigen Test vorhanden sein sollen (Bortz & Döring, 2006). Geeignet ist das Verfahren deswegen vor allem dann, wenn ohnehin zwei Testvarianten benötigt werden.

Die Testhalbierungsreliabilität kann mit dem einmaligen Durchführen des Tests in einer Stichprobe abgeschätzt werden. Nach Durchführung des Tests wird der Test in zwei Hälften geteilt und damit zwei Testergebnisse berechnet. Die Reliabilität entspricht dann wieder der Korrelation zwischen diesen beiden Testwerten. Allgemein unterschätzt diese Methode die Reliabilität, da diese nur für eine Testhälfte bestimmt wird, die Reliabilität aber mit steigender Anzahl an Items zunimmt. Dies kann allerdings im Nachhinein korrigiert werden (Bortz & Döring, 2006). Zu kritisieren ist an dieser Methode, dass es nicht klar ist, in welcher Weise der Test halbiert werden soll und dass die Wahl der Halbierung maßgeblichen Einfluss auf die ermittelte Reliabilität hat.

Um diesem Problem aus dem Weg zu gehen, lässt sich auch die interne Konsistenz eines Tests bestimmen, was einer Weiterentwicklung der Testhalbierungsreliabilität entspricht. Der Test wird bei der Berechnung der internen Konsistenz nicht in Hälften zerteilt, sondern in so viele Teile, wie es im Test Items gibt. Damit entspricht dann nicht mehr die Korrelation zwischen zwei Testhälften der Reliabilität, sondern die Korrelation zwischen den Items. Häufig wird dazu der Wert für Cronbachs Alpha berechnet (siehe z. B. Tab. 5).

„Formal entspricht der Alphakoeffizient der mittleren Testhalbierungsreliabilität eines Tests für alle möglichen Testhalbierungen“ (Bortz & Döring, 2006, S. 198) und wird folgendermaßen berechnet:

$$\alpha_c = \frac{p}{p-1} \cdot \left(1 - \frac{\sum_{i=1}^p s_{Item(i)}^2}{s_{Testwert}^2} \right)$$

Dabei ist p die Anzahl an Items, s_{Item} die Standardabweichung der Items und $s_{Testwert}$ die Standardabweichung der Summe bzw. des Gesamtscores. Bei heterogenen und mehrdimensionalen Tests unterschätzt Alpha die Reliabilität. Da in dieser Studie ein mehrdimensionaler Test eingesetzt wurde, ist nicht davon auszugehen, dass alle Items untereinander hoch korrelieren. Daher kann die interne Konsistenz für jede Subskala des Tests einzeln bestimmt werden.

Bortz und Döring (2006) empfehlen eine Reliabilität von mindestens 0,80. Andere Autoren gehen von einem Alpha von 0,70 als Untergrenze für einen ausreichend reliablen Test aus (Nunnally & Bernstein, 2008). Der hier verwendete fachliche Teil des Tests hat bei 19 Items ein Cronbachs Alpha von 0,84 und ist damit ausreichend reliabel. Die gefundenen Subskalen sind dabei auch jeweils reliabel (siehe Tab. 16 & Tab. 17).

Tab. 16: Reliabilitäten aller im Test verwendeter Skalen (Teil 1).

Testteil	Skalenname	Itemzahl	α_c
Affektive Merkmale zum Fach Physik	Fachinteresse	4	0,87
	Fachspezifisches Selbstkonzept	4	0,93
Modell und Experiment	Einstellungen gegenüber Modellen	7	0,72
	Interesse am Experimentieren	3	0,78
	Interesse am Theoretisieren/Modellieren	5	0,76
Interesse an Computern	Computeraffinität	4	0,72¹⁶

¹⁶ Brell (2008)

Tab. 17: Reliabilitäten aller im Test verwendeter Skalen (Teil 2).

Testteil	Skalename	Itemzahl	α_c
Bewertung des Schülerlabors	Bewertung der Software	5	0,85
	Relevanz des behandelten Themas	3	0,78
	Spaß an der Intervention	4	0,85
Cognitive Load	Intrinsic Cognitive Load	3	0,86¹⁷
	Extraneous Cognitive Load	3	0,55¹⁸
	Germane Cognitive Load	4	0,92¹⁹
Newtonsche Mechanik	Beschleunigungsdiagramme	4	0,91
	1. Newtonsches Gesetz mit Kräftekompensation	4	0,75
	Bestimmen der Kraft bei bekannter Bewegung	5	0,73
	Bestimmen der Bewegung bei bekannter Kraft	6	0,70
	3. Newtonsches Gesetz	4	0,78²⁰

Mit Ausnahme des Cronbachs Alpha für die Skala Extraneous Cognitive Load liegen alle Werte bei mindestens 0,70, wodurch die Reliabilität des gesamten Tests insgesamt als erfüllt angesehen werden kann. Die möglicherweise geringe Genauigkeit der Skala Extraneous Cognitive Load wird in der Auswertung aufgegriffen und berücksichtigt.

¹⁷ Becker et al. (2020b)

¹⁸ Becker et al. (2020b)

¹⁹ Becker et al. (2020b)

²⁰ Schecker und Gerdes (1999b)

Validität

Die Validität gibt an, inwiefern der Test wirklich das misst, was er messen soll. Wenn also die Validität nicht gegeben ist und ein Test für das Messen des Konzeptverständnisses der Newtonschen Mechanik eigentlich etwas anderes misst, ist der Test auch bei einer hohen Reliabilität nicht brauchbar. Die Validität kann als das wichtigste Gütekriterium bezeichnet werden, ist aber auch sehr schwer zu überprüfen. Auch bei der Validität werden verschiedene Arten unterschieden.

Die Inhaltsvalidität misst, ob ein Test, der ein gewisses Konstrukt messen soll, dieses Konstrukt auch vollständig misst. Dies ist nicht erfüllt, wenn einzelne Aspekte des Konstrukts im Test nicht erfasst werden. Ein Test, der vorgibt, das Konzeptverständnis der Newtonschen Gesetze zu messen, hat eine geringe Inhaltsvalidität, wenn das dritte Newtonsche Gesetz nicht Teil des Tests ist. Die Inhaltsvalidität kann nicht berechnet werden und wird meist durch eine Expertenbefragung zugebilligt (Lienert & Raatz, 1998), was hier mit den beteiligten Lehrkräften erfolgte (Tab. 18).

Tab. 18: Zustimmung der Lehrkräfte zu der Aussage, dass das jeweilige Item Inhalte aus der Mechanik der E-Phase testet.

Item	Zustimmung	Item	Zustimmung
1	88,9 %	9	83,3 %
2	100 %	10	95,0 %
3	95,0 %	11	100 %
4	100 %	12	100 %
5	88,9 %	13	95,0 %
6	100 %	14	78,9 %
7	100 %	15	100 %
8	94,4 %	16	89,5 %

In der Expertenbefragung ($N = 19$) wurde die Frage gestellt, ob das Item „Inhalte aus der Mechanik der E-Phase testet“. Die Zustimmung ist bei allen Items sehr hoch, was für eine gute Validität spricht. Nicht gefragt wurde, ob der Test in seiner Gänze die gesamten

Inhalte der Mechanik abdeckt, was aber auch nicht Ziel des Fragebogens war. Vielmehr sollen damit einige Aspekte getestet werden.

Die Kriteriumsvalidität gibt an, inwiefern das Testergebnis, welches etwas über ein latentes Merkmal der Testperson aussagt, mit einem manifesten Merkmal übereinstimmt. Wenn ein Test eingesetzt wird, der die Eignung für einen bestimmten Studiengang messen soll, sollte das Ergebnis mit dem tatsächlichen Studienerfolg korrelieren. Häufig, wie auch in dem hier verwendeten Test, gibt es aber kein manifestes Merkmal, dessen Übereinstimmung man mit dem latenten Merkmal überprüfen könnte. Denkbar wäre es, die Korrelation mit der Note in dem entsprechenden Schulhalbjahr in Physik zu berechnen. Da in die Schulnote aber noch viele andere Dinge einfließen und nicht davon ausgegangen werden kann, dass die Schulnote reliabel und valide ausschließlich das Konzeptverständnis der Mechanik wiedergibt, ist dies nur bedingt sinnvoll. In dieser Studie wurden außerdem nur die zurückliegenden Physiknoten gemessen und nicht die Noten des Halbjahres in der die Intervention stattfand.

Bei der Konstruktvalidität wird kein manifestes Kriterium benötigt, mit dem eine Korrelation berechnet werden kann. Es werden theoriebasiert Hypothesen aus dem Konstrukt abgeleitet. Diese Hypothesen sagen vorher, wie die Testwerte ausfallen sollten. Wenn diese Vorhersage zutrifft, kann von einer hohen Konstruktvalidität ausgegangen werden.

Da die Expertenbefragung dafürspricht, dass die verwendeten Items alle Inhalte aus der Mechanik der E-Phase testen, lässt sich hier von einer guten Validität ausgehen. Die Nähe des Tests zum Force Concept Inventory (Hestenes et al., 1992) spricht ebenfalls für eine ausreichende Validität, da diese für den FCI häufig untersucht und bestätigt wurde (Halloun & Hestenes, 1985).

8.5 Lehrerfragebogen

Parallel zum Fragebogen für die teilnehmenden Schüler*innen wurde auch ein Fragebogen für Lehrkräfte erstellt. Dieser enthält verschiedene Bereiche. Zunächst werden Fragen zum Computereinsatz gestellt, um die Bekanntheit verschiedener Arten von Computereinsatz zu erfahren und zu untersuchen, ob die beiden in den Interventionen genutzten Methoden in den besuchenden Klassen bereits eingesetzt wurden. Dieser Teil des Fragebogens wurde von Wenzel (2018) übernommen. Eine Überprüfung der Qualitätskriterien für diesen Teil war deshalb nicht notwendig. Im nächsten Teil des Fragebogens geht es um die fachlichen Items. Die Lehrkräfte sollten zu jeder Frage beantworten, ob die Items Inhalte aus dem Bereich der Mechanik testen und sie zusätzlich hingehend ihrer Schwierigkeit auf einer siebenstufigen Likert-Skala von „zu leicht“ bis „zu schwierig“ bewerten.

Dabei werden die Items selbst auch von den Lehrkräften beantwortet, sodass auch deren Konzeptverständnis eingeschätzt werden kann. Die Items im fachlichen Teil stimmen dabei mit den Items aus den Schülerfragebogen überein. Daher ist keine eigenständige Pilotierung von Nöten.

8.6 Bildschirmvideos

Um die in Kapitel 7.2.2 beschriebenen Forschungsfragen beantworten zu können, wurden Bildschirmvideos der Partnerarbeit am Computer aufgenommen. Da sowohl die Arbeit mit der Software als auch die dazugehörigen Gespräche der Proband*innen von Interesse waren, wurden Bildschirmvideos in Kombination mit dem Ton der Unterhaltung aufgenommen. Das Filmen der Lernenden selbst war nicht von Interesse.

8.6.1 Auswahl des Versuchs für die Aufnahme der Bildschirmvideos

Insgesamt wurden in der Intervention vier Experimente durchgeführt. Aufgrund der sonst großen Datenmenge einer vollständigen Aufnahme wurde sich gegen eine Aufnahme der kompletten vier Versuche (~3,5 Stunden) entschieden. Daher wurde ein Versuch ausgewählt, der möglichst repräsentativ sein sollte und mit dem es optimal gelingen sollte, die in Kapitel 7.2.2 formulierten Forschungsfragen zu beantworten.

Im ersten Versuch der Intervention (Fall mit Luftreibung) wird im Plenum gearbeitet, wodurch keine bzw. weniger Gespräche zwischen den Lernenden zustande kommen. Der letzte Versuch (Kreisbewegung) ist besonders komplex, sodass dort die Modellierung nicht mehr von allen Proband*innen selbst durchgeführt wird. In diesem wird auch intensiver mit Hilfen gearbeitet, wodurch er sich von den anderen Versuchen unterscheidet. Letztlich wurde der dritte Versuch (schiefer Wurf) gewählt, da die Proband*innen so bereits zuvor ausreichend die Software kennengelernt haben. Es handelt sich bei diesem Versuch um einen schiefen Wurf einer Stahlkugel und einer Styroporkugel, womit auch die Luftreibung thematisiert wird. Zudem handelt es sich im Gegensatz zu Versuch 2 um eine zweidimensionale Bewegung.

8.6.2 Qualitative Inhaltsanalyse

8.6.2.1 *Qualitative vs. quantitative Inhaltsanalyse*

Qualitative Inhaltsanalyse bezeichnet keine konkret umrissene Forschungsmethode – vielmehr handelt es sich um einen Kanon von verschiedenen Forschungsmethoden, die

sich auf einem Spektrum zwischen qualitativer und quantitativer Inhaltsanalyse bewegen (Mayring, 2019), wobei das gesamte Spektrum unter dem Begriff der qualitativen Inhaltsanalyse subsumiert wird. Schreier (2014) unterscheidet bspw. acht verschiedene Varianten der qualitativen Inhaltsanalyse. Dazu zählen u. a. die zwei Basisformen inhaltlich-strukturierende Inhaltsanalyse (Kuckartz, 2018; Mayring, 2015; Schreier, 2012) und die qualitative Inhaltsanalyse durch Extraktion (Gläser & Laudel, 2013). Wenngleich die unterschiedlichen Varianten teilweise unterschiedliche Ziele verfolgen, ist ihnen gemeinsam, dass sie eine „Systematisierung und Interpretation von manifesten und latenten Kommunikationsinhalten unter Verwendung eines Kategoriensystems“ anstreben (Stamann et al., 2016). Qualitative Inhaltsanalyse zeichnet sich allgemein eher dadurch aus, dass das Erkenntnisinteresse häufig explorativ ist und die latente Textbedeutung im Mittelpunkt steht. Damit einhergehend wird die Kategorienbildung häufig induktiv vorgenommen. Quantitative Inhaltsanalyse wird eher eingesetzt, wenn vorher formulierte Hypothesen getestet werden und die manifeste Textbedeutung im Zentrum steht. Damit verbunden ist häufig eine deduktive Kategorienbildung. Dabei gibt es ebenfalls Mischformen bzw. Zwischenstufen dieser beiden Ausprägungen.

In dieser Forschungsarbeit ist das Ziel aus den vorliegenden Tonaufnahmen Antworten auf die Forschungsfragen abzuleiten. Es wird sich daher an der Variante der qualitativen Inhaltsanalyse mittels Extraktion und der strukturierenden qualitativen Inhaltsanalyse orientiert. Da Teile der Untersuchung des Datenmaterials hypothesentestend und andere eher explorativ sind, kommen unterschiedliche Arten der Kategorienbildung zum Tragen. Die Qualität der Zuordnung zu diesen Kategorien wird mit einem Kennwert für die Beurteilerübereinstimmung überprüft.

8.6.2.2 Erstellung eines Kategoriensystems

Aufgrund der Vielfältigkeit der Forschungsmethoden innerhalb der qualitativen Inhaltsanalyse gibt es kein Standardvorgehen und dementsprechend müssen auch die Kategorien des Kategoriensystems und die Kodierregeln an die vorliegenden Daten und das Material angepasst werden. Viele der üblichen Vorgehensweisen der qualitativen Inhaltsanalyse basieren auf dem Auswerten von Interviews, bei denen ein sehr hoher Anteil an Daten für die Forschungsfragen relevant sind. Im hier vorliegenden Fall handelt es sich jedoch um einen Mitschnitt der Arbeitsphasen in Videoform. Diese Tatsache in Kombination mit den Forschungsfragen führt dazu, dass nur ein gewisser Teil der getätigten Aussagen und Gespräche für das Forschungsvorhaben relevant sind, was wiederum heißt, dass ein Großteil der Gespräche nicht kodiert werden muss. Insgesamt muss zwischen einer Kodierung im Video und einer Kodierung im Transkript unterschieden werden. Hier werden alle

Aussagen, die ohne das Video verständlich sind und bei denen die Zeitkomponente keine Rolle spielt, im Transkript kodiert, welches auf Basis der Tonmitschnitte erstellt wurde, während alle Kategorien, in denen die Zeitdauer relevant ist oder es um im Video sichtbare Dinge geht, direkt im Video kodiert werden. Diese Funktionen ermöglicht das Programm „MAXQDA“, was für die gesamte Auswertung genutzt wurde.

Eine weitere grundsätzliche Unterscheidung wird zwischen deduktiver und induktiver Kategorienbildung vorgenommen. Da in den meisten Kategorien hypothesentestend und nicht hypothesengenerierend gearbeitet wurde, wurde auf deduktiv aus der Theorie abgeleitete Kategorien zurückgegriffen. Die Ausnahme stellt hier die Dimension der Schwierigkeiten inhaltlicher Art und mit der Software dar, da es sich um eine explorative Analyse handelt und aus der Theorie nicht bekannt ist, welche Probleme bei der Nutzung der jeweiligen Programme auftreten. Das Kategoriensystem wird in verschiedenen Dimensionen eingeteilt, in denen dann jeweils Kategorien und Subkategorien erstellt werden, da sich die Dimensionen in ihrer Vorgehensweise unterscheiden und so nicht nur ein homogenes Kategoriensystem bilden. Bei der Vorgehensweise wurde sich je nach Forschungsfrage verschiedenen Varianten nach Schreier (2014) bedient. Das vollständige Kategoriensystem ist im Anhang (Kapitel 17) einzusehen.

Dimension 1: Arbeitsphasen

Die Einteilung der Phasen dient dazu herauszufinden, wie zielgerichtet die Lernenden vorgehen. Die von der Lehrkraft gewünschten und in der Intervention eingeübten Phasen (Problemanalyse, Modellbildung, Evaluation etc., siehe Abb. 12) sind als Kategorien vorgegeben (deduktive Kategorienbildung). Dabei unterscheiden sich die Phasen in der Intervention der Modellbildung von derjenigen der Videoanalyse, sodass hier getrennte Kategorien für die beiden Gruppen angewendet werden. Bei der Modellbildung geht es vor allem darum, ob die Proband*innen alle Phasen durchlaufen, zielgerichtet vorgehen und das Modell zur Beantwortung der Abschlussfragen nutzen. Zudem wird das Vorgehen von besonders erfolgreichen und weniger erfolgreichen Gruppen verglichen und analysiert, wieviel Zeit in den in den jeweiligen Phasen verbracht wurde. Bei der Videoanalyse wird vor allem geschaut, welche Funktionen des Programms genutzt werden. Zwar wurden gewisse Darstellungsformen durch die Aufgaben vorgegeben, von Interesse ist aber, wie lange sich Lernenden mit den jeweiligen Darstellungsformen beschäftigen und ob sie die Messdaten für die Beantwortung der Abschlussfragen nutzen oder nach der Analyse eher intuitiv auf die Fragen antworten. Die Reihenfolge und das Vorhandensein der Phasen sind bei der Videoanalyse weniger von Interesse, da sie zwangsläufig aufeinander folgen müssen (Analyse, Darstellung, Interpretation, Auswertung, siehe Abb. 14).

Dimension 2: Schwierigkeiten

Relevant für den Einsatz der Software in der Schule ist, wie viele Schwierigkeiten beim Arbeiten mit der Software entstehen. Die Schwierigkeiten werden dabei nach Häufigkeit und Zeitdauer analysiert und zwischen beiden Gruppen verglichen. Dabei muss unterschieden werden, ob die Schwierigkeiten im Zusammenhang mit der Software stehen (z. B. Eingabe der Kräfte) oder einen physikalischen Hintergrund haben (z. B. Grapheninterpretation). Die Einteilung der Kategorien erfolgt dabei induktiv, da aus der Theorie nicht im Voraus bekannt ist, welche Schwierigkeiten beim Arbeiten mit der Software auftreten. Das Ziel dieser Dimension ist es, Handlungsempfehlungen für die Verbesserung physikdidaktischer Software im Bereich mathematischer Modellbildung und Videoanalyse abzuleiten und Empfehlungen zu geben, auf welche inhaltlichen Bereiche vor der Nutzung der Software besonders Wert gelegt werden sollte, um den Lernerfolg zu maximieren.

Dimension 3: Schülervorstellungen

Über das Vorhandensein von Schülervorstellungen ist in der Mechanik viel bekannt (Kap 3.2). Nicht bekannt ist, ob die beiden verwendeten Vorgehensweisen bestimmte Schülervorstellungen besonders provozieren. Das Vorgehen der Kategorienbildung ist hier deduktiv, da von einer Teilmenge der in der Theorie bekannten Schülervorstellungen ausgegangen wird, die zusätzlich durch den Fragebogen erfragt werden. Es ist nicht auszuschließen, dass auch andere Schülervorstellungen geäußert werden oder eine andere Einteilung möglich wäre. Dies ist aufgrund der Fülle an bekannten Schülervorstellungen im Bereich der Dynamik unvermeidbar. Ein Unterschied in der Häufigkeit der betrachteten Schülervorstellungen zwischen den Gruppen kann dennoch Aufschluss geben, inwiefern die Argumentationsrichtung bzw. Software hier eine entscheidende Rolle spielt.

Dimension 4: Argumentationsrichtung

Wie zuvor argumentiert, ist einer der grundlegenden Unterschiede zwischen den verwendeten Methoden die genutzte Argumentationsrichtung. In einem Fall werden Kräfte modelliert und daraus die Beschleunigung bestimmt. Aus der Beschleunigung wird dann die Geschwindigkeit und der Ort berechnet. Im anderen Fall werden die Messdaten des Orts eines Experiments analysiert, um daraus die Geschwindigkeit und dann die Beschleunigung zu berechnen. Aus der Beschleunigung kann dann die Summe der wirkenden Kräfte abgeleitet werden. Hinter der Analyse der Argumentation der Videos steht die Frage, ob sich dies in den von den Lernenden verwendeten Argumentationen widerspiegelt.

Dazu wird analysiert, wie die Lernenden die Abschlussfragen begründen. Konkret geht es darum, ob die gefragten Größen auf einer höheren, niedrigeren oder auf derselben Ebene begründet werden. Dabei folgt die Einteilung der Reihenfolge von hoch nach niedrig, von Kräften über die Beschleunigung und Geschwindigkeit zum Ort. Unterschieden wird dabei zwischen den kinematischen Größen und den Kräften. Um die Vergleichbarkeit zu gewährleisten, werden dabei nur die Argumentationen berücksichtigt, die beim Ausfüllen der Tabellen am Ende der jeweiligen Versuche angeführt werden, damit die Zahlen nicht durch die zuvor im Arbeitsheft gestellten Fragen verzerrt werden.

Dimension 5: Dynamik

Eng verwandt mit der vorigen Dimension ist die der Dynamik. Dahinter verbergen sich Argumentationen, die Kräfte und Bewegung miteinander in Verbindung bringen. Innerhalb dieser Dimension wird unterschieden, ob von Kräften auf die Bewegung (also das Verhalten des Objekts durch wirkende Kräfte erklärt wird) oder von der Bewegung auf die wirkenden Kräfte geschlossen wird (also das Vorhandensein gewisser Kräfte mit dem Verhalten des Objekts begründet wird). Auch hier ist die Arbeitshypothese, dass diese Reihenfolge durch die Art der Intervention beeinflusst wird, also ein Unterschied zwischen den Gruppen existiert. Dies würde bedeuten, dass Lehrkräfte diese Tatsache im Unterricht beachten und beide Richtungen entsprechend thematisieren müssten.

8.6.2.3 Gütekriterien der Qualitativen Inhaltsanalyse

Auch bei qualitativen Forschungsmethoden gibt es Gütekriterien, die erfüllt und nachgewiesen werden sollten, um die Qualität der Ergebnisse zu erhöhen. Nach einigen Autoren (Krippendorff, 1980; Mayring, 2015; Schreier, 2012) sind hier ebenfalls die Kriterien Objektivität, Reliabilität und Validität entscheidend, die aus dem Bereich der quantitativen Forschung kommen. Andere Autor*innen nennen Gütekriterien wie Verfahrensdokumentation, argumentative Interpretationsabsicherung, Regelgeleitetheit, Nähe zum Gegenstand, kommunikative Validierung und Triangulation (Mayring, 2002). In der Forschungspraxis werden die klassischen Gütekriterien Objektivität, Reliabilität und Validität allerdings am häufigsten herangezogen (Göhner & Krell, 2020). Ein genereller Konsens über die Notwendigkeit bestimmter Gütekriterien besteht nicht (Göhner & Krell, 2020). Die Gewährleistung und der Nachweis dieser Kriterien unterscheiden sich allerdings teilweise stark von anderen Forschungsmethoden, wie z. B. der Arbeit mit Fragebögen, und wird für das vorliegende Kategoriensystem nachfolgend diskutiert.

Objektivität

Objektivität ist das problematischste der drei Gütekriterien, da es fragwürdig ist, ob dies durch eine qualitative Forschungsmethode überhaupt erreicht wird. Manche Autor*innen lehnen die Nutzung dieses Kriterium deshalb für die qualitative Inhaltsanalyse vollständig ab (Schreier, 2012). Oft wird allerdings auch die Intersubjektivität mit der Objektivität gleichgesetzt (Mayring, 2015). Die häufig berichtete Kennzahl für eine InterCoderreliabilität wird von Mayring (2015) als Maß für die Auswerteobjektivität gesehen, während Schreier (2012) dies der Reliabilität zuordnet. Teilweise wird sie sogar der Validität zugeordnet (Göhner & Krell, 2020). Von einem Nachweis der Objektivität wird hier dementsprechend abgesehen. Die InterCoderreliabilität wird im folgenden Kapitel diskutiert.

Reliabilität

Im Bereich der Reliabilität gibt es üblicherweise zwei grundlegende Arten der Reliabilität. Diese sind die InterCoderreliabilität, welche eine Reproduzierbarkeit der Codierungen durch verschiedene Personen darstellt, und die Intracoderreliabilität, welche eine Reproduzierbarkeit der Kodierungen von einer Person in einem zeitlichen Abstand meint. Es gibt verschiedene Koeffizienten, um diese Reliabilitäten zu quantifizieren. Diese Koeffizienten ermöglichen eine Signifikanzprüfung und eine Vergleichbarkeit mit anderen Studien. Diese Vergleichbarkeit ist, ähnlich wie bei Cohens d , jedoch immer nur eingeschränkt gegeben. Die Wahl des Koeffizienten hängt von der Zielsetzung der Studie und von der Wahl der Kategorien bzw. der Art des Ratings ab. Auch wenn die Reliabilität der zugrundeliegenden Skala zugeschrieben wird, drückt sie eher die Zuverlässigkeit der Benutzung dieser Skala aus und ist wiederum von den Beurteilenden abhängig (Wirtz & Caspar, 2002).

Der Koeffizient zur Beurteilung der Reliabilität bzw. Übereinstimmung hängt in erster Linie davon ab, ob Text- bzw. Videostellen nominalskalierten Kategorien zugeordnet oder auf intervallskalierten Ratingskalen eingeordnet werden. Da die angefertigten Videos ausschließlich in Kategorien eingeteilt werden, ist hier konsequenterweise von Übereinstimmungsmaßen (also Beurteilerübereinstimmung) zu reden, da die Interraterreliabilität nur für mindestens intervallskalierte Ratings definiert ist (Wirtz & Caspar, 2002).

Für die Überprüfung der Beurteilerübereinstimmung wurden nach der Schulung eines weiteren Kodierers unabhängig voneinander vier Videos (der insgesamt 21) kodiert und die Ergebnisse verglichen. Da sich die Dimensionen in der Vorgehensweise teilweise stark unterscheiden, werden für die Dimensionen getrennt die Kennwerte für die

Beurteilerübereinstimmung berichtet. Für die Berechnung wurde die Software „MAXQDA 2020“ verwendet. Für die Bestimmung der Beurteilerübereinstimmung muss zunächst die Art der Übereinstimmung festgelegt werden. Hier wird die Überlappung der Codes betrachtet, da dies ein strengeres Kriterium ist als lediglich das Vorhandensein oder die Häufigkeit der Codes zu berücksichtigen. Da die Codesegmente teilweise sehr kurz sind und es dementsprechend einen großen Einfluss hat, wo die Codierung genau beginnt, was speziell im Video (durch Pausen o. ä.) mit einer gewissen Unsicherheit verbunden ist, wurde die Mindestüberlappung von 80 % gewählt, damit eine Übereinstimmung also solche gezählt wird.

Als Kennwert zur Berechnung der Beurteilerübereinstimmung wird das Kappa nach Brennan und Prediger (1981) genutzt. Für die Bestimmung dieses Kennwerts werden in einer Vierfeldertafel alle kodierten Segmente dargestellt. Dabei wird pro Kodierer zwischen kodierten und nicht kodierten Segmenten unterschieden. Ein Feld beinhaltet also alle Segmente, bei denen es eine Übereinstimmung gibt, die über die prozentuale Überlappung ermittelt wird (a), ein Feld beinhaltet alle Segmente, die von Kodierer 1 aber nicht von Kodierer 2 kodiert wurden (b), ein Feld genau umgekehrt (c) und das letzte Feld beinhaltet alle Segmente, die von beiden nicht kodiert wurden (d). Dies ist bei der vorliegenden Berechnung allerdings immer leer, da nur kodierte Segmente in die Berechnung einbezogen werden, da ein Großteil des Materials aus bereits beschriebenen Gründen nicht relevant für die Forschungsfragen ist. Nun wird die Anzahl an Übereinstimmungen durch die Gesamtanzahl an kodierten Segmenten geteilt:

$$P(\text{beobachtet}) = \frac{a}{a + b + c}$$

Damit eine mögliche Übereinstimmung nicht zufällig zustande kommt, wird die Wahrscheinlichkeit einer zufälligen Übereinstimmung in folgender Weise berücksichtigt:

$$P(\text{Zufall}) = \frac{1}{\text{Anzahl Codes}}$$

Folglich ergibt sich der Wert für Kappa:

$$\kappa = \frac{P_b - P_z}{1 - P_z}$$

So ergeben sich für die Übereinstimmung folgende Cohens Kappa (Tab. 19).

Tab. 19: Werte für zufallskorrigierte Cohens Kappa für alle Dimensionen.

Vorgehensweise	Schwierigkeiten	Schülervorstellungen	Argumentationsrichtung	Dynamik
0,70	-	0,83	0,73	0,82

Der Wert für Cohens Kappa sollte nach Greve und Wentura (1997) über 0,60 liegen, um als gut bezeichnet werden zu können, während Werte zwischen 0,40 und 0,60 noch annehmbar sind. Landis und Koch (1977) sprechen bei Werten zwischen 0,40 und 0,60 von einer mittelmäßigen, bei 0,60 bis 0,80 von einer guten und bei über 0,80 von einer fast vollkommenen Übereinstimmung.

Problematisch war insbesondere der Bereich der Vorgehensweise. Hier ergeben sich für die Modellbildung ($\kappa = 0,75$) und für die Videoanalyse ($\kappa = 0,57$) unterschiedliche Werte. Nach anfänglich geringer Übereinstimmung wurde das Kodiermanual überarbeitet, woraufhin eine bessere Übereinstimmung erzielt werden konnte. Da die gewählte notwendige Überlappung für eine Übereinstimmung selbst einer gewissen Willkür unterliegt, sind auch die Kennwerte für die Beurteilerübereinstimmung nur bedingt aussagekräftig. In den Dimensionen „Schülervorstellungen“, „Argumentationsrichtung“ und „Dynamik“ konnten jedoch gute Übereinstimmungen erzielt werden. In der Dimension „Schwierigkeiten“ wurde das gesamte Material zur induktiven Kategorienbildung genutzt und kein Maß für die Beurteilerübereinstimmung berechnet.

Validität

Um das Gütekriterium der Validität zu erfüllen, gibt es an die Kategorien bzw. das Kategoriensystem verschiedene Anforderungen (Vgl. Krippendorff, 1980; Schreier, 2012). Einerseits sollten durch die Kategorien alle für die Forschungsvorhaben relevanten Informationen erfasst werden, während sie an Forschungsgegenstand und Forschungsfrage orientiert sind. Andererseits sollten sie an Limitationen der Studie (wie zeitliche Beschränkungen) angepasst werden. Die Validität wird also durch die Erstellung eines angemessenen Kategoriensystems gewährleistet und lässt sich nicht durch einen bestimmten Kennwert ausdrücken. In Kapitel 8.6.2.2 wird begründet, in welcher Art und Weise das Kategoriensystem erstellt wird, um damit eine möglichst hohe Validität zu erreichen.

Weitere Qualitätskriterien

Wenn man sich nicht an der klassischen Testtheorie orientiert, können weitere Gütekriterien wie z. B. Verfahrensdokumentation, argumentative Interpretationsabsicherung, Regelgeleitetheit, Nähe zum Gegenstand, kommunikative Validierung und Triangulation für die inhaltlich-strukturierende qualitative Inhaltsanalyse formuliert werden (Mayring, 2008). Zur Verfahrensdokumentation gehören z. B. die Entwicklung von Kodierregeln und -manualen und die Begründung der Form der Datenaufnahme (Göhner & Krell, 2020).

Lamnek und Krell (2016) nennen Offenheit, Kommunikativität, Naturalistizität und Interpretativität als Merkmale einer qualitativen Inhaltsanalyse. Diesen Kriterien soll bei der Auswertung und der Diskussion der Ergebnisse sowie durch die Darstellung des Kategoriensystems, so gut es geht, nachgekommen werden. Für die Gewährleistung dieser Qualitätskriterien werden am häufigsten die Modi der Kategorienbildung und Interraterübereinstimmung genannt (Göhner & Krell, 2020), die bereits in den vorigen Kapiteln diskutiert wurden.

Weitere Kriterien, die in der Literatur gefordert werden, sind Eindimensionalität der Kategorien, Eindeutigkeit der Kategorien und dass die Kategorien nicht leer sein sollen. Diese werden teilweise kontrovers diskutiert. Insbesondere das letzte Merkmal ist nicht unumstritten. In dem Fall dieser Studie kommt es vor, dass Kategorien leer bleiben (beispielsweise eine in der Theorie bekannte Schülervorstellung, die in den Gruppen aber nicht auftritt). Die Eindeutigkeit der Kategorien ist hier innerhalb jeder Dimension i. d. R. gegeben, einzelne Textstellen können aber für verschiedene Dimensionen interessant sein und somit in verschiedenen Kategorien auftreten, die sich in unterschiedlichen Dimensionen befinden.

III. Evaluation der Studie

9 Ergebnisse zum Konzeptverständnis

Die Durchführung der Studie fand im ersten Halbjahr 2020 statt und es waren bereits Termine für die Teilnahme von 33 Klassen mit etwa 630 Schüler*innen ausgemacht. Durch die Schulschließungen und die folgenden Beeinträchtigungen durch die Corona-Krise konnten zunächst nur 17 dieser 33 Klassen teilnehmen. Aufgrund dessen wurde die Durchführung der Studie um ein Schuljahr verlängert, sodass im folgenden Schuljahr erneut Klassen an der Studie teilnahmen. Aufgrund der weiter herrschenden Einschränkungen und des hohen Infektionsgeschehens u. a. im Winter 2020/21 kamen allerdings nur vier weitere Klassen hinzu, sodass die Studie mit insgesamt 21 Klassen durchgeführt wurde. Aus diesen 21 Klassen nahmen $N = 274$ Lernende (davon 120 weiblich) vollständig an Vortest, Intervention und Nachtest teil. An Intervention und Nachtest nahmen 312 Schüler*innen teil, Vortests lagen von 309 Schüler*innen vor. Ein Fragebogen wurde dann als vollständig angesehen, wenn nahezu alle Fragen beantwortet wurden. Im fachlichen Teil wurden einzelne ausgelassene Items als falsch bewertet. Wenn die letzten Items (mindestens vier) des Vortests alle nicht beantwortet wurden, wurde der Proband oder die Probandin entfernt und nicht ausgewertet, da dies dafür spricht, dass die Person den Test aus zeitlichen Gründen nicht abschließen konnte. Im Nachtest hatten die Proband*innen beliebig viel Zeit. Im Bereich der affektiven Merkmale wird ein nicht angekreuztes Item als nicht beantwortet angesehen. Bei der Betrachtung der durchschnittlichen Punktzahl einer Skala geht dieses Item nicht in die Berechnung mit ein. Je nach Auswertungsverfahren müssen aber auch solche Proband*innen mit fehlenden Werten („missing values“) teilweise ausgeschlossen werden.

Mit den vollständigen 274 Datensätzen (je $N = 137$ aus den Gruppen der mathematischen Modellbildung und Videoanalyse) sollen die Forschungsfragen in Bezug auf Konzeptverständnis der Newtonschen Mechanik, Ansichten zu Modellen in der Physik, Bewertung der Interventionen und Cognitive Load beantwortet werden. Dabei wird zunächst darauf eingegangen, ob die beiden Interventionen wirksam waren und für das Lernen der Dynamik geeignet sind, bevor auf spezifische Unterschiede zwischen den Gruppen eingegangen werden soll. Außerdem werden zusätzlich die Unterschiede zwischen Grundkursen und Leistungskursen mit in die Analyse einbezogen. An hessischen Gymnasien gibt es Schulen, die in der E-Phase bereits ein Kurssystem haben, und Schulen, bei denen dieses erst in der 12. Klassenstufe (nach G9, Q1) eingeführt wird. Es wird dementsprechend eine Unterteilung zwischen allen Leistungskursen auf der einen Seite und allen Klassen, die noch kein Kurssystem hatten, und Grundkursen auf der anderen Seite vorgenommen. An der Studie nahmen verhältnismäßig viele Leistungskurse (sieben von 21)

teil, da diese durch mehr Unterrichtsstunden schneller im Stoff vorankommen und dadurch früher im Schuljahr, also vor den Schulschließungen, teilnehmen konnten.

9.1 Power Analysis

Da nur knapp die Hälfte der ursprünglich geplanten 630 Schüler*innen an der Studie teilnehmen konnte, soll eine sogenannte „Power Analysis“ zunächst überprüfen, ob man die erwarteten Effekte auf Basis eines 5 %-Signifikanzniveaus mit einer solchen Stichprobengröße überhaupt nachweisen kann. Dabei soll nicht auf jeden Aspekt der Analyse der Daten eingegangen werden. Es soll lediglich für die Haupteffekte eine solche Analyse durchgeführt werden. In der Folge werden verschiedene statistische Methoden angewendet, u. a. t-Tests und ANOVAs (einfaktorielle Varianzanalysen). Um die notwendige Stichprobengröße zu schätzen, wird von üblichen Werten für das Signifikanzniveau α (5 %) und die Teststärke $1 - \beta$ (80 %) ausgegangen. Bei einem zweiseitigen, gepaarten t-Test bräuchte man für diese Randbedingungen eine Stichprobengröße von $N = 33$ um einen Effekt von $d = 0,5$ nachzuweisen. Für einen ungepaarten t-Test mit den sonst gleichen Bedingungen sind $N = 64$ Proband*innen pro Gruppe notwendig. Andersherum kann man mit den $N = 137$ Proband*innen pro Gruppe Effekte von der Größe $d = 0,24$ (gepaart) bzw. $d = 0,34$ (ungepaart) oder größer nachweisen. Letztere Situation sieht graphisch dargestellt folgendermaßen aus (Abb. 23).

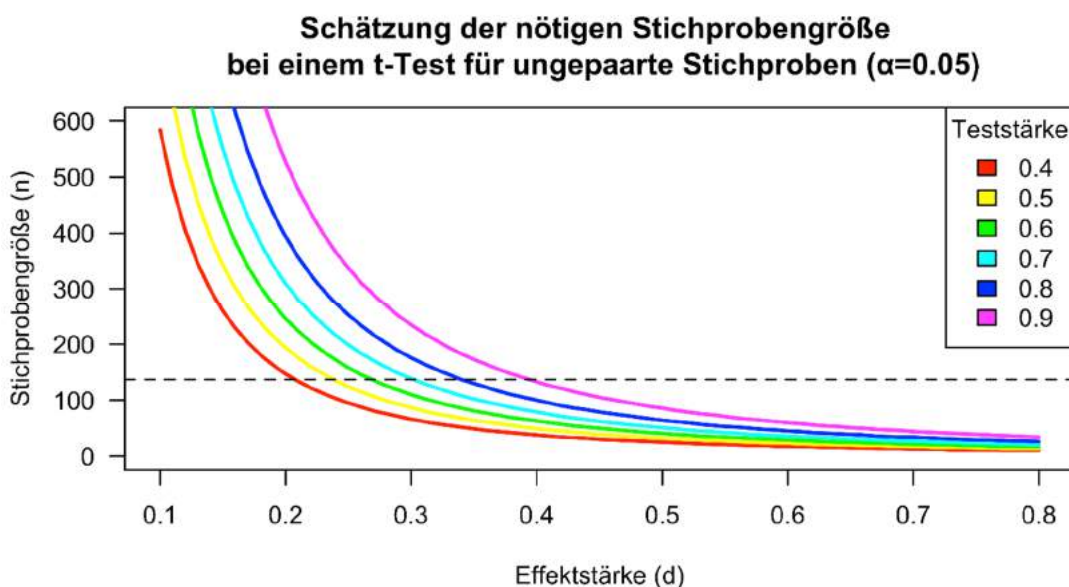


Abb. 23: Darstellung der notwendigen Stichprobengröße in Abhängigkeit der Größe des nachzuweisenden Effekts bei verschiedenen Teststärken bei einem t-Test für ungepaarte Stichproben. Die gestrichelte Linie gibt die hier vorliegende Stichprobengröße von $N=137$ an.

Die gestrichelte Linie gibt die vorliegende Stichprobengröße von $N = 137$ pro Gruppe an. Sichtbar wird, welche Effekte in Abhängigkeit der geforderten Teststärke mit dieser Stichprobengröße mindestens nachgewiesen werden können.

Ähnlich kann bei einer ANOVA die notwendige Stichprobengröße bestimmt werden. Hier spielt die Anzahl der Gruppen eine Rolle, die verglichen werden. In der Auswertung dieser Daten werden bis zu vier Gruppen miteinander verglichen. Bei diesen vier Gruppen, einem Signifikanzniveau von $p = 5\%$ und einer Teststärke von $1 - \beta = 80\%$ werden 45 Proband*innen pro Gruppe benötigt, um einen mittelgroßen Effekt (Effektstärke f für Varianzanalysen (Salkind, 2010), $f = 0,25$) zu erkennen. Umgekehrt kann, wenn man von einer Verteilung in gleichgroße Gruppen ausgeht, mit diesen Randbedingungen ein Effekt von $f = 0,20$ nachgewiesen werden. Bei einem allgemeinen linearen Modell (Effektstärke f^2 für multiple Regressionen (Cohen, 1988)), in dem beispielsweise vier Koeffizienten berücksichtigt werden, kann bei der vorliegenden Stichprobengröße und abermals einem Signifikanzniveau von 5% und einer Teststärke von 80% ein kleiner Effekt von $f^2 = 0,04$ nachgewiesen werden. In der Folge wird nicht an jeder Stelle erneut auf die „Power Analysis“ verwiesen. Die Stichprobengröße scheint aber für die meisten Analysen ausreichend groß zu sein und kann dementsprechend mittlere bis große Effekte gut nachweisen.

9.2 Wirksamkeit

Zum Testen des Konzeptverständnisses im Bereich der Mechanik wurde ein Test mit insgesamt 23 Items (davon drei zweistufig, siehe Kapitel 8.4.2) eingesetzt, wovon vier Items das dritte Newtonsche Gesetz betreffen, welches in der Intervention inhaltlich keine Rolle spielte. Vier weitere Items gehören zu einer kinematischen Skala (Beschleunigungsdiagramme), die restlichen Items teilen sich auf dynamische Skalen auf. Mit diesen Items können also verschiedene Subskalen der Newtonschen Mechanik gemessen werden und durch das Antwortverhalten auch auf Schülervorstellungen geschlossen werden. Bevor ein Vergleich zwischen beiden Interventionen gezogen wird, soll die Wirksamkeit der beiden Interventionen einzeln betrachtet werden.

Um die Wirksamkeit einer Intervention zu messen, ist ein Vergleich von Vor- und Nachtestergebnissen nötig. Dabei sollte nicht nur betrachtet werden, wie groß dieser Unterschied ist und ob sich diese Ergebnisse signifikant unterscheiden, sondern auch mit welcher Effektstärke eine gemessene Veränderung einhergeht und wie groß der relative Lernzuwachs ist. Effektstärken sind dabei vom eingesetzten Test abhängig und lassen sich nicht ohne Weiteres mit anderen Studien vergleichen. Dennoch wurde von Cohen (1988)

eine Einteilung über die Höhe von Effektstärken festgelegt und der Vergleich der Effektstärken mit ähnlichen Studien kann einen Anhaltspunkt geben, wie effektiv eine Intervention war (Cohen, 1988). Die Effektstärke setzt die Änderung zwischen Vor- und Nachtest mit verschiedenen Streuungsmaßen in der Stichprobe in Relation. Der relative Lernzuwachs ist der gemessene Lernzuwachs bezogen auf den maximal möglichen Lernzuwachs. Dieser wurde zuerst von Hake vorgeschlagen (Hake, 1998). Insbesondere auch in der physikdidaktischen Forschung gibt es eine Diskussion darüber, welcher dieser Werte für den Lernerfolg aussagekräftiger und fairer ist (Coletta & Steinert, 2020; Nissen et al., 2018). Demzufolge wird für beide Interventionen sowohl eine Effektstärke nach Cohen als auch ein normalisierter Lernzuwachs nach Hake angegeben.

9.2.1 Mathematische Modellbildung

An der Intervention zur mathematischen Modellbildung nahmen 137 Schüler*innen teil, von denen Vor- und Nachtest vollständig und verwertbar waren. In den 19 Items, die Inhalte testen, die zumindest mittelbar Teil der Intervention waren, wurden im Vortest im Schnitt 6,54 Punkte (34,4 %) und im Nachtest 10,1 Punkte (52,9 %) erreicht. Der durchschnittliche Lernzuwachs betrug 3,50 Punkte (18,4 Prozentpunkte). Da es insbesondere um den Lernerfolg im Bereich Dynamik gehen soll, auf den die Intervention primär abzielte (Beschleunigungsdiagramme wurden zwar genutzt, waren aber nicht Lernziel der Intervention), werden in der Folge lediglich die drei Skalen aus der Dynamik betrachtet, in denen insgesamt 15 Punkte zu erreichen waren. Hier wurden im Vortest durchschnittlich 5,30 Punkte (35,3 %) und im Nachtest durchschnittlich 8,58 Punkte (57,2 %) erreicht. Der durchschnittliche Lernzuwachs betrug 3,28 Punkte (21,9 Prozentpunkte). Der Boxplot der Ergebnisse im Vor- und Nachtest ist in Abb. 24 dargestellt.

Dabei wurde der Boxplot mit den üblichen Regeln erstellt. 50 % aller Werte liegen innerhalb der Box. Der horizontale Strich gibt den Median an, die Antennen zeigen das Maximum und Minimum, es sei denn, es liegen Werte außerhalb des Intervalls $[Q1 - 1,5 \cdot IQR, Q3 + 1,5 \cdot IQR]$. Diese werden gesondert ausgegeben und die Antennen zeigen dann den äußersten Wert innerhalb der genannten Intervallgrenzen an. Dabei stehen Q1 und Q3 für das erste bzw. dritte Quartil und IQR für die Differenz zwischen Q3 und Q1 (Chambers, 1983).

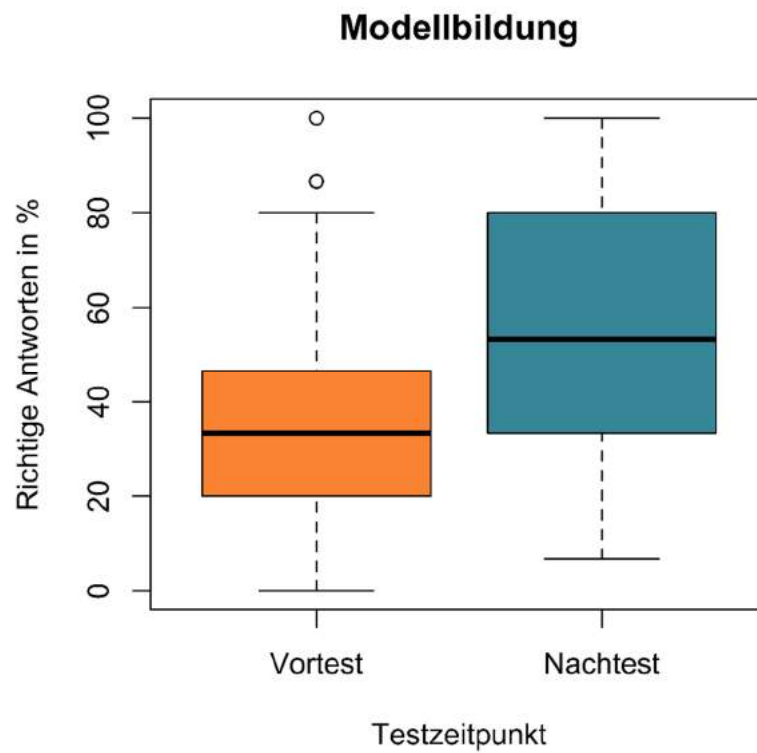


Abb. 24: Boxplot der richtigen Antworten in Prozent der Gruppe Modellbildung im Bereich der Dynamik im Vor- und Nachtest.

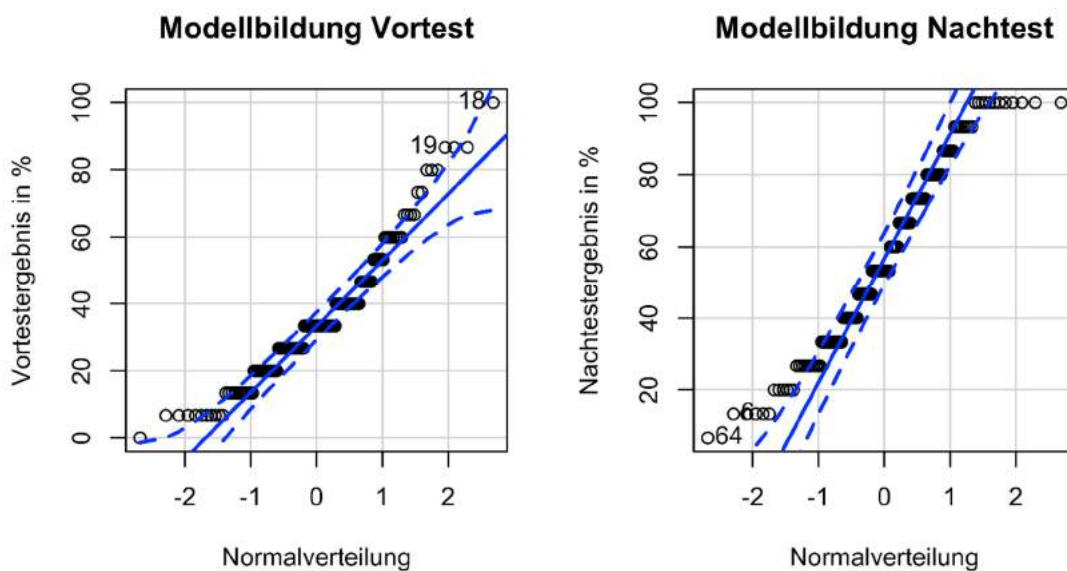


Abb. 25: Q-Q-Plot der prozentualen Punktzahlen in Vor- und Nachtest in der Gruppe der Modellbildung.

Um zu überprüfen, ob die beiden Ergebnisse sich unterscheiden, muss für die Wahl des richtigen Tests zunächst die Normalität der Daten überprüft werden. Da der Shapiro-Wilk-Test bei größeren Stichproben sehr stark auf Ausreißer reagiert und zu Überprüfung der Normalität dementsprechend meist zu streng ist (Field, 2013), wird hier auf einen Q-Q-Plot und eine optische Überprüfung der Daten in einem Histogramm zurückgegriffen. In dem Q-Q-Plot werden die Quantile der vorliegenden Daten über einer theoretischen Verteilung (hier Normalverteilung) geplottet. Wenn diese auf einer annähernd geraden Linie liegen, kann von einer Normalverteilung ausgegangen werden (Abb. 25).

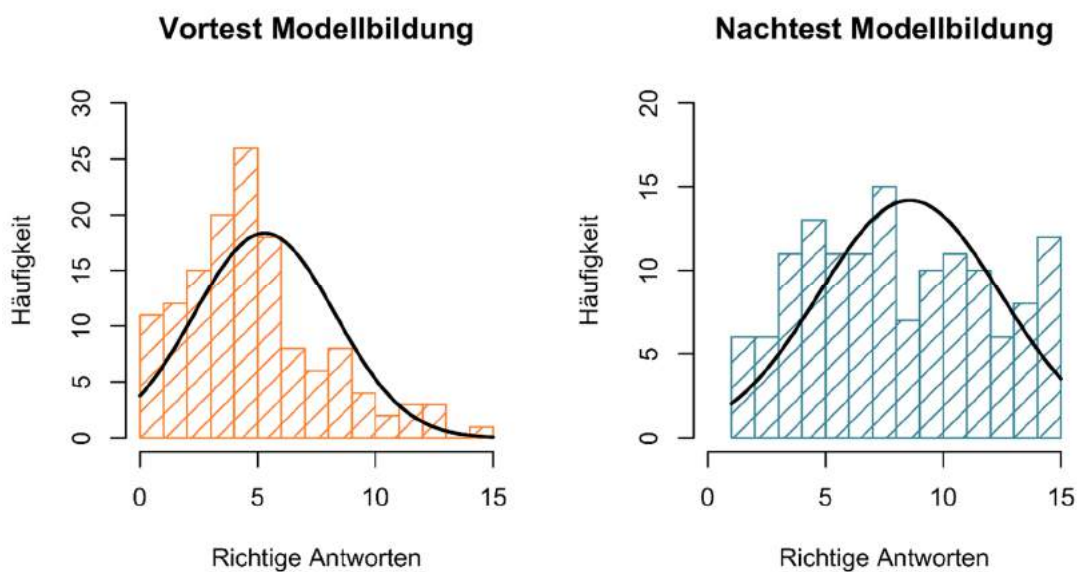


Abb. 26: Histogramm der Punktzahlen (max. 15) in Vor- und Nachtest in der Gruppe der Modellbildung.

Die durchgezogene Linie gibt dabei die Gerade an, auf der eine perfekt normalverteilte Stichprobe liegen würde, die gestrichelten Linien geben 95 %-Konfidenzintervalle an. Eine optische Überprüfung mithilfe eines Histogramms kann den Eindruck der obigen Plots bestätigen und einen ersten Eindruck über die Verteilung der Punktzahlen vermitteln (Abb. 26).

Die Punktzahlen selbst (insbesondere der Nachtest) sind nicht als normalverteilt anzusehen. Die Differenzen der Punkte in Vor- und Nachtest (Residuen) ergeben im Q-Q-Plot und Histogramm folgendes Bild (Abb. 27).

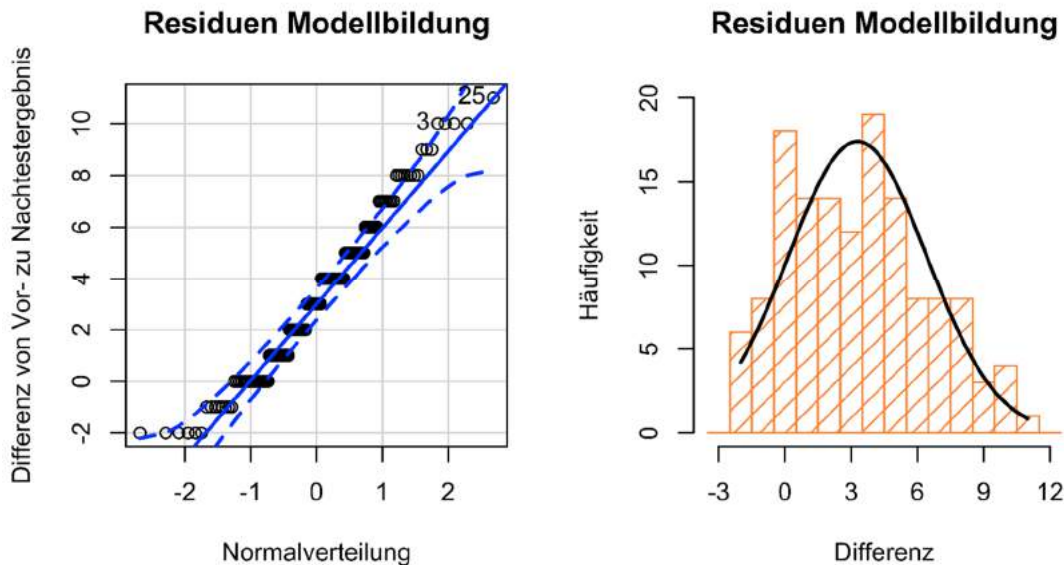


Abb. 27: Q-Q-Plot und Histogramm der Residuen in der Gruppe der Modellbildung.

Da für den gepaarten t-Test nur die Residuen normalverteilt sein müssen (Bortz & Weber, 2005), kann auf Basis dieser Plots von einer ausreichenden Normalverteilung der Daten ausgegangen werden. Die genannte Bedingung ist zudem vor allem für kleine Stichproben (unter 30 Paare) wichtig, wobei diese Auswertung mit $N = 137$ Paaren darüber liegt. Als Absicherung und da der genannte Shapiro-Wilk-Test die Normalverteilungsannahme bei den dargestellten Daten ablehnt, werden aber dennoch auch nicht-parametrische Tests durchgeführt. Bei allen weiteren parametrischen Verfahren, die angewendet werden, wird außerdem überprüft, wie stabil diese gegenüber der Normalverteilungsannahme sind. Also kann zunächst ein t-Test für gepaarte Stichproben für die Überprüfung der Signifikanz genutzt werden. Der t-Test fällt dabei höchst signifikant aus ($t = 1$; $df = 136$; $p < 0,001$), weswegen von einem Unterschied zwischen Vortestergebnis und Nachtestergebnis auszugehen ist.

Um ein Maß dafür zu bekommen, wie groß die Verbesserung von Vor- zu Nachtest war, wird zunächst die Effektstärke nach Cohen (1988) berechnet. Die Effektstärke setzt die Unterschiede im arithmetischen Mittel zweier Stichproben in Relation zu der Standardabweichung. Da die Standardabweichung nicht zu beiden Zeitpunkten identisch ist, kann nach Cohen (1988) die sogenannte gepoolte Standardabweichung verwendet werden. Die Effektstärke ist nach Nissen et al. (2018) das bevorzugte Maß, um die Effektivität einer Intervention zu bestimmen, da nach ihrer Meinung der normalisierte Zuwachs, welcher

danach besprochen wird, zu stark durch den Vortest beeinflusst wird. Cohens d ist nach Cohen (1988) allgemein definiert als:

$$\delta = \frac{\mu_{nach} - \mu_{vor}}{\sigma}$$

Der beste verfügbare Schätzer dafür ist:

$$d = \frac{M_{nach} - M_{vor}}{SD}$$

M_{vor} und M_{nach} stehen für die Mittelwerte der Stichproben in Vor- und Nachtest, SD für die Standardabweichung. Dabei gibt es verschiedene Möglichkeiten die Standardabweichung zu berechnen. Bei gleicher Gruppengröße kann die gepoolte Standardabweichung genutzt werden, die folgendermaßen definiert ist:

$$SD_{pooled} = \sqrt{\frac{(SD_1^2 + SD_2^2)}{2}}$$

Dabei sind SD_1 und SD_2 die Standardabweichungen der beiden Testzeitpunkte. Für die Berechnung der Effektstärke eines gepaarten t-Tests, welcher hier vorliegt, wird von anderen Autor*innen allerdings empfohlen Cohens d_z zu nutzen (Bortz & Döring, 2006). Dieses ist definiert als:

$$d_z = \frac{M_{diff}}{SD_{diff}}$$

Dabei ist M_{diff} der Mittelwert der Differenzen der Testergebnisse von Vor- zu Nachtest der Einzelpersonen und SD_{diff} die Standardabweichung dieser Unterschiede in folgender Form:

$$SD_{diff} = \sqrt{SD_1^2 + SD_2^2 - 2rSD_1SD_2}$$

Dabei steht r für die Korrelation zwischen den beiden Messwertreihen. Diese Effektstärke ist hier $d_z = 1,04$, was einem großen Effekt entspricht (Cohen, 1988).

Nach Morris und DeShon (2002) ist die angemessene Effektstärke für eine Studie mit Messwiederholung d_{rm} . Diese kombiniert die vorher beschriebene Effektstärke und führt einen zusätzlichen Korrekturterm ein. Sie ist definiert als:

$$d_{rm} = \frac{M_{diff}}{\sqrt{SD_1^2 + SD_1^2 - 2r \cdot SD_1 \cdot SD_1}} \cdot \sqrt{2(1-r)}$$

Diese Effektstärke ist hier $d_{rm} = 0,93$, was ebenfalls einem großen Effekt entspricht.

Diese Klassifikation der Effektstärke setzt eine gewisse Vergleichbarkeit voraus, die nicht uneingeschränkt gegeben ist, da die Effektstärke unter anderem von dem eingesetzten Test abhängt. Dennoch wird häufig ein Vergleich zu anderen Studien gezogen.

Coletta und Steinert (2020) argumentieren jedoch für die Berechnung vom normalisierten Zuwachs, um die Effektivität einer Intervention zu überprüfen – insbesondere beim Einsatz der klassischen Tests in der Mechanik, wie FCI oder FMCE, die Ähnlichkeiten zu dem hier verwendeten Test aufweisen. Hake (1998) schlägt dabei zunächst eine Berechnung der Mittelwerte vor, woraus dann der Kennwert g berechnet wird:

$$g = \frac{post - pre}{100\% - pre}$$

Dabei sind pre und $post$ die Mittelwerte auf Kohortenebene zu den beiden Zeitpunkten. Besser ist eine Berechnung des normalisierten Zuwachses auf individueller Ebene (Coletta & Steinert, 2020). Dieser berechnet sich analog wie g , nur gehen die individuellen Prä- und Posttestergebnisse in die Gleichung ein und der Mittelwert wird am Ende berechnet. Dieser wird in der Folge mit $\langle g \rangle$ bezeichnet.

Der normalisierte Zuwachs ergibt hier einen Wert von $g = 0,338$ bzw. $\langle g \rangle = 0,342$. Hake selbst berichtete in einer Metastudie von einem durchschnittlichen $g = 0,23 \pm 0,04$ in klassischen einsemestrigen Physikeinführungskursen an US-amerikanischen Universitäten und von $g = 0,48 \pm 0,14$ für Kurse mit interaktiver Beteiligung, die jeweils den FCI-Test nutzten. Diese Kurse fanden dabei immer über ein ganzes Semester statt und sind auch aufgrund der anderen Zielgruppe nicht vergleichbar.

Eine Alternative zur Berechnung des normalisierten Lernzuwachses auf Klassenebene haben Marx und Cummings (2007) vorgeschlagen, den sie normalisierte Änderung („normalized change“, c) nannten.

Die Berechnung von c ist wie folgt definiert:

$$c = \begin{cases} \frac{post - pre}{100\% - pre} & \text{wenn } post > pre \\ 0 & \text{wenn } post = pre \\ \frac{post - pre}{pre} & \text{wenn } post < pre \end{cases}$$

Dabei werden pre und $post$ der einzelnen Proband*innen genutzt. Zudem werden alle Proband*innen nicht berücksichtigt, die sowohl im Pretest als auch im Posttest entweder 0 % oder 100 % richtig haben. Dabei sind pre und $post$ die jeweiligen Punktzahlen in Prozent. Hier ergibt sich $c = 0,34$.

Außerdem kann statt dem parametrischen t-Test für gepaarte Stichproben auch der nicht-parametrische Wilcoxon-Vorzeichen-Rang-Test für gepaarte Stichproben genutzt werden. Aus dieser Teststatistik lässt sich r als Maß für die Effektstärke berechnen. Für die vorliegenden Daten ergibt sich ein Z von 8,442. Nach Tomczak und Tomczak (2014) kann daraus das r in folgender Weise bestimmt werden:

$$r = \frac{Z}{\sqrt{n}}$$

Zusammengefasst kann der Effekt der Intervention mit folgenden Kennwerten beschrieben werden (Tab. 20):

Tab. 20: Kennzahlen für die Effektstärke des Unterschieds zwischen Vor- und Nachtest in der Gruppe der Videoanalyse.

Kennzahl	Wert	Größe des Effekts
Cohens d_z	1,04	Groß (Cohen, 1988)
Cohens d_{rm}	0,93	Groß (Cohen, 1988)
Hakes $\langle g \rangle$	0,34	Mittel (Nissen et al., 2018)
Normalized change c	0,34	Mittel (Nissen et al., 2018)
Pearsons r	0,74	Groß (Cohen, 1988)

Der Effekt der Intervention ist als groß einzustufen. Auch wenn die Effektstärke immer vom Test abhängig ist, deuten die verschiedenen Kennzahlen auf einen großen

Lernzuwachs hin, der insbesondere erfreulich ist, da die Intervention nur wenige Stunden ging und nach dem eigentlichen Unterricht der Newtonschen Gesetze in der Schule stattfand.

Betrachtung der inhaltlichen Dimensionen

Da die explorative Faktorenanalyse vier Faktoren ergeben hat, zu denen noch eine weitere Skala aus dem FCI-Test übernommen wurde, können nun die damit verbundenen inhaltlichen Dimensionen in Vor- und Nachtest betrachtet werden. Diese Inhaltsbereiche sind: „Beschleunigungsdiagramme“ (Skala 1), „1. Newtonsches Gesetz“ (Skala 2), „Von Kräften zu Bewegung argumentieren“ (Skala 3), „Von der Bewegung zu den Kräften argumentieren“ (Skala 4) und „3. Newtonsches Gesetz“ (Skala 5). Um den richtigen statistischen Test auf Unterschiede zwischen den Mittelwerten zu finden, muss zunächst die Normalverteilung der Residuen überprüft werden (Abb. 28).

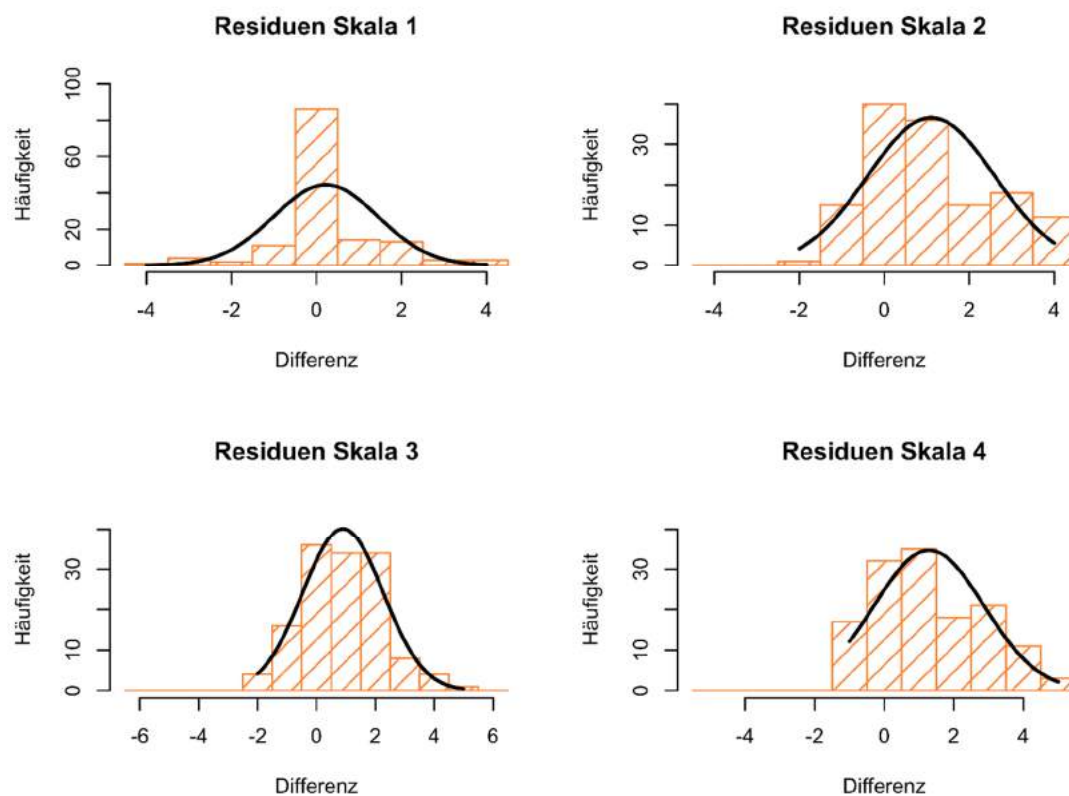


Abb. 28: Residuen der vier inhaltlichen Skalen in der Gruppe der Modellbildung.

Das Nachtestergebnis unterscheidet sich optisch in allen Skalen vom Vortestergebnis (Rechtsverschiebung der Residuen von Null). Während aber im Bereich der Beschleunigungsdiagramme der Großteil der Proband*innen die gleiche Punktzahl im Vortest und

im Nachtest erreichte und der Zuwachs insgesamt am geringsten ist, sieht man bei den anderen Verteilungen eine deutliche Verschiebung nach rechts, was für eine Verbesserung im Test spricht. Da die Normalverteilung allerdings nur bedingt gegeben ist, wird in der Folge mit nicht-parametrischen Tests auf Unterschiede zwischen Vor- und Nachtest getestet und auch eine Effektstärke für nicht-parametrische Verteilungen angegeben. Ergänzt wird dies allerdings für eine bessere Vergleichbarkeit zudem durch ein Cohens d (siehe Tab. 21). Es ergibt sich in allen Dimensionen ein signifikanter Unterschied zwischen Vor- und Nachtest (Abb. 29).

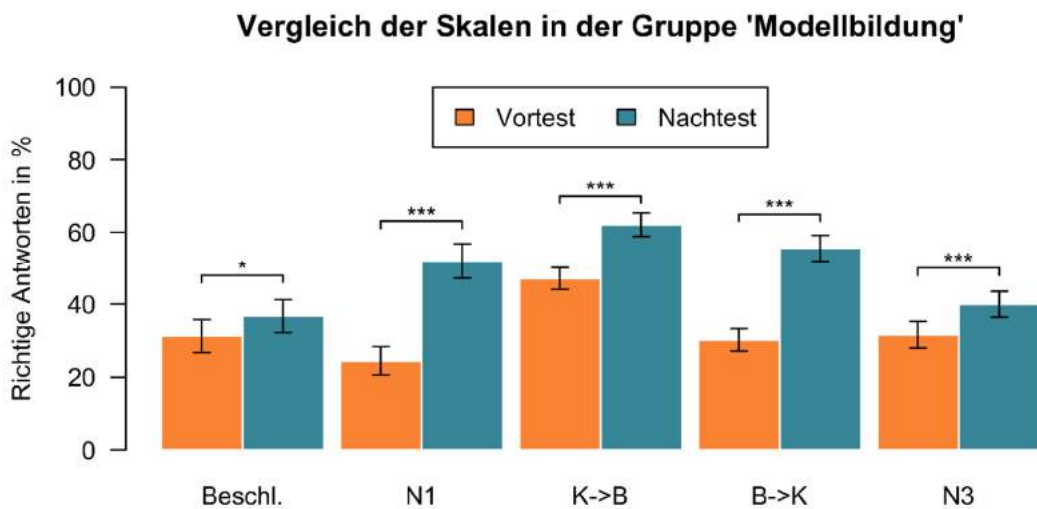


Abb. 29: Relative Punktzahl in Prozent in allen inhaltlichen Skalen („Beschleunigungsdiagramme“, „1. Newtonsches Gesetz“, „Kraft zu Bewegung“, „Bewegung zu Kraft“ und „3. Newtonsches Gesetz“) in der Gruppe der Modellbildung und Vor- und Nachtest mit 95%-Konfidenzintervallen²¹. * $p < 0,05$ *** $p < 0,001$.

Es gibt in allen Skalen einen signifikanten Lernzuwachs – auch in der Skala zum dritten Newtonschen Gesetz. Das verwundert, da diese Skala eine inhaltliche Kontrollskala darstellt, bei der keine Verbesserung zu erwarten ist, da sie kein Teil der Intervention war. Wenn man sich diese Skala „N3“ auf Itemebene anschaut, ergibt sich folgendes Bild (Abb. 30).

²¹ Die hier zugrundeliegende Berechnung der 95 %-Konfidenzintervalle über Quantile setzt eigentlich eine Normalverteilung voraus. Alternativ können bei nicht-parametrischen Verteilungen Konfidenzintervalle über *Bootstrapping* berechnet werden. Da die Verteilungen nahezu normalverteilt sind und die Konfidenzintervalle sich zwischen den Methoden kaum unterscheiden, wird hier und in der Folge auch bei nicht perfekt normalverteilten Daten auf diese Art der Konfidenzintervallberechnung zurückgegriffen.

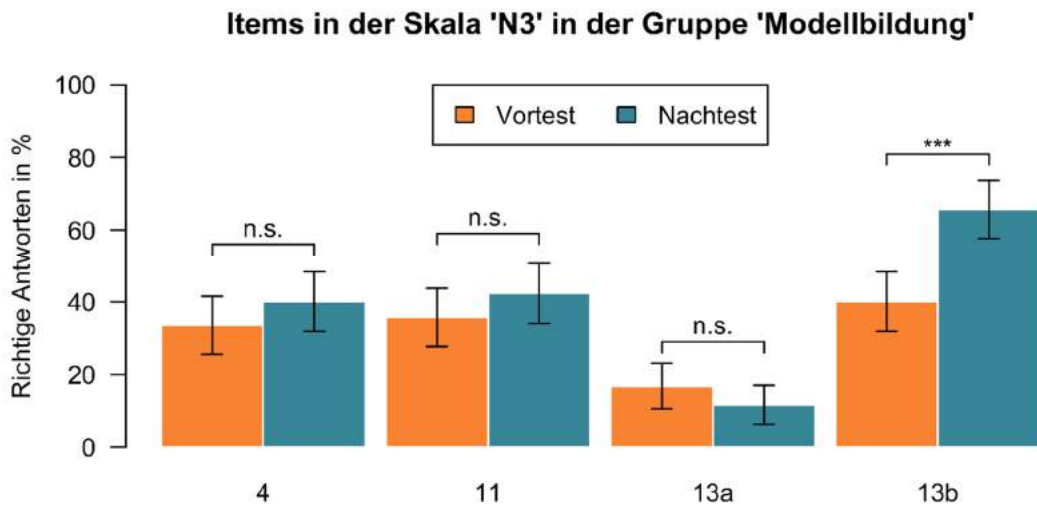


Abb. 30: Relative Punktzahl in Prozent in allen Items zum dritten Newtonschen Gesetz in der Gruppe der Modellbildung und Vor- und Nachtest mit 95%-Konfidenzintervallen. n.s. nicht signifikant, *** $p < 0,001$.

Die Änderung in der Skala „N3“ ist also nur auf die Änderung eines Items zurückzuführen. In dem Item wird nach den Kräften gefragt, die auf ein Automobil und einen LKW wirken, während das Auto den LKW mit einer konstanten Geschwindigkeit vor sich herschiebt (Abb. 31).

In der Intervention wird betont, dass sich alle Kräfte, die auf einen Körper mit konstanter Geschwindigkeit wirken, ausgleichen und zu einer Gesamtkraft von null addieren. Wenn also zwei Kräfte an einem Körper in unterschiedlicher Richtung angreifen, der sich mit konstanter Geschwindigkeit bewegt, müssen diese Kräfte gleichgroß sein. Es wird vermutet, dass diese Tatsache fälschlicherweise auf das Wechselwirkungsprinzip übertragen wurde, bei dem die Kräfte an unterschiedlichen Körpern angreifen. Dies würde hier zu einer korrekten Antwort führen, auch wenn die Überlegung dahinter nicht korrekt ist. Das Item 13a verbessert sich nicht. Vermutlich, da es sich in diesem Fall um eine beschleunigte Bewegung handelt, die auch nach der Intervention nicht mit Kräftegleichheit assoziiert wird. Das Nachtestergebnis ist hier sogar niedriger (wenn auch nicht signifikant), was plausibel ist, da Beschleunigung mit einer wirkenden, von null verschiedenen, Gesamtkraft assoziiert wird und nicht mit Kräftegleichheit. Der falsche Transfer von an einem Körper angreifenden Kräften zum dritten Newtonschen Gesetz müsste im Folgeunterricht berücksichtigt und thematisiert werden. Bestätigen ließe sich diese Vermutung durch eine zweite Stufe bei Item 13b, in der nach einer Begründung für die Antwort gefragt würde. Diese Begründung könnte helfen, die dahinterliegende Schülervorstellung offen zu legen.

Frage 13

Für die folgenden Aufgaben gilt folgende Situation: Ein LKW bleibt mit Motorschaden liegen und wird von einem Kleinwagen (Auto) zur nächsten Tankstelle geschoben.



- a)** Während das Auto beschleunigt, um beim Schieben auf eine bestimmte Geschwindigkeit zu kommen, gilt:
- | | | |
|--|--|--|
| ☐ Die Stärke der Kraft, mit der das Auto gegen den LKW drückt, ist genauso groß wie die Kraft, mit der der LKW gegen das Auto drückt. | ☐ Die Stärke der Kraft, mit der das Auto gegen den LKW drückt, ist kleiner als die Kraft, mit der der LKW gegen das Auto drückt. | ☐ Die Stärke der Kraft, mit der das Auto gegen den LKW drückt, ist größer als die Kraft, mit der der LKW gegen das Auto drückt. |
| ☐ Der Motor des Autos läuft; daher übt das Auto eine Kraft aus, während es gegen den LKW drückt. Aber der LKW-Motor läuft nicht; daher kann der LKW nicht gegen das Auto drücken. | ☐ Weder der LKW noch das Auto üben eine Kraft aus. Der LKW wird einfach deshalb nach vorne geschoben, weil er dem Auto im Wege steht. | |
- b)** Nachdem der Fahrer des Autos die gewünschte Geschwindigkeit erreicht hat, mit der er den Lastwagen konstant vor sich herschieben will, fährt er mit konstanter Geschwindigkeit weiter. Jetzt gilt:
- | | | |
|--|--|--|
| ☐ Die Stärke der Kraft, mit der das Auto gegen den LKW drückt, ist genauso groß wie die Kraft, mit der der LKW gegen das Auto drückt. | ☐ Die Stärke der Kraft, mit der das Auto gegen den LKW drückt, ist kleiner als die Kraft, mit der der LKW gegen das Auto drückt. | ☐ Die Stärke der Kraft, mit der das Auto gegen den LKW drückt, ist größer als die Kraft, mit der der LKW gegen das Auto drückt. |
| ☐ Der Motor des Autos läuft; daher übt das Auto eine Kraft aus, während es gegen den LKW drückt. Aber der LKW-Motor läuft nicht; daher kann der LKW nicht gegen das Auto drücken. | ☐ Weder der LKW noch das Auto üben eine Kraft aus. Der LKW wird einfach deshalb nach vorne geschoben, weil er dem Auto im Wege steht. | |

Abb. 31: Items 13a und 13b aus dem Fragebogen.

Ohne eine Begründung für die Verbesserung der Kontrollskala wäre infrage zu stellen gewesen, ob die generelle Verbesserung im Test lediglich mit einer höheren Konzentration im Nachtest zu tun gehabt hätte oder durch Testeffekte zustande gekommen wäre. Die Ergebnisse in der Kontrollskala deuten so aber auf einen tatsächlichen Effekt der Intervention hin. Der eingesetzte Test ähnelt zudem dem bekannten FCI-Test, weswegen ohnehin nicht von Test-Trainingseffekten auszugehen ist (Schecker & Gerdes, 1999b).

Um die Größe der Verbesserung in den Skalen einschätzen zu können, können wieder verschiedene Parameter für die Effektstärke berechnet werden. In der Folge wird die Effektstärke auf das r (gepaarter Wilcoxon-Test) und das d (gepaarter t-Test) beschränkt.

In den Skalen ergeben sich folgende Effektstärken (Tab. 21):

Tab. 21: Effektstärken für die fünf im Test enthaltenen inhaltlichen Skalen in der Gruppe der Modellbildung.

	Skala 1 (Beschl.)	Skala 2 (1. Newton)	Skala 3 (K->B)	Skala 4 (B->K)	Skala 5 (3. Newton)
p	$p < 0,05$	$p < 0,001$	$p < 0,001$	$p < 0,001$	$p < 0,001$
Pearsons r	0,19 (klein)	0,59 (groß)	0,56 (groß)	0,63 (groß)	0,31 (mittel)
Cohens d	0,18 (klein)	0,71 (mittel)	0,66 (mittel)	0,82 (groß)	0,33 (klein)

In der kinematischen Skala 1 „Beschleunigungsdiagramme“ wird wenig dazu gelernt. Die Effektstärke ist hier sogar noch geringer als in der Kontrollskala. Dies stellte allerdings auch nicht das Lernziel der Intervention dar. Dennoch mussten sich die Proband*innen in der Intervention mit den Diagrammen auseinandersetzen. Der Lerneffekt dieses impliziten Thematisierens von Beschleunigungsdiagrammen ist klein. In Skala 5 wäre kein Zuwachs zu erwarten gewesen. Dieser lässt sich aber, wie zuvor beschrieben, auf ein einzelnes Item und einen falschen Transfer zurückführen. Dieses Item zeigt, dass durch die Intervention ein Zusammenhang zwischen konstanter Geschwindigkeit und Kräftegleichgewicht hergestellt wurde, was Ziel der Intervention war. Deutlicher werden müsste nach der Intervention dann, dass dies nicht für zwei an unterschiedlichen Körpern angreifende Kräfte gilt. Die drei Skalen zur Dynamik weisen mittlere bis große Effektstärken auf (je nach Kennwert).

Analyse der Lernzuwächse auf Itemebene

Aufgeschlüsselt nach Items ergeben sich folgende Werte (Tab. 22) für das arithmetische Mittel in Vor- und Nachtest in den 19 relevanten Items (Dynamik und Beschleunigungsdiagramme, die Items zum dritten Newtonschen Gesetz werden im Folgenden nicht mehr betrachtet).

Tab. 22: Lösungswahrscheinlichkeiten der Items in Vor (P_{Vor})- und Nachtest (P_{Nach}), relativer Zugewinn pro Item (I_g)²² und Signifikanz des Zuwachses in der Gruppe der Modellbildung.

Item	P_{Vor} in %	P_{Nach} in %	I_g in %	Signifikanz	Item	P_{Vor} in %	P_{Nach} in %	I_g in %	Signifikanz
1	20,4	46,0	32,2	$p < 0,001$	10b	67,2	70,8	11,0	$p = 0,52$
2	11,7	56,2	50,4	$p < 0,001$	12	25,5	44,5	25,5	$p < 0,01$
3	25,5	64,2	51,9	$p < 0,001$	14	24,1	65,7	54,8	$p < 0,001$
5a	54,7	59,9	11,5	$p = 0,39$	15a	45,3	55,5	18,6	$p = 0,09$
5b	35,8	41,6	9,0	$p = 0,32$	15b	29,9	32,8	4,1	$p = 0,60$
6	22,6	55,5	42,5	$p < 0,001$	15c	22,6	23,4	1,0	$p = 0,89$
7	40,9	69,3	48,1	$p < 0,001$	15d	27,0	35,0	11,0	$p = 0,15$
8	24,1	43,8	26,0	$p < 0,001$	16a	28,5	51,1	31,6	$p < 0,001$
9	49,6	64,2	29,0	$p < 0,05$	16b	26,3	50,4	32,7	$p < 0,001$
10a	72,3	74,5	7,9	$p = 0,68$					

Der Anteil richtiger Antworten unterscheidet sich in den meisten Items vom Vor- zum Nachtest. Deutlich wird auch, dass die Unterschiede in den Items zu den Beschleunigungsdiagrammen (15a-15d) nicht signifikant sind. Die in der Tabelle dargestellten Werte ergeben folgendes Bild (Abb. 32).

Während die meisten Items im Nachtest deutlich häufiger richtig beantwortet wurden, gibt es auch Items, in denen der Score nahezu gleichbleibt (Item 5a, 5b, 10a, 10b, 15a, 15b, 15c, 15d). Drei der im Test vorhandenen Items sind zweistufig. In diesen Items sollen Schüler*innen eine Begründung für ihre Antwort geben. Diese Items wurden nur als korrekt gewertet, wenn die Antwort und ihre Begründung richtig sind. Die Begründung liefert dabei außerdem zusätzliche Informationen über die zugrundeliegende Schülervorstellung. Diese werden in Kapitel 10.5 näher betrachtet.

²² Girwidz et al. (2003)

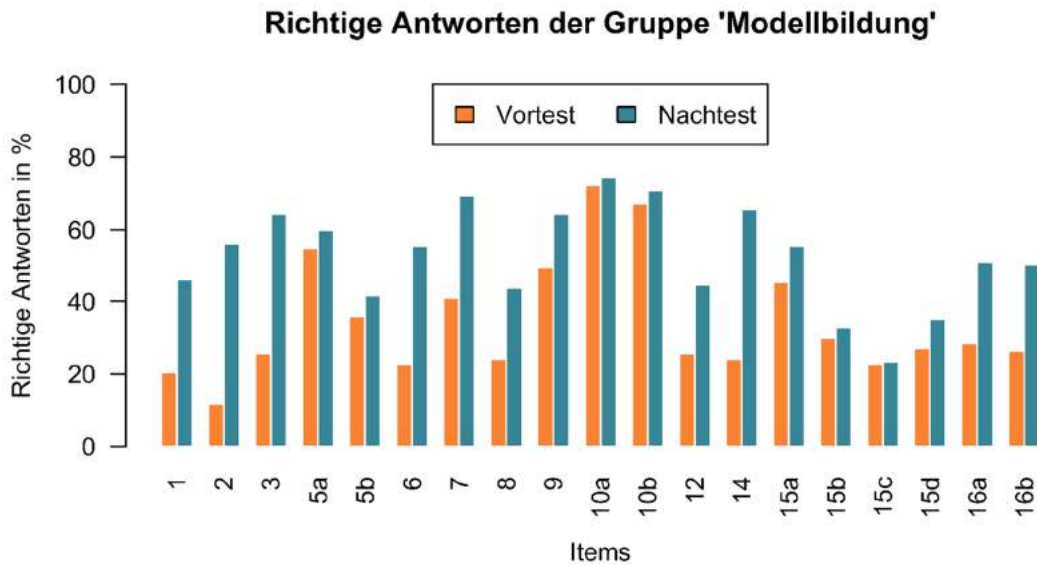


Abb. 32: Itemschwierigkeiten aller relevanten Items in Vor- und Nachtest in der Gruppe der Modellbildung.

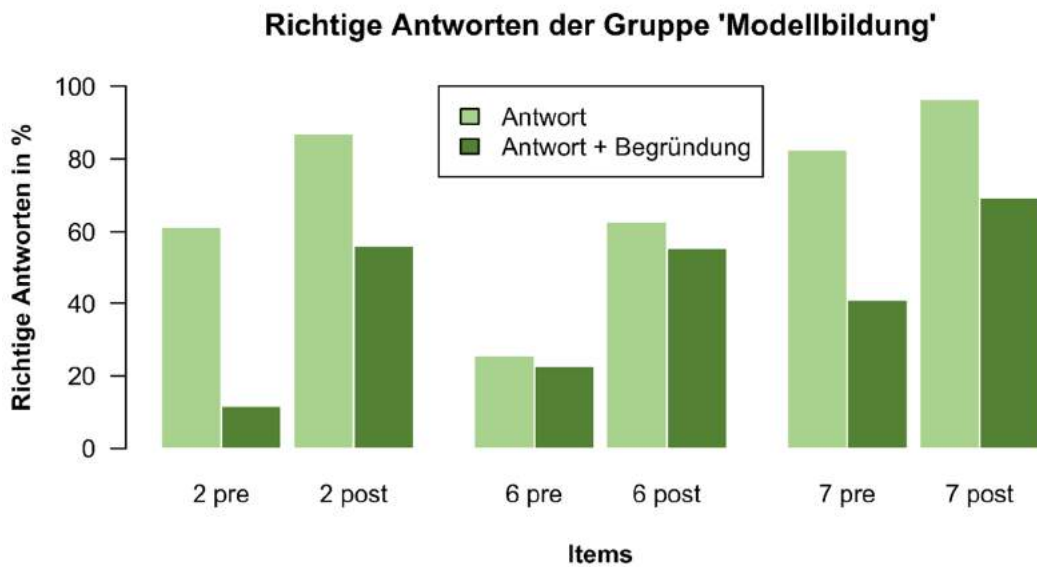
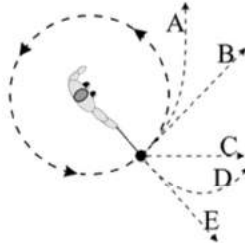


Abb. 33: Itemschwierigkeiten der zweistufigen Items aufgeschlüsselt nach der Antwort und der Begründung dieser Antwort in Vor- und Nachtest in der Gruppe der Modellbildung.

In Abb. 33 fällt auf, dass in Item 2 (Abb. 34) bereits im Vortest häufig die richtige Antwort gegeben wurde – dies aber häufig auf Basis einer falschen Vorstellung passiert. Das kann nur durch das zweistufige Item gezeigt werden. Die Anzahl an richtigen Begründungen hat sich hier besonders stark verbessert.

Frage 2

a) Ein schwerer Ball ist an einem Faden befestigt und wird, wie in der Abbildung gezeigt, im Kreis horizontal herumgeschwungen. An dem gekennzeichneten Punkt reißt jedoch plötzlich der Faden. Der Vorgang wird von oben betrachtet: Welchen Weg nimmt der Ball, nachdem der Faden gerissen ist?



☐ A
☐ D

☐ B
☐ E

☐ C

b) Welche Aussage erklärt deine Antwort am besten? Es geht um Kräfte **nach** Reißen des Fadens.

☐ Die Kugel wird durch die Zentrifugalkraft nach außen geschleudert.

☐ Die Kugel bewegt sich in einem Bogen, da sie vorher im Kreis geschwungen wurde.

☐ Es wirkt nach dem Reißen des Fadens keine Kraft auf die Kugel (in der horizontalen Ebene).

☐ Die Zentrifugalkraft und eine Kraft in Bewegungsrichtung wirken auf den Ball.

☐ Es wirkt nur eine Kraft in Bewegungsrichtung.

Abb. 34: Item 2 aus dem Fragebogen.

In Item 2 wird zunächst also häufig die richtige Flugbahn (B) ausgewählt. Bei den Begründungen ist die fünfte Antwort, dass nach dem Reißen eine Kraft in Bewegungsrichtung wirkt, ein attraktiver Distraktor, der für eine Aktivierung der Schülervorstellung spricht, dass immer eine Kraft in Bewegungsrichtung wirken muss. Nach der Intervention wird viel häufiger die richtige Begründung gewählt, dass unter Vernachlässigung der Reibung in der relevanten Ebene keine Kräfte mehr wirken. In Item 6 (Abb. 35) wurde die richtige Begründung wiederum fast immer gegeben, wenn die korrekte Antwort gegeben wurde, sodass die zweite Stufe hier eine geringere Aussagekraft hat.

Frage 6

Ein Fallschirmspringer bewegt sich mit konstanter Geschwindigkeit in Richtung Erde. Es wirkt auf ihn die nach unten gerichtete Gravitationskraft und eine nach oben gerichtete Luftreibungskraft.

a) Welche Aussage kannst du über die Stärke der Kräfte treffen?

☐ Die Gravitationskraft ist stärker als die Luftreibungskraft.

☐ Die Luftreibungskraft ist stärker als die Gravitationskraft.

☐ Beide Kräfte sind gleich stark.

b) Welche Aussage erklärt deine Antwort am besten?

☐ Bei einer konstanten Geschwindigkeit muss die Summe aller Kräfte Null sein.

☐ Die Kraft in Bewegungsrichtung muss immer am stärksten sein, damit sich etwas bewegen kann.

☐ Die abbremssende Kraft muss größer sein als die antreibende, da der Springer sonst immer schneller werden würde.

Abb. 35: Item 6 aus dem Test.

Das ist nicht erstaunlich, da in diesem Fall nur eine Begründung zu einer bestimmten Antwort passt. Die zweite Stufe wäre hier also nicht nötig gewesen, sie bestätigt aber, dass man aus der Antwort selbst die zugrundeliegende Vorstellung mit einer gewissen Sicherheit ableiten kann.

Zusammenfassung

Die Intervention zur Dynamik mit mathematischer Modellbildung ist als erfolgreich anzusehen. Insbesondere konnten in den Skalen zur Dynamik große Lernzuwächse nachgewiesen werden. Diese sind wahrscheinlich auf die Intervention zurückzuführen, da durch die Kontrollskala und die Art des Tests größere Test-Trainingseffekte auszuschließen sind und zwischen Vor- und Nachtest nichts außer der Intervention stattgefunden hat, was die Lösungswahrscheinlichkeiten der Items beeinflussen sollte. Die Ergebnisse in diesem quasiexperimentellen Prä-Post-Design sind allerdings nur mit einer Kontroll- bzw. Vergleichsgruppe wirklich aussagekräftig. Da die Videoanalyse mittlerweile als etabliertes Werkzeug für den Mechanikunterricht angesehen werden kann und neuere Studien den Erfolg der Videoanalyse gegenüber „herkömmlichem“ Unterricht zeigen, bietet sich diese als Vergleich an.

9.2.2 Videoanalyse

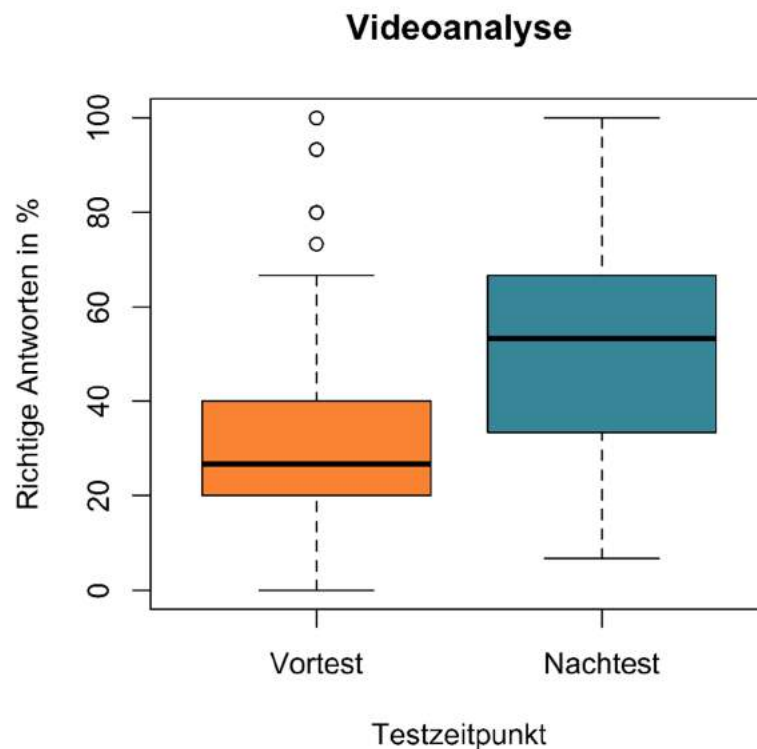


Abb. 36: Boxplots der durchschnittlichen relativen Punktzahlen im Bereich der Dynamik in Vor- und Nachtest in der Gruppe der Videoanalyse.

Analog zur Gruppe der mathematischen Modellbildung wird auch bei der Videoanalyse ($N = 137$) zunächst die Wirksamkeit der Intervention im Hinblick auf das Konzeptverständnis der Newtonschen Mechanik betrachtet, bevor auf einen Vergleich der beiden Interventionen eingegangen wird. In der Gruppe der Videoanalyse wurden von den 19 Items in Dynamik und Beschleunigungsdiagrammen im Schnitt 5,89 Punkte (31,0 %) im Vortest und 9,42 Punkte (49,6 %) im Nachtest erreicht, wobei ein durchschnittlicher Zuwachs von 3,52 Punkten (18,6 Prozentpunkte) zu verzeichnen war. Im Bereich der Dynamik wurden im Schnitt 4,79 Punkte (31,9 %) im Vortest und im Nachtest 7,98 Punkte (53,2 %) erreicht. Der Lernzuwachs betrug durchschnittlich 3,2 Punkte (21,3 Prozentpunkte). Die Boxplots für beide Testzeitpunkte sehen folgendermaßen aus (Abb. 36).

Ein Zuwachs von Vor- zu Nachtest ist in den Boxplots zu erkennen. Um diesen zu quantifizieren, muss wieder die Normalität der Daten überprüft werden. Die Normalität ist in der Gruppe der Videoanalyse insbesondere im Vortest anzuzweifeln (Abb. 37):

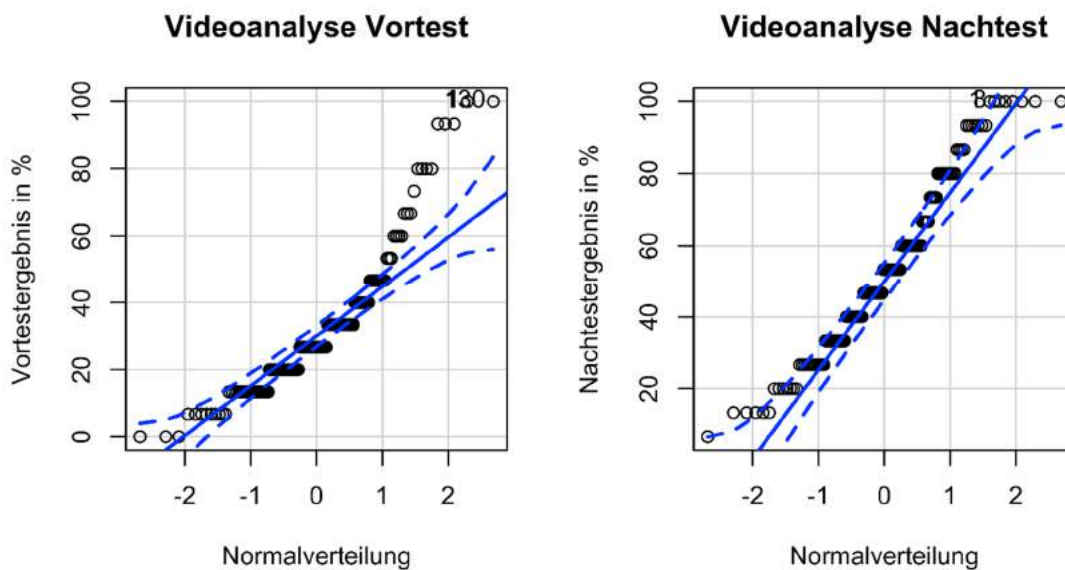


Abb. 37: QQ-Plots der Punktzahlen in Vor- und Nachtest gegenüber einer Normalverteilung in der Gruppe der Videoanalyse.

Dieser Eindruck bestätigt sich beim Blick auf das Histogramm der Testergebnisse, in dem auch der Unterschied zwischen Vor- und Nachtest deutlich wird (Abb. 38).

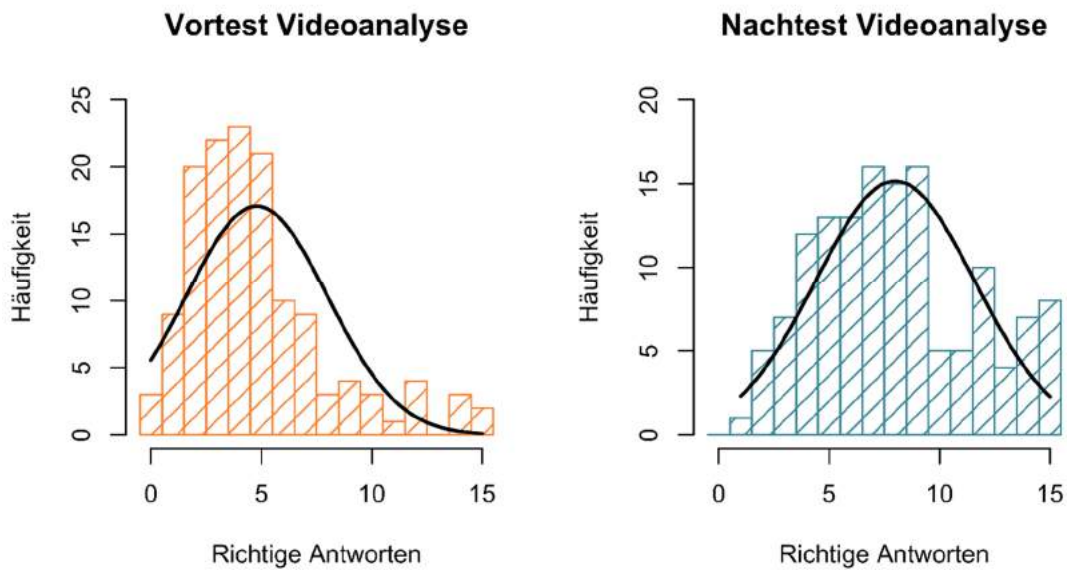


Abb. 38: Histogramme der Punktzahlen in Vor- und Nachtest in der Gruppe der Videoanalyse.

Für die Nutzung des t-Test ist abermals die Normalverteilung der Residuen entscheidend (Abb. 39).

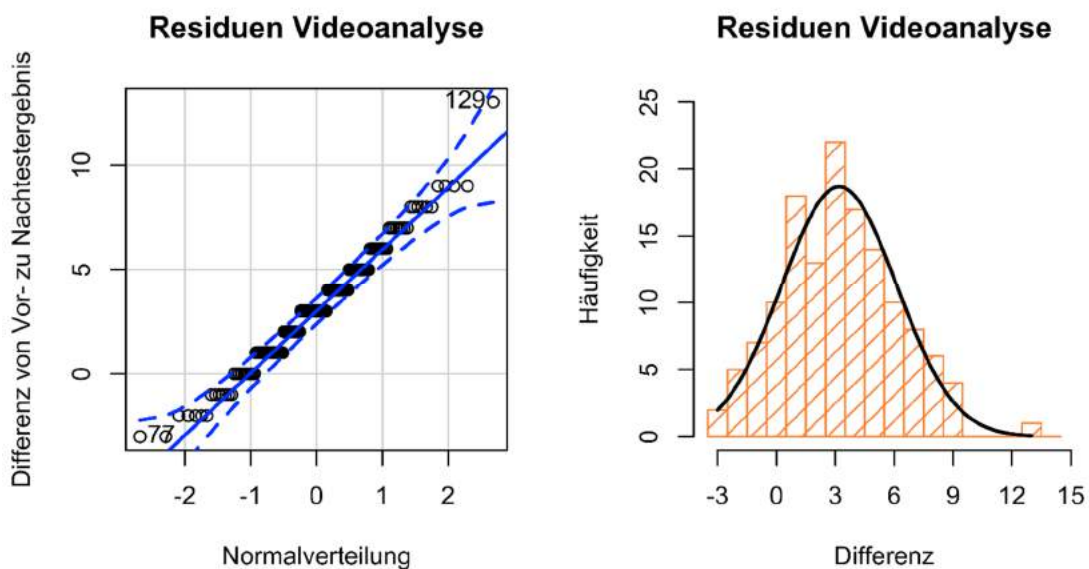


Abb. 39: Q-Q-Plot und Histogramm der Residuen in der Gruppe der Videoanalyse.

Da nach diesen Grafiken die Normalverteilungsannahme als gegeben angesehen werden kann, wird der t-Test für gepaarte Stichproben verwendet. Dieser ergibt einen höchst signifikanten Unterschied zwischen Vor- und Nachtest ($t = 13$; $df = 136$; $p < 0,001$).

Analog zur Gruppe der Modellbildung können auch hier verschiedene Effektstärken berechnet werden (Tab. 23). Deren Diskussion und die Art der Berechnung wird in Kapitel 9.2.1 dargestellt.

Tab. 23: Kennzahlen für die Effektstärke des Unterschieds zwischen Vor- und Nachtest in der Gruppe der Videoanalyse.

Kennzahl	Wert	Größe des Effekts
Cohens d_z	1,07	Groß (Cohen, 1988)
Cohens d_{rm}	0,93	Groß (Cohen, 1988)
Hakes $\langle g \rangle$	0,31	Mittel (Nissen et al., 2018)
Normalized change c	0,33	Mittel (Nissen et al., 2018)
Pearsons r	0,76	Groß (Cohen, 1988)

Ähnlich wie bei der Modellbildung wird deutlich, dass es sich um einen großen Lernzuwachs handelt. Auch hier ist zu erwähnen, dass die Intervention nach dem entsprechenden Schulunterricht stattfand und nur einen Vormittag andauerte.

Betrachtung der inhaltlichen Dimensionen

Betrachtet man die Ergebnisse in den Skalen, ergibt sich ein ähnliches Bild wie bei der Gruppe der Modellbildung (Abb. 40). Auf Unterschiede wurde jeweils mit einem t-Test für gepaarte Stichproben und dem nicht-parametrischen Wilcoxon-Rangsummentest für gepaarte Stichproben getestet (jeweils $p < 0,01$ bzw. $p < 0,001$). Auch hier verbessern sich also die Schüler*innen in der Kontrollskala „N3“ (drittes Newtonsches Gesetz). Auf Itemebene ergibt sich folgendes Bild (Abb. 41).

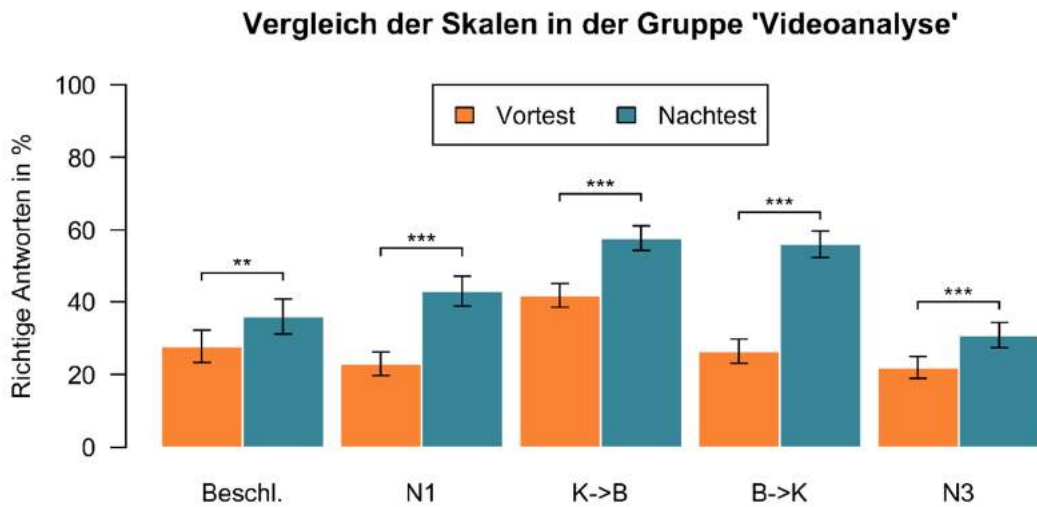


Abb. 40: Durchschnittliche relative Punktzahlen in allen inhaltlichen Skalen („Beschleunigungsdiagramme“, „1. Newtonsches Gesetz“, „Kraft zu Bewegung“, „Bewegung zu Kraft“ und „3. Newtonsches Gesetz“) in Vor- und Nachtest in der Gruppe der Videoanalyse mit Signifikanzniveaus und Konfidenzintervallen. ** $p < 0,01$ *** $p < 0,001$.

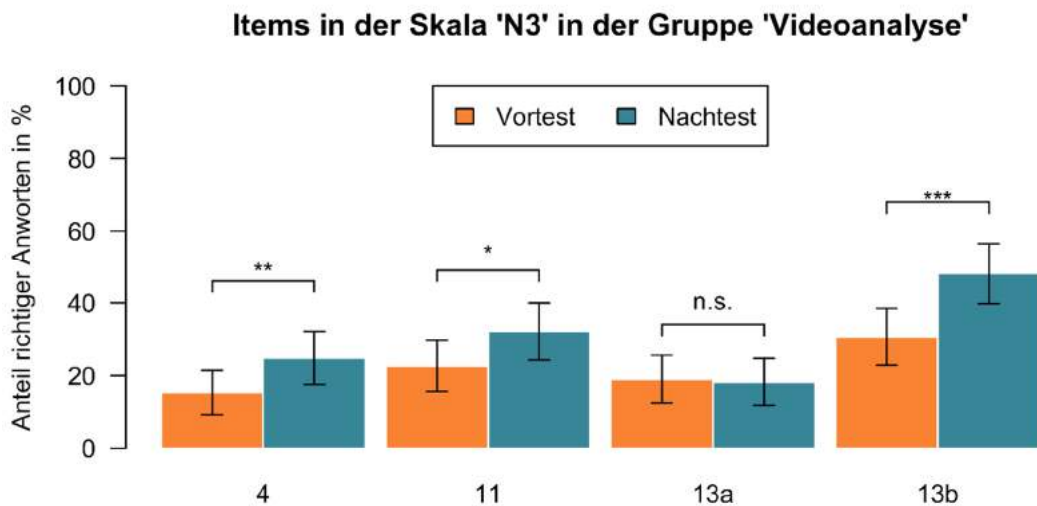


Abb. 41: Relative Punktzahl in Prozent in allen Items zum dritten Newtonschen Gesetz in der Gruppe der Videoanalyse und Vor- und Nachtest mit 95%-Konfidenzintervallen. n.s. nicht signifikant, * $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$.

Wieder ist der Großteil der Änderung auf das Item 13b zurückzuführen. Eine detaillierte Diskussion dieses Sachverhalts ist in Kapitel 9.2.1 zu finden. Minimale Verbesserungen mit kleiner Effektstärke sind auch in zwei weiteren Items zu finden (Item 4 und 11, Abb. 42 und Abb. 43), in denen die richtige Lösung ist, dass Wechselwirkungspartner a auf Wechselwirkungspartner b eine gleich große Kraft ausübt wie b auf a.

Frage 4

Stelle dir einen Frontalzusammenstoß zwischen einem Lastwagen (LKW) und einem Kleinwagen (PKW) vor. Für den Zeitraum des Zusammenpralls gilt:

☐ Der LKW übt eine größere Kraft auf den PKW aus als der PKW auf den LKW.	☐ Der PKW übt eine größere Kraft auf den LKW aus als der LKW auf den PKW.	☐ Die beiden Fahrzeuge üben keine Kräfte aufeinander aus. Der PKW wird einfach deshalb zusammen gedrückt, weil der dem LKW im Wege ist.
☐ Der LKW übt eine Kraft auf den PKW aus, aber der PKW übt keine Kraft auf den LKW aus.	☐ Der LKW übt die gleiche Kraft auf den PKW aus wie der PKW auf den LKW.	

Abb. 42: Item 4 aus dem Fragebogen.

Dies ließe sich einerseits auch durch ein Verinnerlichen des Kräftegleichgewichts und das Übertragen auf eine Situation, bei der die Kräfte an zwei unterschiedlichen Körpern angreifen, erklären. Geringer ist der Effekt bei den beiden Items im Vergleich zu Item 13b, da nicht von einer konstanten Geschwindigkeit die Rede ist. Andererseits ließe sich die Verbesserung möglicherweise auf eine höhere Konzentration im Nachtest oder andere die Umstände betreffende Faktoren zurückführen. Eine zweite Stufe mit einer Begründung zu den jeweiligen Antworten könnte helfen, eine Antwort auf die Frage zu liefern. Die Effektstärken ($d < 0,25$) sind aber klein, weswegen Faktoren, die den Test verzerren könnten, auch als klein einzuschätzen sind.

Frage 11

Zwei Schüler, von denen Schüler "a" die Masse 95 kg und Schüler "b" die Masse 77 kg hat, sitzen sich in zwei gleichen Bürostühlen direkt gegenüber. Schüler "a" stellt seine Füße auf die Knie von Schüler "b" (s. Abbildung). Schüler "b" hat seine Füße angehoben. Plötzlich streckt Schüler "a" seine Beine aus, wodurch beide Stühle in Bewegung versetzt werden. Welche Aussage trifft zu:

- ☐ Keiner der Schüler übt eine Kraft auf den anderen aus.
- ☐ Schüler "a" übt eine Kraft auf "b" aus, aber "b" übt keine Kraft auf "a" aus.
- ☐ Jeder der Schüler übt eine Kraft auf den anderen aus, aber "b" übt die größere Kraft aus.
- ☐ Jeder der Schüler übt eine Kraft auf den anderen aus, aber "a" übt die größere Kraft aus.
- ☐ Jeder der Schüler übt eine gleich starke Kraft auf den anderen aus.

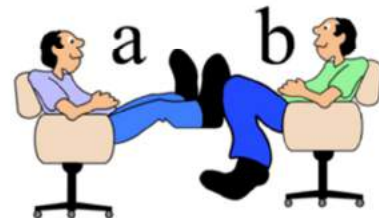


Abb. 43: Item 11 aus dem Fragebogen.

Für die Gruppe der Videoanalyse sollen ebenfalls die Effektstärken bestimmt werden. Dafür ist die Normalverteilung der Residuen in allen Skalen zu überprüfen (Abb. 44).

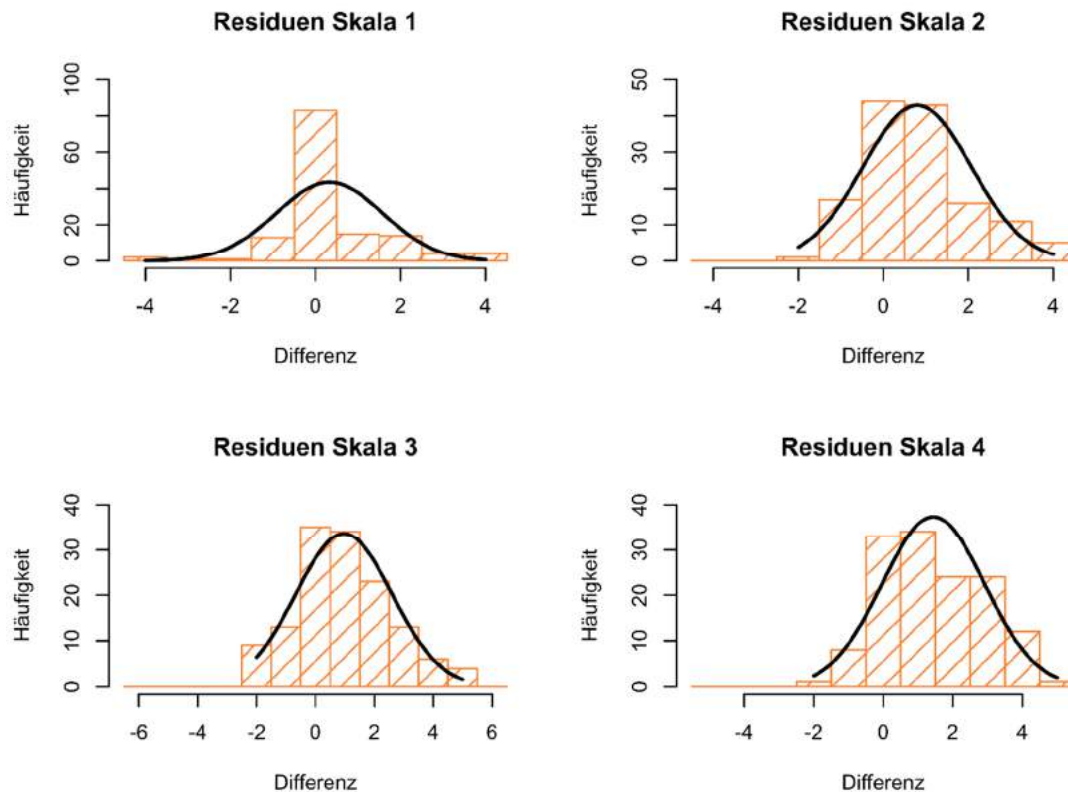


Abb. 44: Histogramme der Residuen in den vier Skalen in der Gruppe der Videoanalyse.

Die Residuen sind ausreichend normalverteilt, um die Effektstärke nach Cohen zu berechnen (gepaarter t-Test). Da zusätzlich mit einem nicht parametrischen Test auf Unterschiede der Mittelwerte getestet wurde, wird jeweils auch das r von Pearson angegeben. Die Effektstärken in den Skalen ergeben sich hier wie folgt (Tab. 24):

Tab. 24: p -Werte und Effektstärken für die im Test enthaltenen Skalen in der Gruppe der Videoanalyse.

	Skala 1 (Beschl.)	Skala 2 (1. Newton)	Skala 3 (K->B)	Skala 4 (B->K)	Skala 5 (3. Newton)
p	$p < 0,01$	$p < 0,001$	$p < 0,001$	$p < 0,001$	$p < 0,001$
Pearsons r	0,25 (klein)	0,50 (groß)	0,51 (groß)	0,73 (groß)	0,34 (mittel)
Cohens d	0,26 (klein)	0,58 (mittel)	0,59 (mittel)	0,98 (groß)	0,36 (klein)

Es ergeben sich ähnliche Effektstärken wie in der Gruppe der Modellbildung. Wie zu erwarten, wird in den Skalen zur Dynamik am meisten gelernt. Geringer ist der Effekt in der Skala Beschleunigungsdiagramme (aber größer als in der Gruppe der Modellbildung), mit denen sich die Proband*innen zwar auseinandergesetzt haben, die aber nicht intensiv besprochen wurden. In der Kontrollskala kam es auch zu einer Verbesserung. Die Ursache dafür wurde bereits zuvor diskutiert. Verdeutlichen müsste man in einem Folgeunterricht in jedem Fall, dass der Zusammenhang zwischen Kräftegleichgewicht und konstanter Geschwindigkeit nur für Kräfte gilt, die an einem einzelnen Körper angreifen und das Wechselwirkungsprinzip etwas anderes aussagt.

Analyse der Lernzuwächse auf Itemebene

Tab. 25: Lösungswahrscheinlichkeiten der Items in Vor (P_{Vor})- und Nachtest (P_{Nach}), relativer Zugewinn pro Item (I_g)²³ und Signifikanz des Zuwachses in der Gruppe der Modellbildung.

Item	P_{Vor} in %	P_{Nach} in %	I_g in %	Signifikanz	Item	P_{Vor} in %	P_{Nach} in %	I_g in %	Signifikanz
1	18,2	45,3	33,1	$p < 0,001$	10b	40,1	48,2	13,5	$p = 0,2$
2	17,5	57,7	47,7	$p < 0,001$	12	21,2	31,4	12,9	$p = 0,06$
3	29,2	58,4	41,2	$p < 0,001$	14	27,7	59,1	43,4	$p < 0,001$
5a	64,2	70,8	18,4	$p = 0,2$	15a	39,4	48,9	15,7	$p = 0,1$
5b	27,7	37,2	13,1	$p = 0,09$	15b	23,4	31,4	10,4	$p = 0,1$
6	18,2	44,5	32,2	$p < 0,001$	15c	21,2	28,5	9,3	$p = 0,2$
7	48,9	73,7	48,5	$p < 0,001$	15d	27,0	35,0	11,0	$p = 0,2$
8	23,4	37,2	18,0	$p < 0,05$	16a	25,5	59,1	45,1	$p < 0,001$
9	35,0	61,3	40,5	$p < 0,001$	16b	29,9	55,5	36,5	$p < 0,001$
10a	51,8	69,1	35,9	$p = 0,2$					

²³ Girwidz et al. (2003)

Auch wenn in der weiteren Analyse auf die Ergebnisse in den Skalen oder das Gesamtergebnis zurückgegriffen wird, kann man sich die Ergebnisse zunächst auf Itemebene anschauen. Dabei ergeben sich für den Anteil richtiger Lösungen folgende Werte (Tab. 25). Grafisch dargestellt ergeben diese Werte folgendes Bild (Abb. 45). Erkennbar ist, dass der Anteil an richtigen Antworten in jedem Item von Vor- zu Nachtest ansteigt, das Maß dieses Anstiegs aber je nach Item sehr unterschiedlich ist. Auch hier kann man die drei zweistufigen Items aufschlüsseln nach der gegebenen Antwort und der dazugehörigen Begründung (Abb. 46).

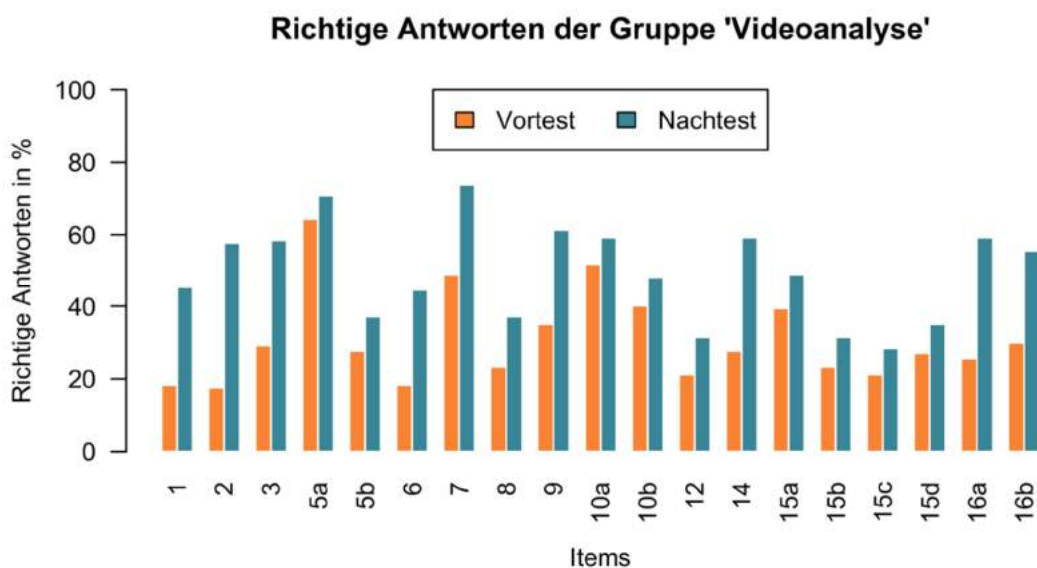


Abb. 45: Itemschwierigkeiten der relevanten Items in Vor- und Nachtest in der Gruppe der Videoanalyse.

Es wird deutlich, dass Item 2 wieder häufig richtig beantwortet aber falsch begründet wird, was insgesamt zu einer falschen Antwort führt. Die zweite Stufe legt also offen, dass viele Schüler*innen mit einer problematischen Schülervorstellung das Item richtig beantworten. Eine richtige Antwort allein spricht also nicht unbedingt für ein angemessenes Konzeptverständnis. Bei Item 6 wird bei richtiger Antwort meist auch die richtige Begründung dazu gegeben. Item 7 liegt in dieser Hinsicht zwischen den anderen Items (Abb. 47).

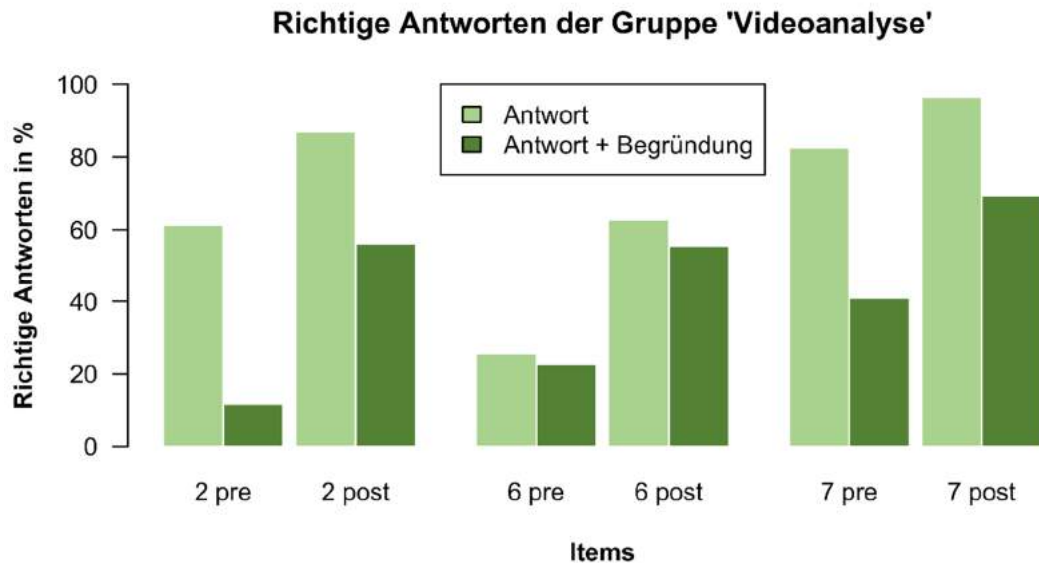
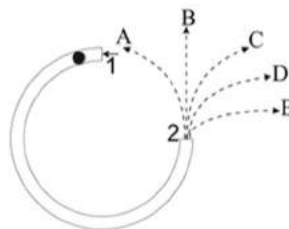


Abb. 46: Lösungswahrscheinlichkeiten der zweistufigen Items aufgeschlüsselt nach Antwort und Antwort mit Begründung in der Gruppe der Videoanalyse.

Frage 7

Die Abbildung zeigt einen von oben betrachteten kreisförmigen Kanal, der in der horizontalen Ebene fest auf dem Tisch verankert ist. Ein Ball tritt bei "1" in den Kanal ein und verlässt ihn bei "2".

a) Welche der in der Abbildung gezeigten Bahnkurven beschreibt am besten den Weg des Balls, wenn er den Kanal bei "2" verlässt und weiter über den Tisch rollt?



☐ A
☐ D

☐ B
☐ E

☐ C

b) Welche Aussage erklärt deine Antwort am besten?

☐ Die Zentrifugalkraft drückt den Ball nach außen.

☐ Die Kugel bewegt sich weiter im Kreis, da sie sich vorher so bewegt hat.

☐ Es wirkt keine Kraft auf die Kugel, also muss sie geradeaus rollen.

☐ Es wirkt eine Kraft in Bewegungsrichtung.

Abb. 47: Item 7 aus dem Fragebogen.

Item 7 ähnelt inhaltlich dem zuvor in Kapitel 9.2.1 diskutierten Item 2. Der Kontext scheint allerdings einerseits dazu zu führen, dass das Item häufiger richtig beantwortet wird und andererseits, dass die Schülervorstellung, dass eine Kraft in Bewegungsrichtung wirken muss, seltener aktiviert wird.

Zusammenfassung

Über die Videoanalyse war bislang vor allem im Bereich der Kinematik bekannt, dass ihr Einsatz zu einem erfolgreichen Unterricht führen kann. Die Daten dieser Arbeit zeigen, dass die Videoanalyse auch zum Lernen der Dynamik geeignet ist. In der kinematischen Skala der Beschleunigungsdiagramme war der Lernzuwachs allerdings am geringsten, da dies nicht Ziel der Intervention war. Damit können keine Rückschlüsse auf die Eignung der Software im Bereich der Kinematik gezogen werden. Ein Vergleich der Effektstärken mit anderen Studien ist schwierig. Dennoch zeigen die Daten, dass beide Methoden wirksam für das Vertiefen der Newtonschen Gesetze sind.

9.3 Vergleich der Interventionen

9.3.1 Kontrollvariablen

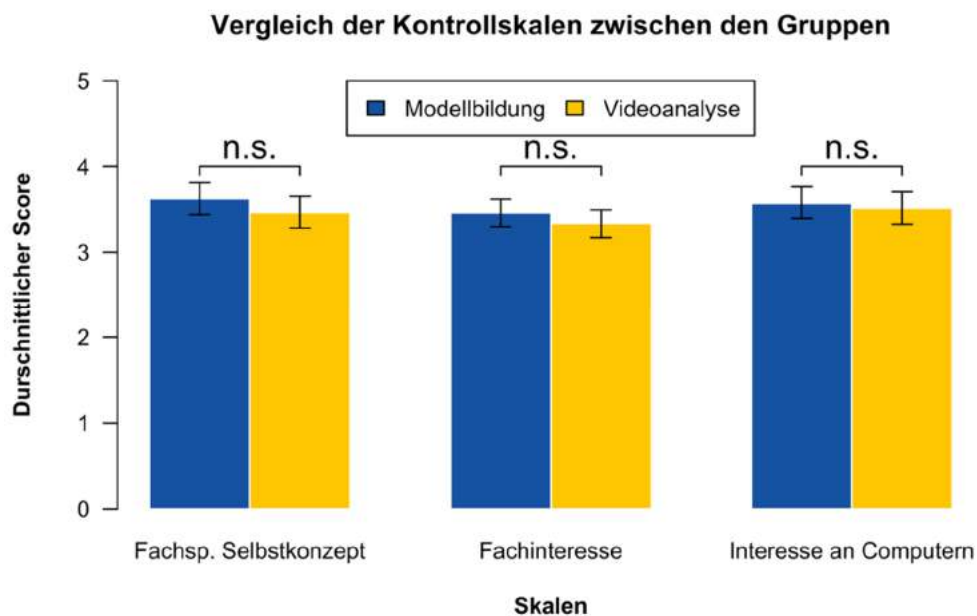


Abb. 48: Durchschnittlicher Score in den Kontrollskalen in beiden Gruppen mit Konfidenzintervall. n.s. nicht signifikant.

Die Interventionen werden verglichen, um herauszufinden, welche der beiden Varianten besser für den Einsatz im Dynamikunterricht geeignet ist und ob es Unterschiede in den verschiedenen inhaltlichen Bereichen gibt. Damit die Ergebnisse in beiden Gruppen verglichen werden können, wird zunächst untersucht, wie verschieden die Gruppen in Bezug auf gewisse Startbedingungen sind. Dazu wurden neben dem Vortest der fachlichen Items

auch weitere Kontrollvariablen erhoben. Es wurden das Fachinteresse und das fachspezifische Selbstkonzept erhoben. Da in der Intervention intensiv mit Computern gearbeitet wird, könnte sich auch ein Unterschied im Interesse an Computern in einem Unterschied im Lernzuwachs äußern (Abb. 48).

Nach dem t-Test für ungepaarte Stichproben gibt es auf Basis eines 5 %-Signifikanzniveaus in den Kontrollskalen keine Unterschiede zwischen den beiden Gruppen, wodurch nicht von einer Beeinflussung dieser Skalen des Vergleichs zwischen den Gruppen auszugehen ist. Es wurden außerdem die letzten Noten in Mathematik und Physik erfragt. Größere Unterschiede in den Noten könnten auf unterschiedliche Voraussetzungen in den beiden Gruppen hindeuten. Dies ist hier nicht der Fall (Abb. 49).

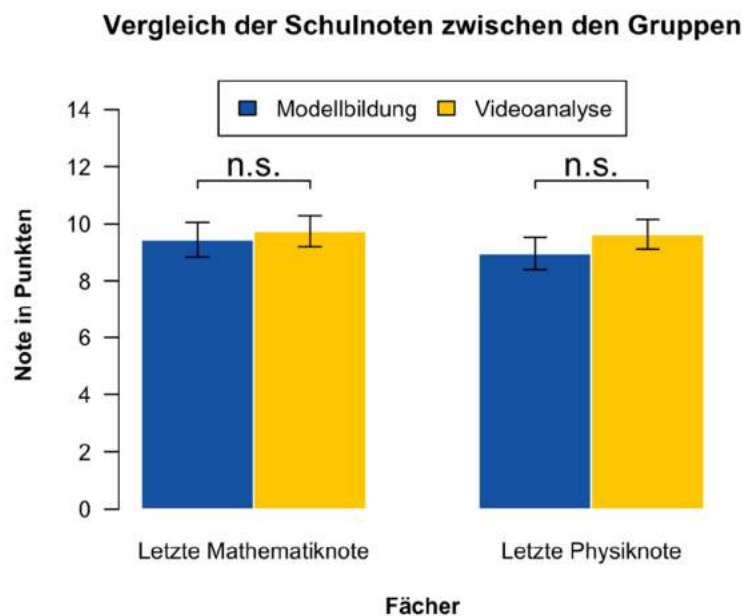


Abb. 49: Durchschnittliche letzte Schulnote in Punkten in Mathematik und Physik in beiden Gruppen mit Konfidenzintervall. n.s. nicht signifikant.

Erhoben wurde zudem das Interesse am Experimentieren im Physikunterricht bzw. das Interesse an theoretischen Zusammenhängen und dem Aufstellen von Hypothesen. Diese Skalen werden zusätzlich Teil einer inhaltlichen Analyse sein, können aber auch dazu dienen, die Gruppen untereinander zu vergleichen (Abb. 50). Es kann also davon ausgegangen werden, dass sich die Gruppen in den Kontrollvariablen nicht signifikant unterscheiden und damit mögliche Unterschiede im Lernzuwachs ausschließlich auf die Intervention zurückzuführen sind. Einen Überblick über alle Kontrollskalen gibt Abb. 51:

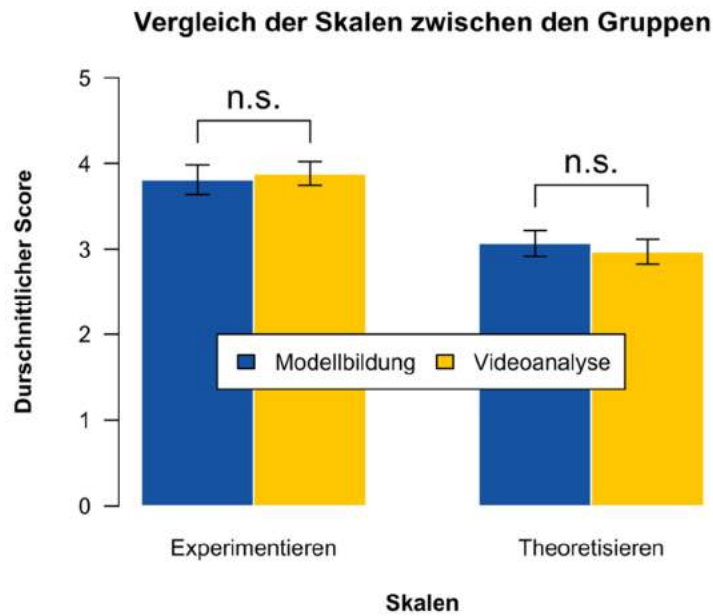


Abb. 50: Durchschnittlicher Score in den Skalen Experimentieren und Theoretisieren in beiden Gruppen mit Konfidenzintervall. n.s. nicht signifikant.

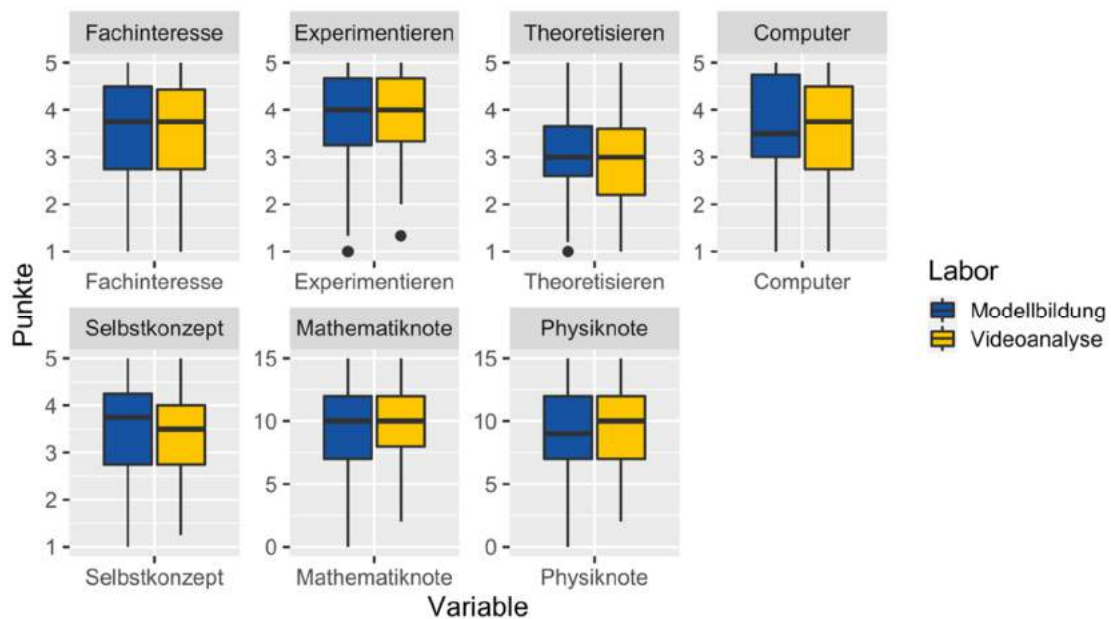


Abb. 51: Überblick über alle Kontrollskalen in beiden Gruppen.

9.3.2 Nachtestergebnisse

Um die beiden Interventionen miteinander vergleichen zu können, kann zunächst das Nachtestergebnis betrachtet werden, da dies die Variable ist, die durch die Intervention beeinflusst werden sollte. Da der Lernzuwachs nur unter Berücksichtigung des Vortestergebnisses betrachtet werden kann, werden die Vortestergebnisse im anschließenden Kapitel analysiert.

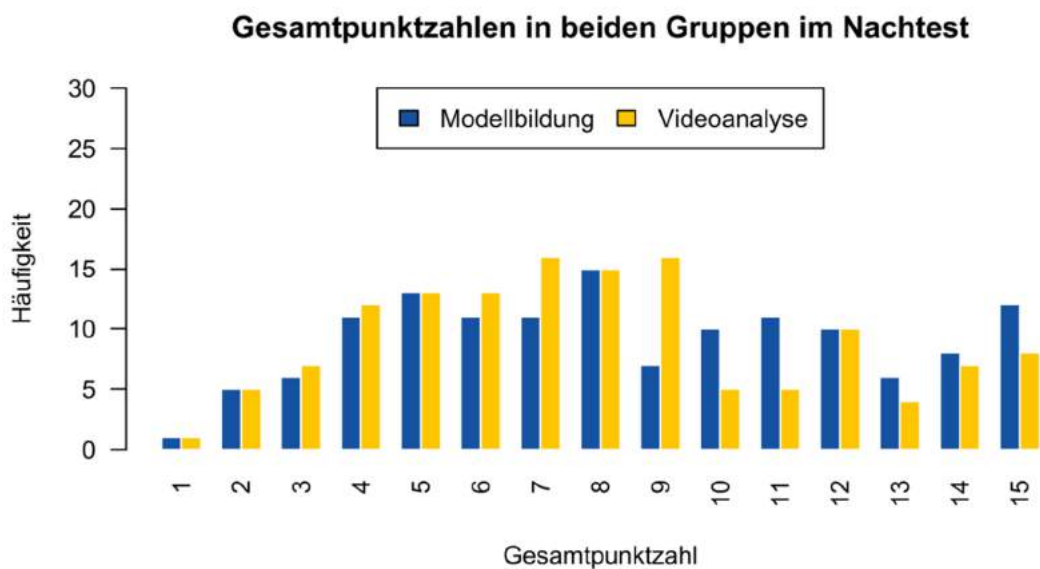


Abb. 52: Histogramm der Gesamtpunktzahlen im Nachtest (von max. 15) in beiden Gruppen.

In dem Plot der Gesamtpunktzahlen sind keine nennenswerten Unterschiede zwischen den Gruppen zu erkennen. Die Punktzahlen sind insgesamt gleichmäßig verteilt. Speziell in der Videoanalyse erreichen viele Schüler*innen vor allem Punktzahlen von bis zu neun Punkten. Die Punktzahlen darüber werden in dieser Gruppe etwas seltener erreicht. Es kann zusätzlich ein Boxplot betrachtet werden (Abb. 53):

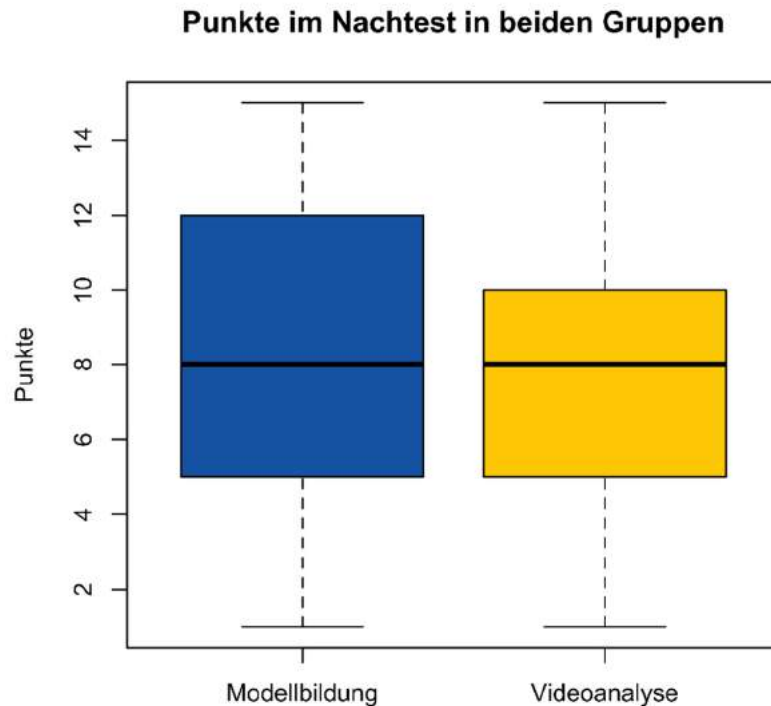


Abb. 53: Boxplot mit Standardeigenschaften der Nachtestpunktzahlen (max. 15) in beiden Gruppen.

Der Median ist in beiden Gruppen gleich, die Maxima und Minima ebenso. Das drittel Quartil ist in der Gruppe der Modellbildung höher. Zum Vergleich der beiden Nachtestergebnisse wird der t-Test für unabhängige Stichproben verwendet. In der Gruppe der Modellbildung wurden im Schnitt 8,58 Punkte und in der Gruppe der Videoanalyse 7,99 Punkte erreicht. Der Unterschied ist nicht signifikant ($p = 0,2$).

9.3.3 Vortestergebnisse

Da die Nachtestergebnisse neben der Intervention vor allem auch von dem Vorwissen abhängen, was durch das Vortestergebnis quantifiziert wird, werden die Vortestergebnisse analog ausgewertet (Abb. 54).

Die Vortestergebnisse sind insgesamt im Vergleich zum Nachtest verschoben. Am häufigsten werden vier oder fünf Punkte im Test erreicht, während die höheren Punktzahlen kaum vorkommen. Auch diese werden im Vortest aber teilweise bereits erreicht. Da das Thema im Schulunterricht bereits behandelt war, kann man sagen, dass der Schulunterricht nicht erfolgreich darin war, den meisten Schüler*innen ein qualitatives Verständnis für den Zusammenhang zwischen Kraft und Bewegung so zu vermitteln, dass sie die Testaufgaben richtig beantworten können.

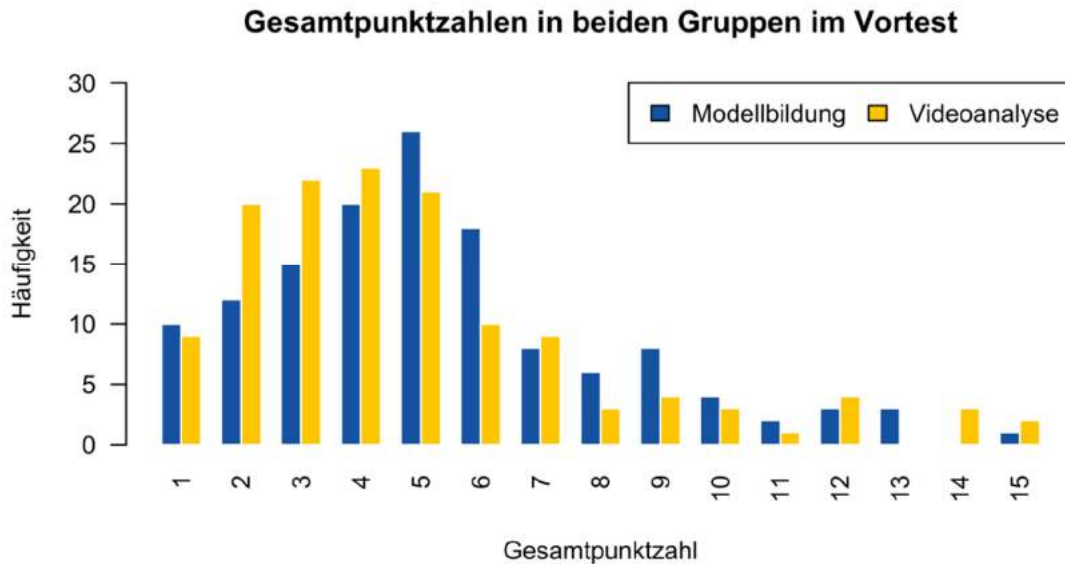


Abb. 54: Histogramm der Gesamtpunktzahlen im Nachtest (von max. 15) in beiden Gruppen.

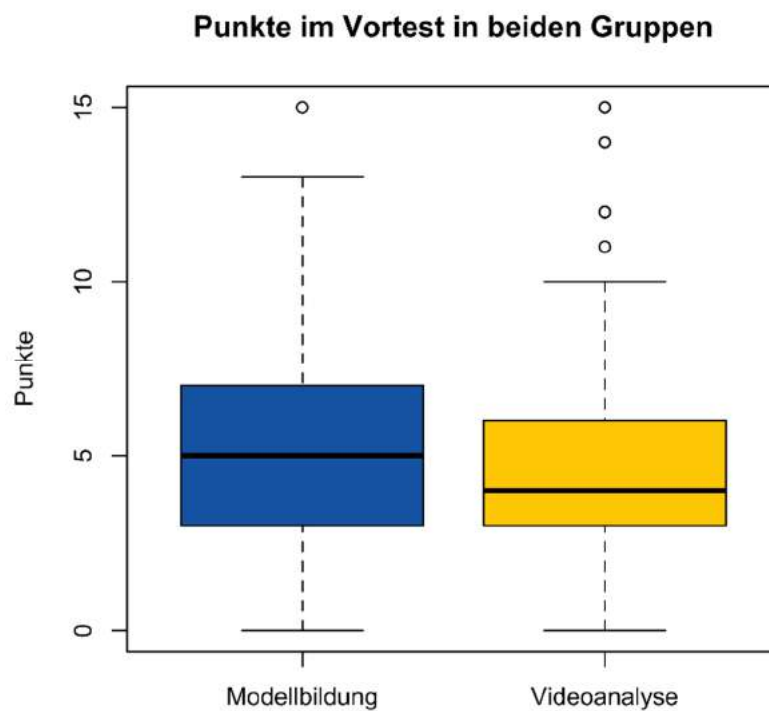


Abb. 55: Boxplot mit Standardeigenschaften der Vortestpunktzahlen (max. 15) in beiden Gruppen.

Zwischen den Gruppen sind keine Unterschiede zu erkennen. Dies kann durch Betrachtung des Boxplots bestätigt werden. Die Vortestergebnisse können im Boxplot dargestellt

werden (Abb. 55). Der Median in der Gruppe Modellbildung ist einen Punkt höher als in der Gruppe der Videoanalyse, die Verteilung ist aber ähnlich. Auch die Vortestergebnisse werden mit einem t-Test für unabhängige Stichproben auf einen Unterschied überprüft. Im Schnitt wurden in der Gruppe der Modellbildung im Vortest 5,29 Punkte und in der Gruppe der Videoanalyse 4,79 Punkte erreicht. Auch hier liegt mit $p = 0,2$ kein signifikanter Unterschied vor.

9.3.4 Absoluter Lernzuwachs

Relevant für die Bewertung der Interventionen ist der absolute Lernzuwachs, also die Differenz zwischen Vor- und Nachtestergebnis. Der durchschnittliche absolute Lernzuwachs in den beiden Gruppen liegt bei 3,28 (Modellbildung) bzw. 3,20 Punkten (Videoanalyse) von 15 erreichbaren Punkten. Die Histogramme geben einen Überblick über die Verteilung und Unterschiede zwischen den Gruppen (Abb. 56):

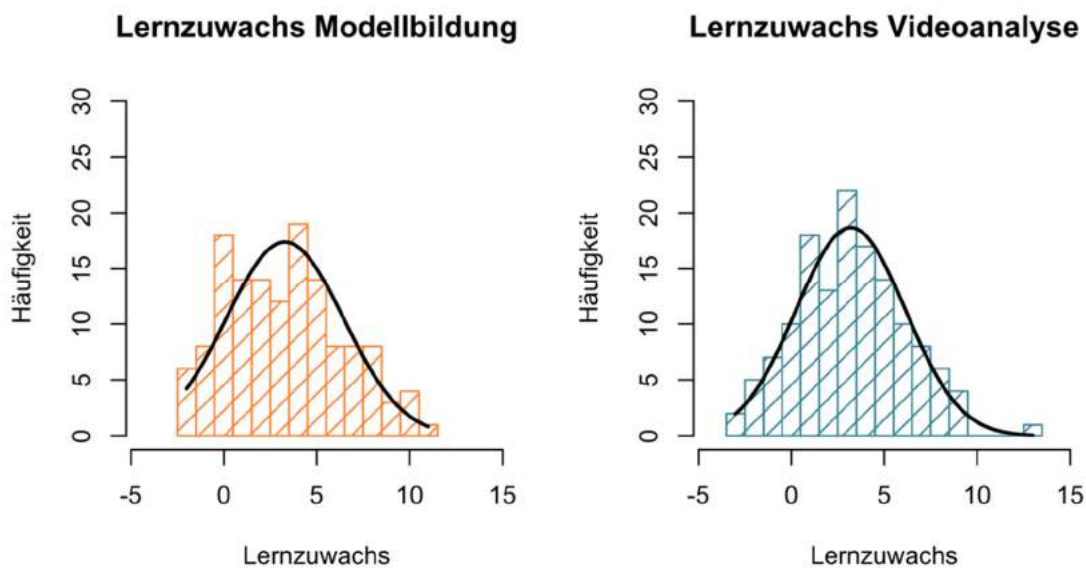


Abb. 56: Histogramm der absoluten Lernzuwächse als Differenz von Vortest zu Nachtest in beiden Gruppen mit Fit einer Normalverteilung.

Der Mittelwert der Lernzuwächse unterscheidet sich in beiden Gruppen signifikant von null wie bereits in Kapitel 9.2 diskutiert wurde. Die Verteilung der Modellbildung ist breiter, was für eine höhere Streuung der Lernzuwächse spricht. Eine höhere Streuung bei ähnlichem Mittelwert würde bedeuten, dass manche Schüler*innen sehr stark von der Methode profitieren, während andere dies sehr wenig tun oder sogar weniger Punkte im Nachtest verzeichnen. Dies könnte mit der Schwierigkeit, einen Zugang zu dieser

Methode zu finden, oder mit dem generellen Interesse daran begründet werden. In der Tat ist die Standardabweichung mit $\sigma = 3,14$ in der Gruppe der Modellbildung im Vergleich zu $\sigma = 2,93$ in der Gruppe der Videoanalyse erhöht. Ein einseitiger F-Test (Levene-Test) zur Überprüfung von Unterschieden in den Varianzen liefert mit $p = 0,2$ ($F = 1$) allerdings keine signifikant größere Streuung in der Gruppe der Modellbildung. Ein t-Test für unabhängige Stichproben liefert außerdem keinen Unterschied zwischen den Mittelwerten der Lernzuwächse zwischen den beiden Gruppen ($p = 0,8$; $t = 0,2$).

Zweifaktorielle Varianzanalyse mit Messwiederholung (ANOVA)

Bortz und Döring (2006) schlagen zum Vergleich des Lernzuwachses für das vorliegende Studiendesign eine zweifaktorielle Varianzanalyse mit Messwiederholung vor. Dabei geht es um die Untersuchung, ob ein signifikanter Interaktionseffekt zwischen Gruppenfaktor und Messwiederholungsfaktor vorliegt. Wenn dies der Fall ist, kann von einem signifikanten Unterschied des Einflusses der Treatments ausgegangen werden. Es wird bei diesem Vorgehen auch von einem gemischten Design gesprochen, da ein „between-subjects-design“ (Vergleich der Gruppen) mit einem „within-subjects-design“ (Testzeitpunkte) gemischt wird. Für die Anwendung des Verfahrens müssen einige Bedingungen erfüllt sein (Rasch et al., 2010):

(1) *Die abhängigen Variablen müssen intervallskaliert sein:*

Bei der abhängigen Variablen handelt es sich in beiden Gruppen um den absoluten Lernzuwachs, der intervallskaliert ist.

(2) *Die Messwerte müssen in allen Stichproben normalverteilt sein:*

Der Lernzuwachs kann in beiden Gruppen als normalverteilt angesehen werden (Abb. 56). Wie bereits in Kapitel 9.2 diskutiert, ist insbesondere in der Gruppe der Videoanalyse in den einzelnen Testzeitpunkten nicht von einer Normalverteilung auszugehen. Die Residuen sind aber ausreichend normalverteilt. Allgemein ist die zweifaktorielle Varianzanalyse nach Rasch et al. (2010) zudem robust gegenüber der Verletzung der Normalitätsbedingung, sodass diese hier dennoch durchgeführt werden kann.

(3) *Es sollte Varianzhomogenität gegeben sein:*

Nach dem Levene-Test sind sowohl im Vortest als auch im Nachtest keine Unterschiede in der Varianz zwischen der Gruppe der Modellbildung und der der Videoanalyse vorhanden. Damit ist diese Bedingung erfüllt.

(4) *Es sollte Sphärizität gegeben sein:*

Diese Bedingung ist nach Rasch et al. (2010, S. 72) bei einem zweistufigen Faktor automatisch gegeben.

Mit Ausnahme der Normalität in allen Messzeitpunkten, dessen Auswirkungen diskutiert wurden, sind alle Voraussetzung für die zweifaktorielle Varianzanalyse mit Messwiederholung gegeben.

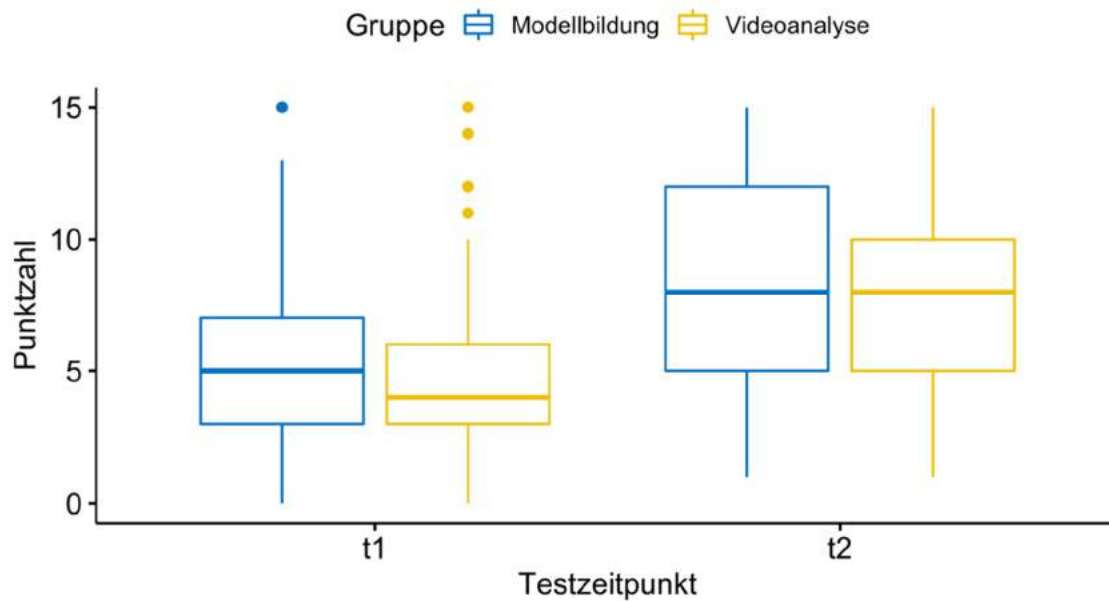


Abb. 57: Direkter Vergleich der Vor- und Nachtestergebnisse in beiden Gruppen anhand von Boxplots.

Die ANOVA ergibt, dass zwar der Faktor Zeit einen höchst signifikanten Einfluss auf das Testergebnis hat (Abb. 57), nicht jedoch die Interaktion zwischen Gruppe und Zeit (Tab. 26). Nach der zweifaktoriellen Varianzanalyse mit Messwiederholung gibt es also keinen Unterschied zwischen den Treatments, was den Lernzuwachs angeht. Dies bestätigt das Ergebnis des einfachen t-Tests, angewendet auf die Lernzuwächse.

Tab. 26: Kennwerte der gemischten ANOVA inkl. Effekt, Freiheitsgraden (DFn u. DFd), F-Statistik (F), p-Wert (p), Signifikanz ($p < .05$) und Effektstärke (η_g^2).

	DFn	DFd	F	p	$p < .05$	η_g^2
Effekt Gruppe	1	272	2,071	0,178	nicht erfüllt	6,00e-03
Effekt Zeit	1	272	303,445	3.65e-46	erfüllt	1,80e-01
Effekt Gruppe:Zeit	1	272	0,101	0,751	nicht erfüllt	7,34e-05

Der einzig signifikante Effekt ist die Zeit, also der Unterschied zwischen Vor- und Nachtest. Der Effekt Zeit ist mit $\eta_g^2 = 0,18$ groß (Cohen, 1988). Der Interaktionseffekt

zwischen Gruppe und Zeit ist klein (nicht signifikant). Grafisch sieht der Lernzuwachs der beiden Gruppen folgendermaßen aus (Abb. 58).

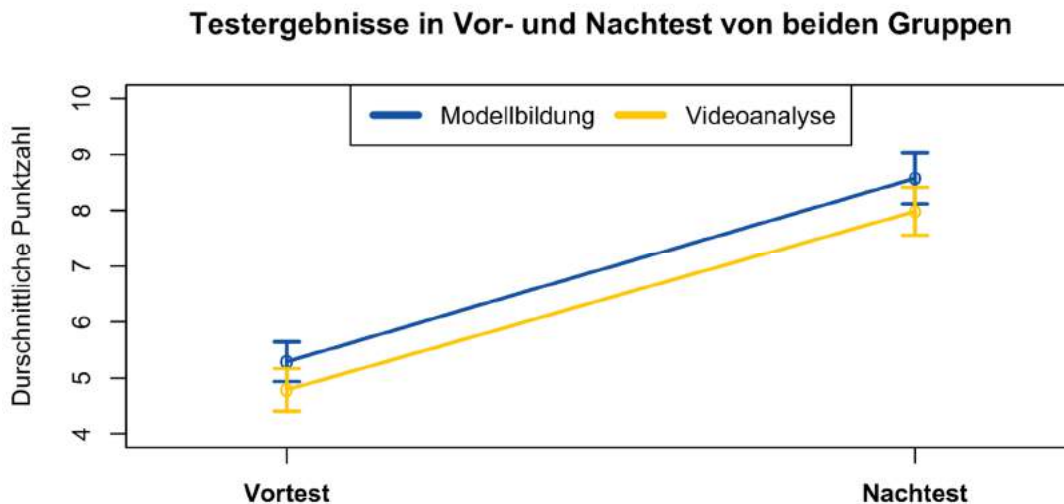


Abb. 58: Punktzahlen in Vor- und Nachtest (max. 15) in beiden Gruppen mit 95%-Konfidenzintervallen.

Die Grafik bestätigt, dass sich das Testergebnis im Nachtest von dem im Vortest unterscheidet, beide Gruppen jedoch scheinbar gleich viel dazu gelernt haben. Es gibt also in dieser Hinsicht keinen Unterschied zwischen den Interventionen.

9.3.5 Lernzuwachs nach Terzilen

Von Interesse ist auch, wie sich der Lernzuwachs von schwächeren Schüler*innen zu denen von stärkeren Schüler*innen verhält. Dabei können verschiedene Effekte eine Rolle spielen. Einerseits besagt der sogenannte Matthäus-Effekt (Rigney, 2010), dass Lernende, die bereits viel wissen, leichter dazu lernen können und damit stärkere Schüler*innen einen höheren Lernzuwachs verzeichnen. Andererseits können Schüler*innen, die im Vortest gut waren, weniger Punkte dazugewinnen als schwächere Schüler*innen. Um zu überprüfen, wie sich das bei diesen Interventionen verhält und ob es einen Unterschied zwischen beiden Gruppen gibt, können die Lernzuwächse nach Terzilen betrachtet werden. Dazu wurden in jeder Gruppe die Proband*innen nach dem Vortestergebnis in drei Gruppen eingeteilt (schlecht, mittel, gut). Es ergibt sich folgendes Bild für die Lernzuwächse (Abb. 59).

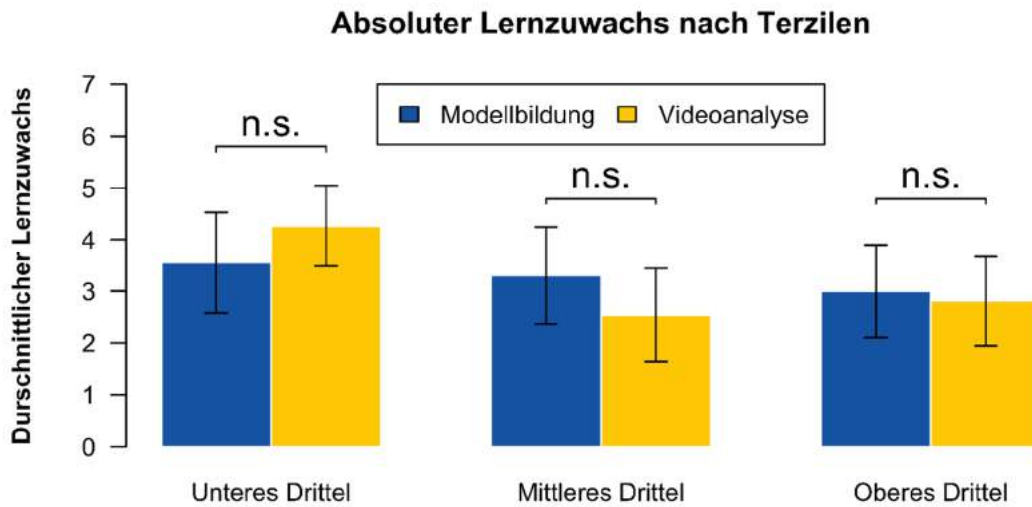


Abb. 59: Absoluter Lernzuwachs nach Terzilen in beiden Gruppen mit 95%-Konfidenzintervallen.

Es kommt nicht zu einem Matthäus-Effekt. Die Korrelation zwischen Vortestergebnis und Lernzuwachs ist in beiden Gruppen mit $r = -0,17$ (Modellbildung) bzw. $r = -0,21$ (Videoanalyse) negativ. Dabei gibt es in den Terzilen keine Unterschiede zwischen den Gruppen. Bessere Schüler*innen haben in absoluten Werten weniger dazugelernt. Das motiviert die Betrachtung des relativen Zuwachses nach Terzilen (Abb. 60).

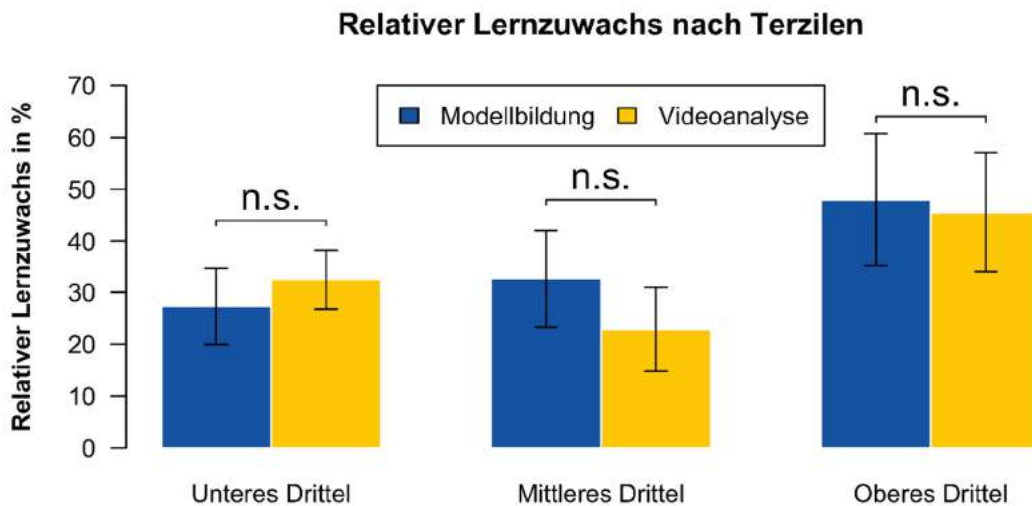


Abb. 60: Relativer Lernzuwachs nach Terzilen in beiden Gruppen mit 95%-Konfidenzintervallen.

Auch hier ist zwischen den Gruppen kein Unterschied zu erkennen. Das kann durch eine ANCOVA (siehe Kapitel 9.3.6) bestätigt werden, in der sich zwar zeigt, dass die Terzile unterschiedlich viel dazulernen, aber kein Interaktionseffekt zwischen Gruppe und Terzil nachgewiesen werden kann. Beim relativen Zuwachs erreicht das obere Terzil die besten Werte. Durch deren hohe Punktzahlen im Vortest sind wenige Punkte Zuwachs auch bereits ein hoher relativer Zugewinn.

9.3.6 Kontrolle der Vortestergebnisse

Auch wenn hier kein signifikanter Unterschied in den Testergebnissen des Vortests vorliegt, gibt es verschiedene statistische Verfahren, mit denen auch kleine Unterschiede in den Vortestergebnissen kontrolliert werden können. Eine Möglichkeit ist die sogenannte Parallelisierung, in denen Paare von Proband*innen aus den beiden Gruppen gebildet werden, die sich möglichst ähnlich sind, deren Nachtestergebnisse daraufhin betrachtet werden. Dies soll zu einer Übereinstimmung in allen wichtigen Variablen der beiden Gruppen führen (inklusive des Vortestergebnisses), sodass das Nachtestergebnis als alleiniger Indikator für den Lernerfolg der Intervention dienen kann. Die Erhöhung der Vergleichbarkeit zwischen den Gruppen geht allerdings mit einem Verlust von Daten einher, da manche Proband*innen gestrichen werden müssen.

Matching nach Vortestergebnissen

Es gibt verschiedene Möglichkeiten, diese *matched samples* zu erreichen. Dabei müssen immer gewisse Proband*innen aus der Stichprobe entfernt werden. Wenn zu einem Probanden bzw. einer Probandin in der einen Gruppe mehrere Proband*innen mit denselben Vortestwerten in der anderen Gruppe existieren, gibt es dabei verschiedene Möglichkeiten die überflüssigen Proband*innen zu entfernen. Dabei sollte man keine einseitige Methode wählen, wie z. B. die Proband*innen mit den größten Lernfortschritten zu entfernen, da dies zum Nachteil der Gruppe mit mehr Proband*innen ist. Bei der Entfernung der Proband*innen mit den niedrigsten Lernfortschritten wäre dies zum Nachteil der Gruppe mit weniger Proband*innen. In jedem Fall sollte also eine Methode gewählt werden, in der symmetrisch um den Median herum aussortiert wird. Dabei ergeben sich die zwei unterschiedlichen Methoden (Vgl. Burde, 2018):

- (1) Entfernung von Schüler*innen mit extremen Lernfortschritten
- (2) Entfernung von Schüler*innen mit mittleren Lernfortschritten

Bei Entfernung der extremen Lernfortschritte ergibt sich bei den beiden Gruppen folgendes Bild (Abb. 61).

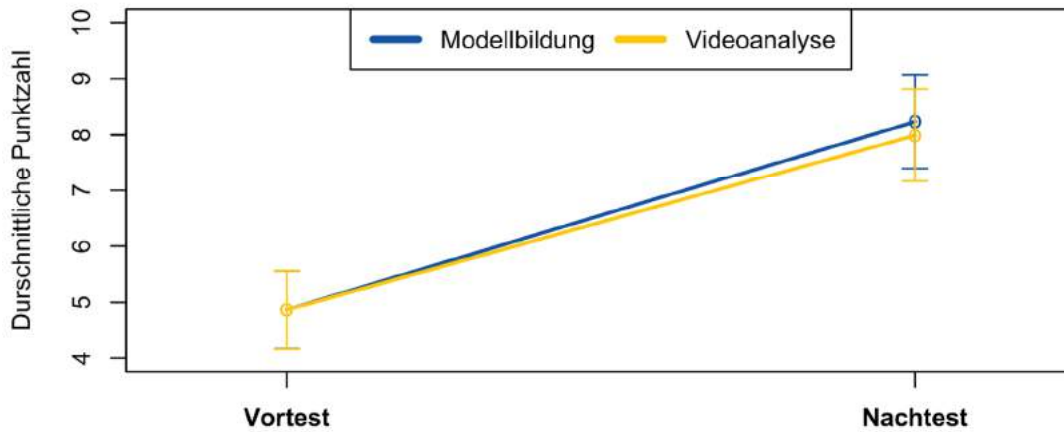
'Matched Samples' durch Entfernung der extremen Lernzuwächse

Abb. 61: Ergebnisse in Vor- und Nachtest in beiden Gruppen (gematcht durch Entfernen der extremen Lernzuwächse, $N=111$ pro Gruppe) mit 95%-Konfidenzintervallen.

Die gematchten Gruppen mit je 111 Teilnehmer*innen haben je eine durchschnittliche Punktzahl von 4,86 Punkten im Vortest. Im Nachtest erreicht die Gruppe der Modellbildung 8,23 Punkte und die Gruppe der Videoanalyse 7,99 Punkte. Es gibt also keinen signifikanten Unterschied zwischen den Gruppen. Wenn die mittleren Lernfortschritte entfernt werden, ergibt sich ein ähnliches Bild (Abb. 62):

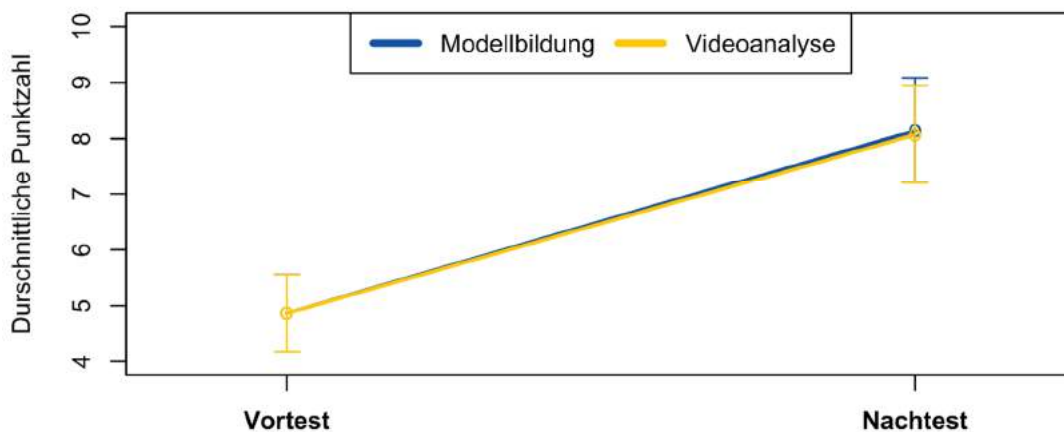
'Matched Samples' durch Entfernung der mittleren Lernzuwächse

Abb. 62: Ergebnisse in Vor- und Nachtest in beiden Gruppen (gematcht durch Entfernen der mittleren Lernzuwächse, $N=111$ pro Gruppe) mit 95%-Konfidenzintervallen.

Die durchschnittliche Punktzahl im Vortest bei je 111 Lernenden ist ebenfalls bei 4,86 Punkten, während nun die Gruppe der Modellbildung im Nachtest durchschnittlich 8,14 Punkte und die Gruppe der Videoanalyse 8,07 Punkte erreicht. Auch beim Vergleich der Gruppen durch eine Parallelisierung zeigt sich kein Unterschied zwischen den Gruppen.

Propensity Score Matching

Eine weitere Möglichkeit des Matchens von Stichproben ist das *Propensity Score Matching* (PSM) (Rosenbaum & Rubin, 1983). Da die bisherigen Varianten ausschließlich auf dem Vortestergebnis basierten, aber weitere Kovariaten gemessen wurden (siehe Kapitel 9.3.1), die das Ergebnis beeinflussen können, sollen diese in das *Matching* einfließen. Auch wenn bereits nachgewiesen wurde, dass sich die Kovariaten im Schnitt in den Gruppen nicht unterscheiden, können diese dennoch Einfluss auf das Nachtestergebnis haben. Da nun aber keine zwei Proband*innen dieselben Werte in allen Kovariaten besitzen, ist eine andere Form des *Matchens* nötig.

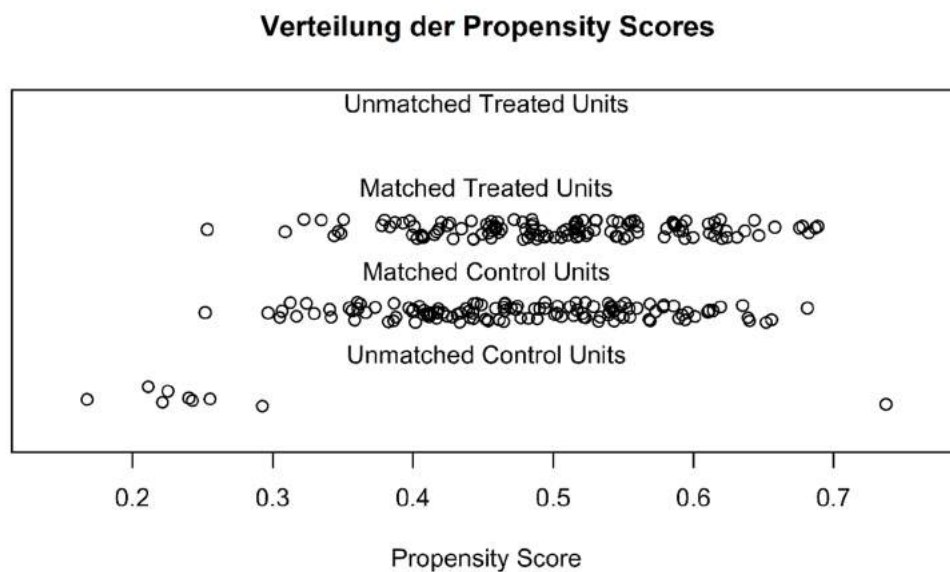


Abb. 63: Propensity Scores für die teilnehmenden Proband*innen für die Gruppe der Modellbildung (Control Units) und die Gruppe der Videoanalyse (Treated Units).

Durch das PSM ist es möglich, die beiden Gruppen auf Basis einer Liste von Kovariaten (Vortestergebnis, Selbstkonzept, Fachinteresse, Computerbezug, Mathematiknote, Physiknote) paarweise zuzuordnen. Die Zuordnung erfolgt durch die „Nächste-Nachbarn-Klassifikation“ (Rosenbaum & Rubin, 1983), d. h. alle Proband*innen werden mit einer Probandin bzw. einem Probanden gematcht, die bzw. der ihm in allen Kovariaten

möglichst ähnlich ist bzw. einen ähnlichen *Propensity Score* besitzt. Es werden also keine Proband*innen gestrichen, wenn kein identisches Match vorliegt, sondern der nächste Nachbar gesucht. Wenn aber z. B. durch unterschiedliche Gruppengröße oder Extremfälle kein*e Partner*in gefunden werden kann, wird diese*r Proband*in aus der Stichprobe entfernt. Durch diese Art des *Matchings* blieben (teilweise durch fehlende Werte in den Kovariaten) 124 Teilnehmer*innen in der Gruppe der Videoanalyse und 133 in der Gruppe der Modellbildung übrig, die gematcht werden (Abb. 63).

Man sieht, dass die *Propensity Scores* bei beiden Gruppen bereits recht ähnlich sind und nur in den Randbereichen Unterschiede bestehen. Die dargestellten *Propensity Scores* stellen einen Gesamtscore dar, der auf den verschiedenen Variablen beruht, die in das *Matching* eingehen. Um die Güte des *Matchings* beurteilen zu können, kann man sich die *Propensity Scores* in den jeweiligen Variablen anschauen (Abb. 64).

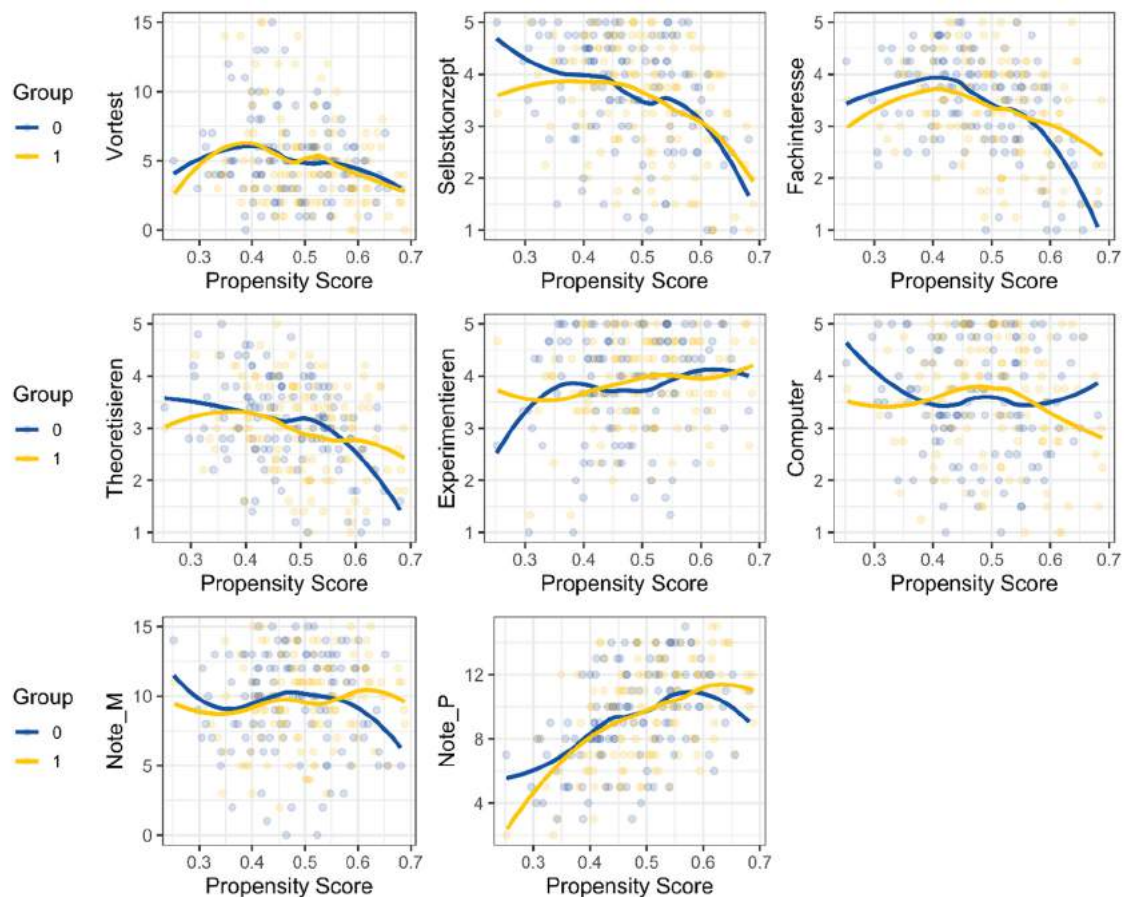


Abb. 64: *Propensity Scores* in den verschiedenen Variablen in beiden Gruppen (0=Modellbildung, 1=Videoanalyse).

In dem Plot sind die errechneten *Propensity Scores* in jeder Kovariate für beide Gruppen aufgetragen. Das *Matching* ist als gut zu bezeichnen, wenn die *Propensity Scores* in allen

Kovariaten für beide Gruppen übereinstimmen, also die Fit-Linien möglichst gleich sind. Dies ist hier nur bedingt gegeben. Die inhaltliche Interpretation der Verläufe der *Propensity Scores* bleibt auch unklar, weswegen das *Matching* selbst nicht als gut einzuschätzen ist.

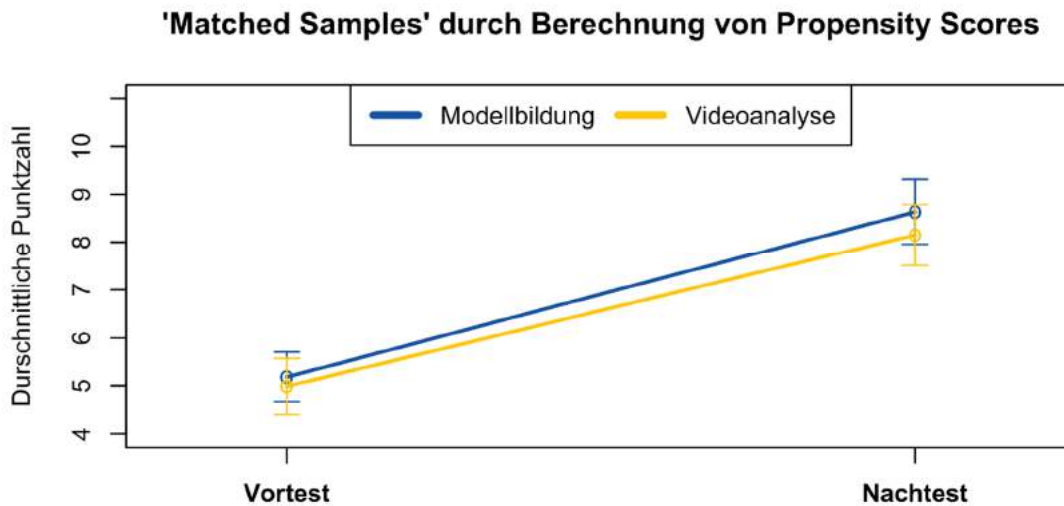


Abb. 65: Ergebnisse in Vor- und Nachtest in beiden Gruppen (gematcht durch „Nächste-Nachbarn“, $N=124$ pro Gruppe) mit 95%-Konfidenzintervallen.

Durch dieses *Matching* wird ebenfalls kein Einfluss der Gruppe sichtbar (Abb. 65). Lediglich der Score in der Skala „Theoretisieren“ und die Physiknoten haben einen signifikanten Einfluss auf das Nachtestergebnis. Aufgrund dessen kann diskutiert werden, ob das *Propensity Score Matching* hier überhaupt einen Mehrwert generiert. Welche Kovariaten insgesamt das Nachtestergebnis beeinflussen, wird ausführlicher in Kapitel 9.3.8 diskutiert, in dem eine Mehrebenenanalyse vorgestellt wird, die die hierarchische Struktur der Proband*innen berücksichtigt.

Kovarianzanalyse

Eine weitere Methode zum Vergleich von Nachtestergebnissen unter Berücksichtigung anderer Kovariaten (z. B. Vortestergebnis) ist die Kovarianzanalyse. Während bei der zweifaktoriellen Varianzanalyse mit Messwiederholung die Unterschiede zwischen Vor- und Nachtestergebnis (also der Lernzuwachs) verglichen werden, testet die Kovarianzanalyse auf Unterschiede im Nachtest unter Kontrolle des Vortestergebnisses. Statt zu fragen, welche Gruppe mehr dazu gelernt hat, wird hier die Frage beantwortet, welche Gruppe mehr dazu gelernt hätte, wenn beide mit einem identischen Vortestergebnis

gestartet wären. Auch für die Durchführung der Kovarianzanalyse müssen (neben den bereits gezeigten) einige Bedingungen erfüllt sein, da sie diese als Annahmen für das erstellte Modell zugrunde legt (Field, 2013):

(1) *Homogenität der Regressionssteigungen:*

Über die Regressionssteigungen kann man sich zunächst einen grafischen Überblick verschaffen. Dazu trägt man die Nachtestergebnisse über den Vortestergebnissen auf und erzeugt für beide Gruppen eine Regressionsgerade (Abb. 66).

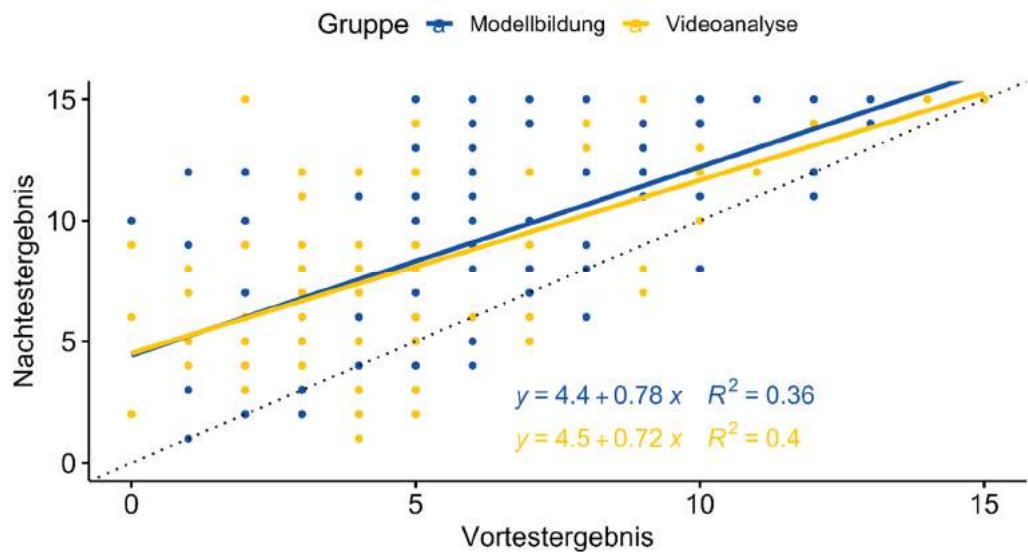


Abb. 66: Nachtestergebnis in Abhängigkeit des Vortestergebnisses für beide Gruppen mit Regressionsgeraden.

Die gestrichelte Linie ist die Winkelhalbierende. Alle Proband*innen unterhalb dieser Linie haben im Nachtest eine geringere Punktzahl erzielt als im Vortest. Umgekehrtes gilt für alle oberhalb liegenden Proband*innen. Dabei scheint die Regressionssteigung nicht durch die Gruppe beeinflusst zu sein. Dies lässt sich durch ein lineares Modell rechnerisch bestätigen. Auch eine ANOVA zeigt, dass die Regressionssteigung unabhängig von der Gruppe ist ($F = 0,291$; $df = 270$; $p = 0,59$). Die Bedingung kann also bestätigt werden.

(2) *Unabhängigkeit der Kovariate vom Gruppeneffekt:*

Das Vortestergebnis ist unabhängig von der Gruppenzugehörigkeit ($t = 1,35$; $df = 270,6$; $p = 0,17$).

Da alle Bedingungen erfüllt sind, kann die ANCOVA durchgeführt werden. Diese bestätigt die vorigen Ergebnisse, da zwar das Vortestergebnis einen höchst signifikanten

Einfluss auf das Nachtestergebnis hat, ein Einfluss der Zugehörigkeit zu einer Gruppe aber nicht zu beobachten ist (Abb. 67).

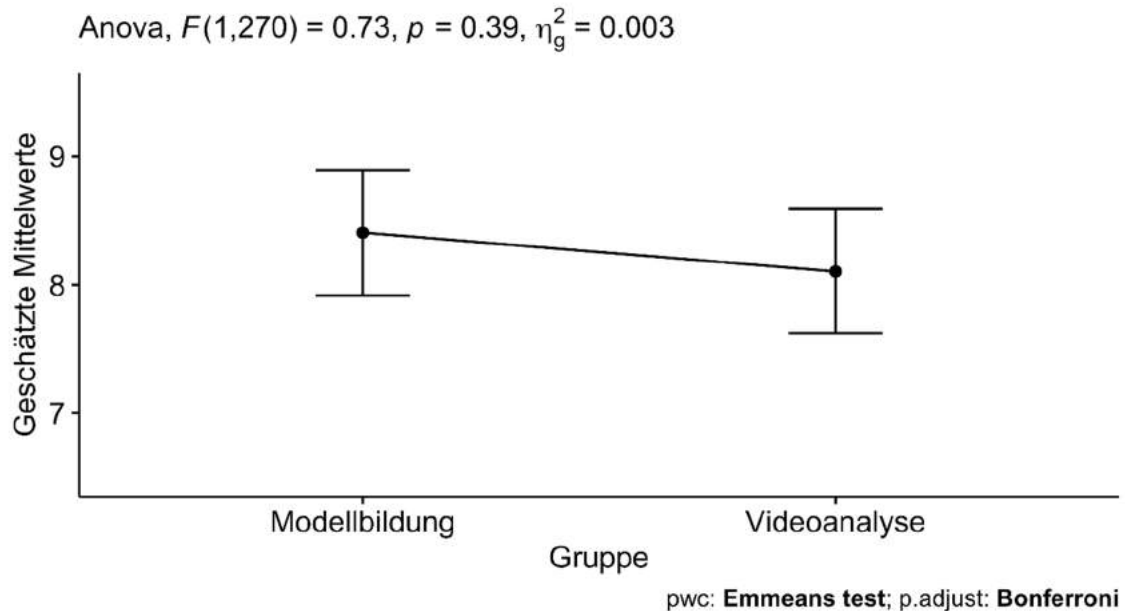


Abb. 67: Geschätzte Mittelwerte für das Nachtestergebnis unter Kontrolle des Vortestergebnisses für beide Gruppen.

Die ANCOVA schätzt das Nachtestergebnis für die Gruppe der Modellbildung auf 8,41 [7,92; 8,89] und das der Videoanalyse auf 8,11 [7,62; 8,59], wenn beide Gruppen mit einem Vortestergebnis von 5,10 starten. Der Unterschied der Nachtestergebnisse ist nicht signifikant.

Vergleich der verwendeten Methoden

Es gibt einige Diskussion, welche dieser Methoden am besten für die Analyse von zwei Gruppen in einer Interventionsstudie im Prä-Post-Design geeignet ist. Diese Diskussion ist so umfangreich, dass hier nicht alle Unterschiede und Vor- und Nachteile erläutert werden können. Dennoch soll kurz auf die verwendeten Methoden und deren Unterschiede eingegangen werden.

Insbesondere die zweifaktorielle Varianzanalyse mit Messwiederholung und die Kovarianzanalyse weisen oberflächlich zunächst gewisse Ähnlichkeiten auf. Es bleibt die Frage, warum beide Verfahren existieren und welche Methode was aussagt. Bei der zweifaktoriellen Varianzanalyse wird betrachtet, welche Variable einen Einfluss auf den Lernzuwachs (also die Differenz zwischen Vor- und Nachtest) hat, während bei der Kovarianzanalyse versucht wird, die Vortestergebnisse zu kontrollieren und unter dieser

Voraussetzung den Einfluss von verschiedenen Variablen auf das Nachtestergebnis zu untersuchen. Das heißt, dass untersucht wird, welche Gruppe einen höheren Nachtest-score hätte, wenn sie mit gleichen Vortestscore gestartet wären. Dabei ist es möglich, dass bei diesen beiden Methoden unterschiedliche Ergebnisse herauskommen, was auch als Lords Paradoxon bezeichnet wird (Lord, 1967b). Dieser Unterschied liegt an einer unterschiedlichen Konzeptualisierung von Änderung. Während manche Autor*innen von einer generellen Unbrauchbarkeit der Differenz zwischen zwei Testergebnissen als Maß für die Änderung ausgehen (Cronbach & Furby, 1970), konnten andere Autor*innen eine hohe Reliabilität dieser Größe nachweisen (z. B. Rogosa, 1988). Auch Lord (1967a) spricht davon, dass die Differenz die natürlichste Größe zur Messung der Änderung ist. Auch die Kovarianzanalyse sah sich bereits einiger Kritik ausgesetzt (Rogosa, 1988). Da in diesem Fall beide Verfahren das gleiche Resultat liefern, muss die Frage der besseren Eignung hier nicht abschließend geklärt werden.

Eine weitere Methode ist die der Parallelisierung bzw. des *Matchings*. Dabei werden die Vortestergebnisse derart kontrolliert, dass Kohorten erzeugt werden, welche in dieser Variablen übereinstimmen. Mit diesen Kohorten kann dann das Nachtestergebnis betrachtet werden. Die Parallelisierung ist die einzige Möglichkeit, wirklich gleiche Startbedingungen (zumindest in Bezug auf eine Variable) zu erzeugen. Der Nachteil ist, dass einige Proband*innen dazu entfernt werden müssen, die Methode der Entfernung das Ergebnis beeinflussen kann und die externe Validität reduziert (Bortz & Döring, 2006). In der stärkeren von beiden Gruppen im Vortest werden eher starke Proband*innen entfernt, während in der schwachen Gruppe eher schwache Proband*innen entfernt werden. Durch die Regression zur Mitte kann es dadurch passieren, dass auch nach der Parallelisierung die ursprünglich bessere Gruppe im Nachtest besser abschneidet, ohne dass ein inhaltlicher Unterschied vorliegt, also beide gleich viel gelernt haben (Bortz & Döring, 2006).

Jede dieser Methoden ist ein Versuch, das Versäumnis einer optimal randomisierten Zuteilung auf verschiedene Weise auszugleichen. Welche Methode am besten geeignet ist, hängt von vielen Parametern ab. Da hier alle Methoden zum gleichen Ergebnis kamen, sind die Unterschiede wenig relevant und es kann davon ausgegangen werden, dass es keinen messbaren Unterschied zwischen den beiden Gruppen gibt.

9.3.7 Unterscheidung zwischen Geschlechtern

In den letzten Jahren wurden Geschlechtsunterschiede in der Physikdidaktik intensiv untersucht. Es konnte dabei unter anderem gezeigt werden, dass es kontextabhängig ist, wann Mädchen bzw. Jungen ein hohes Interesse an der Physik zeigen (Heinicke, 2019).

Dabei konnte auch gezeigt werden, dass Physikunterricht, der den Schülerinnen besonders gefällt, für beide Geschlechter erfolgreich ist (Hoffmann et al., 1998). Auch beim Computereinsatz stellt sich immer wieder die Frage, ob dieser Jungen bevorzugt. Nachfolgend wird nur zwischen weiblich und männlich unterschieden, da nur drei Schüler*innen angaben, einem anderen Geschlecht anzugehören und diese Gruppe damit statistisch nicht auswertbar ist.

Um den Einfluss des Geschlechts auf das Nachtestergebnis (unter Kontrolle des Vortestergebnisses und ohne Beachtung der Gruppenzugehörigkeit) zu ermitteln, wurde eine ANCOVA durchgeführt. Dazu müssen wieder die oben besprochenen Bedingungen erfüllt werden:

(1) *Linearität und Homogenität der Regressionssteigungen:*

In den Daten kann von einer ausreichenden Linearität ausgegangen werden (Abb. 68). Die Regressionssteigungen können ebenfalls als identisch angesehen werden. Es gibt keinen Interaktionseffekt zwischen Geschlecht und Vortestergebnis.

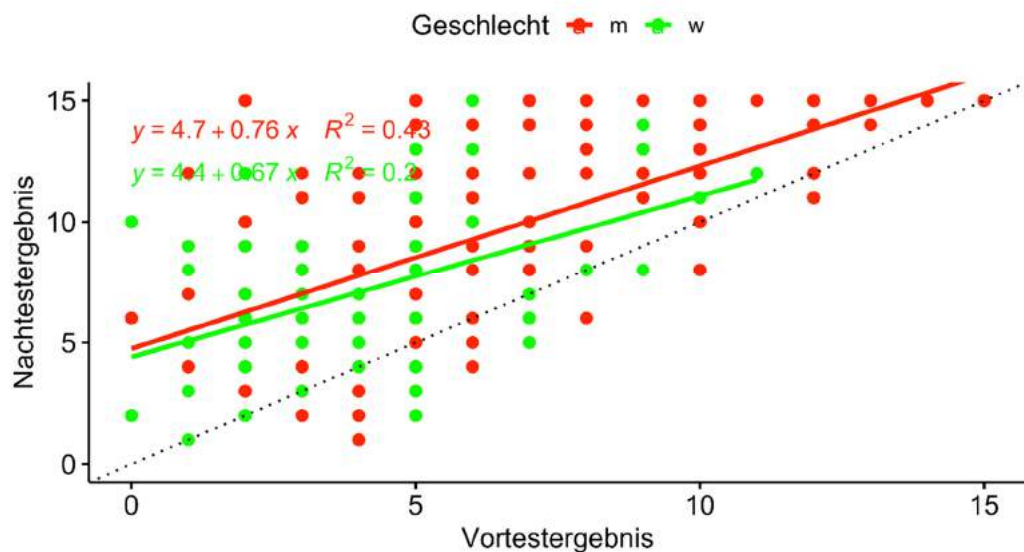


Abb. 68: Nachtestergebnis in Abhängigkeit des Vortestergebnisses für männlich und weiblich mit Regressionsgeraden.

(2) *Normalverteilung der Residuen:*

Die Residuen in beiden Geschlechtern können als ausreichend normalverteilt angesehen werden (Abb. 69).

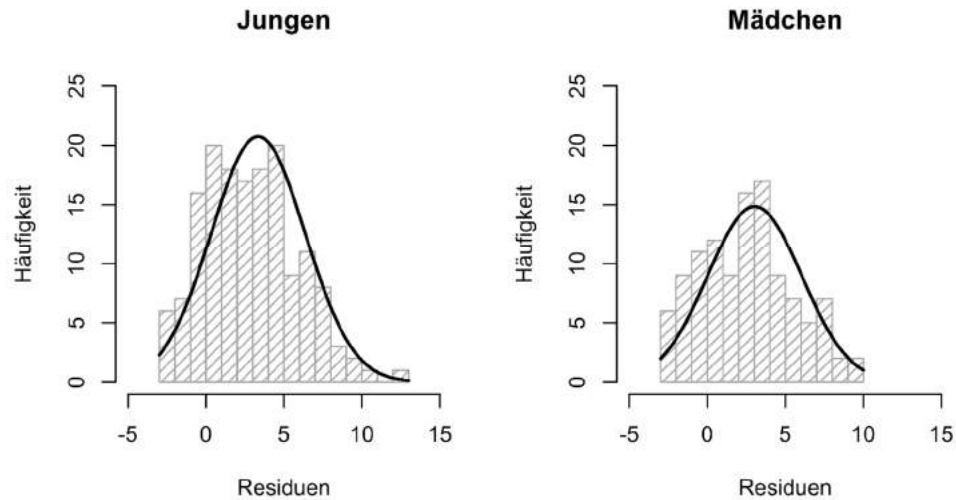


Abb. 69: Residuen in beiden Geschlechtern.

(3) *Unabhängigkeit der Kovariate vom Gruppeneffekt:*

Die Vortestergebnisse sind unterschiedlich in beiden Gruppen ($t = -4,65$; $p < 0,001$), womit die Bedingung verletzt ist. Nach Weber (2009) kann eine ANCOVA dennoch durchgeführt werden, wenn sich die Vortestergebnisse bei der Gruppen ausreichend überlappen. Zwar ist ausreichend hier ein subjektives Kriterium, verhindern soll dies aber die Durchführung einer ANCOVA bei der die Ergebnisse in beiden Gruppen vollkommen unterschiedlich und damit vollkommen von der Gruppe abhängig sind (Miller & Chapman, 2001) und keinerlei bis wenig Überlappung aufweisen. Die Überlappung der Vortestergebnisse in den beiden Gruppen lässt sich Abb. 70 entnehmen.

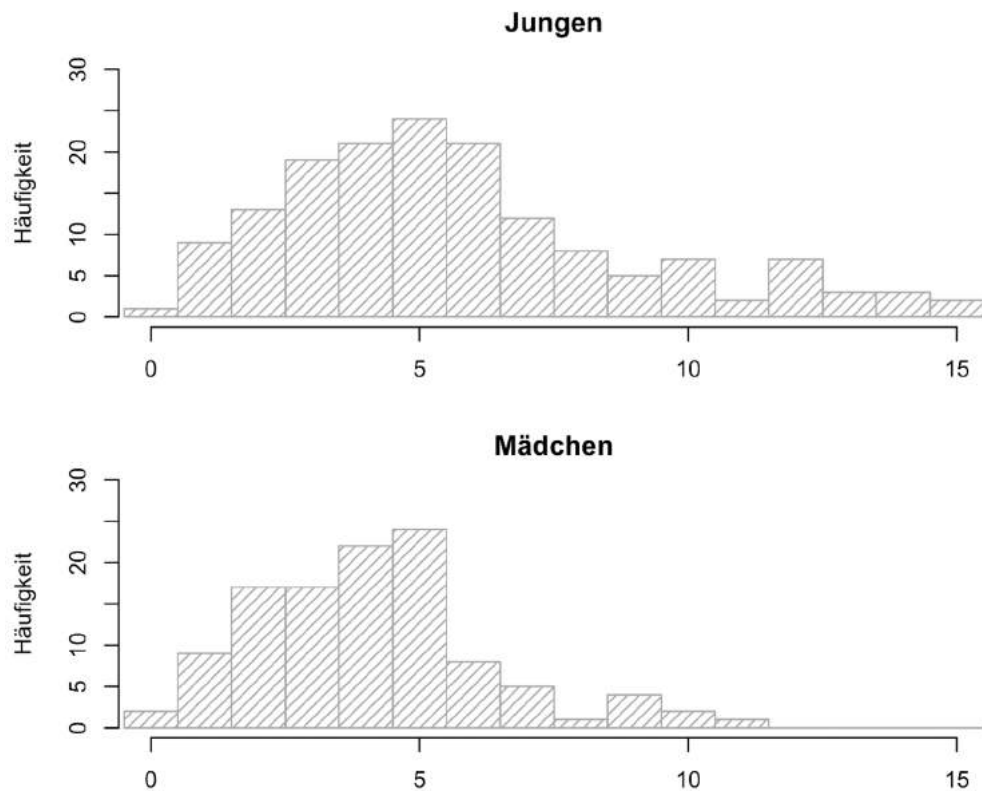


Abb. 70: Histogramm der Vortestergebnisse bei Jungen und Mädchen.

Die Überlappung der Vortestergebnisse lässt sich als ausreichend für die Durchführung einer ANCOVA ansehen.

(4) *Homogenität der Varianzen:*

Der Levene-Test ist mit $p = 0,94$ nicht signifikant. Es kann von einer Homogenität der Varianzen ausgegangen werden.

Insgesamt gesehen kann folglich eine ANCOVA durchgeführt werden. Das Geschlecht hat auch bei Kontrolle des Vortests noch einen signifikanten Einfluss auf das Nachtestergebnis (Abb. 71).

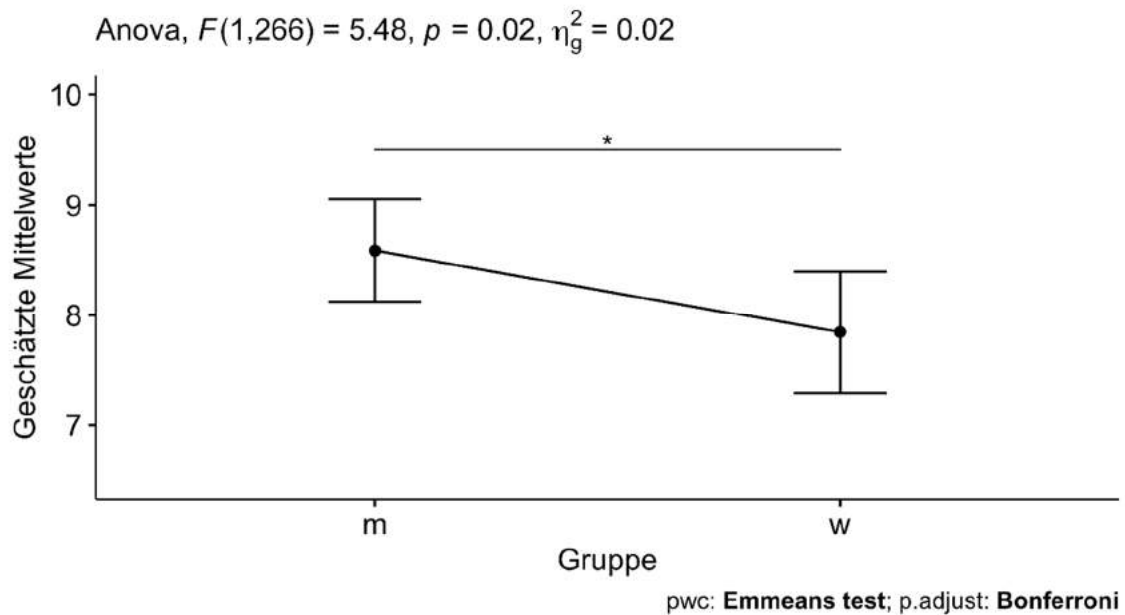


Abb. 71: Geschätzte Mittelwerte für die Nachtestergebnisse von Jungen und Mädchen unter Kontrolle des Vortestergebnisses.

Bei gleichem Vortestergebnis (5,06) schätzt die ANCOVA für Jungen ein durchschnittliches Nachtestergebnis von 8,59 [8,12;9,05] und für Mädchen von 7,85 [7,29;8,40], was einem signifikanten Unterschied mit kleiner Effektstärke entspricht. Eine Interaktion mit der Gruppenzugehörigkeit lässt sich aber nicht beobachten, was bedeutet, dass die Unterschiede unabhängig von der Gruppenzugehörigkeit existieren.

Unterschiedliche Scores von Jungen und Mädchen konnten im Bereich von Leistungstests in den Naturwissenschaften schon häufig beobachtet werden (z. B. Beichner, 1990). Dies wird unter anderem darauf zurückgeführt, dass Naturwissenschaft als etwas männliches angesehen wird (Vockell & Lobonc, 1981) und dass die Geschlechter aufgrund unterschiedlicher Erwartungshaltungen von Lehrkräften unterschiedlich behandelt werden (Brophy & Good, 1970). Dass in der Intervention Jungen möglicherweise mehr lernen als Mädchen, ist zwar nicht erwünscht, aber auch nicht verwunderlich.

9.3.8 Hierarchisches lineares Modell

Die Interventionen dieser Studie fanden im Klassenverband statt (Abb. 72). Das heißt, dass die Schüler*innen in Klassen gegliedert sind, die wiederum von einer Lehrkraft unterrichtet werden. Dabei hat eine Lehrkraft mit zwei Klassen an der Intervention teilgenommen, alle anderen Lehrkräfte mit je einer Klasse.

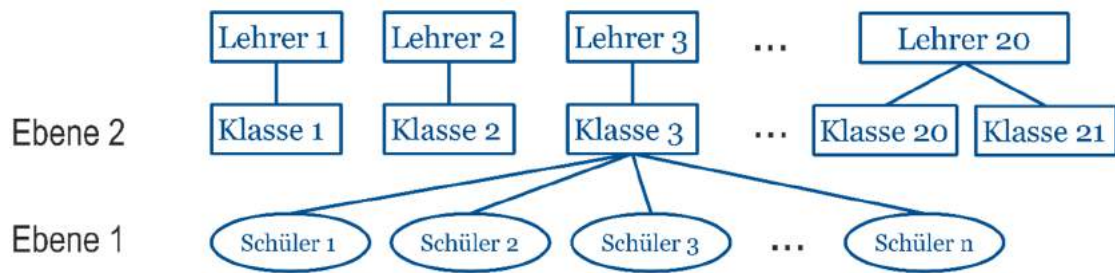


Abb. 72: Darstellungen der Ebenenstruktur der Intervention (Lehrer steht hier und bei der Benennung der Ebenen in diesem Kapitel aus Platzgründen für Lehrer*in und Schüler für Schüler*in).

Auch wenn die Arbeit während der Interventionen ausschließlich in Zweiergruppen (oder im Notfall in Dreiergruppen) erfolgte, die Interventionen nicht in der Schule stattfanden und nicht von der Lehrkraft durchgeführt wurden, lassen sich Klasseneffekte nicht ausschließen, die beispielsweise auf den Unterricht vor der Intervention zurückzuführen sind. Diese Effekte sind hier eher als gering einzustufen, eine Nicht-Beachtung kann aber zu einer Überschätzung der Stichprobengröße führen, was wieder zu einer zu geringen Irrtumswahrscheinlichkeit führt (Kuhn, 2014). Um die Struktur von Schüler*innen in Klassen abbilden und sowohl den Einfluss von Variablen auf Schülerebene (z. B. Vortestergebnis) als auch auf Klassenebene (z. B. Gruppenzugehörigkeit) auf die abhängige Variable untersuchen zu können, eignet sich ein zweistufiges hierarchisches lineares Modell (HLM). Da nur eine Lehrkraft mit zwei Klassen an der Intervention teilnahm, ist es bei den vorliegenden Daten nicht sinnvoll eine dritte Ebene (Lehrerebene) einzuführen. Die Lehrervariablen fließen auf Klassenebene in das Modell ein.

Über die statistischen Voraussetzungen einer solchen Mehrebenenanalyse gibt es keine Einigkeit (Ditton, 1998; Hox, 2010). Dabei reichen die Empfehlungen von zehn Einheiten auf der höchsten modellierten Ebene bis hin zu mindestens 30 oder gar 50 Einheiten. Neuere Untersuchungen zeigen aber, dass die Stabilität der Mehrebenenanalyse ab zehn Einheiten auf der höheren Ebene zufriedenstellend ist (Gollwitzer et al., 2007; Kuhn, 2014; Maas & Hox, 2004; Maas & Hox, 2005).

Für die hier vorliegenden 21 Einheiten auf der höchsten Ebene kann also eine Mehrebenenanalyse durchgeführt werden. Im Folgenden wird zunächst das allgemeine Vorgehen der dieser vorgestellt, wonach die vorliegenden Daten auf ihre Eignung für eine solches Modell überprüft werden. Letztlich werden die schrittweise Erstellung des Modells und das finale Modell vorgestellt, woraus hervorgeht, welche Prädiktoren einen messbaren Einfluss auf das Nachtestergebnis haben.

Allgemeines Vorgehen

In Anlehnung an Kuhn (2014) und Burde (2018) wird die allgemeine Vorgehensweise und die Benennung der Modelle beibehalten. Im vorliegenden Fall werden zwei Ebenen berücksichtigt – die Schülerebene (Ebene 1) und die Klassenebene (Ebene 2). Das lineare hierarchische Modell entspricht einer linearen Regression, die für jede Modellebene eine Gleichung formuliert. Diese Gleichungen sind durch die Regressionskoeffizienten miteinander verknüpft (Tab. 27). Das grundlegendste Modell (das sogenannte Nullmodell) kann folgendermaßen dargestellt werden (Vgl. Burde, 2018):

Tab. 27: Beschreibung der Einflussfaktoren des Nullmodells.

Ebene 1: $Post_{ij} = \beta_{0j} + e_{ij}$	• Der Index i steht für Schüler*in, der Index j für Klassen. • $Post_{ij}$ steht für das Nachtestergebnis des i-ten Lernenden in der j-ten Klasse. • β_{0j}: Regressionskonstante (durchschnittliches Ergebnis der Klasse j). • e_{ij}: Zufallseffekt auf Ebene 1 (im Nullmodell handelt es sich hier um die Differenz des Ergebnisses des i-ten Lernenden aus der j-ten Klasse und dem Durchschnitt der j-ten Klasse).
Ebene 2: $\beta_{0j} = \gamma_{00} + r_{0j}$	• γ_{00}: Mittelwert des Nachtests über alle Schüler*innen und Klassen. • r_{0j}: Zufallseffekt auf Ebene 2 (im Nullmodell die Differenz des durchschnittlichen Nachtestergebnisses in Klasse j und des Gesamtdurchschnitts γ_{00}).

Durch Einsetzen der unteren Gleichung in die obere erhält man das gemischte Modell:

$$Post_{ij} = \gamma_{00} + r_{0j} + e_{ij}$$

Neben dem Gesamtmittelwert befinden sich außerdem die beiden Zufallseffekte in den jeweiligen Ebenen in dem Modell, aus denen die Varianzkomponenten in beiden Ebenen ermittelt werden können. Aus den Varianzkomponenten der Schüler- und Klassenkomponenten kann dann der Intraklassenkorrelationskoeffizient (ICC) bestimmt werden. Dieser ist ein Maß für den Anteil der Varianz auf einer Ebene an der Gesamtvarianz. Er gibt

an, wieviel Varianz des Nachtestergebnisses maximal durch Prädiktoren auf der jeweiligen Ebene erklärt werden kann und berechnet sich für die beiden Ebenen wie folgt:

$$ICC_{Schüler} = \frac{Var(e_{ij})}{Var(e_{ij}) + Var(r_{0j})}$$

und:

$$ICC_{Klasse} = \frac{Var(r_{0j})}{Var(e_{ij}) + Var(r_{0j})}$$

Für die Suche des optimalen Modells wird eine *Step-up*-Strategie nach Raudenbush und Bryk (2010) angewandt. Das heißt, dass zum Nullmodell schrittweise neue Faktoren hinzugefügt und diese auf Signifikanz überprüft werden. Ein Faktor bleibt dann im Modell, wenn das Modell mit diesem Faktor signifikant ($p < 0,05$) besser zu den Daten passt als das Modell ohne diesen Faktor. Da das so gewonnene finale Modell nicht das einzig mögliche ist, werden die Prädiktoren theoriegeleitet eingefügt. Als Faktoren kommen auf den verschiedenen Ebenen folgende Prädiktoren in Frage (Tab. 28 & Tab. 29):

Tab. 28: Überblick über mögliche Prädiktoren für das Nachtestergebnis (Teil 1).

Ebene	Name	Skala	Wertebereich
1	Vortest	Intervall ²⁴	[0;15]
1	Fachinteresse	Intervall (Likert) ²⁵	[1;5]
1	Selbstkonzept	Intervall (Likert)	[1;5]
1	Interesse an Computern	Intervall (Likert)	[1;5]

²⁴ Punkte in einem Leistungstest werden häufig als intervallskaliert betrachtet, was eine äquidistante Verteilung der Punkte voraussetzt. Dies ist durch die klassische Testtheorie eigentlich nicht zu gewährleisten. Diese Problematik kann durch den Einsatz einer Rasch-Analyse behoben werden. Diese wurde hier zur Überprüfungszwecken im Nachhinein durchgeführt. Die Personenfähigkeit der Proband*innen hängt in dieser Studie, mit Ausnahme der Randbereiche (≤ 1 und ≥ 14 Punkte), linear mit den Punktzahlen zusammen. Dies bestätigt, dass die Punktzahlen bei Inkaufnahme eines eher geringen Fehlers als intervallskaliert angesehen werden können.

²⁵ Likert-Skalen sind strenggenommen ebenfalls Ordinalskalen, werden aber häufig als intervallskaliert angesehen (Wu und Leung (2017)). Die Auswirkungen dieser statistischen Ungenauigkeit sind i. d. R. gering und können an dieser Stelle nicht ausführlich diskutiert werden.

Tab. 29: Überblick über mögliche Prädiktoren für das Nachtestergebnis (Teil 2).

Ebene	Name	Skala	Wertebereich
1	Interesse an Theorie	Intervall (Likert)	[1;5]
1	Interesse an Experimenten	Intervall (Likert)	[1;5]
1	Geschlecht	Nominal	[d, m, w]
1	Physiknote	Intervall ²⁶	[0;15]
1	Mathematiknote	Intervall	[0;15]
1	Lernhinderlicher Cognitive Load	Intervall (Likert)	[1;6]
1	Lernrelevanter Cognitive Load	Intervall (Likert)	[1;6]
2	Gruppe	Nominal	[Mod., Vid.]
2	Kursart	Nominal	[GK, LK]
2	Fachwissen Lehrkraft	Intervall	[0;15]
2	Leistung Klasse	Intervall	[0;15]

Bei der Schätzung der Modelle gibt es unterschiedliche Verfahren, über deren Nutzung kein genereller Konsens besteht (Burde, 2018). Bei der *Maximum-Likelihood*-Schätzung (ML) werden sowohl die festen als auch die zufälligen Effekte mittels *Maximum-Likelihood* geschätzt. Davon unterscheidet sich das *Restricted Maximum Likelihood*-Verfahren (REML), bei dem ausschließlich die Zufallseffekte per *Maximum Likelihood*-Schätzung ermittelt werden. Das REML-Verfahren ist nach einigen Autor*innen für die Schätzung solcher Modelle grundsätzlich zu bevorzugen, da REML speziell bei einer geringeren Anzahl von Gruppen bessere Ergebnisse liefert (Browne, 1998). Wenn jedoch ein Modellvergleich durchgeführt werden soll, muss dieser mithilfe des ML-Verfahrens erfolgen (Field, 2007), welches auch im Bereich der festen Effekte Vorteile haben soll (Raudenbush & Bryk, 2010; Twisk, 2006). Beide Verfahren liefern aber ähnliche Ergebnisse, insbesondere bei balancierten Designs und vielen Gruppen. Hier wird für den Fit

²⁶ Ähnlich wie bei Punkten in einem Leistungstest kann bei einer Schulnote ebenfalls nicht von einer intervallskalierten Variablen ausgegangen werden. Ähnlich wie oben beschrieben, wird die Intervallskalierung aber für die Mehrebenenanalyse angenommen.

der Modelle und die Schätzung der Parameter die REML-Methode genutzt, während für die Modellvergleiche die Modelle mithilfe von ML erneut gefittet werden. Daher kann es in der Folge zu Unterschieden in den berichteten Modellparameter kommen.

Für die Beschreibung der Modellgüte gibt es verschiedene Parameter. In der Regel werden AIC (Akaike Informationskriterium), BIC (Bayes'sche Informationskriterium) und Pseudo- R^2 berichtet. Das AIC bewertet die Mächtigkeit des Modells, die abhängige Variable zu erklären, und nutzt dazu den Wert der *Log-Likelihood*. Dabei ist das AIC nie ein absolutes Kriterium, sondern dient lediglich zum Modellvergleich. Bevorzugt wird der niedrigere Wert beim AIC, während die Anzahl an geschätzten Parametern den Wert erhöht, also als Strafterm fungiert. Das BIC ist dem AIC ähnlich. Hier wird lediglich die Stichprobengröße zusätzlich bestraft. Eine andere Art der Bewertung eines Modells stellt das Bestimmtheitsmaß Pseudo- R^2 dar. Es kann von 0 bis 1 reichen, wobei 1 auf eine gute Modellpassung hindeutet. Es gibt verschiedene Ansätze das Pseudo- R^2 zu bestimmen. Häufig wird hier ebenfalls auf den Wert des *Log-Likelihoods* des Modells zurückgegriffen und dieser mit dem des Nullmodells ins Verhältnis gesetzt (Raudenbush & Bryk, 2010). Auch die Devianz ist ein Gütekriterium für die Modelle. Sie kann durch den Wert den *Log-Likelihoods* berechnet werden: $D = -2 \cdot \log(L)$ (Kuhn, 2010). Je besser das Modell zu den Daten passt, desto geringer ist die Devianz (Pötschke, 2019). Für die Modellvergleiche wird das Bayes'sche Informationskriterium herangezogen, während für die Hauptmodelle auch das Pseudo- R^2 berichtet wird.

Erstellen des Modells

M0: Null-Modell (Intercept-Only-Modell)

Das Null-Modell enthält nur die abhängige Variable und dient dazu abzuschätzen, welche Varianz auf den beiden Ebenen erklärbar ist. Dies dient u. a. der Überprüfung der Eignung der Daten für eine Mehrebenenanalyse. Das Nullmodell hat dabei folgende Struktur (Tab. 30):

Tab. 30: Struktur des Nullmodells.

H. Modell	Ebene 1	$Post_{ij} = \beta_{0j} + e_{ij}$
	Ebene 2	$\beta_{0j} = \gamma_{00} + r_{0j}$
Gemischtes Modell	$Post_{ij} = \gamma_{00} + r_{0j} + e_{ij}$	
Modellgüte	$BIC = 1415,46$	

Der sich durch das Modell ergebende Intercept (γ_{00}) von 8,42 entspricht der durchschnittlichen Punktzahl aller teilnehmenden Schüler*innen im Nachtest (Tab. 31).

Tab. 31: Kennwerte der im Modell enthaltenen Parameter.

						95%-Konfidenzintervall	
Parameter	Schätzung	SE	df	t-Wert	Sig.	Untergr.	Obergr.
γ_{00}	8,42	0,48	18,65	17,72	$p < 0,001$	7,47	9,38

Außerdem dargestellt sind in der Tabelle der Standardfehler (SE) der Schätzung für den Parameter, die Freiheitsgrade (df), der t-Wert und das 95 %-Konfidenzintervall. Die Signifikanz des Parameters im Nullmodells ist nicht von großer Bedeutung. Neben den in der obigen Tab. 31 dargestellten festen Effekte des Modells (*Fixed Effects*) kann das Modell auch Auskunft über die Zufallseffekte geben (Tab. 32):

Tab. 32: Darstellung der im Nullmodell enthaltenen Zufallseffekte.

Ebene	Parameter	Varianz	Std.-Abweichung	ICC
Schüler	e_{ij}	10,60	3,26	0,75
Klasse	r_{0j}	3,58	1,89	0,25

Diese Zufallseffekte geben an, welcher Teil der Varianz auf den beiden Ebenen vorliegt. Interessant ist vor allem der Anteil der Varianz auf der Klassenebene an der Gesamtvarianz (Kuhn, 2010).

Bei den vorliegenden Daten ergibt sich ein ICC von $ICC = 0,25$, was bedeutet, dass bis zu 25 % der Varianz durch die Klassenzugehörigkeit erklärbar ist. Damit eignen sich die Daten nach Hartig und Rakoczy (2010) nicht nur für eine Mehrebenenanalyse, diese ist sogar notwendig, um eine Alpha-Fehler-Inflation zu vermeiden. Dies bestätigt sich auch, wenn man das Modell, welches die Klassenstruktur berücksichtigt, mit dem Fit eines Modells vergleicht, was dies nicht tut. Die Zufallseffekte im Nullmodell können graphisch dargestellt werden (Abb. 73).

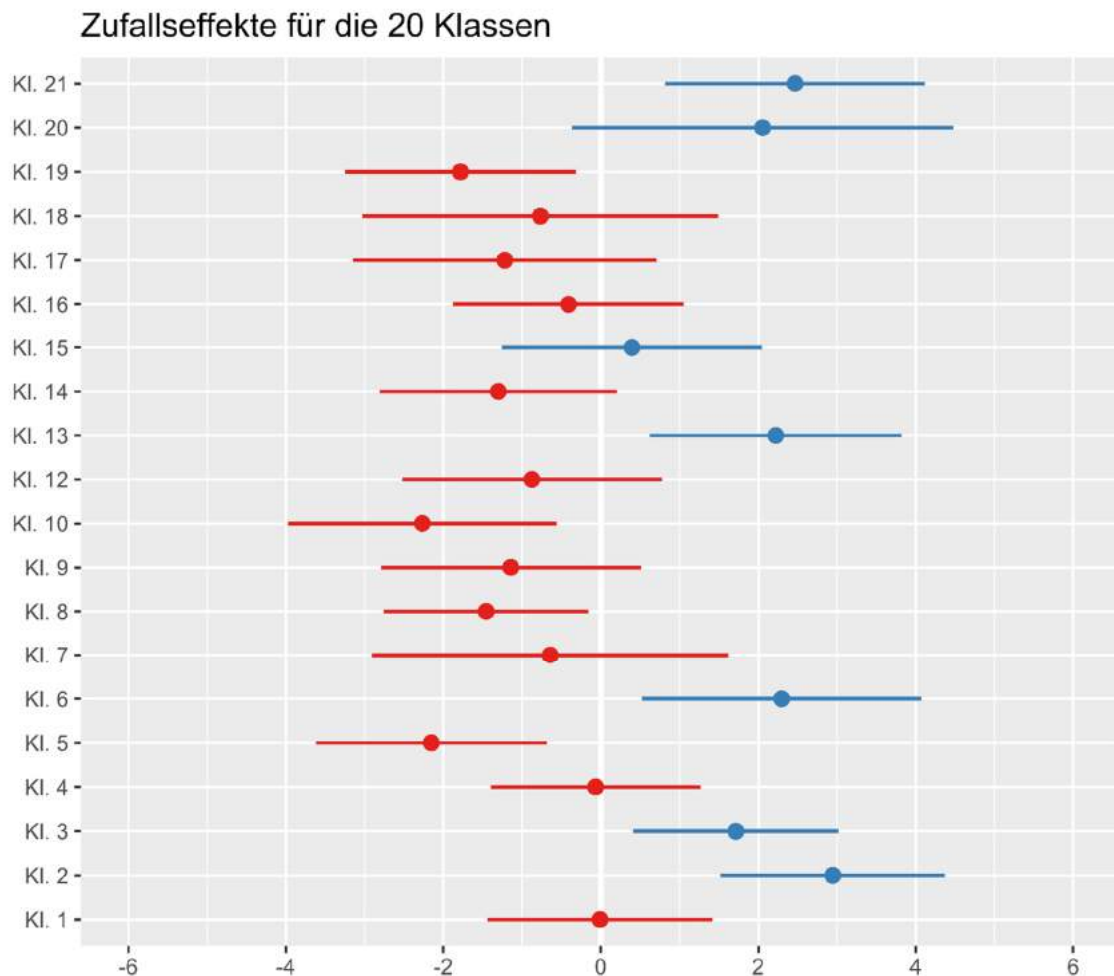


Abb. 73: Zufallseffekte r_{0j} für 20 Klassen. Eine Klasse mit wenigen Schüler*innen ist aufgrund von fehlenden Werten herausgefallen.

In der Grafik sind also r_{01} bis r_{021} dargestellt. Mithilfe eines *Likelihood-Ratio-Tests* kann überprüft werden, ob sich die Varianz auf Klassenebene von Null unterscheidet, die Zufallseffekte also signifikant sind. Die Gleichheit der Varianzen kann auf Basis des Signifikanzniveaus von $p < 0,001$ abgelehnt werden. Die Daten sind insgesamt dafür geeignet, ein Mehrebenenmodell zu erstellen.

M1a: Der Haupteffekt „Vortest“ (*Random-Intercept*)

Der größte Einfluss auf das Nachtestergebnis ist vom Vortestergebnis zu erwarten, da dies in den vorigen Analysen bereits gezeigt wurde und bei einer einmaligen Intervention nach dem Unterricht auch inhaltlich naheliegend ist. Abb. 74 zeigt einen positiven Zusammenhang zwischen Vortest- und Nachtestergebnis, wobei die Regression hier die Klassenstruktur zunächst ignoriert.

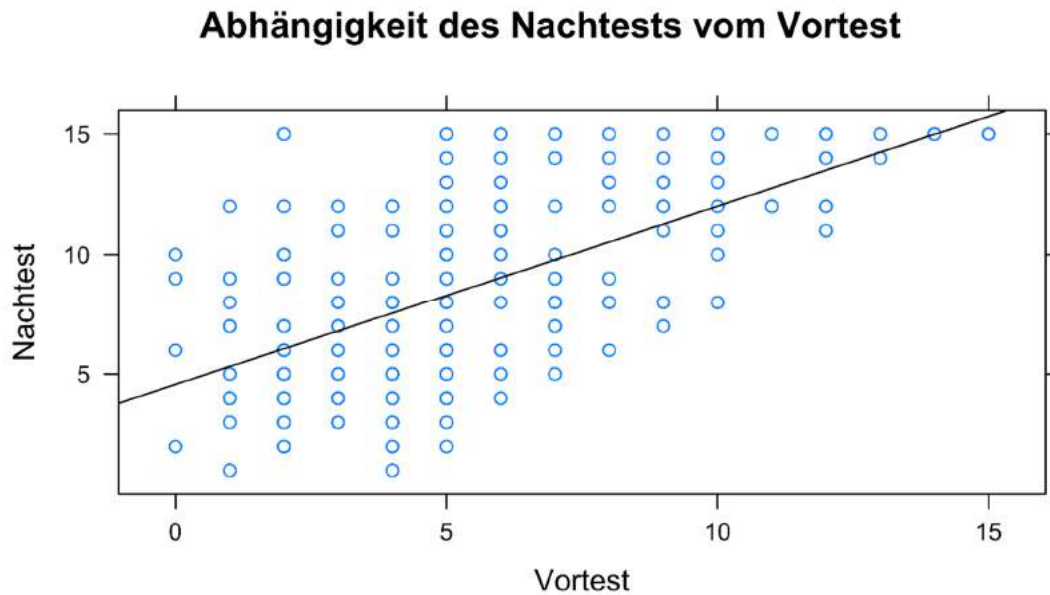


Abb. 74: Nachtestergebnisse in Abhängigkeit der Vortestergebnisse ohne Berücksichtigung der Klassenstruktur mit Regressionsgerade. Ein Punkt steht für ein oder mehrere Schüler*innen.

Die vorige Analyse hat gezeigt, dass die Klassenstruktur in den hier vorliegenden Daten berücksichtigt werden sollte. Dadurch wird der Effekt „Vortest“ als erstes in das lineare hierarchische Modell aufgenommen. Das Modell wird dadurch durch den Faktor γ_{10} ergänzt (Tab. 33).

Tab. 33: Struktur des Modells 1a. Der neue Faktor im gemischten Modell wird fett dargestellt.

H. Modell	Ebene 1	$Post_{ij} = \beta_{0j} + \beta_{1j}Pre + e_{ij}$
	Ebene 2	$\beta_{0j} = \gamma_{00} + r_{0j}$ $\beta_{1j} = \gamma_{10}$
Gemischtes Modell	$Post_{ij} = \gamma_{00} + \gamma_{10}Pre + r_{0j} + e_{ij}$	
Modellgüte	$BIC = 1326,53$	

Durch die Hinzunahme dieses Effekts ergibt sich ein signifikant besser zu den Daten passendes Modell, wie man im Vergleich der beiden Modelle sieht (Tab. 34).

Tab. 34: Vergleich des Modells 1a mit dem Nullmodell.

Modell	Parameter	AIC	BIC	Loglik	Devianz	Chisq	df	p
Modell0	3	1405	1416	-700	1399			
Modell1a	4	1308	1322	-660	1300	99,2	1	<0,001

Die Kennwerte AIC und BIC sind für das Modell 1a besser. Zudem passt das Modell signifikant besser ($p < 0,001$) zu den Daten als das Nullmodell. Der Effekt „Vortest“ wird also im Modell belassen. Weitere in der Tabelle dargestellte Werte sind der *Log-Likelihood* des Modells (Loglik), die Devianz, der Chi-Quadrat-Wert (Chisq), die Freiheitsgrade (df, Differenz der Parameter) und die Signifikanz (p). Die festen Effekte, die das Modell dann schätzt, sind folgende (Tab. 35):

Tab. 35: Kennwerte der im Modell 1a enthaltenen Parameter.

						95%-Konfidenzintervall	
Parameter	Schätzung	SE	df	t-Wert	Sig.	Untergr.	Obergr.
γ_{00}	4,88	0,43	56,37	11,43	$p < 0,001$	4,05	5,74
γ_{10}	0,69	0,06	244,21	11,21	$p < 0,001$	0,57	0,81

Das Modell 1a schätzt, dass der Prädiktor Vortestergebnis signifikant zum Nachtestergebnis beiträgt und dass jeder zusätzliche Punkt im Vortest zu einer Steigerung von 0,69 Punkten im Nachtest führt. Die Zufallseffekte sind im Modell 1a folgende (Tab. 36):

Tab. 36: Darstellung der im Modell 1a enthaltenen Zufallseffekte.

Ebene	Parameter	Varianz	Std.-Abweichung	Ebenenspez. R^2
Schüler	e_{ij}	7,67	2,77	0,276
Klasse	r_{0j}	1,02	1,01	0,715

Der Prädiktor Vortest erklärt also $\frac{10,60-7,67}{10,60} = 27,6\%$ der Varianz auf Ebene 1, $\frac{3,58-1,02}{3,58} = 71,5\%$ auf Ebene 2 und damit $\frac{(10,60+3,58)-(7,67+1,02)}{10,60+3,58} = 38,7\%$ der Gesamtvarianz (Ebenenspez. R^2) und wird im Modell gelassen. Dabei geht es hier um die

Varianz, die der Prädiktor Vortestergebnis in diesem einfachen Modell erklärt. Durch Hinzunahme anderer Prädiktoren kann sich dies in anderen Modellen noch verringern. Um die Varianzaufklärung eines bestimmten Prädiktors im finalen Modell zu bestimmen, muss die Varianzaufklärung dieses finalen Modells mit dem Modell verglichen werden, was sich durch Streichung dieses bestimmten Prädiktors ergibt. Dies wird am Ende des Kapitels genauer diskutiert.

M1b: Zufallseffekt für „Vortest“ (*Random-Intercept* und *-Slope*)

Um zu untersuchen, ob sich der Effekt des Vortests zwischen den Klassen unterscheidet, also manche Klassen von einem guten Vortestergebnis mehr oder weniger profitieren als andere, kann zusätzlich ein Modell mit *Random-Slope* erstellt werden. Abb. 75 zeigt dabei, wie die Regressionsgeraden für den Zusammenhang zwischen Vortestergebnis und Nachtestergebnis in jeder Klasse aussehen.

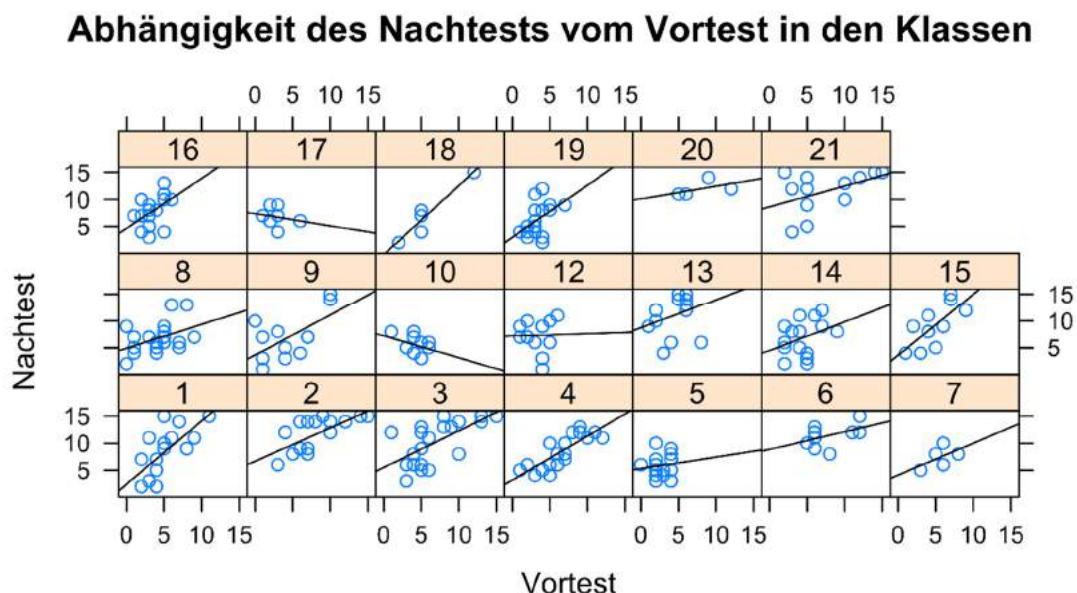


Abb. 75: Nachtestergebnis in Abhängigkeit des Vortestergebnisses in allen Klassen mit Regressionsgeraden.

Die Steigungen der Regressionsgerade unterscheiden sich teilweise stark, was zunächst dafürspricht, dass die Steigung im Modell ebenfalls von der Klasse abhängig sein sollte. Jedoch sind die Regressionen teilweise auf Basis weniger Datenpunkte entstanden. Ob eine unterschiedliche Regressionssteigung in den Klassen tatsächlich besser zu den Daten passt, wird im Folgenden untersucht. Diese Hinzunahme erhöht die zu bestimmenden Parameter auf fünf (Tab. 37):

Tab. 37: Struktur des Modells 1b. Der neue Faktor im gemischten Modell wird fett dargestellt.

H. Modell	Ebene 1	$Post_{ij} = \beta_{0j} + \beta_{1j}Pre + e_{ij}$
	Ebene 2	$\beta_{0j} = \gamma_{00} + r_{0j}$ $\beta_{1j} = \gamma_{10} + r_{1j}$
Gemischtes Modell	$Post_{ij} = \gamma_{00} + \gamma_{10}Pre + r_{0j} + \mathbf{r_{1j}Pre} + e_{ij}$	
Modellgüte	$BIC = 1337,01$	

Das BIC liefert einen schlechteren Wert als im Modell zuvor. Zudem kann das Modell nun mithilfe eines LR-Tests mit dem *Random-Intercept*-Modell verglichen werden. Dies testet die Signifikanz der Level-2-Varianz des *Slopes* und die der Level-2-Kovarianz zwischen *Intercept* und *Slope*. Bei Signifikanz dieser beiden Größen muss ein weiterer Modellvergleich herangezogen werden, um zu überprüfen, welche der beiden Größen signifikant sind. In diesem Fall unterscheidet sich die Steigung trotz des optischen Eindrucks nicht signifikant zwischen den Gruppen, womit das Modell bei einem *Random-Intercept*-Modell für den Faktor Prätestergebnis belassen wird. Die Regressionsgeraden pro Klasse sehen dann wie folgt aus (Abb. 76).

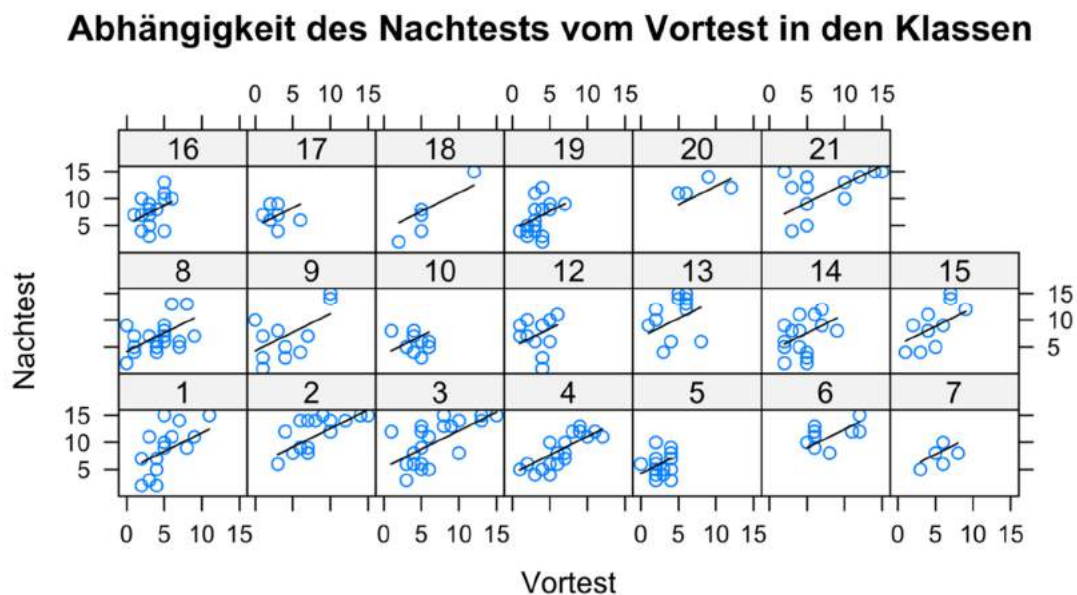


Abb. 76: Nachtestergebnis in Abhängigkeit des Vortestergebnisses in allen Klassen mit Schätzung des linearen hierarchischen Modells.

Zudem kann die Schätzung des Nachtestergebnisses des Modells anhand des Vortestergebnisses mit der Regression zu Beginn verglichen werden (Abb. 77).

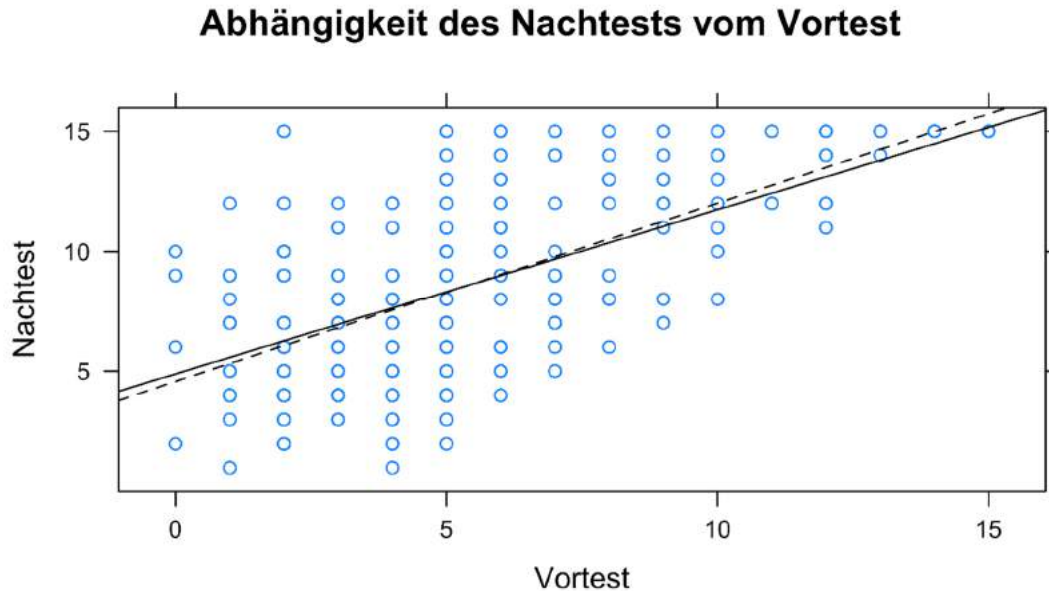


Abb. 77: Nachtestergebnisse in Abhängigkeit der Vortestergebnisse mit Regressionsgerade (gestrichelt) und hierarchischem linearen Modell (durchgezogen).

Unter Berücksichtigung der Mehrebenenstruktur trifft das Modell eine leicht andere Vorhersage bzgl. des Zusammenhangs zwischen Vor- und Nachtest, wenngleich der Unterschied nicht groß ist. Das Mehrebenenmodell schätzt einen geringeren Einfluss des Vortests auf das Nachtestergebnis als die einfache lineare Regression.

M2a: Der Haupteffekt „Gruppe“

Nachdem der Haupteffekt „Vortest“ berücksichtigt wurde, kann auch im Mehrebenenmodell der Frage nachgegangen werden, ob die Gruppenzugehörigkeit einen Einfluss auf das Nachtestergebnis hat. Da in den vorangegangenen Kapiteln dieser Effekt nicht nachweisbar war, lässt sich auch hier davon ausgehen, dass die Gruppenzugehörigkeit keinen signifikanten Einfluss auf das Nachtestergebnis hat. Durch die Berücksichtigung der hierarchischen Struktur sollte dies aber erneut überprüft werden. Die Modellgleichungen werden also um den Faktor „Gruppe“ ergänzt, der auf Ebene 2 in die Gleichungen eingeht (Tab. 38).

Tab. 38: Struktur des Modells 2a. Der neue Faktor im gemischten Modell wird fett dargestellt.

H. Modell	Ebene 1	$Post_{ij} = \beta_{0j} + \beta_{1j}Pre + e_{ij}$
	Ebene 2	$\beta_{0j} = \gamma_{00} + \gamma_{01}Gruppe + r_{0j}$ $\beta_{1j} = \gamma_{10}$
Gemischtes Modell	$Post_{ij} = \gamma_{00} + \gamma_{01}Gruppe + \gamma_{10}Pre + r_{0j} + e_{ij}$	
Modellgüte	$BIC = 1331,32$	

Die Modellgüte (BIC) ist schlechter als im Modell zuvor. Die nicht signifikant bessere Passung des Modells bestätigt sich durch einen direkten Modellvergleich (Tab. 39).

Tab. 39: Vergleich des Modells 2a mit dem Modell 1a.

Modell	Parameter	AIC	BIC	Loglik	Devianz	Chisq	df	p
Modell1a	4	1308	1322	-650	1300			
Modell2a	5	1310	1328	-650	1300	0,01	1	0,91

Das Modell 2a ist auf Basis der Daten abzulehnen. Das zeigt sich auch darin, dass der Faktor „Gruppe“ nicht signifikant wird, wenn man sich die festen Effekte des Modells anschaut (Tab. 40).

Tab. 40: Kennwerte der im Modell 2a enthaltenen Parameter.

						95%-Konfidenzintervall	
Parame-ter	Schät-zung	SE	df	t-Wert	Sig.	Untergr.	Obergr.
γ_{00}	4,92	0,52	37,24	9,40	$p < 0,001$	3,91	5,93
γ_{10}	0,68	0,06	244,65	11,12	$p < 0,001$	0,57	0,81
γ_{01} (VA)	-0,06	0,60	15,94	-0,09	$p = 0,93$		

Das Mehrebenenmodell bestätigt die vorangegangenen Analysen und zeigt erneut, dass es keinen signifikanten Unterschied zwischen den beiden Interventionen in Bezug auf das Konzeptverständnis der Newtonschen Dynamik gibt.

M2b: Interaktion zwischen „Gruppe“ und „Vortest“

Es ist denkbar, dass es eine signifikante Interaktion zwischen der Gruppe und dem Vortestfaktor gibt, ohne dass der Faktor „Gruppe“ selbst signifikant ist. Dies würde bedeuten, dass sich insgesamt zwar keine Unterschiede zwischen den Gruppen zeigen, diese aber je nach Vortestergebnis unterschiedliche Nachtestergebnisse erreichen. Das wäre der Fall, wenn in einer Gruppe vor allem diejenigen Proband*innen mehr dazu lernen, die im Vortest schlecht abgeschnitten haben, während in der anderen Gruppe vor allem diejenigen viel dazu lernen, die im Vortest gut abgeschnitten haben. Auch wenn dies aus der Theorie nicht zu erwarten ist, kann es im Mehrebenenmodell überprüft werden. Ein direkter Modellvergleich zeigt, dass das Modell 2b zwar minimal besser zu den Daten passt als das Modell 2a, dieses vor allem durch die zusätzlichen zwei Faktoren im Vergleich zu Modell 1a aber insgesamt abgelehnt wird (Tab. 41).

Tab. 41: Vergleich des Modells 2b mit den Modellen 1a und 1b.

Modell	Parameter	AIC	BIC	Loglik	Devianz	Chisq	df	p
Modell1a	4	1308	1322	-650	1300			
Modell2a	5	1310	1328	-650	1300	0,01	1	0,91
Modell2b	6	1309	1330	-648	1297	2,88	1	0,09

Es gibt keine signifikante Interaktion zwischen dem Gruppenfaktor und dem Vortestergebnis. Damit wird der Gruppenfaktor aus dem HLM genommen. Es ist aber weiterhin von Interesse, welche zusätzlichen Kovariaten einen Einfluss auf das Nachtestergebnis haben.

M3a: Der Haupteffekt „Theoretisieren“

Eine weitere Forschungsfrage war, ob das Interesse am Theoretisieren bzw. Experimentieren den Erfolg der Intervention beeinflusst und ob dies abhängig von der Gruppenzugehörigkeit ist. Wie in Kapitel 10.4 näher beschrieben wird, scheint vor allem der Faktor „Theoretisieren“ einen Einfluss auf das Nachtestergebnis zu haben. Das Modell wird um den Level-1-Prädiktor „Theorie“ erweitert (Tab. 42).

Tab. 42: Struktur des Modells 3a. Der neue Faktor im gemischten Modell wird fett dargestellt.

H. Modell	Ebene 1	$Post_{ij} = \beta_{0j} + \beta_{1j}Pre + \beta_{2j}Theorie + e_{ij}$
	Ebene 2	$\beta_{0j} = \gamma_{00} + r_{0j}$ $\beta_{1j} = \gamma_{10}$ $\beta_{2j} = \gamma_{20}$
Gemischtes Modell	$Post_{ij} = \gamma_{00} + \gamma_{10}Pre + \gamma_{20}Theorie + r_{0j} + e_{ij}$	
Modellgüte	$BIC = 1305,21$	

Die Modellgüte verbessert sich damit im Vergleich zu Modell 1a. Ein direkter Modellvergleich zeigt, dass Modell 3a besser zu den Daten passt (Tab. 43):

Tab. 43: Vergleich des Modells 3a mit dem Modell 1a.

Modell	Parameter	AIC	BIC	Loglik	Devianz	Chisq	df	p
Modell1a	4	1308	1322	-650	1300			
Modell3a	5	1281	1299	-636	1271	28,4	1	<0,001

Das Interesse an theoretischen Zusammenhängen beeinflusst also auch bei Berücksichtigung des Vortestergebnisses das Nachtestergebnis signifikant (Tab. 44).

Tab. 44: Kennwerte der im Modell 3a enthaltenen Parameter.

						95%-Konfidenzintervall	
Parameter	Schätzung	SE	df	t-Wert	Sig.	Untergr.	Obergr.
γ_{00}	1,93	0,68	162,73	2,78	$p < 0,01$	0,60	3,27
γ_{10}	0,60	0,06	256,98	10,13	$p < 0,001$	0,49	0,72
γ_{20}	1,12	0,20	258,60	6,96	$p < 0,001$	0,72	1,52

Deutlich wird, dass der Parameter γ_{20} einen signifikanten Einfluss auf das Nachtestergebnis hat. Die Zufallseffekte in Modell 3a sind in Tab. 45 dargestellt.

Tab. 45: Darstellung der im Modell 3a enthaltenen Zufallseffekte.

Ebene	Parameter	Varianz	Std.-Abweichung	Ebenenspez. R^2
Schüler	e_{ij}	6,87	2,62	0,103
Klasse	r_{0j}	1,03	1,02	-0,010

Das Interesse an theoretischen Zusammenhängen hat einen hoch signifikanten Einfluss auf das Nachtestergebnis. Das Modell 3a schätzt, dass ein zusätzlicher Punkt in dieser Skala das Nachtestergebnis um 1,12 Punkte erhöht. Dieser Effekt ist dennoch geringer als der des Vortestergebnisses, da sich die Skalen unterscheiden. Ein standardisierter Vergleich aller relevanten Prädiktoren mit Effektstärken wird am Ende des Kapitels vorgenommen.

Neben dem Interesse an theoretischen Zusammenhängen wurde außerdem das Interesse am Experimentieren, das Fachinteresse und das Interesse an Computern erhoben. Diese drei haben keinen signifikanten Einfluss auf das Nachtestergebnis und verbessern das Modell damit nicht. Außerdem wird das Modell durch die Hinzunahme eines *Random Slope* ebenso wenig verbessert wie durch die Hinzunahme des Interaktionsfaktors zwischen dem Faktor „Theorie“ und dem Vortestergebnis.

M4a: Der Haupteffekt „Physiknote“

Tab. 46: Struktur des Modells 4a. Der neue Faktor im gemischten Modell wird fett dargestellt.

H. Modell	Ebene 1	$Post_{ij} = \beta_{0j} + \beta_{1j}Pre + \beta_{2j}Theorie + \beta_{3j}Physiknote + e_{ij}$
	Ebene 2	$\beta_{0j} = \gamma_{00} + r_{0j}$ $\beta_{1j} = \gamma_{10}$ $\beta_{2j} = \gamma_{20}$ $\beta_{3j} = \gamma_{30}$
Gemischtes Modell	$Post_{ij} = \gamma_{00} + \gamma_{10}Pre + \gamma_{20}Theorie + \gamma_{30}Physiknote + r_{0j} + e_{ij}$	
Modellgüte	$BIC = 1302,89$	

Als Kovariate wurden außerdem die Schulnoten in Mathematik und Physik erhoben, welche einen Einfluss auf das Ergebnis in Fachwissenstests haben könnten. Zunächst wird der Einfluss der Physiknote untersucht (Tab. 46).

Die Modellgüte verbessert sich damit im Vergleich zu Modell 3a. Ein direkter Modellvergleich zeigt, dass Modell 4a besser zu den Daten passt (Tab. 47):

Tab. 47: Vergleich des Modells 4a mit dem Modell 3a.

Modell	Parameter	AIC	BIC	Loglik	Devianz	Chisq	df	p
Modell3a	5	1281	1299	-636	1271			
Modell4a	6	1271	1293	-630	259	12	1	<0,001

Die Physiknote beeinflusst also auch bei Berücksichtigung des Vortestergebnisses das Nachtestergebnis signifikant (Tab. 48).

Tab. 48: Kennwerte der im Modell 4a enthaltenen Parameter.

						95%-Konfidenzintervall	
Parameter	Schätzung	SE	df	t-Wert	Sig.	Untergr.	Obergr.
γ_{00}	0,92	0,72	177,21	1,28	$p = 0,20$		
γ_{10}	0,56	0,06	256,06	9,28	$p < 0,001$	0,44	0,68
γ_{20}	0,91	0,21	258,05	4,32	$p < 0,001$	0,49	1,32
γ_{30}	0,20	0,06	258,96	3,47	$p < 0,001$	0,09	0,32

Eine mögliche Begründung für diesen Einfluss wird am Ende des Kapitels diskutiert. Die Zufallseffekte sind im Modell 4a folgende (Tab. 49).

Tab. 49: Darstellung der im Modell 4a enthaltenen Zufallseffekte.

Ebene	Parameter	Varianz	Std.-Abweichung	Ebenenspez. R^2
Schüler	e_{ij}	6,65	2,58	0,032
Klasse	r_{0j}	0,84	0,92	0,184

Die Physiknote hat einen signifikanten Einfluss auf das Nachtestergebnis und wird im Modell belassen. Das Modell schätzt, dass ein Punkt mehr in der Physiknote zu 0,20 Punkten mehr im Nachtestergebnis führt. Die Mathematiknote hat hingegen keinen signifikanten Einfluss auf das Nachtestergebnis.

M5a: Der Haupteffekt „Germane Cognitive Load“

Während in Kapitel 10.3 untersucht wird, inwiefern sich die beiden Interventionen im erlebten Cognitive Load unterscheiden, ist hier von Interesse, wie sich dieser Cognitive Load unabhängig von der Gruppenzugehörigkeit auf das Nachtestergebnis auf Schüler-ebene auswirkt. Dazu wird im nächsten Schritt der Faktor „Germane Cognitive Load“ in das Modell eingefügt, da dies den lernrelevanten bzw. lernförderlichen Cognitive Load darstellt (Tab. 50).

Tab. 50: Struktur des Modells 5a. Der neue Faktor im gemischten Modell wird **fett** dargestellt.

H. Modell	Ebene 1	$Post_{ij} = \beta_{0j} + \beta_{1j}Pre + \beta_{2j}Theorie$ $+ \beta_{3j}Physiknote + \beta_{4j}Germane$ $+ e_{ij}$
	Ebene 2	$\beta_{0j} = \gamma_{00} + r_{0j}$ $\beta_{1j} = \gamma_{10}$ $\beta_{2j} = \gamma_{20}$ $\beta_{3j} = \gamma_{30}$ $\beta_{4j} = \gamma_{40}$
Gemischtes Modell	$Post_{ij} = \gamma_{00} + \gamma_{10}Pre + \gamma_{20}Theorie + \gamma_{30}Physiknote$ $+ \gamma_{40}\textbf{Germane} + r_{0j} + e_{ij}$	
Modellgüte	$BIC = 1310,08$	

Die Modellgüte verbessert sich damit im Vergleich zu Modell 4a nicht, was auch ein direkter Vergleich der beiden Modelle zeigt (Tab. 51).

Tab. 51: Vergleich des Modells 5a mit dem Modell 4a.

Modell	Parameter	AIC	BIC	Loglik	Devianz	Chisq	df	p
Modell4a	6	1271	1293	-630	1259			
Modell5a	7	1273	1298	-630	1259	0,14	1	0,71

Der Faktor „Germane Cognitive Load“ hat keinen signifikanten Einfluss auf das Nachtertergebnis (Tab. 52).

Tab. 52: Kennwerte der im Modell 5a enthaltenen Parameter.

						95%-Konfidenzintervall	
Parameter	Schätzung	SE	df	t-Wert	Sig.	Untergr.	Obergr.
γ_{00}	0,66	0,99	228,30	0,67	$p = 0,50$		
γ_{10}	0,56	0,06	255,14	9,23	$p < 0,001$	0,44	0,68
γ_{20}	0,89	0,21	256,53	4,21	$p < 0,001$	0,48	1,31
γ_{30}	0,20	0,06	257,97	3,46	$p < 0,001$	0,09	0,32
γ_{40}	0,06	0,17	255,35	0,39	$p = 0,70$		

Die Zufallseffekte sind im Modell 5a folgende (Tab. 53):

Tab. 53: Darstellung der im Modell 5a enthaltenen Zufallseffekte.

Ebene	Parameter	Varianz	Std.-Abweichung	Ebenenspez. R^2
Schüler	e_{ij}	6,67	2,58	-0,003
Klasse	r_{0j}	0,85	0,92	-0,012

Der Germane Cognitive Load hat scheinbar keinen Einfluss auf das Nachtertergebnis. Nach manchen Autor*innen in der CLT sollte der Germane Cognitive Load als lernrelevanter Cognitive Load einen positiven Einfluss auf den Lernzuwachs haben (Sweller et al., 1998). Es gibt aber auch Autor*innen, die argumentieren, dass der Germane Cognitive Load nicht sinnvoll von den anderen Arten unterschieden werden kann und die CLT auch

ohne diese Art des Cognitive Load auskommt (Kalyuga, 2011). Eine Unterscheidung zwischen Intrinsic und Extranous Cognitive Load und deren Benutzung wäre demnach ausreichend.

M6a: Der Haupteffekt „Intrinsic und Extranous Cognitive Load“

Die beiden weiteren Arten des Cognitive Load sind der Extranous und der Intrinsic Cognitive Load. Es zeigt sich, dass das Modell mit der größten Varianzaufklärung die Summe der Scores im Extranous und Intrinsic Cognitive Load beinhaltet und nicht die Scores als einzelne Faktoren. Auch nach den Autor*innen der Theorie sind die beiden kognitiven Belastungen additiv (Sweller et al., 2011). Zudem ist bei ähnlicher Varianzaufklärung das Modell mit weniger Faktoren vorzuziehen. Es ergibt sich dadurch folgendes Modell (Tab. 54):

Tab. 54: Struktur des Modells 6a. Der neue Faktor im gemischten Modell wird fett dargestellt.

H. Modell	Ebene 1	$Post_{ij} = \beta_{0j} + \beta_{1j}Pre + \beta_{2j}Theorie + \beta_{3j}Physiknote + \beta_{4j}Load + e_{ij}$
	Ebene 2	$\beta_{0j} = \gamma_{00} + r_{0j}$ $\beta_{1j} = \gamma_{10}$ $\beta_{2j} = \gamma_{20}$ $\beta_{3j} = \gamma_{30}$ $\beta_{4j} = \gamma_{40}$
Gemischtes Modell	$Post_{ij} = \gamma_{00} + \gamma_{10}Pre + \gamma_{20}Theorie + \gamma_{30}Physiknote + \gamma_{40}Load + r_{0j} + e_{ij}$	
Modellgüte	$BIC = 1295,95$	

Die Modellgüte verbessert sich damit im Vergleich zu dem bisher am besten passenden Modell 4a, was auch ein direkter Vergleich der beiden Modelle zeigt (Tab. 55).

Tab. 55: Vergleich des Modells 6a mit dem Modell 4a.

Modell	Parameter	AIC	BIC	Loglik	Devianz	Chisq	df	p
Modell4a	6	1271	1293	-630	1259			
Modell6a	7	1259	1284	-623	1245	14,1	1	<0,001

Der Cognitive Load beeinflusst also auch bei Berücksichtigung des Vortestergebnisses das Nachtestergebnis signifikant (Tab. 56):

Tab. 56: Kennwerte der im Modell 6a enthaltenen Parameter.

						95%-Konfidenzintervall	
Parameter	Schätzung	SE	df	t-Wert	Sig.	Untergr.	Obergr.
γ_{00}	3,54	0,99	245,66	3,58	$p < 0,001$	1,62	5,46
γ_{10}	0,53	0,06	255,90	8,97	$p < 0,001$	0,42	0,65
γ_{20}	0,77	0,21	257,47	3,68	$p < 0,001$	0,36	1,17
γ_{30}	0,19	0,06	257,94	3,35	$p < 0,001$	0,08	0,30
γ_{40}	-0,77	0,20	258,00	-3,77	$p < 0,001$	-1,16	-0,37

Der Einfluss des Faktors „Load“ (γ_{40}) ist negativ. Das ist zu erwarten, da eine höhere kognitive Belastung mit einem geringeren Lernzuwachs assoziiert wird. Die Zufallseffekte von Modell 6a sind in Tab. 57 zu sehen.

Tab. 57: Darstellung der im Modell 6a enthaltenen Zufallseffekte.

Ebene	Parameter	Varianz	Std.-Abweichung	Ebenenspez. R^2
Schüler	e_{ij}	6,35	2,52	0,045
Klasse	r_{0j}	0,74	0,86	0,119

Die Summe aus Intrinsic und Extranous Load hat einen signifikanten Einfluss auf das Nachtestergebnis – auch unter Berücksichtigung der anderen im Modell vorhandenen

Variablen. Ein Punkt in dieser Variablen mehr führt statistisch zu einer Reduktion von 0,77 Punkten im Nachtest. Die Variable wird dementsprechend im Modell belassen.

Auswertung des finalen Modells

Alle weiteren in Tab. 28 dargestellten Prädiktoren hatten keinen signifikanten Einfluss auf das Nachtestergebnis (weder im Achsenabschnitt noch in der Steigung der Regression) und haben das Modell nicht ausreichend verbessert, um eine Hinzunahme zu rechtfertigen. Es konnten außerdem keine signifikanten Wechselwirkungen zwischen den Größen gefunden werden. Der Ablauf der Überprüfung dieser Kovariaten erfolgt auf die gleiche Weise, wie in den vorigen Abschnitten dargestellt, und wird daher nicht im Detail berichtet.

Das finale Modell ist also ein *Random-Intercept-Modell* ohne Prädiktor auf Level 2. Das heißt, dass keine der Variablen auf Klassenebene einen signifikanten Einfluss auf das Nachtestergebnis hatte. Prädiktoren für das Nachtestergebnis sind nach dem Modell das Vortestergebnis, das Interesse an theoretischen Zusammenhängen in der Physik, die Physiknote und der lernhinderliche Cognitive Load (die Summe aus Intrinsic und Extranous Cognitive Load).

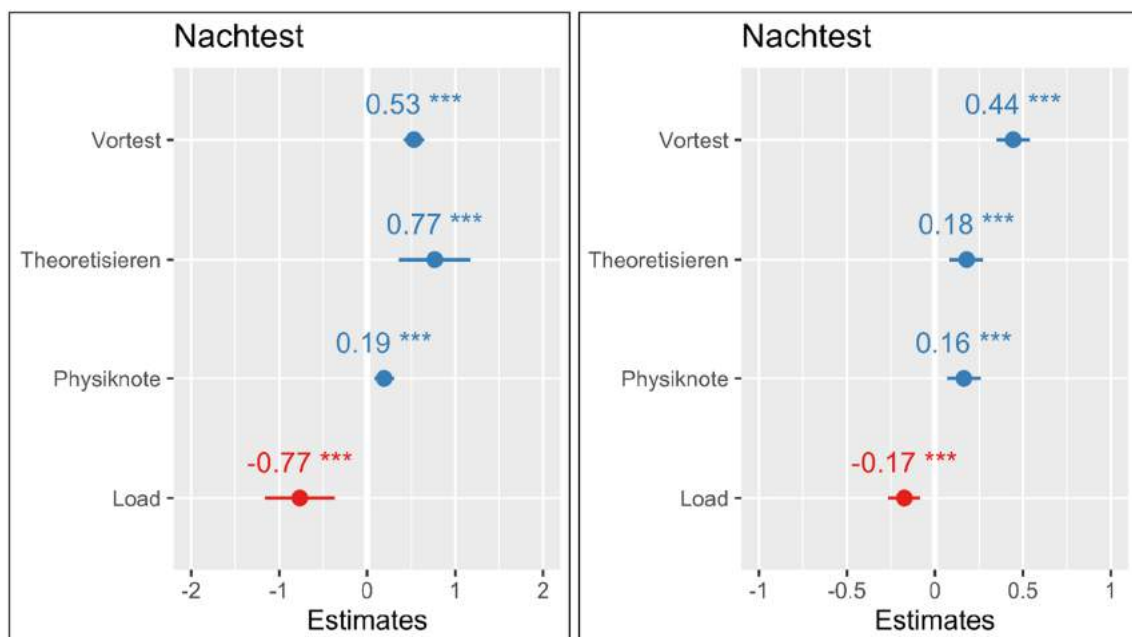


Abb. 78: Feste Effekte des finalen Modells. Dargestellt wird der Einfluss der Prädiktoren auf das Nachtestergebnis. Links in Rohwerten, rechts standardisiert. *** $p < 0,001$ ** $p < 0,01$

Abb. 78 zeigt die Koeffizienten der festen Effekte. In der linken Grafik sind die Koeffizienten nicht standardisiert. Ein Punkt mehr in der Skala „Theoretisieren“ führt zu 0,77 Punkten mehr im Nachtest. Dies bedeutet aufgrund der unterschiedlichen Skalenbereiche (Vortest [0;15], Theoretisieren [1;5]) aber nicht, dass dieser Effekt größer ist als der vom Vortest. In der rechten Grafik sind die Koeffizienten daher standardisiert, was die Vergleichbarkeit möglich macht. Den größten Einfluss hat das Vortestergebnis gefolgt vom Interesse an theoretischen Zusammenhängen („Theoretisieren“) und dem Cognitive Load. Dieser beeinflusst das Nachtestergebnis negativ, was Sinn ergibt, da es sich hier um die Summe aus Extranous und Intrinsic Cognitive Load handelt, die in negativer Weise die kognitive Belastung erhöht. Den im Modell geringsten Einfluss hat die Physiknote.

Übersicht über die Hauptmodelle

Tab. 58 stellt eine Übersicht über alle relevanten Hauptmodelle dar (Vgl. Kuhn, 2014). Dabei wird das f^2 (lokale Effektstärke) folgendermaßen bestimmt (Aiken & West, 2010):

$$f^2 = \frac{R_2^2 - R_1^2}{1 - R_2^2}$$

R_2^2 ist die Varianz, die von dem Modell mit dem jeweiligen Effekt erklärt wird, während R_1^2 die Varianzerklärung des Modells ohne diesen Faktor darstellt. Das Vortestergebnis hat demnach einen großen, die anderen Variablen nur einen kleinen Einfluss auf das Nachtestergebnis (Cohen, 1988). d_{Cohen} kann bei einem solchen linearen hierarchischen Modell folgendermaßen berechnet werden (Burde, 2018; Kuhn, 2010; Tymms et al., 1997; Tymms, 2004):

$$d_{Cohen} = \frac{2 \cdot \beta_N \cdot \sigma_P}{\sigma_{E1}}$$

Dabei ist β_N der Regressionskoeffizient des Prädiktors des Haupteffekts auf Ebene N, σ_P die Standardabweichung des Prädiktors und σ_{E1} die Standardabweichung der Einheiten auf Ebene 1.

Tab. 58: Übersicht über die Hauptmodelle.

	M0	M1a	M3a	M4a	M_final	Effektstärke	
Haupteffekte						d_{Cohen}	f^2
Intercept	8,42***	4,88***	1,93**	0,92	3,54**		
Vortest	-	0,69***	0,60***	0,56***	0,53***	1,31	0,41
Theoretisieren	-	-	1,12***	0,91***	0,77***	0,54	0,06
Phy-siknote	-	-	-	0,20***	0,19***	0,48	0,06
ICL & ECL	-	-	-	-	-0,77***	-0,52	0,08
Restvarianz							
Auf Ebene 1	10,60	7,67	6,87	6,65	6,35		
Auf Ebene 2	3,58	1,02	1,03	0,84	0,74		
Modellgüte							
BIC	1415,5	1326,5	1305,1	1302,9	1295,9		
ΔBIC	-	-88,9***	-21,3***	-2,3***	-6,9***		
R^2 (fixed effects)	0	0,35	0,42	0,45	0,49		
R^2 (total)	0,25	0,42	0,50	0,51	0,54		

In der Tabelle dargestellt sind außerdem die eingefügten Haupteffekte, ihr Einfluss auf das Nachtestergebnis, die jeweils existierende Restvarianz auf beiden Ebenen und verschiedene Parameter für die Modellgüte. Bei der Entwicklung dieses Modells wurde jeweils dargestellt, wieviel zusätzliche Varianz der hinzugefügte Haupteffekt im Vergleich zum vorigen Modell erklären kann. Um die jeweils erklärte Varianz der Parameter im

finalen Modell zu bestimmen, muss allerdings für jeden Haupteffekt ein Modell ohne diesen Effekt erstellt werden und die erklärte Varianz durch die Differenz der beiden Modelle bestimmt werden (Tab. 59).

Tab. 59: Erklärte Varianz der einzelnen Prädiktoren.

Haupteffekte	Erklärte Varianz		
	Ebene 1 (Schüler)	Ebene 2 (Klassen)	Gesamt
Vortest	17,0 %	17,0 %	17,0 %
Theoretisieren	3,3 %	-2,2 %	1,9 %
Physiknote	2,0 %	3,9 %	2,5 %
ECL & ICL	2,8 %	2,8 %	2,8 %

Wie zu erwarten war, ist die erklärte Varianz des Vortests am größten. Das Vortestergebnis vermag jeweils 17 % der Varianz auf beiden Ebenen zu erklären. Der Faktor „Theoretisieren“ hat auf Klassenebene sogar eine negative Varianzaufklärung, was bedeutet, dass das Modell ohne diesen Faktor auf dieser Ebene mehr Varianz aufklärt. Durch die Varianzaufklärung auf der Ebene der Schüler *innen ist die gesamte Varianzaufklärung dennoch positiv.

Diskussion

Das Geschlecht hat nach dem hier berichteten Modell keinen Einfluss auf das Nachtestergebnis. Dies ist zunächst ein Widerspruch zur gerechneten ANCOVA in Kapitel 9.3.6, die einen Einfluss des Geschlechts auf das Nachtestergebnis gezeigt hat. Dieser Einfluss ist auch im Mehrebenenmodell zu sehen, wenn neben dem Vortestergebnis und dem Geschlecht keine weiteren Prädiktoren hinzugenommen werden. Durch die Hinzunahme des Interesses an theoretischen Zusammenhängen und der Physiknote verschwindet der Einfluss des Geschlechts. Das bedeutet, dass sich die Geschlechter in diesen Variablen unterscheiden ($t = 4,61$; $p < 0,001$ für das Interesse an theoretischen Zusammenhängen und $t = 2,81$; $p < 0,01$ für die Physiknote), die genannten Variablen aber eine höhere Varianzaufklärung bieten als das Geschlecht selbst. Die Ursache, dass Jungen auch unter Kontrolle des Vortestergebnisses ein höheres Nachtestergebnis erzielen, liegt also eher in affektiven Variablen als in der Intervention selbst begründet. Nichtsdestotrotz sollte diese Geschlechterungleichheit beachtet werden.

Das Interesse am Computer hat keinen Einfluss auf das Nachtestergebnis. Da eine solche Intervention mit Computereinsatz nicht diejenigen benachteiligen sollte, die privat kein Interesse an der Nutzung von Computern haben, ist dies positiv zu bewerten. Das fachspezifische Selbstkonzept hat keinen signifikanten Einfluss auf das Nachtestergebnis. Das ist aber nur der Fall, wenn die Physiknote inkludiert ist, welche mehr Varianz aufklärt. Das fachspezifische Selbstkonzept scheint also eng an die Physiknote gekoppelt zu sein, wobei hier von einer wechselseitigen Beeinflussung ausgegangen werden kann. In der Tat gibt es hier eine mittlere bis starke Korrelation ($r = 0,49$; $p < 0,001$) zwischen den beiden Größen. Durch die höhere Varianzaufklärung wird aber die Physiknote im Modell belassen. Die Physiknote deckt bei den meisten teilnehmenden Klassen das Kinematik-Halbjahr ab. Sie könnte einen Einfluss auf das Nachtestergebnis haben, da gute Kenntnisse in der Kinematik (z. B. Diagrammkompetenz) hilfreich für die durchgeführte Intervention sind. Gut ist auch, dass die Mathematiknote keine Rolle spielt, da die Intervention bewusst so konzipiert war, dass die mathematische Komplexität trotz der Thematisierung von komplexen Bewegungen gering sein sollte, womit mathematische Fähigkeiten eine geringe Rolle für den Lernerfolg spielen sollten. Das wird durch die Ergebnisse dieses Modells bestätigt. Der zuvor bereits hypothetisierte negative Einfluss der intrinsischen und extrinsischen kognitiven Belastung auf das Nachtestergebnis wurde durch das Modell bestätigt.

9.3.9 Unterschiede in den einzelnen Inhaltsbereichen

Wie bereits diskutiert, wurde das Testinstrument entwickelt, um verschiedene Inhaltsbereiche testen und unterscheiden zu können (siehe Kapitel 8.4.2). Dabei ist vor allem bei den Skalen aus dem Bereich der Dynamik von Interesse, ob es beim Lernzuwachs zwischen den Gruppen Unterschiede gibt. Eine Hypothese war, dass die Gruppe der Modellbildung den Score in Items, in denen von den wirkenden Kräften auf die Bewegung geschlossen werden muss, stärker verbessert als die Gruppe der Videoanalyse, da dies die inhärente Argumentationsweise in der Modellbildung ist. Die Videoanalyse könnte wiederum einen Vorteil in Items haben, in denen von der Bewegung auf die Kraft geschlossen wird.

Von Kraft auf Bewegung schließen

Um zu testen, ob die Gruppe einen Einfluss auf das Ergebnis in der Inhaltsdimension „Von Kraft auf Bewegung schließen“ hat, wird eine ANCOVA gerechnet, in der unter Kontrolle des Vortestergebnisses der Einfluss der Gruppenzugehörigkeit untersucht wird.

Nach Überprüfung der Bedingungen der ANCOVA (siehe Kapitel 9.3.6) können die Nachtestergebnisse für beide Gruppen geschätzt werden (Abb. 79).

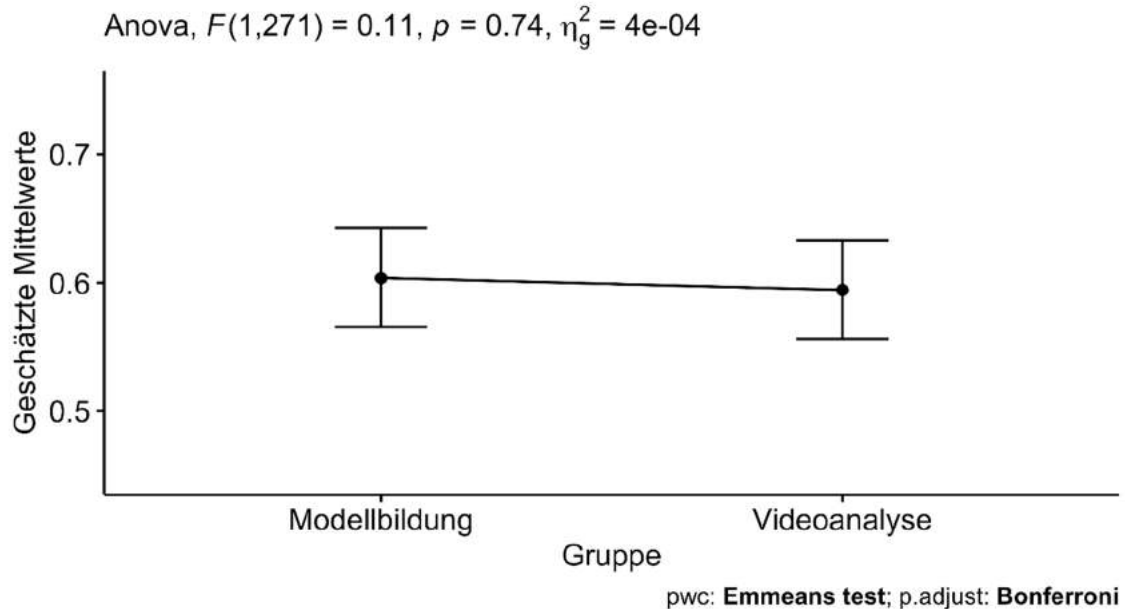


Abb. 79: Geschätzte Mittelwerte im Nachtest für beide Gruppen im Bereich "Von Kraft auf die Bewegung schließen" bei einem Vortestergebnis von 44,4 % mit 95%-Konfidenzintervallen.

Bei einem Vortestergebnis von 44,4 % in dieser Skala schätzt die ANCOVA das Nachtestergebnis auf 60,4 % [56,5 %; 64,3 %] in der Gruppe der Modellbildung und auf 59,5 % [55,6 %; 63,3 %] in der Gruppe der Videoanalyse. Der Einfluss der Gruppe ist nicht signifikant. Die Hypothese, dass die Gruppe der Modellbildung einen Vorteil darin hat, von den Kräften zur Bewegung zu argumentieren, muss verworfen werden.

Von Bewegung auf Kraft schließen

Auf die gleiche Weise kann der Einfluss der Gruppe auf das Ergebnis in der Skala „Von der Bewegung auf die Kräfte schließen“ untersucht werden, in der vermutet werden könnte, dass die Gruppe der Videoanalyse einen Vorteil hat (Abb. 80).

Die ANCOVA geht im Nachtest von 54,9 % [50,3 %; 59,4 %] für die Modellbildung und 56,7 % [52,1 %; 61,2 %] für die Videoanalyse bei einem Vortestergebnis von 28,5 % aus. Der Unterschied ist nicht signifikant. Auch hier wird die Hypothese, dass es einen Unterschied zwischen den Gruppen (mit Vorteil für die Videoanalyse) gibt, verworfen.

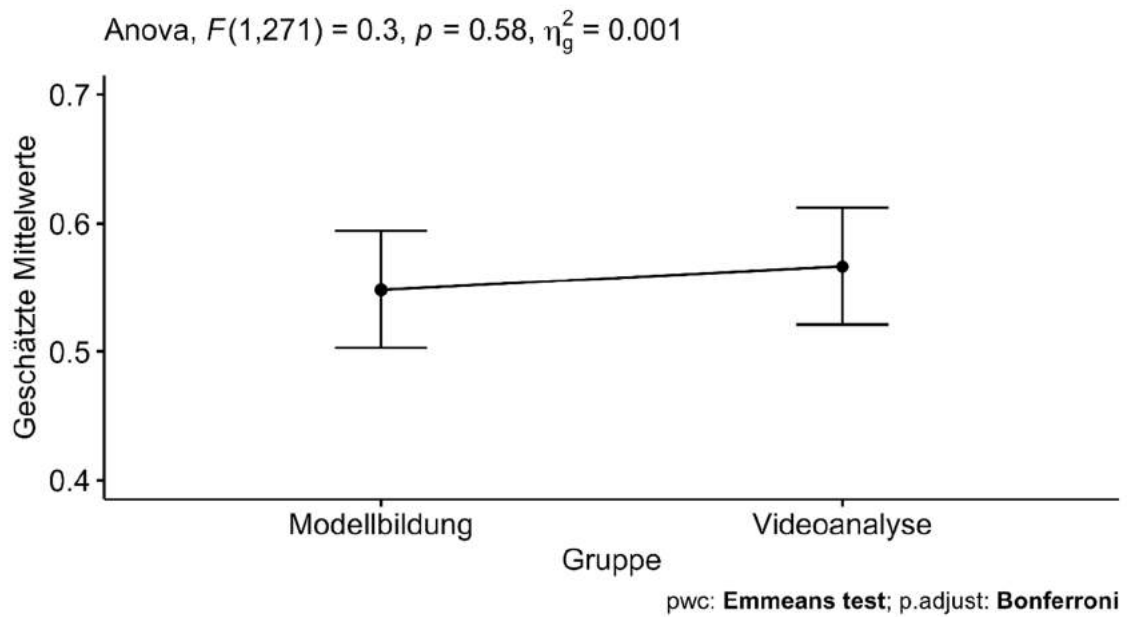


Abb. 80: Geschätzte Mittelwerte im Nachtest für beide Gruppen im Bereich "Von der Bewegung auf die Kräfte schließen" bei einem Vortestergebnis von 28,5 % mit 95%-Konfidenzintervallen.

Erstes Newtonsches Gesetz

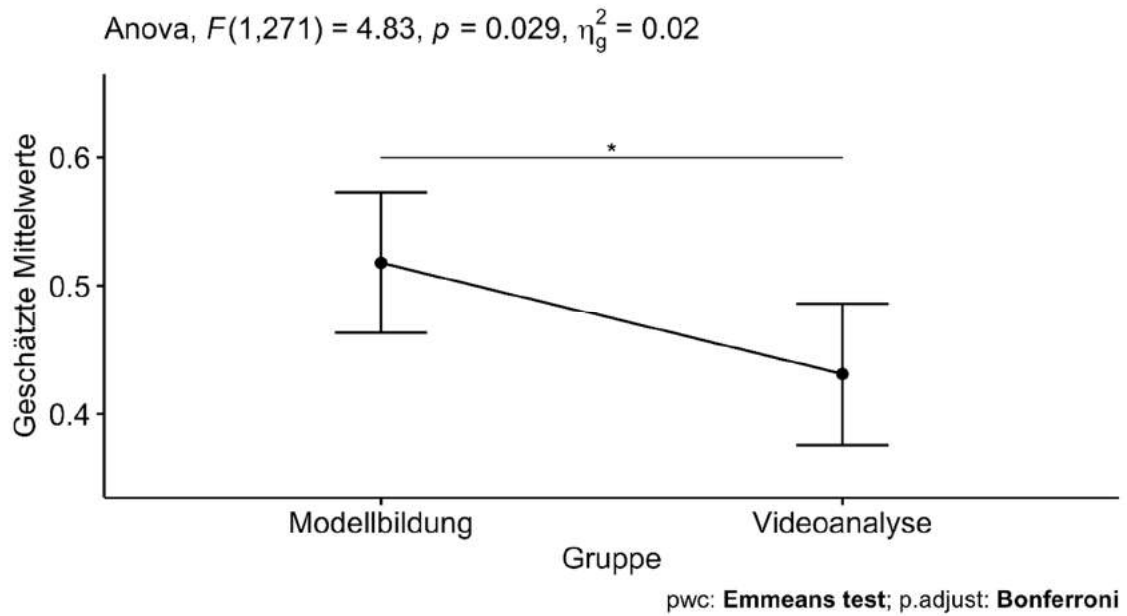


Abb. 81: Geschätzte Mittelwerte im Nachtest für beide Gruppen im Bereich "1. Newtonsches Gesetz" bei einem Vortestergebnis von 24,8 % mit 95%-Konfidenzintervallen.

Die dritte dynamische Inhaltsdimension im Test, die auch Teil der Intervention war, war das erste Newtonsche Gesetz (Abb. 81). In diesem Fall schätzt das Modell die Nachtestergebnisse auf 51,8 % [46,3 %; 57,3 %] in der Gruppe der Modellbildung und 43,1 % [37,6 %; 48,6 %] in der Gruppe der Videoanalyse bei einem Vortestergebnis von 24,8 %. Damit unterscheiden sich die Gruppen hier signifikant²⁷ mit kleiner Effektstärke. Dieses Ergebnis kann nicht mit der Argumentationsrichtung erklärt werden, da alle in dieser Skala enthaltenen Items von einer Situation mit konstanter Geschwindigkeit ausgehen und nach den wirkenden Kräften fragen. Diese Argumentationsrichtung tritt in der Gruppe der Videoanalyse häufiger auf. Möglicherweise steht das bessere Ergebnis von der Gruppe der Modellbildung mit der Reduktion von gewissen Schülervorstellungen in Zusammenhang die hier zu einer falschen Antwort führen würden. Nach Just et al. (2021) ist es für Schüler*innen speziell in Items zum ersten Newtonschen Gesetz schwieriger von der bekannten Geschwindigkeit auf die Kräfte als von den bekannten Kräften auf die Geschwindigkeit zu schließen. Dies begründen die Autor*innen damit, dass bei der erstgenannten Argumentationsrichtung häufig die Schülervorstellung aktiviert wird, dass eine Kraft in Bewegungsrichtung notwendig ist, damit sich ein Objekt bewegen kann. Wenn ein Objekt sich mit einer konstanten Geschwindigkeit bewegt, muss nach dieser Schülervorstellung also eine von Null verschiedene Gesamtkraft vorhanden sein, was zu einer falschen Antwort führt. In der anderen Richtung wird demnach eher die Schülervorstellung aktiviert, dass keine Kraft „keine Änderung“ bedeutet. Diese Assoziation kann zu der Antwort führen, dass sich das Objekt mit einer konstanten Geschwindigkeit bewegt, was in diesem Fall die richtige Antwort ist. Die Schülervorstellung ist also entscheidend für die Lösungswahrscheinlichkeit eines solchen Items. Ob der hier gezeigte Unterschied zwischen den Gruppen in der unterschiedlichen Aktivierung von Schülervorstellungen begründet liegt, wird in Kapitel 10.5 besprochen.

²⁷ Wenn mit der strengen Bonferroni-Korrektur gearbeitet wird, nach der durch die Anzahl aller in dieser Studie getesteten Hypothesen geteilt wird, wäre dieser Unterschied nicht mehr signifikant. Diese strenge Konvention wird auch in der Folge nicht direkt angewandt, sollte aber von den Lesenden berücksichtigt werden.

10 Ergebnisse zu weiteren quantitativen Tests

10.1 Modellverständnis

Eine weitere Forschungsfrage betrifft die Ansichten, die Schüler*innen über Modelle haben, deren Eigenschaften im Unterricht oft nicht thematisiert werden. Auch in der Intervention wurden keine Eigenschaften von Modellen explizit diskutiert, wobei sich die Gruppe der mathematischen Modellbildung direkt mit dem Modellieren und damit indirekt auch mit den Eigenschaften von Modellen beschäftigte. Dazu wurde ein Testinstrument entwickelt und pilotiert (siehe Kapitel 8.4), in dem verschiedene Facetten von Modellen abgefragt werden. Dieses Testinstrument sollte ursprünglich mehrdimensional sein, um eine Differenzierung nach unterschiedlichen Eigenschaften zu ermöglichen. Durch die Ergebnisse der Pilotierung wurde auf einen eindimensionalen Fragebogen umgestellt, indem ein hoher Score auf ein gewünschtes Modellverständnis hindeutet. Dabei muss darauf hingewiesen werden, dass das Testinstrument bewusst kurzgehalten wurde. Das bedeutet auch, dass damit keineswegs die Vielschichtigkeit der Bedeutung von Modellen für die Physik abgedeckt werden kann. Dennoch ist es dazu geeignet zu untersuchen, ob sich durch eine Intervention mit mathematischer Modellbildung eine Änderung der Einstellung zu Modellen ergibt (Abb. 82).

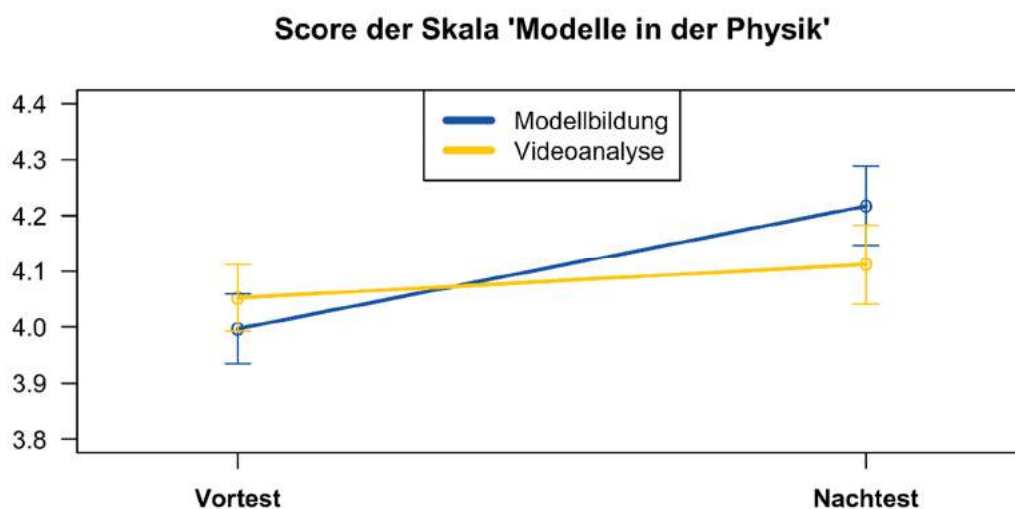


Abb. 82: Ergebnisse der Skala Modellverständnis (von max. 5) in beiden Gruppen in Vor- und Nachtest mit 95%-Konfidenzintervallen.

In der Gruppe der Modellbildung unterscheidet sich der Nachtestscore höchst signifikant ($p < 0,001$) vom Vortestscore mit einer Effektstärke von $d = 0,43$ (mittlerer Effekt),

während es in der Gruppe der Videoanalyse nicht zu einer signifikanten Änderung kommt. Auch eine ANCOVA kommt auf einen signifikanten Einfluss der Gruppenzugehörigkeit auf das Nachtestergebnis in der Skala der Modelle unter Kontrolle des Vortestergebnisses (Abb. 83).

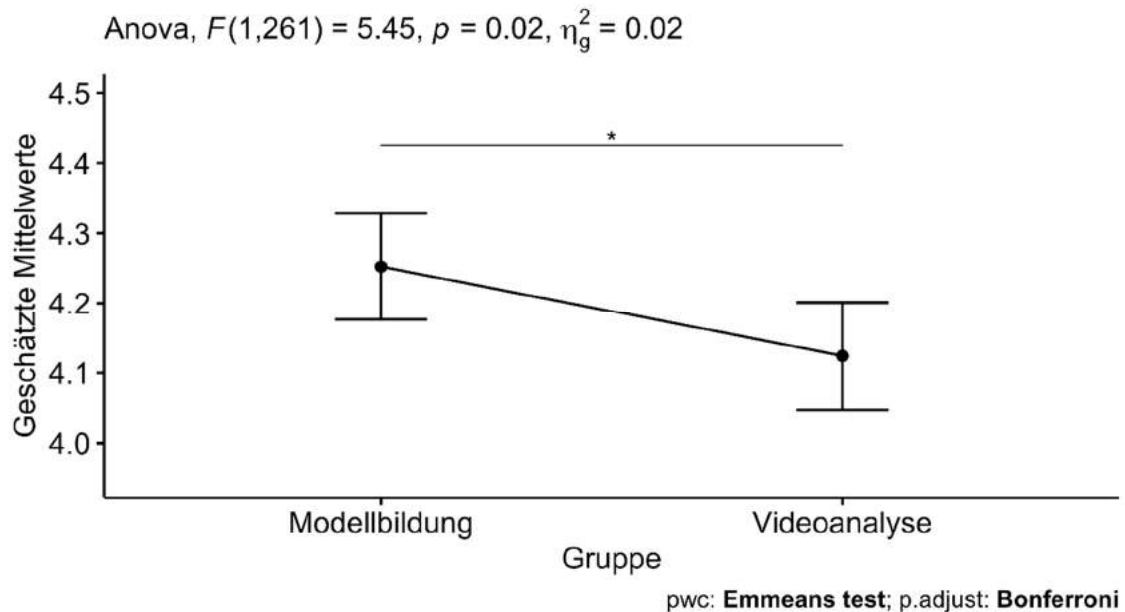


Abb. 83: Geschätzte Mittelwerte für den Nachtest in beiden Gruppen bei einem kontrollierten Vortestergebnis von 4,03 Punkten mit 95%-Konfidenzintervallen.

Auf Basis dieser Daten kann die Hypothese, dass es durch das Modellieren ohne explizite Besprechung von Modelleigenschaften zu einer positiven Entwicklung der Ansichten über Modelle kommen kann, bestätigt werden. Da in der Skala Items unterschiedlicher Art enthalten sind, lohnt sich eine Analyse auf Itemebene (Abb. 84). Es ist erkennbar, dass die Änderung in der Skala in der Gruppe der Modellbildung vor allem auf vier Items (2, 4, 5, 6) zurückzuführen ist, die auch die einzigen mit signifikantem Unterschied sind. Dabei ist die größte Änderung ($d = 0,36$ [0,21; 0,51]) dem Item „Modelle sind idealisierte Beschreibungen der Natur“ zuzuschreiben. Darauf folgt das Item „Modelle in der Physik dienen dazu, neue Sachverhalte zu erschließen“ ($d = 0,35$ [0,18; 0,53]). Die weiteren Items mit signifikantem Unterschied sind „Im naturwissenschaftlichen Erkenntnisprozess sind Modelle wichtig“ ($d = 0,30$ [0,14; 0,48]) und „Theoretische Überlegungen in der Physik dienen dazu, die Natur besser zu verstehen“ ($d = 0,23$ [0,06; 0,41]).

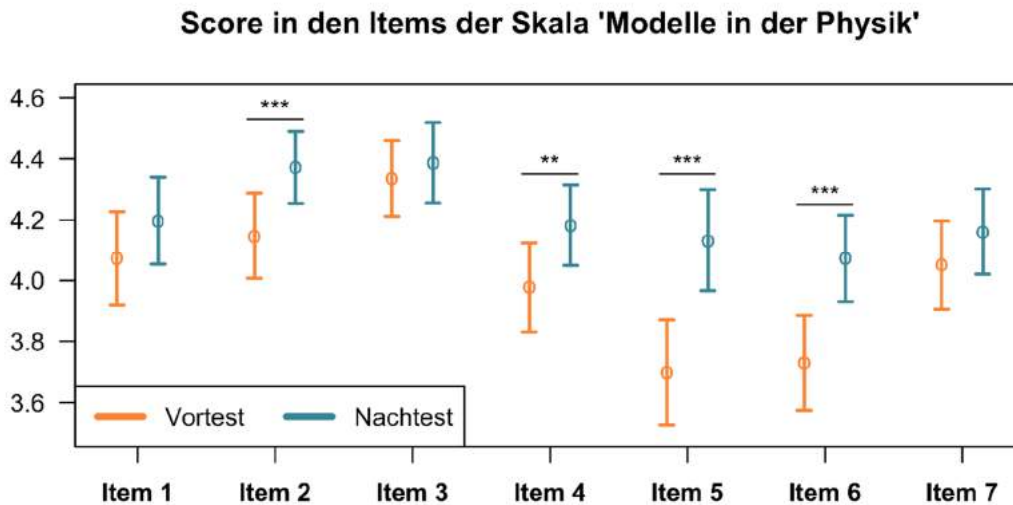


Abb. 84: Durchschnittlicher Score in den einzelnen Items in der Skala "Modelle in der Physik" in Vor- und Nachtest mit 95%-Konfidenzintervallen in der Gruppe der Modellbildung. *** $p < 0,001$, ** $p < 0,01$, sonst nicht signifikant.

Durch das Modellieren scheint eine erwünschte Änderung der Ansichten über Modelle eingetreten zu sein, ohne dass Modelleigenschaften explizit thematisiert wurden. Dies ist insbesondere erfreulich, da Modelle eine große Bedeutung für die Naturwissenschaften haben und dies im Unterricht unterrepräsentiert ist. Ein aktives Modellieren, welches primär fachliche Ziele verfolgt, könnte im Unterricht also auch zu einer positiven Veränderung der Ansichten über Modelle führen. Auch wenn das prozedurale Modellwissen nicht überprüft wurde, liegt der Schluss nahe, dass durch das Modellieren dieser Teil der Modellkompetenz verbessert wurde, was sich auch positiv auf das deklarative Modellwissen ausgewirkt hat (siehe Kapitel 5.1).

10.2 Einschätzung der Lernenden

Die Einschätzung der Proband*innen selbst über die durchgeführte Intervention ist ebenfalls relevant für die Anwendung der Software im Unterricht. Dazu waren im Nachtest Fragen zur Benutzung des Programms, zur Relevanz des behandelten Themas und zum Spaß an der Intervention enthalten (Abb. 85).

Dabei unterscheiden sich die Werte lediglich in der Skala „Software“ ($p < 0,001$) mit einer mittleren Effektstärke von $d = 0,54$ zugunsten der Videoanalyse. Das Programm „measure dynamics“ wurde von den Proband*innen signifikant besser bewertet als das Programm „Newton-II“, wobei dieses immer noch gut bewertet wurde. Eine genauere Analyse erlaubt hier eine Betrachtung der Itemebene (Abb. 86):

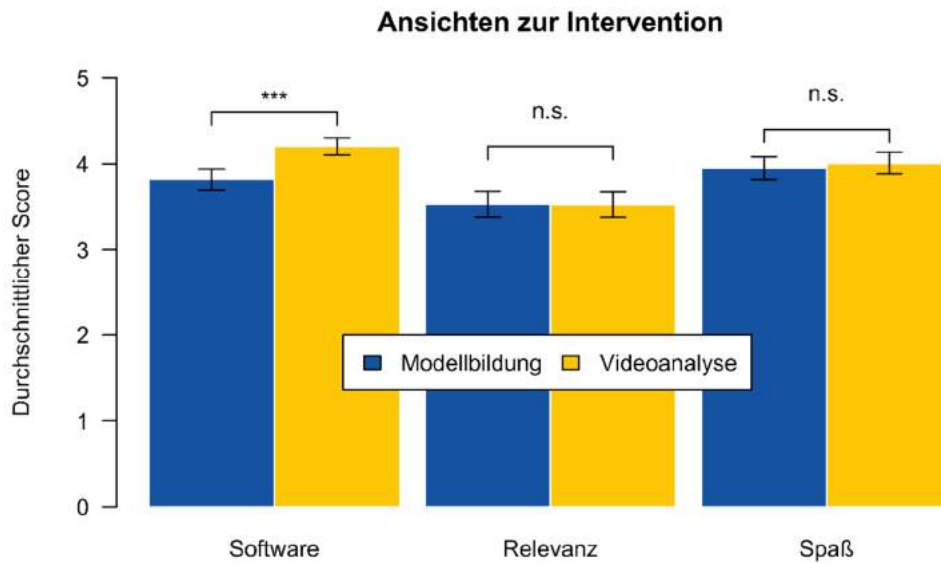


Abb. 85: Durchschnittlicher Score in den Skalen "Software", "Relevanz" und "Spaß" in beiden Gruppen mit 95%-Konfidenzintervallen. *** $p < 0,001$ n.s. nicht signifikant.

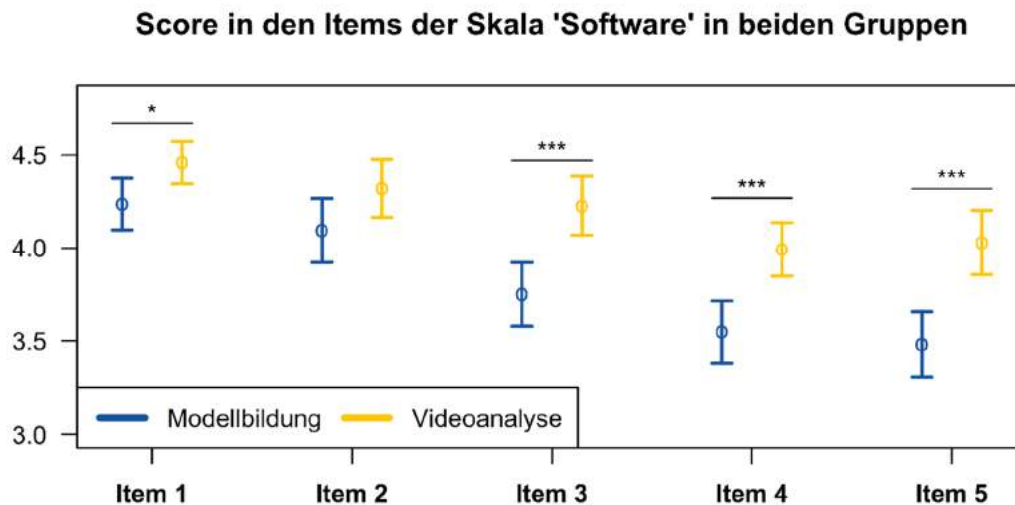


Abb. 86: Durchschnittlicher Score in den einzelnen Items aus der Skala "Software" in beiden Gruppen mit 95%-Konfidenzintervallen. *** $p < 0,001$, * $p < 0,05$, sonst nicht signifikant.

Nahezu in allen Items wird die Software „measure dynamics“ besser bewertet als die Software „Newton-II“. Den größten Unterschied gibt es in dem Item 5 „Ich hatte Schwierigkeiten bei der Nutzung des Programms“ ($p < 0,001$; $d = 0,53$ [0,29; 0,8]). Dieses Item wird umkodiert dargestellt, sodass ein höherer Score mit weniger Schwierigkeiten einhergeht. Ähnlich große Unterschiede gibt es in den Items 3 „Ich konnte mich schnell

in das Programm einarbeiten“ ($p < 0,001$; $d = 0,49$ [0,27; 0,76]) und 4 „Das Programm ist übersichtlich und leicht zu bedienen“ ($p < 0,001$; $d = 0,48$ [0,22; 0,73]). Den Items 1 „Das Programm ist gut geeignet, um Physik zu lernen“ ($p < 0,05$; $d = 0,30$ [0,08; 0,53]) und 2 „Ich würde das Programm gerne auch im Schulunterricht einsetzen“ ($p = 0,054$) wird in beiden Gruppen sehr zugestimmt.

Der Unterschied in der Gesamtskala kann durch Schwierigkeiten mit dem Modellieren mit „Newton-II“ erklärt werden. Die Proband*innen würden aber auch diese Software gerne in der Schule einsetzen. „Measure dynamics“ wird sehr gut angenommen und kommt scheinbar auch mit weniger Schwierigkeiten bei der Benutzung daher. Eine inhaltliche Diskussion der tatsächlich vorhandenen Schwierigkeiten erfolgt in Kapitel 11.1.2.

10.3 Cognitive Load

Die Cognitive Load Theory (CLT) liefert eine Theorie, die den Erfolg von Lernumgebungen erklären kann. Sie postuliert eine Begrenztheit des Arbeitsgedächtnisses und unterscheidet zwischen drei unterschiedlichen Arten kognitiver Belastung (siehe auch Kapitel 4.1). Die beiden Interventionen sollen im Hinblick auf den Cognitive Load verglichen werden. Die Hypothese war, dass sich die höhere Komplexität der mathematischen Modellbildung in einem höheren Intrinsic Cognitive Load äußert. Offen ist, ob diese möglicherweise höhere intrinsische Belastung zu einer Reduktion des Germane Cognitive Load führt oder ob die andere Art der Auseinandersetzung mit Kräften möglicherweise sogar zu einem gesteigerten Germane Cognitive Load führt. Der Cognitive Load wurde durch einen Fragebogen im Nachtest erhoben (siehe Kapitel 8.4.2). Dabei ergaben sich für die beiden Gruppen folgendes Bild (Abb. 87).

Erfreulich ist der in beiden Gruppen niedrige Extranous Cognitive Load, der vor allem von der Lernumgebung abhängt, und der hohe Germane Cognitive Load, der lernrelevant ist. Das spricht in beiden Fällen für einen guten Aufbau der Software und gut gestaltete Materialien und erklärt den hohen Lernzuwachs in beiden Gruppen.

Der Intrinsic Cognitive Load unterscheidet sich erwartungsgemäß signifikant zwischen den Gruppen, allerdings mit einer kleinen Effektstärke von $d = 0,40$. Das kann u. a. darauf zurückzuführen sein, dass Reibungseinflüsse in dieser Gruppe explizit (in Gleichungsform) besprochen wurden und dies die gefühlte Schwierigkeit der Thematik im Vergleich zur Videoanalyse erhöht. Eine Betrachtung auf Itemebene bestätigt dies (Abb. 88).

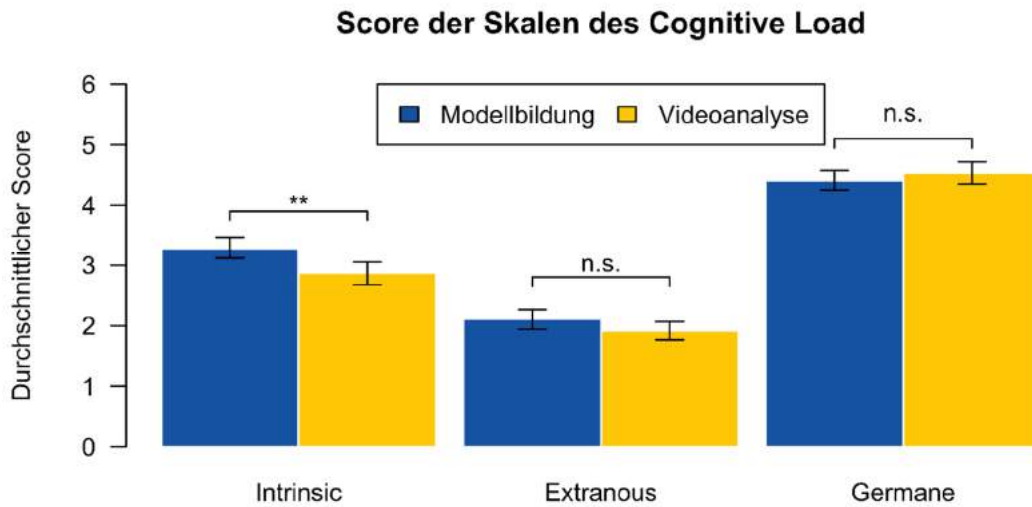


Abb. 87: Durchschnittliche Punktzahlen in den Skalen zum Cognitive Load in beiden Gruppen mit 95%-Konfidenzintervallen. ** $p < 0,01$ n.s. nicht signifikant.

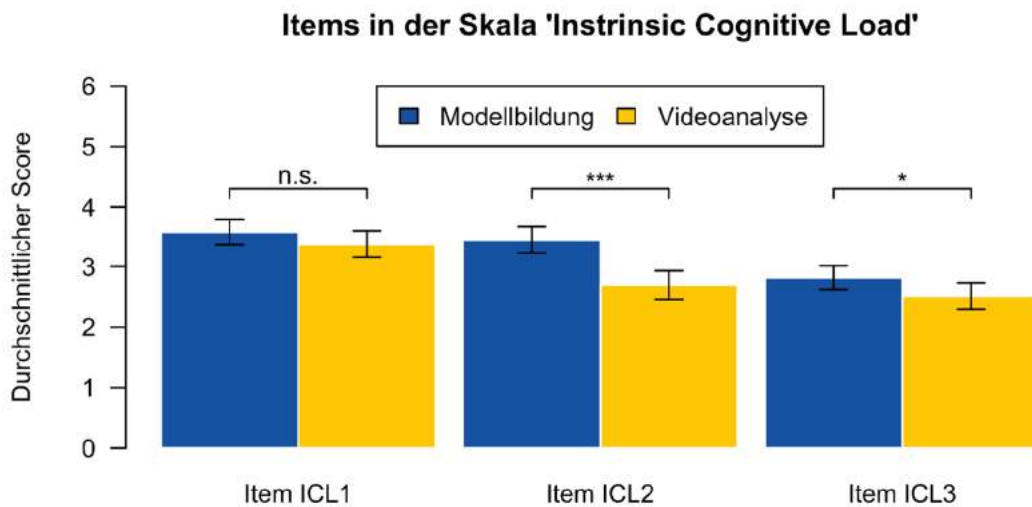


Abb. 88: Durchschnittliche Punktzahlen in den einzelnen Items (max. 6) innerhalb der Skala Intrinsic Cognitive Load in beiden Gruppen mit 95%-Konfidenzintervall. *** $p < 0,001$ * $p < 0,05$ n.s. nicht signifikant.

Die hier dargestellten Items lauten:

- ICL1: „Das Thema ‚Dynamik‘ war sehr komplex.“
- ICL2: „Die Formeln zur ‚Dynamik‘ empfand ich als sehr komplex.“
- ICL3: „Die physikalischen Konzepte und Definitionen zur ‚Dynamik‘ empfand ich als sehr komplex.“

Das Item, auf das der größte Unterschied in der Skala zurückzuführen ist (mittlere Effektstärke von $d = 0,55$), fragt also nach den in der Intervention vorkommenden Formeln, die in der Modellbildung komplexer waren, da Reibung explizit modelliert werden musste und nicht nur qualitativ besprochen wurde. Insgesamt wird dem Thema Dynamik von der Gruppe der Modellbildung aber keine höhere Komplexität zugesprochen, was dafürsprechen könnte, dass trotz der komplexeren Themen ein ähnliches Verständnis für die Grundaussagen erzeugt werden könnte wie in der Gruppe der Videoanalyse.

10.4 Theoretisierende vs. Experimentierende

Eine weitere Fragestellung war, ob die Interventionen je nach Interesse der Proband*innen am Theoretisieren bzw. am Experimentieren erfolgreicher waren. Dazu wurden diese beiden Konzepte im Test abgefragt (siehe Kapitel 8.4.2). Zunächst wird der Score in der Skala „Theoretisieren“ über dem Score in der Skala „Experimentieren“ aufgetragen, um sich ein Bild über die Verteilung zu machen (Abb. 89).

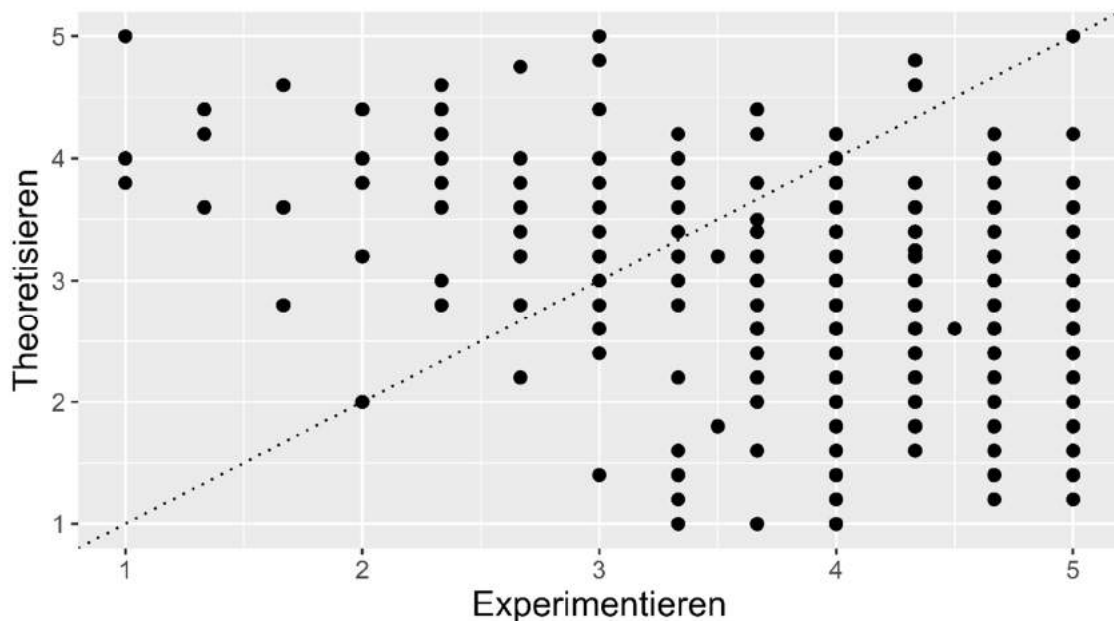


Abb. 89: Score in der Skala "Theoretisieren" über dem Score in der Skala "Experimentieren" mit Winkelhalbierender (gepunktet).

In beiden Skalen waren Werte von 1-5 zu erreichen. Die Items in beiden Skalen waren analog formuliert, um eine Vergleichbarkeit zu ermöglichen. Man sieht, dass es verschiedene Ausprägungen gibt: Proband*innen, die großes Interesse am Theoretisieren aber nicht am Experimentieren haben und umgekehrt. Zudem gibt es generell sehr Interessierte und Teilnehmer*innen, die an beidem wenig Interesse haben. Eingezeichnet ist zudem

eine gepunktete Winkelhalbierende. Alle Proband*innen, deren Punkte sich oberhalb dieser Linie befinden, haben mehr Interesse an theoretischen Zusammenhängen als am Experimentieren und bei allen unterhalb dieser Linie ist es umgekehrt. Man erkennt an der Winkelhalbierenden, dass es insgesamt eine Neigung in Richtung „Experimentieren“ gibt. Kontrolliert werden konnte dabei nicht, was die Proband*innen unter Experimentieren verstehen – also ob Vorbereitung, Durchführung, Auswertung und Interpretation oder nur einzelne Phasen davon.

Als erster Ansatz für eine Kategorisierung wurde die Differenz zwischen den beiden Scores gebildet und jedem Probanden dadurch ein Score zwischen -4 und 4 zugeteilt. Ein Histogramm der Differenz der beiden Scores zeigt erneut die Verschiebung in Richtung „Experimentieren“ (Abb. 90).

Es wäre möglich, die Proband*innen auf Basis dieser eindimensionalen Größe in zwei Gruppen zu teilen – in Personen, die eher zum Theoretisieren neigen, und Personen, die eher zum Experimentieren neigen. Diese Reduktion auf eine Dimension erlaubt aber keine Unterscheidung zwischen Personen, die an beidem gleichermaßen niedriges oder gleichermaßen hohes Interesse haben. Besser ist also eine Einteilung unter Einbezug beider Dimensionen. Zudem hat die Faktorenanalyse getrennte Faktoren ergeben. Für die Einteilung bietet sich dann eine Clusteranalyse an.

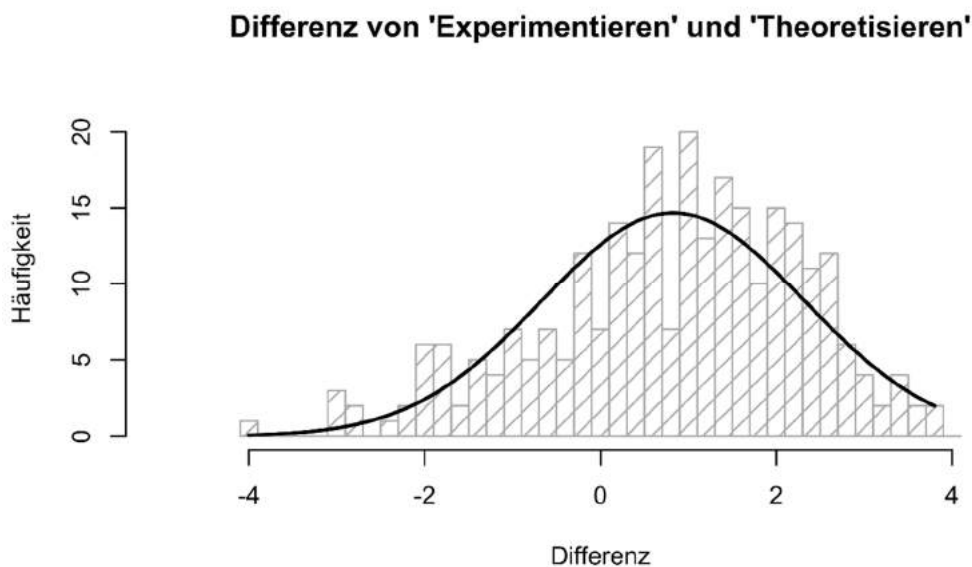


Abb. 90: Differenz aus dem Score in der Skala "Experimentieren" und "Theoretisieren". Ein positiver Score bedeutet ein höherer Score in der Skala "Experimentieren".

Clusteranalysen geben neben der Frage, ob überhaupt eine signifikante Clusterung vorliegt, Antworten auf zwei weitere Fragen. Eine Frage betrifft die optimale Anzahl von Clustern, die nächste betrifft die optimale Einteilung der Einheiten in diese Cluster. Für die Anzahl an Clustern gibt es, ähnlich wie bei der explorativen Faktorenanalyse, kein eindeutiges Kriterium. Ein Blick auf das sogenannte Dendrogramm gibt einen ersten Eindruck (Abb. 91).

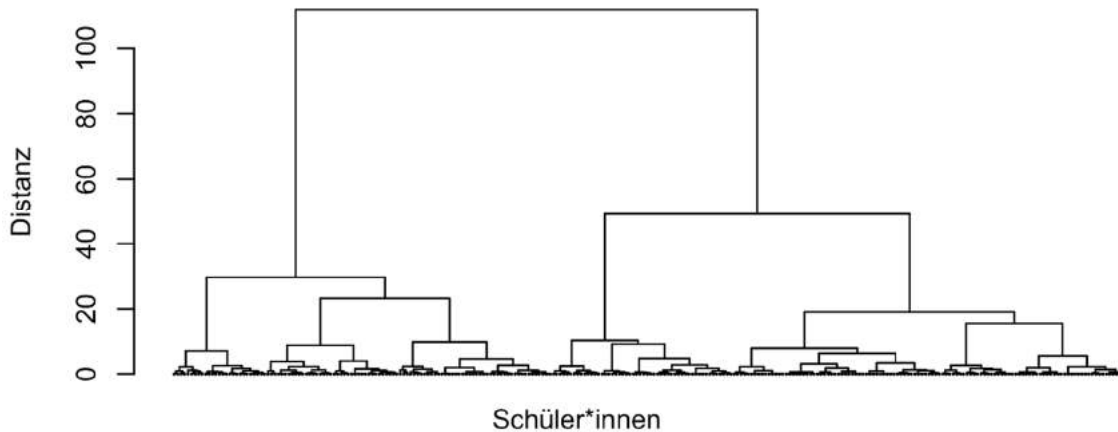


Abb. 91: Dendrogramm der Einheiten.

Auf der x-Achse sind die einzelnen Proband*innen aufgetragen. Mithilfe einer Distanzmetrik (hier euklidisch) werden in einem fortlaufenden Prozess Nachbarn zu einem Cluster hinzugefügt. Diese werden dann wieder mit einem benachbarten Cluster zusammengefügt usw. Anhand der Struktur lassen sich erste Tendenzen im Hinblick auf die Art und die Anzahl der Cluster erkennen. Es gibt allerdings kein festes Kriterium anhand dessen man diese Fragen zweifelsfrei entscheiden könnte. In diesem Fall lägen zwei oder drei Cluster am nächsten. Die weiteren Methoden zur Bestimmung von Cluster ergeben bei den hier vorliegenden Daten in der Regel drei Cluster (Salvador & Chan, 2004; Sugar & James, 2003; Thorndike, 1953; Tibshirani et al., 2001). Bei drei Clustern liefert die beste Einteilung in die beiden Cluster nach dem k-Means-Algorithmus (MacQueen, 1967) folgendes Bild (Abb. 92):

Das k-Means-Verfahren versucht dabei eine möglichst große Ähnlichkeit innerhalb der Cluster und eine möglichst geringe Ähnlichkeit zwischen den Clustern herzustellen. Verwendet wurde dabei der Algorithmus von Hartigan und Wong (1979) unter Nutzung der euklidischen Distanzmetrik. Diese Art der Clusterung führt zu einer Teilung zwischen „Theoretiker*innen“, „Experimentierer*innen“ und Proband*innen, die generell interessiert sind (mit Neigung zum Experimentieren). Inhaltlich zu kritisieren ist dabei, dass die Proband*innen, welche in beiden Skalen eher niedrige Scores verzeichnen, zu den

Theoretiker*innen gezählt werden. Es ist bei der Entscheidung für eine Anzahl an Clustern auch möglich neben statistischen auch inhaltliche Argumente heranzuziehen. Vier Cluster erscheinen hier inhaltlich sinnvoller (Abb. 93).

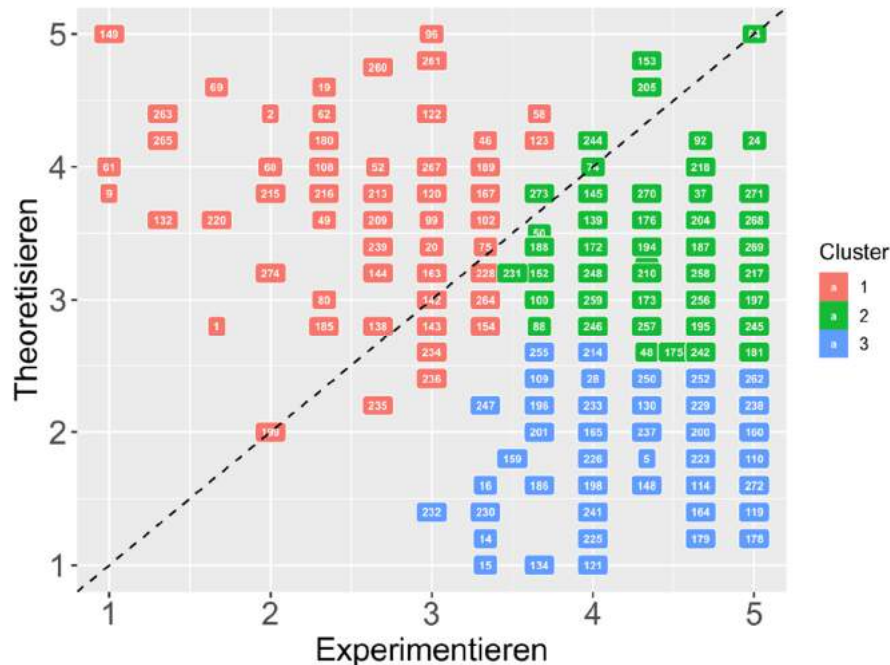


Abb. 92: Score in der Skala "Theoretisieren" über dem Score in der Skala "Experimentieren" mit Winkelhalbierender. Dabei sind die Proband*innen in drei Cluster nach dem k-Means-Verfahren eingeteilt.

Unter Einbezug der Schiefe in den Daten können die Cluster folgendermaßen benannt werden: „Interessierte“ (C1), „Desinteressierte“ (C2), „Experimentierende“ (C3), „Theoretisierende“ (C4), wobei bei allen eine Tendenz zum Interesse am Experimentieren zu sehen ist. Der Silhouettenkoeffizient ergibt hier eine schwache Strukturierung der Daten (Ester & Sander, 2000). Eine ANCOVA zeigt einen unter Kontrolle des Vortestergebnisses signifikanten Einfluss der Cluster auf das Nachtestergebnis (Abb. 94).

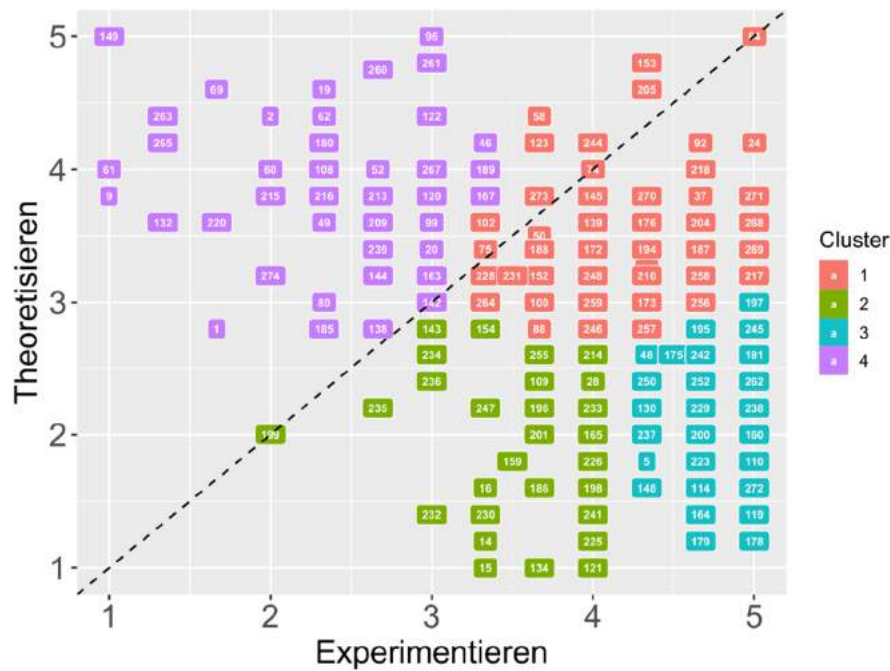


Abb. 93: Score in der Skala "Theoretisieren" über dem Score in der Skala "Experimentieren" mit Winkelhalbierender. Dabei sind die Proband*innen in vier Cluster nach dem k-Means-Verfahren eingeteilt.

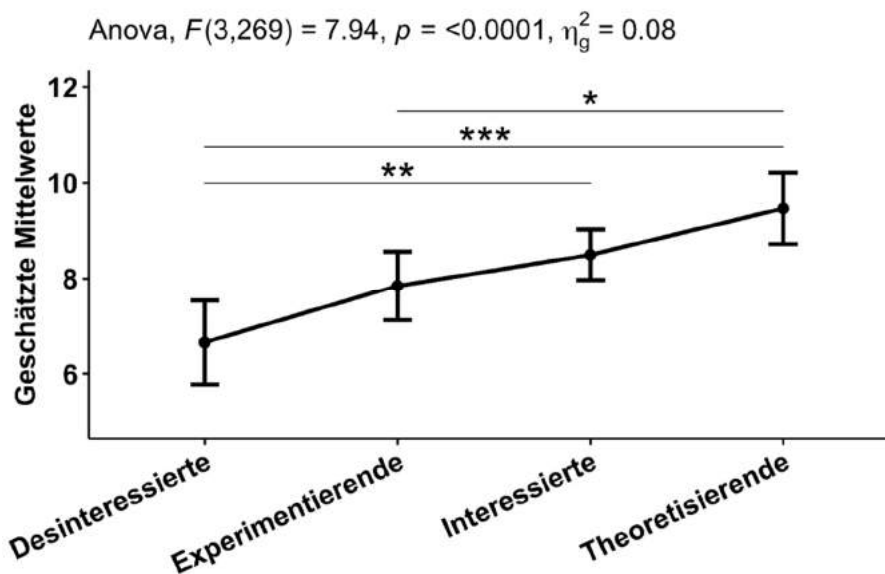


Abb. 94: Nachtestergebnisse mit 95%-Konfidenzintervallen für die vier Cluster bei Kontrolle des Vortestergebnisses.

Dieser Effekt ist mit einem generalisierten Eta-Quadrat von $\mu_g^2 = 0,08$ (Bakeman, 2005) als mittel bis hoch einzustufen. Diese Effektstärke korrespondiert nach Lenhard und Lenhard (2017) mit einer Effektstärke nach Cohen von $d = 0,59$. Diejenigen Schüler*innen ($N = 50$), die primär an theoretischen Zusammenhängen interessiert sind, erreichen

im Nachtest im Schnitt 5,48 Punkte mehr als im Vortest (bei 15 möglichen Punkten), während dieser Zuwachs bei den „Desinteressierten“ lediglich 1,59 Punkte beträgt. „Experimentierende“ und „Interessierte“ legen im Schnitt 3,38 bzw. 3,91 zu. Die Cluster haben also einen mittleren bis großen Einfluss auf das Nachtestergebnis unter Kontrolle des Vortests. An theoretischen Zusammenhängen interessierte Proband*innen profitieren dabei am meisten von der Intervention. Dies zeigte sich bereits durch den in Kapitel 9.3.8 dargestellten Einfluss des Interesses an theoretischen Zusammenhängen auf das Nachtestergebnis im hierarchischen linearen Modell.

Eine weitere Hypothese war, dass die Cluster nicht nur einen Einfluss auf den Lernzuwachs haben, sondern darüber hinaus eine Wechselwirkung mit den Gruppen besteht – also, dass „Theoretisierende“ in der Gruppe der mathematischen Modellbildung mehr lernen und „Experimentierende“ in der Gruppe der Videoanalyse. Dies konnte nicht bestätigt werden – auch weil eine Unterteilung der Stichproben pro Gruppe in je vier Cluster zu kleinen und damit wenig aussagekräftigen Stichprobengrößen führt.

10.5 Schülervorstellungen

Bisher wurde der Fragebogen lediglich dichotom ausgewertet und es wurde keine Auswertung der unterschiedlichen falschen Antwortmöglichkeiten vorgenommen. Eine differenzierte Auswertung des Fragebogens erlaubt einen Rückschluss auf die zugrundeliegenden Schülervorstellungen. Ein Ziel der Intervention war die Verbesserung von gängigen Schülervorstellungen. Der bisherige Vergleich der beiden Gruppen legt nahe, dass die Veränderung der Schülervorstellungen in beiden Gruppen gleichermaßen vonstattengegangen ist. In dem Fragebogen sind vor allem Schülervorstellung zur Dynamik enthalten, allerdings sind auch kinematische Schülervorstellungen bei den Antworten einiger Items erkennbar. Die Schülervorstellungen wurden in Kapitel 3 inhaltlich vorgestellt und diskutiert. Folgende Schülervorstellungen sind in den Antwortmöglichkeiten der Items enthalten²⁸:

- (SV1) Zur Aufrechterhaltung der Bewegung bedarf es einer Kraft in Bewegungsrichtung.
- (SV2) Bei Kreisbewegungen wirkt die Zentrifugalkraft.
- (SV3) Die Bewegungsform prägt sich dem bewegenden Objekt ein.

²⁸ Die Schülervorstellungen sind dabei u. a. aus Schecker et al. (2018) übernommen.

- (SV4) Ein Objekt kann sich nur mit konstanter Geschwindigkeit bewegen, wenn keine Kräfte wirken. Das Beharrungsprinzip gilt nur für Spezialfälle.
- (SV5) Die Kräfte, die zwischen zwei Körpern wirken, sind abhängig von deren Masse: „Je schwerer desto stärker“.
- (SV6) Eine Kraft braucht eine Zeit, bis sie sich durchsetzen kann. Die stärkere Kraft gewinnt.
- (SV7) Ein bewegter Körper hat Kraft.
- (SV8) Bei einer gleichförmigen Kreisbewegung wirken keine Kräfte.
- (SV9) Nur aktive Körper können Kräfte ausüben, passive leisten Widerstand.
- (SV10) Alle Körper kommen natürlicherweise zum Stillstand.
- (SV11) Undifferenziertheit von Geschwindigkeit und Beschleunigung.

Die Einteilung der Schülervorstellungen selbst ist nicht eindeutig. Die Schülervorstellungen sind in unterschiedlich vielen Items enthalten und kommen auch unterschiedlich häufig vor. In der Folge werden nur die fünf am häufigsten auftretenden Vorstellungen betrachtet (1, 5, 7, 10 und 11). Die Schülervorstellungen werden dabei zunächst dichotom ausgewertet. Das heißt, dass eine Schülervorstellung entweder vorhanden oder nicht vorhanden ist. Sie wird als vorhanden betrachtet, wenn mindestens ein Item so beantwortet wird, dass es für das Vorhandensein der Schülervorstellung spricht. Außerdem wird der Ausprägungsgrad betrachtet. Der Anteil der Schüler*innen, der in Vor- und Nachtest nach der entsprechenden Schülervorstellung geantwortet hat, ist in Abb. 95 dargestellt.

Die Vorstellung, dass eine Kraft in Bewegungsrichtung nötig ist, damit sich ein Objekt bewegen kann (SV1), tritt am häufigsten auf. Sie reduziert sich aber im Nachtest auch signifikant (mit Effektstärke $r = 0,34$). In SV7 ist ebenfalls eine große Änderung zu sehen ($r = 0,51$), während diese in SV5 ($r = 0,19$) kleiner ausfällt. In SV10 und vor allem SV11 ist keine Änderung zu sehen. SV11 ist die einzige kinematische Schülervorstellung. Es ist nicht verwunderlich, dass hier am wenigsten passiert. Die Interventionen waren insbesondere darauf ausgelegt, SV1 zu reduzieren, da diese Schülervorstellung sehr verbreitet und hartnäckig ist und eine typische Fehlerquelle darstellt. Aufgrund dessen wird diese Schülervorstellung SV1 nun im Vergleich der beiden Gruppen betrachtet (Abb. 96).

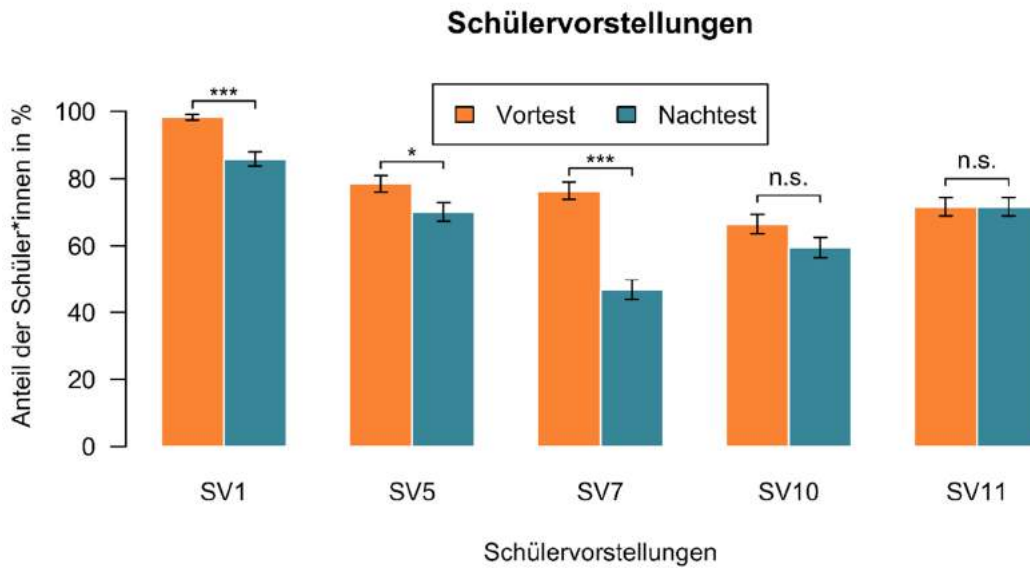


Abb. 95: Die häufigsten in den Schülerantworten im Test vorkommenden Schülervorstellungen in Vor- und Nachtest mit Standardfehlern. *** $p < 0,001$, * $p < 0,05$, n.s. nicht signifikant.

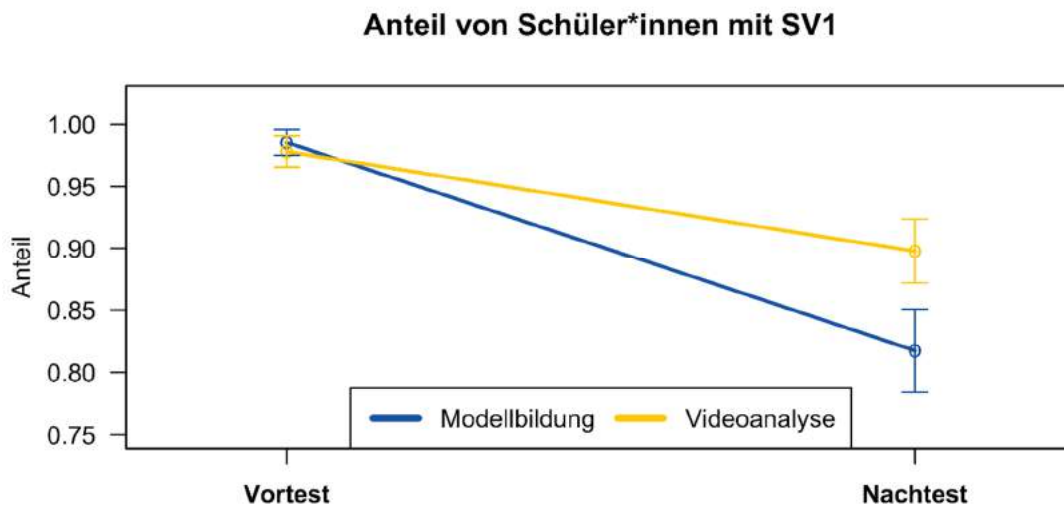


Abb. 96: Anteil an Schüler*innen, die mindestens einmal so antworten, als sei SV1 aktiviert worden in Vor- und Nachtest mit Standardfehler in beiden Gruppen in Vor- und Nachtest. SV1: Eine Kraft in Bewegungsrichtung ist notwendig.

Der Anteil an Schüler*innen, der diese Schülervorstellung mindestens einmal aktiviert, reduziert sich in der Gruppe der Modellbildung stärker als in der Gruppe der Videoanalyse (Wilcoxon-Rangsummentest, $p < 0,05$), während es in den anderen Schülervorstellungen keine Unterschiede zwischen den Gruppen gibt. In SV7, in der es ebenfalls eine große Änderung von Vor- zu Nachtest gibt, sieht die Entwicklung wie in Abb. 97 dargestellt aus. Der im Bild sichtbare Unterschied, dass sich der Anteil an Schüler*innen, die

mindestens einmal so antworten, als sei SV7 aktiviert worden, in der Gruppe der Modellbildung stärker reduziert, ist nicht signifikant.

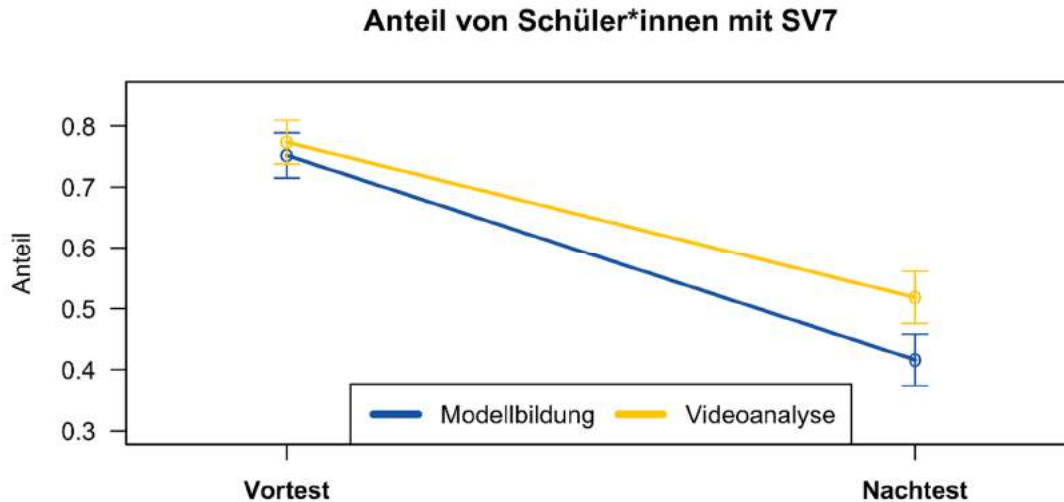


Abb. 97: Anteil an Schüler*innen, die mindestens einmal so antworten, als sei SV7 aktiviert worden in Vor- und Nachtest mit Standardfehler in beiden Gruppen in Vor- und Nachtest. SV7: Ein bewegter Körper hat Kraft.

Bislang wurde der Anteil an Schüler*innen betrachtet, bei dem die entsprechende Schülervorstellung auftritt. Alternativ kann man auch das prozentuale Auftreten der Schülervorstellung in den Items analysieren. Dann ergibt sich für die erste Schülervorstellung das in Abb. 98 dargestellte Bild. Erkennbar ist eine stärkere Reduktion des Auftretens der Schülervorstellung in der Gruppe der Modellbildung. Der Unterschied ist aber nicht signifikant. Die ANCOVA schätzt hier bei einem Vorkommen von 48,3 % im Vortest ein Vorkommen von 30,6 % [27,0 %; 34,3 %] in der Gruppe der Modellbildung und 33,6 % [29,8 %; 37,1 %] in der Gruppe der Videoanalyse.

Zusammengefasst werden kann, dass die meisten Schülervorstellungen, die die Dynamik betreffen, durch die Interventionen reduziert werden. Während es in den meisten Schülervorstellungen keinen Unterschied zwischen den Gruppen gibt, reduziert sich der Anteil der Schüler*innen, die die am häufigsten vorkommende Schülervorstellung aktivieren, in der Gruppe der Modellbildung scheinbar stärker. Dies könnte die Begründung dafür sein, dass der Lernzuwachs in Items zum ersten Newtonschen Gesetz in der Gruppe der Modellbildung größer ist. Der Frage, warum sich diese Schülervorstellung möglicherweise stärker reduziert, wird in der Auswertung der qualitativen Daten nachgegangen (Kapitel 11.3).

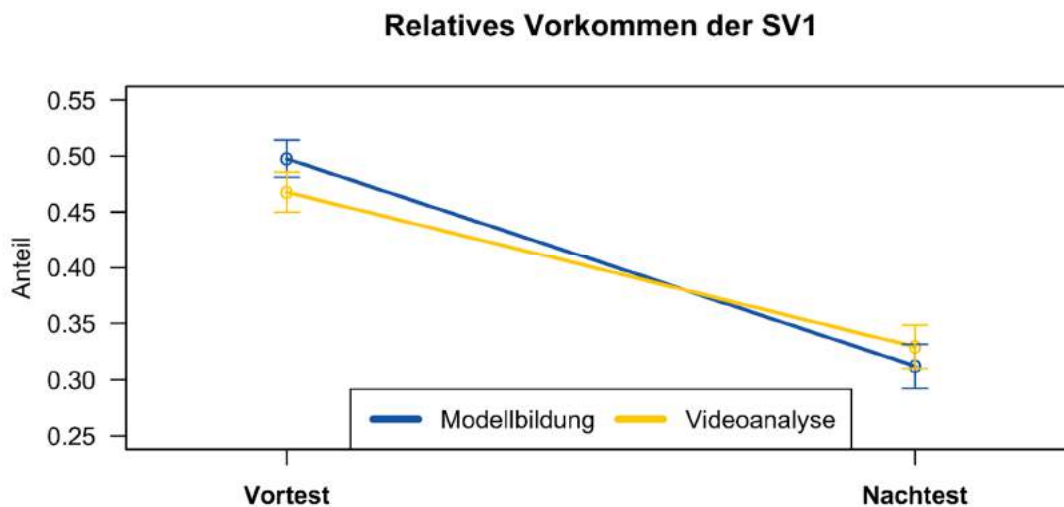


Abb. 98: Anteil der Items, bei denen so geantwortet wird, als sei SV1 aktiviert worden, an den Items, bei denen es einen Distraktor gibt, der für eine Aktivierung dieser Schülervorstellung spricht mit Standardfehler in beiden Gruppen in Vor- und Nachtest.

10.6 Auswertung des Lehrerfragebogens

Während der Interventionen füllten auch die Lehrkräfte der Kurse einen Fragebogen aus, der zwei Teile einhielt. Einerseits ist von Interesse, welcher Art von Computereinsatz in der Klasse vor der Durchführung der Intervention bereits stattgefunden hat. Andererseits sollten die Lehrkräfte Items zum Konzeptverständnis beantworten. Die Items, die die Schüler*innen beantworten mussten, waren dabei ein Teil der Items des Lehrerfragebogens. Zu diesen Items sollte von den Lehrkräften beantwortet werden, ob sie Unterrichtsstoff aus der Mechanik der E-Phase testen. Zusätzlich sollte die Schwierigkeit der Items eingeschätzt werden. Auf Basis dieser Daten können die auch bei Lehrkräften noch vorhandenen Fehlvorstellungen identifiziert und die geschätzte Schwierigkeit mit der tatsächlichen Schwierigkeit für Schüler*innen verglichen werden. Bei großen Unterschieden kann auf eine mögliche Ursache eingegangen werden. Insgesamt liegen von 19 der 20 teilnehmenden Lehrkräfte ausgefüllte Fragebögen vor (20 Lehrkräfte waren mit je einer und eine Lehrkraft mit zwei Klassen zu Besuch).

10.6.1 Computereinsatz im Unterricht

Um den Einsatz von Computern, Tablets und Smartphones im Unterricht einschätzen zu können, wurde im Fragebogen nach der Häufigkeit ihres Einsatzes gefragt. Inkludiert waren dabei auch andere Medien, wie Tafel, Schulbuch oder Arbeitsblatt (Abb. 99).

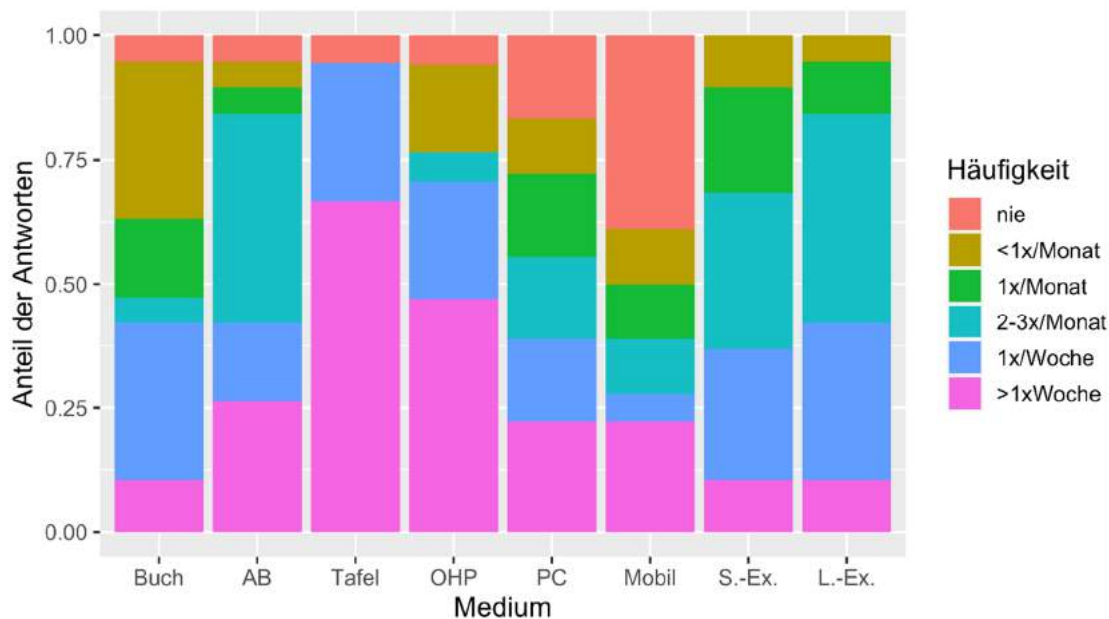


Abb. 99: Anteile der Antworten zu den Items zum Medieneinsatz im Unterricht. Buch: Schulbuch, AB: Arbeitsblatt, Tafel: Tafel u. interaktives Whiteboard (nur zum Anschreiben), OHP: Overhead-Folien u. Dokumentenkamera, PC: PC u. Laptop, Mobil: Tablet u. Smartphone, S.-Ex.: Schülerexperimente, L.-Ex.: Lehrerexperimente.

Die Nutzung von klassischen Computern, die in der Intervention eingesetzt wurden, ist im Vergleich zu den anderen Medien im mittleren Bereich anzusiedeln. Mobilgeräte wie Tablets und Smartphones werden seltener eingesetzt. Die klassischen Medien Tafel, Overhead-Projektor und Arbeitsblatt werden immer noch am häufigsten im Unterricht eingesetzt. Um wiederum zwischen den unterschiedlichen Arten des Computereinsatzes unterscheiden zu können, wurde hier konkret nachgefragt (Abb. 100).

Am häufigsten werden das Internet als Informationsmedium und als Kommunikationsmedium und Simulationen und Animationen genutzt. Am wenigsten werden Mind-Maps/Concept-Maps und das Arbeiten mit kompletten Lerneinheiten genutzt. Mathematische Modellbildung wird von den Lehrkräften, die zu Besuch im Schülerlabor waren, weniger genutzt (Median: noch nie genutzt) als die Videoanalyse (Median: schon einmal genutzt). In einer Befragung von Wenzel (2018) nutzten die Lehrkräfte im Vergleich Videoanalyse etwas seltener und mathematische Modellbildung häufiger als in dieser Studie. Insgesamt sind die Ergebnisse aber ähnlich. Konkret sagen hier zehn der Lehrkräfte, dass sie keine Modellbildungssoftware nutzen. Die anderen neun Lehrkräfte nutzen dafür Tabellenkalkulationsprogramme, zwei Lehrkräfte nennen in dem offenen Antwortfeld zusätzlich die Software „Geogebra“. „Newton-II“ wurde von keiner Lehrkraft bereits genutzt. Bei der Videoanalyse sagen sieben Lehrkräfte, dass sie keine Software für die Videoanalyse nutzen, vier Mal wird „measure dynamics“ genannt, drei Mal „Viana“ und

ein Mal „Video Analyzer“. In dem offenen Antwortfeld werden an Softwareprogrammen zusätzlich das Videoanalyseprogramm „NewtonDV“, die Videoanalyse-App „Coach’s Eye“, die Messwerterfassungs-App „Phyphox“, das Messwerterfassungsprogramm „Cassy“ („Leyboldt“), die Dynamische.-Geometrie-Software „GeoGebra“, die Simulationen von Phet.colorado.edu und die manuelle Videoanalyse per Beamer und Tafel genannt. Die Häufigkeit des vorigen Computereinsatzes hat in der Mehrebenenanalyse jedoch keinen Einfluss auf die Nachtestergebnisse der Schüler*innen.

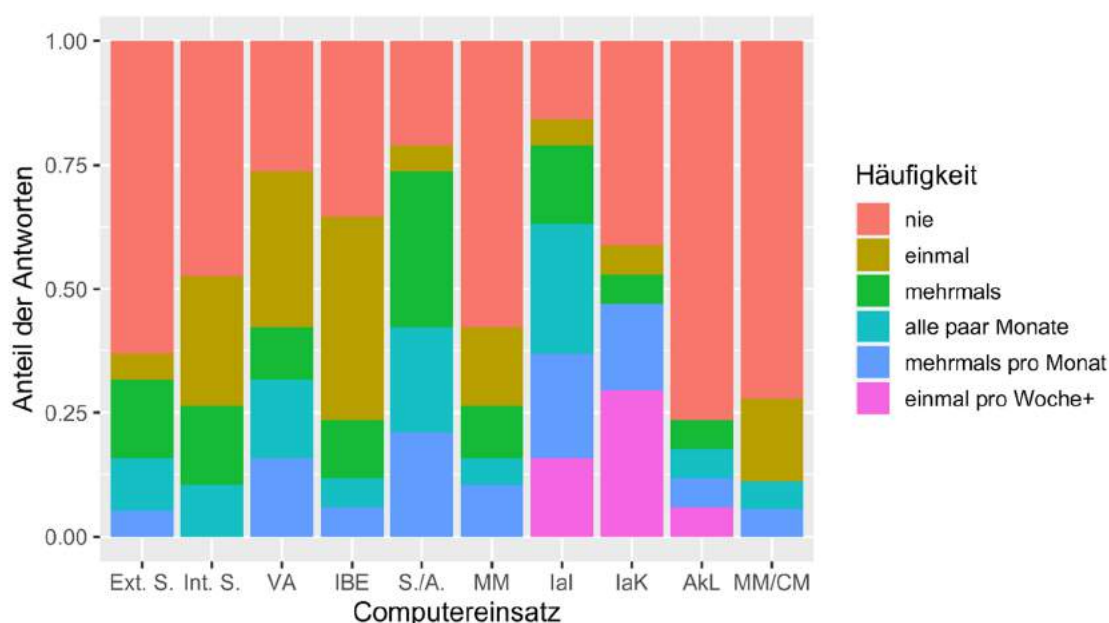


Abb. 100: Boxplots der Antworten in den Items zum Computereinsatz im Unterricht: Ext. S.: Messwerterfassung mit externen Sensoren, Int. S.: Messwerterfassung mit internen Sensoren von Tablets/Smartphones, VA: Videoanalyse von Bewegungen, IBE: Interaktive Bildschirmexperimente, S./A.: Arbeiten mit Simulationen oder Animationen, MM: Mathematische Modellbildung, IaI: Internet als Informationsmedium, IaK: Internet als Kommunikationsmedium, AkL: Arbeiten mit kompletten Lerneinheiten, MM/CM: Mindmaps u. Concept Maps.

10.6.2 Bewertung der Items

Um die Validität des eingesetzten Tests zu bestätigen, wurden die Lehrkräfte danach gefragt, ob die verwendeten Items Inhalte aus der Mechanik der E-Phase testen. Teilweise gehören im Test zwei Items zu einer Frage (z. B. 13a und 13b). Da die Zustimmung nur auf Basis der Frage erhoben wurde und die Items innerhalb einer Frage inhaltlich ähnlich sind, wird die Zustimmung unten auf Fragenebene angegeben (Tab. 60).

Tab. 60: Zustimmung zur Frage "Testet das Item Inhalte aus der Mechanik der E-Phase?" in jedem Item.

Frage	Zustimmung	Frage	Zustimmung	Frage	Zustimmung
1	88,2 %	7	100,0 %	12	100,0 %
2	100,0 %	8	94,1 %	13 (a+b)	94,7 %
3	94,7 %	9	82,4 %	14	77,8 %
4	100,0 %	10 (a+b)	94,7 %	15 (a+b+c+d)	100,0 %
5 (a+b)	88,2 %	11	100,0 %	16 (a+b)	88,9 %
6	100,0 %				

Die Zustimmung der Lehrkräfte zu der Frage, ob das besagte Item Inhalte aus der Mechanik der E-Phase testet, ist im Allgemeinen sehr hoch. Am niedrigsten ist die Zustimmung bei Frage 14 (Abb. 101):

Frage 14

Ein Golfball bewegt sich längs eines "Fairway" mit folgender Flugbahn durch die Luft.



Welche Kraft wirkt, bzw. welche Kräfte wirken, während der gesamten Flugphase auf den Ball, also **nachdem** der Ball den Schläger verlassen hat?

1. Gravitationskraft
2. Abschlagskraft
3. Luftwiderstandskraft

☐ Nur 1
☐ 1 und 3

☐ 1 und 2
☐ 2 und 3

☐ 1, 2 und 3

Abb. 101: Frage 14 aus dem Testinstrument.

Dies liegt möglicherweise an der Art und Weise, wie nach den wirkenden Kräften gefragt wird. Andererseits ist der schräge Wurf in Grundkursen kein Standardthema mehr. Möglicherweise ist auch die genannte Luftreibung der entscheidende Punkt, warum einige Lehrer*innen die Aussage für dieses Item abgelehnt haben. Da alle weiteren Items eine Zustimmung von über 80 % haben, kann davon ausgegangen werden, dass der Fragebogen relevante Inhalte der Mechanik der E-Phase überprüft. Nicht hervorgeht aus der

Befragung, ob der Fragebogen den gesamten Bereich der Mechanik oder auch der Dynamik abdeckt.

Zusätzlich wurden die Lehrkräfte darum gebeten, die Schwierigkeit der Items auf einer siebenstufigen Likert-Skala von „zu leicht“ bis „zu schwer“ einzuschätzen (Abb. 102). Die Items kommen dabei auf folgende Scores (geordnet nach dem arithmetischen Mittel):

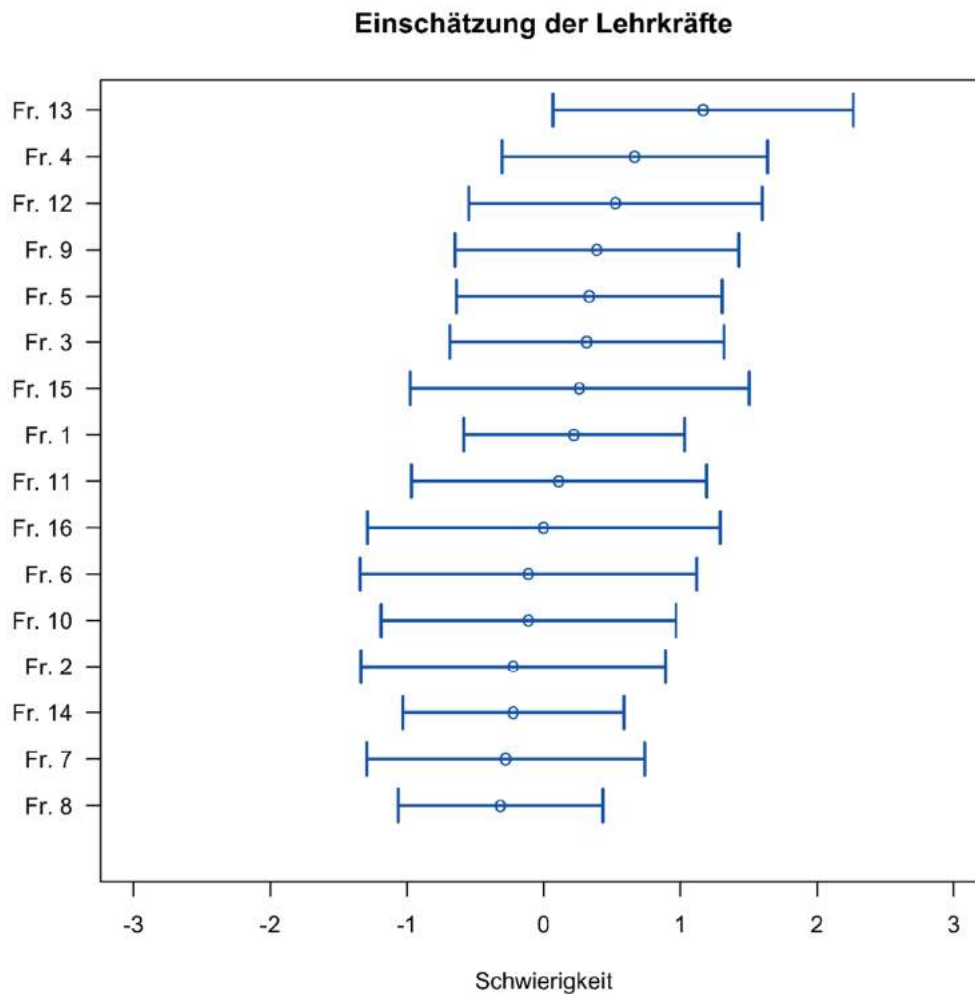


Abb. 102: Plots für die Einschätzung der Schwierigkeit für jedes Item von "zu leicht" (-3) bis "zu schwer" (3) mit arithmetischem Mittel und Standardabweichung. 0 steht dementsprechend für eine angemessene Schwierigkeit.

Der Median liegt bei allen Items bei 0 oder 1. Die Items haben laut Lehrkräften also eine angemessene Schwierigkeit. Sie werden dennoch als unterschiedlich schwierig eingeschätzt, was für einen solchen Fragebogen erwünscht ist. Das am schwierigsten eingeschätzte Item ist folgendes (Abb. 103).

Frage 13

Für die folgenden Aufgaben gilt folgende Situation: Ein LKW bleibt mit Motorschaden liegen und wird von einem Kleinwagen (Auto) zur nächsten Tankstelle geschoben.



- a) Während das Auto beschleunigt, um beim Schieben auf eine bestimmte Geschwindigkeit zu kommen, gilt:
- ☐ Die Stärke der Kraft, mit der das Auto gegen den LKW drückt, ist genauso groß wie die Kraft, mit der der LKW gegen das Auto drückt.
 - ☐ Die Stärke der Kraft, mit der das Auto gegen den LKW drückt, ist kleiner als die Kraft, mit der der LKW gegen das Auto drückt.
 - ☐ Die Stärke der Kraft, mit der das Auto gegen den LKW drückt, ist größer als die Kraft, mit der der LKW gegen das Auto drückt.
 - ☐ Der Motor des Autos läuft; daher übt das Auto eine Kraft aus, während es gegen den LKW drückt. Aber der LKW-Motor läuft nicht; daher kann der LKW nicht gegen das Auto drücken.
 - ☐ Weder der LKW noch das Auto üben eine Kraft aus. Der LKW wird einfach deshalb nach vorne geschoben, weil er dem Auto im Wege steht.

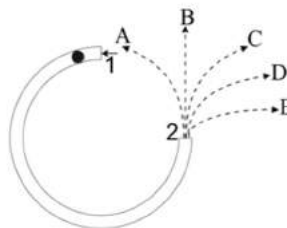
- b) Nachdem der Fahrer des Autos die gewünschte Geschwindigkeit erreicht hat, mit der er den Lastwagen konstant vor sich herschieben will, fährt er mit konstanter Geschwindigkeit weiter. Jetzt gilt:
- ☐ Die Stärke der Kraft, mit der das Auto gegen den LKW drückt, ist genauso groß wie die Kraft, mit der der LKW gegen das Auto drückt.
 - ☐ Die Stärke der Kraft, mit der das Auto gegen den LKW drückt, ist kleiner als die Kraft, mit der der LKW gegen das Auto drückt.
 - ☐ Die Stärke der Kraft, mit der das Auto gegen den LKW drückt, ist größer als die Kraft, mit der der LKW gegen das Auto drückt.
 - ☐ Der Motor des Autos läuft; daher übt das Auto eine Kraft aus, während es gegen den LKW drückt. Aber der LKW-Motor läuft nicht; daher kann der LKW nicht gegen das Auto drücken.
 - ☐ Weder der LKW noch das Auto üben eine Kraft aus. Der LKW wird einfach deshalb nach vorne geschoben, weil er dem Auto im Wege steht.

Abb. 103: Frage 13 aus dem Testinstrument.

Frage 7

Die Abbildung zeigt einen von oben betrachteten kreisförmigen Kanal, der in der horizontalen Ebene fest auf dem Tisch verankert ist. Ein Ball tritt bei "1" in den Kanal ein und verlässt ihn bei "2".

- a) Welche der in der Abbildung gezeigten Bahnkurven beschreibt am besten den Weg des Balls, wenn er den Kanal bei "2" verlässt und weiter über den Tisch rollt?



- ☐ A
 - ☐ B
 - ☐ C
 - ☐ D
 - ☐ E
- b) Welche Aussage erklärt deine Antwort am besten?
- ☐ Die Zentrifugalkraft drückt den Ball nach außen.
 - ☐ Die Kugel bewegt sich weiter im Kreis, da sie sich vorher so bewegt hat.
 - ☐ Es wirkt keine Kraft auf die Kugel, also muss sie geradeaus rollen.
 - ☐ Es wirkt eine Kraft in Bewegungsrichtung.

Abb. 104: Frage 7 aus dem Testinstrument.

Neben den langen Antworten, die ein gutes Leseverständnis erfordern, ist auch das physikalische Konzept erfahrungsgemäß schwierig. Die Lehrkräfte selbst beantworten das Item 13a nur zu 55 % korrekt (siehe Kapitel 10.6.3). Das am zweitleichtesten bewertete Item ist in Abb. 104 zu sehen (Frage 8, welche am leichtesten bewertet wurde, wird danach diskutiert):

Dieses Item wird auch tatsächlich von vielen Schüler*innen korrekt beantwortet. Interessant ist ein allgemeiner Vergleich der Einschätzungen der Lehrkräfte mit den tatsächlichen Lösungswahrscheinlichkeiten der Items (Abb. 105):

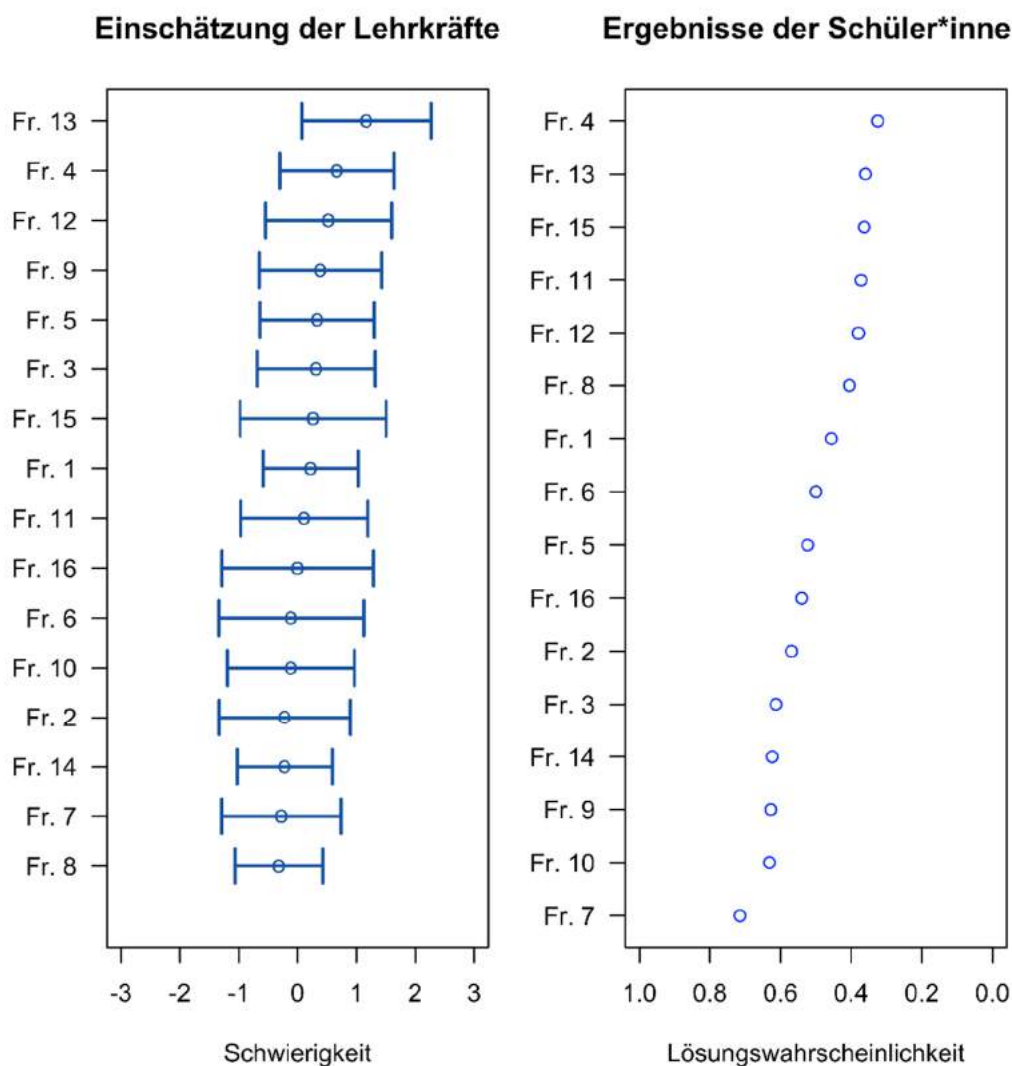


Abb. 105: Plots der Einschätzung der Lehrkräfte mit Standardabweichung im Vergleich zur tatsächlichen Lösungswahrscheinlichkeit der Items.

Die Einschätzung der Lehrkräfte ist als sehr gut einzuordnen. Als schwierig eingeschätzte Items werden von den Schüler*innen in der Tat weniger häufig korrekt beantwortet als die als einfach eingeschätzten Items. Die größten Unterschiede (Differenz zwischen Position in der linken und rechten Grafik) zwischen der Einschätzung der Lehrkräfte und den Ergebnissen der Schüler*innen liegen bei den Fragen 8 (wird zu einfach geschätzt, Abb. 106) und 9 (wird zu schwierig geschätzt, Abb. 107) vor.

Frage 8

Ein Mädchen fährt mit dem Fahrrad auf einem ebenen Weg mit einem konstanten Tempo geradeaus. Warum muss sie dafür in die Pedalen treten?

☐ Wenn auf das Fahrrad keine Kraft wirken würde, würde es langsamer werden und irgendwann stehen bleiben.	☐ Es muss eine nach vorne gerichtete Kraft auf das Fahrrad erzeugt werden, um die Reibungskräfte zu überwinden. Dabei muss die nach vorne gerichtete Kraft größer sein als die Reibungskräfte.	☐ Es muss eine nach vorne gerichtete Kraft auf das Fahrrad erzeugt werden, um die Reibungskräfte auszugleichen. Dabei muss die nach vorne gerichtete Kraft genau so groß sein, wie die Reibungskräfte.

Abb. 106: Frage 8 aus dem Testinstrument.

Frage 8 wird von den Lehrkräften als relativ einfach eingeschätzt, hat aber eine geringe Lösungswahrscheinlichkeit. Dies könnte daran liegen, dass die beschriebene Situation einfach ist, viele Schüler*innen dann aber die Schülervorstellung aktivieren, dass eine Kraft in Bewegungsrichtung wirken muss und damit die falsche Antwort wählen. Interessant ist, dass viele Lehrkräfte diese Frage selbst falsch beantworten, obwohl sie sie als leicht einschätzen (siehe Kapitel 10.6.3). Das könnte auch daran liegen, dass die Antwortmöglichkeiten so verstanden werden, dass es um die Kraft geht, die das Mädchen ausübt. Die erste Antwortmöglichkeit wäre bspw. richtig, wenn es um die Kraft vom Mädchen gehen würde und nicht generell um die wirkende Kraft.

Frage 9

Ein Junge wirft eine Stahlkugel schräg nach oben. Welche der folgenden Kräfte wirken während der Flugphase auf die Kugel?

A: Eine stetig abnehmende Kraft nach oben.		
B: Eine stetig zunehmende Gewichtskraft nach unten.		
C: Eine Kraft, die in Bewegungsrichtung der Kugel zeigt.		
D: Eine der Bewegungsrichtung entgegenwirkende Luftreibungskraft.		
E: Eine konstante nach unten gerichtete Gewichtskraft.		
☐ A und B	☐ Nur C	☐ D und E
☐ C und E	☐ Es wirken keine Kräfte. Die Kugel fällt auf die Erde zurück, weil das ihrem natürlichen Verhalten entspricht.	

Abb. 107: Frage 9 aus dem Testinstrument.

Frage 9 wird wiederum schwieriger eingeschätzt als sie ist. Über 60 % der Schüler*innen konnten diese Frage im Nachtest korrekt beantworten.

10.6.3 Konzeptverständnis

Bereits in Kapitel 9.3.8 wurde diskutiert, dass das Fachwissen bzw. Konzeptverständnis der Lehrkräfte in die Mehrebenenanalyse auf dem Klassenlevel eingefügt wurde. Dieses hatte keinen signifikanten Einfluss auf das Nachtestergebnis der Schüler*innen. Eine genauere Analyse der Ergebnisse der Lehrkräfte im Test bietet sich dennoch an – insbesondere da es wenig Ergebnisse zum Konzeptverständnis bei Lehrkräften gibt. Diese liegen von 17 von den 20 Lehrkräften vor. Betrachtet man alle Items, die auch im Schülerfragebogen enthalten sind, ergibt im Boxplot folgendes Bild (Abb. 108).

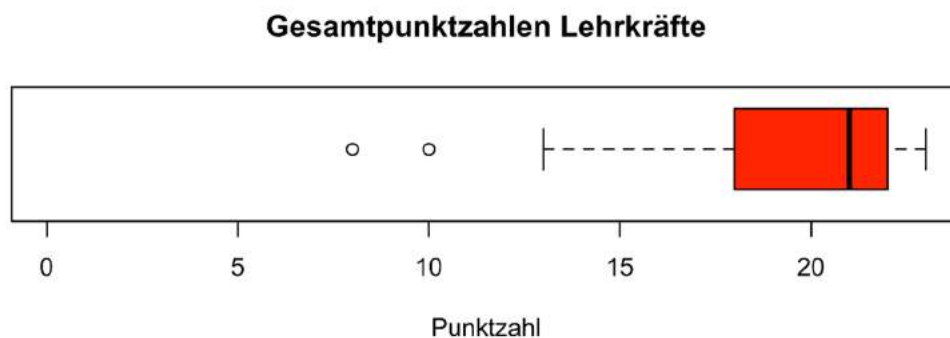


Abb. 108: Boxplot der Punktzahlen (max. 23) der Lehrkräfte im Fragebogen zur Mechanik.

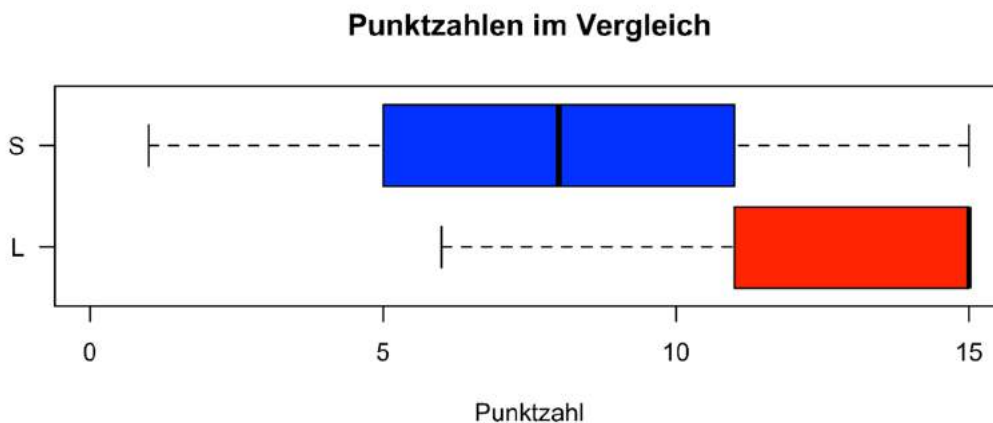


Abb. 109: Gesamtpunktzahlen im Bereich der ersten beiden Newtonschen Gesetze (max. 15) der und Schüler*innen („S“, oben) und Lehrkräfte („L“, unten)

Es gibt einige Lehrkräfte, die ebenfalls Schwierigkeiten mit den Items haben. Wenn man die Schwelle von 80 % richtiger Antworten für „Newtonsche Denker“ aus dem FCI (Hestenes et al., 1992) auf diesen Test überträgt, liegt die Schwelle hier bei 19 Punkten. Fünf der 17 Lehrkräfte liegen unterhalb dieser Schwelle. Zwei der Lehrkräfte liegen sogar

unter der Schwelle von 60 %, die als Ziel für die Schüler*innen nach dem Unterricht angesehen wird („Eintrittsschwelle in die Newtonsche Physik“ nach Hestenes und Halloun (1995)). Alternativ können die Items ausgewertet werden, die auch in der Auswertung der Schüler*innen herangezogen wurden (also alle 15 Items aus dem Bereich der ersten beiden Newtonschen Gesetze, Abb. 109).

Für diesen Fall liegen sechs der 17 Lehrkräfte unterhalb der Schwelle zu „Newtonschen Denkern“. Genauere Auskunft gibt eine Analyse der einzelnen Items (Abb. 110).

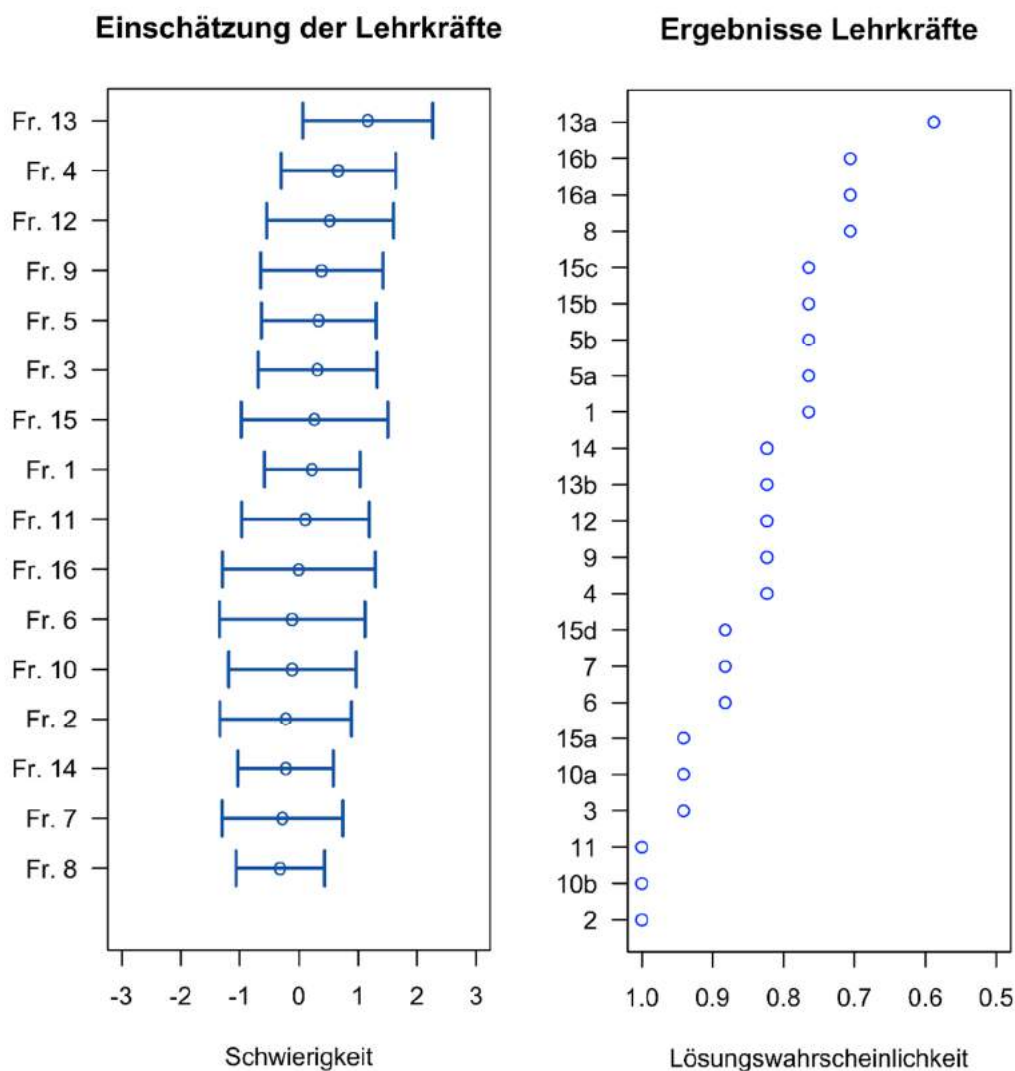


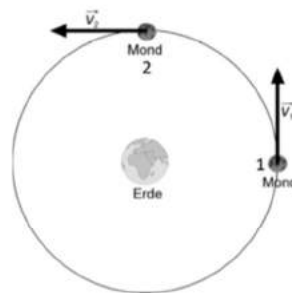
Abb. 110: Vergleich der eingeschätzten Schwierigkeit der Fragen mit Standardabweichung und der tatsächlichen Schwierigkeit der Items bei den Lehrkräften.

Wie schon angedeutet, ist Item 13a das schwierigste. Dieses wird auch korrekterweise als schwierig eingeschätzt. Darauf folgen die Items 16a und 16b, wobei Frage 16 (Abb. 111) insgesamt von den Lehrkräften nicht als schwierig eingeschätzt wird.

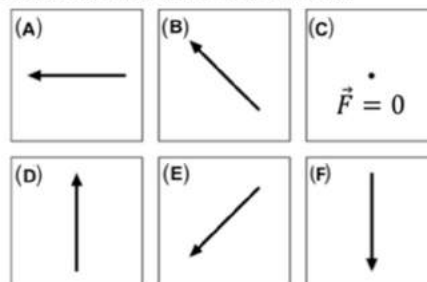
Eine noch größere Diskrepanz gibt es zwischen dem ebenfalls bereits besprochenen Item 8. Insgesamt zeigt die Erhebung unter Lehrkräften, dass sich die Alltagsvorstellungen, die ein angemessenes Verständnis der Newtonschen Mechanik erschweren, auch bei Physiklehrkräften weiterhin anzutreffen sind.

Frage 16

Das Bild stellt die Bewegung des Mondes um die Erde dar. Der Mond ist zu den Zeitpunkten "1" und "2" dargestellt.



Wähle in den folgenden beiden Fragen aus den gegebenen Möglichkeiten A-F jeweils einen Pfeil aus, der die **Kraft**, die auf den Mond wirkt, zu dem Zeitpunkt am besten darstellt.



a) Zeitpunkt 1

- ☐ A
☐ D

- ☐ B
☐ E

- ☐ C
☐ F

b) Zeitpunkt 2

- ☐ A
☐ D

- ☐ B
☐ E

- ☐ C
☐ F

Abb. 111: Frage 16 aus dem Testinstrument.

10.7 Zusammenfassung zum quantitativen Teil

Aus den quantitativen Daten des Fragebogens lassen sich einige Erkenntnisse ableiten. Es gibt im gesamten Konzeptverständnis keinen nennenswerten Unterschied zwischen den beiden Gruppen im Lernzuwachs in der Intervention. Lediglich in der Skala „1. Newtonsches Gesetz“ gibt es einen Unterschied zwischen den Gruppen. Hier lernt die Gruppe der Modellbildung mehr dazu. Dies kann durch eine Analyse der Schülervorstellungen

begründet werden. Der Anteil an Schüler*innen mit der Schülervorstellung, dass eine Kraft in Bewegungsrichtung wirken muss, wird in der Gruppe der Modellbildung stärker reduziert. Da diese Schülervorstellung in Fragen zum ersten Newtonschen Gesetz zu einer falschen Antwort führt, kann der höhere Lernzuwachs in der Gruppe der Modellbildung auf die stärkere Reduktion der Schülervorstellung zurückgeführt werden. Einen Einblick, warum die Schülervorstellung in der Gruppe der Modellbildung stärker reduziert wird, soll eine Analyse der Bildschirmvideos im nächsten Kapitel beleuchten. Die Gruppen unterscheiden sich nicht in Skalen mit unterschiedlicher Argumentationsrichtung, obwohl sie mehrheitlich gegenteilige Argumentationsrichtungen nutzen. Einfluss auf das Nachtestergebnis haben nach der Mehrebenenanalyse das Vortestergebnis, das Interesse an theoretischen Zusammenhängen, die negative kognitive Belastung und die Physiknote. Das Interesse an theoretischen Zusammenhängen ist dabei unabhängig von der Intervention lernförderlich, während das Interesse am Experimentieren keinen Einfluss hat.

In affektiven Variablen unterscheiden sich die Gruppen durchaus. Die Software „measure dynamics“ wird besser bewertet als die Software „Newton-II“. Die intrinsische kognitive Belastung wird in der Intervention der Modellbildung als höher eingeschätzt. In der Gruppe der Modellbildung ändert sich die Ansicht zu Modellen in der Physik zum Positiven, obwohl die Eigenschaften von Modellen in der Intervention nicht explizit thematisiert wurden, während diese in der Gruppe der Videoanalyse unverändert bleiben. Auch Lehrkräfte lassen sich durch die im Test vorhandenen Disktraktoren zu falschen Antworten verleiten. Sie sind aber gut in der Lage, die Schwierigkeiten der Items einzuschätzen.

11 Ergebnisse des qualitativen Teils der Studie

Um einen Einblick zu bekommen, wie mit den jeweiligen Programmen gearbeitet wird, und die Resultate der quantitativen Erhebung besser interpretieren zu können, wurden zudem die Bildschirmvideos ausgewertet. Dabei werden, wie in Kapitel 8.6.2.2 erläutert, verschiedene Inhaltsdimensionen unterschieden. In der Gruppe der Modellbildung wurden dazu elf Videos ausgewertet, in denen 22 Schüler*innen zu hören sind, während in der Gruppe der Videoanalyse 23 Schüler*innen analysiert wurden, die sich in zehn Gruppen zusammensetzen. In der Gruppe der Videoanalyse wurde Videomaterial mit einer Länge von insgesamt 10 Stunden und 12 Minuten (0:55:38 pro Video) analysiert. In der Gruppe der Videoanalyse hatten die Videos eine Länge von insgesamt 8 Stunden und 36 Minuten (0:51:36 pro Video).

11.1 Schwierigkeiten

Die Schüler*innen arbeiten in den Arbeitsphasen nach einer kurzen Einführung in das Programm selbstständig mit der Software. Damit diese Arbeit im Physikunterricht zielführend ist, müssen Schwierigkeiten, die während der Arbeit auftreten, analysiert und, wenn möglich, abgestellt bzw. verhindert werden. Dabei soll zwischen fachlichen und softwarebedingten Schwierigkeiten unterschieden werden. Während erste vor allem für Lehrkräfte relevant sind, um den Einsatz der diskutierten Software im Dynamikunterricht zu planen, sind Schwierigkeiten mit der Software zusätzlich für Entwickler*innen von Interesse. Letztere können möglicherweise durch eine entsprechende Hilfestellung (der Lehrkraft) oder eine Weiterentwicklung der Programme behoben werden. Allgemein werden in diesem Kapitel nur Schwierigkeiten diskutiert, die nicht der Erreichung der Lernziele dienen bzw. diese eher verhindern. Wenn es also inhaltliche Schwierigkeiten bei der Interpretation der Diagramme gibt, weswegen die Zusammenhänge der Newtonschen Dynamik möglicherweise schlechter erlernt werden, ist dies eine Schwierigkeit, die verhindert werden sollte. Wenn es wiederum Verständnisschwierigkeiten bei der Diskussion von Kräften gibt, werden die in diesem Kapitel nicht erfasst, da diese Auseinandersetzung mit dem zentralen Lernobjekt als lernförderlich eingestuft wird. Nicht immer sind diese unterschiedlichen Arten von Schwierigkeiten jedoch eindeutig voneinander zu trennen.

11.1.1 Inhaltliche Schwierigkeiten

Die häufigste inhaltliche Schwierigkeit lag in der Interpretation der Diagramme. Dass dies für Schüler*innen generell eine große Schwierigkeit ist, ist bekannt (Beichner, 1990).

Da beide Programme viel mit Diagrammen arbeiten („Newton-II“ vollständig, „measure dynamics“ neben dynamisch ikonischen Repräsentationen teilweise), ist ein mangelndes Verständnis dieser ein Problem für den Einsatz in der Dynamik.

In der Gruppe der Videoanalyse konnten sechs Stellen in fünf unterschiedlichen Videos identifiziert werden, in denen ein mangelndes Verständnis zu einer Behinderung des Arbeitsablaufes geführt haben. Diese Stellen nahmen insgesamt knapp über acht Minuten in Anspruch. Negative Auswirkungen auf die weitere Arbeitsphase über diese acht Minuten hinaus können ebenfalls identifiziert werden, sind aber schwierig zu quantifizieren. In der Gruppe der Modellbildung wurden zehn Stellen in fünf verschiedenen Videos identifiziert, welche insgesamt über 17 Minuten dauerten. Auch hier lassen sich die Auswirkungen danach weiter beobachten. Konkret äußern sich die Schwierigkeiten darin, dass sinnvolle Modelle verworfen oder ungünstige Modelle angenommen werden, da die Daten, die in einem Diagramm dargestellt sind, falsch interpretiert werden. Den Schüler*innen, bei denen die entsprechenden Schwierigkeiten identifiziert wurden, ist selten bewusst, welche Größen sinnvoll gegeneinander aufgetragen werden können. Gesucht wird außerdem häufig das Diagramm für den Weg in Abhängigkeit der Zeit. Die zugehörige Größe s gibt es bei „measure dynamics“ nie und bei „Newton-II“ nur im eindimensionalen Fall. Besonders problematisch ist in beiden Gruppen auch, dass zwischen der horizontalen und der vertikalen Richtung unterschieden werden muss. In beiden Programmen wird standardmäßig zu Beginn der Ort in x-Richtung in Abhängigkeit der Zeit angezeigt, was in dem betrachteten Versuch ein linear steigender Graph ist. Häufig wird in diesem Diagramm bereits eine Parabel bzw. ein Bogen erwartet, der aber nur zu sehen ist, wenn der Ort in y-Richtung über der Zeit oder dem Ort in x-Richtung aufgetragen ist. Eine mögliche Lösung wäre neben einer vorherigen Stärkung der Diagrammkompetenz, dass zu Beginn von der Software immer das x-y-Diagramm angezeigt wird, da dieses von Schüler*innen am häufigsten erwartet wird. Nach bewusster Einstellung der Achsen (zum Beispiel Zeit auf der horizontalen Achse) könnte das Missverstehen der Diagramme geringer sein. Dies deutet sich in den Aufzeichnungen an. Aber auch die in die Koordinatenachsen aufgeteilten Geschwindigkeiten verursachen bei den Proband*innen Schwierigkeiten.

Es zeigt sich sowohl in der Häufigkeit als auch in der Dauer, dass das Problem der Diagramminterpretation in der Gruppe der Videoanalyse geringer, wenngleich immer noch zu groß ist. Die Anzahl an Zweiergruppen, in denen die Probleme auftraten, waren aber in beiden Kohorten gleichverteilt. Die Unterschiede in der Häufigkeit können verschiedene Gründe haben. In der Gruppe der Modellbildung findet die vollständige Evaluation diagrammbasiert statt, während in der Videoanalyse auch andere Darstellungen (vor allem Pfeile) genutzt werden. Dies führt aufgrund der geringeren Beschäftigungszeit mit

Diagrammen in der Videoanalyse erwartbar zu selteneren Problemen. Zusätzlich können die anderen Darstellungsformen (inklusive des Videos selbst) dazu führen, dass die Diagramme besser verstanden werden. Es ist bekannt, dass sich die Videoanalyse gut zum Verbessern des Diagrammverständnisses eignet (z. B. Beichner, 1996). In der Modellbildung muss diskutiert werden, ob eine zusätzliche Form der Darstellung der modellierten Daten didaktisch sinnvoll und notwendig ist. Es gibt bereits Software, die auch Animationen ausgibt, was sich allerdings wieder negativ auf die Komplexität und Bedienbarkeit der Software auswirkt (Weber & Wilhelm, 2020a). Eine andere Möglichkeit ist es, das erstellte Modell in einem Video mitsamt der realen Bewegung anzuzeigen, wozu man aber ein Video benötigt. Diese Funktion ermöglicht beispielsweise die Software „Tracker“. Da Diagrammkompetenz aber auch ohne die Nutzung von Modellierungssoftware ein Ziel des naturwissenschaftlichen Unterrichts ist, sollte für den erfolgreichen Einsatz dieser Software in jedem Fall darauf geachtet werden, dass bereits zuvor ein sicherer Umgang mit Diagrammen vorhanden ist.

Weitere inhaltliche Schwierigkeiten zeigten sich vor allem in der Gruppe der Modellbildung, was nochmal deren höhere inhaltliche Komplexität zeigt. Hauptsächlich zu nennen sind hier Probleme bei der Formulierung und Eingabe der Kräfte, was an neun Stellen in insgesamt fünf Gruppen identifiziert werden konnte und insgesamt 31 Minuten in Anspruch genommen hat. Teilweise hängen die Probleme mit der Formulierung und Eingabe der Reibungskräfte zusammen, da hier Reibungskräfte in beiden Dimensionen definiert werden müssen und es schwierig ist, die Richtung der Luftreibung über trigonometrische Zusammenhänge herzuleiten. Hierfür gab es in einem Zusatzheft Hilfestellungen, die von den Schüler*innen genutzt werden konnten. Problematischer ist, dass auch bei der Formulierung und Eingabe der Gewichtskraft der fliegenden Kugel häufig Schwierigkeiten bestanden, obwohl die Gewichtskraft aus dem Unterricht bereits bekannt sein sollte und in den ersten beiden Versuchen ebenfalls modelliert werden musste. Häufig wird statt $m \cdot g$ für die Gewichtskraft nur z. B. nur g eingegeben. Es scheint also keine ausreichend klar abgegrenzte Vorstellung von Kraft zu bestehen, da aus den Gesprächen hervorgeht, dass die Masse nicht nur vergessen wurde, sondern dass der Ortsfaktor häufig als Kraft angesehen wird. Eine inhaltliche Diskussion dieser Vorstellung wird sich in Kapitel 11.3 anschließen. Teilweise versuchten die Proband*innen auch die allgemeine Definition $F = m \cdot a$ zu nutzen, um die Kraft zu modellieren. Dies spricht dafür, dass die Logik der Software, die die Definition $a = F/m$ nutzt, um aus den wirkenden Kräften und der Masse die Beschleunigung zu berechnen und daher die einzelnen Kräfte benötigt, die auf das Objekt wirken, noch nicht vollständig verstanden wurde. Für die Beschleunigung wurde dann eine spezielle Berechnungsformel wie $a = v/t$ genutzt. Auch wenn dieser Fall selten eintrat, sollte dies von der Lehrkraft berücksichtigt und durch eine erneute

Besprechung der Funktionsweise der Berechnung ausgeräumt werden. Obwohl die Formulierung von Kräften auch als Lernziel der Intervention bezeichnet werden könnte, können Fehlvorstellungen in diesem Bereich die Modellierung behindern.

Die Nutzung von speziellen Bewegungsformeln ist ein weiteres Problem, das mehrfach zu beobachten war. Dies wurde an sechs Stellen in vier Gruppen identifiziert und nahm 13 Minuten in Anspruch. Wie in Kapitel 5.4 thematisiert, soll die Modellbildung gerade dazu führen, dass erkannt wird, welche Zusammenhänge zwischen den Größen fundamental sind und welche Gleichungen lediglich spezielle Berechnungsformeln für Einzelfälle unter bestimmten Bedingungen sind. Man merkt, dass viele Schüler*innen im Unterricht verschiedene Dinge mit diesen Formeln berechnet haben und dies dementsprechend die erste Überlegung ist, wenn eine Bewegung betrachtet wird. Teilweise wird auch in den Unterlagen aus dem Unterricht nachgeschaut, welche Bewegungsformeln für den waagerechten Wurf im Unterricht genutzt wurden. Dies ist der Fall, obwohl das Vorgehen (also die Modellierung von Kräften, Masse und Anfangsbedingung und das Ausrechnenlassen der Bewegung durch die Software) zuvor bereits zwei Mal durchgeführt wurde. In den meisten Fällen führt entweder der Misserfolg in der Nutzung der Bewegungsformeln oder ein Hinweis des Dozenten dazu, dass die Gruppen im Anschluss korrekt die Kräfte modellieren. Diese Beobachtung spricht dafür, dass Modellbildung im Unterricht eingesetzt werden sollte, da so ein Fokus auf die wichtigen Zusammenhänge gelegt werden kann. Diese Analyse kann aber keine Aussage dazu treffen, ob sich dies durch die Nutzung der Modellbildung tatsächlich verbessert hat und die Proband*innen bei einer erneuten Nutzung weniger auf spezielle Bewegungsformeln zurückgreifen würden. Für ein unzureichendes Verständnis der Abläufe und der Ziele der mathematischen Modellbildung und damit auch der Newtonschen Mechanik allgemein spricht auch, dass eine Gruppe nach der erfolglosen Anwendung von Bewegungsformeln diskutiert, wie man dem Programm sagt, dass die Kugel am Anfang hoch und dann runter geht. Die Bewegung ist also bei dieser Gruppe noch nicht als Folge der wirkenden Kräfte in Kombination mit den Anfangsbedingungen internalisiert.

Eine weitere identifizierte Schwierigkeit ist eine unnötige Anpassung der erstellten Modelle nach einer theoriebasierten Evaluation, welche in drei Fällen beobachtet wurde. Das heißt, dass sinnvolle Modelle nach Betrachtung der Diagramme negativ geändert wurden. Diese Schwierigkeit könnte auch an mangelndem Diagrammverständnis liegen, wobei dieser Zusammenhang nicht in jeder Situation deutlich wurde. Teilweise wurde von den Schüler*innen auch versucht, das „Schema“ aus der Aufgabe zuvor (z. B. das Modellieren von Reibungskräften) auf diese Aufgabe zu übertragen, ohne dass überhaupt evaluiert wird, ob das Modell möglicherweise schon ausreichend ist. In diesen Fällen müsste noch

deutlicher gemacht werden, dass die Lernenden das Modell zunächst evaluieren sollen und nur bei signifikanten Unterschieden zur Realität eine Änderung vornehmen. In weiteren Einzelfällen, die im Vergleich zu den anderen berichteten Schwierigkeiten aber eine geringe Relevanz haben, wurden die Koordinatenrichtungen x und y verwechselt, was zu einer Verzögerung der Arbeit geführt hat. Eine Gruppe hat außerdem versucht, die für die Modellierung vermeintlich notwendige Geschwindigkeit manuell auszurechnen.

Die hier beschriebene Analyse ist nicht vollständig objektivierbar und es ist nicht immer eindeutig, was als Schwierigkeit und was als produktive Auseinandersetzung mit dem Lerngegenstand anzusehen ist. Dennoch ermöglicht die Analyse Lehrenden eine Vorbereitung auf entstehende Probleme bei der Auseinandersetzung mit Bewegungsphänomenen. Auch eine Unterscheidung zwischen inhaltlich begründeten Schwierigkeiten und Schwierigkeiten mit der Software ist nicht immer eindeutig, da sich inhaltliche Schwächen auch negativ auf den Umgang mit der Software auswirken können. Dennoch werden nun die Schwierigkeiten, die primär bzw. unmittelbar mit der Software zusammenhängen, diskutiert.

11.1.2 Schwierigkeiten mit der Software

In Summe wurden in den beiden Gruppen 13 verschiedene Schwierigkeiten identifiziert, die sich an 56 Stellen bei 20 Gruppen zeigten und etwa 85 Minuten in Anspruch genommen haben. Da die Schwierigkeiten unmittelbar von der Software abhängen, werden diese in unterschiedlichen Kapiteln dargestellt. Bei „Newton-II“ wurden 23 Schwierigkeiten in zehn Gruppen beobachtet, die etwa 46 Minuten in Anspruch nahmen, während bei „measure dynamics“ 34 Stellen in ebenfalls zehn Gruppen identifiziert wurden, die etwa 38 Minuten kosteten (die Abweichung zu 85 Minuten ergibt sich durch Rundungen).

Newton-II

Bei der Software „Newton-II“ (Version 3.04) war die am häufigsten zu beobachtende Schwierigkeit, dass die modellierten Reibungskräfte in das Modellierungsfenster geschrieben, aber nicht in das Modell einbezogen wurden. In der Praxis sah dies oft folgendermaßen aus (Abb. 112):

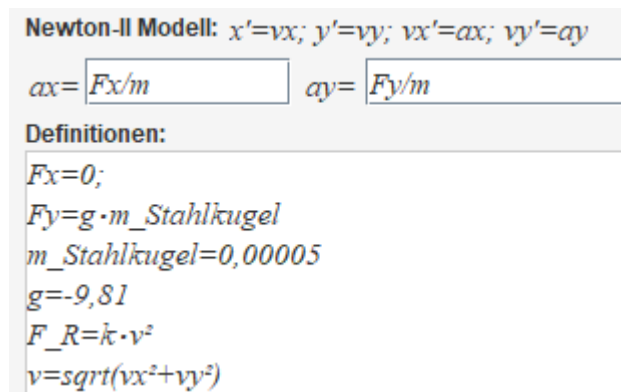


Abb. 112: Fehler bei der Modellierung von Reibungskräften.

Die Zweiergruppe, von der das obige Bild stammt, hat mit $F_R = k \cdot v^2$ einen Ansatz für eine Reibungskraft formuliert, diese aber lediglich in das Fenster der Definitionen geschrieben, ohne die Kraft mit dem Modell zu verknüpfen. Da das Modell aus den Kräften in die beiden Koordinatenrichtungen die Beschleunigung bestimmt und die Reibungskraft nicht in einer der beiden ersten Definitionen eingefügt wurde, hat sie in diesem Fall keinen Einfluss auf die Bewegung. Bei dieser Schwierigkeit kann auch diskutiert werden, ob sie den inhaltlichen Schwierigkeiten zuzuordnen ist, da nicht abschließend geklärt werden kann, ob die Probleme mit einem mangelnden Verständnis der physikalischen Zusammenhänge oder mit einer unzureichenden Kenntnis der Funktionsweise der Software einhergeht. Eine Abmilderung dieses Problems könnte möglicherweise dadurch erreicht werden, dass die Definitionen für die Kräfte besonders hinterlegt oder von den anderen Definitionen getrennt wären. Dies hat aber den großen Nachteil, dass die Software dadurch auch an Flexibilität verliert. Ein sinnvoller Umgang damit wäre, dies im Plenum aktiv zu thematisieren und die Problematik damit abzumildern.

Weiterhin gab es mehrfach Probleme mit dem Einstellen der Diagramme, wo die Schüler*innen die anzuzeigende Größe explizit eingeben mussten. Zum Beispiel wurden die Variablen im eindimensionalen Fall s , v , und a genannt, während im zweidimensionalen Fall nur x und y , vx und vy und ax und ay zur Verfügung standen. Obwohl dies im Plenum zuvor angekündigt wurde, versuchten viele Schüler*innen die Variablen des eindimensionalen Falls einzutragen und erhielten so keine Diagramme. Beheben ließe sich dies entweder durch eine Umbenennung des eindimensionalen Falls in x , vx , ax oder, noch besser, durch die Erstellung eines Dropdown-Menüs, in dem die Möglichkeiten direkt angeklickt werden können. Vereinzelt gab es auch Schwierigkeiten mit dem Zoom und der Skalierung der Achsen.

Auch das Einlesen der externen Messdaten bereitete einigen Schüler*innen auch im dritten Durchgang noch Schwierigkeiten. An dieser Stelle könnte die Benutzerfreundlichkeit dadurch verbessert werden, dass die Daten auch über die rechte Maustaste und die Entfernen-Taste gelöscht werden können und nicht nur über die Schaltfläche in der Software. In einigen Fällen waren auch Probleme mit der Definition der verwendeten Größen zu erkennen. So wurde bspw. g genutzt, ohne dem Programm mitzuteilen, welchen Wert g hat. Ähnlich sieht es mit dem Tempo v aus, welches im zweidimensionalen Fall erstmal nicht definiert ist und vom Nutzer als $v = \sqrt{v_x^2 + v_y^2}$ in das Feld der Definitionen geschrieben werden muss, falls es benötigt wird. In beiden Fällen wäre es denkbar, die Definition bzw. den Zahlenwert standardmäßig in der Software zu hinterlegen, wobei die selbstständige Definition von genutzten Größen ein eigenes Lernziel der Modellierung ist und dementsprechend von den Lernenden auch selbst durchgeführt werden sollte. Auch die Reibungskonstante wurde von verschiedenen Gruppen nicht definiert oder ein ungünstiger Bereich für den Schieberegler der Konstante festgelegt.

Weitere Probleme, die aber nur jeweils einmal beobachtet wurden, sind Schwierigkeiten bei der Eingabe der Anfangsbedingungen und Fehler bei der Eingabe einer bedingten Variablen (die hier ohnehin nicht nötig gewesen wäre). Eine Gruppe war mit dem Modell zufrieden, ohne dass sie überhaupt Kräfte modelliert hat, da sie sich das $x(t)$ -Diagramm anschaute und mit den Messdaten verglich. Weil in x -Richtung keine Kräfte wirken, passt das Modell hier. Eine Betrachtung der anderen Koordinatenrichtung erfolgte nicht.

Die am häufigsten auftretenden Probleme lassen sich vermutlich durch kleine Änderungen an der Software beheben. Dies ist zum Zeitpunkt des Verfassens dieser Arbeit teilweise auch schon erfolgt, da die aktuelle Version (Version 3.6.1) der Software bereits einige Änderungen im Vergleich zu der in der Studie genutzten Software aufweist. So wurde das vorgeschlagene Dropdown-Menü z. B. bereits eingepflegt.

measure dynamics

Auch wenn die Software „measure dynamics“ von den Schüler*innen besser bewertet wurde als „Newton- II“ (siehe Kapitel 10.2), gab es auch hier einige Schwierigkeiten bei der Benutzung. Das häufigste Problem, was an elf Stellen in neun Gruppen identifiziert werden konnte, hängt mit der Skalierung für die Videoanalyse zusammen. Hier ist die Problematik aber eher in dem Video selbst als in der Software oder den Schüler*innen begründet. In dem Video ist ein Meterstab zu sehen, der allerdings teilweise verdeckt ist (Abb. 113).

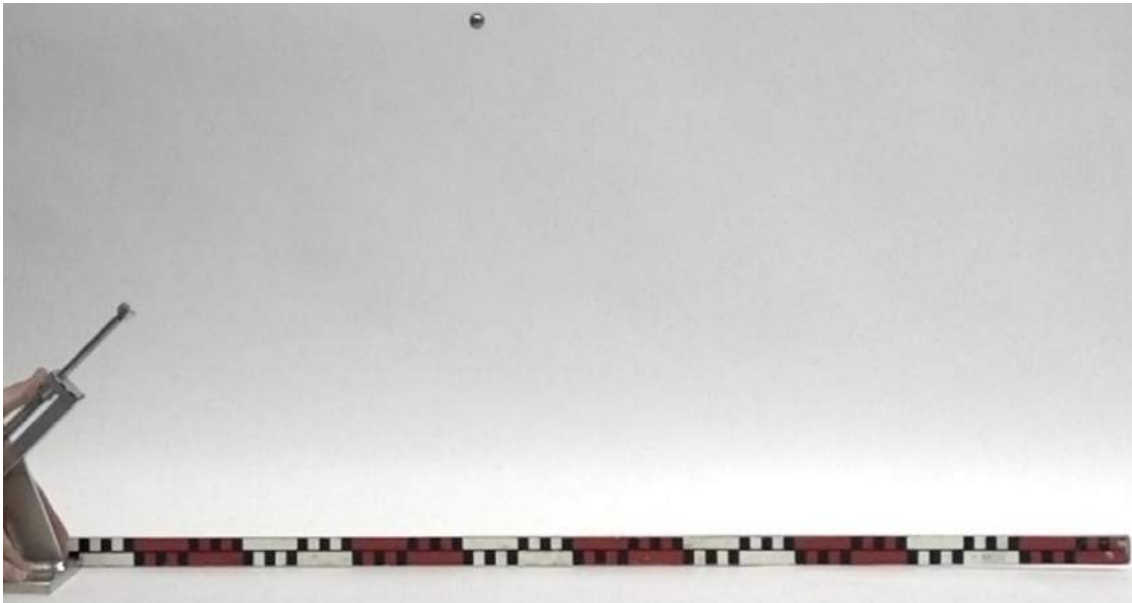


Abb. 113: Screenshot des Videos von dem Flug der Stahlkugel mit nicht vollständig sichtbarem Meterstab.

Viele Proband*innen versuchten, das Ende des Stabs auf der linken Seite abzuschätzen, statt die Skalierung bis zur letzten sichtbaren Grenze durchzuführen und dann 0,9 m statt 1 m für die entsprechende Länge anzugeben. Das Problem lässt sich beheben, indem man darauf achtet, dass der Stab vollständig sichtbar ist oder noch deutlicher macht, dass beliebige Längen genommen werden können und jeder farbige Abschnitt auf dem Stab eine Länge von 10 cm hat. Dies wurde zwar zu Beginn gesagt, kam aber nicht ausreichend bei den Schüler*innen an. Außerdem könnte dies softwareseitig verbessert werden, in dem nicht bereits 1m in dem entsprechenden Feld eingetragen ist, sondern dies z. B. als Vorschlag darunter steht. Wenn Schüler*innen die Länge selbst eintragen müssen, dürfte klarer werden, dass beliebige Längen (und nicht nur ein Meter) möglich sind.

Häufiger als bei „Newton-II“ kam es zu Problemen bei der Einstellung der Diagramme. Bei „measure dynamics“ besteht zwar nicht das Problem, dass beliebige Größen für die Achsen eingetragen werden können, die möglicherweise nicht definiert sind, eine Umstellung der Achsen ist aber umständlich. Dazu muss ein Doppelklick auf das Diagramm erfolgen, woraufhin man die alten Daten erst löschen muss. Danach kann man einstellen, welche Größe man auf der horizontalen und welche auf der vertikalen Achse haben möchte, muss die richtige Tabelle auswählen, die die Daten enthält, und kann diese Daten dann hinzufügen, bevor man die Änderungen bestätigt. Häufig wurden die alten Daten nicht gelöscht oder die neuen nicht hinzugefügt. Teilweise gab es hier auch Probleme, die nicht auf die Software zurückzuführen sind, sondern eher für ein mangelndes Verständnis der vorkommenden Größen sprechen. Dennoch könnte der Ablauf der Umstellung der

Größen für die Diagramme einfacher gestaltet werden. Einige zoomten auch ungewollt, sodass der Graph nicht mehr sichtbar war. Es war für Schüler*innen dann schwierig, die Einstellung zu finden, bei der man dies wieder rückgängig machen konnte. Bei einer Aufgabe war es notwendig, die automatische Skalierung der Achsen umzustellen. Hier hatten mehrere Gruppen Schwierigkeiten. Möglicherweise wären die Probleme geringer, wenn der dargestellte Bereich direkt an den Achsen eingestellt werden könnte. Eine regelmäßige Nutzung der Software würde hier wahrscheinlich ebenfalls zu Verbesserungen führen.

Eine weitere Fehlerquelle, die etwas seltener vorkommt, ist die Videoanalyse selbst. Meist war dabei das Problem, dass die automatische Analyse die Kugel gar nicht oder nicht in jedem Bild erkannt hat. Dies ließe sich durch ein besseres Video lösen. Die Software wählt nach einer Voruntersuchung selbstständig den Algorithmus aus, der zur Verfolgung des Objekts genutzt wird (Bewegungserkennung, Mustererkennung, Farberkennung). Bei demselben Video kommt es dabei teilweise zu unterschiedlichen Ergebnissen. Auch die Schrittweite, die Position, die Vergrößerung und der Frame, in denen das Objekt angeklickt wird, spielen eine Rolle dafür, ob die automatische Analyse gelingt. Selbst mit einem guten Video kann nicht garantiert werden, dass die Analyse immer problemlos funktioniert. Eine sinnvolle Funktion der Software ist, dass man einzelne Punkte, die in der automatischen Analyse nicht richtig gesetzt wurden, manuell ergänzen kann. Dies ist in jedem Fall eine Technik, die den Schüler*innen gezeigt werden könnte, um das genannte Problem zu verringern. Hilfreich ist in diesem konkreten Fall auch das Anklicken des Objekts in einem späteren Frame.

In manchen Fällen (an vier verschiedenen Stellen in zwei Gruppen) wurde außerdem versucht, die Daten anzuzeigen, ohne dass vorher eine Videoanalyse durchgeführt wurde. Dies ließe sich softwareseitig beheben, indem den Nutzer*innen eine Meldung angezeigt wird, dass noch keine Daten erhoben wurden, falls diese auf „Anzeige“ klicken, ohne vorher eine Videoanalyse durchgeführt zu haben. In zwei anderen Fällen missglückte die Videoanalyse so, dass sich das Objekt gar nicht zu bewegen schien, was in der Anzeige der Daten nicht gemerkt wurde. Dieses Problem ließe sich durch die oben beschriebene Meldung nicht lösen, da in einem solchen Fall Daten vorhanden sind – diese allerdings unbrauchbar sind. Nur einmal gab es Probleme beim Einzeichnen von Pfeilen in das Video.

Vergleich

Vergleicht man die beiden Programme wird deutlich, dass bei „Newton-II“ mehr unterschiedliche Problemarten auftreten, die auch insgesamt mehr Zeit in Anspruch nahmen, wenngleich bei der Videoanalyse an insgesamt mehr Stellen diese Probleme identifiziert werden konnten. Zwei der drei größten Probleme bei der Nutzung von „measure dynamics“ ließen sich aber durch ein besseres Video des Experiments zumindest teilweise lösen. Da die Videos bei den anderen Experimenten in dieser Hinsicht besser waren, sollten diese Probleme dort geringer sein. Auch wenn dies nicht durch eine Datenanalyse bestätigt werden kann, da die Bildschirmvideos nur für ein Experiment aufgezeichnet wurden, stützt der subjektive Eindruck bei der teilnehmenden Beobachtung während der Intervention diese These. Berücksichtigt man diese Tatsache, decken sich die Häufigkeiten der Schwierigkeiten weitestgehend mit der Einschätzung der Software durch die Schüler*innen. Nicht alle Schwierigkeiten lassen sich unmittelbar der Software zuordnen. Es kann argumentiert werden, dass der Lerngegenstand der mathematischen Modellbildung komplexer ist, wodurch auch die Einschätzung zur Software beeinflusst sein könnte.

11.2 Argumentationsrichtung

Es gibt in der Literatur Hinweise darauf, dass es für Schüler*innen einen Unterschied macht, ob von den Kräften argumentativ auf die Eigenschaften der Bewegung geschlossen wird oder andersherum (siehe Just et al., 2021). Zudem wurde in Kapitel 8.4.3 bereits gezeigt, dass die im Testinstrument verwendeten Items mit der gleichen Argumentationsrichtung stärker miteinander korrelieren und damit bei der Faktorenanalyse jeweils auf einen Faktor laden. Das spricht dafür, dass dieses Merkmal einer Testaufgabe für Schüler*innen bei der Beantwortung relevant ist. Auch wenn es zwischen den Ergebnissen der beiden Gruppen in beiden Argumentationsrichtungen keinen Unterschied gab, kann untersucht werden, in welcher Weise sich die Erkenntnisse von Just et al. (2021), die die Effekte in Single-Choice-Testitems untersuchten, auf natürliche Gespräche über physikalische Größen übertragen lassen. Dafür wurden die aufgezeichneten Bildschirmvideos unter diesem Gesichtspunkt ausgewertet. Grundsätzlich wäre es möglich, dass man alle Situationen betrachtet, in denen überhaupt eine argumentative Verbindung von wirkenden Kräften und Bewegungen vollzogen wird. Da die Aufgaben im Arbeitsheft teilweise explizit nach einer solchen Argumentation verlangen und dies die Ergebnisse verzerren würde, werden hier aber nur Aussagen untersucht, die bei der Besprechung der Abschlusstabelle des jeweiligen Versuchs getätigt wurden. Somit ist eine möglichst hohe Vergleichbarkeit gegeben. Dabei ist nicht vorgegeben, wie die Schüler*innen argumentieren

sollen, allerdings ist die Reihenfolge der genannten Größen an die Argumentationslogik der jeweiligen Methode der Intervention angepasst.



Welche Aussagen zur Kugel lassen sich über folgende Größen bei der Bewegung treffen? **(Ohne Luftreibung)**

x-Richtung

Die Summe aller Kräfte...	☐ ist konstant ($\neq 0$). ☐ ist Null. ☐ ändert sich.
Die Beschleunigung...	☐ ist konstant ($\neq 0$). ☐ ist Null. ☐ ändert sich.
Die Geschwindigkeit...	☐ ist konstant ($\neq 0$). ☐ ist Null. ☐ ändert sich.

Abb. 114: Abschlusstabelle der Aufgabe in der Gruppe der Modellbildung.



Welche Aussagen zur Kugel lassen sich über folgende Größen bei der Bewegung treffen? **(Ohne Luftreibung)**

x-Richtung

Die Geschwindigkeit...	☐ ist konstant ($\neq 0$). ☐ ist Null. ☐ ändert sich.
Die Beschleunigung...	☐ ist konstant ($\neq 0$). ☐ ist Null. ☐ ändert sich.
Die Summe aller Kräfte...	☐ ist konstant ($\neq 0$). ☐ ist Null. ☐ ändert sich.

Abb. 115: Abschlusstabelle der Aufgabe in der Gruppe der Videoanalyse.

Während in der Gruppe der Modellbildung dementsprechend erst nach der Summe aller Kräfte gefragt wird (Abb. 114), aus der die Beschleunigung und danach die Geschwindigkeit abgeleitet werden soll, steht in der Gruppe der Videoanalyse die

Geschwindigkeit an erster Stelle, auf die die Beschleunigung und dann die Summe aller Kräfte folgen (Abb. 115).

Da die Interventionen für eine Verbesserung des Konzeptverständnisses im Bereich der Dynamik konzipiert wurden und die Ergebnisse von Just et al. (2021) sich ebenfalls auf den Zusammenhang von Kräften und Bewegung beziehen, werden zunächst alle Aussagen analysiert, die den Zusammenhang zwischen den wirkenden Kräften und irgendeiner Eigenschaft der Bewegung betreffen. Argumentationen zwischen Geschwindigkeit und Beschleunigung werden dementsprechend zunächst nicht berücksichtigt. Tab. 61 zeigt die Häufigkeit der Argumentationen in beiden Gruppen mit der Angabe, welcher Anteil der Argumentationen insgesamt zu einer korrekten Antwort führte.

Tab. 61: Anzahl an Argumentationen pro Richtung zwischen Kräften und Bewegung in beiden Gruppen.

	Modellbildung (Anteil richtig)	Videoanalyse (Anteil richtig)	Gesamt (Anteil richtig)
Kraft zu Bewegung	8 (63 %)	12 (100 %)	20 (85 %)
Bewegung zu Kraft	5 (40 %)	7 (71 %)	12 (58 %)
Gesamt	13 (54 %)	19 (89 %)	32 (75 %)

Daraus wird ersichtlich, dass Schüler*innen es in einer freien Diskussion über physikalische Größen präferieren, von den wirkenden Kräften auf die Bewegung zu schließen bzw. die Eigenschaften der Bewegung mit den wirkenden Kräften zu begründen. Zusätzlich scheint diese Art der Argumentation auch in einer Diskussion erfolgreicher zu sein, was die zuvor berichteten Ergebnisse von Just et al. (2021) stützt. Die Kohorte der Videoanalyse nutzt den Zusammenhang zwischen Kräften und Bewegung insgesamt häufiger, was vor allem an einer Zweiergruppe lag, welche sehr viel argumentierte, sodass nicht von nennenswerten Unterschieden zwischen den Kohorten ausgegangen werden kann. Diese Zweiergruppe verzerrte das Verhältnis zwischen den beiden Argumentationsrichtungen aber nicht. Insgesamt argumentiert die Kohorte der Videoanalyse häufiger korrekt.

Tab. 62: Anzahl der Argumentationen in verschiedenen Richtungen zwischen allen Größen in beiden Kohorten.

	Modellbildung (Anteil richtig)	Videoanalyse (Anteil richtig)	Gesamt (Anteil richtig)
Begründung auf höherer Ebene	10 (70 %)	15 (100 %)	25 (88 %)
Begründung auf gleicher Ebene	21 (57 %)	16 (81 %)	37 (68 %)
Begründung auf niedrigerer Ebene	10 (40 %)	10 (70 %)	20 (55 %)
Antwort ohne Begründung	35 (69 %)	47 (70 %)	82 (70 %)
Gesamt	76 (62 %)	88 (77 %)	164 (70 %)

Erweiterbar ist diese Analyse, indem die kinematischen Größen zusätzlich untersucht werden (Tab. 62). Nun werden auch Stellen betrachtet, in denen mithilfe der Beschleunigung die Geschwindigkeit begründet wird (Begründung auf höherer Ebene) oder andersherum (Begründung auf niedrigerer Ebene). Insgesamt sind die Größen Kraft, Beschleunigung, Geschwindigkeit und Ort in dieser Reihenfolge von hoch nach niedrig sortiert, ohne dass dies eine inhaltliche Wertung impliziert.

Die Schüler*innen argumentieren in den meisten Fällen nicht und nennen die Antwort lediglich. Das kann daran liegen, dass sie die Größen im vergangenen Verlauf des Versuchs bereits ausreichend diskutiert oder einfach geraten haben. In manchen Gruppen kommen auch keine der in der Tabelle erfassten Begründungen vor, was wiederum daran liegen kann, dass die Tabellen gar nicht ausgefüllt oder nicht besprochen und still in Einzelarbeit ausgefüllt wurden. Von denen, die die Größen besprochen und begründet haben, nutzen wiederum die meisten die Begründung auf gleicher Ebene. Die entsprechende Größe wird also selbst beschrieben und damit ihre Eigenschaften abgeleitet (bspw. durch Betrachtung des Graphen oder Beschreibung der Bewegung aus der Erinnerung). Die Begründung auf höherer und niedrigerer Ebene wird seltener genutzt, da diese auch mit einer höheren Komplexität verbunden ist. Von den beiden letztgenannten wird häufiger auf einer höheren Ebene argumentiert. Diese Variante ist von allen die erfolgreichste, was erneut die Ergebnisse aus der vorigen Tabelle bestätigt. Ein Unterschied zwischen den beiden Interventionen ist nicht zu erkennen.

Schüler*innen scheinen also, wenn sie eine andere physikalische Größe zur Erklärung nutzen, eher auf der höheren Ebene zu argumentieren bzw. von Kräften auf die Bewegung zu schließen. Diese Argumentation führt häufiger zu einem richtigen Schluss. Diese Beobachtung machten Just et al. (2021) vor allem beim ersten Newtonschen Gesetz, welches

beim analysierten Versuch eine Rolle spielt, da bei der Stahlkugel in x-Richtung keine (relevante) Kraft wirkt und sich die Kugel deshalb mit einer konstanten Geschwindigkeit fortbewegt. Erklärt wurde dies von den Autor*innen durch eine unterschiedliche Aktivierung von Schülervorstellungen in den beiden Argumentationsrichtungen. Wenn eine konstante Geschwindigkeit bekannt ist und von dieser auf die wirkenden Kräfte geschlossen werden soll, wird häufig die Schülervorstellung aktiviert, dass eine Kraft in Bewegungsrichtung wirken müsse, damit sich etwas überhaupt bewegen kann. Dadurch wird auf eine konstante Kraft geschlossen. Andererseits könnte, wenn von einer resultierenden Kraft von null ausgegangen wird, eher die Vorstellung aktiviert werden, dass keine Kraft keine Änderung bedeutet. Dies könnte dann auf keine Änderung der Geschwindigkeit übertragen werden, was zu einer korrekten Antwort führen würde (Just et al., 2021). Ein Blick auf die in den jeweiligen Gruppen aktivierten Schülervorstellungen ist also auch in den Videos sinnvoll.

11.3 Schülervorstellungen

Welche Schülervorstellungen aktiviert werden, ist stark kontextabhängig (Kapitel 3.3). Eine Kenntnis über die Aktivierung von Schülervorstellung bei einem bestimmten Versuch oder einer Unterrichtsmethode kann wichtig für den erfolgreichen Umgang mit Lernschwierigkeiten bei Schüler*innen sein. Um die in den Interventionen aufgetretenen Schülervorstellungen zu untersuchen, wurden die vorhandenen Schülergespräche in den Videos vollständig analysiert und die geäußerten Vorstellungen den bekannten Schülervorstellungen zugeordnet. Dabei stellten sich vor allem drei Vorstellungen als relevant heraus. Alle weiteren wurden insgesamt maximal zweimal geäußert und spielen deshalb eine untergeordnete Rolle und werden nicht in die Analyse einbezogen (Tab. 63).

Tab. 63: Vorkommen der häufigsten Schülervorstellungen in beiden Gruppen.

Vorstellung	Modellbildung	Videoanalyse	Gesamt
Es muss eine Kraft in Bewegungsrichtung wirken.	34	17	51
Kraft ist eine universelle Wirkungsfähigkeit.	13	10	23
Eine Kraft braucht eine Zeit, bis sie sich durchsetzen kann: Die stärkere Kraft gewinnt.	10	6	16

Am häufigsten kommt die Schülervorstellung vor, dass eine Kraft in Bewegungsrichtung wirken muss. Dies zeigt zwar nicht, dass es sich insgesamt um die häufigste Vorstellung handelt, da deren Aktivierung kontextabhängig ist und an dem besprochenen Experiment gelegen haben könnte. Die Häufigkeit deckt sich aber mit der Auswertung der Antworten im Schülerfragebogen. Zwischen den Gruppen scheint es, im Gegensatz zu den anderen Vorstellungen, bei dieser auch einen Unterschied zu geben. In der letztgenannten Vorstellung kann durch die geringe Häufigkeit keine zuverlässige Aussage über Gruppenunterschiede getroffen werden. Eine Erklärung dafür, dass die erstgenannte Vorstellung in der Gruppe der Modellbildung häufiger geäußert wird, könnte sein, dass sich die Proband*innen beim Modellieren zwangsläufig intensiver mit den wirkenden Kräften auseinandersetzen, da sie diese aktiv eingeben müssen. Dadurch entsteht ein ausgeprägter kognitiver Konflikt, da das Modell ohne diese Kraft in Bewegungsrichtung zur Realität zu passen scheint und von den Schüler*innen oft auch kein mathematischer Ausdruck für die vermeintlich vorhandene Kraft gefunden wird. Dieser kognitive Konflikt könnte wiederum dafür verantwortlich sein, dass sich die Vorstellung im Vergleich zur Gruppe der Videoanalyse durch die Intervention signifikant stärker reduziert (siehe Kapitel 10.5).

Es gibt eine ausgeprägte Diskussion über den Einsatz von kognitiven Konflikten im Rahmen eines Lernprozesses (Kapitel 3.5). Der vorgestellte Konsens, nachdem auch die Interventionen gestaltet waren, ist, dass erst dann mit kognitiven Konflikten gearbeitet werden sollte, wenn eine alternative Theorie vorliegt. Dies ist in der vorliegenden Studie der Fall, da die Proband*innen erst nach dem Unterricht der Newtonschen Gesetze an der Intervention teilnahmen. Das berichtete Teilergebnis könnte dafürsprechen, dass der von der aktiven Modellierung erzeugte kognitive Konflikt an dieser Stelle sinnvoll und hilfreich ist und zu einer Reduktion der unerwünschten Schülervorstellung führt. Dies deckt sich mit der Einschätzung, dass insbesondere bei häufig vorkommenden Schülervorstellungen mit kognitiven Konflikten gearbeitet werden kann, da diese ohnehin häufig aktiviert werden. Eine seltene Schülervorstellung kann dagegen durch einen kognitiven Konflikt teilweise sogar verstärkt werden. Von Vorteil ist hier auch, dass der Computer in der Gruppe der Modellbildung als neutraler Part die Rückmeldungen über die getätigten Hypothesen übernimmt und dies nicht durch die Lehrkraft geschehen muss.

Die Vorstellung „Kraft ist eine universelle Wirkungsfähigkeit“ äußerte sich vor allem darin, dass Schüler*innen z. B. Energie oder Impuls nannten, wenn nach den Kräften gefragt wurde. Auch Beschleunigung und Kraft wurde teilweise synonym verwendet. Da dies eine bekannte Schülervorstellung ist und es hier keinen Unterschied zwischen den beiden Gruppen gibt, wird diese nicht näher diskutiert. Ähnlich ist es bei der Vorstellung „Eine Kraft braucht eine Zeit, bis sie sich durchsetzen kann“. In dem analysierten

Experiment fliegt die Kugel einen Bogen, der einer nach unten geöffneten Parabel gleicht. Die Lernenden gehen davon aus, dass die Erdanziehungskraft sich erst am Hochpunkt der Bewegung gegen die „Bewegungskraft“ des Objekts durchsetzen kann und die Kugel dann anfängt zu fallen. Auch hier bringt eine inhaltliche Diskussion der Schülervorstellung keinen Mehrwert.

11.4 Vorgehensweise

Während in den vorigen Kapiteln das Vorkommen gewisser Argumentationen oder Schülervorstellungen untersucht und quantifiziert wurde, geht es in diesem Kapitel um das Vorgehen der Proband*innen, während sie mit der Software arbeiten. Da es hier große Unterschiede zwischen den Programmen gibt, werden diese separat betrachtet.

11.4.1 Gruppe Modellbildung

Beim Modellieren mit der Software „Newton-II“ wurde in der ersten Aufgabe von der Lehrperson ein gewisser Ablauf der Modellierung verfolgt und motiviert, dass sich die Schüler*innen an diesem orientieren. Dieser Ablauf war an Theorien des Modellierens angelehnt (siehe auch Kapitel 5). Begonnen werden sollte mit einer Problemanalyse, in der diskutiert wird, welche Ausgangssituation vorliegt und was das Ziel der Modellbildung ist. Daran schließt sich dann die Modellbildung selbst an. Hierunter zu verstehen sind alle Überlegungen und Diskussionen zu der zugrundeliegenden „Newton-Maschine“, vorgenommenen Abstraktionen, wirkenden Kräften, etc. Nach der Modellbildung folgt die Implementation der getroffenen Annahmen in die Software. Diese beiden Phasen können parallel stattfinden, wenn die Annahmen während der Diskussion direkt in die Software eingegeben werden. Die Simulation erfolgt je nach Implementation automatisch oder per Mausklick und wird nicht als eigene Phase identifiziert. Nach der Simulation sollte eine erste Evaluation erfolgen. Je nach Ergebnis der Evaluation sollte dann das Modell verbessert werden, also eine erneute Modellbildung, Implementation und Simulation stattfinden, und dann erneut evaluiert werden. Die Realdaten sollten erst hinzugefügt werden, wenn man mit der Struktur des Modells zufrieden ist. Nach Vergleich mit den Realdaten erfolgt eine weitere Evaluation und damit möglicherweise eine erneute Änderung des Modells. Die Modellierenden entscheiden selbst, wann das Modell ausreichend gut zur Realität passt. Wenn dies der Fall ist, werden mit dem Modell Fragen beantwortet, die in der Problemanalyse aufgeworfen wurden. Um zu überprüfen, ob sich die Schüler*innen bei der Modellierung an diesem Ablauf orientiert haben und um einen Einblick zu bekommen, welche Phasen besonders betont und welche eher weggelassen

wurden, wurden diese im Video identifiziert und kodiert. Eine erste Analyse der Häufigkeiten der Vorkommen der verschiedenen Phasen ergibt folgendes Bild (Tab. 64):

Tab. 64: Häufigkeit und Dauer der identifizierten Phasen in den elf Videos der Gruppe der Modellbildung.

Phase	Häufigkeit In Klammern pro Gruppe	Dauer (h:mm:ss) In Klammern pro Gruppe
Problemanalyse	0	0
Modellbildung	34 (3,1)	1:15:45 (0:06:53)
Implementation	26 (2,4)	0:50:01 (0:04:33)
Diagrammbasierte Evaluation	14 (1,3)	0:05:34 (0:00:30)
Überarbeitung des Modells	59 (5,4)	1:49:09 (0:09:55)
Einlesen der Daten	19 (1,7)	0:16:18 (0:01:29)
Datenbasierte Evaluation	58 (5,3)	0:38:20 (0:03:29)
Formulierung von Ergebnissen bzw. Beantwortung der Abschlussfragen		
Ohne Modell	15 (1,4)	0:32:41 (0:02:58)
Mit Modell	18 (1,6)	0:39:53 (0:03:38)
Gesamt	243 (22,1)	6:07:41 (0:33:26)

Diese Daten wurden aus den elf Videos der Gruppe der Modellbildung mit $N = 22$ Proband*innen gewonnen. Zu Beginn wurde der Flug einer Stahlkugel und anschließend der einer Styroporkugel modelliert. Der zweite Teil war dabei eine binnendifferenzierende Zusatzaufgabe, sodass nicht alle Gruppen diesen Teil vollständig gelöst haben. Dies erklärt, warum beispielsweise in elf Gruppen 19-mal Daten eingelesen wurden. Eine echte Problemanalyse wurde nicht beobachtet. Dies kann daran liegen, dass das Vorgehen bei allen vier Versuchen analog war und die Videos im dritten Versuch aufgezeichnet wurden. Da die Proband*innen das Vorgehen also einerseits kannten und dies andererseits durch das Arbeitsheft vorgegeben war, war eine echte Problemanalyse, in der motiviert wird, warum eine Modellierung überhaupt durchgeführt werden sollte, nicht notwendig. Die Phase der Überarbeitung des Modells unterscheidet sich von der Modellbildung nur darin, dass sie nach einer Evaluation stattfindet – egal, ob dies nach der

diagrammbasierten oder der datenbasierten Evaluation erfolgte. Das häufige Vorkommen dieser Überarbeitung ist nicht negativ zu deuten, da nicht erwartet werden kann, dass das Modell auf Anhieb ausreichend gut ist. Es ist eher positiv, dass die Schüler*innen auf die Rückmeldungen des Programms reagieren und ihr Modell anhand dieser Rückmeldungen anpassen. Bei den Evaluationen wird deutlich häufiger im Vergleich zu den Realdaten evaluiert als nur auf Basis des Diagramms, welches sich direkt nach Erstellen des Modells ergibt. Dies könnte die Ursache haben, dass die Evaluation auf Basis des entstehenden Diagramms eine höhere Diagrammkompetenz erfordert, da bekannt sein muss, wie der Verlauf in dem jeweiligen Diagramm eigentlich aussehen müsste. Beim Vergleich des Modells mit den Realdaten muss nur verglichen werden, ob der reale Verlauf ähnlich zum modellierten Verlauf ist, ohne die verschiedenen Diagramme verstehen zu müssen.

Einige der Gruppen (vier von elf) lesen bei einer der Modellierungen direkt nach der Modellbildung die Daten ein, während zwei Gruppen dies sogar mindestens an einer Stelle vor der Modellbildung tun, um die Messdaten als Vorlage zu verwenden. Dies birgt die Gefahr, dass das Modellieren nur inhaltloses Ausprobieren ist, bis die entstehenden Diagramme zu den Daten passen. Auch durch dieses Vorgehen wäre es zwar möglich, etwas über die wirkenden Kräfte zu lernen, ein zielgerichteteres Vorgehen erscheint aber sinnvoller.

Nur etwa die Hälfte der Proband*innen nutzt das Modell, um die Fragen aus dem Arbeitsheft zu beantworten. Dies kann einerseits bedeuten, dass durch die Modellierung die Bewegung umfassend verstanden wurde, sodass die Modellierung selbst nicht mehr benötigt wird. Dagegen spricht andererseits, dass einige der formulierten Ergebnisse falsch sind und die Betrachtung der Modellierung geholfen hätte. Eine andere Erklärung wäre, dass die Modellierung nur vorgenommen wird, da dies Teil der Aufgabe ist und nicht klar genug wird, dass die Modellierung wiederum Fragen über die Bewegung selbst beantworten kann. Um die formulierten Fragestellungen näher zu beleuchten, können die entsprechenden Stellen der Schülergespräche in den Videos genauer betrachtet werden. Mehrere Gruppen diskutieren z. B., welche Kräfte auf das Objekt wirken, während die korrekte und vollständige Modellierung am PC zu sehen ist. Dabei schauen sie nicht nach, was bei den Kräften modelliert wurde. Ebenso wird die Chance häufig nicht genutzt, auch die kinematischen Größen in der Software betrachten zu können, nach denen am Ende der Aufgabe explizit gefragt wird. Nicht deutlich wird aus den Videos, warum einige das Modell nicht nutzen. Nur eine Gruppe formuliert alle Ergebnisse ohne das Modell. Die anderen Gruppen nutzen bei einigen Ergebnissen das Modell, bei anderen, in denen sie dies auch gebrauchen könnten, aber nicht. Es sollte durch die Lehrperson nochmals

motiviert werden, dass das Modell für die Formulierung der Endergebnisse genutzt werden sollte.

11.4.2 Gruppe Videoanalyse

Ähnlich wie bei „Newton-II“ kann auch bei „measure dynamics“ die Arbeit in unterschiedliche Phasen geteilt werden, die sich aber stark von den Arbeitsphasen bei der mathematischen Modellbildung unterscheiden.

Tab. 65: Häufigkeit und Dauer der identifizierten Phasen in den elf Videos der Gruppe der Videoanalyse.

Phase	Häufigkeit In Klammer pro Gruppe	Dauer (h:mm:ss) In Klammern pro Gruppe
Erzeugung der Messdaten	26 (2,6)	0:39:25 (0:03:57)
Darstellung und Interpretation der Daten		
Stroboskopbild	3 (0,3)	0:03:06 (0:00:19)
Diagramm		
Ort	7 (0,7)	0:05:48 (0:00:35)
Geschwindigkeit	23 (2,3)	0:44:29 (0:04:27)
Beschleunigung	16 (1,6)	0:40:54 (0:04:05)
Pfeile		
Geschwindigkeit	8 (0,8)	0:06:36 (0:00:40)
Beschleunigung	24 (2,4)	0:40:36 (0:04:04)
Formulierung von Ergebnissen bzw. Beantwortung der Abschlussfragen		
Ohne die Messdaten	19 (1,9)	0:23:24 (0:02:20)
Mit den Messdaten	6 (0,6)	0:17:24 (0:01:44)
Gesamt	132 (13,2)	3:41:42 (0:22:10)

Das Vorgehen im ersten Versuch, welcher im Plenum durchgeführt wurde, teilte sich in Erzeugung der Messdaten, Darstellung der Messdaten, Interpretation der Messdaten und

Formulierung von Ergebnissen auf. Die Interpretation und Diskussion der Messdaten erfolgt immer parallel zu deren Darstellung, sodass diese Arbeitsphase gemeinsam mit der Darstellung kodiert wird. Die Phase der Darstellung der Messdaten wird dabei für die unterschiedlichen Darstellungsmöglichkeiten der Software (Diagramme der verschiedenen kinematischen Größen, Pfeile der verschiedenen kinematischen Größen, Stroboskopbild etc.) mehrfach durchlaufen (Tab. 65). Auch hier handelt es sich um die Analyse der Stahlkugel und als Zusatzaufgabe die der Styroporkugel.

Die genannten Phasen müssen hier, anders als bei der mathematischen Modellbildung, alle durchlaufen werden. Bei der Darstellung der Daten werden die Proband*innen durch das Arbeitsheft beeinflusst, in dem Beschleunigungspfeile und Geschwindigkeits- und Beschleunigungsdiagramme gefordert werden. Diese sind dementsprechend am häufigsten vertreten und nehmen auch die meiste Zeit in Anspruch. Stroboskopbilder, Ortsdiagramme und Geschwindigkeitspfeile werden von den Schüler*innen zum besseren Verständnis ebenfalls genutzt, auch wenn diese nicht vorgegeben sind. Weiterhin fällt auf, dass die Proband*innen bei der Formulierung der Ergebnisse noch seltener auf die Daten zurückgreifen, als dies bei den Modellierenden der Fall war. Auch hier könnten die Schüler*innen es durch die vorige Analyse nicht unbedingt nötig haben, für die Formulierung der Ergebnisse in die Daten zu schauen. Dennoch sind wieder einige Ergebnisse falsch. Auch hier sollte betont werden, dass die Analyse durchgeführt wird, um Erkenntnisse über die Bewegung zu erlangen und die gewonnenen Daten damit auch genutzt werden sollten.

11.5 Zusammenhänge mit dem Lernerfolg

Im linearen hierarchischen Modell aus Kapitel 9.3.8 wurden Prädiktoren für den Lernerfolg in den Interventionen identifiziert. Gemäß eines Mixed-Methods-Ansatzes bietet es sich zusätzlich an, die Vorgehensweise und die Schwierigkeiten der Lernenden aus den vergangenen Kapiteln zu dieser Betrachtung hinzuzuziehen. Da es sich um eine qualitative Analyse der Arbeitsweise der Proband*innen handelt, ist es aufgrund der hohen Datenmenge nicht möglich, die Arbeitsweise aller Teilnehmenden zu untersuchen. Daher wurden insgesamt $N = 45$ Lernende aufgenommen und analysiert. Dementsprechend kann die Vorgehensweise nicht in das erstellte hierarchische lineare Modell eingepflegt werden. Vielmehr soll auf qualitativer Ebene untersucht werden, welche Arbeitsweisen zu einem eher großen oder eher kleinen Lernzuwachs führen. Um herauszufinden, welche Vorgehensweisen erfolgreich sind, werden die beobachteten Gruppen auf Basis des Lernzuwachses in drei Terzile (hoher, mittlerer und niedriger Lernzuwachs) eingeteilt. Der Lernzuwachs ist die absolute Differenz aus Nachtestergebnis und Vortestergebnis (eine

Einteilung nach dem relativen Lernzuwachs wäre in diesem Fall identisch). Dieser wird auf Schülerebene berechnet und dann der Durchschnitt der Zweier- oder Dreiergruppe gebildet. Eine Herausforderung stellen dabei fehlende Daten dar. Da die Proband*innen randomisiert für die Aufnahme der Bildschirmvideos ausgewählt wurden, konnte nicht garantiert werden, dass von diesen Proband*innen auch Vortestergebnisse vorliegen. In diesem Fall basiert der Lernzuwachs dann ausschließlich auf dem Partner bzw. der Partnerin, von dem bzw. der die Daten vorliegen. Eine Gruppe fällt aus der Analyse gänzlich raus, da in dieser beide Proband*innen nicht am Vortest teilgenommen haben. In sechs weiteren Gruppen fehlte das Vortestergebnis von je einer Person. Andererseits gibt es auch eine Gruppe, in der ein Teilnehmer krankheitsbedingt nach Aufzeichnung des Videos aber vor der Durchführung des Nachtests das Schülerlabor verlassen musste. In diesem Fall wird ebenfalls nur der Lernzuwachs einer Person berücksichtigt. Für die Analyse werden die Vorgehensweisen des oberen Terzils mit denen des unteren Terzils verglichen und einzelne Aspekte herausgearbeitet, die möglicherweise zu einer erfolgreichen Arbeit mit der Software führen. Dazu wird das gesamte Kategoriensystem genutzt, was zwischen Schwierigkeiten, Schülervorstellungen, Argumentationen und den unterschiedlichen Arbeitsphasen unterscheidet. Da zunächst nur Korrelationen gefunden werden können, wird jeweils darauf eingegangen, inwiefern eine Kausalität zwischen den gefundenen Unterschieden und dem Lernzuwachs vorliegen könnte. Da die Vorgehensweise bei der mathematischen Modellbildung grundverschieden von der der Videoanalyse ist, werden für beide Varianten Terzile gebildet und innerhalb einer Variante verglichen. Damit sind bei beiden Varianten je vier Gruppen im oberen und vier im unteren Terzil. Dies reduziert die Datenmenge und damit die Generalisierbarkeit der Aussagen. Die Unterscheidung zwischen Gruppen ist jedoch notwendig und der Vergleich zwischen oberem und unterem Terzil stellt sicher, dass überhaupt ausreichende Unterschiede im Lernzuwachs zwischen diesen Kohorten vorliegen.

11.5.1 Gruppe Modellbildung

In der Gruppe der mathematischen Modellbildung wurden Videos von elf Gruppen ($N = 22$ Proband*innen) angefertigt. Daher werden vier Gruppen in das untere, drei in das mittlere und vier in das obere Terzil eingeteilt. Das untere Terzil hat einen mittleren Lernzuwachs von 1,3 Punkten, das obere Terzil von 7,4 Punkten (bei jeweils 15 maximal erreichbaren). Tab. 66 gibt eine Übersicht über die Häufigkeit und die Dauer der aufgetretenen Schwierigkeiten. Außer der Tatsache, dass die inhaltlichen Schwierigkeiten beim unteren Terzil etwas mehr Zeit in Anspruch nehmen, ist kein Unterschied zwischen den Kohorten zu erkennen.

Tab. 66: Schwierigkeiten im Vergleich zwischen oberem und unterem Terzil in der Gruppe der Modellbildung.

Schwierigkeiten	Oberes Terzil	Unteres Terzil
	Häufigkeit (Dauer in h:mm:ss)	Häufigkeit (Dauer in h:mm:ss)
Inhaltliche Schwierigkeiten	7 (0:14:49)	10 (0:20:34)
Schwierigkeiten mit der Software	9 (0:16:43)	7 (0:17:12)

Weder inhaltliche Schwierigkeiten noch Probleme bei der Bedienung der Software beeinflussen den Lernzuwachs nachweisbar negativ. Ähnlich ist es bei der Äußerung von Schülervorstellungen (Tab. 67).

Tab. 67: Häufigkeit der vorkommenden Schülervorstellungen im Vergleich zwischen oberem und unterem Terzil in der Gruppe der Modellbildung.

Schülervorstellung	Oberes Terzil	Unteres Terzil
Es muss eine Kraft in Bewegungsrichtung wirken.	13	14
Kraft ist eine universelle Wirkungsfähigkeit.	7	1
Eine Kraft braucht eine Zeit, bis sie sich durchsetzen kann: Die stärkere Kraft gewinnt.	5	5

Auch hier lassen sich keine nennenswerten Unterschiede erkennen. Diese Daten stellen außerdem die zuvor geäußerte These infrage, dass ein häufigeres Vorkommen der Schülervorstellung, dass eine Kraft in Bewegungsrichtung wirken muss, durch den entstehenden kognitiven Konflikt zu einem höheren Lernzuwachs führt. Der kognitive Konflikt könnte dennoch bei einigen Schüler*innen zu einem höheren Lernzuwachs führen, während bei anderen trotz des Widerspruchs zwischen der geäußerten Vorstellung und der physikalischen Theorie kein kognitiver Konflikt auftritt. Der hier angestellte Vergleich von jeweils vier Gruppen ist statistisch allerdings wenig belastbar. Zudem bezog sich diese These nur auf den Lernzuwachs im Bereich des ersten Newtonschen Gesetzes, der hier nicht einzeln betrachtet wird.

Wenn man die Terzile nach Vortestergebnis einteilt, wird deutlich, dass die Schüler*innen, die mit einem niedrigeren Vortestergebnis in die Intervention kommen, alle betrachteten Schülervorstellung häufiger äußern als die besseren Schüler*innen. Dies ist plausibel, da die genannte Vorstellung im Test eher zu falschen Antworten führen.

Die Argumentationen ausschließlich zwischen Kraft und Bewegung während der Arbeit mit der Software sind bei dieser Einschränkung der Daten so selten, dass ein Vergleich wenig Sinn ergibt. Betrachtet werden können aber alle Argumentationen (inklusive derer zwischen kinematischen Größen), die bei den Formulierungen der Ergebnisse gemacht wurden (Tab. 68):

Tab. 68: Häufigkeit der vorkommenden Argumentationen im Vergleich zwischen oberem und unterem Terzil in der Gruppe der Modellbildung.

Argumentationen	Oberes Terzil Häufigkeit (Anteil richtig)	Unteres Terzil Häufigkeit (Anteil richtig)
Begründung auf höherer Ebene	3 (100 %)	7 (57 %)
Begründung auf gleicher Ebene	12 (75 %)	9 (33 %)
Begründung auf niedrigerer Ebene	8 (50 %)	2 (0 %)
Antwort ohne Begründung	18 (78 %)	7 (43 %)
Summe	41 (73 %)	25 (40 %)

Im oberen Terzil werden insgesamt mehr Ergebnisse diskutiert und anteilig häufiger das richtige Ergebnis geschlussfolgert. Nur der Begründung auf höherer Ebene bedient sich das untere Terzil häufiger. Es liegt nahe, dass eine ausgeprägte Diskussion der Ergebnisse positiv zum Zuwachs des Konzeptverständnisses beiträgt. Es könnte andersherum aber auch sein, dass erfolgreichere Schülergruppen mehr diskutieren. Eine gewisse Aufklärung kann die Einteilung der Terzile nach Vortestergebnis liefern. Bei dieser Einteilung werden im oberen Terzil 24 Größen (54 % korrekt) und im unteren Terzil 30 Größen (67 % korrekt) besprochen. Dies stützt die These, dass das Besprechen der Größen unabhängig vom vorigen Leistungsstand den Lernzuwachs steigert. Es kann auch ein Vergleich der Arbeitsphasen zwischen den beiden Terzilen herangezogen werden (Tab. 69).

Im unteren Terzil werden die Ergebnisse seltener und in geringerer Dauer formuliert. Am größten ist der Unterschied bei der Nutzung des Modells. Es könnte also zielführend sein, die Phase der selbstständigen Ergebnisfindung und -sicherung noch mehr zu betonen. Zudem sollte das Modell herangezogen werden, um die Ergebnisse zu formulieren. Weiterhin ist ein Unterschied in der Überarbeitung des Modells zu erkennen. Das untere Terzil verbringt hier deutlich mehr Zeit, was daran liegen könnte, dass das zunächst erstellte Modell schlechter ist oder dass es in diesen Gruppen länger dauert, die Fehler zu beheben.

In jedem Fall fehlt diese Zeit dann an anderer Stelle (z. B. bei der Formulierung der Endergebnisse oder der Evaluation des Modells).

Tab. 69: Häufigkeit und Dauer der identifizierten Phasen in den elf Videos im Vergleich zwischen oberem und unterem Terzil in der Gruppe der Modellbildung.

Phase	Oberes Terzil Häufigkeit (Dauer in h:mm:ss)	Unteres Terzil Häufigkeit (Dauer in h:mm:ss)
Problemanalyse	0	0
Modellbildung	11 (0:29:21)	10 (0:25:08)
Implementation	10 (0:19:09)	8 (0:19:58)
Diagrammbasierte Evaluation	7 (0:03:50)	5 (0:01:25)
Überarbeitung des Modells	14 (0:25:33)	28 (0:47:29)
Einlesen der Daten	6 (0:05:10)	6 (0:05:23)
Datenbasierte Evaluation	19 (0:14:51)	23 (0:11:05)
Formulierung von Ergebnissen bzw. Beantwortung der Abschlussfragen		
Ohne Modell	9 (0:15:16)	4 (0:14:01)
Mit Modell	10 (0:22:36)	3 (0:10:37)

Wenn mathematische Modellbildung im Unterricht eingesetzt wird, sollte also darauf geachtet werden, dass in den Gruppen ausreichend Zeit für die Ergebnisbesprechung genutzt wird. Die Schüler*innen sollten dazu motiviert werden, über die relevanten Größen zu diskutieren und das erstellte Modell für die Beantwortung der Fragen zu nutzen. Diese Faktoren scheinen den größten Einfluss auf den Lernerfolg zu haben und sind, anders als z. B. das Interesse an theoretischen Zusammenhängen, unmittelbar durch die Art des Einsatzes der Software beeinflussbar. Weitere Untersuchungen über die Richtung der Kausalität wären angebracht.

11.5.2 Gruppe Videoanalyse

In der Gruppe der Videoanalyse wurden Videos von zehn Gruppen ($N = 23$ Proband*innen) angefertigt. Um auch hier eine Symmetrie in den Terzilen und Vergleichbarkeit mit der Gruppe der mathematischen Modellbildung zu haben, werden vier Gruppen in das

untere, zwei Gruppen in das mittlere und vier Gruppen in das obere Terzil eingeordnet. Obwohl es sich also nicht um gleichgroße Terzile sondern um drei unterschiedlich große Gruppen handelt, die nach dem Lernzuwachs eingeteilt wurden, wird dieser Begriff hier verwendet. Das untere Terzil hat einen mittleren Lernzuwachs von -0,5 Punkten (also einen Verlust), das obere Terzil von 4,1 Punkten. Obwohl sich Kohorten der mathematischen Modellbildung und Videoanalyse insgesamt nicht unterscheiden in ihrem Lernzuwachs (siehe Kapitel 9.3) sind die Stichproben, die für die Bildschirmvideos ausgewählt wurden, in dieser Hinsicht unterschiedlich. Wie bei „Newton-II“ können die beiden Terzile miteinander verglichen werden. Zunächst werden die auftretenden Schwierigkeiten betrachtet (Tab. 70):

Tab. 70: Identifizierte Schwierigkeiten im Vergleich zwischen oberem und unterem Terzil in der Gruppe der Videoanalyse.

Schwierigkeiten	Oberes Terzil	Unteres Terzil
	Häufigkeit (Dauer in h:mm:ss)	Häufigkeit (Dauer in h:mm:ss)
Inhaltliche Schwierigkeiten	1 (0:00:48)	4 (0:06:37)
Schwierigkeiten mit der Software	10 (0:07:20)	20 (0:27:56)

Anders als bei der Modellbildung treten in der Gruppe der Videoanalyse mehr Schwierigkeiten im unteren Terzil auf. Zudem nehmen diese einen deutlich längeren Zeitraum in Anspruch. Diese Zeit fehlt dann an anderer Stelle. Die Dauer, die die Gruppen mit Schwierigkeiten zubringen, könnte einen Einfluss auf den Lernzuwachs zu haben. Eine Einteilung der Terzile nach Vortestergebnissen zeigt keinen Unterschied zwischen den Terzilen, sodass nicht das Vortestergebnis dafür verantwortlich zu sein scheint, wie viele Schwierigkeiten auftreten. In ähnlicher Weise können die geäußerten Schülervorstellungen betrachtet werden (Tab. 71):

Tab. 71: Häufigkeit der vorkommenden Schülervorstellungen im Vergleich zwischen oberem und unterem Terzil in der Gruppe der Videoanalyse.

Schülervorstellung	Oberes Terzil	Unteres Terzil
Es muss eine Kraft in Bewegungsrichtung wirken.	6	5
Kraft ist eine universelle Wirkungsfähigkeit.	1	6
Eine Kraft braucht eine Zeit, bis sie sich durchsetzen kann: Die stärkere Kraft gewinnt.	2	3

Ähnlich wie in der Gruppe der mathematischen Modellbildung gibt es keinen nennenswerten Unterschied zwischen den Gruppen. Zwar wird die Schülervorstellung „Kraft ist eine universelle Wirkungsfähigkeit“, hinter der ein diffuser Kraftbegriff steckt, im unteren Terzil deutlich häufiger genannt, dies kann aber ebenso Zufall sein. Das Vorhandensein und Äußern von Schülervorstellungen scheint den Lernzuwachs hier nicht zu beeinflussen. Dies ist plausibel, da Schülervorstellungen einerseits lernhinderlich sein können, die Äußerung dieser aber auch zu einem Umdenken und dem Abstellen der Schülervorstellung führen kann, was wiederum zu einem Lernzuwachs führen würde. Weitere Erkenntnisse liefern die von den Schüler*innen getätigten Argumentationen (Tab. 72).

Tab. 72: Häufigkeit der vorkommenden Argumentationen im Vergleich zwischen oberem und unterem Terzil in der Gruppe der Videoanalyse.

Argumentationen	Oberes Terzil Häufigkeit (Anteil richtig)	Unteres Terzil Häufigkeit (Anteil richtig)
Begründung auf höherer Ebene	5 (100 %)	1 (100 %)
Begründung auf gleicher Ebene	10 (80 %)	2 (50 %)
Begründung auf niedrigerer Ebene	3 (67 %)	0 (-)
Antwort ohne Begründung	28 (71 %)	7 (29 %)
Summe	46 (76 %)	10 (40 %)

Das obere Terzil bespricht wesentlich mehr Ergebnisse (46 vs. 10) und liefert mehr Begründungen (18 vs. 3). Der Anteil richtiger Antworten ist in jeder der Kategorien beim oberen Terzil höher. Da dies bereits in der mathematischen Modellbildung der Fall war, verstärkt sich die These, dass die Ergebnissicherung eine wichtige Phase ist und den Lernzuwachs positiv beeinflusst oder dass erfolgreichere Gruppen mehr diskutieren. Besonders groß ist der Unterschied zwischen den Terzilen darin, wie viele Begründungen für die Größen geliefert werden. Eine Kraft oder kinematische Größe mit einer anderen zu begründen, erfordert ein intensives Nachdenken über die vorhandenen Zusammenhänge. Die Gesamtdauer der Formulierung von Ergebnissen im unteren Terzil ist nicht geringer (Tab. 73). Diese Phase scheint im oberen Terzil produktiver genutzt zu werden, da mehr Größen diskutiert werden (Tab. 72). Auch hier ist der Unterschied zwischen Terzilen, die nach Vortestergebnissen eingeteilt werden kaum vorhanden (40 vs. 33).

Zudem sind Zeiten innerhalb der Formulierung von Ergebnissen ungleich verteilt (Tab. 73). Im oberen Terzil werden häufiger die Messdaten genutzt, um daraus Ergebnisse

abzuleiten, was wiederum zu dem oben berichteten höheren Anteil richtiger Antworten geführt haben könnte. Das untere Terzil verbringt mehr Zeit mit der Erzeugung von Messdaten, welche dann an anderer Stelle fehlt. Außerdem beschäftigt sich das obere Terzil mehr mit Diagrammen zur Betrachtung der Daten und das untere Terzil mehr mit Pfeilen. Dies muss aber nicht heißen, dass Pfeile schlechter für das Verständnis geeignet sind. Das Meiden von Diagrammen könnte für Unsicherheiten im Bereich der Diagrammkompetenz sprechen. Dies könnte im unteren Terzil ausgeprägter sein, sodass mehr auf leichter verständliche Pfeile zurückgegriffen wird.

Tab. 73: Häufigkeit und Dauer der identifizierten Phasen in den elf Videos im Vergleich zwischen oberem und unterem Terzil in der Gruppe der Videoanalyse.

Phase	Oberes Terzil Häufigkeit (Dauer in h:mm:ss)	Unteres Terzil Häufigkeit (Dauer in h:mm:ss)
Erzeugung der Messdaten	10 (0:11:21)	12 (0:21:20)
Darstellung und Interpretation der Daten		
Stroboskopbild	2 (0:01:35)	1 (0:01:31)
Diagramm		
Ort	5 (0:03:52)	2 (0:02:00)
Geschwindigkeit	9 (0:22:07)	8 (0:14:13)
Beschleunigung	6 (0:21:12)	4 (0:08:49)
Pfeile		
Geschwindigkeit	3 (0:01:20)	2 (0:02:32)
Beschleunigung	9 (0:11:14)	11 (0:18:52)
Formulierung von Ergebnissen bzw. Beantwortung der Abschlussfragen		
Ohne die Messdaten	6 (0:06:34)	11 (0:13:55)
Mit den Messdaten	4 (0:04:13)	1 (0:00:57)

Schwierigkeiten mit der Software könnten den Lernzuwachs negativ beeinflussen. Es sollten also noch mehr Hilfen zur Benutzung der Software bereitgestellt werden. Es ist

aber auch zu erwarten, dass sich diese Schwierigkeiten bei wiederholtem Einsatz der Software in der Schule reduzieren. Auch in der Gruppe der Videoanalyse ist ein entscheidender Faktor, wie viel über die vorhandenen Größen diskutiert wird und wie diese begründet werden. Dies sollte in der Arbeit mit dem Programm gefördert werden. Auch die Nutzung der Messdaten für die Formulierung von Ergebnissen könnte nützlich sein.

11.5.3 Gesamtbetrachtung

Die Terzile in der Gruppe der mathematischer Modellbildung unterscheiden sich im Lernzuwachs deutlich von denen der Videoanalyse (siehe Kapitel 11.5.1 und 11.5.2). Durch diese Unterschiede lassen sich aus der zuvor getätigten Einteilung möglicherweise keine allgemeinen Aussagen ableiten. Wenn man unabhängig der Software das obere Terzil aller betrachteten Schülergruppen mit dem unteren vergleicht, befinden sich vor allem Proband*innen aus der Gruppe der mathematischen Modellbildung im oberen Terzil (fünf von sieben). Im unteren Terzil sind drei von sieben Gruppen aus der mathematischen Modellbildung. Mit Ausnahme der Arbeitsphase können die betrachteten Analysen auch zwischen diesen Terzilen verglichen werden. Mit dieser Einteilung haben die oberen sieben Gruppen (oberes Terzil) einen durchschnittlichen Lernzuwachs von 6,43 Punkten, die unteren sieben Gruppen (unteres Terzil) verbessern sich durchschnittlich um 0,07 Punkte. Beim Vergleich der unterschiedlichen Dimensionen ergibt sich ein ähnliches Bild wie beim Vergleich innerhalb der Gruppen (Tab. 74):

Tab. 74: Identifizierte Schwierigkeiten im Vergleich zwischen oberem und unterem Terzil.

Schwierigkeiten	Oberes Terzil	Unteres Terzil
	Häufigkeit (Dauer in h:mm:ss)	Häufigkeit (Dauer in h:mm:ss)
Inhaltliche Schwierigkeiten	15 (0:27:21)	14 (0:27:11)
Schwierigkeiten mit der Software	18 (0:30:27)	26 (0:42:43)

Es gibt nur geringe Unterschiede zwischen den Terzilen im Auftreten von Schwierigkeiten. Schwierigkeiten mit der Software treten aber häufiger im unteren Terzil auf und nehmen auch mehr Zeit in Anspruch. Bei den Schülervorstellungen ergibt sich kein eindeutiges Bild (Tab. 75). Zwei der drei Schülervorstellungen werden im oberen Terzil häufiger geäußert bzw. aktiviert und die andere im unteren Terzil.

Tab. 75: Häufigkeit der vorkommenden Schülervorstellungen im Vergleich zwischen oberem und unterem Terzil.

Schülervorstellung	Oberes Terzil	Unteres Terzil
Es muss eine Kraft in Bewegungsrichtung wirken.	20	18
Kraft ist eine universelle Wirkungsfähigkeit.	12	7
Eine Kraft braucht eine Zeit, bis sie sich durchsetzen kann.	5	8

Im oberen Terzil wird eindeutig mehr argumentiert und diese Argumentationen führen häufiger zur korrekten Antwort (Tab. 76). Die Arbeitsphase der Formulierung von Ergebnissen und das Sprechen über die vorkommenden Größen könnten also wichtig für den Lernzuwachs der Schüler*innen zu sein. Wie bereits diskutiert, wäre es aber auch denkbar, dass Gruppen, die erfolgreicher sind und mehr dazu lernen, mehr diskutieren, ohne dass dies der Grund für den höheren Lernzuwachs ist. Die Gesamtbetrachtung bestätigt die Ergebnisse des Vergleichs innerhalb der zwei Gruppen.

Tab. 76: Häufigkeit der vorkommenden Argumentationen im Vergleich zwischen oberem und unterem Terzil.

Argumentationen	Oberes Terzil Häufigkeit (Anteil richtig)	Unteres Terzil Häufigkeit (Anteil richtig)
Begründung auf höherer Ebene	4 (100 %)	8 (63 %)
Begründung auf gleicher Ebene	15 (73 %)	11 (36 %)
Begründung auf niedrigerer Ebene	10 (60 %)	2 (0 %)
Antwort ohne Begründung	36 (81 %)	14 (36 %)
Summe	65 (77 %)	35 (40 %)

11.6 Zusammenfassung zum qualitativen Teil

In der Analyse der Schülergespräche der Bildschirmvideos wurden die am häufigsten auftretenden Schwierigkeiten deutlich. Inhaltlich war das größte Hindernis für eine erfolgreiche Beschäftigung mit den Lerninhalten die Diagrammkompetenz. Dies sollte von den Lehrkräften bedacht werden. Zudem zeigten sich einige Schwierigkeiten im Umgang mit der Software (vor allem beim Einstellen der Diagramme). Verbesserungen an der Software und den Materialien könnten diese Schwierigkeiten reduzieren. Die Analyse der

Schülervorstellungen zeigte Unterschiede im Auftreten der Vorstellung, dass eine Kraft in Bewegungsrichtung wirken muss, damit sich etwas bewegen kann. Diese Vorstellung trat in der Gruppe der Modellbildung häufiger auf, was mit der direkteren Konfrontation durch die Modellierung erklärt werden kann.

Die Lernenden nutzten häufiger Kräfte, um die Bewegung zu begründen bzw. zu erklären, als dass sie die Bewegung nutzten, um damit auf wirkende Kräfte zu schließen. Dies war in beiden Gruppen der Fall und scheint damit unabhängig von der Unterrichtsmethode zu sein. Gleichzeitig führte ein größerer Teil der Argumentationen von Kräften auf Bewegung zu einem richtigen Ergebnis als in der anderen Richtung. Es erscheint also sinnvoll, dass Lernende von sich aus eher diese Argumentationsrichtung nutzen. Deutlich häufiger, als auf einer anderen Ebene zu argumentieren, liefern Schüler*innen keine Begründung für ihre Ergebnisse oder begründen diese durch Beschreibung der Größe selbst, ohne diese mit einer anderen zu verknüpfen.

Die Phase der Diskussion über Ergebnisse scheint wichtig für den Lernerfolg der Gruppe zu sein. Erfolgreichere Gruppen diskutierten deutlich häufiger und länger über die Größen und nutzten alle Arten der Argumentation intensiver. Während in der Videoanalyse ein negativer Zusammenhang zwischen auftretenden Schwierigkeiten mit der Software und dem Lernzuwachs beobachtet werden konnte, war dieser Zusammenhang in der Gruppe der Modellbildung nicht vorhanden. Auch die geäußerten Schülervorstellungen konnten nicht in Zusammenhang mit dem Lernerfolg gebracht werden. Erfolgreicher scheint es aber zu sein, wenn das erstellte Modell bzw. die erhobenen Messdaten für die Formulierung der Ergebnisse genutzt werden.

12 Kombination der Ergebnisse

Der hier beschriebene Mixed-Methods-Ansatz soll dazu führen ein ganzheitliches Bild der Stärken und Schwächen der mathematischen Modellbildung und der Videoanalyse zu generieren. Dementsprechend werden die Ergebnisse der beiden vorigen Kapitel hier zusammengeführt und miteinander verknüpft. Während die unterschiedlichen Forschungsmethoden teilweise eigenständige Ergebnisse liefern, lassen sich insbesondere in zwei Bereichen die Ergebnisse der quantitativen Daten mit der qualitativen Untersuchung verknüpfen. Einerseits könnte die Verbesserung im ersten Newtonschen Gesetz in einem kognitiven Konflikt begründet liegen, der sich im Auftreten von Schülervorstellungen niederschlägt, andererseits können die Prädiktoren für einen hohen Lernzuwachs aus beiden Erhebungen kombiniert werden.

12.1 Verbesserung der Ergebnisse im ersten Newtonschen Gesetz

In Kapitel 9.3.9 wurde gezeigt, dass die Gruppe der mathematischen Modellbildung im Inhaltsbereich des ersten Newtonschen Gesetzes mit Kräftekompensation, in dem von der Geschwindigkeit auf die wirkenden Kräfte geschlossen werden musste, einen größeren Lernzuwachs verzeichnen konnte. Gleichzeitig wurde eine stärkere Reduktion der Schülervorstellung, dass eine Kraft in Bewegungsrichtung wirken muss, zwischen Vor- und Nachtest festgestellt. Es wurde diskutiert, dass die Reduktion der Schülervorstellung Ursache für die Verbesserung in der Skala sein könnte. Diese mögliche Verbindung wurde bereits von Just et al. (2021) hergestellt. Diese Verbindung erklärt allerdings noch nicht, warum bei der mathematischen Modellbildung die entsprechende Schülervorstellung stärker reduziert wird. Die Analyse der Bildschirmvideos und Tonaufnahmen zeigt, dass die Schülervorstellung in der Arbeitsphase in der Gruppe der Modellbildung deutlich häufiger genannt wird. Dies passiert, da die Proband*innen die Kraft in horizontaler Richtung bei einem reibungsfreien schrägen Wurf selbst modellieren müssen. An dieser Stelle wird überlegt, welche Kraft wirkt, wie man diese benennen könnte und wie sie in einer mathematischen Formel auszudrücken ist. Teilweise scheitert es an der Benennung, teilweise an der Mathematisierung, teilweise wird eine solche Kraft auch in irgendeiner Form modelliert. In jedem Fall erhalten die Nutzenden bei Vergleich des Modells mit den Messdaten die Rückmeldung, dass eine solche Kraft nicht existiert. Es ist also davon auszugehen, dass der im Vergleich zur Videoanalyse stärkere kognitive Konflikt mit einer anschließenden Auflösung zu einer Reduzierung dieser Schülervorstellung führt. Diese Reduktion verbessert wiederum die Ergebnisse in Items zum ersten Newtonschen Gesetz mit Kräftekompensation bei Argumentation von Bewegung zu Kräften.

12.2 Prädiktoren für einen erfolgreichen Einsatz der genutzten Software

Sowohl in den Fragebögen als auch in den Schülergesprächen der Bildschirmvideos wurde nach Faktoren geschaut, die den Lernerfolg positiv beeinflussen, um diese für das Design eines erfolgreichen Unterrichts mit der entsprechenden Methode zu berücksichtigen. Dabei kann es sich einerseits um Personenmerkmale handeln, die nicht unmittelbar beeinflusst werden können, aber auch um Gestaltungsmerkmale der Intervention.

In der Mehrebenenanalyse stellten sich neben dem Vortestergebnis das Interesse an theoretischen Zusammenhängen, die kognitive Belastung und die Physiknote als Einflussfaktoren auf das Nachtestergebnis heraus. Das Interesse an theoretischen Zusammenhängen lässt sich langfristig durch den Physikunterricht möglicherweise verbessern, hat aber keinen Einfluss auf die unmittelbare Implementation von mathematischer Modellbildung oder Videoanalyse im Unterricht. Die Physiknote ist u. a. ein Maß für vergangene Leistungen im Fach Physik, erhöht möglicherweise aber auch das Selbstvertrauen im Kontext Physik. Das Modell zeigt einen Zusammenhang zwischen Physiknote und fachspezifischem Selbstkonzept. Direkt beeinflussbar ist aber die kognitive Belastung, da diese u. a. von der Lernumgebung abhängt. Die Analyse der Videos zeigte vor allem, dass Schwierigkeiten beim Umgang mit der Software verringert werden müssen und dass die Phase der Ergebnissicherung wichtig für ein erfolgreiches Lernen ist. Bei letzterem sollte vor allem motiviert werden, die vorhandenen Größen zu besprechen und die formulierten Ergebnisse zu begründen.

Nach diesen Ergebnissen sollte darauf geachtet werden, dass die Lernumgebung und die Aufgaben klar und verständlich sind und Hilfen zur Software angeboten werden, um die kognitive Belastung und Schwierigkeiten mit der Software zu reduzieren. Darüber hinaus wurden auch konkrete Hinweise für die Verbesserung der Software gegeben. Außerdem sollte sichergestellt werden, dass die Schüler*innen ausreichend Zeit mit Diskussion, Begründung und Formulierung der Ergebnisse verbringen, bevor diese im Plenum besprochen werden. Die Art des Computereinsatzes spielt in dieser Studie keine Rolle für den Lernerfolg.

12.3 Argumentationsrichtungen

Der erstellte Fragebogen (siehe Kapitel 8.4.3) ist in der Lage, zwischen Items zu unterscheiden, in denen von Kräften auf die Bewegung geschlossen wird, und Items, in denen von der Bewegung auf die Kräfte geschlossen wird. Dass diese Items in einer Faktorenanalyse jeweils auf einen Faktor laden und somit stärker miteinander korrelieren als mit anderen Items, spricht dafür, dass es für Schüler*innen relevant ist, in welche Richtung

argumentiert werden muss. Die Items, in denen von den Kräften auf die Bewegung geschlossen wird, werden häufiger richtig beantwortet als Items, in denen in die andere Richtung argumentiert werden muss. Die Schwierigkeit der Items konnte hier aber nicht kontrolliert werden, sodass dieser Unterschied auch an den Items selbst und nicht an der Argumentationsrichtung liegen könnte. Es liegt aber nahe, davon auszugehen, dass eine Methode, die primär eine dieser beiden Argumentationsrichtungen nutzt, diese auch besonders fördert und verbessert. Die Daten aus den Fragebögen zeigen aber, dass die Ergebnisse in den beiden Inhaltsbereichen unabhängig von der eingesetzten Software sind (siehe Kapitel 9.3.9). Die bevorzugte Argumentationsrichtung von der jeweiligen Software wird also nicht in besonderem Maße trainiert. Eine Analyse der Schülergespräche in den aufgezeichneten Videos kann darüber hinaus zeigen, welche Argumentationen Schüler*innen selbst häufiger nutzen, welche häufiger erfolgreich sind und ob dies von der genutzten Methode abhängt. Es zeigt sich, dass Schüler*innen häufiger Kräfte nutzen, um die Bewegung zu begründen als andersherum. Gleichzeitig ist diese Richtung häufiger erfolgreich, was die Ergebnisse aus dem Fragebogen stützt. Auch hier ist weder in der Häufigkeit noch im Anteil richtiger Begründungen ein Unterschied zwischen den Gruppen festzustellen, was erneut zeigt, dass die eingesetzte Methode keinen Einfluss auf die bevorzugte Argumentationsrichtung hat. Die Daten aus den Videos bestätigen die Ergebnisse aus dem Fragebogen und ergänzen, dass auch die Häufigkeit der Nutzung der verschiedenen Argumentationsrichtungen nicht von der Software abhängt.

IV. Zusammenfassung

13 Zusammenfassung und Diskussion

Schüler*innen haben große Schwierigkeiten beim Erlernen der Newtonschen Gesetze. Dies liegt unter anderem an der Nähe dieses Lerngegenstands zu alltäglichen Erfahrungen und der damit verbundenen Hartnäckigkeit der vorhandenen Schülervorstellungen. Eine Unterrepräsentation von Reibungskräften im Unterricht aufgrund ihrer mathematischen Komplexität könnte ein Grund für unbefriedigende Resultate des Mechanikunterrichts sein, da so die Newtonschen Gesetze im Alltag nicht erkannt werden können und Schüler*innen eher bei einer aristotelischen Sichtweise verharren. Dadurch motiviert wurde in dieser Studie ein Computereinsatz, der eine aktive Auseinandersetzung mit Reibungskräften ermöglicht, im Dynamikunterricht auf seine Wirksamkeit untersucht.

Nach einer kurzen Einführung des fachlichen Hintergrunds der Newtonschen Mechanik (Kapitel 2) wurden die bekannten Schülervorstellungen näher beleuchtet (Kapitel 3). Dazu wurde auch auf Konzeptwechselstrategien eingegangen und aus diesem Wissen erfolgversprechende Faktoren für einen Unterrichtsgang abgeleitet. Da sich die Arbeit mit Computereinsatz in der Newtonschen Dynamik beschäftigt, wird das bekannte Wissen über das Lernen mit neuen Medien im Allgemeinen dargestellt (Kapitel 4). Insbesondere wurden die Cognitive Load Theory und die Multimedia Theory of Cognitive Learning vorgestellt und auch hieraus Design-Empfehlungen für den Computereinsatz im Unterricht gefolgert.

Anschließend wurden bisher durchgeführte Studien mit mathematischer Modellbildung und Videoanalyse vorgestellt (Kapitel 5 und 6). Ein Ergebnis vergangener Studien war, dass frühere Modellbildungsprogramme die Erwartungen teilweise nicht erfüllt haben, es allerdings keine bzw. kaum Studien zu neueren, gleichungsbasierten Modellbildungsprogrammen wie „Newton-II“ gibt. Videoanalyseprogramme wurden bisher fast ausschließlich im Bereich der Kinematik, nicht aber im Bereich der Dynamik eingesetzt.

Ziel der Arbeit war es, die Wirksamkeit des Einsatzes von mathematischer Modellbildung mit aktueller Software zu untersuchen. Da die Videoanalyse zwar insgesamt etablierter im Unterricht ist, aber kaum Daten im Bereich der Dynamik vorliegen, sollte auch die Wirksamkeit des Einsatzes der Videoanalyse im Bereich der Dynamik untersucht werden. Gleichzeitig dient diese Untersuchung auch als Vergleich für die mathematische Modellbildung. Dazu wurden Interventionen entworfen, die mathematische Modellbildung und Videoanalyse für die Vertiefung der Dynamik einsetzen (Kapitel 8). Diese Interventionen sind bis auf die Nutzung der Software analog und an die Kenntnisse von Forschungen im Bereich der Schülervorstellungen, der Konzeptwechselstrategien, der Lerntheorien mit

neuen Medien und der spezifischen Forschung im Bereich der Software angepasst. Sie decken zudem verschiedene Arten von Bewegungen ab (eindimensionale Bewegungen, schräger Wurf, Kreisbewegung), beginnen immer mit einem Realexperiment und thematisieren die vorhandenen Reibungskräfte aktiv (Kapitel 8.3.2). Um die Wirksamkeit in Bezug auf Konzeptverständnis, Modellverständnis, kognitiver Belastung, Sicht der Schüler*innen etc. vergleichen zu können, wurde neben der Nutzung von in der Literatur vorhandenen Testteilen auch ein eigener Test entwickelt und pilotiert (Kapitel 8.4). Der entwickelte Test im Bereich des Konzeptverständnisses ähnelt dabei dem *Force Concept Inventory* (FCI), ist aber in der Lage reliabel zwischen verschiedenen Inhaltsbereichen zu unterscheiden, die relevant für die Forschungsfragen dieser Studie sind („1. Newtonsches Gesetz“, „Von Kräften auf die Bewegung schließen“ und „Von der Bewegung auf die Kräfte schließen“).

Darüber hinaus wurden Bildschirmvideos von einigen der Proband*innen mit Aufzeichnung der Gespräche angefertigt, um einen Einblick in die Arbeitsweise der Schüler*innen zu erhalten und Schwierigkeiten im Umgang mit der Software zu identifizieren (Kapitel 8.6). Erhoben wurden die Daten in der elften Klassenstufe des hessischen Gymnasiums, die nach dem Unterricht der Newtonschen Gesetze in das Schülerlabor der Goethe-Universität kamen und einen Vormittag an einer entsprechenden Intervention teilnahmen.

13.1 Wirksamkeit der durchgeführten Interventionen

Es zeigte sich, dass beide Interventionen wirksam für das Vertiefen der Newtonschen Dynamik sind. Die Gruppe, die mathematische Modellbildung nutzte ($N = 137$), erreichte durchschnittlich 5,29 Punkte (von max. 15) im Vortest und 8,58 Punkte im Nachtest. Dieser Lernzuwachs ist höchst signifikant von null verschieden und entspricht einer großen Effektstärke von $d_z = 1,04$. Dies ist erfreulich, da die Interventionen nach dem Unterricht der Newtonschen Gesetze stattfanden. Die Schüler*innen verbesserten sich in allen Inhaltsbereichen der Dynamik („1. Newtonsches Gesetz“, „Von Kraft auf Bewegung schließen“, „Von Bewegung auf Kraft schließen“) mit jeweils mittleren bis großen Effektstärken. Auch im Inhaltsbereich „Beschleunigungsdiagramme“ trat eine Verbesserung ein, die allerdings erwartungsgemäß geringer ausfiel als in den anderen Inhaltsbereichen, da die Dynamik der Fokus der Intervention war (Kapitel 9.2).

In der Gruppe der Videoanalyse verbesserte sich das Ergebnis von durchschnittlich 4,79 Punkten im Vortest zu 7,99 Punkten im Nachtest. Dies entspricht ebenfalls einer signifikant, von null verschiedenen Steigerung und einer großen Effektstärke von $d_z = 1,07$. Auch die Gruppe der Videoanalyse verbesserte sich in allen Inhaltsbereichen (s. o.). Auch

hier war die Verbesserung in dem Bereich der Beschleunigungsdiagramme am geringsten. Die Interventionen, die an die bekannten Schülervorstellungen angepasst waren und Reibung aktiv mithilfe des Computers thematisierten, lassen sich als erfolgreich für das Vertiefen der Newtonschen Dynamik einstufen.

13.2 Vergleich der durchgeführten Interventionen

Ein Vergleich der beiden Interventionen zeigt, dass es keinen Unterschied zwischen den Gruppen im Konzeptverständnis gibt. Die beiden Gruppen mit je $N = 137$ Proband*innen unterscheiden sich zu Beginn in keiner der erhobenen Kontrollvariablen (Fachinteresse, fachspezifisches Selbstkonzept, Interesse an Computer, Interesse am Experimentieren, Interesse an theoretischen Zusammenhängen, Physiknote, Mathematiknote) und auch die Vortestergebnisse sind ähnlich. Ein direkter Vergleich der Nachtestergebnisse deutet daraufhin, dass es keinen Unterschied im Lernzuwachs gibt. Verschiedene Methoden (ANOVA mit Messwiederholung, Parallelisierung, *Propensity Score Matching*, ANCOVA, hierarchisches lineares Modell) bestätigen dies (Kapitel 9.3). Die ANCOVA, in der das Nachtestergebnis bei gleichem Vortestergebnis durch ein lineares Modell geschätzt wird, geht bei einem Vortestergebnis von 5,10 Punkten in beiden Gruppen von einem Nachtestergebnis von 8,41 Punkten in der Gruppe der mathematischen Modellbildung und einem Ergebnis von 8,11 Punkten in der Gruppe der Videoanalyse aus. Es ist also nicht von einem signifikanten Unterschied zwischen den Gruppen im Gesamtergebnis auszugehen.

Es gibt aber einen Unterschied in einem der Inhaltsbereiche. Während die Bereiche „Von Kraft auf Bewegung schließen“ und „Von Bewegung auf Kraft schließen“ sich gleichermaßen verbessern, erhöht sich die Punktzahl im Bereich „1. Newtonsches Gesetz“ in der Gruppe der Modellbildung stärker. In diesem Inhaltsbereich schätzt die ANCOVA, dass die Gruppe der Modellbildung im Nachtest 51,8 % der Fragen richtig beantwortet, während die Gruppe der Videoanalyse 43,1 % der Items korrekt ankreuzt – bei einem jeweiligen Vortestergebnis von 24,8 %. Bei näherer Betrachtung wird deutlich, dass sich der Anteil der Schüler*innen, die im Test die häufige Schülervorstellung aktivieren, dass eine Kraft in Bewegungsrichtung wirken muss, damit sich etwas bewegen kann, in der Gruppe der Modellbildung stärker zurückgeht. Dies könnte die Begründung für die größere Verbesserung im Inhaltsbereich „1. Newtonsches Gesetz“ sein, da hier stets von der bekannten Geschwindigkeit auf die Kräfte geschlossen werden muss und die genannte Vorstellung hier eine falsche Antwort hervorruft. Die gleiche Verbesserung in den anderen beiden Inhaltsbereichen spricht dafür, dass es keine Rolle spielt, welche Argumentationsrichtung die verwendete Methode nutzt. Obwohl bei mathematischer Modellbildung eher

von den wirkenden Kräften ausgehend die Bewegung erklärt wird und bei der Videoanalyse ausgehend von den Messdaten die Kräfte hergeleitet werden, scheint sich dies nicht auf die Lösungswahrscheinlichkeiten der entsprechenden Items im Test zu übertragen (Kapitel 9.3.9).

Zudem wurde das Modellverständnis der Proband*innen im Vor- und im Nachtest erhoben. Die Items bilden dabei einen kleinen Teil des gesamten Modellverständnisses ab, was wiederum Teil der Modellkompetenz ist. In der Gruppe der Modellbildung verändert sich das Modellverständnis höchst signifikant mit einer mittleren Effektstärke von $d_z = 0,43$ in die gewünschte Richtung, während sich in der Gruppe der Videoanalyse erwartungsgemäß nichts ändert. Die Eigenschaften von Modellen sind in keiner der Interventionen explizites Thema. In der Modellbildung führt der Umgang mit Modellen aber scheinbar zu einer Änderung der Ansichten über Modelle (Kapitel 10.1).

Vergangene Forschungen brachten die Videoanalyse in Verbindung mit einer Reduktion der kognitiven Belastung im Vergleich zu herkömmlichem Unterricht, während in Studien zum Einsatz von mathematischer Modellbildung teilweise eher von Schwierigkeiten berichtet wurde – wenngleich im zweiten Fall nicht die kognitive Belastung erhoben wurde. Ein Vergleich der beiden Interventionen bestätigt die Annahme, dass die wahrgenommene kognitive Belastung bei den Schüler*innen geringer war, die an der Intervention zur Videoanalyse teilgenommen haben. Die Unterschiede sind dabei nur in der intrinsischen kognitiven Belastung signifikant ($d = 0,40$), während es bei der extrinsischen und der lernbezogenen kognitiven Belastung keine Unterschiede zwischen den Gruppen gab (Kapitel 10.3).

In der Wahrnehmung der Schüler*innen war außerdem die Software „measure dynamics“ für die Videoanalyse besser als „Newton-II“ für die Modellbildung, was sich auch durch die Häufigkeit und Dauer der Schwierigkeiten in den beobachteten Arbeitsphasen bestätigte. In den Bereichen „Spaß an der Intervention“ und „Relevanz des behandelten Themas“ gab es allerdings keine Unterschiede zwischen den Gruppen (Kapitel 10.2).

13.3 Einflussfaktoren auf die Änderung des Konzeptverständnisses

Da verschiedene Kovariablen erhoben wurden, konnten auch Modelle erstellt werden, die den Einfluss dieser Variablen auf das Nachtestergebnis ermitteln (Kapitel 9.3.8). Wie bereits beschrieben hat die Gruppenzugehörigkeit keinen Einfluss auf das Nachtestergebnis und damit auf den Lernzuwachs. In einem linearen hierarchischen Modell wurde untersucht, welche anderen Variablen auf Schüler- und auf Klassenebene zur Varianzaufklärung des Nachtestergebnisses beitragen. Den größten Einfluss auf das Nachtestergebnis

hat das Vortestergebnis, was bei einer eintägigen Intervention erwartbar ist ($f^2 = 0,41$). Weiterhin haben nach diesem Modell auch das Interesse an theoretischen Zusammenhängen ($f^2 = 0,06$), die Physiknote ($f^2 = 0,06$) und die kognitive Belastung ($f^2 = 0,08$) einen Einfluss auf das Nachtestergebnis ($f^2 \geq 0,02$ kleiner Effekt, $f^2 \geq 0,15$, mittlerer Effekt, $f^2 \geq 0,35$ großer Effekt (Cohen, 1988)). Die kognitive Belastung mit der größten Varianzaufklärung ist die Kombination aus der intrinsischen und der extrinsischen kognitiven Belastung. Diese beeinflusst das Nachtestergebnis negativ.

Es gibt weitere Größen, deren Einfluss auf das Nachtestergebnis nicht nachgewiesen werden konnte, was ebenfalls nennenswert ist. Zunächst scheint weder das Interesse am Experimentieren noch das Interesse an Computern oder das allgemeine Fachinteresse das Nachtestergebnis zu beeinflussen. Insbesondere das Interesse an Computern sollte beim Computereinsatz in der Schule nach Möglichkeit nicht relevant sein, um Schüler*innen, die privat keinen Computer nutzen, nicht zu benachteiligen. Das Geschlecht beeinflusst das Nachtestergebnis unter Kontrolle des Vortestergebnisses nur dann, wenn keine weiteren Variablen in das Modell hinzugefügt werden. Bei Hinzunahme der genannten Variablen spielt das Geschlecht keine Rolle mehr. Es zeigt sich, dass sich die Geschlechter insbesondere in fachspezifischem Selbstkonzept und Physiknote unterscheiden, die beide eine bessere Varianzaufklärung liefern als das Geschlecht selbst. Dass die Mathematiknote keinen Einfluss auf das Nachtestergebnis hat, ist erfreulich, da der Computereinsatz insbesondere von der Mathematik entlasten und diese keine Gelingensbedingung darstellen sollte. Zudem hatte keine der erhobenen Lehrervariablen einen Einfluss.

Ergänzt wurde die beschriebene Mehrebenenanalyse durch die Analyse der Schülergespräche in den aufgenommenen Bildschirmvideos mit Ton ($N = 45$). Ein Vergleich zwischen Schüler*innen, die viel dazu lernten, und Schüler*innen, die wenig lernten, machte deutlich, dass vor allem die Phase der Ergebnisbesprechung wichtig ist und dass eine intensive Diskussion der relevanten Größen mit einem höheren Lernzuwachs korreliert. In der Gruppe der Videoanalyse beeinflussten zudem Schwierigkeiten mit der Software den Lernzuwachs negativ. Keinen Einfluss scheinen inhaltliche Schwierigkeiten und die Äußerung von Schülervorstellungen zu haben. Die Datenmenge ist bei der Analyse der Videos allerdings geringer (Kapitel 11.5).

13.4 Weitere Beobachtungen aus den aufgezeichneten Schüleraktivitäten

Über die Einflussfaktoren auf den Lernzuwachs hinausgehend wurden in den Bildschirmvideos zusätzliche Aspekte beobachtet und analysiert. Zunächst wurde untersucht, welche Schwierigkeiten sowohl in Zusammenhang mit der Software als auch inhaltlich auftreten.

Die inhaltlich größte lernhinderliche Schwierigkeit in beiden Interventionen war eine mangelnde Diagrammkompetenz. Da ein Verständnis von Diagrammen für die Nutzung von Programmen dieser Art notwendig ist, sollte dieses nach Möglichkeit bereits vor der Nutzung vorhanden sein oder mit der Software verbessert werden. Zu den aufgetretenen inhaltlichen Schwierigkeiten wurden in den meisten Fällen Verbesserungsvorschläge für die Lehrkraft bzw. die Programmerstellenden gemacht. Es wurden keine so fundamentalen Schwierigkeiten festgestellt, dass der Einsatz der Software dadurch verhindert werden würde (Kapitel 11.1).

Ergänzend zu den Schülervorstellungen, die im Fragebogen auftraten, wurde auch in den Videos nach Schülervorstellungen geschaut. Bestätigt wurde die Analyse aus dem Fragebogen, dass die Schülervorstellung „eine Kraft muss in Bewegungsrichtung wirken“ am häufigsten aktiviert wird. Bei dieser Schülervorstellung wurde auch ein Unterschied zwischen den Gruppen festgestellt. Proband*innen der mathematischen Modellbildung äußerten diese Schülervorstellung deutlich häufiger. Da die Schülervorstellung sich in dieser Gruppe von Vor- zu Nachtest gleichzeitig stärker reduziert, liegt die Vermutung nahe, dass die Reduktion in der aktiveren Auseinandersetzung mit der Vorstellung in der mathematischen Modellbildung begründet liegt. Der Notwendigkeit der mathematischen Formulierung dieser vermeintlich in Bewegungsrichtung wirkenden Kraft und der nachfolgenden Rückmeldung, dass das Modell ohne diese Kraft optimal zu den Daten passt, scheint die Schülervorstellung stärker zu reduzieren, als das bei der Videoanalyse der Fall ist (Kapitel 11.3).

Außerdem wurden die Argumentationen der Schüler*innen während der Arbeitsphase analysiert, um herauszufinden, ob sich die Argumentationsrichtung der eingesetzten Methode auf die Proband*innen überträgt. Insgesamt wird häufiger von Kraft in Richtung Bewegung argumentiert als andersherum. Es gibt aber keinen Unterschied zwischen den Gruppen. Die Hypothese, dass dies von der Methode abhängt, kann also nicht bestätigt werden. Es bestätigt sich allerdings, dass es Schüler*innen leichter zu fallen scheint, von Kräften zur Bewegung zu argumentieren, da dies freiwillig häufiger gewählt wird und auch anteilig häufiger zum richtigen Ergebnis führt (Kapitel 11.2). Darüber hinaus wurde in Kapitel 11.4 die Vorgehensweise der Proband*innen analysiert und beschrieben. Während die meisten Gruppen beim Modellieren einer sinnvollen Vorgehensweise folgen, gibt es auch Gruppen, die die externen Messdaten zu Beginn einlesen und das Modellieren eher einem willkürlichen Ausprobieren als zielgerichtetem Vorgehen gleicht. Sowohl bei der Videoanalyse als auch bei der Modellbildung werden häufig die am Computer generierten Daten nicht zur Beantwortung der Abschlussfragen genutzt.

13.5 Diskussion der Ergebnisse

Die Ergebnisse dieser Studie fügen sich in Teilen in bereits bekannte Forschungsergebnisse ein. Da bekannt ist, dass sich die Videoanalyse gut für den Einsatz im Bereich der Kinematik eignet, war dies auch für die Dynamik zu erwarten. Dies konnte durch diese Arbeit bestätigt werden. Videoanalyse scheint also in allen Bereichen der Mechanik eine sinnvolle Ergänzung für den Physikunterricht zu sein. Bekannt war außerdem, dass Videoanalyse auf Mobilgeräten gut für den Mechanikunterricht geeignet ist. Dies scheint auch für Laptops der Fall zu sein. Dass Videoanalyse einen vergleichsweise geringen Cognitive Load verursacht und von Lernenden äußerst positiv bewertet wird, deckt sich ebenfalls mit der Literatur.

Unklarer war, ob andere als die bisher untersuchte Software im Bereich der mathematischen Modellbildung dazu führen kann, dass die Schwierigkeiten bei der Benutzung reduziert werden können und mit der Software ein erfolgreicher Unterricht möglich ist. Dies konnte durch die Studie gezeigt werden. Es ist davon auszugehen, dass die von anderen Autor*innen berichteten Schwierigkeiten beim Einsatz von mathematischer Modellbildung in der Schule mit Software wie „Newton-II“ geringer sein könnten. Hinsichtlich des Konzeptverständnisses ist die mathematische Modellbildung in dieser Studie ebenso gut wie die Videoanalyse. Im Gegensatz zu anderen Studien in der Vergangenheit wurde hier aber nur nach dem Konzeptverständnis der Dynamik und nicht nach überfachlichen Kompetenzen wie etwa dem systemischen Denken geschaut. Nichtsdestotrotz verursacht die mathematische Modellbildung auch in dieser Studie mehr unterschiedliche Schwierigkeiten mit der Software, welche zudem schlechter bewertet wurde als die Videoanalysesoftware. Die kognitive Belastung war bei der mathematischen Modellbildung ebenfalls höher, was vor allem an der Komplexität der verwendeten Formeln (für die Reibung) lag.

Die Durchführung der Studie war ursprünglich mit etwa 630 Schüler*innen geplant, mit denen bereits Termine ausgemacht waren. Durch die Schulschließungen im Schuljahr 2020/2021 im Zuge der Covid-19-Pandemie konnten nur etwa die Hälfte teilnehmen. Durch fehlende oder unvollständig ausgefüllte Fragebögen blieben so noch 274 Schüler*innen, deren Daten ausgewertet werden konnten. Diese Reduktion verringert auch die statistische Belastbarkeit der gefundenen Ergebnisse. Mit dieser Stichprobengröße lassen sich jedoch mittlere bis große Effektstärken gut nachweisen (Kapitel 9.1). Subgruppenanalysen, wie sie teilweise auch durchgeführt wurden, reduzieren die Daten aber weiterhin. Bildschirmvideos wurden nur von einer Teilgruppe der Teilnehmenden ausgewertet. Die in dieser Kategorie berichteten Ergebnisse unterliegen einerseits einer größeren

Subjektivität, was in der Forschungsmethode begründet liegt, und sind andererseits noch mehr abhängig von der zufälligen Auswahl der Proband*innen. Dass sich viele der Ergebnisse mit denen aus den Fragebögen decken bzw. diese ergänzen, bestätigt deren Validität.

In der vorgestellten Studie handelt es sich um eine quasi-experimentelle Laborstudie. Weiterer Forschungsbedarf besteht darin, zu untersuchen, ob sich die Ergebnisse in einer Feldstudie, die sich durch eine höhere externe Validität auszeichnet, reproduzieren lassen. Dazu müsste ein Unterrichtskonzept, welches konsequent die untersuchten Methoden einsetzt und im normalen Physikunterricht stattfindet, mit herkömmlichen Unterrichtsmethoden verglichen werden. Sinnvoll wäre es zudem, entsprechende Software auf Mobilgeräten zu untersuchen, da diese in Schulklassen in der Regel eher verfügbar sind. Ein signifikanter Unterschied der Ergebnisse bei der Nutzung von Tablets ist jedoch nicht zu erwarten.

13.6 Fazit

In der Studie ist es gelungen, erfolgreiche Interventionen mit mathematischer Modellbildung und Videoanalyse durchzuführen. Die untersuchten Methoden unterschieden sich nicht im Lernzuwachs des Konzeptverständnisses der Dynamik. Aufgrund der Größe der Veränderung von Vor- zu Nachtest lassen sich beide Methoden als erfolgreich und sinnvoll für das Vertiefen der Newtonschen Dynamik einordnen. Da speziell mathematische Modellbildung mit den entsprechenden Programmen seltener in der Schule vertreten ist und das Modellieren eine wichtige naturwissenschaftliche Kompetenz ist, wird dafür plädiert, entsprechende Software häufiger im Unterricht einzusetzen. Zusätzlich verbessert sich das Modellverständnis der Schüler*innen bei Nutzung des mathematischen Modellbildungsprogramms ohne explizite Thematisierung von Modelleigenschaften und der Anteil der Schüler*innen mit aristotelischer Sichtweise verringert sich stärker. Mathematische Modellbildung verursacht allerdings auch eine höhere kognitive Belastung bei Schüler*innen, die den Lernerfolg negativ beeinflusst. Dies muss von der Lehrkraft berücksichtigt werden. Daraus lässt sich nicht schließen, dass die mathematische Modellbildung nur mit leistungsstärkeren Schüler*innen genutzt werden sollte. Entsprechende Hilfestellungen scheinen jedoch notwendig. Der Einsatz im Unterricht vor allem auch bei Erlernen und nicht nur Vertiefen der Newtonschen Gesetze sollte gesondert untersucht werden.

Es ist bekannt, dass sich die Videoanalyse gut im Bereich der Kinematik eignet. Da diese meist vor der Dynamik unterrichtet wird, bietet es sich an, bereits hier die Nutzung der Software einzuüben und diese dann ebenfalls im Bereich der Dynamik einzusetzen.

Videoanalyse und mathematische Modellbildung lassen sich ebenfalls gut kombinieren und somit im gesamten Schuljahr einsetzen. Unabhängig von den Unterschieden zwischen den Methoden scheint ein Unterricht mit Computereinsatz, der auf bekannte Schülervorstellungen ausgelegt ist und Reibungskräfte aktiv thematisiert, erfolgreich darin zu sein, das Verständnis der Newtonschen Gesetze deutlich zu verbessern und unerwünschte Schülervorstellungen zu reduzieren.

14 Summary and Discussion

Students have many difficulties understanding the Newtonian laws of motion. This is partly due to the fact, that Newtonian mechanics explain motions that are observable in everyday life. Therefore, many preconceptions have formed over the course of many years and are known to be especially persistent. One reason for the disappointing results of physics lessons in this area could be due to disregarding friction because of its mathematical complexity. If friction is not accounted for, the Newtonian laws of motion are not visible in everyday life. Students therefore are more likely to keep their Aristotelian views. Motivated by that, the presented study analyses two different modes of using computers in physics lessons, which both allow active discussion and analysis of frictional forces without significantly increasing mathematical complexity.

This thesis briefly presents the relevant aspects of Newtonian dynamics (Chapter 2) before discussing the known preconceptions in the field of mechanics (Chapter 3). Closely linked to the preconceptions, conceptual change strategies are presented with the intention to derive requirements for a successful conceptual change. Due to the focus on computer usage, the following chapter summarises the research around digital media (Chapter 4), especially focusing on the Cognitive Load Theory and the Multimedia Theory of Cognitive Learning to derive design choices for the interventions used in this research.

The research on computational modelling and video motion analysis is presented in the following two chapters (Chapters 5 and 6). Studies, where computational modelling was used in German schools, showed some problems with the software used and not all the expectations were met. There are only few studies that used equation-based modelling software like “Newton-II”, though. Video motion analysis on the other hand is associated with an increase in learning gain and a reduction in cognitive load compared to classical approaches. Video motion analysis was mostly researched in kinematics and not in dynamics.

The goal of the presented study is, therefore, to investigate the effectiveness of using computational modelling with newer software. Although video motion analysis is more established in German schools, there is little data regarding its effects on the learning gain in dynamics. Thus, the use of video motion analysis was investigated as well. At the same time, it acts as a comparative group to the computational modelling cohort.

Two interventions were designed that make use of computational modelling and video motion analyses respectively and are carried out after the Newtonian laws of motion were discussed in school. Both interventions cover identical physical situations and have

analogous tasks for learners to complete, only differing in the software usage. They are designed based on the findings linked to alternative conceptions, conceptual change, learning theories with multimedia tools and the specific use of computational modelling and video motion analysis. Different types of motion (i.e. one-dimensional motions, projectile motion, circular motion) are being discussed in the interventions and every one of the four parts starts with an experiment, which is then modelled or analysed by the students in groups of two.

To analyse the effect of the two methods on conceptual understanding, the views about models, cognitive load and other affective variables, a test was created and piloted that consisted in part of already established tests and in part on newly created items (Chapter 8). The conceptual understanding part of the test is partly based on the Force Concept Inventory but is able to reliably distinguish between different content areas (e.g. “Newtons 1. law”, “Arguing from forces to motion”, “Arguing from motion to forces”).

The research questions (Chapter 7.2) are in part answered by analysing the answers given in the test. Additionally, the screens of some of the students in combination with their conversations were recorded to be able to study the way the participants use the software and to identify difficulties (Chapter 8.6). The sample of students consist of $N = 274$ 11th graders from Hessian secondary schools that visit the university after their lessons of Newton’s laws of motion.

14.1 Effect of the interventions

The study shows that both interventions were successful in improving the conceptual understanding of Newtonian dynamics. The cohort that used computational modelling scored on average 5,29 points (from max. 15 points) in the pre-test and 8,58 points in the post-test. This corresponds to a highly significant learning gain with a large effect size of $d_z = 1,04$. This is especially notable since the intervention took place after the respective lesson in school. The improvement was greatest in the content areas of dynamics (i.e. “Newtons 1. law”, “Arguing from forces to motion”, “Arguing from motion to forces”) with medium to large effect sizes. The conceptual understanding of acceleration graphs improved with a smaller effect size. This was expected since acceleration graphs are used to analyse the model and the data from the experiments but were not the focus of the intervention.

The cohort that used video motion analysis improved from an average score of 4,79 points in the pre-test to an average score of 7,99 points in the post-test. This corresponds to a highly significant improvement with a large effect size of $d_z = 1,07$. The cohort also

improved in all content areas and the improvement was largest in areas regarding Newtonian dynamics and smaller in the area from kinematics (i.e. “acceleration graphs”). It can be summarised that both interventions, which are adapted to the common misconceptions in the field of Newtonian dynamics and use the computer to actively discuss frictional forces, are successful in improving the conceptual understanding.

14.2 Comparison of the two interventions

A comparison shows no difference between the groups regarding the overall conceptual understanding of Newtonian dynamics. The two cohorts with $N = 137$ participants respectively are not different in the measured control variables (i.e. interest in physics, physics related self-concept, interest in computers, interest in experimenting, interest in theoretical relationships, physics grade, maths grade). There is also no difference in the average score in conceptual understanding in the pre-test. The similar scores in the post-test therefore suggest that there is no difference in the learning gain between the two groups. This result was confirmed with different statistical methods (i.e. ANOVA with repeated measures, matched samples, propensity score matching, ANCOVA, linear hierarchical model; Chapter 9.3). The ANCOVA estimates an average score of 8,41 points in the post-test for the computational modelling cohort and 8,11 points for the video motion analysis cohort, if both groups started with a score of 5,10 points in the pre-test. This difference is not significant. Thus, the data presents no reason to expect a difference in learning gain between the two methods in the field of Newtonian dynamics.

Though there are no significant differences overall, differences in one of the content areas were found. The computational modelling intervention improved learners’ score in items regarding Newton’s first law of motion significantly more than the video analysis intervention. The ANCOVA estimates a mean of 51,8 % correct answers in the post-test in the computational modelling cohort and 43,1 % in the video motion analysis starting with 24,8 % in the pre-test, which is a significant difference with a small effect size. A further analysis shows that the share of students who activate the preconception that a force in the direction of motion is necessary, is reduced further in those who used computational modelling. This may explain the larger improvement in items regarding Newton’s first law, since students had to derive the acting forces knowing the constant velocity in all items. An Aristotelian view on forces would generate false answers in this case. There were no differences in the content areas, where students had to argue from forces to motion or vice versa. Though there are more correct answers overall in items of the first direction, no significant differences between the groups were found. The preferred direction of argumentation (computational modelling starts with the acting forces and explains

the motion whereas video motion analysis uses the data of the motion to derive the acting forces) does not influence the performance of the students in the respective items.

The understanding of models of the participants was also analysed and was found to positively change in the computational modelling cohort with an effect size of $d_z = 0,43$, while it remained nearly unchanged in the video motion analysis cohort. The properties of models were not explicitly discussed in either of the two interventions. The process of modelling nevertheless changed the students' views about models positively.

Research already linked video motion analysis to a reduction in cognitive load compared to a traditional approach in kinematics. Computational modelling on the other hand caused more difficulties. Although the cognitive load was not measured in these cases an increase of the cognitive load due to the reported difficulties would be possible. It is therefore sensible to compare the two methods regarding the cognitive load. The data suggests that the cognitive load is indeed significantly higher in the computational modelling cohort. Especially the intrinsic cognitive load is different with an effect size of $d = 0,40$, whereas the extraneous and germane cognitive load were not significantly different.

The video motion analysis software “measure dynamics” was rated significantly higher by the participating students than the software “Newton-II” for computational modelling and students seem to have more problems using “Newton-II”. This impression is confirmed in the analysis of the screen recordings. There were no differences between the cohorts in the areas “relevance of the topic” and “fun during the intervention”.

14.3 Predictors of learning gain in conceptual understanding

Several covariates were measured in the test. With those, a hierarchical linear model (HLM) was created that takes the hierarchical structure of students grouped into classes into account and estimates which covariates can explain a significant portion of the variance in the model (Chapter 9.3.8). As discussed, the group variable (computational modelling/video motion analysis) has no influence on the outcome of the post-test result. The pre-test result has the biggest influence ($f^2 = 0,41$). This is no surprise considering that these are one-time interventions. The HLM predicts that also the interest in theoretical relationships/concepts ($f^2 = 0,06$), the last physics grade ($f^2 = 0,06$) and the cognitive load ($f^2 = 0,08$) influence the post-test score ($f^2 \geq 0,02$ small effect, $f^2 \geq 0,15$, medium effect, $f^2 \geq 0,35$ large effect (Cohen, 1988)). The cognitive load that explains the most variance is the combination of intrinsic and extraneous cognitive load and influences the score negatively.

No other covariate has a significant influence on the post-test score. Neither the interest in experiments, nor the interest in physics or the computer affinity influences the post-test score. An influence of the interest in computers on the post-test score would favour students that use a computer in their spare time which is not desirable. Learners' gender is only relevant for the learning gain if no other variable is included in the model. The physics grade and the physics-related self-concept better explained the variance in the model. The different genders differ in these variables. Differences in the learning gain are therefore not best explained by gender itself, but rather by the differences in certain variables that are relevant for the learning gain. The maths grade on the other hand does not influence learning gain. The used software tried to reduce mathematical complexity so that students can focus on the underlying physics. The maths grade should therefore not have an influence on the performance in the interventions, which is confirmed by the model.

Additional to the hierarchical linear model, the screen recordings in combination with a recording of the dialogues were analysed to compare successful students (in term of learning gain) with less successful ones. It was shown that successful students spent more time discussing the relevant quantities and more often use the model or the data to formulate results. In the video motion analysis cohort, difficulties using the software seem to influence the learning gain negatively. The activation of certain preconceptions and difficulties regarding the underlying physics (such as graph comprehension) did not have a measurable influence on the result. The number of learners for which screen recordings were analysed is less than the test scores, see Chapter 11.5.

14.4 Further findings from the screen recordings and conversations

Additional to possible predictors of a large learning gain in conceptual understanding, more aspects were analysed in the screen recordings. Firstly, difficulties with the underlying physics and the software were identified. The most common difficulty regarding physics was the misinterpretation of the different graphs that must be analysed during the intervention. These difficulties were common in both groups. Although there were no differences between successful and less successful students in that regard, a lack of understanding of graphs could hinder the success of such an intervention. Therefore, the ability to read and understand graphs should be trained prior to the usage of the software, or with it (especially with video motion analysis software). Recommendations were made to developers to improve parts of the software, and to teachers to take the difficulties of students during the lesson into account to reduce them. No serious difficulties were found that would prohibit the use of the software in school entirely.

Additional to the preconceptions that were studied by examining the answers in the test, the dialogues between the students were analysed. The most common preconception was once again, that a force in the direction of motion is necessary for an object to move. A difference in the frequency of the preconception was found between the two groups. Students using computational modelling software express this preconception more often. This can be explained by the modelling process where the forces must be mathematically and explicitly formulated. In the case of the recorded projectile motion with non-significant friction there is no force in the horizontal direction. Students have, therefore, difficulties formulating the “movement force” and see that the model fits the real-world data well without a force in the horizontal direction. The preconception is therefore being expressed more often, and probably therefore changed to a higher degree due to the cognitive conflict, than in the video motion analysis cohort. This explains the larger improvement in items regarding Newton’s first law discussed in the previous chapter.

The way students argue when talking about physical quantities was also analysed. Similar to the data of the test, there were no differences between the groups in that regard. Overall, students more frequently used forces to derive properties of the motion than the other way around. A larger share of this kind of argumentation leads to the correct answer as well. The way students work with the respective software is also analysed in Chapter 11.4. Although most of the students used a sensible approach while modelling, some groups first imported the external data and more randomly guessed while modelling rather than using a target-oriented approach. In both cohorts, many students did not use the model or the data to answer the questions in the end.

14.5 Discussion

Many of the results of this study tie in older research results. It was known that video motion analysis is successful in teaching kinematics. This study showed that this is also the case for dynamics. Video motion analysis seems to be a sensible addition physics class regardless of the area of Newtonian mechanics and the type of device it is used on. This study also confirms older research in finding a lower cognitive load in the video motion analysis cohort.

It was unclear though, if newer computational modelling software can solve some of the issues that arose in the past. The study confirmed that this is the case. Regarding learning gain in conceptual understanding of Newtonian dynamics, computational modelling was as good as video motion analysis. In contrast to some prior studies, no overarching competences like systemic thinking were analysed, though. Computational modelling caused,

however, more different types of difficulties linked to the software, which was also rated lower than the video motion analysis software by participating students. Additionally, the cognitive load was higher in the computational modelling cohort, which was mainly caused by the complexity of the used formulas (for friction).

The study was originally planned and already scheduled with around 630 students. Due to school closures during the Covid-19 pandemic in the school year 2020/2021, only about half of those could participate. Due to missing or incomplete tests, there were $N = 274$ participants remaining. This reduction in the number of participants also reduces statistical power and robustness of the data. As discussed in a power analysis, medium to large effect sizes can still be shown well with the number of participants. Some analyses of subgroups further reduce the available data, though. Due to time constraints, the number of screen recordings that could be analysed, was even smaller. Many of the findings in those screen recordings correspond well to the findings from the test, though, confirming their validity to some degree.

The study is designed as a quasi-experimental laboratory study. It would be reasonable to examine in a future study if similar results are found in a field study, which has a higher external validity. For that, a teaching concept would have to be designed, which consistently uses the software and takes place in the physics class in school. The results of this concept could then be compared to a classical approach. It would be sensible to use mobile devices such as tablets since they are normally much more available than desktop computers or laptops. It is to be expected that the results would not vastly differ in comparison to the usage of laptops.

14.6 Conclusion

The study showed that it is possible to design successful interventions using computational modelling and video motion analysis that are based on known preconceptions and actively discuss frictional forces to bridge the gap between Newtonian mechanics and everyday life of students. From the changes between pre-test and post-test score, both methods can be seen as successful in improving conceptual understanding of Newtonian dynamics. Because especially computational modelling is used less frequently in Hessian schools and modelling itself is an important scientific skill, an increased usage of such software is recommended. Additionally, the views about models change while using computational modelling software without explicitly discussing them. The share of students, which activated the most common preconception, that a force in the direction of motion is necessary for an object to move at all, is reduced further in the cohort of computational

modelling than in the cohort of video motion analysis, which is another reason for its usage. It does, however, cause a higher cognitive load in students, which negatively affects the learning gain. This should be considered by the teacher. The data does not imply though, that computational modelling is only beneficial for high-performing students. Certain tips and aids are necessary, especially if students are supposed to work with the software autonomously. The usage of the software in the introductory phase of Newtonian dynamics must be studied separately since this study took place after the initial lessons.

Because video motion analysis is successful in teaching kinematics and dynamics it is reasonable to use it throughout the whole school year, which would also reduce the proportion of time to get accustomed to the software. Video motion analysis can also be combined with computational modelling. Regardless of the specific method, dynamics lessons that use computers to include frictional forces and are based on known preconceptions seem to be successful in improving the conceptual understanding of Newtons laws of motion.

15 Danksagung

Abschließend möchte ich mich bei einigen Personen bedanken, die direkt oder indirekt zur Entstehung dieser Arbeit beigetragen haben.

- Dankbar bin ich meinen Eltern, Irene und Frank Weber, die mich stets bei allem unterstützten.
- Ich bedanke mich bei Linda Carstens für die generelle Unterstützung in den letzten Jahren und für das Korrekturlesen der Arbeit.
- Vielen Dank an alle Lehrkräfte, die mit ihren Klassen an der Studie teilgenommen und so dafür gesorgt haben, dass eine empirische Überprüfung der erstellten Interventionen stattfinden könnte.
- Ebenso großer Dank gilt auch den Schüler*innen, die bei den Interventionen mitgemacht und damit die Forschung ermöglicht haben.
- Außerdem danken möchte ich allen Kolleg*innen und dem gesamten Institut für eine angenehme Arbeitsatmosphäre und viele interessante Gespräche und Diskussionen.
- Weiterer Dank gilt der gemeinnützigen Stiftung Giersch, die die Forschung durch ihre finanzielle Unterstützung des Physik-Schülerlabors der Goethe-Universität Frankfurt in der Form erst möglich gemacht hat.
- Ein großer Dank geht vor allem an Prof. Dr. Thomas Wilhelm, welcher bei Fragen immer erreichbar war und sich zum Beantworten dieser viel Zeit nahm. Die fachliche und kollegiale Unterstützung und Betreuung waren stets außergewöhnlich.

16 Literaturverzeichnis

- Adams, W. K., Perkins, K. K., Podolefsky, N. S., Dubson, M., Finkelstein, N. D. & Wieman, C. E. (2006). New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey. *Physical Review Special Topics - Physics Education Research*, 2(1), 157. <https://doi.org/10.1103/PhysRevSTPER.2.010101>
- Aiken, J. M., Caballero, M. D., Douglas, S. S., Burk, J. B., Scanlon, E. M., Thoms, B. D. & Schatz, M. F. (2012). Understanding Student Computational Thinking with Computational Modeling. *AIP Conference Proceedings* 1513, 46–49.
- Aiken, L. S. & West, S. G. (2010). *Multiple regression: Testing and interpreting interactions* [Reprinted]. Sage.
- Alonzo, A. C. & Aufschnaiter, C. von (2018). Moving Beyond Misconceptions: Learning Progressions as a Lens for Seeing Progress in Student Thinking. *The Physics Teacher*, 56(7), 470–473. <https://doi.org/10.1119/1.5055332>
- Ambrosis, A. de, Malgieri, M., Mascheretti, P. & Onorato, P. (2015). Investigating the role of sliding friction in rolling motion: a teaching sequence based on experiments and simulations. *European Journal of Physics*, 36(3), 35020. <https://doi.org/10.1088/0143-0807/36/3/035020>
- Amin, T., Smith, G. & Wiser, M. (2014). Student conceptions and conceptual change. In N. G. Lederman & S. K. Abell (Hrsg.), *Handbook of research on science education*. Routledge.
- Araujo, I. S., Veit, E. A. & Moreira, M. A. (2008). Physics students' performance using computational modelling activities to improve kinematics graphs interpretation. *Computers & Education*, 50(4), 1128–1140. <https://doi.org/10.1016/j.compedu.2006.11.004>
- Arganbright, D. E. (1984). The Electronic Spreadsheet and Mathematical Algorithms. *The Two-Year College Mathematics Journal*, 15(2), 148–157. <https://doi.org/10.1080/00494925.1984.11972767>
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2016). *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung* (14., überarbeitete und aktualisierte Auflage). Springer Gabler. <https://doi.org/10.1007/978-3-662-46076-4>

- Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior research methods*, 37(3), 379–384. <https://doi.org/10.3758/bf03192707>
- Baker, J. & Sugden, S. (2003). Spreadsheets in Education –The First 25 Years. *Spreadsheets in Education (eJSiE)*, 1(1), 18–43.
- Baron-Cohen, S., Knickmeyer, R. C. & Belmonte, M. K. (2005). Sex differences in the brain: implications for explaining autism. *Science*, 310(5749), 819–823. <https://doi.org/10.1126/science.1115455>
- Bartlett, M. S. (1937). Properties of Sufficiency and Statistical Tests. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 160(901), 268–282. <http://www.jstor.org/stable/96803> (Datum des letzten Abrufs: 02.09.2022).
- Becker, S., Klein, P., Gößling, A. & Kuhn, J. (2019a). Förderung von Konzeptverständnis und Repräsentationskompetenz durch Tablet-PC-gestützte Videoanalyse. *Zeitschrift für Didaktik der Naturwissenschaften*, 25(1), 1–24. <https://doi.org/10.1007/s40573-019-00089-4>
- Becker, S., Klein, P., Gößling, A. & Kuhn, J. (2019b, 29. August). *Using Mobile Devices to Augment Inquiry-Based Learning Processes with Multiple Representations*. <https://arxiv.org/pdf/1908.11281> (Datum des letzten Abrufs: 02.09.2022).
- Becker, S., Klein, P., Gößling, A. & Kuhn, J. (2020a). Investigating Dynamic Visualizations of Multiple Representations Using Mobile Video Analysis in Physics Lessons. *Zeitschrift für Didaktik der Naturwissenschaften*, 1–20. <https://doi.org/10.1007/s40573-020-00116-9>
- Becker, S., Klein, P., Gößling, A. & Kuhn, J. (2020b). Reduktion kognitiver Belastung und Förderung von Konzeptverständnis durch Tablet-PC-gestützte Videoanalyse: Hauptstudie zur Lernwirksamkeit im Mechanikunterricht der Sekundarstufe 2. *Zeitschrift für Didaktik der Naturwissenschaften*, 26(1) (submitted).
- Beichner, R. J. (1990). The effect of simultaneous motion presentation and graph generation in a kinematics lab. *Journal of Research in Science Teaching*, 27(8), 803–815. <https://doi.org/10.1002/tea.3660270809>
- Beichner, R. J. (1994). Testing student interpretation of kinematics graphs. *American Journal of Physics*, 62(8), 750–762. <https://doi.org/10.1119/1.17449>

- Beichner, R. J. (1996). The impact of video motion analysis on kinematics graph interpretation skills. *American Journal of Physics*, 64(10), 1272–1277. <https://doi.org/10.1119/1.18390>
- Benacka, J. (2015a). Projectile general motion in a vacuum and a spreadsheet simulation. *Physics Education*, 50(1), 58–63. <https://doi.org/10.1088/0031-9120/50/1/58>
- Benacka, J. (2015b). Spreadsheet application showing the proper elevation angle, points of shot and impact of a projectile. *Physics Education*, 50(3), 342–347. <https://doi.org/10.1088/0031-9120/50/3/342>
- Benacka, J. (2016). Numerical Modelling with Spreadsheets as a Means to Promote STEM to High School Students. *EURASIA Journal of Mathematics, Science & Technology Education*, 12(7). <https://doi.org/10.12973/eurasia.2016.1236a>
- Bethge, T. (1992). Mechanik in der Sekundarstufe II: Ein Kurskonzept unter Nutzung von Software-Werkzeugen. In K. H. Wiebel (Hrsg.), *Zur Didaktik der Physik und Chemie. Probleme und Perspektiven* (S. 152–154). Leuchtturm-Verlag.
- Billington, J., Baron-Cohen, S. & Wheelwright, S. (2007). Cognitive style predicts entry into physical sciences and humanities: Questionnaire and performance tests of empathy and systemizing. *Learning and Individual Differences*, 17(3), 260–268. <https://doi.org/10.1016/j.lindif.2007.02.004>
- Bork, A. (1980). Physics in the Irvine Educational Technology Center. In A. Bork (Hrsg.), *Computer Assisted Learning in Physics Education* (S. 37–57). Elsevier Science. <https://doi.org/10.1016/B978-0-08-025812-6.50009-1>
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation: Für Human- und Sozialwissenschaftler ; mit 87 Tabellen* (4. Aufl.). Springer-Lehrbuch Bachelor, Master. Springer-Medizin-Verl.
- Bortz, J. & Weber, R. (2005). *Statistik für Human- und Sozialwissenschaftler: Mit 242 Tabellen* (6. Aufl.). Springer-Lehrbuch. Springer Medizin. <https://doi.org/10.1007/b137571>
- Bossel, H. (1992). *Modellbildung und Simulation: Konzepte, Verfahren und Modelle Zum Verhalten Dynamischer Systeme*. Springer Vieweg. in Springer Fachmedien Wiesbaden GmbH. <https://doi.org/10.1007/978-3-322-83658-8>

- Boyd, A. & Rubin, A. (1996). Interactive video: A bridge between motion and math. *International Journal of Computers for Mathematical Learning*, 1(1). <https://doi.org/10.1007/BF00191472>
- Brell, C. (2008). *Lernmedien und Lernerfolg - reale und virtuelle Materialien im Physikunterricht: Empirische Untersuchungen in achten Klassen an Gymnasien (Laborstudie) zum Computereinsatz mit Simulation und IBE*. Zugl.: Bremen, Univ., Diss., 2007. *Studien zum Physik- und Chemielernen: Bd. 74*. Logos-Verl.
- 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>
- Brewe, E., Kramer, L. & O'Brien, G. (2009). Modeling instruction: Positive attitudinal shifts in introductory physics measured with CLASS. *Physical Review Special Topics - Physics Education Research*, 5(1), 121. <https://doi.org/10.1103/PhysRevSTPER.5.013102>
- Brewe, E., Sawtelle, V., Kramer, L. H., O'Brien, G. E., Rodriguez, I. & Pamelá, P. (2010). Toward equity through participation in Modeling Instruction in introductory university physics. *Physical Review Special Topics - Physics Education Research*, 6(1), 20. <https://doi.org/10.1103/PhysRevSTPER.6.010106>
- Brophy, J. E. & Good, T. L. (1970). Teachers' communication of differential expectations for children's classroom performance: Some behavioral data. *Journal of Educational Psychology*, 61(5), 365–374. <https://doi.org/10.1037/h0029908>
- Brown, D. E. (1989). Students' concept of force: the importance of understanding Newton's third law. *Physics Education*, 24(6), 353–358. <https://doi.org/10.1088/0031-9120/24/6/007>
- Brown, D. & Cox, A. J. (2009). Innovative Uses of Video Analysis. *The Physics Teacher*, 47(3), 145–150. <https://doi.org/10.1119/1.3081296>
- Browne, W. J. (1998). *Applying MCMC methods to multi-level models* [Dissertation]. University of Bath, Bath. <https://users.soe.ucsc.edu/~draper/browne-PhD-dissertation-1999.pdf> (Datum des letzten Abrufs: 02.09.2022).
- Brungardt, J. B. & Zollman, D. (1995). Influence of interactive videodisc instruction using simultaneous-time analysis on kinematics graphing skills of high school

- physics students. *Journal of Research in Science Teaching*, 32(8), 855–869.
<https://doi.org/10.1002/tea.3660320808>
- Bryan, J. A. (2010). Investigating the conservation of mechanical energy using video analysis: four cases. *Physics Education*, 45(1), 50–57.
<https://doi.org/10.1088/0031-9120/45/1/005>
- Buffler, A., Pillay, S., Lubben, F. & Fearick, R. (2008). A model-based view of physics for computational activities in the introductory physics course. *American Journal of Physics*, 76(4), 431–437. <https://doi.org/10.1119/1.2835045>
- Bunge, M. (1973). *Method, Model and Matter. Synthese Library, Monographs on Epistemology, Logic, Methodology, Philosophy of Science, Sociology of Science and of Knowledge, and on the Mathematical Methods of Social and Behavioral Sciences: Bd. 44.* Springer. <https://doi.org/10.1007/978-94-010-2519-5>
- Burde, J.-P. (2018). *Konzeption und Evaluation eines Unterrichtskonzepts zu einfachen Stromkreisen auf Basis des Elektronengasmodells.* Dissertation. *Studien zum Physik- und Chemielernen: Band 259.* Logos-Verl.
- Burde, J.-P. & Wilhelm, T. (2020). Teaching electric circuits with a focus on potential differences. *Physical Review Physics Education Research*, 16(2).
<https://doi.org/10.1103/PhysRevPhysEducRes.16.020153>
- Caballero, M. D., Burk, J. B., Aiken, J. M., Thoms, B. D., Douglas, S. S., Scanlon, E. M. & Schatz, M. F. (2014). Integrating Numerical Computation into the Modeling Instruction Curriculum. *The Physics Teacher*, 52(1), 38–42.
<https://doi.org/10.1119/1.4849153>
- Caballero, M. D., Kohlmyer, M. A. & Schatz, M. F. (2012). Implementing and assessing computational modeling in introductory mechanics. *Physical Review Special Topics - Physics Education Research*, 8(2), 23.
<https://doi.org/10.1103/PhysRevSTPER.8.020106>
- Caballero, M. D. & Merner, L. (2018). On the Prevalence and Nature of Computational Instruction in Undergraduate Physics Programs across the United States. *Physical Review Physics Education Research*, 14(2), 1.
<https://doi.org/10.1103/PhysRevPhysEducRes.14.020129>

- Caballero, M. D. & Pollock, S. J. (2014). A model for incorporating computation without changing the course: An example from middle-division classical mechanics. *American Journal of Physics*, 82(3), 231–237. <https://doi.org/10.1119/1.4837437>
- Carnes, E. R., Lindbeck, J. S. & Griffin, C. F. (1987). Effects of group size and advance organizers on learning parameters when using microcomputer tutorials in kinematics. *Journal of Research in Science Teaching*, 24(9), 781–789. <https://doi.org/10.1002/tea.3660240902>
- Castro Kurtz dos Santos, A. de. (1992). *Computational modelling in science education : a study of students' ability to manage some different approaches to modelling* [Dissertation]. University of London, London. https://discovery.ucl.ac.uk/id/eprint/10018818/1/149052_Redacted.pdf (Datum des letzten Abrufs: 02.09.2022).
- Cattell, R. B. (1966). The Scree Test For The Number Of Factors. *Multivariate behavioral research*, 1(2), 245–276. https://doi.org/10.1207/s15327906mbr0102_10
- Chabay, R. & Sherwood, B. (2008). Computational physics in the introductory calculus-based course. *American Journal of Physics*, 76(4), 307–313. <https://doi.org/10.1119/1.2835054>
- Chabay, R. W. & Sherwood, B. A. (2015). *Matter & interactions* (4. Aufl.). Wiley.
- Chambers, J. M. (1983). *Graphical methods for data analysis. Chapman & Hall statistics series*. Wadsworth & Brooks/Cole.
- Chen, J., Wang, M., Kirschner, P. A. & Tsai, C.-C. (2018). The Role of Collaboration, Computer Use, Learning Environments, and Supporting Strategies in CSCL: A Meta-Analysis. *Review of Educational Research*, 88(6), 799–843. <https://doi.org/10.3102/0034654318791584>
- Chi, M. T.H., Slotta, J. D. & Leeuw, N. de (1994). From things to processes: A theory of conceptual change for learning science concepts. *Learning and Instruction*, 4(1), 27–43. [https://doi.org/10.1016/0959-4752\(94\)90017-5](https://doi.org/10.1016/0959-4752(94)90017-5)
- Claessens, T. (2017). Analyzing Virtual Physics Simulations with Tracker. *The Physics Teacher*, 55(9), 558–560. <https://doi.org/10.1119/1.5011834>

- Clark, R. C. & Mayer, R. E. (2016). *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning* (Fourth edition). Wiley. <https://doi.org/10.1002/9781119239086>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). L. Erlbaum Associates. <https://doi.org/10.4324/9780203771587>
- Coletta, V. P. & Steinert, J. J. (2020). Why normalized gain should continue to be used in analyzing preinstruction and postinstruction scores on concept inventories. *Physical Review Physics Education Research*, 16(1). <https://doi.org/10.1103/PhysRevPhysEducRes.16.010108>
- Crawford, B. & Cullin, M. (2005). Dynamic Assessments of Preservice Teachers' Knowledge of Models and Modelling. In K. Boersma, H. Eijkelhof, M. Goedhart & O. Jong (Hrsg.), *SpringerLink Bücher. Research and the Quality of Science Education* (S. 309–323). Springer Netherlands. https://doi.org/10.1007/1-4020-3673-6_25
- Crawford, B. A. & Cullin, M. J. (2004). Supporting prospective teachers' conceptions of modelling in science. *International Journal of Science Education*, 26(11), 1379–1401. <https://doi.org/10.1080/09500690410001673775>
- Cronbach, L. J. & Furby, L. (1970). How we should measure "change": Or should we? *Psychological bulletin*, 74(1), 68–80. <https://doi.org/10.1037/h0029382>
- diSessa, A. A. (1993). Toward an Epistemology of Physics. *Cognition and Instruction*, 10(2-3), 105–225. <https://doi.org/10.1080/07370008.1985.9649008>
- Ditton, H. (1998). *Mehrebenenanalyse: Grundlagen und Anwendungen des hierarchisch linearen Modells. Juventa-Paperback*. Juventa-Verl.
- Dreyfus, A., Jungwirth, E. & Eliovitch, R. (1990). Applying the “cognitive conflict” strategy for conceptual change—some implications, difficulties, and problems. *Science Education*, 74(5), 555–569. <https://doi.org/10.1002/sce.3730740506>
- Ducas, T. (1993). Active Video: The Promise of AVID Learning. *Journal of College Science Teaching*, 23(3), 166–172.
- Duit, R. (1993). Schülervorstellungen – von Lerndefiziten zu neuen Unterrichtsansätzen. *Naturwissenschaften im Unterricht. Physik*, 4(16), 4–10.

- Duit, R. (2009). STCSE: Students' and Teachers' Conceptions and Science Education (Bibliografie). Erhältlich von Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik. <http://archiv.ipn.uni-kiel.de/stcse/>
- Eisberg, R. M. & Eisberg, R. M. (1978). *Mathematische Physik für die Benutzer programmierbarer Taschenrechner: Mit 16 Programmen, 29 Beispielen, 92 Aufgaben* (1. Aufl.). Oldenbourg.
- Escalada, L., Grabhorn, R. & Zollman, D. (1996). Applications of Interactive Digital Video in a Physics Classroom. *Journal of Educational Multimedia and Hypermedia*, 5(1), 73–97.
- Ester, M. & Sander, J. (2000). *Knowledge Discovery in Databases: Techniken und Anwendungen. Springer eBook Collection Computer Science and Engineering*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-58331-5>
- Ferreira, A., Lemmer, M. & Gunstone, R. (2019). Alternative Conceptions: Turning Adversity into Advantage. *Research in Science Education*, 49(3), 657–678. <https://doi.org/10.1007/s11165-017-9638-y>
- Feynman, R. P., Leighton, R. B. & Sands, M. L. (2011). *The Feynman lectures on physics* (New millennium ed.). Basic Books.
- Field, A. (2003). *Discovering statistics using SPSS for windows: Advanced techniques for the beginner. Introducing statistical methods*. Sage.
- Field, A. (2007). *Discovering statistics using SPSS: (and sex, drugs and rock'n'roll)* (2. Aufl.). *ISM, Introducing statistical methods*. SAGE Publ.
- Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*. Sage.
- Filsinger, S. & Wilhelm, T. (2020). Einsatzmöglichkeiten des Computerspiels "Minecraft" im Physikunterricht. In H. Grötzebauch (Vorsitz), *DPG-Frühjahrstagung*. Symposium im Rahmen der Tagung von DPG, Bonn.
- Flores, S., Kanim, S. E. & Kautz, C. H. (2004). Student use of vectors in introductory mechanics. *American Journal of Physics*, 72(4), 460–468. <https://doi.org/10.1119/1.1648686>
- Forrester, J. W. (1968). *Market Growth as Influenced by Capital Investment. Reprint*. Industrial Management Review.

- Gerdes, J. & Schecker, H. (1999). Der Force Concept Inventory. Ein diagnostischer Test zu Schülervorstellungen in der Mechanik. *Der mathematische und naturwissenschaftliche Unterricht*, 52(5), 283–288.
- Gilbert, J. K. & Justi, R. (2016). Models of Modelling. In R. Justi (Hrsg.), *Models and Modeling in Science Education Ser: v.9. Modelling-Based Teaching in Science Education* (Bd. 9, S. 17–40). Springer International Publishing AG. https://doi.org/10.1007/978-3-319-29039-3_2
- Girwidz, R., Kurz, G. & Kautz, C. (2003). Zum Verständnis der newtonschen Mechanik bei Studienanfängern: Der Test "Force Concept Inventory - FCI". In V. Nordmeier (Vorsitz), *Didaktik der Physik: Beiträge der Frühjahrstagung der DPG*. Symposium im Rahmen der Tagung von DPG, Augsburg.
- Gläser, J. & Laudel, G. (2013). Life With and Without Coding: Two Methods for Early-Stage Data Analysis in Qualitative Research Aiming at Causal Explanations. Vorab-Onlinepublikation. <https://doi.org/10.17169/fqs-14.2.1886> (Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, Vol 14, No 2 (2013).
- Göhner, M. & Krell, M. (2020). Qualitative Inhaltsanalyse in naturwissenschaftsdidaktischer Forschung unter Berücksichtigung von Gütekriterien: Ein Review. *Zeitschrift für Didaktik der Naturwissenschaften*, 26(1), 207–225. <https://doi.org/10.1007/s40573-020-00111-0>
- Gollwitzer, M., Banse, R., Eisenbach, K. & Naumann, A. (2007). Effectiveness of the Vienna Social Competence Training on Explicit and Implicit Aggression. *European Journal of Psychological Assessment*, 23(3), 150–156. <https://doi.org/10.1027/1015-5759.23.3.150>
- Greve, W. & Wentura, D. (1997). *Wissenschaftliche Beobachtung: Eine Einführung* [2. Aufl.]. Beltz.
- Grosslight, L., Unger, C., Jay, E. & Smith, C. L. (1991). Understanding models and their use in science: Conceptions of middle and high school students and experts. *Journal of Research in Science Teaching*, 28(9), 799–822. <https://doi.org/10.1002/tea.3660280907>
- Guicciardini, N. (2003). *Reading the principia: The debate on Newton's mathematical methods for natural philosophy from 1687 to 1736* (1. paperback ed.). Cambridge Univ. Press.

- Gulliksen, H. (1950). *Theory of mental tests*. John Wiley & Sons Inc.
<https://doi.org/10.1037/13240-000>
- Gunstone, R. & White, R. (1981). Understanding of gravity. *Science Education*, 65(3), 291–299. <https://doi.org/10.1002/sce.3730650308>
- Guzdial, M. (1994). *Approches to Classroom-Based Computational Science*.
<https://files.eric.ed.gov/fulltext/ED396683.pdf> (Datum des letzten Abrufs: 02.09.2022).
- Hair, J. F. (2010). *Multivariate data analysis* (7. ed.). Pearson Prentice Hall.
- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66(1), 64–74. <https://doi.org/10.1119/1.18809>
- Halloun, I. A. & Hestenes, D. (1985). The initial knowledge state of college physics students. *American Journal of Physics*, 53(11), 1043–1055.
<https://doi.org/10.1119/1.14030>
- Hartig, J. & Rakoczy, K. (2010). Mehrebenenanalyse - Multilevel Analysis. In H. Holling (Hrsg.), *Handbuch der Psychologie: Bd. 13. Handbuch Statistik, Methoden und Evaluation* (S. 538–547). Hogrefe.
- Hartigan, J. A. & Wong, M. A. (1979). Algorithm AS 136: A K-Means Clustering Algorithm. *Applied Statistics*, 28(1), 100. <https://doi.org/10.2307/2346830>
- Hassel, D. J. (1987). *The roll of Modelling avticipities in the humanities curriculum, with special reference to geography: an investigative study* [Dissertation]. King's College, London.
- Hawkins, N., Obsniuk, M. J., Irving, P. W. & Caballero, M. D. (2017). Examining Thematic Variation in a Phenomenographical Study on Computational Physics. In L. Ding, A. Traxler & Y. Cao (Hrsg.), *2017 Physics Education Research Conference Proceedings* (S. 168–171). American Association of Physics Teachers.
<https://doi.org/10.1119/perc.2017.pr.037>
- Hedges, L. V. (1981). Distribution Theory for Glass's Estimator of Effect Size and Related Estimators. *Journal of Educational Statistics*, 6(2), 107.
<https://doi.org/10.2307/1164588>

- Heinicke, S. (2019). Physikunterricht aus Perspektive von Mädchen – und Jungen. In D. Duchardt, A. B. Bossmann & C. Denz (Hrsg.), *Vielfältige Physik: Wissenschaftlerinnen schreiben über ihre Forschung* (1. Aufl., S. 27–40). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-58035-6_3
- Helmke, A. (1992). Determinanten der Schulleistung: Forschungsstand und Forschungsdefizite. In K. Ingenkamp, R. Jäger, H. Petillon & B. Wolf (Hrsg.), *Empirische Pädagogik 1970-1990. Eine Bestandsaufnahme der Forschung in der Bundesrepublik Deutschland*. (S. 595–602). Deutscher Studien Verlag.
- Herman, R. (2009). Spreadsheet physics: Examples in meteorology and planetary science. *American Journal of Physics*, 77(12), 1124–1129. <https://doi.org/10.1119/1.3230033>
- Hermann, M. (2010). Numerische Verfahren für gewöhnliche Differentialgleichungen. In D. Fey (Hrsg.), *eXamen.press. Grid-Computing: Eine Basistechnologie für Computational Science* (S. 149–205). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-79747-0_8
- Hessisches Kultusministerium. (2016). *Kerncurriculum gymnasiale Oberstufe*. Hessisches Kultusministerium. <https://kultusministerium.hessen.de/sites/kultusministerium.hessen.de/files/2021-07/kcgo-m.pdf> (Datum des letzten Abrufs: 02.09.2022).
- Hestenes, D. & Halloun, I. (1995). Interpreting the force concept inventory: A response to March 1995 critique by Huffman and Heller. *The Physics Teacher*, 33(8), 502. <https://doi.org/10.1119/1.2344278>
- Hestenes, D., Wells, M. & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher*, 30(3), 141–158. <https://doi.org/10.1119/1.2343497>
- Hewson, M. G. & Hewson, P. W. (1983). Effect of instruction using students' prior knowledge and conceptual change strategies on science learning. *Journal of Research in Science Teaching*, 20(8), 731–743. <https://doi.org/10.1002/tea.3660200804>
- Hillmayr, D., Reinhold, F., Ziernwald, L. & Reiss, K. (2017). *Digitale Medien im mathematisch-naturwissenschaftlichen Unterricht der Sekundarstufe: Einsatzmöglichkeiten, Umsetzung und Wirksamkeit* [1. Auflage]. Waxmann.

- <https://www.waxmann.com/?eID=texte&pdf=3766Volltext.pdf&typ=zusatztext>
(Datum des letzten Abrufs: 02.09.2022).
- Hillmayr, D., Ziernwald, L., Reinhold, F., Hofer, S. I. & Reiss, K. M. (2020). The potential of digital tools to enhance mathematics and science learning in secondary schools: A context-specific meta-analysis. *Computers & Education*, 153, 103897. <https://doi.org/10.1016/j.compedu.2020.103897>
- Hochberg, K., Becker, S., Louis, M., Klein, P. & Kuhn, J. (2020). Using Smartphones as Experimental Tools—a Follow-up: Cognitive Effects by Video Analysis and Reduction of Cognitive Load by Multiple Representations. *Journal of Science Education and Technology*, 29(2), 303–317. <https://doi.org/10.1007/s10956-020-09816-w>
- Hockicko, P., Trpišová, B. & Ondruš, J. (2014). Correcting Students' Misconceptions about Automobile Braking Distances and Video Analysis Using Interactive Program Tracker. *Journal of Science Education and Technology*, 23(6), 763–776. <https://doi.org/10.1007/s10956-014-9510-z>
- Hoffmann, L., Häußler, P. & Lehrke, M. (1998). *Die IPN-Interessenstudie Physik* (IPN IPN 158). Kiel.
- Hollender, N., Hofmann, C., Deneke, M. & Schmitz, B. (2010). Integrating cognitive load theory and concepts of human–computer interaction. *Computers in Human Behavior*, 26(6), 1278–1288. <https://doi.org/10.1016/j.chb.2010.05.031>
- Hopf, M. & Wilhelm, T. (2018). Conceptual Change: Entwicklung physikalischer Vorstellungen. In H. Schecker, T. Wilhelm, M. Hopf & R. Duit (Hrsg.), *Schülervorstellungen und Physikunterricht* (S. 23–37). Springer Berlin Heidelberg.
- Horn, J. (1965). A Rationale and Test for the Number Of Factors in Factor Analysis. *Psychometrika*, 30, 179–185. <https://doi.org/10.1007/bf02289447>
- Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2. Aufl.). *Quantitative methodology series*. Routledge.
- Hsiao, F. S.T. (1985). Micros in mathematics education—Uses of spreadsheets in CAL. *International Journal of Mathematical Education in Science and Technology*, 16(6), 705–713. <https://doi.org/10.1080/0020739850160606>

- Hucke, L. (2000). *Handlungsregulation und Wissenserwerb in traditionellen und computergestützten Experimenten des physikalischen Praktikums*. Zugl.: Dortmund, Univ., Diss., 1999. *Studien zum Physiklernen: Bd. 8*. Logos-Verl.
- Hucke, L. & Fischer, H. (1999). Die Verbindung von Theorie und Praxis in traditionellen und in computergestützten Praktikumsexperimenten. Fallstudien zur Effektivität des physikalischen Anfängerpraktikums Anfängerpraktikums. In R. Brechel (Hrsg.), *Zur Didaktik der Physik und Chemie, Probleme und Perspektiven, Vorträge auf der Tagung für Didaktik der Physik/Chemie in Essen* (S. 250–252). Leuchtturm-Verlag.
- Huffman, D. & Heller, P. (1995). What does the force concept inventory actually measure? *The Physics Teacher*, 33(3), 138–143. <https://doi.org/10.1119/1.2344171>
- Irving, P. W., Obsniuk, M. J. & Caballero, M. D. (2017). P 3 : a practice focused learning environment. *European Journal of Physics*, 38(5), 55701. <https://doi.org/10.1088/1361-6404/aa7529>
- Jackson, J., Dukerich, L. & Hestenes, D. (2008). Modeling Instruction. An Effective Model for Science Education.
- Jung, W. (1981). Vorstellungen über Kraft und Stoß bei Schülern vom 8. bis 11. Schuljahr. In W. Jung, H. Wiesner & P. Engelhardt (Hrsg.), *Texte zur mathematisch-naturwissenschaftlich-technischen Forschung und Lehre: Bd. 8. Vorstellungen von Schülern über Begriffe der Newtonschen Mechanik: Empirische Untersuchungen und Ansätze zu didaktisch-methodischen Folgerungen* (S. 63–111). Franzbecker.
- Just, A. M., Aufschnaiter, C. von & Vorholzer, A. (2021). Effects of conceptual and contextual task characteristics on students' activation of mechanics conceptions. *European Journal of Physics*, 42(2), 25702. <https://doi.org/10.1088/1361-6404/abd229>
- Justi, R. & Gilbert, J. (2003). Teachers' views on the nature of models. *International Journal of Science Education*, 25(11), 1369–1386. <https://doi.org/10.1080/0950069032000070324>
- Justi, R. S. & Gilbert, J. K. (2002). Modelling, teachers' views on the nature of modelling, and implications for the education of modellers. *International Journal of Science Education*, 24(4), 369–387. <https://doi.org/10.1080/09500690110110142>

- Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23(3), 187–200. <https://doi.org/10.1007/BF02289233>
- Kaiser, H. F. (1960). The Application of Electronic Computers to Factor Analysis. *Educational and Psychological Measurement*, 20(1), 141–151. <https://doi.org/10.1177/001316446002000116>
- Kaiser, H. F. (1970). A second generation little jiffy. *Psychometrika*, 35(4), 401–415. <https://doi.org/10.1007/BF02291817>
- Kaiser, H. F. & Rice, J. (1974). Little Jiffy, Mark Iv. *Educational and Psychological Measurement*, 34(1), 111–117. <https://doi.org/10.1177/001316447403400115>
- Kalyuga, S. (2011). Cognitive Load Theory: How Many Types of Load Does It Really Need? *Educational Psychology Review*, 23(1), 1–19. <https://doi.org/10.1007/s10648-010-9150-7>
- Kane, D. & Sherwood, B. (1980). A computer-based course in classical mechanics. *Computers & Education*, 4(1), 15–36. [https://doi.org/10.1016/0360-1315\(80\)90006-8](https://doi.org/10.1016/0360-1315(80)90006-8)
- Kanim, S. E. & Subero, K. (2010). Introductory labs on the vector nature of force and acceleration. *American Journal of Physics*, 78(5), 461–466. <https://doi.org/10.1119/1.3373923>
- Kelemen, D. & Rosset, E. (2009). The human function compunction: teleological explanation in adults. *Cognition*, 111(1), 138–143. <https://doi.org/10.1016/j.cognition.2009.01.001>
- Kircher, E. (2007). Modellbegriff und Modellbildung in der Physikdidaktik. In E. Kircher, R. Girwidz & P. Häussler (Hrsg.), *SpringerLink Bücher. Physikdidaktik: Theorie und Praxis* (S. 679–706). Springer Berlin Heidelberg.
- Kircher, E., Girwidz, R. & Häussler, P. (Hrsg.). (2007). *SpringerLink Bücher. Physikdidaktik: Theorie und Praxis*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-540-34091-1>
- Klein, P., Kuhn, J. & Müller, A. (2018). Förderung von Repräsentationskompetenz und Experimentbezug in den vorlesungsbegleitenden Übungen zur Experimentalphysik. *Zeitschrift für Didaktik der Naturwissenschaften*, 24(1), 17–34. <https://doi.org/10.1007/s40573-018-0070-2>

- Klieme, E. & Maichle, U. (1994). *Modellbildung und Simulation im Unterricht der Sekundarstufe I - Auswertung von Unterrichtsversuchen mit dem Modellbildungssystem MODUS*, Bonn.
- Krippendorff, K. (1980). Validity in Content Analysis. In E. Mochmann (Hrsg.), *Computerstrategien für die Kommunikationsanalyse*. Campus Verlag.
- Kuckartz, U. (2018). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung* (4. Aufl.). *Grundlagentexte Methoden*. Beltz.
- Kuhn, J. (2010). *Authentische Aufgaben im theoretischen Rahmen von Instruktions- und Lehr-Lern-Forschung: Optimierung von Ankermedien für eine neue Aufgabenkultur im Physikunterricht*. Vieweg+Teubner Verlag / GWV Fachverlage GmbH Wiesbaden.
- Kuhn, J. (2014). Mehrebenenanalyse am Beispiel der Lernwirkung von Aufgaben. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Springer eBook Collection. Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 297–310). Springer Spektrum.
- Kunter, M., Schümer, G., Artelt, C., Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillmann, K.-J. & Weiß, M. (2002). *PISA 2000: Dokumentation der Erhebungsinstrumente. Materialien aus der Bildungsforschung: Nr. 72*. Max-Planck-Inst. für Bildungsforschung. <http://hdl.handle.net/hdl:11858/00-001M-0000-0023-9987-C> (Datum des letzten Abrufs: 02.09.2022).
- Lamnek, S. & Krell, C. (2016). *Qualitative Sozialforschung: Mit Online-Materialien* (6., überarbeitete Auflage). Beltz.
- Lance, C. E. & Vandenberg, R. J. (2009). *Statistical and methodological myths and urban legends: Doctrine, verity and fable in the organizational and social sciences. Statistical and Methodological Myths and Urban Legends: Doctrine, Verity and Fable in the Organizational and Social Sciences*. Routledge.
- Landis, J. R. & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159. <https://doi.org/10.2307/2529310>
- Laukenmann, M., Bleicher, M., Fuß, M., Gläser-Zikuda, M., Mayring, P. & Rhöneck, C. von (2000). Eine Untersuchung zum Einfluss emotionaler Faktoren auf das

- Lernen im Physikunterricht. *Zeitschrift für Didaktik der Naturwissenschaften*, 6, 139–155.
- Laws, P. W. (1991a). Calculus-Based Physics without Lectures. *Physics Today*, 44(12), 24. <https://doi.org/10.1063/1.881276>
- Laws, P. W. (1991b). Workshop Physics: Learning Introductory Physics by Doing It. *Change: The Magazine of Higher Learning*, 23(4), 20–27. <https://doi.org/10.1080/00091383.1991.9940581>
- Laws, P. W., Boyle, R. J., Cooney, P. J., Luetzelschwab, J. W., Sokoloff, D. R. & Thornton, R. K. (2004). *Workshop physics® activity guide*. Wiley.
- Laws, P. W. & Pfister, H. (1998). Using digital video analysis in introductory mechanics projects. *The Physics Teacher*, 36(5), 282–287. <https://doi.org/10.1119/1.880068>
- Laws, P. W., Willis, M. C. & Sokoloff, D. R. (2015). Workshop Physics and Related Curricula: A 25-Year History of Collaborative Learning Enhanced by Computer Tools for Observation and Analysis. *The Physics Teacher*, 53(7), 401–406. <https://doi.org/10.1119/1.4931006>
- Leisner-Bodenthin, A. (2006). Zur Entwicklung von Modellkompetenz im Physikunterricht. *Zeitschrift für Didaktik der Naturwissenschaften*, 12, 91–109.
- Lemmer, M. (2013). Nature, Cause and Effect of Students' Intuitive Conceptions Regarding Changes in Velocity. *International Journal of Science Education*, 35(2), 239–261. <https://doi.org/10.1080/09500693.2011.647110>
- Lenhard, W. & Lenhard, A. (2017). *Computation of Effect Sizes*. <https://doi.org/10.13140/RG.2.2.17823.92329>
- Leppink, J., Paas, F., van der Vleuten, C. P. M., van Gog, T. & van Merriënboer, J. J. G. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods*, 45(4), 1058–1072. <https://doi.org/10.3758/s13428-013-0334-1>
- Lienert, G. A. & Raatz, U. (1998). *Testaufbau und Testanalyse* (6. Aufl.). *Grundlagen Psychologie*. Beltz.
- Lombrozo, T., Kelemen, D. & Zaitchik, D. (2007). Inferring design: evidence of a preference for teleological explanations in patients with Alzheimer's disease.

- Psychological Science*, 18(11), 999–1006. <https://doi.org/10.1111/j.1467-9280.2007.02015.x>
- Lord, F. (1967a). Elementary models for measuring change. In C. Harris (Hrsg.), *Problems in measuring change* (S. 199–211). University of Wisconsin Press.
- Lord, F. (1967b). A paradox in the interpretation of group comparisons. *Psychological bulletin*, 68(5), 304–305. <https://doi.org/10.1037/h0025105>
- Lück, S. & Wilhelm, T. (2011). Modellierung physikalischer Vorgänge am Computer. Modellbildungssysteme als Unterstützung zum Verständnis physikalischer Strukturen. *Unterricht Physik*, 22(122), 26–31.
- Maas, C. J. M. & Hox, J. J. (2005). Sufficient Sample Sizes for Multilevel Modeling. *Methodology*, 1(3), 86–92. <https://doi.org/10.1027/1614-2241.1.3.86>
- Maas, C. J.M. & Hox, J. J. (2004). The influence of violations of assumptions on multi-level parameter estimates and their standard errors. *Computational Statistics & Data Analysis*, 46(3), 427–440. <https://doi.org/10.1016/j.csda.2003.08.006>
- MacDonald, W. M., Redish, E. F. & Wilson, J. M. (1988). The M.U.P.P.E.T. Manifesto. *Computers in Physics*, 2(4), 23. <https://doi.org/10.1063/1.4822749>
- MacQueen, J. (1967). Some Methods for Classification and Analysis of Multivariate Observations. In L. Le Cam & Neymanm Jerzy (Hrsg.), *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability: Volume 1: Statistics* (S. 281–297). University of California Press.
- Marx, J. D. & Cummings, K. (2007). Normalized change. *American Journal of Physics*, 75(1), 87–91. <https://doi.org/10.1119/1.2372468>
- Mayer, R. E. (1999). Multimedia aids to problem-solving transfer. *International Journal of Educational Research*, 31(7), 611–623. [https://doi.org/10.1016/S0883-0355\(99\)00027-0](https://doi.org/10.1016/S0883-0355(99)00027-0)
- Mayer, R. E. (2003). The promise of multimedia learning: using the same instructional design methods across different media. *Learning and Instruction*, 13(2), 125–139. [https://doi.org/10.1016/S0959-4752\(02\)00016-6](https://doi.org/10.1016/S0959-4752(02)00016-6)
- Mayer, R. E. (2009). *Multimedia learning* (Second edition). Cambridge University Press. <https://doi.org/10.1017/CBO9780511811678>

- Mayer, R. E. (2014). *The Cambridge handbook of multimedia learning*. Cambridge handbooks in psychology. Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369>
- Mayer, R. E. (2017). Using multimedia for e-learning. *Journal of Computer Assisted Learning*, 33(5), 403–423. <https://doi.org/10.1111/jcal.12197>
- Mayer, R. E. & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology*, 93(2), 390–397. <https://doi.org/10.1037/0022-0663.93.2.390>
- Mayer, R. E. & Moreno, R. (2003). Nine Ways to Reduce Cognitive Load in Multimedia Learning. *Educational Psychologist*, 38(1), 43–52. https://doi.org/10.1207/S15326985EP3801_6
- Mayring, P. (2002). *Einführung in die qualitative Sozialforschung* (5. Aufl.). Beltz Studium. Beltz.
- Mayring, P. (2008). *Einführung in die qualitative Sozialforschung: Eine Anleitung zu qualitativem Denken* (5. Aufl.). Beltz Studium. Beltz.
- Mayring, P. (2015). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (12. Aufl.). Beltz Pädagogik. Beltz.
- Mayring, P. (2019). Qualitative Content Analysis: Demarcation, Varieties, Developments, 20(3). <https://doi.org/10.17169/FQS-20.3.3343> (Forum Qualitative Sozialforschung / Forum: Qualitative Social Research).
- McDermott, L. C., Rosenquist, M. L. & van Zee, E. H. (1987). Student difficulties in connecting graphs and physics: Examples from kinematics. *American Journal of Physics*, 55(6), 503–513. <https://doi.org/10.1119/1.15104>
- Meisert, A. (2008). Vom Modellwissen zum Modellverständnis - Elemente einer umfassenden Modellkompetenz und deren Fundierung durch lernerseitige Kriterien zur Klassifikation von Modellen: From knowing models to an understanding of models - elements of comprehensive model competence and its well founded learner-based development criteria to identify models. *Zeitschrift für Didaktik der Naturwissenschaften*, 14, 243-261. http://archiv.ipn.uni-kiel.de/zfdn/pdf/14_013_Meisert.pdf (Datum des letzten Abrufs: 02.09.2022).

- Mikelskis-Seifert, S. & Kasper, L. (2011). Modellieren in der Physik, im Alltag und im Unterricht. Hintergründe und unterrichtliche Orientierung zum Thema Modelle. *Naturwissenschaften im Unterricht. Physik*, 22(122), 4–12.
- Mikelskis-Seifert, S. & Leisner, A. (2005). Transferfähigkeit einer Modellkompetenz zur Teilchenstruktur der Materie. In A. Pitton (Hrsg.), *Relevanz fachdidaktischer Forschungsergebnisse für die Lehrerbildung: Gesellschaft für Didaktik der Chemie und Physik* (S. 454–456). LIT-Verlag.
- Mikelskis-Seifert, S., Thiele, M. & Wünscher, T. (2005). Modellieren - Schlüsselfähigkeit für physikalische Forschungs- und Lernprozesse. *PhyDid A - Physik und Didaktik in Schule und Hochschule*, 4(1), 30–46.
- Miller, G. A. & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology*, 110(1), 40–48. <https://doi.org/10.1037//0021-843X.110.1.40>
- Moreno, R. & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology*, 91(2), 358–368. <https://doi.org/10.1037/0022-0663.91.2.358>
- Morris, S. B. & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological methods*, 7(1), 105–125. <https://doi.org/10.1037/1082-989X.7.1.105>
- Muis, K. R. & Gierus, B. (2014). Beliefs About Knowledge, Knowing, and Learning: Differences Across Knowledge Types in Physics. *The Journal of Experimental Education*, 82(3), 408–430. <https://doi.org/10.1080/00220973.2013.813371>
- Mummendey, H. D. & Grau, I. (2014). *Die Fragebogen-Methode: Grundlagen und Anwendung in Persönlichkeits-, Einstellungs- und Selbstkonzeptforschung* (6. Aufl.). Hogrefe eLibrary. Hogrefe.
- Neves, R. G., Silva, J. C. & Teodoro, V. D. (2010, 24. Juni). *Computational Modelling with Modellus: An Enhancement Vector for the General University Physics Course*. <http://arxiv.org/pdf/1006.4662v1> (Datum des letzten Abrufs: 02.09.2022).
- Newton, I. (1687). *Philosophiae Naturalis Principia Mathematica* (1. Aufl.). Jussu Societatis Regiae ac typis Josephi Streater. <http://cudl.lib.cam.ac.uk/view/PR-ADV-B-00039-00001/40> (Datum des letzten Abrufs: 02.09.2022).

- Newton, I. (1963). *Mathematische Prinzipien der Naturlehre* (Unveränd. fotomechan. Nachdr. d. Ausg. Berlin, 1872).
- Niedderer, H. & Schecker, H. (1992). Towards an explicit description of cognitive systems for research in physics learning. In R. Duit, F. Goldberg & H. Niedderer (Hrsg.), *Research in Physics Learning: Theoretical Issues and Empirical Studies* (S. 74–98). IPN.
- Nissen, J. M., Talbot, R. M., Nasim Thompson, A. & van Dusen, B. (2018). Comparison of normalized gain and Cohen's d for analyzing gains on concept inventories. *Physical Review Physics Education Research*, 14(1). <https://doi.org/10.1103/PhysRevPhysEducRes.14.010115>
- Nunnally, J. C. & Bernstein, I. H. (2008). *Psychometric theory* (3. Aufl.). *McGraw-Hill series in psychology*. McGraw-Hill.
- Obsniuk, M. J., Irving, P. W. & Caballero, M. D. (2015, 17. November). *A Case Study: Novel Group Interactions through Introductory Computational Physics*. <http://arxiv.org/pdf/1511.05457v1> (Datum des letzten Abrufs: 02.09.2022).
- Ossimitz, G. (1996). Können Schüler im Mathematikunterricht systemisch denken lernen? In K. P. Müller (Hrsg.), *Beiträge zum Mathematikunterricht - Vorträge auf der 30. Bundestagung für Didaktik der Mathematik* (S. 337–340). Franzbecker-Verlag.
- Paas, F. G. W. C. & van Merriënboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of Educational Psychology*, 86(1), 122–133. <https://doi.org/10.1037/0022-0663.86.1.122>
- Palazzo, D. & Schools, C. (2008). Video analysis: the next physics laboratory? *Proceedings of the engineering education: 21st century frontiers*.
- Pallant, J. (2016). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS* (6th edition).
- Palmer, D. (1997). The effect of context on students' reasoning about forces. *International Journal of Science Education*, 19(6), 681–696. <https://doi.org/10.1080/0950069970190605>

- Pappas, J., Koleza, E., Rizos, Ioannis, John & Skordoulis, C. (2002). Using interactive digital video and motion analysis to bridge abstract mathematical notions with concrete everyday experiences. Paper presented at the 2nd. *2nd International Conference on Teaching Mathematics*.
- Peters, H. J. (1980). The CONDUIT series in physics. *Computers & Education*, 4(1), 1–9. [https://doi.org/10.1016/0360-1315\(80\)90004-4](https://doi.org/10.1016/0360-1315(80)90004-4)
- Pintrich, P. R., Marx, R. W. & Boyle, R. A. (1993). Beyond Cold Conceptual Change: The Role of Motivational Beliefs and Classroom Contextual Factors in the Process of Conceptual Change. *Review of Educational Research*, 63(2), 167–199. <https://doi.org/10.3102/00346543063002167>
- Plass, J. L. & Kaplan, U. (2016). Emotional Design in Digital Media for Learning. In S. Y. Tettegah & M. Gartmeier (Hrsg.), *Emotions and technology. Emotions, technology, design, and learning* (S. 131–161). Elsevier Academic Press. <https://doi.org/10.1016/B978-0-12-801856-9.00007-4>
- Posner, G. J., Strike, K. A., Hewson, P. W. & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66(2), 211–227. <https://doi.org/10.1002/sce.3730660207>
- Pötschke, M. (2019). Mehrebenenmodelle. In M. Apelt, I. Bode, R. Hasse, U. Meyer, V. von Groddeck, M. Wilkesmann & A. Windeler (Hrsg.), *Living reference work. Handbuch Organisationssoziologie* (S. 1–30). Springer VS. https://doi.org/10.1007/978-3-658-16937-4_29-1
- Potvin, P. (2017). The coexistence claim and its possible implications for success in teaching for conceptual “change”. *European Journal of Science and Mathematics Education*, 5(1), 55–66. <https://doi.org/10.30935/scimath/9497>
- Potvin, P., Malenfant-Robichaud, G., Cormier, C. & Masson, S. (2020). Coexistence of Misconceptions and Scientific Conceptions in Chemistry Professors: A Mental Chronometry and fMRI Study. *Frontiers in Education*, 5, Artikel 542458. <https://doi.org/10.3389/educ.2020.542458>
- Potvin, P., Masson, S., Lafortune, S. & Cyr, G. (2015). PERSISTENCE OF THE INTUITIVE CONCEPTION THAT HEAVIER OBJECTS SINK MORE: A REACTION TIME STUDY WITH DIFFERENT LEVELS OF INTERFERENCE.

- International Journal of Science and Mathematics Education*, 13(1), 21–43.
<https://doi.org/10.1007/s10763-014-9520-6>
- Potvin, P., Nenciovici, L., Malenfant-Robichaud, G., Thibault, F., Sy, O., Mahhou, M. A., Bernard, A., Allaire-Duquette, G., Blanchette Sarrasin, J., Brault Foisy, L.-M., Brouillette, N., St-Aubin, A.-A., Charland, P., Masson, S., Riopel, M., Tsai, C.-C., Bélanger, M. & Chastenay, P. (2020). Models of conceptual change in science learning: establishing an exhaustive inventory based on support given by articles published in major journals. *Studies in Science Education*, 56(2), 157–211.
<https://doi.org/10.1080/03057267.2020.1744796>
- Przyborski, A. & Wohlrab-Sahr, M. (2014). *Qualitative Sozialforschung: Ein Arbeitsbuch* (4. Aufl.). *Lehr- und Handbücher der Soziologie*. Oldenbourg Verlag.
<https://doi.org/10.1524/9783486719550>
- Rasch, B., Frieze, M., Hofmann, W. & Naumann, E. (2010). *Quantitative Methoden I - Einführung in die Statistik für Psychologen und Sozialwissenschaftler: Deskriptive Statistik, Inferenzstatistik, t-Test, Korrelationstechniken, Regressionsanalyse, Formelsammlung, Glossar, Verteilungstabellen* (3. Aufl.). *Quantitative Methoden*. Springer.
- Raudenbush, S. W. & Bryk, A. S. (2010). *Hierarchical linear models: Applications and data analysis methods* (2. Aufl.). *Advanced quantitative techniques in the social sciences: Bd. 1*. SAGE Publ.
- Ray, S. & Fröhlich, J. (2015). An analytic solution to the equations of the motion of a point mass with quadratic resistance and generalizations. *Archive of Applied Mechanics*, 85(4), 395–414. <https://doi.org/10.1007/s00419-014-0919-x>
- Redish, E. F. & Wilson, J. M. (1993). Student programming in the introductory physics course: M.U.P.P.E.T. *American Journal of Physics*, 61(3), 222–232.
<https://doi.org/10.1119/1.17295>
- Rigney, D. (2010). *The Matthew effect: How advantage begets further advantage*. Columbia University Press.
- Rodrigues, M. & Simeão Carvalho, P. (2013). Teaching physics with Angry Birds : exploring the kinematics and dynamics of the game. *Physics Education*, 48(4), 431–437. <https://doi.org/10.1088/0031-9120/48/4/431>

- Rodrigues, S., Pearce, J. & Livett, M. (2001). Using Video Analysis or Data loggers During Practical Work in First Year Physics. *Educational Studies*, 27(1), 31–43. <https://doi.org/10.1080/03055690020002116>
- Rogosa, D. (1988). Myths about longitudinal research. In K. Schaie, R. Campbell, W. Meredith & S. Rawlings (Hrsg.), *Methodological issues in aging research* (S. 171–209). Springer Publishing Company.
- Rosenbaum, P. & Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55. <https://doi.org/10.1093/biomet/70.1.41>
- Salkind, N. (2010). Cohen's f Statistic. In N. Salkind (Hrsg.), *Encyclopedia of research design*. Sage. <https://doi.org/10.4135/9781412961288.n59>
- Salvador, S. & Chan, P. (2004). Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms. In *16th IEEE International Conference on Tools with Artificial Intelligence: Proceedings : ICTAI 2004 : 15-17 November, 2004, Boca Raton, Florida* (S. 576–584). IEEE Computer Society. <https://doi.org/10.1109/ICTAI.2004.50>
- Sander, F. (2000). *Verbindung von Theorie und Experiment im physikalischen Praktikum: Eine empirische Untersuchung zum handlungsbezogenen Vorverständnis und dem Einsatz grafikorientierter Modellbildung im Praktikum*. Zugl.: Bremen, Univ., Diss., 1999. *Studien zum Physiklernen: Bd. 13*. Logos-Verl.
- Sander, F., Schecker, H. & Niedderer, H. (2002). Wirkungen des Einsatzes grafikorientierter Modellbildung im physikalischen Praktikum. *Zeitschrift für Didaktik der Naturwissenschaften*, 7, 147–165.
- Savinainen, A. & Scott, P. (2002). The Force Concept Inventory: a tool for monitoring student learning. *Physics Education*, 37(1), 45–52. <https://doi.org/10.1088/0031-9120/37/1/306>
- Sawtelle, V., Brewe, E., Kramer, L. H., Singh, C., Sabella, M. & Rebello, S. (2010). Positive Impacts of Modeling Instruction on Self-Efficacy. In C. Singh, M. Sabella & S. Rebello (Hrsg.), *AIP Conference Proceedings, 2010 Physics Education Research Conference* (S. 289–292). AIP. <https://doi.org/10.1063/1.3515225>

- Schecker, H. (1983). Schülerinteressen und Schülervorstellungen zur Mechanik. In H. Mikalskis (Vorsitz), *Zur Didaktik der Physik und Chemie, Probleme und Perspektive, Vorträge auf der Tagung für Chemie/Physik in Siegen*, Siegen.
- Schecker, H. (1985). *Das Schülervorverständnis zur Mechanik: Eine Untersuchung in der Sekundarstufe II unter Einbeziehung historischer und wissenschaftstheoretischer Aspekte* [Dissertation]. Universität Bremen, Bremen.
- Schecker, H. (1992). Computer als Werkzeuge im Physikunterricht. In K. H. Wiebel (Hrsg.), *Zur Didaktik der Physik und Chemie. Probleme und Perspektiven* (S. 149–151). Leuchtturm-Verlag.
- Schecker, H. (1993a). The Didactic Potential of Computer Aided Modeling for Physics Education. In D. L. Ferguson (Hrsg.), *Advanced Educational Technologies for Mathematics and Science* (S. 165–207). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-02938-1_6
- Schecker, H. (1993b). Learning physics by making models. *Physics Education*, 28(2), 102–106. <https://doi.org/10.1088/0031-9120/28/2/007>
- Schecker, H. (1998a). *Physik - Modellieren: Grafikorientierte Modellbildungssysteme im Physikunterricht* (1. Aufl.). Naturwissenschaftliche Reihe. Klett.
- Schecker, H. (1998b). Physiklernen mit Modellbildungssystemen - Forschungskonzeption. In H. Berendt (Hrsg.), *Zur Didaktik der Physik und Chemie. Probleme und Perspektiven* (S. 230–232). Leuchtturm-Verlag.
- Schecker, H. & Gerdes, J. (1998). Interviews über Experimente zu Bewegungsvorgängen. *Zeitschrift für Didaktik der Naturwissenschaften*, 4(3), 61–74.
- Schecker, H. & Gerdes, J. (1999a). Messung der Konzeptualisierungsfähigkeit in der Mechanik: Zur Aussagekraft des Force Concept Inventory. *Zeitschrift für Didaktik der Naturwissenschaften*, 5(1), 75–89.
- Schecker, H. & Gerdes, J. (1999b). Messung von Konzeptualisierungsfähigkeit in der Mechanik - Zur Aussagekraft des Force Concept Inventory. *Zeitschrift für Didaktik der Naturwissenschaften*, 5(1), 75–89.
- Schecker, H., Klieme, E., Niedderer, H., Ebach, J. & Gerdes, J. (1999). Abschlussbericht zum DFG-Projekt - Physiklernen mit Modellbildungssystemen: Förderung

- physikalischer Kompetenz und systemischen Denkens durch computergestützte Modellbildungssysteme.
- Schecker, H. & Niedderer, H. (1991). Über den Beitrag des Computereinsatzes zur Entwicklung physikalischer Kompetenz. In K. H. Wiebel (Hrsg.), *Zur Didaktik der Physik und Chemie. Probleme und Perspektiven* (S. 147–149). Leuchtturm-Verlag.
- Schecker, H. & Wilhelm, T. (2018). Schülervorstellungen in der Mechanik. In H. Schecker, T. Wilhelm, M. Hopf & R. Duit (Hrsg.), *Schülervorstellungen und Physikunterricht* (S. 63–88). Springer Berlin Heidelberg.
- Schecker, H., Wilhelm, T., Hopf, M. & Duit, R. (Hrsg.). (2018). *Schülervorstellungen und Physikunterricht*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-57270-2>
- Scherer, D., Dubois, P. & Sherwood, B. (2000). VPython: 3D interactive scientific graphics for students. *Computing in Science & Engineering*, 2(5), 56–62. <https://doi.org/10.1109/5992.877397>
- Schnotz, W. (Hrsg.). (1999). *Advances in learning and instruction series. New perspectives on conceptual change*. Pergamon.
- Schreier, M. (2012). *Qualitative content analysis in practice*. Sage.
- Schreier, M. (2014). Ways of Doing Qualitative Content Analysis: Disentangling Terms and Terminologies, 15(1). <https://doi.org/10.17169/FQS-15.1.2043> (Forum Qualitative Sozialforschung / Forum: Qualitative Social Research).
- Skorsetz, N., Wagner, S., Wilhelm, T. & Priemer, B. (2021). Brain Type und Nawi-Interesse von Grundschullehramtsstudierenden. In S. Habig (Hrsg.), *Naturwissenschaftlicher Unterricht und Lehrerbildung im Umbruch?* (S. 101–104). Universität Duisburg-Essen.
- Stamann, C., Janssen, M. & Schreier, M. (2016). Searching for the Core: Defining Qualitative Content Analysis, 17(3). <https://doi.org/10.17169/FQS-17.3.2581> (Forum Qualitative Sozialforschung / Forum: Qualitative Social Research).
- Stoer, J. & Bulirsch, R. (2005). *Numerische Mathematik 2: Eine Einführung — unter Berücksichtigung von Vorlesungen von F.L. Bauer*. SpringerLink Bücher. Springer Berlin Heidelberg. <https://doi.org/10.1007/b137272>

- Struck, W. & Yerrick, R. (2010). The Effect of Data Acquisition-Probeware and Digital Video Analysis on Accurate Graphical Representation of Kinetics in a High School Physics Class. *Journal of Science Education and Technology*, 19(2), 199–211. <https://doi.org/10.1007/s10956-009-9194-y>
- Sugar, C. A. & James, G. M. (2003). Finding the Number of Clusters in a Dataset. *Journal of the American Statistical Association*, 98(463), 750–763. <https://doi.org/10.1198/016214503000000666>
- Suleder, M. (2010). *Videoanalyse und Physikunterricht: Technik, Didaktik, Unterrichtspraxis. Unterrichtshilfen Naturwissenschaften*. Aulis-Verl.
- Suleder, M. (2020). Eine kurze Geschichte der Videoanalyse. *Plus Lucis*, 28(1), 4–6. https://www.pluslucis.org/ZeitschriftenArchiv/2020-1_PL.pdf (Datum des letzten Abrufs: 02.09.2022).
- Sung, Y.-T., Chang, K.-E. & Liu, T.-C. (2016). The effects of integrating mobile devices with teaching and learning on students' learning performance: A meta-analysis and research synthesis. *Computers & Education*, 94, 252–275. <https://doi.org/10.1016/j.compedu.2015.11.008>
- Suri, V. (2021). *Vergleich von Videoanalyse-Apps auf Tablet für den Physikunterricht* [Examensarbeit für das erste Staatsexamen]. Goethe-Universität Frankfurt, Frankfurt am Main. www.thomas-wilhelm.net/arbeiten/Videoanalyseapps.pdf (Datum des letzten Abrufs: 02.09.2022).
- Sweller, J. (2003). Evolution of human cognitive architecture. In *Psychology of Learning and Motivation* (Bd. 43, S. 215–266). Academic Press. [https://doi.org/10.1016/S0079-7421\(03\)01015-6](https://doi.org/10.1016/S0079-7421(03)01015-6)
- Sweller, J., Ayres, P. & Kalyuga, S. (2011). *Cognitive Load Theory* (1. Aufl.). *Explorations in the Learning Sciences, Instructional Systems and Performance Technologies: Bd. 1*. Springer Science+Business Media LLC. <https://doi.org/10.1007/978-1-4419-8126-4>
- Sweller, J. & Chandler, P. (1994). Why Some Material Is Difficult to Learn. *Cognition and Instruction*, 12(3), 185–233. https://doi.org/10.1207/s1532690xc1203_1
- Sweller, J., van Merriënboer, J. J. G. & Paas, F. G. W. C. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review*, 10(3), 251–296. <https://doi.org/10.1023/A:1022193728205>

- Teodoro, V. D. (2004). Playing Newtonian games with Modellus. *Physics Education*, 39(5), 421–428. <https://doi.org/10.1088/0031-9120/39/5/005>
- Teodoro, V. D. & Neves, R. G. (2011). Mathematical modelling in science and mathematics education. *Computer Physics Communications*, 182(1), 8–10. <https://doi.org/10.1016/j.cpc.2010.05.021>
- Thorndike, R. L. (1953). Who belongs in the family? *Psychometrika*, 18(4), 267–276. <https://doi.org/10.1007/BF02289263>
- Thornton, R. K., Kuhl, D., Cummings, K. & Marx, J. (2009). Comparing the force and motion conceptual evaluation and the force concept inventory. *Physical Review Special Topics - Physics Education Research*, 5(1). <https://doi.org/10.1103/PhysRevSTPER.5.010105>
- Thornton, R. K. & Sokoloff, D. R. (1998). Assessing student learning of Newton's laws: The Force and Motion Conceptual Evaluation and the Evaluation of Active Learning Laboratory and Lecture Curricula. *American Journal of Physics*, 66(4), 338–352. <https://doi.org/10.1119/1.18863>
- Tibshirani, R., Walther, G. & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2), 411–423. <https://doi.org/10.1111/1467-9868.00293>
- Tinker, R. F. (1993). Modelling and Theory Building: Technology in Support of Student Theorizing. In D. L. Ferguson (Hrsg.), *Advanced Educational Technologies for Mathematics and Science* (S. 91–113). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-02938-1_3
- Tomczak, M. & Tomczak, E. (2014). The need to report effect size estimates revisited. An overview of some recommended measures of effect size. *Trends in Sport Sciences*, 21(1), 19–25.
- Trier, U. & Upmeyer zu Belzen, A. (2009). "Wissenschaftler nutzen Modelle, um etwas Neues zu entdecken, und in der Schule lernt man einfach nur, dass es so ist." Schülervorstellungen zu Modellen. *Erkenntnisweg Biologiedidaktik*(8), 23–38. http://www.biologie.fu-berlin.de/arbeitsgruppen/didaktik/Erkenntnisweg/2009/2009-02_Trier-Upmeyer_zu_Belzen.pdf (Datum des letzten Abrufs: 02.09.2022).

- Twisk, J. W. R. (2006). *Applied multilevel analysis: A practical guide. Practical guides to biostatistics and epidemiology*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511610806>
- Tymms, P. (2004). Effect sizes in multilevel models. In I. Schagen & K. Elliot (Hrsg.), *But what does it mean? The use of effect sizes in educational research*. National Foundation for Educational Research.
- Tymms, P., Merrell, C. & Henderson, B. (1997). The First Year at School: A Quantitative Investigation of the Attainment and Progress of Pupils*. *Educational Research and Evaluation*, 3(2), 101–118. <https://doi.org/10.1080/1380361970030201>
- Upmeyer zu Belzen, A. & Krüger, D. (2010). Modellkompetenz im Biologieunterricht. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 41–57.
- van Dongen, P. (2021). *Klassische Mechanik: Von der Newton'schen Mechanik zur Relativitätstheorie in drei Postulaten* (1. Aufl. 2021). Springer Berlin Heidelberg.
- Vockell, E. L. & Lobonc, S. (1981). Sex-role stereotyping by high school females in science. *Journal of Research in Science Teaching*, 18(3), 209–219. <https://doi.org/10.1002/tea.3660180304>
- Vollmer, G. (2008). *Die Erkenntnis der Natur: Beiträge zur modernen Naturphilosophie ; mit 15 Tabellen* (4. Aufl.). *Was können wir wissen? / Gerhard Vollmer: Bd. 2*. Hirzel.
- Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction*, 4(1), 45–69. [https://doi.org/10.1016/0959-4752\(94\)90018-3](https://doi.org/10.1016/0959-4752(94)90018-3)
- Waltner, C., Tobias, V., Wiesner, H. & Wilhelm, T. (2010). Ein Unterrichtskonzept zur Einführung in die Dynamik in der Mittelstufe. *Praxis der Naturwissenschaften - Physik in der Schule*, 59(7), 9–22.
- Warren, J. (1979). *Understanding Force*. Übersetzt aus dem Englischen von Backhaus, U. & Schneider, T. (1998). Universität Koblenz.
- Weber, E. (2009). Quantifying Student Learning: How to Analyze Assessment Data. *Bulletin of the Ecological Society of America*, 90(4), 501–511. <https://doi.org/10.1890/0012-9623-90.4.501>
- Weber, J. (2017). *Kombination von Videoanalyse mit mathematischer Modellbildung* [Wissenschaftliche Hausarbeit im Rahmen der Ersten Staatsprüfung für das

- Lehramt an Gymnasien im Fach Physik]. Goethe-Universität Frankfurt, Frankfurt am Main. <http://www.thomas-wilhelm.net/arbeiten/modellbildung+videoanalyse.htm> (Datum des letzten Abrufs: 02.09.2022).
- Weber, J. & Wilhelm, T. (2018). Vergleich von modellierten Daten mit Videoanalysedaten mit verschiedener Software. *Plus Lucis*, 26(4), 18–25. https://www.pluslucis.org/ZeitschriftenArchiv/2018-4_PL.pdf (Datum des letzten Abrufs: 02.09.2022).
- Weber, J. & Wilhelm, T. (2020a). The benefit of computational modelling in physics teaching: A historical overview. *European Journal of Physics*(41), 034003. <https://doi.org/10.1088/1361-6404/ab7a7f>
- Weber, J. & Wilhelm, T. (2020b). Videoanalyse in Echtzeit. *Plus Lucis*, 28(1), 32–38. https://www.pluslucis.org/ZeitschriftenArchiv/2020-1_PL.pdf (Datum des letzten Abrufs: 02.09.2022).
- Wee, L. K., Chew, C., Goh, G. H., Tan, S. & Lee, T. L. (2012). Using Tracker as a pedagogical tool for understanding projectile motion. *Physics Education*, 47(4), 448–455. <https://doi.org/10.1088/0031-9120/47/4/448>
- Wee, L. K., Tan, K. K., Leong, T. K. & Tan, C. (2015). Using Tracker to understand ‘toss up’ and free fall motion: a case study. *Physics Education*, 50(4), 436–442. <https://doi.org/10.1088/0031-9120/50/4/436>
- Weidenmann, B. (2002). Multicodierung und Multimedialität im Lernprozess. In L. J. Issing & P. Klimsa (Hrsg.), *Beltz PVU. Information und Lernen mit Multimedia und Internet: Lehrbuch für Studium und Praxis* (3. Aufl.). Beltz PVU.
- Weise, G. (1975). *Intelligenz, Konzentration, spezielle Fähigkeiten. Psychologische Leistungstests: ein Handbuch für Studium und Praxis / von Georg Weise ; Bd. 1.* Hogrefe.
- Wenzel, M. (2018). *Computereinsatz in Schule und Schülerlabor*. Dissertation. *Studien zum Physik- und Chemielernen: Band 251*. Logos-Verl.
- Wiesner, H., Tobias, V., Waltner, C., Hopf, M., Wilhelm, T. & Sen, A. (2010). Dynamik in den Mechanikunterricht. *PhyDid B - Didaktik der Physik - Beiträge zur DPG-Frühjahrstagung*.

- Wilhelm, T. (2000). Der alte Fallkegel - modern behandelt. Quantitative Auswertung eines klassischen Freihandversuchs. *Praxis der Naturwissenschaften. Physik*, 49(6), 28–31.
- Wilhelm, T. (2005a). *Konzeption und Evaluation eines Kinematik/Dynamik-Lehrgangs zur Veränderung von Schülervorstellungen mit Hilfe dynamisch ikonischer Repräsentationen und graphischer Modellbildung*. Dissertation. *Studien zum Physik- und Chemielernen: Bd. 46*. Logos-Verl.
- Wilhelm, T. (2005b). Verständnis der Newtonschen Mechanik bei bayerischen Elftklässlern - Ergebnisse beim Test "Force Concept Inventory" in herkömmlichen Klassen und im Würzburger Kinematik-/Dynamikunterricht. *PhyDid A - Physik und Didaktik in Schule und Hochschule*, 2(4), 47–56.
- Wilhelm, T. (2007). Vektorverständnis von Studienanfängern. In V. Nordmeier, A. Oberländer & H. Grötzebauch (Hrsg.), *Didaktik der Physik: Regensburg 2007*. Lehmanns Media LOB.de.
- Wilhelm, T. (2018a). Was ist eine gute Elementarisierung? In T. Wilhelm (Hrsg.), *Stolpersteine überwinden im Physikunterricht: Anregungen für fachgerechte Elementarisierungen* (S. 6–8). Aulis/Friedrich.
- Wilhelm, T. (2018b). Zur Geschichte der mathematischen Modellbildung im Physikunterricht. *Plus Lucis*, 26(4), 4–10. https://www.pluslucis.org/ZeitschriftenArchiv/2018-4_PL.pdf (Datum des letzten Abrufs: 02.09.2022).
- Wilhelm, T. (Hrsg.). (2020). *Stolpersteine überwinden im Physikunterricht: Anregungen zu fachgerechten Elementarisierungen* (2. Auflage). Aulis Verlag in Friedrich Verlag GmbH.
- Wilhelm, T. & Gemici, B. (2017). Beschleunigungsverständnis in der Oberstufe. In H. Grötzebauch (Vorsitz), *DPG-Frühjahrstagung*, Dresden.
- Wilhelm, T. & Schecker, H. (2018). Strategien für den Umgang mit Schülervorstellungen. In H. Schecker, T. Wilhelm, M. Hopf & R. Duit (Hrsg.), *Schülervorstellungen und Physikunterricht* (S. 39–61). Springer Berlin Heidelberg.
- Wilhelm, T. & Schecker, H. (2021). Unterrichtskonzeptionen zur Numerischen Physik. In T. Wilhelm, H. Schecker & M. Hopf (Hrsg.), *Unterrichtskonzeptionen für den Physikunterricht: Ein Lehrbuch für Studium, Referendariat und Unterrichtspraxis* (1. Aufl., S. 121–150). Springer Berlin Heidelberg.

- Wilhelm, T., Schecker, H. & Hopf, M. (Hrsg.). (2021). *Unterrichtskonzeptionen für den Physikunterricht: Ein Lehrbuch für Studium, Referendariat und Unterrichtspraxis* (1. Aufl. 2021). Springer Berlin Heidelberg.
- Wirtz, M. A. & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität: Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen*. Hogrefe Verl. für Psychologie.
- Wu, H. & Leung, S.-O. (2017). Can Likert Scales be Treated as Interval Scales?—A Simulation Study. *Journal of Social Service Research*, 43(4), 527–532. <https://doi.org/10.1080/01488376.2017.1329775>
- Zumann, J. P. & Weaver, S. L. (1988). Tools for teaching problem solving: An evaluation of a Modelling and systems thinking approach: Paper presented at the 1988 Annual NARST Conference.

17 Anhang

17.1 Vortest

MUSTER

EvaSys	Vortest Hauptstudie	Electric Paper Druckkopierpapier
Goethe-Universität Frankfurt	Jannis Weber	
Institut für Didaktik der Physik	Schülerlabor Dynamik	

Bitte so markieren: ☐ ☒ ☐ ☐ ☐ Bitte verwenden Sie einen Kugelschreiber oder nicht zu starken Filzstift. Dieser Fragebogen wird maschinell erfasst.
 Korrektur: ☐ ☒ ☐ ☐ Bitte beachten Sie im Interesse einer optimalen Datenerfassung die links gegebenen Hinweise beim Ausfüllen.

Allgemeines

Hallo! Vielen Dank, für das Ausfüllen des Fragebogens. Du nimmst dadurch an einem Forschungsprojekt zum Thema Mechanik teil, das den Schulunterricht in Mechanik verbessern soll. Einen ähnlichen Fragebogen wirst du nach dem Besuch des Schülerlabors der Uni Frankfurt noch einmal bekommen.
 Es ist wichtig, dass du den Test konzentriert und ehrlich ausfüllst. In jeder Teilfrage ist nur ein Kreuz zu setzen. Die Daten werden anonym ausgewertet. Der persönliche Code ist nur dafür da, das Vortest-Ergebnis mit dem Nachtest-Ergebnis vergleichen zu können.

Trage zunächst die Buchstaben bzw. Ziffern mit folgenden Regeln in die Felder ein:
 1. Erste und zweite Stelle: Die ersten beiden Buchstaben des Vornamens deiner Mutter.
 2. Dritte und vierte Stelle: Die ersten beiden Buchstaben des Vornamens deines Vaters.
 3. Fünfte und sechste Stelle: Die ersten beiden Ziffern deines Geburtstages.

Beispiel:
 Name der Mutter: Svenja
 Name des Vaters: Peter
 Geburtstag: 22.04.2003
 Code: "SVPE22"

Geschlecht		
☐ m	☐ w	☐ d
Alter		
☐ <14	☐ 14	☐ 15
☐ 16	☐ 17	☐ >17

Trage bitte deine letzte Schulnote in Mathematik in Punkten ein:

Trage bitte deine letzte Schulnote in Physik in Punkten ein:

Teil I

Denke an das Fach Physik im Allgemeinen.

Ich traue mir im Fach Physik viel zu.	trifft gar nicht zu ☐ ☐ ☐ ☐ ☐ trifft völlig zu
Für mich ist an der Physik besonders interessant, Versuche durchzuführen und auszuwerten.	trifft gar nicht zu ☐ ☐ ☐ ☐ ☐ trifft völlig zu
Mir macht Physikunterricht Spaß.	trifft gar nicht zu ☐ ☐ ☐ ☐ ☐ trifft völlig zu
Ich mag es, im Physikunterricht Vorhersagen aufzustellen und Abläufe zu erklären.	trifft gar nicht zu ☐ ☐ ☐ ☐ ☐ trifft völlig zu
Es gibt viele Themen im Physikunterricht, die mir egal sind.	trifft gar nicht zu ☐ ☐ ☐ ☐ ☐ trifft völlig zu

F18035U0P1PL0V0
15.01.2020, Seite 1/10

MUSTER

MUSTER

EvaSys	Vortest Hauptstudie	Electric Paper
Teil I [Fortsetzung]		
Ich experimentiere lieber, als mir zu überlegen, warum es so ist.	trifft gar nicht zu	☐ ☐ ☐ ☐ ☐ trifft völlig zu
Das Fach Physik liegt mir nicht.	trifft gar nicht zu	☐ ☐ ☐ ☐ ☐ trifft völlig zu
Ich würde im Physikunterricht gerne mehr Hypothesen über physikalische Zusammenhänge aufstellen.	trifft gar nicht zu	☐ ☐ ☐ ☐ ☐ trifft völlig zu
Ich lerne in Physik etwas, das für mich sehr wichtig ist.	trifft gar nicht zu	☐ ☐ ☐ ☐ ☐ trifft völlig zu
Ich würde im Physikunterricht gerne mehr experimentieren.	trifft gar nicht zu	☐ ☐ ☐ ☐ ☐ trifft völlig zu
Für das Fach Physik habe ich einfach keine Begabung.	trifft gar nicht zu	☐ ☐ ☐ ☐ ☐ trifft völlig zu
Ich denke lieber nach, warum es so ist, als eine Messung durchzuführen.	trifft gar nicht zu	☐ ☐ ☐ ☐ ☐ trifft völlig zu
Das Fach Physik werde ich nie richtig verstehen.	trifft gar nicht zu	☐ ☐ ☐ ☐ ☐ trifft völlig zu
Ich mag es, im Physikunterricht Experimente durchzuführen.	trifft gar nicht zu	☐ ☐ ☐ ☐ ☐ trifft völlig zu
Physikunterricht finde ich sehr interessant.	trifft gar nicht zu	☐ ☐ ☐ ☐ ☐ trifft völlig zu
Für mich ist an der Physik besonders interessant, mir theoretische Überlegungen über physikalische Zusammenhänge zu machen.	trifft gar nicht zu	☐ ☐ ☐ ☐ ☐ trifft völlig zu
Teil II		
Kreuze auf der Skala an, inwiefern du der jeweiligen Aussage zur Physik zustimmst.		
Modelle in der Physik dienen dazu, die Natur besser zu verstehen.	trifft gar nicht zu	☐ ☐ ☐ ☐ ☐ trifft völlig zu
Im naturwissenschaftlichen Erkenntnisprozess sind Modelle wichtig.	trifft gar nicht zu	☐ ☐ ☐ ☐ ☐ trifft völlig zu
Modelle werden für einen bestimmten Zweck erschaffen.	trifft gar nicht zu	☐ ☐ ☐ ☐ ☐ trifft völlig zu
Theoretische Überlegungen in der Physik dienen dazu die Natur besser zu verstehen.	trifft gar nicht zu	☐ ☐ ☐ ☐ ☐ trifft völlig zu
Modelle sind idealisierte Beschreibungen der Natur.	trifft gar nicht zu	☐ ☐ ☐ ☐ ☐ trifft völlig zu
Modelle in der Physik dienen dazu, neue Sachverhalte zu erschließen.	trifft gar nicht zu	☐ ☐ ☐ ☐ ☐ trifft völlig zu
Modelle in der Physik dienen dazu, bekanntes Wissen zu beschreiben.	trifft gar nicht zu	☐ ☐ ☐ ☐ ☐ trifft völlig zu
Teil III		
Ich beschäftige mich gerne mit dem Computer, Computer machen Spaß.	trifft gar nicht zu	☐ ☐ ☐ ☐ ☐ trifft völlig zu
Im Physikunterricht sollte mehr mit Computern gemacht werden.	trifft gar nicht zu	☐ ☐ ☐ ☐ ☐ trifft völlig zu
Ich denke, ich kann mit Computern gut lernen.	trifft gar nicht zu	☐ ☐ ☐ ☐ ☐ trifft völlig zu
Ich lerne Physik gut, wenn ich den Computer nutze.	trifft gar nicht zu	☐ ☐ ☐ ☐ ☐ trifft völlig zu

MUSTER

MUSTER

EvaSys

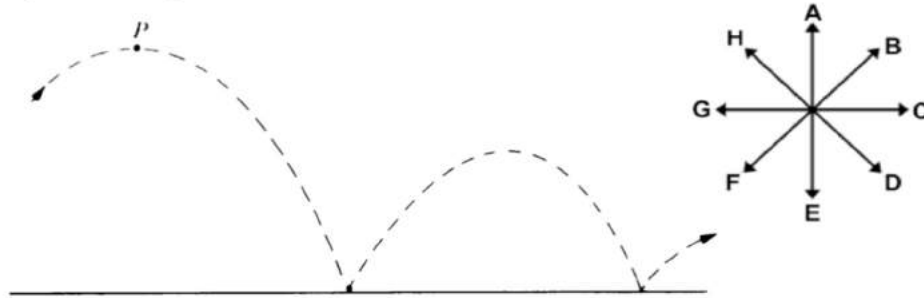
Vortest Hauptstudie

 Electric Paper
elektronische Notizblätter

Teil IV

Frage 1

Betrachte das untere Bild. In welche Richtung zeigt im Punkt P die auf den Körper wirkende **Gesamtkraft**? Reibungskräfte sollen ignoriert werden. Wähle aus den Pfeilen A bis H einen aus.



Welchem Pfeil entspricht die Gesamtkraft im Punkt P?

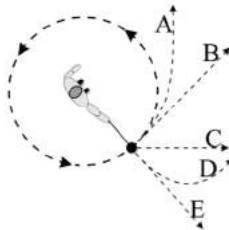
- ☐ A
☐ D
☐ G

- ☐ B
☐ E
☐ H

- ☐ C
☐ F

Frage 2

a) Ein schwerer Ball ist an einem Faden befestigt und wird, wie in der Abbildung gezeigt, im Kreis horizontal herumgeschwungen. An dem gekennzeichneten Punkt reißt jedoch plötzlich der Faden. Der Vorgang wird von oben betrachtet: Welchen Weg nimmt der Ball, nachdem der Faden gerissen ist?



- ☐ A
☐ D

- ☐ B
☐ E

- ☐ C

b) Welche Aussage erklärt deine Antwort am besten? Es geht um Kräfte **nach** Reißen des Fadens.

- ☐ Die Kugel wird durch die Zentrifugalkraft nach außen geschleudert.

- ☐ Die Kugel bewegt sich in einem Bogen, da sie vorher im Kreis geschwungen wurde.

- ☐ Es wirkt nach dem Reißen des Fadens keine Kraft auf die Kugel (in der horizontalen Ebene).

- ☐ Die Zentrifugalkraft und eine Kraft in Bewegungsrichtung wirken auf den Ball.

- ☐ Es wirkt nur eine Kraft in Bewegungsrichtung.

MUSTER

MUSTER

EvaSys

Vortest Hauptstudie



Frage 3

Ein Vogel fliegt mit konstantem Tempo geradeaus. Was kann man über die auf den Vogel wirkenden Kräfte sagen?

- ☐ Man hat nicht genug Informationen, um etwas über die Kräfte sagen zu können.
 ☐ Es kann keine Kraft auf den Vogel wirken.
 ☐ Alle wirkenden Kräfte müssen sich zu Null addieren. Die Gesamtkraft ist also Null.
- ☐ Die Gesamtkraft muss in Bewegungsrichtung wirken.

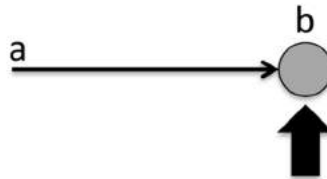
Frage 4

Stelle dir einen Frontalzusammenstoß zwischen einem Lastwagen (LKW) und einem Kleinwagen (PKW) vor. Für den Zeitraum des Zusammenpralls gilt:

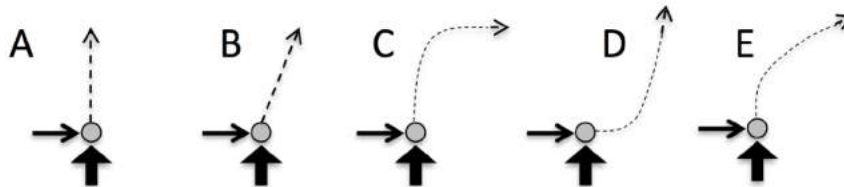
- ☐ Der LKW übt eine größere Kraft auf den PKW aus als der PKW auf den LKW.
 ☐ Der PKW übt eine größere Kraft auf den LKW aus als der LKW auf den PKW.
 ☐ Die beiden Fahrzeuge üben keine Kräfte aufeinander aus. Der PKW wird einfach deshalb zusammen gedrückt, weil der dem LKW im Wege ist.
- ☐ Der LKW übt eine Kraft auf den PKW aus, aber der PKW übt keine Kraft auf den LKW aus.
 ☐ Der LKW übt die gleiche Kraft auf den PKW aus wie der PKW auf den LKW.

Frage 5

Den folgenden Fragen a) und b) liegt folgende Situation zugrunde: Die Abbildung zeigt eine Scheibe, die mit konstanter Geschwindigkeit auf einer reibungsfreien horizontalen Oberfläche (Tisch) von Punkt „a“ nach Punkt „b“ gleitet. Wenn die Scheibe „b“ erreicht, erhält sie einen kurzzeitigen horizontalen „Kick“ in Richtung des dicken Pfeils.



a) Welche der gezeigten Bahnkurven beschreibt die Bewegung der Scheibe nach dem "Kick"?



- ☐ A
 ☐ B
 ☐ C
- ☐ D
 ☐ E

b) Wie entwickelt sich das Tempo der Scheibe bei einer reibungsfreien Bewegung **nach** dem "Kick"?

- ☐ Keine Änderung
 ☐ Stetig zunehmend
 ☐ Stetig abnehmend
- ☐ Eine Zeit lang zunehmend, danach abnehmend
 ☐ Eine Zeit lang konstant, dann abnehmend

MUSTER

MUSTER

EvaSys

Vortest Hauptstudie



Frage 6

Ein Fallschirmspringer bewegt sich mit konstanter Geschwindigkeit in Richtung Erde. Es wirkt auf ihn die nach unten gerichtete Gravitationskraft und eine nach oben gerichtete Luftreibungskraft.

a) Welche Aussage kannst du über die Stärke der Kräfte treffen?

- ☐ Die Gravitationskraft ist stärker als die Luftreibungskraft. ☐ Die Luftreibungskraft ist stärker als die Gravitationskraft. ☐ Beide Kräfte sind gleich stark.

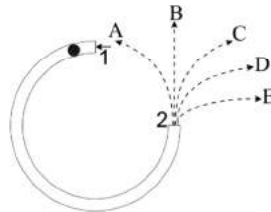
b) Welche Aussage erklärt deine Antwort am besten?

- ☐ Bei einer konstanten Geschwindigkeit muss die Summe aller Kräfte Null sein. ☐ Die Kraft in Bewegungsrichtung muss immer am stärksten sein, damit sich etwas bewegen kann. ☐ Die abbremssende Kraft muss größer sein als die antreibende, da der Springer sonst immer schneller werden würde.

Frage 7

Die Abbildung zeigt einen von oben betrachteten kreisförmigen Kanal, der in der horizontalen Ebene fest auf dem Tisch verankert ist. Ein Ball tritt bei "1" in den Kanal ein und verlässt ihn bei "2".

a) Welche der in der Abbildung gezeigten Bahnkurven beschreibt am besten den Weg des Balls, wenn er den Kanal bei "2" verlässt und weiter über den Tisch rollt?



- ☐ A ☐ B ☐ C
☐ D ☐ E

b) Welche Aussage erklärt deine Antwort am besten?

- ☐ Die Zentrifugalkraft drückt den Ball nach außen. ☐ Die Kugel bewegt sich weiter im Kreis, da sie sich vorher so bewegt hat. ☐ Es wirkt keine Kraft auf die Kugel, also muss sie geradeaus rollen.
☐ Es wirkt eine Kraft in Bewegungsrichtung.

Frage 8

Ein Mädchen fährt mit dem Fahrrad auf einem ebenen Weg mit einem konstanten Tempo geradeaus. Warum muss sie dafür in die Pedalen treten?

- ☐ Wenn auf das Fahrrad keine Kraft wirken würde, würde es langsamer werden und irgendwann stehen bleiben. ☐ Es muss eine nach vorne gerichtete Kraft auf das Fahrrad erzeugt werden, um die Reibungskräfte zu überwinden. Dabei muss die nach vorne gerichtete Kraft größer sein als die Reibungskräfte. ☐ Es muss eine nach vorne gerichtete Kraft auf das Fahrrad erzeugt werden, um die Reibungskräfte auszugleichen. Dabei muss die nach vorne gerichtete Kraft genau so groß sein, wie die Reibungskräfte.

MUSTER

MUSTER

EvaSys

Vortest Hauptstudie



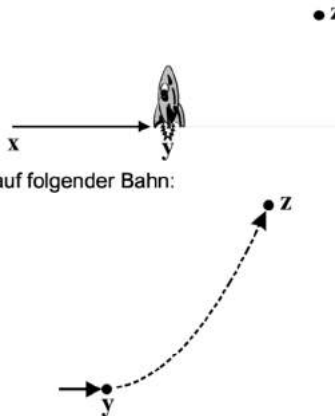
Frage 9

Ein Junge wirft eine Stahlkugel schräg nach oben. Welche der folgenden Kräfte wirken während der Flugphase auf die Kugel?

- A: Eine stetig abnehmende Kraft nach oben.
 B: Eine stetig zunehmende Gewichtskraft nach unten.
 C: Eine Kraft, die in Bewegungsrichtung der Kugel zeigt.
 D: Eine der Bewegungsrichtung entgegenwirkende Luftreibungskraft.
 E: Eine konstante nach unten gerichtete Gewichtskraft.
- ☐ A und B ☐ Nur C ☐ D und E
☐ C und E ☐ Es wirken keine Kräfte. Die Kugel fällt auf die Erde zurück, weil das ihrem natürlichen Verhalten entspricht.

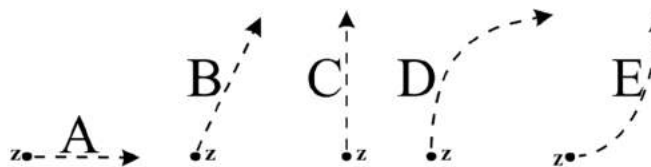
Frage 10

Für die Fragen a) und b) liegt folgende Situation zugrunde: Ein Raumschiff, das sich im Weltraum von Position "x" nach Position "y" bewegt, wird zunächst nicht von Kräften beeinflusst. Am Punkt "y" wird der Raketenmotor gezündet. Er erzeugt eine konstante Kraft im rechten Winkel zu Linie "xy". Nachdem der Punkt "z" erreicht ist, wird der Motor wieder abgeschaltet.



Dabei bewegt sich das Raumschiff auf folgender Bahn:

a) Bei "z" wird der Raketenmotor abgeschaltet. Welche Bahnkurve wird das Raumschiff danach beschreiben?



- ☐ A ☐ B ☐ C
☐ D ☐ E
- b) Hinter dem Punkt "z"...
- ☐ ... bleibt das Tempo konstant. ☐ ... nimmt das Tempo stetig ab. ☐ ... nimmt das Tempo stetig zu.
☐ ... nimmt das Tempo eine Zeit lang zu und bleibt dann konstant. ☐ ... bleibt das Tempo eine Zeit lang konstant und nimmt dann ab.

MUSTER

MUSTER

EvaSys

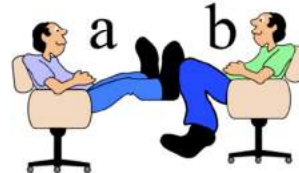
Vortest Hauptstudie



Frage 11

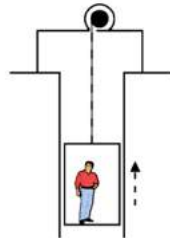
Zwei Schüler, von denen Schüler "a" die Masse 95 kg und Schüler "b" die Masse 77 kg hat, sitzen sich in zwei gleichen Bürostühlen direkt gegenüber. Schüler "a" stellt seine Füße auf die Knie von Schüler "b" (s. Abbildung). Schüler "b" hat seine Füße angehoben. Plötzlich streckt Schüler "a" seine Beine aus, wodurch beide Stühle in Bewegung versetzt werden. Welche Aussage trifft zu:

- ☐ Keiner der Schüler übt eine Kraft auf den anderen aus.
- ☐ Schüler "a" übt eine Kraft auf "b" aus, aber "b" übt keine Kraft auf "a" aus.
- ☐ Jeder der Schüler übt eine Kraft auf den anderen aus, aber "b" übt die größere Kraft aus.
- ☐ Jeder der Schüler übt eine Kraft auf den anderen aus, aber "a" übt die größere Kraft aus.
- ☐ Jeder der Schüler übt eine gleich starke Kraft auf den anderen aus.



Frage 12

Ein Fahrstuhl wird in einem Fahrstuhlschacht über ein Stahlseil heraufgezogen. Die Kabine bewegt sich während des Heraufziehens mit konstanter Geschwindigkeit. Reibungskräfte sind hier so gering, dass sie vernachlässigt werden können.



Welche der folgenden Aussagen beschreibt die Bewegung korrekt? Dabei geht es nur um die Kräfte, die wirken, während sich die Kabine mit konstanter Geschwindigkeit bewegt.

- ☐ Die nach oben gerichtete Kraft durch das Stahlseil ist größer als die nach unten gerichtete Gravitationskraft.
- ☐ Die nach oben gerichtete Kraft durch das Stahlseil ist genauso stark wie die nach unten gerichtete Gravitationskraft.
- ☐ Die nach oben gerichtete Kraft durch das Stahlseil ist kleiner als die nach unten gerichtete Gravitationskraft.
- ☐ Die Kabine bewegt sich nach oben, weil das Seil kürzer wird. Es wirken keine Kräfte auf die Kabine.

MUSTER

MUSTER

EvaSys

Vortest Hauptstudie



Frage 13

Für die folgenden Aufgaben gilt folgende Situation: Ein LKW bleibt mit Motorschaden liegen und wird von einem Kleinwagen (Auto) zur nächsten Tankstelle geschoben.



- a) Während das Auto beschleunigt, um beim Schieben auf eine bestimmte Geschwindigkeit zu kommen, gilt:
- ☐ Die Stärke der Kraft, mit der das Auto gegen den LKW drückt, ist genauso groß wie die Kraft, mit der der LKW gegen das Auto drückt.
 - ☐ Die Stärke der Kraft, mit der das Auto gegen den LKW drückt, ist kleiner als die Kraft, mit der der LKW gegen das Auto drückt.
 - ☐ Die Stärke der Kraft, mit der das Auto gegen den LKW drückt, ist größer als die Kraft, mit der der LKW gegen das Auto drückt.
 - ☐ Der Motor des Autos läuft; daher übt das Auto eine Kraft aus, während es gegen den LKW drückt. Aber der LKW-Motor läuft nicht; daher kann der LKW nicht gegen das Auto drücken.
 - ☐ Weder der LKW noch das Auto üben eine Kraft aus. Der LKW wird einfach deshalb nach vorne geschoben, weil er dem Auto im Wege steht.

b) Nachdem der Fahrer des Autos die gewünschte Geschwindigkeit erreicht hat, mit der er den Lastwagen konstant vor sich herschieben will, fährt er mit konstanter Geschwindigkeit weiter. Jetzt gilt:

- ☐ Die Stärke der Kraft, mit der das Auto gegen den LKW drückt, ist genauso groß wie die Kraft, mit der der LKW gegen das Auto drückt.
- ☐ Die Stärke der Kraft, mit der das Auto gegen den LKW drückt, ist kleiner als die Kraft, mit der der LKW gegen das Auto drückt.
- ☐ Die Stärke der Kraft, mit der das Auto gegen den LKW drückt, ist größer als die Kraft, mit der der LKW gegen das Auto drückt.
- ☐ Der Motor des Autos läuft; daher übt das Auto eine Kraft aus, während es gegen den LKW drückt. Aber der LKW-Motor läuft nicht; daher kann der LKW nicht gegen das Auto drücken.
- ☐ Weder der LKW noch das Auto üben eine Kraft aus. Der LKW wird einfach deshalb nach vorne geschoben, weil er dem Auto im Wege steht.

Frage 14

Ein Golfball bewegt sich längs eines "Fairway" mit folgender Flugbahn durch die Luft.



Welche Kraft wirkt, bzw. welche Kräfte wirken, während der gesamten Flugphase auf den Ball, also **nachdem** der Ball den Schläger verlassen hat?

1. Gravitationskraft
2. Abschlagskraft
3. Luftwiderstandskraft

- ☐ Nur 1
- ☐ 1 und 3

- ☐ 1 und 2
- ☐ 2 und 3

- ☐ 1, 2 und 3

MUSTER

MUSTER

EvaSys

Vortest Hauptstudie

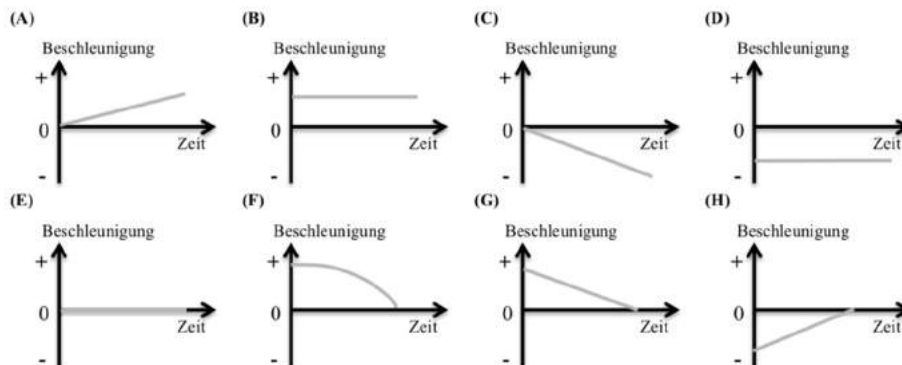
 Electric Paper
elektronisches Notizbuch

Frage 15

Die folgenden Fragen beziehen sich auf ein Spielzeugauto, das sich nach rechts oder links entlang einer horizontalen Linie (der x-Achse des Koordinatensystems) bewegen kann.



Unterschiedliche Bewegungen des Autos sind unten beschrieben. Wähle den einen Buchstaben der folgenden Möglichkeiten (A bis H) des Beschleunigungs-Zeit-Graphen, der zu der Bewegung des Autos passt, die in jeder Aufgabe beschrieben ist.



a) Das Auto bewegt sich mit konstanter Geschwindigkeit nach rechts (in positiver Richtung).

☐ A	☐ B	☐ C
☐ D	☐ E	☐ F
☐ G	☐ H	☐ Keiner der Graphen.

b) Das Auto bewegt sich nach rechts (in positiver Richtung) und wird gleichmäßig immer schneller.

☐ A	☐ B	☐ C
☐ D	☐ E	☐ F
☐ G	☐ H	☐ Keiner der Graphen.

c) Das Auto bewegt sich nach rechts (in positiver Richtung) und wird gleichmäßig immer langsamer.

☐ A	☐ B	☐ C
☐ D	☐ E	☐ F
☐ G	☐ H	☐ Keiner der Graphen.

d) Das Auto bewegt sich mit konstanter Geschwindigkeit nach links (in negativer Richtung).

☐ A	☐ B	☐ C
☐ D	☐ E	☐ F
☐ G	☐ H	☐ Keiner der Graphen.

MUSTER

MUSTER

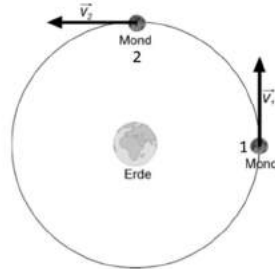
EvaSys

Vortest Hauptstudie

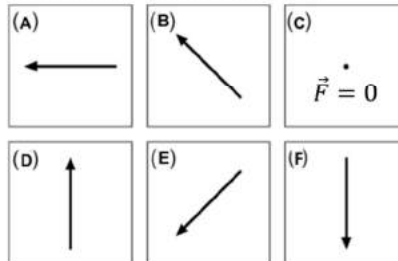
Electric Paper

Frage 16

Das Bild stellt die Bewegung des Mondes um die Erde dar. Der Mond ist zu den Zeitpunkten "1" und "2" dargestellt.



Wähle in den folgenden beiden Fragen aus den gegebenen Möglichkeiten A-F jeweils einen Pfeil aus, der die **Kraft**, die auf den Mond wirkt, zu dem Zeitpunkt am besten darstellt.



a) Zeitpunkt 1

☐ A☐ D☐ B☐ E☐ C☐ F

b) Zeitpunkt 2

☐ A☐ D☐ B☐ E☐ C☐ F

Vielen Dank!

17.2 Nachtest

EvaSys		Nachtest Hauptstudie		Electric Paper elektronisches Papier
Goethe-Universität Frankfurt		Jannis Weber		
Institut für Didaktik der Physik		Schülerlabor Dynamik		

Bitte so markieren: ☐ ☒ ☐ ☐ Bitte verwenden Sie einen Kugelschreiber oder nicht zu starken Filzstift. Dieser Fragebogen wird maschinell erfasst.
 Korrektur: ☐ ☒ ☐ ☐ Bitte beachten Sie im Interesse einer optimalen Datenerfassung die links gegebenen Hinweise beim Ausfüllen.

Allgemeines

Hallo! Vielen Dank, dass du an dem Schülerlabor teilgenommen hast. Du wirst nun erneut einen ähnlichen Fragebogen erhalten, wie du ihn bereits kennst.
 Es ist wichtig, dass du den Fragebogen konzentriert und ehrlich ausfüllst. In jeder Teilfrage ist nur ein Kreuz zu setzen. Die Daten werden anonym ausgewertet. Der persönliche Code ist nur dafür da, das Vortest-Ergebnis mit dem Nachtest-Ergebnis vergleichen zu können.

Trage zunächst die Buchstaben bzw. Ziffern mit folgenden Regeln in die Felder ein:

1. Erste und zweite Stelle: Die ersten beiden Buchstaben des Vornamens deiner Mutter.
2. Dritte und vierte Stelle: Die ersten beiden Buchstaben des Vornamens deines Vaters.
3. Fünfte und sechste Stelle: Die ersten beiden Ziffern deines Geburtstages.

Beispiel:
 Name der Mutter: Svenja
 Name des Vaters: Peter
 Geburtstag: 22.04.2003
 Code: "SVPE22"

Geschlecht		
☐ m	☐ w	☐ d
Alter		
☐ <14	☐ 14	☐ 15
☐ 16	☐ 17	☐ >17

Teil I

Denke an die heutige Physikeinheit.

Der Tag hat mir Freude gemacht.	trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu nicht zu
Ich fand das Thema wichtig.	trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu nicht zu
Das Programm ist gut geeignet, um Physik zu lernen.	trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu nicht zu
Ich war mit dem Tag zufrieden.	trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu nicht zu
Ich würde das Programm gerne auch im Schulunterricht einsetzen.	trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu nicht zu
Der Unterricht hat mir Spaß gemacht.	trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu nicht zu
Ich konnte mich schnell in das Programm einarbeiten.	trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu nicht zu
Was ich über das Thema erfahren habe, bringt mir was.	trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu nicht zu
Das Programm ist übersichtlich und leicht zu bedienen.	trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu nicht zu
Es war für mich eine gute Stunde.	trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu nicht zu

F17942U0P1PL0V0
15.01.2020, Seite 1/10

MUSTER

MUSTER

EvaSys	Nachtest Hauptstudie					

Teil I [Fortsetzung]

Ich möchte mehr über das Thema erfahren. trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu
nicht zu

Ich hatte Schwierigkeiten bei der Nutzung des Programms. trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu
nicht zu

Teil II

Kreuze auf der Skala an, inwiefern du der jeweiligen Aussage zur Physik zustimmst.

Modelle in der Physik dienen dazu, die Natur besser zu verstehen. trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu
nicht zu

Im naturwissenschaftlichen Erkenntnisprozess sind Modelle wichtig. trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu
nicht zu

Modelle werden für einen bestimmten Zweck erschaffen. trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu
nicht zu

Theoretische Überlegungen in der Physik dienen dazu die Natur besser zu verstehen. trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu
nicht zu

Modelle sind idealisierte Beschreibungen der Natur. trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu
nicht zu

Modelle in der Physik dienen dazu, neue Sachverhalte zu erschließen. trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu
nicht zu

Modelle in der Physik dienen dazu, bekanntes Wissen zu beschreiben. trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu
nicht zu

Teil III

Kreuze auf der Skala an, inwiefern du der jeweiligen Aussage zum heutigen Tag zustimmst.

Das Thema „Dynamik“ war sehr komplex. trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu
nicht zu

Die Bearbeitung der Aufgaben hat wirklich mein Verständnis des Themas „Dynamik“ verbessert. trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu
nicht zu

Die Formeln zur „Dynamik“ empfand ich als sehr komplex. trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu
nicht zu

Die Bearbeitung der Aufgaben hat wirklich mein Verständnis der Physik der „Dynamik“ verbessert. trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu
nicht zu

Die Aufgaben und Arbeitsaufträge waren sehr unklar formuliert. trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu
nicht zu

Die Bearbeitung der Aufgaben hat wirklich mein Verständnis der Formeln zur „Dynamik“ verbessert. trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu
nicht zu

Die Aufgaben und Arbeitsaufträge waren zum Lernen sehr ineffektiv. trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu
nicht zu

Die Aufgaben und Arbeitsaufträge waren voll von unklaren Begriffen. trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu
nicht zu

Die physikalischen Konzepte und Definitionen zur „Dynamik“ empfand ich als sehr komplex. trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu
nicht zu

Die Bearbeitung der Aufgaben hat wirklich mein Verständnis der physikalischen Konzepte und Definitionen zur „Dynamik“ verbessert. trifft gar ☐ ☐ ☐ ☐ ☐ trifft völlig zu
nicht zu

MUSTER

MUSTER

EvaSys

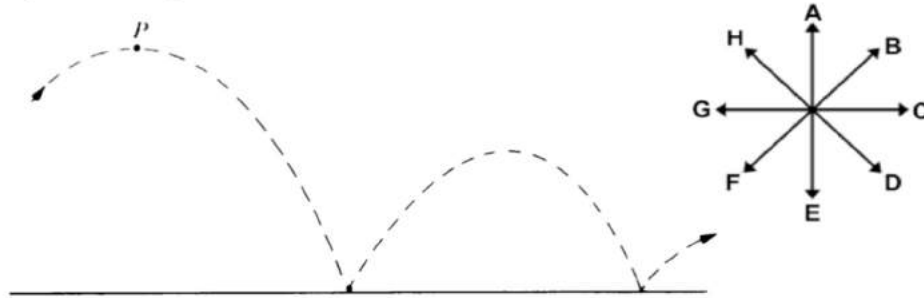
Nachtest Hauptstudie

 Electric Paper
elektronische Notizblätter

Teil IV - Mechanik

Frage 1

Betrachte das untere Bild. In welche Richtung zeigt im Punkt P die auf den Körper wirkende **Gesamtkraft**? Reibungskräfte sollen ignoriert werden. Wähle aus den Pfeilen A bis H einen aus.



Welchem Pfeil entspricht die Gesamtkraft im Punkt P?

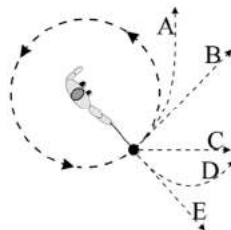
- ☐ A
☐ D
☐ G

- ☐ B
☐ E
☐ H

- ☐ C
☐ F

Frage 2

a) Ein schwerer Ball ist an einem Faden befestigt und wird, wie in der Abbildung gezeigt, im Kreis horizontal herumgeschwungen. An dem gekennzeichneten Punkt reißt jedoch plötzlich der Faden. Der Vorgang wird von oben betrachtet: Welchen Weg nimmt der Ball, nachdem der Faden gerissen ist?



- ☐ A
☐ D

- ☐ B
☐ E

- ☐ C

b) Welche Aussage erklärt deine Antwort am besten? Es geht um Kräfte **nach** Reißen des Fadens.

☐ Die Kugel wird durch die Zentrifugalkraft nach außen geschleudert.

☐ Die Kugel bewegt sich in einem Bogen, da sie vorher im Kreis geschwungen wurde.

☐ Es wirkt nach dem Reißen des Fadens keine Kraft auf die Kugel (in der horizontalen Ebene).

☐ Die Zentrifugalkraft und eine Kraft in Bewegungsrichtung wirken auf den Ball.

☐ Es wirkt nur eine Kraft in Bewegungsrichtung.

MUSTER

MUSTER

EvaSys

Nachtest Hauptstudie

Electric Paper
Technik

Frage 3

Ein Vogel fliegt mit konstantem Tempo geradeaus. Was kann man über die auf den Vogel wirkenden Kräfte sagen?

- ☐ Man hat nicht genug Informationen, um etwas über die Kräfte sagen zu können.
 ☐ Es kann keine Kraft auf den Vogel wirken.
 ☐ Alle wirkenden Kräfte müssen sich zu Null addieren. Die Gesamtkraft ist also Null.
- ☐ Die Gesamtkraft muss in Bewegungsrichtung wirken.

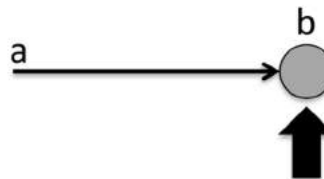
Frage 4

Stelle dir einen Frontalzusammenstoß zwischen einem Lastwagen (LKW) und einem Kleinwagen (PKW) vor. Für den Zeitraum des Zusammenpralls gilt:

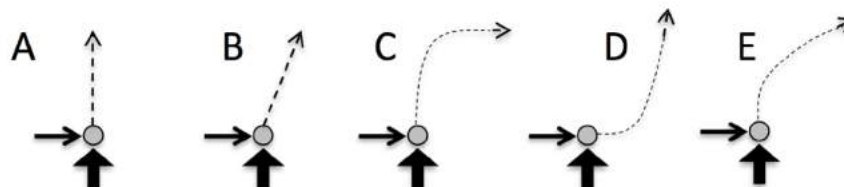
- ☐ Der LKW übt eine größere Kraft auf den PKW aus als der PKW auf den LKW.
 ☐ Der PKW übt eine größere Kraft auf den LKW aus als der LKW auf den PKW.
 ☐ Die beiden Fahrzeuge üben keine Kräfte aufeinander aus. Der PKW wird einfach deshalb zusammen gedrückt, weil der dem LKW im Wege ist.
- ☐ Der LKW übt eine Kraft auf den PKW aus, aber der PKW übt keine Kraft auf den LKW aus.
 ☐ Der LKW übt die gleiche Kraft auf den PKW aus wie der PKW auf den LKW.

Frage 5

Den folgenden Fragen a) und b) liegt folgende Situation zugrunde: Die Abbildung zeigt eine Scheibe, die mit konstanter Geschwindigkeit auf einer reibungsfreien horizontalen Oberfläche (Tisch) von Punkt „a“ nach Punkt „b“ gleitet. Wenn die Scheibe „b“ erreicht, erhält sie einen kurzzeitigen horizontalen „Kick“ in Richtung des dicken Pfeils.



a) Welche der gezeigten Bahnkurven beschreibt die Bewegung der Scheibe nach dem "Kick"?



- ☐ A
 ☐ B
 ☐ C
- ☐ D
 ☐ E

b) Wie entwickelt sich das Tempo der Scheibe bei einer reibungsfreien Bewegung **nach** dem "Kick"?

- ☐ Keine Änderung
 ☐ Stetig zunehmend
 ☐ Stetig abnehmend
- ☐ Eine Zeit lang zunehmend, danach abnehmend
 ☐ Eine Zeit lang konstant, dann abnehmend

MUSTER

EvaSys

Nachtest Hauptstudie



Frage 6

Ein Fallschirmspringer bewegt sich mit konstanter Geschwindigkeit in Richtung Erde. Es wirkt auf ihn die nach unten gerichtete Gravitationskraft und eine nach oben gerichtete Luftreibungskraft.

a) Welche Aussage kannst du über die Stärke der Kräfte treffen?

- ☐ Die Gravitationskraft ist stärker als die Luftreibungskraft. ☐ Die Luftreibungskraft ist stärker als die Gravitationskraft. ☐ Beide Kräfte sind gleich stark.

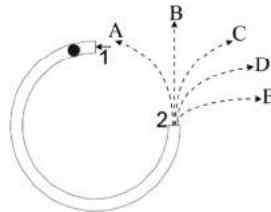
b) Welche Aussage erklärt deine Antwort am besten?

- ☐ Bei einer konstanten Geschwindigkeit muss die Summe aller Kräfte Null sein. ☐ Die Kraft in Bewegungsrichtung muss immer am stärksten sein, damit sich etwas bewegen kann. ☐ Die abbremssende Kraft muss größer sein als die antreibende, da der Springer sonst immer schneller werden würde.

Frage 7

Die Abbildung zeigt einen von oben betrachteten kreisförmigen Kanal, der in der horizontalen Ebene fest auf dem Tisch verankert ist. Ein Ball tritt bei "1" in den Kanal ein und verlässt ihn bei "2".

a) Welche der in der Abbildung gezeigten Bahnkurven beschreibt am besten den Weg des Balls, wenn er den Kanal bei "2" verlässt und weiter über den Tisch rollt?



- ☐ A ☐ B ☐ C
☐ D ☐ E

b) Welche Aussage erklärt deine Antwort am besten?

- ☐ Die Zentrifugalkraft drückt den Ball nach außen. ☐ Die Kugel bewegt sich weiter im Kreis, da sie sich vorher so bewegt hat. ☐ Es wirkt keine Kraft auf die Kugel, also muss sie geradeaus rollen.
☐ Es wirkt eine Kraft in Bewegungsrichtung.

Frage 8

Ein Mädchen fährt mit dem Fahrrad auf einem ebenen Weg mit einem konstanten Tempo geradeaus. Warum muss sie dafür in die Pedalen treten?

- ☐ Wenn auf das Fahrrad keine Kraft wirken würde, würde es langsamer werden und irgendwann stehen bleiben. ☐ Es muss eine nach vorne gerichtete Kraft auf das Fahrrad erzeugt werden, um die Reibungskräfte zu überwinden. Dabei muss die nach vorne gerichtete Kraft größer sein als die Reibungskräfte. ☐ Es muss eine nach vorne gerichtete Kraft auf das Fahrrad erzeugt werden, um die Reibungskräfte auszugleichen. Dabei muss die nach vorne gerichtete Kraft genau so groß sein, wie die Reibungskräfte.

MUSTER

MUSTER

EvaSys

Nachtest Hauptstudie



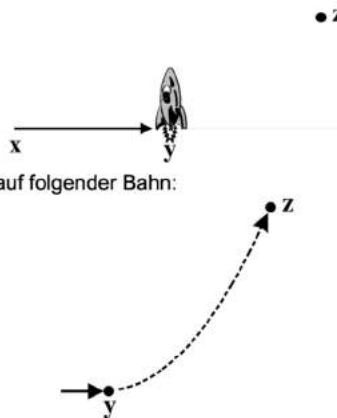
Frage 9

Ein Junge wirft eine Stahlkugel schräg nach oben. Welche der folgenden Kräfte wirken während der Flugphase auf die Kugel?

- A: Eine stetig abnehmende Kraft nach oben.
 B: Eine stetig zunehmende Gewichtskraft nach unten.
 C: Eine Kraft, die in Bewegungsrichtung der Kugel zeigt.
 D: Eine der Bewegungsrichtung entgegenwirkende Luftreibungskraft.
 E: Eine konstante nach unten gerichtete Gewichtskraft.
- ☐ A und B ☐ Nur C ☐ D und E
☐ C und E ☐ Es wirken keine Kräfte. Die Kugel fällt auf die Erde zurück, weil das ihrem natürlichen Verhalten entspricht.

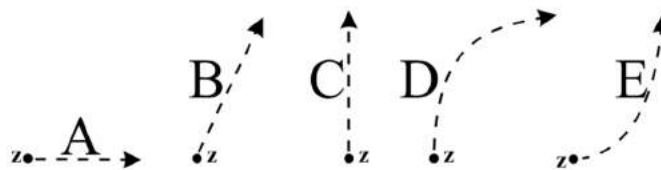
Frage 10

Für die Fragen a) und b) liegt folgende Situation zugrunde: Ein Raumschiff, das sich im Weltraum von Position "x" nach Position "y" bewegt, wird zunächst nicht von Kräften beeinflusst. Am Punkt "y" wird der Raketenmotor gezündet. Er erzeugt eine konstante Kraft im rechten Winkel zu Linie "xy". Nachdem der Punkt "z" erreicht ist, wird der Motor wieder abgeschaltet.



Dabei bewegt sich das Raumschiff auf folgender Bahn:

a) Bei "z" wird der Raketenmotor abgeschaltet. Welche Bahnkurve wird das Raumschiff danach beschreiben?



- ☐ A ☐ B ☐ C
☐ D ☐ E
- b) Hinter dem Punkt "z"...
☐ ... bleibt das Tempo konstant. ☐ ... nimmt das Tempo stetig ab. ☐ ... nimmt das Tempo stetig zu.
☐ ... nimmt das Tempo eine Zeit lang zu und bleibt dann konstant. ☐ ... bleibt das Tempo eine Zeit lang konstant und nimmt dann ab.

MUSTER

MUSTER

EvaSys

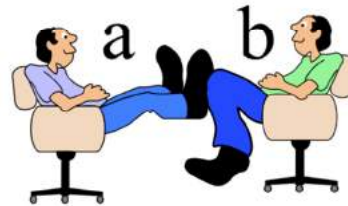
Nachtest Hauptstudie



Frage 11

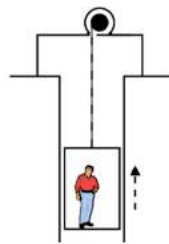
Zwei Schüler, von denen Schüler "a" die Masse 95 kg und Schüler "b" die Masse 77 kg hat, sitzen sich in zwei gleichen Bürostühlen direkt gegenüber. Schüler "a" stellt seine Füße auf die Knie von Schüler "b" (s. Abbildung). Schüler "b" hat seine Füße angehoben. Plötzlich streckt Schüler "a" seine Beine aus, wodurch beide Stühle in Bewegung versetzt werden. Welche Aussage trifft zu:

- ☐ Keiner der Schüler übt eine Kraft auf den anderen aus.
- ☐ Ein Schüler "a" übt eine Kraft auf "b" aus, aber "b" übt keine Kraft auf "a" aus.
- ☐ Jeder der Schüler übt eine Kraft auf den anderen aus, aber "b" übt die größere Kraft aus.
- ☐ Jeder der Schüler übt eine Kraft auf den anderen aus, aber "a" übt die größere Kraft aus.
- ☐ Jeder der Schüler übt eine gleich starke Kraft auf den anderen aus.



Frage 12

Ein Fahrstuhl wird in einem Fahrstuhlschacht über ein Stahlseil heraufgezogen. Die Kabine bewegt sich während des Heraufziehens mit konstanter Geschwindigkeit. Reibungskräfte sind hier so gering, dass sie vernachlässigt werden können.



Welche der folgenden Aussagen beschreibt die Bewegung korrekt? Dabei geht es nur um die Kräfte, die wirken, während sich die Kabine mit konstanter Geschwindigkeit bewegt.

- ☐ Die nach oben gerichtete Kraft durch das Stahlseil ist größer als die nach unten gerichtete Gravitationskraft.
- ☐ Die nach oben gerichtete Kraft durch das Stahlseil ist genauso stark wie die nach unten gerichtete Gravitationskraft.
- ☐ Die nach oben gerichtete Kraft durch das Stahlseil ist kleiner als die nach unten gerichtete Gravitationskraft.
- ☐ Die Kabine bewegt sich nach oben, weil das Seil kürzer wird. Es wirken keine Kräfte auf die Kabine.

MUSTER

MUSTER

EvaSys

Nachtest Hauptstudie



Frage 13

Für die folgenden Aufgaben gilt folgende Situation: Ein LKW bleibt mit Motorschaden liegen und wird von einem Kleinwagen (Auto) zur nächsten Tankstelle geschoben.



- a) Während das Auto beschleunigt, um beim Schieben auf eine bestimmte Geschwindigkeit zu kommen, gilt:
- ☐ Die Stärke der Kraft, mit der das Auto gegen den LKW drückt, ist genauso groß wie die Kraft, mit der der LKW gegen das Auto drückt.
 - ☐ Die Stärke der Kraft, mit der das Auto gegen den LKW drückt, ist kleiner als die Kraft, mit der der LKW gegen das Auto drückt.
 - ☐ Die Stärke der Kraft, mit der das Auto gegen den LKW drückt, ist größer als die Kraft, mit der der LKW gegen das Auto drückt.
 - ☐ Der Motor des Autos läuft; daher übt das Auto eine Kraft aus, während es gegen den LKW drückt. Aber der LKW-Motor läuft nicht; daher kann der LKW nicht gegen das Auto drücken.
 - ☐ Weder der LKW noch das Auto üben eine Kraft aus. Der LKW wird einfach deshalb nach vorne geschoben, weil er dem Auto im Wege steht.
- b) Nachdem der Fahrer des Autos die gewünschte Geschwindigkeit erreicht hat, mit der er den Lastwagen konstant vor sich herschieben will, fährt er mit konstanter Geschwindigkeit weiter. Jetzt gilt:
- ☐ Die Stärke der Kraft, mit der das Auto gegen den LKW drückt, ist genauso groß wie die Kraft, mit der der LKW gegen das Auto drückt.
 - ☐ Die Stärke der Kraft, mit der das Auto gegen den LKW drückt, ist kleiner als die Kraft, mit der der LKW gegen das Auto drückt.
 - ☐ Die Stärke der Kraft, mit der das Auto gegen den LKW drückt, ist größer als die Kraft, mit der der LKW gegen das Auto drückt.
 - ☐ Der Motor des Autos läuft; daher übt das Auto eine Kraft aus, während es gegen den LKW drückt. Aber der LKW-Motor läuft nicht; daher kann der LKW nicht gegen das Auto drücken.
 - ☐ Weder der LKW noch das Auto üben eine Kraft aus. Der LKW wird einfach deshalb nach vorne geschoben, weil er dem Auto im Wege steht.

Frage 14

Ein Golfball bewegt sich längs eines "Fairway" mit folgender Flugbahn durch die Luft.



Welche Kraft wirkt, bzw. welche Kräfte wirken, während der gesamten Flugphase auf den Ball, also **nachdem** der Ball den Schläger verlassen hat?

1. Gravitationskraft
2. Abschlagskraft
3. Luftwiderstandskraft

- ☐ Nur 1
☐ 1 und 3

- ☐ 1 und 2
☐ 2 und 3

- ☐ 1, 2 und 3

MUSTER

MUSTER

EvaSys

Nachtest Hauptstudie

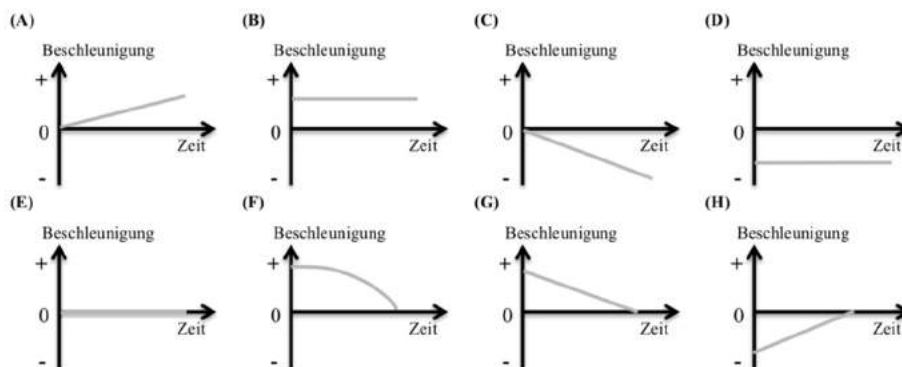
Electric Paper

Frage 15

Die folgenden Fragen beziehen sich auf ein Spielzeugauto, das sich nach rechts oder links entlang einer horizontalen Linie (der x-Achse des Koordinatensystems) bewegen kann.



Unterschiedliche Bewegungen des Autos sind unten beschrieben. Wähle den einen Buchstaben der folgenden Möglichkeiten (A bis H) des Beschleunigungs-Zeit-Graphen, der zu der Bewegung des Autos passt, die in jeder Aufgabe beschrieben ist.



a) Das Auto bewegt sich mit konstanter Geschwindigkeit nach rechts (in positiver Richtung).

☐ A	☐ B	☐ C
☐ D	☐ E	☐ F
☐ G	☐ H	☐ Keiner der Graphen.

b) Das Auto bewegt sich nach rechts (in positiver Richtung) und wird gleichmäßig immer schneller.

☐ A	☐ B	☐ C
☐ D	☐ E	☐ F
☐ G	☐ H	☐ Keiner der Graphen.

c) Das Auto bewegt sich nach rechts (in positiver Richtung) und wird gleichmäßig immer langsamer.

☐ A	☐ B	☐ C
☐ D	☐ E	☐ F
☐ G	☐ H	☐ Keiner der Graphen.

d) Das Auto bewegt sich mit konstanter Geschwindigkeit nach links (in negativer Richtung).

☐ A	☐ B	☐ C
☐ D	☐ E	☐ F
☐ G	☐ H	☐ Keiner der Graphen.

MUSTER

MUSTER

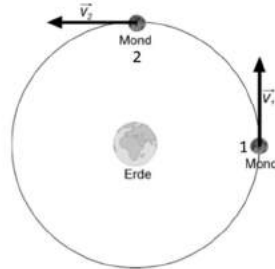
EvaSys

Nachtest Hauptstudie

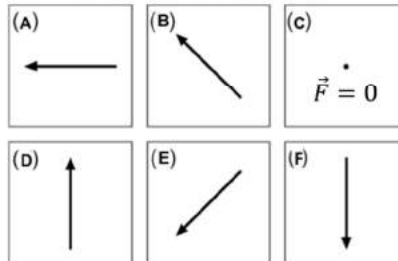
Electric Paper

Frage 16

Das Bild stellt die Bewegung des Mondes um die Erde dar. Der Mond ist zu den Zeitpunkten "1" und "2" dargestellt.



Wähle in den folgenden beiden Fragen aus den gegebenen Möglichkeiten A-F jeweils einen Pfeil aus, der die **Kraft**, die auf den Mond wirkt, zu dem Zeitpunkt am besten darstellt.



a) Zeitpunkt 1

- ☐ A
☐ D

- ☐ B
☐ E

- ☐ C
☐ F

b) Zeitpunkt 2

- ☐ A
☐ D

- ☐ B
☐ E

- ☐ C
☐ F

Vielen Dank!

17.3 Kodiermanual

Dimension 1: Vorgehen

Vorgehen: Deduktiv

Art der Kodierung: Kodierung im Video (Event-Sampling)

Hinweis: Anker-Beispiele werden hier nicht genannt, da nicht der Text, sondern Stellen im Video kodiert wurden, deren Zugehörigkeit zu einer Arbeitsphase sich nicht durch einzelne Textstelle ergibt.

Modellbildung

Phase	Definition	Beschreibung/Kodierregeln	Anmerkungen
1	Problemanalyse	Die Personen sprechen über die Ziele der Modellbildung. Sie sprechen über das Experiment oder was zu tun ist, damit ein Modell erstellt werden kann bzw. wozu das Modell erstellt wird.	Die Problemanalyse wird in der Regel im Plenum besprochen. Zudem folgt die aufgenommene Aufgabe 3 demselben Schema, wie die Aufgaben zuvor. Es ist dementsprechend nicht zu erwarten, dass bei vielen Gruppen eine aktive Phase der Problemanalyse zu erkennen ist. Ein Kodieren dieser Phase ist nicht zwingend notwendig.
2	Modellbildung	Die Personen unterhalten sich über die physikalischen Größen, die für die Erstellung des mathematischen Modells benötigt werden. Insbesondere die Diskussion über die wirkenden Kräfte ist hier zu nennen, aber auch die Nennung von weiteren relevanten Größen, wie Masse oder Anfangsbedingungen.	Die Modellbildung kann ohne oder mit Nutzung des Programms stattfinden. Durch Anleitung im Heft sollen sich die Lernenden zunächst Gedanken über die wirkenden Kräfte machen. Wenn diese direkt eingegeben werden, verläuft die Implementation parallel zur Modellbildung.

3	Implementa- tion (+Simulation)	Die Personen geben die zuvor besprochenen Größen in das Programm ein. Dazu gehören Kräfte, Masse, Anfangsbedingungen und möglicherweise weitere Definitionen.	Die Implementation betrifft nur die aktive Eingabe der Modellgrößen in das Programm. Dies kann aber parallel zur Modellbildung erfolgen, falls die Überlegungen direkt in das Programm implementiert werden. Die Berechnung des Modells erfolgt in manchen Fällen automatisch, sodass auch die Simulation des Modells zu dieser Phase zu zählen ist und nicht einzeln aufgeführt wird.
4	Diagrammba- sierte Evalua- tion	Die Personen analysieren die Diagramme, die das Programm nach der Simulation anzeigt, und überprüfen somit das Modell auf seine Plausibilität. Dabei werden noch keine Messdaten hinzugezogen, sondern nur auf Basis der Graphen des Modells analysiert.	In der Evaluation geht es darum, ob die erste Implementation des Modells plausibel ist. Bewegt sich das Objekt verglichen mit dem Experiment in die richtige Richtung und in der richtigen Art und Weise können Fragen sein, die sich die Lernenden damit beantworten.
5	Überarbei- tung des Mo- dells	Nach der Analyse verbessern die Personen ihr Modell, falls ihnen Fehler aufgefallen sind oder das Modell unzureichend ist.	Diese Phase ist nur notwendig, wenn das Modell nicht zufriedenstellend ist. Ein Fehlen der Phase kann also auf eine fehlende Evaluation oder ein bereits gutes Modell hindeuten. Dies erschließt sich im Kontext. Inhaltlich unterscheidet sich diese Phase kaum von der der Modellbildung. Der einzige Unterschied ist, dass vorher bereits eine Evaluation der Daten stattgefunden hat.

6	Einlesen der Vergleichsdaten	Die Messdaten aus dem Experiment werden in das Programm eingelesen und in die Diagramme eingebunden.	Die Lernenden werden dazu aufgefordert, die Messdaten hinzuzuziehen. Von Interesse ist, an welcher Stelle im Arbeitsprozess dies passiert und wie lang sie dafür benötigen.
7	Datenbasierte Evaluation	Die Personen evaluieren ihr Modell im Vergleich zu den realen Messdaten in demselben Diagramm. Dabei kann auf mögliche Unterschiede eingegangen werden.	Ähnlich wie in Phase 4 wird das Modell evaluiert. Als Vergleich dienen aber nun die bereitgestellten Messwerte. Das führt dazu, dass auch bei einer geringeren Diagrammkompetenz ein Fehler im Modell durch Unterschiede in den Graphen unmittelbar erkannt werden kann.
8	Formulierung der Ergebnisse	Die Personen beantworten die Abschlussfragen über die relevanten Größen des Experiments (Kräfte, Beschleunigung, Geschwindigkeit).	Das Arbeitsheft hält die Lernenden dazu an, am Ende der Aufgabe Aussagen über die Summe aller Kräfte, die Beschleunigung und die Geschwindigkeit in x- und y-Richtung zu tätigen. Dies kann mithilfe des angefertigten Modells oder ohne passieren. Diese beiden Fälle werden unterschieden.
8a	Formulierung der Ergebnisse ohne Modellnutzung	Die Personen formulieren Endergebnisse, ohne dabei auf das Modell zurückzugreifen.	Eine in dem Moment fehlende Nutzung kann auch heißen, dass die Eigenschaften des Modells verinnerlicht wurden. Insbesondere aber wenn das Modell richtig erstellt wurde und die Antworten falsch gegeben werden zeigt dies, dass die Bedeutung des Modells für die Bearbeitung der Aufgaben nicht erfasst wurde.

8b	Formulierung der Ergebnisse mit Modellnutzung	Die Personen formulieren Endergebnisse und greifen dabei aktiv auf die aus dem Modell gewonnenen Kenntnisse zurück und/oder nutzen die verschiedenen Diagramme, die die Software anbietet, um die Fragen beantworten zu können.	Ziel der Erstellung des Modells ist Aussagen über das gezeigte Experiment abzuleiten. Darüber hinaus können bei Änderungen der Parameter Voraussagen über ähnliche Bewegungen getätigt werden. Zur Beantwortung der Abschlussfragen sollten die Erkenntnisse des Modells jedoch in jedem Fall herangezogen werden.

Videoanalyse

Code	Definition	Beschreibung/Kodierregeln	Anmerkungen
1	Erzeugung der Messdaten	Die Personen führen eine Videoanalyse durch. Diese kann automatisch oder manuell erfolgen. Dazu gehört auch die Skalierung und vorbereitende Einstellungen.	Dies Arbeitsphase erfolgt gezwungenermaßen zu Beginn. Falls die Schüler*innen diese Phase vergessen oder hier Schwierigkeiten haben, hat die Software keine Daten zur Verfügung, die in der Folge analysiert werden könnten. Eine Fehlermeldung gibt dann einen Hinweis, dass erst die Videoanalyse erfolgen muss.

2	Darstellung der Messdaten und Interpretation derer	Die Personen stellen die Messdaten auf irgendeine Weise dar und diskutieren darüber. Dazu gehören dynamisch ikonische Repräsentationen, Diagramme und Stroboskopbilder. Diese werden in den folgenden Subkategorien unterschieden.	Nach der Erzeugung von Daten erfolgt deren Darstellung. Die Software bietet verschiedene Darstellungsmöglichkeiten, von denen durch das Arbeitsheft manche explizit gefordert werden, andere aber auch möglich sind. Bspw. Stroboskopbilder müssen bei dieser Aufgabe nicht angefertigt werden, sind aus den vorigen Aufgaben aber bekannt, sodass sie dennoch genutzt werden könnten. Da die Darstellungen der Daten immer mit einer Diskussion darüber und einer Interpretation verbunden sind, werden diese zusammen kodiert.
2a	Stroboskopbild	Die Personen nutzen die Funktion der Stroboskopbilder, um die Bewegung zu analysieren.	Wenn auch nicht erfordert, kann ein Stroboskopbild der Bewegung einen Überblick über die Ortskurve und das Tempo der Bewegung liefern. Falls das Stroboskopbild mit Einblendung von Pfeilen kombiniert wird, wird es beiden Kategorien zugeordnet.
2b	Diagramme	Die Personen nutzen die Funktion der Diagramme, um die Bewegung zu analysieren. Dabei gibt es unterschiedliche Arten der Diagramme, die in der Folge nochmals in Subkategorien unterteilt werden.	Die Software ist in der Lage jede der gemessenen bzw. berechneten Größen gegen jede andere aufzutragen. Mit Ausnahme der Ortskurve ist in der Regel die Zeit auf der Ordinate, sodass die Diagramme eindeutig Ort, Geschwindigkeit oder Beschleunigung zugeordnet werden können. Für den unwahrscheinlichen Fall, dass eine Gruppe bspw. die Geschwindigkeit in x-Richtung in Abhängigkeit der Beschleunigung in y-Richtung darstellt o. ä., wird dies beiden Diagrammartentypen zugeordnet.

2bi	Ort	Die Personen nutzen und diskutieren über Ortsdiagramme in jeglicher Form.	Zu den Ortsdiagrammen gehören Diagramme, die den Ort des Objekts in einer Koordinatenrichtung in Abhängigkeit der Zeit darstellen und Diagramme, die die Ortskurve darstellen.
2bii	Geschwindigkeit	Die Personen nutzen und diskutieren über Geschwindigkeitsdiagramme in jeglicher Form.	Zu den Geschwindigkeitsdiagrammen gehören Diagramme, die die Geschwindigkeit des Objekts in einer Koordinatenrichtung in Abhängigkeit der Zeit (oder einer anderen Größe) darstellen.
2biii	Beschleunigung	Die Personen nutzen und diskutieren über Beschleunigungsdiagramme in jeglicher Form.	Zu den Geschwindigkeitsdiagrammen gehören Diagramme, die die Beschleunigung des Objekts in einer Koordinatenrichtung in Abhängigkeit der Zeit (oder einer anderen Größe) darstellen.
2c	Einblendungen (dynamisch ikonische Repräsentationen)	Die dritte Art der Darstellung von Messdaten ist die Einblendung von Pfeilen in das Video. Dabei wird zwischen Geschwindigkeit und Beschleunigung unterschieden.	Das Programm ist in der Lage noch weitere Größen in das Video einzublenden. Diese werden aber nicht erfordert und wurden zu Beginn auch nicht gezeigt. Es ist somit nicht zu erwarten, dass Schüler*innen andere Größen als die Geschwindigkeits- und Beschleunigungspfeile nutzen.
2ci	Geschwindigkeitspfeile	Die Personen nutzen und diskutieren über Geschwindigkeitspfeile.	Falls mehrere Pfeilarten gleichzeitig genutzt werden, wird dies beiden Kategorien zugeordnet.
2cii	Beschleunigungspfeile	Die Personen nutzen und diskutieren über Beschleunigungspfeile.	Falls mehrere Pfeilarten gleichzeitig genutzt werden, wird dies beiden Kategorien zugeordnet.
3	Formulierung der Ergebnisse	Die Personen formulieren Endergebnisse. Dabei können die Messdaten genutzt, müssen es aber nicht.	Diese Kategorie umfasst alles, was mit dem Formulieren von Endergebnissen zusammenhängt, die aus der Analyse des Experiments mithilfe der Messdaten hervorgehen.

3a	Formulierung der Ergebnisse ohne Nutzung der Messdaten.	Die Personen formulieren Endergebnisse, ohne dabei auf die Messdaten zurückzugreifen.	Die Analyse des Experiments dient dazu, Informationen über den Ablauf der Bewegung zu erhalten. Dementsprechend sollen die Daten zur Formulierung der Endergebnisse möglichst genutzt werden. In dieser Kategorie werden Stellen kodiert, in denen dies nicht erkennbar ist. Das kann bedeuten, dass die Größen aus der vorigen Analyse verinnerlicht sind oder auch, dass diese durch Diskussion oder direkte Beobachtung des Experiments klar sind. Insbesondere falsche Ergebnisse, die ohne Rückgriff auf die Messdaten formuliert wurden, sprechen dafür, dass nicht verstanden wurde, wozu die Videoanalyse durchgeführt wurde.
3b	Formulierung der Ergebnisse mit Nutzung der Messdaten.	Die Personen formulieren Endergebnisse und greifen dabei aktiv auf die aus den Messdaten gewonnenen Kenntnisse zurück und/oder nutzen die verschiedenen Diagramme, die die Software anbietet, um die Fragen beantworten zu können. Dabei reicht es auch, wenn die entsprechenden Diagramme oder Einblendungen in der Software nicht erneut aufgerufen, aber diskutiert oder beschrieben werden.	Anders als in der vorigen Kategorie werden hier die Erkenntnisse der Arbeit mit der Software genutzt, um die Endergebnisse zu formulieren.

Dimension 2: Schwierigkeiten

Vorgehen: InduktivArt der Kodierung: Kodierung im Video (Event-Sampling)

Hinweis: Da es sich hier um Kodierungen im Video handelt, die nicht zwangsläufig an Aussagen geknüpft, welche auch nur im Kontext zu verstehen wären, werden hier keine Ankerbeispiele genannt.

Code	Definition	Beschreibung/Kodierregeln	Anmerkungen
1	Inhaltliche Schwierigkeiten	In diese Kategorie fallen alle Schwierigkeiten, die die Schüler*innen in ihrem Arbeitsprozesse merklich hindern und fachlicher Natur sind. Diese werden induktiv in Subkategorien unterteilt.	Es geht hier vor allem um Schwierigkeiten, die einerseits den Arbeitsablauf hindern, andererseits aber nicht den Lerngegenstand bzw. die Lernziele direkt betreffen. Schwierigkeiten mit Diagrammen werden dementsprechend in die Kategorie fallen, während Schwierigkeiten mit der Formulierung von Kräften dies nicht tun.
1a	Mangelnde Diagrammkompetenz	Die Personen haben Schwierigkeiten, die sich ergebenden Diagramme zu verstehen, wodurch sie nicht ableiten können, ob das Modell sinnvoll ist bzw. wie sich das Objekt bewegt.	Dies ist die einzige Subkategorie, die bei den inhaltlichen Schwierigkeiten gefunden wurde. Alle anderen Schwierigkeiten, die auftraten, betrafen direkt den Lerngegenstand, sodass diese Schwierigkeiten zu einem fruchtbaren Lernprozess führen können.

2	Schwierigkeiten mit der Software	In diese Kategorie fallen alle Schwierigkeiten, die die Lernenden in deren Arbeitsprozess hindern und die an der (Bedienung der) Software liegen. Auch hier werden induktiv Subkategorien erstellt.	Die Unterteilung zwischen Schwierigkeiten mit der Software und inhaltlichen Schwierigkeiten ist nicht in allen Fällen eindeutig (siehe oben). Hier werden aber vor allem solche Schwierigkeiten kodiert, die primäre mit der Nutzung der Software zu tun haben.
2a	Newton-II	Alle Schwierigkeiten, die bei Benutzung der Software „Newton-II“ aufgetreten sind.	„Newton-II“ und „measure dynamics“ sind in der Bedienung sehr unterschiedlich, sodass auch unterschiedliche Schwierigkeiten zu erwarten sind.
2ai	Eingegebene Kräfte werden nicht mit dem Modell verknüpft.	Hierzu gehören Schwierigkeiten, die damit zusammenhängen, dass Kräfte zwar formuliert, aber nicht mit in die Berechnung des Modells hinzugefügt werden.	Insbesondere Reibungskräfte wurden häufig in das Modellierungsfenster eingegeben, aber nicht mit dem Modell verknüpft. Die kann entweder passieren, in dem die Reibung im Eingabefenster zu den Kräften in x- bzw. y-Richtung hinzugefügt wird oder sie müssen in dem Feld hinzugeführt werden, in dem die Berechnung der Beschleunigung definiert wird.
2aii	Definition von genutzten Größen	Schwierigkeiten, die damit zusammenhängen, dass Größen falsch definiert werden und von der Software nicht berechnet werden können.	Im Modellierungsfenster können eigene Definitionen frei gewählt werden. Wenn im zweidimensionalen Fall das Tempo benötigt wird, muss dieses bspw. als $v = \sqrt{vx^2 + vy^2}$ definiert werden.
2aiii	Probleme mit Messdaten	Schwierigkeiten, die mit dem Einlesen, Anzeigen, Löschen etc. der Messdaten zu tun haben.	Wenn das Einlesen der Messdaten Schwierigkeiten bereitet, die den Arbeitsablauf stören, wird dies hier kodiert.

2aiv	Einstellung der Diagramme	Schwierigkeiten, die mit dem Einstellen der Diagramme zu tun haben. Dazu gehörten die Skalierung und die Wahl der dargestellten Größen.	In den Diagrammen können an beiden Achsen beliebige Größen eingegeben werden. Auch das Fenster bzw. der Zoom kann manuell eingestellt werden.
2av	Eingabe der Anfangsbedingungen	Schwierigkeiten, die mit der Eingabe der Anfangsbedingungen, also Masse, Ort und Anfangsgeschwindigkeit, zusammenhängen.	Die Anfangsbedingungen werden nicht in dem allgemeinen Modellierungsfenster, sondern in dafür vorgesehene Felder darunter eingegeben. Um das Modell mit der Realität vergleichen zu können, sind diese Anfangsbedingungen notwendig.
2avi	Bedingte Variable	Schwierigkeiten, die mit der Erstellung einer bedingten Variablen zu tun haben.	Eine bedingte Variable wurde bei dieser Aufgabe nicht benötigt. In einer der vorigen Aufgaben war sie jedoch notwendig, sodass eine Gruppe auch hier versuchte, eine solche Variable einzugeben.
2avii	Kräfte werden nicht modelliert	Das Modell wird ohne Modellierung von Kräften betrachtet.	Wenn keine Kräfte aber Anfangsbedingungen modelliert werden, wird der entsprechende Graph dargestellt. In dieser Kategorie werden Stellen kodiert, in denen dies dann für das vollständige Modell gehalten wird.
2b	measure dynamics	Alle Schwierigkeiten, die die bei Benutzung der Software „measure dynamics“ aufgetreten sind.	„Newton-II“ und „measure dynamics“ sind in der Bedienung sehr unterschiedlich, sodass auch unterschiedliche Schwierigkeiten zu erwarten sind.
2bi	Skalierung	Schwierigkeiten, die mit der Skalierung vor der Videoanalyse zu tun haben.	Vor der Videoanalyse muss eine Skalierung durchgeführt werden, damit die Software Pixel in Meter umrechnen kann. Dies geschieht mit einem im Bild vorhandenen Meterstab.

2bii	Einstellung der Diagramme	Schwierigkeiten, die mit der Einstellung der Diagramme zu tun haben.	Auch bei der Videoanalyse müssen die Skalierung der Diagramme und die Größen, die dargestellt werden sollen, selbstständig eingestellt werden. Alle Probleme, die damit in Verbindung stehen, werden hier kodiert.
2biii	Durchführung der Videoanalyse	Schwierigkeiten, die mit der Durchführung der Videoanalyse selbst zu tun haben.	Hierzu gehören Probleme, der automatischen Objekterkennung durch Muster-, Farb- oder Bewegungserkennung. Auch das korrekte Anklicken im richtigen Bild spielt eine Rolle.
2biv	Anzeigen der Daten ohne Videoanalyse	Stellen, in denen versucht wird, die Daten zu betrachten, ohne dass diese zuvor durch eine Videoanalyse gewonnen wurden.	Falls keine Videoanalyse durchgeführt wurde, sind auch keine Daten vorhanden. Manche Proband*innen versuchen dennoch, diese nicht vorhandenen Daten in unterschiedlicher Weise darzustellen, ehe sie dann doch die Videoanalyse durchführten.
2bv	Einzeichnen von Pfeilen	Schwierigkeiten, die mit dem Einzeichnen von Pfeilen zur Veranschaulichung der Daten zu tun haben.	Hierzu gehören alle Probleme, die mit dem Darstellen, Skalieren etc. von Pfeilen zusammenhängen.

Dimension 3: Schülervorstellungen

Vorgehen: deduktivArt der Kodierung: Kodierung im Transkript

Hinweis: Da es sich hier um bekannte Schülervorstellungen handelt, die auch in Kapitel 3 besprochen wurden, bedarf es keiner weiteren Anmerkungen bezüglich der jeweiligen Vorstellungen. Diese Spalte ist hier dementsprechend nicht vorhanden. Die Tabelle gibt wieder, nach welchen Schülervorstellungen mit dem Kategoriensystem geschaut wurde. Nach Schülervorstellungen zur Kreisbewegung oder zum dritten Newtonschen Gesetz wurde bspw. nicht geschaut, da diese in dem aufgezeichneten Versuch nicht zu erwarten waren. Nicht alle Schülervorstellungen traten auch auf, was an den teilweise fehlenden Ankerbeispielen sichtbar wird.

Code	Definition	Beschreibung/Kodierregeln	Ankerbeispiele
1	Kraft ist ein Sammelbegriff	Schüleraussagen, die Kraft als Cluster- oder Sammelbegriff auffassen, also nicht oder nicht ausreichend zwischen Kraft, Energie, Impuls etc. unterscheiden.	2: „Dann Kräfte in y-Richtung, (..) vielleicht Geschwindigkeit? Also wie beim Wagen.“
2	Zur Aufrechterhaltung einer Bewegung bedarf es einer Kraft in Bewegungsrichtung.	In diese Kategorie gehören alle Aussagen, die darauf hindeuten, dass die Schüler*innen davon ausgehen, dass immer eine Kraft in Bewegungsrichtung wirken muss, damit sich ein Objekt bewegen kann. Implizit wird dann zusätzlich oft von einer Proportionalität zwischen Kraft und Geschwindigkeit ausgegangen.	2: „Und in die Richtung muss doch auch was wirken.“ 1: „Das ist die Kraft, die am Anfang wirkt. Also die die von Anfang an hat.“
3	Ein bewegter Körper hat Kraft.	Aussagen, die damit im Zusammenhang stehen, dass Körper Kraft besitzen oder innehaben. Diese Kraft kann dann z. B. „verbraucht“ werden.	1: „Sie fliegt mit ner eigentlich konstanten Geschwindigkeit, nur die Gewichtskraft zieht es halt runter bis der fast keine Kraft mehr hat von dem Katapult.“

4	Alle Körper kommen natürlicherweise zum Stillstand.	Aussagen, die daraufhin deuten, dass die Schüler*innen der Meinung sind, dass alle Körper natürlicherweise (also ohne Krafteinwirkung) langsamer werden und zum Stillstand kommen.	-
5	Die Bewegungsform prägt sich dem bewegendem Objekt ein.	Aussagen, die vermuten lassen, dass die Schüler*innen davon ausgehen, dass die Art der Bewegung im Objekt gespeichert ist und den weiteren Verlauf der Bewegung beeinflusst.	-
6	Ein Körper kann sich nur mit konstanter Geschwindigkeit bewegen, wenn keine Kräfte wirken.	In diese Kategorie fallen alle Aussagen, die darauf hindeuten, dass die Schüler*innen denken, dass eine konstante Geschwindigkeit ausschließlich durch Kräftefreiheit (und nicht durch ein Kräftegleichgewicht) erreicht werden kann.	-
7	Es gibt unterschiedliche Wertigkeiten von Kräften (z. B. aktiv/passiv).	In diese Kategorie fallen alle Schülervorstellungen, die damit zu tun haben, dass es unterschiedliche Wertigkeiten von Kräften gibt. Dies kann eine Unterteilung in aktive Kräfte oder Körper oder auch Haupt- und Nebenkkräfte oder ähnliches sein.	1: „Nein das sind ja die Gegenkräfte, die wirken.“ 2: „Ja stimmt, wir brauchen die Hauptkräfte.“
8	Eine Kraft, die auf einen Körper wirkt, braucht eine Zeit, bis sie sich durchsetzen kann.	Alle Aussagen, die darauf hindeuten, dass Kräfte erst nach einiger Zeit einen Effekt haben. Beispielsweise, dass die Erdanziehungskraft sich erst nach einiger Zeit gegen die „Bewegungskraft“ durchsetzt.	1: „Also das heißt der Stoß Dingsda, ist stärker und ab dort wirkt dann wieder die Anziehungskraft, also gleich null. Ab diesem Punkt und ab dort überwiegt dann wieder die Anziehungskraft würde ich sagen.“

9	Undifferenziertheit zwischen Geschwindigkeit und Beschleunigung.	Während Geschwindigkeit eine für Schüler*innen häufig noch greifbare Größe ist, ist das Konzeptverständnis von Beschleunigung in der Regel deutlich schlechter. Das führt zu einer fehlenden Differenzierung der beiden Größen. In diese Kategorie gehören Aussagen, die Geschwindigkeit und Beschleunigung als ähnliche oder gleiche Größen beschreiben.	1: „Geschwindigkeit... Ist konstant bei 9,81.“

Dimension 4: Argumentationsrichtung

Vorgehen: DeduktivArt der Kodierung: Kodierung im Transkript

Code	Defini- tion	Beschreibung/Ko- dierregeln	Ankerbeispiele	Anmerkungen
1	Argu- menta- tion auf höherer Ebene	Bei der Beantwor- tung von Fragen bzw. der Beschrei- bung von gewissen Größen, wird auf eine höhere Ebene zur Begründung zurückgegriffen.	2: „So das passt jetzt. Also, wie verändert sich die Beschleunigung in x-Richtung? Also in x- Richtung ist die unge- fähr Null würde ich sa- gen einfach.“ 3: „Ja.“ 2: „Weil da halt...“ 3: „Ja.“ 2: „Keine Kraft wirkt.“	Hier liegt die Ein- teilung zugrunde, dass sich die Kräfte auf der höchsten Ebene befinden. Darauf folgen Be- schleunigung, Ge- schwindigkeit und der Ort.
2	Argu- menta- tion auf niedri- gerer Ebene	Bei der Beantwor- tung von Fragen bzw. der Beschrei- bung von gewissen Größen, wird auf eine niedrigere Ebene zur Begrün- dung zurückgegrif- fen.	1: „Die Summe aller Kräfte ist Null, weil Be- schleunigung Null ist.“	Hier liegt die Ein- teilung zugrunde, dass sich die Kräfte auf der höchsten Ebene befinden. Darauf folgen Be- schleunigung, Ge- schwindigkeit und der Ort.

3	Argumentation auf gleicher Ebene	Bei der Beantwortung von Fragen bzw. der Beschreibung von gewissen Größen wird mit derselben Größe argumentiert.	2: „Die Summe aller Kräfte... (.)“ 1: „Ändert sich, weil die Antriebskraft ist ja erst stärker und dann ist sie schwächer. Weil die ist ja nicht konstant. Weißt du, weil die ist ja... erst am Anfang ist sie stärker als die Erdanziehungskraft und irgendwann nimmt die halt ab und dann wird die Erdanziehungskraft halt... zieht den Ball halt wieder runter oder die Kugel. Also es ändert sich.“	Hier liegt die Einteilung zugrunde, dass sich die Kräfte auf der höchsten Ebene befinden. Darauf folgen Beschleunigung, Geschwindigkeit und der Ort.
4	Antwort ohne Begründung	Bei der Beantwortung von Fragen bzw. der Beschreibung von gewissen Größen wird keine Begründung geliefert.	2: „Summe aller Kräfte ist 0.“	Hier liegt die Einteilung zugrunde, dass sich die Kräfte auf der höchsten Ebene befinden. Darauf folgen Beschleunigung, Geschwindigkeit und der Ort.

Dimension 5: Dynamik

Vorgehen: DeduktivArt der Kodierung: Kodierung im Transkript

Code	Defini- tion	Beschreibung/Ko- dierregeln	Ankerbei- spiele	Anmerkungen
1	Kraft → Bewe- gung	Bei der Begründung des Verhaltens der kinematischen Größen wird mit Kräften argumentiert.	2: „Also Beschleunigung und Summe aller Kräfte muss ja immer gleich sein.“ 1: „Letztes ändert sich. Die Summe aller Kräfte.“	Ähnlich wie in der vorigen Dimension wird hier nach den Ebenen der Argumentation geschaut. Der Fokus ist hier aber ausschließlich auf den Argumentationen zwischen Kräften und Bewegung, um die Grundgleichung der Mechanik zu fokussieren. In diesem Fall geht es um Argumentationen in eine Richtung der Gleichung.
2	Bewe- gung → Kraft	Bei der Begründung des Wirkens der Kräfte wird mit dem Verhalten der kinematischen Größen argumentiert.	1: „Wenn die Geschwindigkeit konstant ist, sind die Summe aller Kräfte null, oder?“	In dieser Kategorie geht es um Argumentationen in die andere Richtung der Gleichung.

Bisher erschienene Bände der Reihe „*Studien zum Physik- und Chemielernen*“

ISSN 1614-8967 (vormals *Studien zum Physiklernen* ISSN 1435-5280)

- 1 Helmut Fischler, Jochen Peuckert (Hrsg.): Concept Mapping in fachdidaktischen Forschungsprojekten der Physik und Chemie
ISBN 978-3-89722-256-4 40.50 EUR
- 2 Anja Schoster: Bedeutungsentwicklungsprozesse beim Lösen algorithmischer Physikaufgaben. *Eine Fallstudie zu Lernprozessen von Schülern im Physiknachhilfeunterricht während der Bearbeitung algorithmischer Physikaufgaben*
ISBN 978-3-89722-045-4 40.50 EUR
- 3 Claudia von Aufschnaiter: Bedeutungsentwicklungen, Interaktionen und situatives Erleben beim Bearbeiten physikalischer Aufgaben
ISBN 978-3-89722-143-7 40.50 EUR
- 4 Susanne Haerberlen: Lernprozesse im Unterricht mit Wasserstromkreisen. *Eine Fallstudie in der Sekundarstufe I*
ISBN 978-3-89722-172-7 40.50 EUR
- 5 Kerstin Haller: Über den Zusammenhang von Handlungen und Zielen. *Eine empirische Untersuchung zu Lernprozessen im physikalischen Praktikum*
ISBN 978-3-89722-242-7 40.50 EUR
- 6 Michaela Horstendahl: Motivationale Orientierungen im Physikunterricht
ISBN 978-3-89722-227-4 50.00 EUR
- 7 Stefan Deylitz: Lernergebnisse in der Quanten-Atomphysik. *Evaluation des Bremer Unterrichtskonzepts*
ISBN 978-3-89722-291-5 40.50 EUR
- 8 Lorenz Hucke: Handlungsregulation und Wissenserwerb in traditionellen und computergestützten Experimenten des physikalischen Praktikums
ISBN 978-3-89722-316-5 50.00 EUR
- 9 Heike Theyßen: Ein Physikpraktikum für Studierende der Medizin. *Darstellung der Entwicklung und Evaluation eines adressatenspezifischen Praktikums nach dem Modell der Didaktischen Rekonstruktion*
ISBN 978-3-89722-334-9 40.50 EUR
- 10 Annette Schick: Der Einfluß von Interesse und anderen selbstbezogenen Kognitionen auf Handlungen im Physikunterricht. *Fallstudien zu Interessenhandlungen im Physikunterricht*
ISBN 978-3-89722-380-6 40.50 EUR
- 11 Roland Berger: Moderne bildgebende Verfahren der medizinischen Diagnostik. *Ein Weg zu interessanterem Physikunterricht*
ISBN 978-3-89722-445-2 40.50 EUR

- 12 Johannes Werner: Vom Licht zum Atom. *Ein Unterrichtskonzept zur Quantenphysik unter Nutzung des Zeigermodells*
ISBN 978-3-89722-471-1 40.50 EUR
- 13 Florian Sander: Verbindung von Theorie und Experiment im physikalischen Praktikum. *Eine empirische Untersuchung zum handlungsbezogenen Vorverständnis und dem Einsatz grafikorientierter Modellbildung im Praktikum*
ISBN 978-3-89722-482-7 40.50 EUR
- 14 Jörn Gerdes: Der Begriff der physikalischen Kompetenz. *Zur Validierung eines Konstruktes*
ISBN 978-3-89722-510-7 40.50 EUR
- 15 Malte Meyer-Arndt: Interaktionen im Physikpraktikum zwischen Studierenden und Betreuern. *Feldstudie zu Bedeutungsentwicklungsprozessen im physikalischen Praktikum*
ISBN 978-3-89722-541-1 40.50 EUR
- 16 Dietmar Höttecke: Die Natur der Naturwissenschaften historisch verstehen. *Fachdidaktische und wissenschaftshistorische Untersuchungen*
ISBN 978-3-89722-607-4 40.50 EUR
- 17 Gil Gabriel Mavanga: Entwicklung und Evaluation eines experimentell- und phänomenorientierten Optikcurriculums. *Untersuchung zu Schülervorstellungen in der Sekundarstufe I in Mosambik und Deutschland*
ISBN 978-3-89722-721-7 40.50 EUR
- 18 Meike Ute Zastrow: Interaktive Experimentieranleitungen. *Entwicklung und Evaluation eines Konzeptes zur Vorbereitung auf das Experimentieren mit Messgeräten im Physikalischen Praktikum*
ISBN 978-3-89722-802-3 40.50 EUR
- 19 Gunnar Friege: Wissen und Problemlösen. *Eine empirische Untersuchung des wissenszentrierten Problemlösens im Gebiet der Elektrizitätslehre auf der Grundlage des Experten-Novizen-Vergleichs*
ISBN 978-3-89722-809-2 40.50 EUR
- 20 Erich Starauschek: Physikunterricht nach dem Karlsruher Physikkurs. *Ergebnisse einer Evaluationsstudie*
ISBN 978-3-89722-823-8 40.50 EUR
- 21 Roland Paatz: Charakteristika analogiebasierten Denkens. *Vergleich von Lernprozessen in Basis- und Zielbereich*
ISBN 978-3-89722-944-0 40.50 EUR
- 22 Silke Mikelskis-Seifert: Die Entwicklung von Metakzepten zur Teilchenvorstellung bei Schülern. *Untersuchung eines Unterrichts über Modelle mithilfe eines Systems multipler Repräsentationsebenen*
ISBN 978-3-8325-0013-9 40.50 EUR
- 23 Brunhild Landwehr: Distanzen von Lehrkräften und Studierenden des Sachunterrichts zur Physik. *Eine qualitativ-empirische Studie zu den Ursachen*
ISBN 978-3-8325-0044-3 40.50 EUR

- 24 Lydia Murmann: Physiklernen zu Licht, Schatten und Sehen. *Eine phänomenografische Untersuchung in der Primarstufe*
ISBN 978-3-8325-0060-3 40.50 EUR
- 25 Thorsten Bell: Strukturprinzipien der Selbstregulation. *Komplexe Systeme, Elementarisierungen und Lernprozessstudien für den Unterricht der Sekundarstufe II*
ISBN 978-3-8325-0134-1 40.50 EUR
- 26 Rainer Müller: Quantenphysik in der Schule
ISBN 978-3-8325-0186-0 40.50 EUR
- 27 Jutta Roth: Bedeutungsentwicklungsprozesse von Physikerinnen und Physikern in den Dimensionen Komplexität, Zeit und Inhalt
ISBN 978-3-8325-0183-9 40.50 EUR
- 28 Andreas Saniter: Spezifika der Verhaltensmuster fortgeschrittener Studierender der Physik
ISBN 978-3-8325-0292-8 40.50 EUR
- 29 Thomas Weber: Kumulatives Lernen im Physikunterricht. *Eine vergleichende Untersuchung in Unterrichtsgängen zur geometrischen Optik*
ISBN 978-3-8325-0316-1 40.50 EUR
- 30 Markus Rehm: Über die Chancen und Grenzen moralischer Erziehung im naturwissenschaftlichen Unterricht
ISBN 978-3-8325-0368-0 40.50 EUR
- 31 Marion Budde: Lernwirkungen in der Quanten-Atom-Physik. *Fallstudien über Resonanzen zwischen Lernangeboten und SchülerInnen-Vorstellungen*
ISBN 978-3-8325-0483-0 40.50 EUR
- 32 Thomas Reyer: Oberflächenmerkmale und Tiefenstrukturen im Unterricht. *Exemplarische Analysen im Physikunterricht der gymnasialen Sekundarstufe*
ISBN 978-3-8325-0488-5 40.50 EUR
- 33 Christoph Thomas Müller: Subjektive Theorien und handlungsleitende Kognitionen von Lehrern als Determinanten schulischer Lehr-Lern-Prozesse im Physikunterricht
ISBN 978-3-8325-0543-1 40.50 EUR
- 34 Gabriela Jonas-Ahrend: Physiklehrervorstellungen zum Experiment im Physikunterricht
ISBN 978-3-8325-0576-9 40.50 EUR
- 35 Dimitrios Stavrou: Das Zusammenspiel von Zufall und Gesetzmäßigkeiten in der nicht-linearen Dynamik. *Didaktische Analyse und Lernprozesse*
ISBN 978-3-8325-0609-4 40.50 EUR
- 36 Katrin Engeln: Schülerlabors: authentische, aktivierende Lernumgebungen als Möglichkeit, Interesse an Naturwissenschaften und Technik zu wecken
ISBN 978-3-8325-0689-6 40.50 EUR
- 37 Susann Hartmann: Erklärungsvielfalt
ISBN 978-3-8325-0730-5 40.50 EUR

- 38 Knut Neumann: Didaktische Rekonstruktion eines physikalischen Praktikums für Physiker
ISBN 978-3-8325-0762-6 40.50 EUR
- 39 Michael Späth: Kontextbedingungen für Physikunterricht an der Hauptschule. *Möglichkeiten und Ansatzpunkte für einen fachübergreifenden, handlungsorientierten und berufsorientierten Unterricht*
ISBN 978-3-8325-0827-2 40.50 EUR
- 40 Jörg Hirsch: Interesse, Handlungen und situatives Erleben von Schülerinnen und Schülern beim Bearbeiten physikalischer Aufgaben
ISBN 978-3-8325-0875-3 40.50 EUR
- 41 Monika Hüther: Evaluation einer hypermedialen Lernumgebung zum Thema Gasgesetze. *Eine Studie im Rahmen des Physikpraktikums für Studierende der Medizin*
ISBN 978-3-8325-0911-8 40.50 EUR
- 42 Maike Tesch: Das Experiment im Physikunterricht. *Didaktische Konzepte und Ergebnisse einer Videostudie*
ISBN 978-3-8325-0975-0 40.50 EUR
- 43 Nina Nicolai: Skriptgeleitete Eltern-Kind-Interaktion bei Chemiehausaufgaben. *Eine Evaluationsstudie im Themenbereich Säure-Base*
ISBN 978-3-8325-1013-8 40.50 EUR
- 44 Antje Leisner: Entwicklung von Modellkompetenz im Physikunterricht
ISBN 978-3-8325-1020-6 40.50 EUR
- 45 Stefan Rumann: Evaluation einer Interventionsstudie zur Säure-Base-Thematik
ISBN 978-3-8325-1027-5 40.50 EUR
- 46 Thomas Wilhelm: Konzeption und Evaluation eines Kinematik/Dynamik-Lehrgangs zur Veränderung von Schülervorstellungen mit Hilfe dynamisch ikonischer Repräsentationen und graphischer Modellbildung – mit CD-ROM
ISBN 978-3-8325-1046-6 45.50 EUR
- 47 Andrea Maier-Richter: Computerunterstütztes Lernen mit Lösungsbeispielen in der Chemie. *Eine Evaluationsstudie im Themenbereich Löslichkeit*
ISBN 978-3-8325-1046-6 40.50 EUR
- 48 Jochen Peuckert: Stabilität und Ausprägung kognitiver Strukturen zum Atombegriff
ISBN 978-3-8325-1104-3 40.50 EUR
- 49 Maik Walpuski: Optimierung von experimenteller Kleingruppenarbeit durch Strukturierungshilfen und Feedback
ISBN 978-3-8325-1184-5 40.50 EUR
- 50 Helmut Fischler, Christiane S. Reiners (Hrsg.): Die Teilchenstruktur der Materie im Physik- und Chemieunterricht
ISBN 978-3-8325-1225-5 34.90 EUR
- 51 Claudia Eysel: Interdisziplinäres Lehren und Lernen in der Lehrerbildung. *Eine empirische Studie zum Kompetenzerwerb in einer komplexen Lernumgebung*
ISBN 978-3-8325-1238-5 40.50 EUR

- 52 Johannes Günther: Lehrerfortbildung über die Natur der Naturwissenschaften. *Studien über das Wissenschaftsverständnis von Grundschullehrkräften*
ISBN 978-3-8325-1287-3 40.50 EUR
- 53 Christoph Neugebauer: Lernen mit Simulationen und der Einfluss auf das Problemlösen in der Physik
ISBN 978-3-8325-1300-9 40.50 EUR
- 54 Andreas Schnirch: Gendergerechte Interessen- und Motivationsförderung im Kontext naturwissenschaftlicher Grundbildung. *Konzeption, Entwicklung und Evaluation einer multimedial unterstützten Lernumgebung*
ISBN 978-3-8325-1334-4 40.50 EUR
- 55 Hilde Köster: Freies Explorieren und Experimentieren. *Eine Untersuchung zur selbstbestimmten Gewinnung von Erfahrungen mit physikalischen Phänomenen im Sachunterricht*
ISBN 978-3-8325-1348-1 40.50 EUR
- 56 Eva Heran-Dörr: Entwicklung und Evaluation einer Lehrerfortbildung zur Förderung der physikdidaktischen Kompetenz von Sachunterrichtslehrkräften
ISBN 978-3-8325-1377-1 40.50 EUR
- 57 Agnes Szabone Varnai: Unterstützung des Problemlösens in Physik durch den Einsatz von Simulationen und die Vorgabe eines strukturierten Kooperationsformats
ISBN 978-3-8325-1403-7 40.50 EUR
- 58 Johannes Rethfeld: Aufgabenbasierte Lernprozesse in selbstorganisationsoffenem Unterricht der Sekundarstufe I zum Themengebiet ELEKTROSTATIK. *Eine Feldstudie in vier 10. Klassen zu einer kartenbasierten Lernumgebung mit Aufgaben aus der Elektrostatik*
ISBN 978-3-8325-1416-7 40.50 EUR
- 59 Christian Henke: Experimentell-naturwissenschaftliche Arbeitsweisen in der Oberstufe. *Untersuchung am Beispiel des HIGHSEA-Projekts in Bremerhaven*
ISBN 978-3-8325-1515-7 40.50 EUR
- 60 Lutz Kasper: Diskursiv-narrative Elemente für den Physikunterricht. *Entwicklung und Evaluation einer multimedialen Lernumgebung zum Erdmagnetismus*
ISBN 978-3-8325-1537-9 40.50 EUR
- 61 Thorid Rabe: Textgestaltung und Aufforderung zu Selbsterklärungen beim Physiklernen mit Multimedia
ISBN 978-3-8325-1539-3 40.50 EUR
- 62 Ina Glemnitz: Vertikale Vernetzung im Chemieunterricht. *Ein Vergleich von traditionellem Unterricht mit Unterricht nach Chemie im Kontext*
ISBN 978-3-8325-1628-4 40.50 EUR
- 63 Erik Einhaus: Schülerkompetenzen im Bereich Wärmelehre. *Entwicklung eines Testinstruments zur Überprüfung und Weiterentwicklung eines normativen Modells fachbezogener Kompetenzen*
ISBN 978-3-8325-1630-7 40.50 EUR

- 64 Jasmin Neuroth: Concept Mapping als Lernstrategie. *Eine Interventionsstudie zum Chemielernen aus Texten*
ISBN 978-3-8325-1659-8 40.50 EUR
- 65 Hans Gerd Hegeler-Burkhart: Zur Kommunikation von Hauptschülerinnen und Hauptschülern in einem handlungsorientierten und fächerübergreifenden Unterricht mit physikalischen und technischen Inhalten
ISBN 978-3-8325-1667-3 40.50 EUR
- 66 Karsten Rincke: Sprachentwicklung und Fachlernen im Mechanikunterricht. *Sprache und Kommunikation bei der Einführung in den Kraftbegriff*
ISBN 978-3-8325-1699-4 40.50 EUR
- 67 Nina Strehle: Das Ion im Chemieunterricht. *Alternative Schülervorstellungen und curriculare Konsequenzen*
ISBN 978-3-8325-1710-6 40.50 EUR
- 68 Martin Hopf: Problemorientierte Schülerexperimente
ISBN 978-3-8325-1711-3 40.50 EUR
- 69 Anne Beerenwinkel: Fostering conceptual change in chemistry classes using expository texts
ISBN 978-3-8325-1721-2 40.50 EUR
- 70 Roland Berger: Das Gruppenpuzzle im Physikunterricht der Sekundarstufe II. *Eine empirische Untersuchung auf der Grundlage der Selbstbestimmungstheorie der Motivation*
ISBN 978-3-8325-1732-8 40.50 EUR
- 71 Giuseppe Colicchia: Physikunterricht im Kontext von Medizin und Biologie. *Entwicklung und Erprobung von Unterrichtseinheiten*
ISBN 978-3-8325-1746-5 40.50 EUR
- 72 Sandra Winheller: Geschlechtsspezifische Auswirkungen der Lehrer-Schüler-Interaktion im Chemieanfangsunterricht
ISBN 978-3-8325-1757-1 40.50 EUR
- 73 Isabel Wahser: Training von naturwissenschaftlichen Arbeitsweisen zur Unterstützung experimenteller Kleingruppenarbeit im Fach Chemie
ISBN 978-3-8325-1815-8 40.50 EUR
- 74 Claus Brell: Lernmedien und Lernerfolg - reale und virtuelle Materialien im Physikunterricht. *Empirische Untersuchungen in achten Klassen an Gymnasien (Laborstudie) zum Computereinsatz mit Simulation und IBE*
ISBN 978-3-8325-1829-5 40.50 EUR
- 75 Rainer Wackermann: Überprüfung der Wirksamkeit eines Basismodell-Trainings für Physiklehrer
ISBN 978-3-8325-1882-0 40.50 EUR
- 76 Oliver Tepner: Effektivität von Aufgaben im Chemieunterricht der Sekundarstufe I
ISBN 978-3-8325-1919-3 40.50 EUR

- 77 Claudia Geyer: Museums- und Science-Center-Besuche im naturwissenschaftlichen Unterricht aus einer motivationalen Perspektive. *Die Sicht von Lehrkräften und Schülerinnen und Schülern*
ISBN 978-3-8325-1922-3 40.50 EUR
- 78 Tobias Leonhard: Professionalisierung in der Lehrerbildung. *Eine explorative Studie zur Entwicklung professioneller Kompetenzen in der Lehrererstausbildung*
ISBN 978-3-8325-1924-7 40.50 EUR
- 79 Alexander Kauertz: Schwierigkeitserzeugende Merkmale physikalischer Leistungstestaufgaben
ISBN 978-3-8325-1925-4 40.50 EUR
- 80 Regina Hübinger: Schüler auf Weltreise. *Entwicklung und Evaluation von Lehr-/Lernmaterialien zur Förderung experimentell-naturwissenschaftlicher Kompetenzen für die Jahrgangsstufen 5 und 6*
ISBN 978-3-8325-1932-2 40.50 EUR
- 81 Christine Waltner: Physik lernen im Deutschen Museum
ISBN 978-3-8325-1933-9 40.50 EUR
- 82 Torsten Fischer: Handlungsmuster von Physiklehrkräften beim Einsatz neuer Medien. *Fallstudien zur Unterrichtspraxis*
ISBN 978-3-8325-1948-3 42.00 EUR
- 83 Corinna Kieren: Chemiehausaufgaben in der Sekundarstufe I des Gymnasiums. *Fragebogenerhebung zur gegenwärtigen Praxis und Entwicklung eines optimierten Hausaufgabendesigns im Themenbereich Säure-Base*
978-3-8325-1975-9 37.00 EUR
- 84 Marco Thiele: Modelle der Thermohalinen Zirkulation im Unterricht. *Eine empirische Studie zur Förderung des Modellverständnisses*
ISBN 978-3-8325-1982-7 40.50 EUR
- 85 Bernd Zinn: Physik lernen, um Physik zu lehren. *Eine Möglichkeit für interessanteren Physikunterricht*
ISBN 978-3-8325-1995-7 39.50 EUR
- 86 Esther Klaes: Außerschulische Lernorte im naturwissenschaftlichen Unterricht. *Die Perspektive der Lehrkraft*
ISBN 978-3-8325-2006-9 43.00 EUR
- 87 Marita Schmidt: Kompetenzmodellierung und -diagnostik im Themengebiet Energie der Sekundarstufe I. *Entwicklung und Erprobung eines Testinventars*
ISBN 978-3-8325-2024-3 37.00 EUR
- 88 Gudrun Franke-Braun: Aufgaben mit gestuften Lernhilfen. *Ein Aufgabenformat zur Förderung der sachbezogenen Kommunikation und Lernleistung für den naturwissenschaftlichen Unterricht*
ISBN 978-3-8325-2026-7 38.00 EUR
- 89 Silke Klos: Kompetenzförderung im naturwissenschaftlichen Anfangsunterricht. *Der Einfluss eines integrierten Unterrichtskonzepts*
ISBN 978-3-8325-2133-2 37.00 EUR

- 90 Ulrike Elisabeth Burkard: Quantenphysik in der Schule. *Bestandsaufnahme, Perspektiven und Weiterentwicklungsmöglichkeiten durch die Implementation eines Medienservers*
ISBN 978-3-8325-2215-5 43.00 EUR
- 91 Ulrike Gromadecki: Argumente in physikalischen Kontexten. *Welche Geltungsgründe halten Physikanfänger für überzeugend?*
ISBN 978-3-8325-2250-6 41.50 EUR
- 92 Jürgen Bruns: Auf dem Weg zur Förderung naturwissenschaftsspezifischer Vorstellungen von zukünftigen Chemie-Lehrenden
ISBN 978-3-8325-2257-5 43.50 EUR
- 93 Cornelius Marsch: Räumliche Atomvorstellung. *Entwicklung und Erprobung eines Unterrichtskonzeptes mit Hilfe des Computers*
ISBN 978-3-8325-2293-3 82.50 EUR
- 94 Maja Brückmann: Sachstrukturen im Physikunterricht. *Ergebnisse einer Videostudie*
ISBN 978-3-8325-2272-8 39.50 EUR
- 95 Sabine Fechner: Effects of Context-oriented Learning on Student Interest and Achievement in Chemistry Education
ISBN 978-3-8325-2343-5 36.50 EUR
- 96 Clemens Nagel: eLearning im Physikalischen Anfängerpraktikum
ISBN 978-3-8325-2355-8 39.50 EUR
- 97 Josef Riese: Professionelles Wissen und professionelle Handlungskompetenz von (angehenden) Physiklehrkräften
ISBN 978-3-8325-2376-3 39.00 EUR
- 98 Sascha Bernholt: Kompetenzmodellierung in der Chemie. *Theoretische und empirische Reflexion am Beispiel des Modells hierarchischer Komplexität*
ISBN 978-3-8325-2447-0 40.00 EUR
- 99 Holger Christoph Stawitz: Auswirkung unterschiedlicher Aufgabenprofile auf die Schülerleistung. *Vergleich von Naturwissenschafts- und Problemlöseaufgaben der PISA 2003-Studie*
ISBN 978-3-8325-2451-7 37.50 EUR
- 100 Hans Ernst Fischer, Elke Sumfleth (Hrsg.): nwu-essen – 10 Jahre Essener Forschung zum naturwissenschaftlichen Unterricht
ISBN 978-3-8325-3331-1 40.00 EUR
- 101 Hendrik Härtig: Sachstrukturen von Physikschulbüchern als Grundlage zur Bestimmung der Inhaltsvalidität eines Tests
ISBN 978-3-8325-2512-5 34.00 EUR
- 102 Thomas Grüß-Niehaus: Zum Verständnis des Löslichkeitskonzeptes im Chemieunterricht. *Der Effekt von Methoden progressiver und kollaborativer Reflexion*
ISBN 978-3-8325-2537-8 40.50 EUR

- 103 Patrick Bronner: Quantenoptische Experimente als Grundlage eines Curriculums zur Quantenphysik des Photons
ISBN 978-3-8325-2540-8 36.00 EUR
- 104 Adrian Voßkühler: Blickbewegungsmessung an Versuchsaufbauten. *Studien zur Wahrnehmung, Verarbeitung und Usability von physikbezogenen Experimenten am Bildschirm und in der Realität*
ISBN 978-3-8325-2548-4 47.50 EUR
- 105 Verena Tobias: Newton'sche Mechanik im Anfangsunterricht. *Die Wirksamkeit einer Einführung über die zweidimensionale Dynamik auf das Lehren und Lernen*
ISBN 978-3-8325-2558-3 54.00 EUR
- 106 Christian Rogge: Entwicklung physikalischer Konzepte in aufgabenbasierten Lernumgebungen
ISBN 978-3-8325-2574-3 45.00 EUR
- 107 Mathias Ropohl: Modellierung von Schülerkompetenzen im Basiskonzept Chemische Reaktion. *Entwicklung und Analyse von Testaufgaben*
ISBN 978-3-8325-2609-2 36.50 EUR
- 108 Christoph Kulgemeyer: Physikalische Kommunikationskompetenz. *Modellierung und Diagnostik*
ISBN 978-3-8325-2674-0 44.50 EUR
- 109 Jennifer Olszewski: The Impact of Physics Teachers' Pedagogical Content Knowledge on Teacher Actions and Student Outcomes
ISBN 978-3-8325-2680-1 33.50 EUR
- 110 Annika Ohle: Primary School Teachers' Content Knowledge in Physics and its Impact on Teaching and Students' Achievement
ISBN 978-3-8325-2684-9 36.50 EUR
- 111 Susanne Mannel: Assessing scientific inquiry. *Development and evaluation of a test for the low-performing stage*
ISBN 978-3-8325-2761-7 40.00 EUR
- 112 Michael Plomer: Physik physiologisch passend praktiziert. *Eine Studie zur Lernwirksamkeit von traditionellen und adressatenspezifischen Physikpraktika für die Physiologie*
ISBN 978-3-8325-2804-1 34.50 EUR
- 113 Alexandra Schulz: Experimentierspezifische Qualitätsmerkmale im Chemieunterricht. *Eine Videostudie*
ISBN 978-3-8325-2817-1 40.00 EUR
- 114 Franz Boczianowski: Eine empirische Untersuchung zu Vektoren im Physikunterricht der Mittelstufe
ISBN 978-3-8325-2843-0 39.50 EUR
- 115 Maria Ploog: Internetbasiertes Lernen durch Textproduktion im Fach Physik
ISBN 978-3-8325-2853-9 39.50 EUR

- 116 Anja Dhein: Lernen in Explorier- und Experimentiersituationen. *Eine explorative Studie zu Bedeutungsentwicklungsprozessen bei Kindern im Alter zwischen 4 und 6 Jahren*
ISBN 978-3-8325-2859-1 45.50 EUR
- 117 Irene Neumann: Beyond Physics Content Knowledge. *Modeling Competence Regarding Nature of Scientific Inquiry and Nature of Scientific Knowledge*
ISBN 978-3-8325-2880-5 37.00 EUR
- 118 Markus Emden: Prozessorientierte Leistungsmessung des naturwissenschaftlich-experimentellen Arbeitens. *Eine vergleichende Studie zu Diagnoseinstrumenten zu Beginn der Sekundarstufe I*
ISBN 978-3-8325-2867-6 38.00 EUR
- 119 Birgit Hofmann: Analyse von Blickbewegungen von Schülern beim Lesen von physikbezogenen Texten mit Bildern. *Eye Tracking als Methodenwerkzeug in der physikdidaktischen Forschung*
ISBN 978-3-8325-2925-3 59.00 EUR
- 120 Rebecca Knobloch: Analyse der fachinhaltlichen Qualität von Schüleräußerungen und deren Einfluss auf den Lernerfolg. *Eine Videostudie zu kooperativer Kleingruppenarbeit*
ISBN 978-3-8325-3006-8 36.50 EUR
- 121 Julia Hostenbach: Entwicklung und Prüfung eines Modells zur Beschreibung der Bewertungskompetenz im Chemieunterricht
ISBN 978-3-8325-3013-6 38.00 EUR
- 122 Anna Windt: Naturwissenschaftliches Experimentieren im Elementarbereich. *Evaluation verschiedener Lernsituationen*
ISBN 978-3-8325-3020-4 43.50 EUR
- 123 Eva Kölbach: Kontexteinflüsse beim Lernen mit Lösungsbeispielen
ISBN 978-3-8325-3025-9 38.50 EUR
- 124 Anna Lau: Passung und vertikale Vernetzung im Chemie- und Physikunterricht
ISBN 978-3-8325-3021-1 36.00 EUR
- 125 Jan Lamprecht: Ausbildungswege und Komponenten professioneller Handlungskompetenz. *Vergleich von Quereinsteigern mit Lehramtsabsolventen für Gymnasien im Fach Physik*
ISBN 978-3-8325-3035-8 38.50 EUR
- 126 Ulrike Böhm: Förderung von Verstehensprozessen unter Einsatz von Modellen
ISBN 978-3-8325-3042-6 41.00 EUR
- 127 Sabrina Dollny: Entwicklung und Evaluation eines Testinstruments zur Erfassung des fachspezifischen Professionswissens von Chemielehrkräften
ISBN 978-3-8325-3046-4 37.00 EUR
- 128 Monika Zimmermann: Naturwissenschaftliche Bildung im Kindergarten. *Eine integrative Längsschnittstudie zur Kompetenzentwicklung von Erzieherinnen*
ISBN 978-3-8325-3053-2 54.00 EUR

- 129 Ulf Saballus: Über das Schlussfolgern von Schülerinnen und Schülern zu öffentlichen Kontroversen mit naturwissenschaftlichem Hintergrund. *Eine Fallstudie*
ISBN 978-3-8325-3086-0 39.50 EUR
- 130 Olaf Krey: Zur Rolle der Mathematik in der Physik. *Wissenschaftstheoretische Aspekte und Vorstellungen Physiklernender*
ISBN 978-3-8325-3101-0 46.00 EUR
- 131 Angelika Wolf: Zusammenhänge zwischen der Eigenständigkeit im Physikunterricht, der Motivation, den Grundbedürfnissen und dem Lernerfolg von Schülern
ISBN 978-3-8325-3161-4 45.00 EUR
- 132 Johannes Börlin: Das Experiment als Lerngelegenheit. *Vom interkulturellen Vergleich des Physikunterrichts zu Merkmalen seiner Qualität*
ISBN 978-3-8325-3170-6 45.00 EUR
- 133 Olaf Uhden: Mathematisches Denken im Physikunterricht. *Theorieentwicklung und Problemanalyse*
ISBN 978-3-8325-3170-6 45.00 EUR
- 134 Christoph Gut: Modellierung und Messung experimenteller Kompetenz. *Analyse eines large-scale Experimentiertests*
ISBN 978-3-8325-3213-0 40.00 EUR
- 135 Antonio Rueda: Lernen mit ExploMultimedial in kolumbianischen Schulen. *Analyse von kurzzeitigen Lernprozessen und der Motivation beim länderübergreifenden Einsatz einer deutschen computergestützten multimedialen Lernumgebung für den naturwissenschaftlichen Unterricht*
ISBN 978-3-8325-3218-5 45.50 EUR
- 136 Krisztina Berger: Bilder, Animationen und Notizen. *Empirische Untersuchung zur Wirkung einfacher visueller Repräsentationen und Notizen auf den Wissenserwerb in der Optik*
ISBN 978-3-8325-3238-3 41.50 EUR
- 137 Antony Crossley: Untersuchung des Einflusses unterschiedlicher physikalischer Konzepte auf den Wissenserwerb in der Thermodynamik der Sekundarstufe I
ISBN 978-3-8325-3275-8 40.00 EUR
- 138 Tobias Viering: Entwicklung physikalischer Kompetenz in der Sekundarstufe I. *Validierung eines Kompetenzentwicklungsmodells für das Energiekonzept im Bereich Fachwissen*
ISBN 978-3-8325-3277-2 37.00 EUR
- 139 Nico Schreiber: Diagnostik experimenteller Kompetenz. *Validierung technologiegestützter Testverfahren im Rahmen eines Kompetenzstrukturmodells*
ISBN 978-3-8325-3284-0 39.00 EUR
- 140 Sarah Hundertmark: Einblicke in kollaborative Lernprozesse. *Eine Fallstudie zur reflektierenden Zusammenarbeit unterstützt durch die Methoden Concept Mapping und Lernbegleitbogen*
ISBN 978-3-8325-3251-2 43.00 EUR

- 141 Ronny Scherer: Analyse der Struktur, Messinvarianz und Ausprägung komplexer Problemlösekompetenz im Fach Chemie. *Eine Querschnittstudie in der Sekundarstufe I und am Übergang zur Sekundarstufe II*
ISBN 978-3-8325-3312-0 43.00 EUR
- 142 Patricia Heitmann: Bewertungskompetenz im Rahmen naturwissenschaftlicher Problemlöseprozesse. *Modellierung und Diagnose der Kompetenzen Bewertung und analytisches Problemlösen für das Fach Chemie*
ISBN 978-3-8325-3314-4 37.00 EUR
- 143 Jan Fleischhauer: Wissenschaftliches Argumentieren und Entwicklung von Konzepten beim Lernen von Physik
ISBN 978-3-8325-3325-0 35.00 EUR
- 144 Nermin Özcan: Zum Einfluss der Fachsprache auf die Leistung im Fach Chemie. *Eine Förderstudie zur Fachsprache im Chemieunterricht*
ISBN 978-3-8325-3328-1 36.50 EUR
- 145 Helena van Vorst: Kontextmerkmale und ihr Einfluss auf das Schülerinteresse im Fach Chemie
ISBN 978-3-8325-3321-2 38.50 EUR
- 146 Janine Cappell: Fachspezifische Diagnosekompetenz angehender Physiklehrkräfte in der ersten Ausbildungsphase
ISBN 978-3-8325-3356-4 38.50 EUR
- 147 Susanne Bley: Förderung von Transferprozessen im Chemieunterricht
ISBN 978-3-8325-3407-3 40.50 EUR
- 148 Cathrin Blaes: Die übungsgestützte Lehrerpräsentation im Chemieunterricht der Sekundarstufe I. *Evaluation der Effektivität*
ISBN 978-3-8325-3409-7 43.50 EUR
- 149 Julia Suckut: Die Wirksamkeit von piko-OWL als Lehrerfortbildung. Eine Evaluation zum Projekt *Physik im Kontext* in Fallstudien
ISBN 978-3-8325-3440-0 45.00 EUR
- 150 Alexandra Dorschu: Die Wirkung von Kontexten in Physikkompetenztestaufgaben
ISBN 978-3-8325-3446-2 37.00 EUR
- 151 Jochen Scheid: Multiple Repräsentationen, Verständnis physikalischer Experimente und kognitive Aktivierung: *Ein Beitrag zur Entwicklung der Aufgabenkultur*
ISBN 978-3-8325-3449-3 49.00 EUR
- 152 Tim Plasa: Die Wahrnehmung von Schülerlaboren und Schülerforschungszentren
ISBN 978-3-8325-3483-7 35.50 EUR
- 153 Felix Schoppmeier: Physikkompetenz in der gymnasialen Oberstufe. *Entwicklung und Validierung eines Kompetenzstrukturmodells für den Kompetenzbereich Umgang mit Fachwissen*
ISBN 978-3-8325-3502-5 36.00 EUR

- 154 Katharina Groß: Experimente alternativ dokumentieren. *Eine qualitative Studie zur Förderung der Diagnose- und Differenzierungskompetenz in der Chemielehrerbildung*
ISBN 978-3-8325-3508-7 43.50 EUR
- 155 Barbara Hank: Konzeptwandelprozesse im Anfangsunterricht Chemie. *Eine quasixperimentelle Längsschnittstudie*
ISBN 978-3-8325-3519-3 38.50 EUR
- 156 Katja Freyer: Zum Einfluss von Studieneingangsvoraussetzungen auf den Studienerfolg Erstsemesterstudierender im Fach Chemie
ISBN 978-3-8325-3544-5 38.00 EUR
- 157 Alexander Rachel: Auswirkungen instruktionaler Hilfen bei der Einführung des (Ferro-)Magnetismus. *Eine Vergleichsstudie in der Primar- und Sekundarstufe*
ISBN 978-3-8325-3548-3 43.50 EUR
- 158 Sebastian Ritter: Einfluss des Lerninhalts Nanogrößeneffekte auf Teilchen- und Teilchenmodellvorstellungen von Schülerinnen und Schülern
ISBN 978-3-8325-3558-2 36.00 EUR
- 159 Andrea Harbach: Problemorientierung und Vernetzung in kontextbasierten Lernaufgaben
ISBN 978-3-8325-3564-3 39.00 EUR
- 160 David Obst: Interaktive Tafeln im Physikunterricht. *Entwicklung und Evaluation einer Lehrerfortbildung*
ISBN 978-3-8325-3582-7 40.50 EUR
- 161 Sophie Kirschner: Modellierung und Analyse des Professionswissens von Physik Lehrkräften
ISBN 978-3-8325-3601-5 35.00 EUR
- 162 Katja Stief: Selbstregulationsprozesse und Hausaufgabenmotivation im Chemieunterricht
ISBN 978-3-8325-3631-2 34.00 EUR
- 163 Nicola Meschede: Professionelle Wahrnehmung der inhaltlichen Strukturierung im naturwissenschaftlichen Grundschulunterricht. *Theoretische Beschreibung und empirische Erfassung*
ISBN 978-3-8325-3668-8 37.00 EUR
- 164 Johannes Maximilian Barth: Experimentieren im Physikunterricht der gymnasialen Oberstufe. *Eine Rekonstruktion übergeordneter Einbettungsstrategien*
ISBN 978-3-8325-3681-7 39.00 EUR
- 165 Sandra Lein: Das Betriebspraktikum in der Lehrerbildung. *Eine Untersuchung zur Förderung der Wissenschafts- und Technikbildung im allgemeinbildenden Unterricht*
ISBN 978-3-8325-3698-5 40.00 EUR
- 166 Veranika Maiseyenka: Modellbasiertes Experimentieren im Unterricht. *Praxistauglichkeit und Lernwirkungen*
ISBN 978-3-8325-3708-1 38.00 EUR

- 167 Christoph Stolzenberger: Der Einfluss der didaktischen Lernumgebung auf das Erreichen geforderter Bildungsziele am Beispiel der W- und P-Seminare im Fach Physik
ISBN 978-3-8325-3708-1 38.00 EUR
- 168 Pia Altenburger: Mehrebenenregressionsanalysen zum Physiklernen im Sachunterricht der Primarstufe. *Ergebnisse einer Evaluationsstudie.*
ISBN 978-3-8325-3717-3 37.50 EUR
- 169 Nora Ferber: Entwicklung und Validierung eines Testinstruments zur Erfassung von Kompetenzentwicklung im Fach Chemie in der Sekundarstufe I
ISBN 978-3-8325-3727-2 39.50 EUR
- 170 Anita Stender: Unterrichtsplanung: Vom Wissen zum Handeln.
Theoretische Entwicklung und empirische Überprüfung des Transformationsmodells der Unterrichtsplanung
ISBN 978-3-8325-3750-0 41.50 EUR
- 171 Jenna Koenen: Entwicklung und Evaluation von experimentunterstützten Lösungsbeispielen zur Förderung naturwissenschaftlich-experimenteller Arbeitsweisen
ISBN 978-3-8325-3785-2 43.00 EUR
- 172 Teresa Henning: Empirische Untersuchung kontextorientierter Lernumgebungen in der Hochschuldidaktik. *Entwicklung und Evaluation kontextorientierter Aufgaben in der Studieneingangsphase für Fach- und Nebenfachstudierende der Physik*
ISBN 978-3-8325-3801-9 43.00 EUR
- 173 Alexander Pusch: Fachspezifische Instrumente zur Diagnose und individuellen Förderung von Lehramtsstudierenden der Physik
ISBN 978-3-8325-3829-3 38.00 EUR
- 174 Christoph Vogelsang: Validierung eines Instruments zur Erfassung der professionellen Handlungskompetenz von (angehenden) Physiklehrkräften. *Zusammenhangsanalysen zwischen Lehrerkompetenz und Lehrerperformanz*
ISBN 978-3-8325-3846-0 50.50 EUR
- 175 Ingo Brebeck: Selbstreguliertes Lernen in der Studieneingangsphase im Fach Chemie
ISBN 978-3-8325-3859-0 37.00 EUR
- 176 Axel Eghtessad: Merkmale und Strukturen von Professionalisierungsprozessen in der ersten und zweiten Phase der Chemielehrerbildung. *Eine empirisch-qualitative Studie mit niedersächsischen Fachleiter_innen der Sekundarstufenlehrkräfte*
ISBN 978-3-8325-3861-3 45.00 EUR
- 177 Andreas Nehring: Wissenschaftliche Denk- und Arbeitsweisen im Fach Chemie. Eine kompetenzorientierte Modell- und Testentwicklung für den Bereich der Erkenntnisgewinnung
ISBN 978-3-8325-3872-9 39.50 EUR
- 178 Maike Schmidt: Professionswissen von Sachunterrichtslehrkräften. Zusammenhangsanalyse zur Wirkung von Ausbildungshintergrund und Unterrichtserfahrung auf das fachspezifische Professionswissen im Unterrichtsinhalt „Verbrennung“
ISBN 978-3-8325-3907-8 38.50 EUR

- 179 Jan Winkelmann: Auswirkungen auf den Fachwissenszuwachs und auf affektive Schülermerkmale durch Schüler- und Demonstrationsexperimente im Physikunterricht
ISBN 978-3-8325-3915-3 41.00 EUR
- 180 Iwen Kobow: Entwicklung und Validierung eines Testinstrumentes zur Erfassung der Kommunikationskompetenz im Fach Chemie
ISBN 978-3-8325-3927-6 34.50 EUR
- 181 Yvonne Gramzow: Fachdidaktisches Wissen von Lehramtsstudierenden im Fach Physik. Modellierung und Testkonstruktion
ISBN 978-3-8325-3931-3 42.50 EUR
- 182 Evelin Schröter: Entwicklung der Kompetenzerwartung durch Lösen physikalischer Aufgaben einer multimedialen Lernumgebung
ISBN 978-3-8325-3975-7 54.50 EUR
- 183 Inga Kallweit: Effektivität des Einsatzes von Selbsteinschätzungsbögen im Chemieunterricht der Sekundarstufe I. *Individuelle Förderung durch selbstreguliertes Lernen*
ISBN 978-3-8325-3965-8 44.00 EUR
- 184 Andrea Schumacher: Paving the way towards authentic chemistry teaching. *A contribution to teachers' professional development*
ISBN 978-3-8325-3976-4 48.50 EUR
- 185 David Woitkowski: Fachliches Wissen Physik in der Hochschulausbildung. *Konzeptualisierung, Messung, Niveaubildung*
ISBN 978-3-8325-3988-7 53.00 EUR
- 186 Marianne Korner: Cross-Age Peer Tutoring in Physik. *Evaluation einer Unterrichtsmethode*
ISBN 978-3-8325-3979-5 38.50 EUR
- 187 Simone Nakoinz: Untersuchung zur Verknüpfung submikroskopischer und makroskopischer Konzepte im Fach Chemie
ISBN 978-3-8325-4057-9 38.50 EUR
- 188 Sandra Anus: Evaluation individueller Förderung im Chemieunterricht. *Adaptivität von Lerninhalten an das Vorwissen von Lernenden am Beispiel des Basiskonzeptes Chemische Reaktion*
ISBN 978-3-8325-4059-3 43.50 EUR
- 189 Thomas Roßbegalle: Fachdidaktische Entwicklungsforschung zum besseren Verständnis atmosphärischer Phänomene. *Treibhauseffekt, saurer Regen und stratosphärischer Ozonabbau als Kontexte zur Vermittlung von Basiskonzepten der Chemie*
ISBN 978-3-8325-4059-3 45.50 EUR
- 190 Kathrin Steckenmesser-Sander: Gemeinsamkeiten und Unterschiede physikbezogener Handlungs-, Denk- und Lernprozesse von Mädchen und Jungen
ISBN 978-3-8325-4066-1 38.50 EUR
- 191 Cornelia Geller: Lernprozessorientierte Sequenzierung des Physikunterrichts im Zusammenhang mit Fachwissenserwerb. *Eine Videostudie in Finnland, Deutschland und der Schweiz*
ISBN 978-3-8325-4082-1 35.50 EUR

- 192 Jan Hofmann: Untersuchung des Kompetenzaufbaus von Physiklehrkräften während einer Fortbildungsmaßnahme
ISBN 978-3-8325-4104-0 38.50 EUR
- 193 Andreas Dickhäuser: Chemiespezifischer Humor. *Theoriebildung, Materialentwicklung, Evaluation*
ISBN 978-3-8325-4108-8 37.00 EUR
- 194 Stefan Korte: Die Grenzen der Naturwissenschaft als Thema des Physikunterrichts
ISBN 978-3-8325-4112-5 57.50 EUR
- 195 Carolin Hülsmann: Kurswahlmotive im Fach Chemie. Eine Studie zum Wahlverhalten und Erfolg von Schülerinnen und Schülern in der gymnasialen Oberstufe
ISBN 978-3-8325-4144-6 49.00 EUR
- 196 Caroline Körbs: Mindeststandards im Fach Chemie am Ende der Pflichtschulzeit
ISBN 978-3-8325-4148-4 34.00 EUR
- 197 Andreas Vorholzer: Wie lassen sich Kompetenzen des experimentellen Denkens und Arbeitens fördern? *Eine empirische Untersuchung der Wirkung eines expliziten und eines impliziten Instruktionsansatzes*
ISBN 978-3-8325-4194-1 37.50 EUR
- 198 Anna Katharina Schmitt: Entwicklung und Evaluation einer Chemielehrerfortbildung zum Kompetenzbereich Erkenntnisgewinnung
ISBN 978-3-8325-4228-3 39.50 EUR
- 199 Christian Maurer: Strukturierung von Lehr-Lern-Sequenzen
ISBN 978-3-8325-4247-4 36.50 EUR
- 200 Helmut Fischler, Elke Sumfleth (Hrsg.): Professionelle Kompetenz von Lehrkräften der Chemie und Physik
ISBN 978-3-8325-4523-9 34.00 EUR
- 201 Simon Zander: Lehrerfortbildung zu Basismodellen und Zusammenhänge zum Fachwissen
ISBN 978-3-8325-4248-1 35.00 EUR
- 202 Kerstin Arndt: Experimentierkompetenz erfassen.
Analyse von Prozessen und Mustern am Beispiel von Lehramtsstudierenden der Chemie
ISBN 978-3-8325-4266-5 45.00 EUR
- 203 Christian Lang: Kompetenzorientierung im Rahmen experimentalchemischer Praktika
ISBN 978-3-8325-4268-9 42.50 EUR
- 204 Eva Cauet: Testen wir relevantes Wissen? *Zusammenhang zwischen dem Professionswissen von Physiklehrkräften und gutem und erfolgreichem Unterrichten*
ISBN 978-3-8325-4276-4 39.50 EUR
- 205 Patrick Löffler: Modellanwendung in Problemlöseaufgaben. *Wie wirkt Kontext?*
ISBN 978-3-8325-4303-7 35.00 EUR

- 206 Carina Gehlen: Kompetenzstruktur naturwissenschaftlicher Erkenntnisgewinnung im Fach Chemie
ISBN 978-3-8325-4318-1 43.00 EUR
- 207 Lars Oettinghaus: Lehrerüberzeugungen und physikbezogenes Professionswissen. *Vergleich von Absolventinnen und Absolventen verschiedener Ausbildungswege im Physikreferendariat*
ISBN 978-3-8325-4319-8 38.50 EUR
- 208 Jennifer Petersen: Zum Einfluss des Merkmals Humor auf die Gesundheitsförderung im Chemieunterricht der Sekundarstufe I. *Eine Interventionsstudie zum Thema Sonnenschutz*
ISBN 978-3-8325-4348-8 40.00 EUR
- 209 Philipp Straube: Modellierung und Erfassung von Kompetenzen naturwissenschaftlicher Erkenntnisgewinnung bei (Lehramts-) Studierenden im Fach Physik
ISBN 978-3-8325-4351-8 35.50 EUR
- 210 Martin Dickmann: Messung von Experimentierfähigkeiten. *Validierungsstudien zur Qualität eines computerbasierten Testverfahrens*
ISBN 978-3-8325-4356-3 41.00 EUR
- 211 Markus Bohlmann: Science Education. Empirie, Kulturen und Mechanismen der Didaktik der Naturwissenschaften
ISBN 978-3-8325-4377-8 44.00 EUR
- 212 Martin Draude: Die Kompetenz von Physiklehrkräften, Schwierigkeiten von Schülerinnen und Schülern beim eigenständigen Experimentieren zu diagnostizieren
ISBN 978-3-8325-4382-2 37.50 EUR
- 213 Henning Rode: Prototypen evidenzbasierten Physikunterrichts. *Zwei empirische Studien zum Einsatz von Feedback und Blackboxes in der Sekundarstufe*
ISBN 978-3-8325-4389-1 42.00 EUR
- 214 Jan-Henrik Kechel: Schülerschwierigkeiten beim eigenständigen Experimentieren. *Eine qualitative Studie am Beispiel einer Experimentieraufgabe zum Hooke'schen Gesetz*
ISBN 978-3-8325-4392-1 55.00 EUR
- 215 Katharina Fricke: Classroom Management and its Impact on Lesson Outcomes in Physics. *A multi-perspective comparison of teaching practices in primary and secondary schools*
ISBN 978-3-8325-4394-5 40.00 EUR
- 216 Hannes Sander: Orientierungen von Jugendlichen beim Urteilen und Entscheiden in Kontexten nachhaltiger Entwicklung. *Eine rekonstruktive Perspektive auf Bewertungskompetenz in der Didaktik der Naturwissenschaft*
ISBN 978-3-8325-4434-8 46.00 EUR
- 217 Inka Haak: Maßnahmen zur Unterstützung kognitiver und metakognitiver Prozesse in der Studieneingangsphase. *Eine Design-Based-Research-Studie zum universitären Lernzentrum Physiktreff*
ISBN 978-3-8325-4437-9 46.50 EUR

- 218 Martina Brandenburger: Was beeinflusst den Erfolg beim Problemlösen in der Physik?
Eine Untersuchung mit Studierenden
ISBN 978-3-8325-4409-6 42.50 EUR
- 219 Corinna Helms: Entwicklung und Evaluation eines Trainings zur Verbesserung der Erklärqualität von Schülerinnen und Schülern im Gruppenpuzzle
ISBN 978-3-8325-4454-6 42.50 EUR
- 220 Viktoria Rath: Diagnostische Kompetenz von angehenden Physiklehrkräften. *Modellierung, Testinstrumentenentwicklung und Erhebung der Performanz bei der Diagnose von Schülervorstellungen in der Mechanik*
ISBN 978-3-8325-4456-0 42.50 EUR
- 221 Janne Krüger: Schülerperspektiven auf die zeitliche Entwicklung der Naturwissenschaften
ISBN 978-3-8325-4457-7 45.50 EUR
- 222 Stefan Mutke: Das Professionswissen von Chemiereferendarinnen und -referendaren in Nordrhein-Westfalen. *Eine Längsschnittstudie*
ISBN 978-3-8325-4458-4 37.50 EUR
- 223 Sebastian Habig: Systematisch variierte Kontextaufgaben und ihr Einfluss auf kognitive und affektive Schülerfaktoren
ISBN 978-3-8325-4467-6 40.50 EUR
- 224 Sven Liepertz: Zusammenhang zwischen dem Professionswissen von Physiklehrkräften, dem sachstrukturellen Angebot des Unterrichts und der Schülerleistung
ISBN 978-3-8325-4480-5 34.00 EUR
- 225 Elina Platova: Optimierung eines Laborpraktikums durch kognitive Aktivierung
ISBN 978-3-8325-4481-2 39.00 EUR
- 226 Tim Reschke: Lesegeschichten im Chemieunterricht der Sekundarstufe I zur Unterstützung von situationalem Interesse und Lernerfolg
ISBN 978-3-8325-4487-4 41.00 EUR
- 227 Lena Mareike Walper: Entwicklung der physikbezogenen Interessen und selbstbezogenen Kognitionen von Schülerinnen und Schülern in der Übergangsphase von der Primar- in die Sekundarstufe. *Eine Längsschnittanalyse vom vierten bis zum siebten Schuljahr*
ISBN 978-3-8325-4495-9 43.00 EUR
- 228 Stefan Anthofer: Förderung des fachspezifischen Professionswissens von Chemielehramtsstudierenden
ISBN 978-3-8325-4498-0 39.50 EUR
- 229 Marcel Bullinger: Handlungsorientiertes Physiklernen mit instruierten Selbsterklärungen in der Primarstufe. *Eine experimentelle Laborstudie*
ISBN 978-3-8325-4504-8 44.00 EUR
- 230 Thomas Amenda: Bedeutung fachlicher Elementarisierungen für das Verständnis der Kinematik
ISBN 978-3-8325-4531-4 43.50 EUR

- 231 Sabrina Milke: Beeinflusst *Priming* das Physiklernen?
Eine empirische Studie zum Dritten Newtonschen Axiom
ISBN 978-3-8325-4549-4 42.00 EUR
- 232 Corinna Erfmann: Ein anschaulicher Weg zum Verständnis der elektromagnetischen Induktion. *Evaluation eines Unterrichtsvorschlags und Validierung eines Leistungsdiagnoseinstruments*
ISBN 978-3-8325-4550-5 49.50 EUR
- 233 Hanne Rautenstrauch: Erhebung des (Fach-)Sprachstandes bei Lehramtsstudierenden im Kontext des Faches Chemie
ISBN 978-3-8325-4556-7 40.50 EUR
- 234 Tobias Klug: Wirkung kontextorientierter physikalischer Praktikumsversuche auf Lernprozesse von Studierenden der Medizin
ISBN 978-3-8325-4558-1 37.00 EUR
- 235 Mareike Bohrmann: Zur Förderung des Verständnisses der Variablenkontrolle im naturwissenschaftlichen Sachunterricht
ISBN 978-3-8325-4559-8 52.00 EUR
- 236 Anja Schödl: FALKO-Physik – Fachspezifische Lehrerkompetenzen im Fach Physik. *Entwicklung und Validierung eines Testinstruments zur Erfassung des fachspezifischen Professionswissens von Physiklehrkräften*
ISBN 978-3-8325-4553-6 40.50 EUR
- 237 Hilda Scheuermann: Entwicklung und Evaluation von Unterstützungsmaßnahmen zur Förderung der Variablenkontrollstrategie beim Planen von Experimenten
ISBN 978-3-8325-4568-0 39.00 EUR
- 238 Christian G. Strippel: Naturwissenschaftliche Erkenntnisgewinnung an chemischen Inhalten vermitteln. *Konzeption und empirische Untersuchung einer Ausstellung mit Experimentierstation*
ISBN 978-3-8325-4577-2 41.50 EUR
- 239 Sarah Rau: Durchführung von Sachunterricht im Vorbereitungsdienst. *Eine längsschnittliche, videobasierte Unterrichtsanalyse*
ISBN 978-3-8325-4579-6 46.00 EUR
- 240 Thomas Plotz: Lernprozesse zu nicht-sichtbarer Strahlung. *Empirische Untersuchungen in der Sekundarstufe 2*
ISBN 978-3-8325-4624-3 39.50 EUR
- 241 Wolfgang Aschauer: Elektrische und magnetische Felder. *Eine empirische Studie zu Lernprozessen in der Sekundarstufe II*
ISBN 978-3-8325-4625-0 50.00 EUR
- 242 Anna Donhauser: Didaktisch rekonstruierte Materialwissenschaft. *Aufbau und Konzeption eines Schülerlabors für den Exzellenzcluster Engineering of Advanced Materials*
ISBN 978-3-8325-4636-6 39.00 EUR

- 243 Katrin Schüßler: Lernen mit Lösungsbeispielen im Chemieunterricht. *Einflüsse auf Lernerfolg, kognitive Belastung und Motivation*
ISBN 978-3-8325-4640-3 42.50 EUR
- 244 Timo Fleischer: Untersuchung der chemischen Fachsprache unter besonderer Berücksichtigung chemischer Repräsentationen
ISBN 978-3-8325-4642-7 46.50 EUR
- 245 Rosina Steininger: Concept Cartoons als Stimuli für Kleingruppendiskussionen im Chemieunterricht. *Beschreibung und Analyse einer komplexen Lerngelegenheit*
ISBN 978-3-8325-4647-2 39.00 EUR
- 246 Daniel Rehfeldt: Erfassung der Lehrqualität naturwissenschaftlicher Experimentalpraktika
ISBN 978-3-8325-4590-1 40.00 EUR
- 247 Sandra Puddu: Implementing Inquiry-based Learning in a Diverse Classroom: Investigating Strategies of Scaffolding and Students' Views of Scientific Inquiry
ISBN 978-3-8325-4591-8 35.50 EUR
- 248 Markus Bliersbach: Kreativität in der Chemie. *Erhebung und Förderung der Vorstellungen von Chemielehramtsstudierenden*
ISBN 978-3-8325-4593-2 44.00 EUR
- 249 Lennart Kimpel: Aufgaben in der Allgemeinen Chemie. *Zum Zusammenspiel von chemischem Verständnis und Rechenfähigkeit*
ISBN 978-3-8325-4618-2 36.00 EUR
- 250 Louise Bindel: Effects of integrated learning: explicating a mathematical concept in inquiry-based science camps
ISBN 978-3-8325-4655-7 37.50 EUR
- 251 Michael Wenzel: Computereinsatz in Schule und Schülerlabor. *Einstellung von Physiklehrkräften zu Neuen Medien*
ISBN 978-3-8325-4659-5 38.50 EUR
- 252 Laura Muth: Einfluss der Auswertephase von Experimenten im Physikunterricht. *Ergebnisse einer Interventionsstudie zum Zuwachs von Fachwissen und experimenteller Kompetenz von Schülerinnen und Schülern*
ISBN 978-3-8325-4675-5 36.50 EUR
- 253 Annika Fricke: Interaktive Skripte im Physikalischen Praktikum. *Entwicklung und Evaluation von Hypermedien für die Nebenfachausbildung*
ISBN 978-3-8325-4676-2 41.00 EUR
- 254 Julia Haase: Selbstbestimmtes Lernen im naturwissenschaftlichen Sachunterricht. *Eine empirische Interventionsstudie mit Fokus auf Feedback und Kompetenzerleben*
ISBN 978-3-8325-4685-4 38.50 EUR
- 255 Antje J. Heine: Was ist Theoretische Physik? *Eine wissenschaftstheoretische Betrachtung und Rekonstruktion von Vorstellungen von Studierenden und Dozenten über das Wesen der Theoretischen Physik*
ISBN 978-3-8325-4691-5 46.50 EUR

- 256 Claudia Meinhardt: Entwicklung und Validierung eines Testinstruments zu Selbstwirksamkeitserwartungen von (angehenden) Physiklehrkräften in physikdidaktischen Handlungsfeldern
ISBN 978-3-8325-4712-7 47.00 EUR
- 257 Ann-Kathrin Schlüter: Professionalisierung angehender Chemielehrkräfte für einen Gemeinsamen Unterricht
ISBN 978-3-8325-4713-4 53.50 EUR
- 258 Stefan Richtberg: Elektronenbahnen in Feldern. Konzeption und Evaluation einer webbasierten Lernumgebung
ISBN 978-3-8325-4723-3 49.00 EUR
- 259 Jan-Philipp Burde: Konzeption und Evaluation eines Unterrichtskonzepts zu einfachen Stromkreisen auf Basis des Elektronengasmodells
ISBN 978-3-8325-4726-4 57.50 EUR
- 260 Frank Finkenberg: Flipped Classroom im Physikunterricht
ISBN 978-3-8325-4737-4 42.50 EUR
- 261 Florian Treisch: Die Entwicklung der Professionellen Unterrichtswahrnehmung im Lehr-Lern-Labor Seminar
ISBN 978-3-8325-4741-4 41.50 EUR
- 262 Desiree Mayr: Strukturiertheit des experimentellen naturwissenschaftlichen Problemlöseprozesses
ISBN 978-3-8325-4757-8 37.00 EUR
- 263 Katrin Weber: Entwicklung und Validierung einer Learning Progression für das Konzept der chemischen Reaktion in der Sekundarstufe I
ISBN 978-3-8325-4762-2 48.50 EUR
- 264 Hauke Bartels: Entwicklung und Bewertung eines performanznahen Videovignetten-tests zur Messung der Erklärfähigkeit von Physiklehrkräften
ISBN 978-3-8325-4804-9 37.00 EUR
- 265 Karl Marniok: Zum Wesen von Theorien und Gesetzen in der Chemie. *Begriffsanalyse und Förderung der Vorstellungen von Lehramtsstudierenden*
ISBN 978-3-8325-4805-6 42.00 EUR
- 266 Marisa Holzapfel: Fachspezifischer Humor als Methode in der Gesundheitsbildung im Übergang von der Primarstufe zur Sekundarstufe I
ISBN 978-3-8325-4808-7 50.00 EUR
- 267 Anna Stolz: Die Auswirkungen von Experimentiersituationen mit unterschiedlichem Öffnungsgrad auf Leistung und Motivation der Schülerinnen und Schüler
ISBN 978-3-8325-4781-3 38.00 EUR
- 268 Nina Ulrich: Interaktive Lernaufgaben in dem digitalen Schulbuch eChemBook. *Einfluss des Interaktivitätsgrads der Lernaufgaben und des Vorwissens der Lernenden auf den Lernerfolg*
ISBN 978-3-8325-4814-8 43.50 EUR

- 269 Kim-Alessandro Weber: Quantenoptik in der Lehrerfortbildung. *Ein bedarfsgeprägtes Fortbildungskonzept zum Quantenobjekt „Photon“ mit Realexperimenten*
ISBN 978-3-8325-4792-9 55.00 EUR
- 270 Nina Skorsetz: Empathisierer und Systematisierer im Vorschulalter. *Eine Fragebogen- und Videostudie zur Motivation, sich mit Naturphänomenen zu beschäftigen*
ISBN 978-3-8325-4825-4 43.50 EUR
- 271 Franziska Kehne: Analyse des Transfers von kontextualisiert erworbenem Wissen im Fach Chemie
ISBN 978-3-8325-4846-9 45.00 EUR
- 272 Markus Elsholz: Das akademische Selbstkonzept angehender Physiklehrkräfte als Teil ihrer professionellen Identität. *Dimensionalität und Veränderung während einer zentralen Praxisphase*
ISBN 978-3-8325-4857-5 37.50 EUR
- 273 Joachim Müller: Studienerfolg in der Physik. *Zusammenhang zwischen Modellierungskompetenz und Studienerfolg*
ISBN 978-3-8325-4859-9 35.00 EUR
- 274 Jennifer Dörschelln: Organische Leuchtdioden. *Implementation eines innovativen Themas in den Chemieunterricht*
ISBN 978-3-8325-4865-0 59.00 EUR
- 275 Stephanie Strelow: Beliefs von Studienanfängern des Kombi-Bachelors Physik über die Natur der Naturwissenschaften
ISBN 978-3-8325-4881-0 40.50 EUR
- 276 Dennis Jaeger: Kognitive Belastung und aufgabenspezifische sowie personenspezifische Einflussfaktoren beim Lösen von Physikaufgaben
ISBN 978-3-8325-4928-2 50.50 EUR
- 277 Vanessa Fischer: Der Einfluss von Interesse und Motivation auf die Messung von Fach- und Bewertungskompetenz im Fach Chemie
ISBN 978-3-8325-4933-6 39.00 EUR
- 278 René Dohrmann: Professionsbezogene Wirkungen einer Lehr-Lern-Labor-Veranstaltung. *Eine multimethodische Studie zu den professionsbezogenen Wirkungen einer Lehr-Lern-Labor-Blockveranstaltung auf Studierende der Bachelorstudiengänge Lehramt Physik und Grundschulpädagogik (Sachunterricht)*
ISBN 978-3-8325-4958-9 40.00 EUR
- 279 Meike Bergs: Can We Make Them Use These Strategies? *Fostering Inquiry-Based Science Learning Skills with Physical and Virtual Experimentation Environments*
ISBN 978-3-8325-4962-6 39.50 EUR
- 280 Marie-Therese Hauerstein: Untersuchung zur Effektivität von Strukturierung und Binnendifferenzierung im Chemieunterricht der Sekundarstufe I. *Evaluation der Strukturierungshilfe Lernleiter*
ISBN 978-3-8325-4982-4 42.50 EUR

- 281 Verena Zucker: Erkennen und Beschreiben von formativem Assessment im naturwissenschaftlichen Grundschulunterricht. *Entwicklung eines Instruments zur Erfassung von Teilfähigkeiten der professionellen Wahrnehmung von Lehramtsstudierenden*
ISBN 978-3-8325-4991-6 38.00 EUR
- 282 Victoria Telser: Erfassung und Förderung experimenteller Kompetenz von Lehrkräften im Fach Chemie
ISBN 978-3-8325-4996-1 50.50 EUR
- 283 Kristine Tschirschky: Entwicklung und Evaluation eines gedächtnisorientierten Aufgabendesigns für Physikaufgaben
ISBN 978-3-8325-5002-8 42.50 EUR
- 284 Thomas Elert: Course Success in the Undergraduate General Chemistry Lab
ISBN 978-3-8325-5004-2 41.50 EUR
- 285 Britta Kalthoff: Explizit oder implizit? *Untersuchung der Lernwirksamkeit verschiedener fachmethodischer Instruktionen im Hinblick auf fachmethodische und fachinhaltliche Fähigkeiten von Sachunterrichtsstudierenden*
ISBN 978-3-8325-5013-4 37.50 EUR
- 286 Thomas Dickmann: Visuelles Modellverständnis und Studienerfolg in der Chemie. *Zwei Seiten einer Medaille*
ISBN 978-3-8325-5016-5 44.00 EUR
- 287 Markus Sebastian Feser: Physiklehrkräfte korrigieren Schülertexte. *Eine Explorationsstudie zur fachlich-konzeptuellen und sprachlichen Leistungsfeststellung und -beurteilung im Physikunterricht*
ISBN 978-3-8325-5020-2 49.00 EUR
- 288 Matylda Dudzinska: Lernen mit Beispielaufgaben und Feedback im Physikunterricht der Sekundarstufe 1. *Energieerhaltung zur Lösung von Aufgaben nutzen*
ISBN 978-3-8325-5025-7 47.00 EUR
- 289 Ines Sonnenschein: Naturwissenschaftliche Denk- und Arbeitsprozesse Studierender im Labor
ISBN 978-3-8325-5033-2 52.00 EUR
- 290 Florian Simon: Der Einfluss von Betreuung und Betreuenden auf die Wirksamkeit von Schülerlaborbesuchen. *Eine Zusammenhangsanalyse von Betreuungsqualität, Betreuermerkmalen und Schülerlaborzielen sowie Replikationsstudie zur Wirksamkeit von Schülerlaborbesuchen*
ISBN 978-3-8325-5036-3 49.50 EUR
- 291 Marie-Annette Geyer: Physikalisch-mathematische Darstellungswechsel funktionaler Zusammenhänge. *Das Vorgehen von SchülerInnen der Sekundarstufe 1 und ihre Schwierigkeiten*
ISBN 978-3-8325-5047-9 46.50 EUR
- 292 Susanne Digel: Messung von Modellierungskompetenz in Physik. *Theoretische Herleitung und empirische Prüfung eines Kompetenzmodells physikspezifischer Modellierungskompetenz*
ISBN 978-3-8325-5055-4 41.00 EUR

- 293 Sönke Janssen: Angebots-Nutzungs-Prozesse eines Schülerlabors analysieren und gestalten. *Ein design-based research Projekt*
ISBN 978-3-8325-5065-3 57.50 EUR
- 294 Knut Wille: Der Productive Failure Ansatz als Beitrag zur Weiterentwicklung der Aufgabenkultur
ISBN 978-3-8325-5074-5 49.00 EUR
- 295 Lisanne Kraeva: Problemlösestrategien von Schülerinnen und Schülern diagnostizieren
ISBN 978-3-8325-5110-0 59.50 EUR
- 296 Jenny Lorentzen: Entwicklung und Evaluation eines Lernangebots im Lehramtsstudium Chemie zur Förderung von Vernetzungen innerhalb des fachbezogenen Professionswissens
ISBN 978-3-8325-5120-9 39.50 EUR
- 297 Micha Winkelmann: Lernprozesse in einem Schülerlabor unter Berücksichtigung individueller naturwissenschaftlicher Interessenstrukturen
ISBN 978-3-8325-5147-6 48.50 EUR
- 298 Carina Wöhlke: Entwicklung und Validierung eines Instruments zur Erfassung der professionellen Unterrichtswahrnehmung angehender Physiklehrkräfte
ISBN 978-3-8325-5149-0 43.00 EUR
- 299 Thomas Schubatzky: Das Amalgam Anfangs-Elektrizitätslehreunterricht. *Eine multiperspektivische Betrachtung in Deutschland und Österreich*
ISBN 978-3-8325-5159-9 50.50 EUR
- 300 Amany Annaggar: A Design Framework for Video Game-Based Gamification Elements to Assess Problem-solving Competence in Chemistry Education
ISBN 978-3-8325-5150-6 52.00 EUR
- 301 Alexander Engl: CHEMIE PUR – Unterrichten in der Natur: *Entwicklung und Evaluation eines kontextorientierten Unterrichtskonzepts im Bereich Outdoor Education zur Änderung der Einstellung zu „Chemie und Natur“*
ISBN 978-3-8325-5174-2 59.00 EUR
- 302 Christin Marie Sajons: Kognitive und motivationale Dynamik in Schülerlaboren. *Kontextualisierung, Problemorientierung und Autonomieunterstützung der didaktischen Struktur analysieren und weiterentwickeln*
ISBN 978-3-8325-5155-1 56.00 EUR
- 303 Philipp Bitzenbauer: Quantenoptik an Schulen. *Studie im Mixed-Methods Design zur Evaluation des Erlanger Unterrichtskonzepts zur Quantenoptik*
ISBN 978-3-8325-5123-0 59.00 EUR
- 304 Malte S. Ubben: Typisierung des Verständnisses mentaler Modelle mittels empirischer Datenerhebung am Beispiel der Quantenphysik
ISBN 978-3-8325-5181-0 43.50 EUR
- 305 Wiebke Kuske-Janßen: Sprachlicher Umgang mit Formeln von LehrerInnen im Physikunterricht am Beispiel des elektrischen Widerstandes in Klassenstufe 8
ISBN 978-3-8325-5183-4 47.50 EUR

- 306 Kai Bliesmer: Physik der Küste für außerschulische Lernorte. *Eine Didaktische Rekonstruktion*
ISBN 978-3-8325-5190-2 58.00 EUR
- 307 Nikola Schild: Eignung von domänenspezifischen Studieneingangsvariablen als Prädiktoren für Studienerfolg im Fach und Lehramt Physik
ISBN 978-3-8325-5226-8 42.00 EUR
- 308 Daniel Averbeck: Zum Studienerfolg in der Studieneingangsphase des Chemiestudiums. *Der Einfluss kognitiver und affektiv-motivationaler Variablen*
ISBN 978-3-8325-5227-5 51.00 EUR
- 309 Martina Strübe: Modelle und Experimente im Chemieunterricht. *Eine Videostudie zum fachspezifischen Lehrerwissen und -handeln*
ISBN 978-3-8325-5245-9 45.50 EUR
- 310 Wolfgang Becker: Auswirkungen unterschiedlicher experimenteller Repräsentationen auf den Kenntnisstand bei Grundschulkindern
ISBN 978-3-8325-5255-8 50.00 EUR
- 311 Marvin Rost: Modelle als Mittel der Erkenntnisgewinnung im Chemieunterricht der Sekundarstufe I. *Entwicklung und quantitative Dimensionalitätsanalyse eines Testinstruments aus epistemologischer Perspektive*
ISBN 978-3-8325-5256-5 44.00 EUR
- 312 Christina Kobl: Förderung und Erfassung der Reflexionskompetenz im Fach Chemie
ISBN 978-3-8325-5259-6 41.00 EUR
- 313 Ann-Kathrin Beretz: Diagnostische Prozesse von Studierenden des Lehramts – *eine Videostudie in den Fächern Physik und Mathematik*
ISBN 978-3-8325-5288-6 45.00 EUR
- 314 Judith Breuer: Implementierung fachdidaktischer Innovationen durch das Angebot materialgestützter Unterrichtskonzeptionen. *Fallanalysen zum Nutzungsverhalten von Lehrkräften am Beispiel des Münchener Lehrgangs zur Quantenmechanik*
ISBN 978-3-8325-5293-0 50.50 EUR
- 315 Michaela Oettle: Modellierung des Fachwissens von Lehrkräften in der Teilchenphysik. *Eine Delphi-Studie*
ISBN 978-3-8325-5305-0 57.50 EUR
- 316 Volker Brüggemann: Entwicklung und Pilotierung eines adaptiven Multistage-Tests zur Kompetenzerfassung im Bereich naturwissenschaftlichen Denkens
ISBN 978-3-8325-5331-9 40.00 EUR
- 317 Stefan Müller: Die Vorläufigkeit und soziokulturelle Eingebundenheit naturwissenschaftlicher Erkenntnisse. *Kritische Reflexion, empirische Befunde und fachdidaktische Konsequenzen für die Chemielehrer*innenbildung*
ISBN 978-3-8325-5343-2 63.00 EUR
- 318 Laurence Müller: Alltagsentscheidungen für den Chemieunterricht erkennen und Entscheidungsprozesse explorativ begleiten
ISBN 978-3-8325-5379-1 59.00 EUR

- 319 Lars Ehlert: Entwicklung und Evaluation einer Lehrkräftefortbildung zur Planung von selbstgesteuerten Experimenten
ISBN 978-3-8325-5393-71 41.50 EUR
- 320 Florian Seiler: Entwicklung und Evaluation eines Seminarkonzepts zur Förderung der experimentellen Planungskompetenz von Lehramtsstudierenden im Fach Chemie
ISBN 978-3-8325-5397-5 47.50 EUR
- 321 Nadine Boele: Entwicklung eines Messinstruments zur Erfassung der professionellen Unterrichtswahrnehmung von (angehenden) Chemielehrkräften hinsichtlich der Lernunterstützung
ISBN 978-3-8325-5402-6 46.50 EUR
- 322 Franziska Zimmermann: Entwicklung und Evaluation digitalisierungsbezogener Kompetenzen von angehenden Chemielehrkräften
ISBN 978-3-8325-5410-1 49.50 EUR
- 323 Lars-Frederik Weiß: Der Flipped Classroom in der Physik-Lehre. *Empirische Untersuchungen in Schule und Hochschule*
ISBN 978-3-8325-5418-7 51.00 EUR
- 324 Tilmann Steinmetz: Kumulatives Lehren und Lernen im Lehramtsstudium Physik. *Theorie und Evaluation eines Lehrkonzepts*
ISBN 978-3-8325-5421-7 51.00 EUR
- 325 Kübra Nur Celik: Entwicklung von chemischem Fachwissen in der Sekundarstufe I. *Validierung einer Learning Progression für die Basiskonzepte „Struktur der Materie“, „Chemische Reaktion“ und „Energie“ im Kompetenzbereich „Umgang mit Fachwissen“*
ISBN 978-3-8325-5431-6 55.00 EUR
- 326 Matthias Ungermann: Förderung des Verständnisses von Nature of Science und der experimentellen Kompetenz im Schüler*innen-Labor Physik in Abgrenzung zum Regelunterricht
ISBN 978-3-8325-5442-2 55.50 EUR
- 327 Christoph Hoyer: Multimedial unterstütztes Experimentieren im webbasierten Labor zur Messung, Visualisierung und Analyse des Feldes eines Permanentmagneten
ISBN 978-3-8325-5453-8 45.00 EUR
- 328 Tobias Schüttler: Schülerlabore als interesselördernde authentische Lernorte für den naturwissenschaftlichen Unterricht nutzen
ISBN 978-3-8325-5454-5 50.50 EUR
- 329 Christopher Kurth: Die Kompetenz von Studierenden, Schülerschwierigkeiten beim eigenständigen Experimentieren zu diagnostizieren
ISBN 978-3-8325-5457-6 58.50 EUR
- 330 Dagmar Michna: Inklusiver Anfangsunterricht Chemie *Entwicklung und Evaluation einer Unterrichtseinheit zur Einführung der chemischen Reaktion*
ISBN 978-3-8325-5463-7 49.50 EUR
- 331 Marco Seiter: Die Bedeutung der Elementarisierung für den Erfolg von Mechanikunterricht in der Sekundarstufe I
ISBN 978-3-8325-5471-2 66.00 EUR

- 332 Jörn Hägele: Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten. *Videobasierte Analysen zu Aktivitäten und Vorstellungen von Schülerinnen und Schülern der gymnasialen Oberstufe bei der Bearbeitung von fachmethodischer Instruktion*
ISBN 978-3-8325-5476-7 56.50 EUR
- 333 Erik Heine: Wissenschaftliche Kontroversen im Physikunterricht. *Explorationsstudie zum Umgang von Physiklehrkräften und Physiklehramtsstudierenden mit einer wissenschaftlichen Kontroverse am Beispiel der Masse in der Speziellen Relativitätstheorie*
ISBN 978-3-8325-5478-1 48.50 EUR
- 334 Simon Goertz: Module und Lernzirkel der Plattform FLexKom zur Förderung experimenteller Kompetenzen in der Schulpraxis *Verlauf und Ergebnisse einer Design-Based Research Studie*
ISBN 978-3-8325-5494-1 66.50 EUR
- 335 Christina Toschka: Lernen mit Modellexperimenten *Empirische Untersuchung der Wahrnehmung und des Denkens in Analogien beim Umgang mit Modellexperimenten*
ISBN 978-3-8325-5495-8 50.00 EUR
- 336 Alina Behrendt: Chemiebezogene Kompetenzen in der Übergangsphase zwischen dem Sachunterricht der Primarstufe und dem Chemieunterricht der Sekundarstufe I
ISBN 978-3-8325-5498-9 40.50 EUR
- 337 Manuel Daiber: Entwicklung eines Lehrkonzepts für eine elementare Quantenmechanik *Formuliert mit In-Out Symbolen*
ISBN 978-3-8325-5507-8 48.50 EUR
- 338 Felix Pawlak: Das Gemeinsame Experimentieren (an-)leiten *Eine qualitative Studie zum chemiespezifischen Classroom-Management*
ISBN 978-3-8325-5508-5 46.50 EUR
- 339 Liza Dopatka: Konzeption und Evaluation eines kontextstrukturierten Unterrichtskonzeptes für den Anfangs-Elektrizitätslehreunterricht
ISBN 978-3-8325-5514-6 69.50 EUR
- 340 Arne Bewersdorff: Untersuchung der Effektivität zweier Fortbildungsformate zum Experimentieren mit dem Fokus auf das Unterrichtshandeln
ISBN 978-3-8325-5522-1 39.00 EUR
- 341 Thomas Christoph Münster: Wie diagnostizieren Studierende des Lehramtes physikbezogene Lernprozesse von Schüler*innen? Eine Videostudie zur Mechanik
ISBN 978-3-8325-5534-4 44.50 EUR
- 342 Ines Komor: Förderung des symbolisch-mathematischen Modellverständnisses in der Physikalischen Chemie
ISBN 978-3-8325-5546-7 46.50 EUR
- 343 Verena Petermann: Überzeugungen von Lehrkräften zum Lehren und Lernen von Fachinhalten und Fachmethoden und deren Beziehung zu unterrichtsnahem Handeln
ISBN 978-3-8325-5545-0 47.00 EUR

- 344 Jana Heinze: Einfluss der sprachlichen Konzeption auf die Einschätzung der Qualität instruktionaler Unterrichtserklärungen im Fach Physik
ISBN 978-3-8325-5545-0 47.00 EUR
- 345 Jannis Weber: Mathematische Modellbildung und Videoanalyse zum Lernen der Newtonschen Dynamik im Vergleich
ISBN 978-3-8325-5566-5 68.00 EUR
- 346 Fabian Sterzing: Zur Lernwirksamkeit von Erklärvideos in der Physik *Eine Untersuchung in Abhängigkeit von ihrer fachdidaktischen Qualität und ihrem Einbettungsformat*
ISBN 978-3-8325-5576-4 52.00 EUR
- 347 Lars Greitemann: Wirkung des Tablet-Einsatzes im Chemieunterricht der Sekundarstufe I unter besonderer Berücksichtigung von Wissensvermittlung und Wissenssicherung
ISBN 978-3-8325-5580-1 50.00 EUR
- 348 Fabian Poensgen: Diagnose experimenteller Kompetenzen in der laborpraktischen Chemielehrer*innenbildung
ISBN 978-3-8325-5587-0 48.00 EUR
- 349 William Lindlahr: Virtual-Reality-Experimente *Entwicklung und Evaluation eines Konzepts für den forschend-entwickelnden Physikunterricht mit digitalen Medien*
ISBN 978-3-8325-5595-5 49.00 EUR
- 350 Bert Schlüter: Teilnahmemotivation und situationales Interesse von Kindern und Eltern im experimentellen Lernsetting KEMIE
ISBN 978-3-8325-5598-6 43.00 EUR
- 351 Katharina Nave: Charakterisierung situativer mentaler Modellkomponenten in der Chemie und die Bildung von Hypothesen *Eine qualitative Studie zur Operationalisierung mentaler Modell-komponenten für den Fachbereich Chemie*
ISBN 978-3-8325-5599-3 43.00 EUR

Alle erschienenen Bücher können unter der angegebenen ISBN direkt online (<http://www.logos-verlag.de>) oder per Fax (030 - 42 85 10 92) beim Logos Verlag Berlin bestellt werden.

Studien zum Physik- und Chemielernen

Herausgegeben von Martin Hopf, Hans Niedderer, Mathias Ropohl und Elke Sumfleth

Die Reihe umfasst inzwischen eine große Zahl von wissenschaftlichen Arbeiten aus vielen Arbeitsgruppen der Physik- und Chemiedidaktik und zeichnet damit ein gültiges Bild der empirischen physik- und chemiedidaktischen Forschung im deutschsprachigen Raum.

Die Herausgeber laden daher Interessenten zu neuen Beiträgen ein und bitten sie, sich im Bedarfsfall an den Logos-Verlag oder an ein Mitglied des Herausgeberteams zu wenden.

Kontaktadressen:

Univ.-Prof. Dr. Martin Hopf
Universität Wien,
Österreichisches Kompetenzzentrum
für Didaktik der Physik,
Porzellangasse 4, Stiege 2,
1090 Wien, Österreich,
Tel. +43-1-4277-60330,
e-mail: martin.hopf@univie.ac.at

Prof. Dr. Hans Niedderer
Inst. f. Didaktik der Naturwissenschaften,
Abt. Physikdidaktik,
FB Physik/ Elektrotechnik,
Universität Bremen,
Postfach 33 04 40, 28334 Bremen
Tel. 0421-218 4695 (Sekretariat),
e-mail: niedderer@physik.uni-bremen.de

Prof. Dr. Mathias Ropohl
Didaktik der Chemie,
Fakultät für Chemie,
Universität Duisburg-Essen,
Schützenbahn 70, 45127 Essen,
Tel. 0201-183 2704,
e-mail: mathias.ropohl@uni-due.de

Prof. Dr. Elke Sumfleth
Didaktik der Chemie,
Fakultät für Chemie,
Universität Duisburg-Essen,
Schützenbahn 70, 45127 Essen
Tel. 0201-183 3757/3761,
e-mail: elke.sumfleth@uni-due.de

Die zentralen Aussagen der Newtonschen Mechanik sind für Lernende nur schwer zu verstehen. Nicht zur physikalischen Theorie passende Schülervorstellungen erschweren den Aufbau von Konzeptverständnis. Dies liegt u. a. an der wahrgenommenen Unvereinbarkeit der Newtonschen Gesetze mit Alltagserfahrungen der Lernenden. Oft vorgeschlagen wird hier Computereinsatz zur Bearbeitung authentischer Bewegungen mit Reibungseinflüssen.

Die vorliegende Mixed-Methods-Studie ($N = 274$) umfasst die Entwicklung und Pilotierung eines Testinstruments, die theoriegeleitete Konzeption zweier Interventionen zur Newtonschen Dynamik für die gymnasiale Oberstufe (mathematische Modellbildung bzw. Videoanalyse von Bewegungen) sowie deren quantitative und qualitative Evaluation. Das Ziel war, ein angemessenes Kraftverständnis bei den Lernenden aufzubauen.

Mit Blick auf den gesamten Lernzuwachs ergibt sich kein Unterschied zwischen den Varianten, wohl aber in einzelnen Inhaltsbereichen und im Modellverständnis der Lernenden. Eine Mehrebenenanalyse zeigt, dass u. a. das Interesse an theoretischen Zusammenhängen und die kognitive Belastung den Lernzuwachs beeinflussen. Die Analyse von Bildschirmaufnahmen deutet an, dass die Phase der Ergebnissicherung und die Nutzung von Modell- bzw. Messdaten wichtig für den Lernerfolg sein könnten. Die Thematisierung von komplexen Bewegungen mithilfe des Computers ist erfolgreich und führt in beiden Gruppen zu einer starken Verbesserung im Kraftverständnis.

Logos Verlag Berlin

ISBN 978-3-8325-5566-5