

Steffens, Ulrich

Was ist das Wichtigste beim Lernen? Zu Ertrag und Grenzen der Hattie-Studie "Visible Learning"

Steffens, Ulrich [Hrsg.]; Messner, Rudolf [Hrsg.]: *Unterrichtsqualität. Konzepte und Bilanzen gelingenden Lehrens und Lernens*. Münster ; New York : Waxmann 2019, S. 251-277. - (Grundlagen der Qualität von Schule; 3)



Quellenangabe/ Reference:

Steffens, Ulrich: Was ist das Wichtigste beim Lernen? Zu Ertrag und Grenzen der Hattie-Studie "Visible Learning" - In: Steffens, Ulrich [Hrsg.]; Messner, Rudolf [Hrsg.]: *Unterrichtsqualität. Konzepte und Bilanzen gelingenden Lehrens und Lernens*. Münster ; New York : Waxmann 2019, S. 251-277 - URN: urn:nbn:de:0111-pedocs-329854 - DOI: 10.25656/01:32985

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-329854>

<https://doi.org/10.25656/01:32985>

in Kooperation mit / in cooperation with:



WAXMANN
www.waxmann.com

<http://www.waxmann.com>

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

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

Ulrich Steffens

Was ist das Wichtigste beim Lernen? – Zu Ertrag und Grenzen der Hattie-Studie „Visible Learning“

Ulrich Steffens,
Rudolf Messner (Hrsg.)

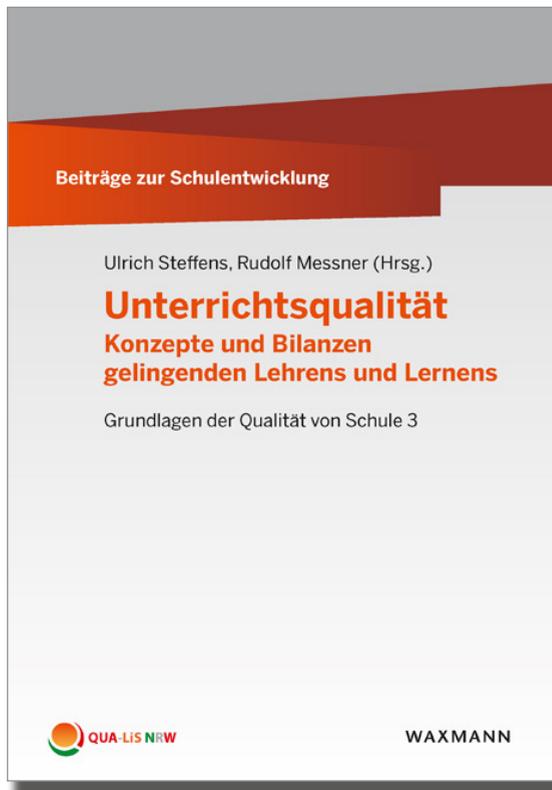
Unterrichtsqualität

Konzepte und Bilanzen
gelingenden
Lehrens und Lernens

Grundlagen der Qualität
von Schule 3

Beiträge zur Schulentwicklung,
2019, 424 Seiten, br., 39,90 €,
ISBN 978-3-8309-3937-5

E-Book: 35,99 €,
ISBN 978-3-8309-8937-0



Waxmann Verlag GmbH, 2019

Alle Rechte vorbehalten. Nachdruck, auch auszugsweise, verboten. Kein Teil dieses Werkes darf ohne schriftliche Genehmigung des Verlages in irgendeiner Form reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.



WAXMANN

Steinfurter Str. 555
48159 Münster

Fon 02 51 – 2 65 04-0
Fax 02 51 – 2 65 04-26

info@waxmann.com
www.waxmann.com

Mehr zum Buch [hier](#).

Ulrich Steffens

Was ist das Wichtigste beim Lernen? – Zu Ertrag und Grenzen der Hattie-Studie „Visible Learning“

Der Neuseeländer John Hattie, Professor für Erziehungswissenschaft und derzeit Direktor des Educational Research Institute der Universität Melbourne, hat die pädagogischen Diskussionen in Wissenschaft und Praxis maßgeblich beeinflusst. Von manchen wird er gar als „einflussreichster Bildungsforscher der Welt“ angesehen. Grund dafür ist vor allem seine Forschungsbilanz mit dem bezeichnenden Titel „Visible Learning“ (Hattie, 2009; auf Deutsch: Hattie, 2013).

1. Anliegen und Anlage der Studie

Hattie geht es in seiner Forschungsarbeit um eine Identifizierung und Bilanzierung der zentralen Einflussgrößen für den Lernerfolg von Schülerinnen und Schülern. Dabei handelt es sich nicht um irgendeine Bilanz, vielmehr geht es um einen Überblick über die internationale Lehr- und Lernforschung. Vor allem aus zwei Gründen ist diese Studie einzigartig: Erstens wird in singulärer Weise eine Forschungsarbeit über das breite Spektrum von 138 Einflussfaktoren (bzw. Konzepte und Programme) zum Lernerfolg vorgelegt. Und zweitens wird zum ersten Mal der Versuch einer Gesamtschau aller Studien unternommen, die zu diesen Einflussfaktoren (in englischer Sprache) vorliegen. Mit 52.637 Studien (vgl. Hattie, 2013, S. 18) handelt es sich um die größte Datenbasis zur Unterrichtsforschung, die jemals in einer Arbeit ausgewertet wurde.

Obwohl Hattie an der Studie 15 Jahre gearbeitet hatte, wäre für ihn eine Lektüre aller dieser Studien und eine darauf basierende inhaltsanalytische Auswertung nicht möglich gewesen. Vielmehr wurde ein statistisches Verfahren angewandt, bei dem die zentralen Ergebnisse einzelner Studien zu einem Untersuchungsbereich erfasst und miteinander in Beziehung gesetzt werden können. Hattie konnte auf über 700 solcher sogenannter „Metaanalysen“, denen die erwähnten ca. 50.000 Studien zugrunde liegen, zurückgreifen.¹ Hattie arbeitet nach wie vor an einer Fortführung der Synthese. Inzwi-

1 Hattie selbst gibt die Zahl der Metaanalysen meistens mit 815, manchmal auch mit 816 oder mit „über 800“ an. Der akribische Leser kommt zu einem anderen Ergebnis. Die Übersetzer der Studie haben nochmals nachgerechnet und ermitteln 736 Metaanalysen (Beywl & Zierer, 2013, S. XI); Rolf Schulmeister und Jörn Loviscach, die sich mit der Hattie-Studie in erster Linie methodenkritisch auseinandersetzen, identifizieren 724 verarbeitete Metaanalysen (Schulmeister & Loviscach, 2104, S. 129). Der Grund für die Diskrepanzen dürfte v.a. darin zu sehen sein, dass einige Metaanalysen für die Einflussgrößen mehrfach berücksichtigt wurden. Trotz-

schen (Stand: August 2017) beträgt die Grundlage der Auswertungen 1.400 Metaanalysen. Die Datenbasis beträgt nunmehr ca. 80.000 Primärstudien. Dabei wurde auch das Spektrum der Einflussfaktoren von 138 auf 250 erhöht.² Ein solch ambitionierter Versuch ist nur durch ein methodisches Verfahren möglich, das es erlaubt, über ganz verschiedene Studien hinweg Ergebnisse miteinander zu vergleichen.

Dazu bedient man sich sogenannter „Effektmaße“, die etwas über die praktische Wirksamkeit von Einflussfaktoren aussagen. Bei den Effektmaßen geht es um Zusammenhänge zwischen zwei Variablen, im Regelfall auf der Basis von Mittelwerten. Setzt man beispielsweise ein Programm zur Förderung von Leseverständnis ein und will nach einem Jahr seine Wirkungen feststellen, so dürfte ein Lesetest zum Einsatz kommen, bei dem die Programmgruppe mit einer Kontrollgruppe verglichen wird. Für den Lesetest werden Mittelwerte der beiden Gruppen berechnet und miteinander verglichen. Das dabei ermittelte Differenzmaß gibt Auskunft über die Ähnlichkeit der beiden Mittelwerte. Je mehr sich die Mittelwerte unterscheiden, umso größer fällt das Effektmaß aus (wobei auch die Standardabweichungen der Mittelwerte bei den Berechnungen berücksichtigt werden).³ Als Effektstärkemaß wird oft „Cohens d“ verwendet, auch von Hattie. Maße mit $d < .20$ gelten als unbedeutend; bei einem $d \geq .20$ wird von geringen Effekten gesprochen, $d \geq .40$ gelten als moderate Effekte und $d \geq .60$ werden als große Effekte gewertet. Um eine Vorstellung über die praktische Bedeutung von Effektmaßen zu erhalten, wird gerne darauf hingewiesen, dass ein Effektmaß von $d = .30$ bis $d = .40$, je nach Untersuchungsgegenstand, den durchschnittlichen Lernzuwachs von einem Schuljahr ausdrückt.

In den Metaanalysen werden die in den Primärstudien ermittelten Effektmaße in einem Mittelwert über alle berücksichtigten Studien eines Untersuchungsbereichs bzw. eines Einflussfaktors zusammengefasst. Hatties „Synthese“ stellt eine Zusammenfassung aller Mittelwerte der Effektmaße d zu den einzelnen Metaanalysen eines Einflussfaktors in Form eines weiteren Mittelwerts dar. Es handelt sich sozusagen um eine Meta-Metaanalyse. Die so gebildeten Mittelwerte der 138 Einflussgrößen stellen das Grundgerüst seiner Analysen dar. Grundlage dafür sind im Regelfall eine Reihe von Metaanalysen (oftmals 5–10), die sich einem Faktor zurechnen lassen. Die Spannweite reicht von einer Metaanalyse bei den Faktoren „Kognitive Entwicklungsstufen nach

dem überrascht die ungenaue Feststellung in einer Forschungsarbeit, in der Zahlen eine große Rolle spielen.

2. Leider werden die 112 neuen Faktoren nicht erläutert, so dass sich der bisherige Nutzen von Hatties Ausweitung seiner Analysen für Außenstehende in Grenzen hält.
3. Hattie hat die Effektstärke wie folgt berechnet: Mittelwert der Interventionsgruppe („treatment“) minus Mittelwert der Kontrollgruppe („control“), geteilt durch die gemeinsame Standardabweichung der beiden Gruppen. Das Effektmaß kommt auch bei Vorher-Nachher-Vergleichen zur Anwendung – bei folgender Berechnungsweise: Mittelwert am Ende der Intervention minus Mittelwert zu Beginn einer Intervention, geteilt durch die gepoolte Standardabweichung der beiden Gruppen (Hattie 2009 S. 8). Statt der gemeinsamen Standardabweichung wird auch die Streuung der Kontrollgruppe verwendet. Bei ungleichen Gruppengrößen und unterschiedlichen Varianzen sowie bei der Berechnung von Effektstärken mehrerer Studien kommen etwas abgewandelte Formeln zur Anwendung. „Diese Maße drücken hinsichtlich der abhängigen Variablen (z.B. Lesekompetenz) den Verteilungsabstand von Interventions- und Kontrollgruppe (z.B. Teilnahme vs. Nichtteilnahme an einer Sprachfördermaßnahme) in Einheiten der Standardabweichung aus.“ (Pant, 2014, S. 84) Für eine genauere Beschreibung des Effektmaßes vgl. Köller, Möller & Möller, 2013, S. 18f.

Piaget“ ($d = 1.28$) und „Kreativität“ ($d = .37$) bis hin zu 81 Metaanalysen beim Faktor „Computerunterstützung“ ($d = .37$). Auf diese Weise können die Effektmaße für die untersuchten Einflussfaktoren über viele Studien hinweg identifiziert und miteinander verglichen werden.

So beeindruckend die Vielzahl der in seinen Analysen berücksichtigten Studien und Einflussfaktoren ist, so macht Hattie keine Angaben darüber, welche Kriterien bei der Auswahl der Metaanalysen und Faktoren für den Lernerfolg zugrunde lagen – obwohl dies bei wissenschaftlichen Arbeiten eigentlich üblich ist, sei es, um auf methodenkritische Implikationen einzugehen oder um den theoretischen Zugang zu erläutern. Das gilt auch für die sechs Untersuchungsbereiche, denen die 138 Einflussfaktoren zugeordnet werden. Es handelt sich dabei um folgende Bereiche:

- „Die Beiträge der Lernenden“ (mit 7 Einflussfaktoren unter Berücksichtigung von 35 Metaanalysen),
- „Die Beiträge des Elternhauses“ (19 Faktoren, 139 Metaanalysen),
- „Die Beiträge der Schule“ (28 Faktoren, 101 Metaanalysen),
- „Die Beiträge der Lehrperson“ (10 Faktoren, 31 Metaanalysen),
- „Die Beiträge der Curricula“ (25 Faktoren, 144 Metaanalysen),
- „Die Beiträge des Unterrichtens“ (49 Faktoren, 365 Metaanalysen) – in zwei Kapiteln unterteilt.

In Hatties Forschungsbilanz werden die 138 Einflussfaktoren im Regelfall nicht definiert; oftmals sind entsprechende Anhaltspunkte aus der Darstellung der aufgeführten Befunde zu erschließen. Insofern setzt die Lektüre Vorkenntnisse aus der Lehr- und Lernforschung voraus. Für eine Bewertung der Ergebnisse gilt es zu beachten, dass als Erfolgskriterium für Schülerleistungen in erster Linie Fachleistungen herangezogen werden, die vorrangig in Leistungstests in den klassischen Hauptfächern gewonnen wurden.

Die zugrundeliegenden Primärstudien bzw. Metaanalysen beziehen sich auf Lehr- und Lernprozesse über die gesamte Lebensspanne, angefangen von frühkindlichen Lernprozessen, über Bildungsprozesse in Hochschulen bis hin ins Seniorenalter. Da oftmals nicht klar ersichtlich ist, auf welche Adressatengruppe bzw. Bildungsinstitution sich die vorgestellten Effekte beziehen, ist mit möglichen Missverständnissen zu rechnen. Beispielsweise beeindruckt das Effektmaß $d = .49$ für „Lernen in Kleingruppen“ (vgl. Hattie, 2013, S. 113); allerdings sollte der Rezipient dabei beachten, dass sich die betreffende Metaanalyse auf den Hochschulbereich bezieht.

Hattie bezieht in seine Forschungsbilanz nur Studien ein, die in englischer Sprache verfasst wurden. Insofern ist davon auszugehen, dass die meisten Studien in Schulsystemen des anglo-amerikanischen Raums durchgeführt wurden. Deshalb besteht bei der Forschungsbilanz auch ein Interpretationsproblem, inwiefern sich die Ergebnisse auf Schulsysteme in deutschsprachigen Räumen übertragen lassen. Hier ist man sich im Übrigen in Fachkreisen weitgehend einig, dass sich die meisten Befunde, die sich auf Lehr- und Lernprozesse beziehen, auch auf Schulsysteme des deutschsprachigen Raums übertragen ließen.

Durch die verwendeten Effektmaße ermöglichen Metaanalysen Vergleiche über eine Vielzahl von Studien hinweg. Allerdings kann dabei immer nur *eine* Variable im Fokus stehen, beispielsweise die Klassengröße. Wie dabei in großen und kleinen Klassen unterrichtet wird, ob beispielsweise sehr effektiv oder beispielsweise sehr schülerorientiert, wird nicht gleichzeitig mit untersucht. Insofern können bei Metaanalysen Merkmalskonfigurationen und wesentliche Wirkmechanismen nicht erhellt werden. Methodenbedingt ist dies der Preis für die Vergleichbarkeit der Ergebnisse aus ganz verschiedenen Studien. Will man wissen, unter welchen Konfigurationen die Klassengröße eine Rolle spielt, dann können dafür nicht Metaanalysen zu Rate gezogen werden, vielmehr muss man sich auf entsprechende Primärstudien konzentrieren. Methodenbedingt haben Metaanalysen noch weitere Nachteile, insbesondere das hohe Alter der ihnen zugrundeliegenden Primärstudien. Denn bevor eine Metaanalyse durchgeführt werden kann, muss erst eine bestimmte Anzahl an auswertbaren Studien vorliegen. Insofern führen Hatties Meta-Metaanalysen zu einem weiteren Zeitverzug. Zudem sollte methodenbedingt beachtet werden, dass die bei Metaanalysen verwendeten Effektmaße Zusammenhänge zwischen Variablen anzeigen, aber keine Hinweise auf Kausalitäten liefern können. (Denn die entsprechenden Maße zeigen rein rechnerische Zusammenhänge zwischen Variablen auf, ohne dass etwas über deren ursächliche Variationen bekannt ist.)

Bei der Bewertung der Ergebnisse orientiert sich Hattie an den erwähnten Effektmaßen. Dabei geht es ihm in erster Linie um die Ergebnisse jener 138 Faktoren, die besonders hohe Effektstärken erzielen. Mit anderen Worten: Ihn interessiert weniger die Frage: „What works?“, als vielmehr die Frage „What works *best*?“. Deshalb schenkt er in erster Linie nur solchen Einflussgrößen weitere Beachtung, die ein Effektmaß von $d \geq .40$ erzielen. Er nennt diesen Punkt „hinge point“ („Umschlagpunkt“) und erklärt ihn zur ‚Richtschnur‘ pädagogischer Maßnahmen. Er begründet dies damit, dass die durchschnittliche Effektstärke $d = .40$ beträgt: „Dieser Durchschnitt fasst den typischen Effekt aller denkbaren Einflüsse im Bildungsbereich zusammen und sollte als Benchmark oder Referenzmaßstab verwendet werden, um Effekte im Bildungsbereich zu beurteilen.“ (vgl. Hattie, 2013, S. 20) An anderer Stelle weist Hattie darauf hin, dass der Effekt eines Schulbesuchs zwischen $d = .15$ und $d = .40$ pro Schuljahr liegt. „Jeder Einfluss in dieser Zone ist vergleichbar mit dem, was eine Lehrperson in einem typischen Schuljahr erzielen kann.“⁴ Die Zone zwischen $d = 0.0$ und $d = 0.15$ entspricht dem, was Lernende vermutlich erreichen könnten, wenn kein Schulbesuch stattfinden würde [...] Allein schon aufgrund von Reifungsprozessen lässt sich ein Großteil der Lernverbesserungen zurückführen [...].“ (Hattie, 2013, S. 24). Nach John Hattie können deshalb „alle Effekte unterhalb von $d = 0,15$ als potenziell schädlich angesehen werden und sollten nicht umgesetzt werden“ (ebd.).

Vor diesem Hintergrund identifiziert Hattie wirksame und weniger wirkungsmächtige Einflussfaktoren, wobei er sich, wie gesagt, auf die besonders positiven Effekte konzentriert. Von den 138 Einflussfaktoren befinden sich in Hatties Forschungsbilanz 66 Faktoren über dem erwähnten „Umschlagpunkt“ von $d \geq .40$ und 18 Einflussgrö-

4 An anderer Stelle präzisiert Hattie den „typischen Lehrpersoneneffekt“ mit $d = 0,25$ bis $d = 0,40$ (Hattie, 2013, S. 215).

ßen, die unterhalb des genannten Schwellenwertes von $d > .15$ liegen. Für diese Ausrichtung ist laut John Hattie die Beobachtung leitend, dass die pädagogische Literatur eine Vielzahl von Innovationen empfiehlt, jeweils mit der Erwartung, eine Verbesserung des Lernens anzustoßen. All diese Ansätze berichten von positiven Erfahrungen und gehen von der Wirksamkeit dieser Maßnahmen aus. Offenbar ist hierbei fast alles wirksam, so Hattie. Wenn aber so Vieles und so Unterschiedliches positiv wirksam ist, so stellt sich für ihn die Frage nach besonders nachhaltig wirksamen Maßnahmen. Diese können – laut Hattie – jedoch nur durch eine entsprechende empirische Forschung begründet werden. Es bedarf hierzu einer plausiblen Erklärung für die Schlüsseinflüsse eines erfolgreichen Unterrichtens und Lernens – und nicht nur immer neuer Rezepte mit behaupteter Wirksamkeit (Hattie 2013, S. 6).

Mit diesem Anliegen verbindet sich auch der Buchtitel: Der Band „Visible Learning“ ist der Versuch, anhand von Einflussgrößen die Wirksamkeit von Lernprozessen „sichtbar“ zu machen (und zwar im Sinne von erkennbar, thematisierbar, einsehbar, belegbar, einsichtig und verhandelbar) – im Interesse einer Empirie gestützten Bildungsplanung und Schulentwicklung. Erkennbares Lehren und Lernen („visible teaching and learning“) finden – so Hatties Schlussfolgerungen aus den Studienergebnissen – dann statt,

- wenn das aktive Lernen jedes einzelnen Lernenden das explizite Ziel ist,
- wenn es Schülerinnen und Schüler angemessen herausfordert,
- wenn Lehrende und Lernende (auf ihren unterschiedlichen Wegen) überprüfen, ob und auf welchem Niveau die Ziele auch wirklich erreicht werden,
- wenn es eine bewusste Praxis gibt, die auf eine gute Qualität der Zielerreichung gerichtet ist,
- wenn Feedback gegeben und nachgefragt wird und
- wenn – wie sich Hattie nicht zu sagen scheut – aktive, leidenschaftliche und engagierte Menschen am Akt des Lernens teilnehmen (vgl. Hattie, 2013, S. 27).

Für Hattie ist erfolgreiches Lehren und Lernen im Klassenraum auch unmittelbar sichtbar. Je mehr der Lernende dabei selbst zum Lehrenden und der Lehrende zum Lernenden werden, desto erfolgreicher verlaufen – nach Hatties Überzeugung – die jeweiligen Lernprozesse (vgl. S. 26). Für Hattie steht hinter der Forschungsbilanz eine klare „overall message“ (S. 22): „Wie bereits erwähnt, ist sichtbares Lehren und sichtbares Lernen entscheidend“ (Hattie, 2013, S. 45).

2. ‚Hauptlinien‘ der Ergebnisse

In Anbetracht der Ergebnisvielfalt und der umfangreichen, aber dennoch komprimierten Berichterstattung der Forschungsbilanz lassen sich die Ergebnisse eigentlich nicht in knapper Form zusammenfassen. Besondere Aufmerksamkeit müssten darüber hinaus die Folgerungen erfahren, die aus den Forschungsergebnissen zu ziehen wären und die nachhaltige Auswirkungen für Handlungsperspektiven einzunehmen hätten. Im Vorliegenden lassen sich deshalb nur einige ‚Hauptlinien‘ zu den Ergebnis-

sen und zu den Handlungsperspektiven aufzeigen, und zwar solche, die eng mit Unterricht korrespondieren. (Für eine ausführlichere Darstellung der Hattie-Studie wird verwiesen auf Steffens & Höfer, 2016).

2.1 Worauf es besonders ankommt

Betrachtet man die Faktoren über dem „Schwellenwert“ von $d \geq .40$ (moderater Effekt), so zeigen sich gleich auf den ersten Blick zwei Auffälligkeiten. Erstens: Die wichtigsten Faktoren zur Vorhersage von Lernerfolg sind Vorwissen und kognitive Grundfähigkeiten, die in der Regel mit weiteren Faktoren wie sozioökonomischer Status, Anregungsgehalt und Engagement des Elternhauses verknüpft sind. Zweitens: Unter den wirksamen Einflussfaktoren („What works best?“) befinden sich viele Variablen, die den Unterricht und das Lehrerverhalten betreffen, und nur wenige Faktoren, die sich auf strukturelle und organisatorische Maßnahmen beziehen. Dieses Ergebnismuster verweist auf die herausragende Bedeutung der Lehrperson für den Lernerfolg, und zwar sowohl hinsichtlich personaler Merkmale (Einstellungen, Haltungen) als auch hinsichtlich konkreter unterrichtlicher Verhaltensweisen (Unterrichtsskripte).

Dementsprechend ist es nicht überraschend, dass die strukturbezogenen Einflussgrößen überwiegend niedrige Effektwerte erzielen. Klassenwiederholung (Sitzenbleiben) erzielt das Effektmaß $d = -.16$, die Klassengröße ein d von $.21$. Merkmale, die auf eine klassen-, schultyp- oder schulformbezogene Eingruppierung der Schülerinnen und Schüler nach Rasse, Geschlecht, Religion oder Leistungsfähigkeit verweisen, liegen mit niedrigen Effektmaßen im unteren Quartil der Rangreihe und unterhalb von $d = .20$.

Diesen Erkenntnisstand vor Augen könnte man auch von einem *Primat personaler Einflussgrößen vor strukturellen Einflussgrößen* sprechen. Für Hattie sind dabei zwei Aspekte von entscheidender Bedeutung: Erstens wünscht er sich eine Unterrichtsgestaltung *mit den Augen* der Lernenden: „Wenn man die Brille der Lehrperson so ändern kann, dass sie das Lernen mit den Augen der Lernenden sieht, wäre dies schon einmal ein exzellenter Anfang.“ (Hattie, 2013, S. 297f.) Zweitens haben für ihn evaluative Herangehensweisen beim Lehren und Lernen einen zentralen Stellenwert: Alle Informationen, die Auskunft über Lernmöglichkeiten, Lernstand, Lernprozesse und Lernerträge der Schülerinnen und Schüler liefern, sind von besonderem Interesse. Folgende Feedback-Fragestellungen sind dabei für Hattie konstitutiv: „Wohin gehe ich?“ (Lernintentionen/Ziele/Erfolgskriterien), „Wie komme ich voran?“ (Selbstbewertung und Selbsteinschätzung) und „Wohin geht es danach?“ (Fortschreiten, neue Ziele).“ (Hattie, 2013, S. 210) Dabei hat er die Empirie auf seiner Seite: „Formative Evaluation“ und „Feedback“ erzielen in Hatties „Synthese“ der Metaanalysen hohe Effektmaße.⁵

5 In Hatties Forschungsbilanz von 2009 (vgl. Hattie, 2009 bzw. 2013) wird für „Formative Evaluation“ ein d -Wert von $.90$ und für „Feedback“ ein d -Wert von $.73$ genannt. Aufgrund methodenkritischer Hinweise (vgl. u.a. Gold, 2015, S. 108), auch von den Übersetzern der Studie, Wolfgang Beywl und Klaus Zierer, wurde das Effektmaß für „Formative Evaluation“ inzwischen abgeändert: „Diese Effektstärke wird aufgrund von veränderten Zuordnungen der Meta-

Die herausgehobene Bedeutung von Lehrpersonen wird in der Rezeption der Hattie-Befunde gerne zu der Aussage verdichtet „Auf den Lehrer kommt es an!“. Viele Rezipienten assoziieren damit, dass es die Lehrerpersönlichkeit ist, auf die es letztendlich ankommen würde. Das ist jedoch zu einfach. So spielen zwar die pädagogisch-professionellen Einstellungen von Lehrerinnen und Lehrern eine entscheidende Rolle für erfolgreiches Handeln im Beruf, allerdings sind diese professionellen Haltungen nur so viel wert, wie sie im pädagogischen Handeln auch einen praktischen Niederschlag erfahren: Die den Lernenden zugewandten und unterstützenden Haltungen müssen für die Schülerinnen und Schüler sichtbar werden, damit sie Wirkung erzielen können. Es geht also nicht so sehr um die gutgemeinten erzieherischen Absichten von Lehrpersonen, vielmehr geht es darum, wie sie mit den Lernenden arbeiten; es geht um ihren Unterricht.

Dementsprechend formuliert John Hattie: „Die zentrale Botschaft ist einfach: Was Lehrpersonen *tun*, ist wichtig. Dies ist allerdings zu einem Klischee geworden, das die Tatsache verdeckt, dass die größte Quelle für die Varianz in unserem System bei den Lehrpersonen liegt. Lehrerhandeln kann sich in ausschlaggebender Weise voneinander unterscheiden. Es muss hier folgender Zusatz gemacht werden: Was ‚*einige*‘ Lehrpersonen tun, ist wichtig – insbesondere was jene tun, die auf besonders überlegte und sichtbare Weise lehren.“ (Hattie, 2013, S. 27f.; Hervorhebung durch den Autor) Hattie verdeutlicht die großen Unterschiede zwischen guten und „schwachen“ Lehrpersonen mit Bezug auf eine Studie von Sanders und Rivers aus dem Jahr 1996. „(Sie) kommen zu dem Ergebnis, dass die am wenigsten effektiven Lehrpersonen für durchschnittliche Fortschritte der Lernenden in Höhe von rund 14 Prozentpunkten pro Jahr sorgen. Hingegen können die effektivsten Lehrpersonen durchschnittliche Fortschritte in Höhe von 53 Prozentpunkten pro Jahr bewirken.“ (Hattie, 2013, S. 141) Das Bemerkenswerte im Zusammenhang dieser Befunde ist noch Folgendes: „Ineffektive Lehrpersonen sind für alle Lernenden in der Klasse ineffektiv und es gibt ‚wenig Hinweise für einen kompensierenden Effekt durch effektivere Lehrpersonen in späteren Klassen‘ (S. 6). Die Effekte einer schlechten Lehrerqualität halten auch noch Jahre, nachdem Lernende diesen Unterricht besucht haben, an (Sanders & Rivers, 1996).“ (ebd.)

Auch wenn John Hattie immer wieder die Lehrperson in den Fokus seiner Betrachtungen rückt, so ist es letztendlich doch das *unterrichtliche Handeln* von Lehrerinnen und Lehrern, worauf es ankommt. Vor dem Hintergrund der geschilderten Befunde wird verständlich, warum Hilbert Meyer Hatties Formel „Auf den Lehrer kommt es an!“ für zu kurz gegriffen hält. Er präzisiert treffend: „Auf den Unterricht kommt es an!“ (Meyer, 2014, S. 117) Insgesamt gesehen, betont John Hattie die zentrale Bedeutung des Lehrerhandelns und des Unterrichts für den Lernerfolg der Schülerinnen und Schüler. In dieser Perspektive lassen sich aus der Vielzahl seiner Un-

analysen in der nächsten Auflage von John Hatties Visible Learning deutlich nach unten korrigiert werden.“ (lernensichtbarmachen.ch, 2018) Auf persönliche Nachfrage war von Wolfgang Beywl zu erfahren, dass das korrigierte Effektmaß für formative Evaluation nunmehr $d = .48$ beträgt. Auch wenn das Effektmaß für „Feedback“ aufgrund inhaltlich unpassender Zuordnungen ebenfalls deutlich niedriger ausfallen müsste (vgl. Gold, 2015, S. 109, sowie Schulmeister & Loviscach, 2014, S. 128), so dürfte es trotzdem immer noch moderate mittlere Effekte anzeigen. Allerdings liegen bislang keine Hinweise darüber vor, ob John Hattie eine Nachberechnung der Einflussgröße Feedback vornehmen wird.

tersuchungsergebnisse vor allem fünf Felder hervorheben, denen eine entscheidende Bedeutung zuzusprechen ist:

- (1) den Aufbau einer Evaluations- und Feedbackkultur,
- (2) die fachdidaktische Qualität unterrichtlicher Lehr-Lernstrategien in Lernprozessen,
- (3) der Primat der „Direkten Instruktion“,
- (4) ein lernförderliches Klassenklima,
- (5) einen Blick der Lehrpersonen auf das Unterrichtsgeschehen aus der Perspektive der Lernenden.

Im Folgenden werden diese Felder näher erläutert, wobei dabei zwar von den zentralen Befunden ausgegangen wird, allerdings Hatties Folgerungen aus den Befunden in den Mittelpunkt gerückt werden.

2.2 Aufbau einer Evaluations- und Feedback-Kultur

Es sind zwei Einflussfaktoren, die in Hatties Forschungsbilanz eine ganz zentrale Rolle spielen: Feedback und formative Evaluation.

Feedback in Lernprozessen wird verstanden als eine Information über Aspekte von Leistung, und zwar in Ausrichtung auf den Lernstand einer Person, gemessen an einem zu erstrebenden Zielzustand (vgl. Hattie, 2013, S. 206). Im Wesentlichen geht es dabei um eine beständige Überprüfung der Lernfortschritte von Schülerinnen und Schülern. Hattie sieht den Hauptzweck von Feedback darin, „die Diskrepanzen zwischen dem aktuellen Verständnis und der Leistung auf der einen Seite und einer Lernintention oder einem Ziel auf der anderen Seite zu verringern“ (Hattie, 2013, S. 210). Es geht dabei um „das Schließen der Lücke zwischen dem Ort, wo die Lernenden stehen, und dem Ort, wo sie hinwollen, das zur Wirksamkeit des Feedbacks führt“ (a.a.O.).

Hattie verwendet „Feedback“ in einer doppelten Perspektive: Zunächst – und im Schwerpunkt – geht es darum, dass Lehrpersonen ihren Schülerinnen und Schülern regelmäßig Feedback zum Stand ihrer Lernentwicklung geben. Dieses Feedback soll auf die aktuellen Unterrichtsinhalte und die dabei erfolgenden Lernprozesse direkt bezogen sein. Für Hattie ist allerdings das Feedback, das die Lehrerinnen und Lehrer den Lernenden geben, nicht die einzige Form relevanter Rückmeldungen im unterrichtlichen Geschehen. Vielmehr tritt er entschieden dafür ein, auch das Feedback, welches die Lernenden ihren Lehrpersonen geben, als ein wertvolles Instrument der Unterrichtsentwicklung zu nutzen. Bei Hattie steht Feedback in einer schülerorientierten Wirkungsperspektive. Dabei sollen die Lehrpersonen die Sicht der Lernenden einnehmen, indem sie sich empathisch in die Schülerinnen und Schüler hineinversetzen und ihren Unterricht mit den Augen der Kinder und Jugendlichen gestalten.

Beim Feedbackgeben geht es nach Hattie immer um sachbezogene Informationen zum Lernprozess und seinen Resultaten; Sanktionen wie Lob und Tadel sind darunter nicht zu verstehen. Feedback korrespondiert mit formativer Leistungsmessung, bei

der es um die nächsten Schritte im weiteren Unterricht geht (Black & Wiliam, 2009, S. 9; nach Maier, 2010, S. 299). Dem Aufbau einer umfassenden Feedback-Kultur kommt in Hatties Vorstellung von einer guten Schule und einem erfolgreichen Unterrichten eine zentrale Stellung zu. Im Feedback sieht er einen der wirkungsmächtigsten Faktoren des schulischen Handelns überhaupt.

Besonders hohe Wirkungsgrade hat Hattie bei seinen empirischen Untersuchungen für ein lernbegleitendes Feedback der Lehrpersonen an die Schülerinnen und Schüler gefunden (vgl. bereits Hattie & Timperley, 2007). Wichtig erscheint es hierbei, dass die Rückmeldungen regelmäßig und individuell erfolgen. Die Grundlage dazu bilden Informationen im Sinne individueller Lerndiagnosen. Diese erfassen zunächst die Lernausgangslage, später auch wesentliche Schritte im Prozess der Zielerreichung. Dazu muss für die Lernenden ein in seiner Struktur und in seinen Schritten erkennbarer Arbeitsplan mit Erfolgskriterien für die Zielerreichung vorliegen. Das diesbezügliche Feedback soll ebenfalls formativ – also im laufenden Prozess –, ausgerichtet sein und sachliche Hinweise auf den Lernprozess enthalten – also keineswegs bewertend im Sinne von Noten erfolgen.

Die zweite Dimension der angesprochenen Feedback-Kultur mit herausgehobener Bedeutung besteht in der Rückmeldung der Lernenden an ihre Lehrerinnen und Lehrer. Lehrpersonen, die erfolgreich arbeiten wollen, müssen möglichst genau wissen, ob und wie ihr Unterrichtshandeln die Lernenden erreicht. So haben sie bessere Möglichkeiten, zeitnah auf die Lernprozesse und eventuelle Lernschwierigkeiten in ihren Klassen und Kursen zu reagieren, nötigenfalls das Lerntempo zu verändern, innere Differenzierungen vorzunehmen oder Erklärungsvarianten einzubringen und Wiederholungen einzufügen. Lehrerinnen und Lehrer sollen – nach Hattie – für Hinweise, auch für Kritik, offen sein und das Feedback ihrerseits aktiv nachfragen, da es ihnen helfen kann, ihren Unterricht besser an den Bedürfnissen und Notwendigkeiten ihrer Schülerinnen und Schüler zu orientieren.

Feedback steht in einem engen Zusammenhang zur „Formativen Evaluation“. Bei ihr geht es um eine regelmäßige Überprüfung von Lernprozessen und ihrer Wirksamkeit. Bei formativer Evaluation handelt es sich nach Fuchs & Fuchs (1986, S. 200) um eine „datenbasierte Rückmeldung, welche die Lehrperson zu ihrem Unterrichten und ihrem Einfluss auf das Lernen der Schülerinnen und Schüler erhält. Solche Rückmeldungen werden vorzugsweise durch die Lernenden gegeben, beispielsweise über mündliches oder schriftliches Beantworten von auf das Unterrichten bezogenen Fragen (= direkt) oder über Beobachtung von Lernaktivitäten oder Auswertung von Klassentests (= indirekt). Auf dieser Datenbasis bewertet die Lehrperson ihren eigenen Unterricht und entwickelt ihn systematisch weiter, wenn möglich in Zusammenarbeit mit anderen Lehrpersonen.“ (lernensichtbarmachen.ch, 2018)

2.3 Die fachdidaktische Qualität unterrichtlicher Lehr- und Lernprozesse

Unterrichtsanalysen haben nach Hattie häufig gezeigt, dass Lernprozesse vor allem auf die Vermittlung neuen Wissens hin angelegt sind. Er kritisiert diesbezüglich, dass dabei zu oft lediglich die Oberfläche der Sachinformationen („surface“) angesprochen wird, es sich somit auf Seiten der Lernenden lediglich um ein kurzfristiges Auswendiglernen handelt. Notwendig ist es jedoch, zu den diesen Informationen zugrundeliegenden Tiefenstrukturen („deep“) vorzudringen. Dies bedeutet, dass die neu erlernten Fakten nicht nur zutreffend wiedergegeben, sondern auch in ihren Zusammenhängen und Begründungen differenzierend verstanden werden sollen. Ein solches komplexes Verstehen wiederum ist nur möglich, wenn die neuen Lerninhalte an zuvor bereits vorhandene Wissens- und Verstehensbestände angeschlossen und mit schon bestehenden Konzepten – Theorien und Weltbildern – verbunden und in diese integriert werden können („conceptual“). (Vgl. dazu ausführlicher Hattie, 2013, S. 32ff., insbesondere S. 34f.)

Nur wenn es gelingt, die neu erlernten Inhalte in der beschriebenen Weise zu vernetzen, kann die notwendige Nachhaltigkeit des Lernens gesichert werden. Dies wiederum ist eine unverzichtbare Voraussetzung für eine erfolgreiche und flexible Nutzung des Gelernten in unterschiedlichen Praxiszusammenhängen. Auch eine reflektierte Verwendung in Transfersituationen zu einer weiterführenden Problemerkennung und Problemlösung in künftigen Lernfeldern ist hieran gebunden. Nach Hattie kann somit ein erfolgreicher Aufbau relevanter Kompetenzen nur erfolgen, wenn es den Lernenden möglich ist, komplexe Modelle aus Oberflächenwissen, Tiefenverstehen und konzeptueller Vernetzung zu entwickeln und eigenständig zu nutzen.

Insgesamt betrachtet, kommt es auf folgende Kompetenzen von Lehrpersonen an (vgl. Hattie, 2013, S. 238ff.):

- kompetentes Wissen und inhaltliches Verständnis eines Faches besitzen,
- fachlich relevante Konzepte und Prinzipien vermitteln können,
- Lernintentionen und Erfolgskriterien für die einzelnen Lehrsequenzen kennen,
- Bescheid wissen, wie die Intentionen in den einzelnen Sequenzen alle Lernenden erreichen können,
- stärkere Ausrichtung auf die „Tiefenstrukturen“ und die Konstruktion des Wissens statt auf das „Oberflächenwissen“ (vgl. dazu auch Messners Beitrag zur „Tiefendidaktik“ im vorliegenden Band).

2.4 Primat der „Direkten Instruktion“

Nach Hattie hat die „Direkte Instruktion“ zu Unrecht einen schlechten Ruf, obwohl doch alle Forschungsbefunde zeigen, wie wirksam sie ist. Ein Wesenszug der Direkten Instruktion ist die lehrerzentrierte Lenkung des Unterrichtsgeschehens. Die Lehrperson ist in allen Lernprozessen präsent. Ein solcher Unterricht darf jedoch nicht mit einem lehrerzentrierten und fragengeleiteten Frontalunterricht verwechselt werden. Er

ist vielmehr sehr anspruchsvoll und eröffnet den Schülerinnen und Schülern vielfältige Lerngelegenheiten, über deren Nutzung und Nutzen die Lehrperson ‚wacht‘. Sie übernimmt sozusagen die Verantwortung dafür, dass und wie gelernt wird. Die „Direkte Instruktion“ besteht nach Hattie aus sieben Schritten: aus

- klaren Zielsetzungen und Erfolgskriterien, die für die Lernenden transparent sind;
- der aktiven Einbeziehung der Schülerinnen und Schüler in die Lernprozesse;
- einem genauen Verständnis der Lehrperson, wie die Lerninhalte zu vermitteln und zu erklären sind;
- einer permanenten Überprüfung im Unterrichtsprozess, ob die Kinder bzw. Jugendlichen das Gelernte richtig verstanden haben, bevor im Lernprozess weiter vorangegangen wird;
- einem angeleiteten Üben unter der Aufsicht der Lehrperson;
- einer Bilanzierung des Gelernten auf eine für die Lernenden verständliche Weise, bei der die wesentlichen Gedanken bzw. Schlüsselbegriffe in einem größeren Zusammenhang eingebunden werden;
- einer wiederkehrenden, gegebenenfalls modifizierten Anwendung des Gelernten in verschiedenen Kontexten (vgl. Hattie, 2013, S. 243f.).

Aus diesen Schritten lässt sich ein Unterrichtsarrangement herauslesen, das durch Strukturierung, Regelklarheit und Klassenführung, durch eine kognitive Aktivierung mit Blick auf die „Tiefenstrukturen“ beim Lernen, d. h. auf das Verstehen komplexer Zusammenhänge sowie durch evaluative Lehr- und Lernhaltungen gekennzeichnet ist.

Werden die wirksamsten Einflussgrößen der Hattieschen Forschungsbilanz in Augenschein genommen, so ist auffallend, dass sich unter ihnen vor allem solche befinden, die sich auf die schulischen Primärprozesse, also auf die unterrichtlichen Lehr- und Lernstrategien beziehen. Die betreffenden Einflussgrößen reihen sich ein in den aktuellen Erkenntnisstand der Lehr-Lernforschung. So entsprechen diese unterrichtlichen Merkmale den von Eckhard Klieme et al. (2006, S. 131) in einer Forschungsbilanz zusammengestellten „Basisdimensionen“ des Unterrichts:

- „strukturierte, klare und störungspräventive Unterrichtsführung“;
- „unterstützendes, schülerorientiertes Sozialklima“;
- „kognitive Aktivierung“ (z. B. offene Aufgaben, diskursiver Umgang mit Fehlern).

Zu ähnlichen Bilanzierungen kommen auch Mareike Kunter und Jürgen Baumert (siehe dazu Baumert & Kunter, 2006 bzw. Baumert & Kunter, 2011) sowie Lipowsky & Bleck im vorliegenden Band.

Für Hattie steht außer Zweifel, dass die Forschungssynopse auf einen erfolgreichen Unterricht verweist, bei dem die Lehrperson im Zentrum des Geschehens steht und diese die Lernsequenzen initiiert und situiert. Dabei kommt es vor allem auf angeleitete Lernprozesse an, und zwar in Form von gut strukturierten Erklärungen mit Lösungsbeispielen sowie Übungen – angepasst an das Vorwissen der Lernenden. Ein solcher Unterricht der „Direkten Instruktion“ ist – wie den Ergebnissen der Hattie-Studie zu entnehmen – offenen Lernmethoden wie einem entdeckenden, problemorientierten, induktiven, außerschulischen, forschenden oder experimentierenden Lernen

überlegen. Die Ergebnisse zeigen, dass aktiver und von Lehrpersonen gelenkter Unterricht effektiver ist als ein Unterricht, bei dem die Lehrenden als Lernbegleiter und Lernunterstützer nur indirekt in das Geschehen eingreifen (vgl. Hattie, 2013, S. 286). Und er spitzt zu: „only minimal guidance during instruction does not work“ (Hattie, 2009, S. 243).

Um diese Auffassung zu belegen, vergleicht er in einer Tabelle Einflussfaktoren, die für ihn zum einen für eine Lehrperson als „activator“ („Regisseur“) und zum anderen für eine Lehrperson als „facilitator“ („Moderator“) stehen (vgl. Hatties Tabelle 21):

Tabelle 21: Effekte der Lehrperson als Regisseur und Moderator⁶ (Hattie, 2013, S. 287)

Lehrperson als Regisseur („Teacher as activator“)	Effekt- maß d	Lehrperson als Moderator („Teacher as facilitator“)	Effekt- maß d
Reziprokes Lehren	.74	Simulationen und Simulationsspiele	.33
Feedback	.73	Forschendes Lernen	.31
Lautes Denken	.64	Individualisierung	.23
Metakognitive Strategien	.69	Problembasiertes Lernen	.15
Direkte Instruktion	.59	Webbasiertes Lernen	.18
Mastery-Learning	.58	Ganzheits-Methoden	.06
Herausfordernde Ziele	.56	Induktives Vorgehen	.33
Taktung von Leistungstests	.34		
Verhaltensziele / Advance Organizers	.41		
Durchschnitt Lehrperson als Regisseur	.59	Durchschnitt Lehrperson als Moderator	.23

In dieser pointierten Gegenüberstellung erscheinen die Faktoren eines lehrerzentrierten Vorgehens denen eines lehrermoderierten Vorgehens deutlich überlegen. Mit dieser Auflistung bezieht Hattie zugleich eine klare Position zum Verhältnis von Direkter Instruktion und offenen Lernarrangements. Darauf wird in Kap. 3 noch näher eingegangen.

2.5 Lernförderliches Klassenklima

John Hattie geht davon aus, dass der Aufbau und die Entwicklung eines lernförderlichen Klassenklimas zu den zentralen Aufgaben jeder Lehrperson gehören. Hierzu bedarf es bestimmter grundlegender Einstellungen und eines dementsprechenden

6 In der Originalausgabe der Forschungsbilanz sind einige Effektmaße falsch wiedergegeben (genauer gesagt: 10 der 20 Maße; vgl. Hattie, 2009, S. 243, in drei Fällen in erheblichem Umfang. Diese Unstimmigkeiten wurden von den Übersetzern in Absprache mit John Hattie in der deutschsprachigen Ausgabe (Hattie, 2013, S. 287) bereinigt. Ferner wurde in der deutschsprachigen Übersetzung auf eine Zuordnung der ursprünglich dem Typus „Teacher as facilitator“ zugeordneten Variablen „Smaller class sizes“ und „Different teaching for boys and girls“ aus inhaltlichen Gründen verzichtet.

Verhaltens und Handelns. Damit eine solche Atmosphäre entstehen kann, hat er insbesondere drei Erwartungen an die Lehrperson.

Eine positive Beziehung zwischen Lehrende und Lernende ist eine unverzichtbare Grundlage für ein erfolgreiches Lernen. Hattie erwartet deshalb erstens von den Lehrerinnen und Lehrern Wärme, Einfühlungsvermögen und die Bereitschaft zur Ermutigung. Wichtig ist ihm hierbei, dass die Lernenden ein entsprechendes Lehrerverhalten auch tatsächlich wahrnehmen können. Aus diesem Grund gehört es zu den Aufgaben der Lehrpersonen, sich ihrer tatsächlichen Wirkungen durch das Einholen von Feedback zu vergewissern.

Hattie erwartet zweitens, dass die Lehrpersonen ihre Schülerinnen und Schüler durchgehend mit Respekt behandeln; dieser Respekt darf nicht an Bedingungen geknüpft sein. Er ist vielmehr die immer geltende Grundlage eines angemessenen Lehrerverhaltens, auch und gerade dann, wenn die Lernenden Fehler machen, Unzulänglichkeiten erkennen lassen oder selbst ein unangemessenes Verhalten zeigen. Indem die Lehrperson allen ihren Schülerinnen und Schülern vermittelt, dass sie sie für wertvolle Persönlichkeiten hält, legt sie die Grundlage für ein respektvolles Miteinander.

Die dritte Erwartung, die Hattie an Lehrpersonen richtet, ist die eines grundlegenden Vertrauens in die Fähigkeiten ihrer Schülerinnen und Schüler. Die Lernenden sollen deutlich erkennen können, dass die Lehrpersonen ihnen zutrauen, konstruktiv miteinander kooperieren, anspruchsvolle Aufgaben lösen und hohe Erwartungen erfüllen zu können. Der hierbei zugrundeliegende Optimismus soll für die Lernenden klar erkennbar sein.

2.6 Lehrerhandeln aus der Perspektive der Lernenden

Lehrpersonen sind nach Hattie lernfähige Expertinnen und Experten („adaptive experts“), die wissen, wo ihre Schülerinnen und Schüler in ihrer Lernentwicklung stehen. Sie wissen, ob sie gerade lernen oder nicht und wohin die Lernwege als Nächstes führen sollen. Lehrerinnen und Lehrer wirken zudem darauf hin, ein Klassenklima zu erzeugen, das erfolgreiche Lernprozesse fördert und unterstützt. Als lernfähige Experten verfügen sie über ein hohes Maß an Empathie und sehen die Lernprozesse auch aus der Perspektive der Lernenden. Hattie betont, dass es ihm nicht um routiniertes Expertenwissen geht. Vielmehr fordert er eine pädagogische Expertise, die sich an die jeweilige Handlungssituation anpasst („adaptive expertise“).

Lehrpersonen sind in der Lage, vielfältige Wege des Verstehens und differenzierte Formen des Interagierens zu vermitteln, wobei sie zugleich unterschiedliche Gelegenheiten zum Handeln eröffnen. Eine „kognitive Flexibilität“ der Lehrerinnen und Lehrer ist bei der Ausgestaltung unterschiedlicher Verstehenswege in hohem Maß erforderlich. Sie stellt ein wesentliches Kriterium für guten Unterricht dar und nimmt mit der Anzahl der gewählten Perspektiven, Fakten, zu berücksichtigenden Fähigkeiten, Vorgehensweisen und tiefgreifenden konzeptuellen Prinzipien zu.

Bei den Interaktionswegen ist zu beachten, dass die Lernenden auch darin unterrichtet werden, wie Lernprozesse strukturiert sind und erfolgreich vorangebracht

werden können. Komplexes Denken und Argumentieren wird im Unterricht beispielsweise dadurch gefördert, dass durch herausfordernde Problemvorgaben eine Verunsicherung über das bisher Gelernte bei den Schülerinnen und Schülern ausgelöst wird. Das bewusste Hervorrufen von Hindernissen, Widersprüchen und Konflikten sollte als grundlegendes Gestaltungsprinzip von Lernprozessen gesehen und dementsprechend von den Lehrpersonen thematisiert werden.

John Hattie empfiehlt Lehrerinnen und Lehrern, die Lernprozesse in ihrer Klasse stets aus der Perspektive der Lernenden zu sehen („to see learning through the eyes of students“). Dabei gilt es, in einem dialogischen Prozess sicherzustellen, dass auch die Lernenden ihre Lehrpersonen mit all jenen Informationen versorgen, die diese benötigen, um den Unterricht sachgerecht planen und durchführen zu können.

2.7 Fazit: Kernpunkte professionellen Handelns von Lehrerinnen und Lehrern

- Das konkrete Handeln der Lehrpersonen ist entscheidend für den Erfolg des Unterrichts und den individuellen Lernerfolg der Schülerinnen und Schüler.
- Lehrerinnen und Lehrer sind aktive Gestalter unterrichtlicher Prozesse; wesentlich für den Erfolg sind die didaktische Qualität des Unterrichts und der Aufbau eines lernförderlichen Klassenklimas.
- Lernen ist ein individueller Prozess der Schülerinnen und Schüler; dessen Stand und Verlauf müssen Lehrpersonen möglichst genau kennen.
- Lehrerinnen und Lehrer müssen lernen, den Unterricht „mit den Augen der Lernenden“ zu sehen und ihn entsprechend zu gestalten; dazu bedarf es der Entwicklung einer umfassenden Feedback-Kultur.
- Alle Formen des Unterrichts und alle Programme der Unterrichtsentwicklung müssen sich an ihrer Wirksamkeit messen lassen.
- Um diese Ziele erreichen zu können, bedarf es einer sachgerechten Lehrerausbildung und einer praxisnahen Lehrerfortbildung.

Diese Perspektiven erfordern eine wirksame Professionalisierung des Lehrpersonals. Die Forschungsbilanz verweist auf den außerordentlichen Ertrag kompetenten Personals (s.o.) und zeigt, wie wirksam Lehrerfortbildung dabei sein kann ($d = .62$).

Aus Platzgründen musste eine Auswahl getroffen werden, wobei vor allem solche Ergebnisse im Mittelpunkt standen, die sich auf den *Unterricht* beziehen und denen Hattie selbst immer wieder eine herausgehobene Bedeutung zugesprochen hat.

Gerade mit Blick auf praktische Konsequenzen würde es sich lohnen, weitere unterrichtsbezogene Ergebnisse zur Kenntnis zu nehmen, z.B. das Potenzial der Gleichaltrigen für Lernprozesse bzw. Befunde zum kooperativen Lernen ($d = .59$), die hohe Wirksamkeit curricularer Programme (z.B. Leseförderprogramme: $d = .67$), die jedoch in deutschsprachigen Schulsystemen bislang kaum eine Anwendung finden, ferner vorschulische Maßnahmen ($d = .45$), gar Erlebnispädagogik ($d = .52$) oder die relative Bedeutung von computergestütztem Unterricht ($d = .37$).

Auch Faktoren, die die Schulsystemsteuerung betreffen, würden eine gesonderte Abhandlung und Auseinandersetzung lohnen, beispielsweise die mäßigen Effekte für schulweite Reformvorgaben „von oben“ ($d = .22$) oder die Ergebnisse zu Privatschulen in verschiedenen Varianten ($d = .23$ bzw. $.20$). Auch zu Schulstrukturfragen liefert die Forschungsbilanz aufschlussreiche Ergebnisse, sei es im Hinblick auf Schullaufbahnen (Sitzenbleiben: $d = -.16$; Schulwechsel: $d = -.34$)⁷ oder zur Einteilung von Schülerinnen und Schülern in leistungshomogene Klassen („ability grouping“ bzw. „tracking“, in Deutschland mit Schulformen vergleichbar, $d = .12$). Große Diskussionen haben in Hatties Forschungsbilanz immer wieder die Ergebnisse zur Klassengröße ($d = .21$) ausgelöst, die leicht falsch interpretiert werden könnten.⁸

3. Hatties „Visible Learning“ in der Diskussion

3.1 Zur Resonanz der Studie

Noch nie hat eine Untersuchung aus der Lehr- und Lernforschung eine solche breite Resonanz ausgelöst, mit der selbst die Diskussionen um die PISA-Studie kaum mithalten können. So schaffte es die Hattie-Veröffentlichung kurzzeitig sogar, in den „Top Ten“ der Spiegel-Sachbuchliste geführt zu werden (vgl. Beywl, 2015, S. 14) – einmalig für eine Studie aus der Bildungsforschung. Von Kritikern wird weitgehend übereinstimmend der besondere Wert von Hatties Forschungsbilanz in der einmaligen Übersicht gesehen, selbst dann, wenn man mit den Ergebnissen im Einzelnen nicht übereinstimmen mag oder von der Art der Ergebnisaufbereitung nicht überzeugt zu sein scheint. Denn noch nie wurden Forschungsergebnisse zu den Lernleistungen von Schülerinnen und Schülern in diesem Umfang bilanziert. Offenbar sind viele Rezipienten allein schon von den Zahlen beeindruckt: 138 Einflussfaktoren, 816 Metaanalysen, 52.649 Studien, 146.626 Effekte, 83.033.433 Untersuchungsbeteiligte (vgl. Hat-

7 Die Minuszeichen sind so zu lesen, dass die Zusammenhänge negativ ausfallen; bei Nichtversetzungen bedeutet dies, dass die Fachleistungen bei Klassenwiederholern schlechter ausfallen als bei vergleichbaren Lernenden, die nicht wiederholen. Zu beachten ist hier, dass es in Einzelfällen sehr sinnvoll sein kann, dass ein Schüler bzw. eine Schülerin eine Klasse wiederholt (z.B. bei längerfristiger Erkrankung); die Ergebnisse sind – wie bei allen Effektmaßen – Durchschnittswerte, die natürlich Einzelfällen nicht gerecht werden können.

8 Bei einer Rezeption wird unwillkürlich unterstellt, dass mit kleinen Klassen in systematischer Weise ein anderer Unterricht einhergeht, etwa durch mehr Aufmerksamkeit für den Einzelnen bzw. durch mehr Schülerorientierung oder durch effektivere Methoden. Das muss aber in den zugrundeliegenden Studien gar nicht der Fall sein. Oftmals handelt es sich bei solchen Untersuchungen nicht um eine systematische Variation unterrichtlicher Settings (also um guten Unterricht in kleinen Klassen), vielmehr können die Gründe für die in einer Studie gegebenen kleinen und großen Klassen ganz unterschiedlich sein. Sofern es sich bei der Klassengröße nicht um eine pädagogisch-konzeptionelle Maßnahme handelt, ist davon auszugehen, dass Lehrpersonen in kleinen Klassen ihr unterrichtliches Vorgehen beibehalten und den gleichen Unterricht durchführen wie in großen Klassen. Die entsprechenden Verhaltensweisen (wie lehrerzentrierter Frontalunterricht) sind offenbar dermaßen stark ausgeprägt, dass es schwierig ist, selbst unter günstigen Umständen andere Wege zu beschreiten.

tie, 2013, S. 22)⁹ – so etwas gab es noch nie. Insofern könnte damit die Wahrnehmung verbunden sein, dass solche Ergebnisse nicht einfach ignoriert werden können.

Seit Hatties Studie werden sowohl in Wissenschaft als auch Praxis Stellungnahmen zu Schule und Unterricht oftmals mit Bezug auf die Forschungsbilanz begründet. Es ist deshalb zu vermuten, dass mit dieser Studie auch in Praxiskreisen die Bedeutung empirischer Schul- und Unterrichtsforschung inzwischen an erheblicher Akzeptanz gewonnen hat (vgl. Steffens & Höfer, 2016, Kap. 7.3). Immer mehr scheint die Einsicht Raum zu ergreifen, dass empirische Forschung auch für pädagogische Arbeitszusammenhänge hilfreich sein kann und dass sich selbst pädagogische Prozesse messen lassen und etwas mit Wirksamkeit zu tun haben. Vielleicht ist das sogar der wichtigste Effekt dieser Studie.

Ewald Terhart, einer der ersten deutschen Erziehungswissenschaftler, der sich mit der Forschungsbilanz von Hattie eingehend befasst hat (vgl. Terhart, 2011; s.a. Terhart, 2014), begreift die Hattie-Studie als „einen Meilenstein in der Debatte um die Voraussetzungen und Bedingungen erfolgreichen Lernens in der Schule“ (Terhart, 2014, S. 10). Er würdigt Hatties Gesamtschau von Forschungsbefunden als „einzigartige Leistung“ (S. 11). Das besondere Verdienst sieht er darin, dass Hattie die Flut an Effektkräften in „prägnanten Quintessenzen“ zu bündeln vermag und damit eine „wahre Fundgrube“ für interessierte Forschende bietet sowie „wichtige empirische Resultate und weiterführende Anregungen“ vermittelt (S. 20). Auch Andreas Gold betrachtet Hatties Analysen als „Meilensteine der empirischen Schul- und Unterrichtsforschung“ und stellt fest, dass seit Erscheinen der Hattie-Studie „intensiv wie schon lange nicht mehr über die Bedeutsamkeit wissenschaftlicher Erkenntnisse für die Unterrichtspraxis und die Lehrerbildung diskutiert (wird)“ (Gold, 2014, S. 332). Auch Andreas Helmke findet keine treffendere auszeichnende Charakterisierung als „Meilenstein“. Er sieht den Vorteil solcher Forschungsbilanzen darin, dass sie „den gesamten empirischen Forschungsstand zu einer Fragestellung übersichtlich zusammen(fassen)“ und dabei einer „hermeneutischen, notwendigerweise stark subjektiven Berichterstattung klar überlegen“ sind (S. 11).

Olaf Köller, Johanna Möller und Jens Möller sehen als größtes Verdienst der Studie, dass sie „die Schulwirklichkeitsforschung enorm belebt hat“ (Köller, Möller & Möller, 2013, S. 46) und dass nunmehr die Befunde der empirischen Bildungsforschung in der Praxis mehr wahrgenommen werden (S. 47). Bei der Würdigung der Forschungsarbeit heben sie Hatties „Machbarkeitsvorstellungen“ hervor: „Sein Credo ist, dass Schule Vieles schaffen kann [...]“ (S. 46). Schule ist folglich dann wirksam, wenn Lehrpersonen guten Unterricht machen. Hatties Anliegen ist es dabei, einen solchen Unterricht mit Bezugnahme auf die Forschungsergebnisse zu beschreiben.

9 Zur Anzahl der berücksichtigten Metaanalysen siehe auch Fußnote 1. Auch über die weiteren Zahlen werden von Hattie unterschiedliche Angaben gemacht, z.B. auf S. 18: Berücksichtigte Studien: 52.637, festgestellte Effekte: 146.142, Anzahl Lernende: ca. 236 Millionen (geschätzt). Die große Differenz bei den Lernenden (zwischen 83 und 236 Millionen) erklärt sich dadurch, dass sich die Angaben auf S. 22 auf ermittelbare Zahlen auf der Basis von 286 Metaanalysen beziehen; da zu den übrigen Analysen keine Zahlen bekannt sind, wurden sie auf der Grundlage vorliegender Daten auf 236.000.000 geschätzt.

Möglicherweise ist die positive Resonanz, die die Studie sowohl in Wissenschafts- als auch Praxiskreisen ausgelöst hat, auch darauf zurückzuführen, dass die Befunde – neben ihrer außergewöhnlichen Bilanz – zugleich in ungewöhnlicher Eintracht einerseits der empirischen Forschung in pädagogischen Handlungszusammenhängen zu einer größeren Verbreitung verholfen haben und andererseits „viele altbekannte Ansätze in der Pädagogik bestätigen“ (Nix & Wollmann, 2015, S. 138).

Dennoch scheiden sich an der Hattie-Studie die Geister. Während die einen von der Ergebnisfülle und Differenziertheit der Befunde oder von den klaren Ergebnislinien mit eindeutigen Effektmaßen und Rangreihen und darauf bezugnehmenden prägnanten Aussagen beeindruckt sind, betonen die anderen den fehlenden seriösen und differenzierenden Umgang mit den Daten sowie die isolierte Darstellung einzelner Variablen und die fehlenden Zusammenhangsanalysen – insbesondere in Anbetracht des Wissens darüber, dass guter Unterricht eine wohlkonfigurierte Prozessgestaltung durch verschiedene Komponenten (Faktorenbündel) voraussetzt. Der kritische Leser bzw. die kritische Leserin gewinnt den Eindruck, dass Hattie die Ergebnisse aus den Metaanalysen ganz bewusst pointiert-prägnant darstellt, in dem er die Effektmaße zur nahezu alleinigen Richtschnur erklärt und auf dieser Grundlage Rangplätze verteilt. Die damit verbundenen apodiktischen Aussagen – im Sinne von „Was hilft richtig?“ „Was schadet wirklich?“ – wirken dann doch in ihrer verbreiteten Anwendung effekt-hascherisch.¹⁰

Insofern verwundert es nicht, dass die Übersetzer der Hattie-Studie ins Deutsche, Wolfgang Beywl und Klaus Zierer, eine verkürzte, an den Rangplätzen orientierte Lesart als „Fast-Food-Hattie“ wahrnehmen (Beywl & Zierer, 2013, S. X), die der Komplexität der Materie nicht gerecht wird. Deshalb ist es nachvollziehbar, dass Wolfgang Beywl und Klaus Zierer in dem großen Wert der Studie gleichzeitig auch ihre größte Gefahr sehen: „Der Versuch, Komplexität zu reduzieren.“ (Beywl & Zierer, 2014, S. 155) Dazu haben auch die Ranglisten der Einflussfaktoren beigetragen, wie sie mit Verweis auf Hatties eigene Worte ausführen: Die Ranglisten seien „eine der größten Quellen für die Ablenkung von der eigentlichen Aufgabe (vgl. Zeit-Interview mit John Hattie, 02.05.2013): der Suche nach besseren, nach optimalen Kombinationen von Einflussmöglichkeiten [...]“ (ebd., S. 161). An dieser Stelle ist ein Hattie zu erkennen, dem die Resonanz auf den von ihm selbst initiierten „Hitlistenfetisch“ nicht mehr ganz geheuer ist. Demgegenüber scheint eine ‚Lesart‘ von Wolfgang Beywl und Klaus Zierer ein erfreulich angemessener Umgang mit den Forschungsergebnissen zu sein: „Vielfach liegt der Schlüssel für den Erfolg nicht in der Auswahl weniger ‚starker‘ Faktoren, sondern in einer wohlüberlegten, auf den jeweiligen schulischen Kontext abgestimmten Kombination, auch mit schwächeren Faktoren“ (Beywl & Zierer, 2014, S. 154). (Für eine ausführlichere Auseinandersetzung mit der Hattie-Studie wird verwiesen auf Steffens & Höfer, 2016).

¹⁰ Aufgrund vielfacher Kritik verzichtet John Hattie inzwischen bei der Weiterführung seiner Arbeiten auf sein umstrittenes Ranking der Einflussfaktoren (vgl. Beywl & Strasser, 2017).

3.2 Hatties Vorgehen in der Kritik

Zum methodischen Vorgehen Hatties

Ein weiteres Manko in Hatties Berichterstattung über die Forschungsbefunde besteht darin, dass die 138 *Einflussfaktoren im Regelfall nicht definiert* und erläutert werden. Das begünstigt immer dann Missverständnisse, wenn es für eine bestimmte Bezeichnung verbreitete Bedeutungshöfe gibt, die jedoch mit dem, was Hattie darunter versteht, wenig zu tun haben. Ein Beispiel dafür ist der Faktor „Individualisierung“ (im Original: „Individualized Instruction“) mit dem Effektmaß $d = .22$. Hierunter sind auch Metaanalysen zum programmierten Unterricht subsumiert, die insofern nur teilweise das erfassen, was unter individueller Förderung gerade im deutschsprachigen Raum verstanden wird. Ein anderes Beispiel ist offener Unterricht (im Original: „Open vs. Traditional“) ($d = .01$), der auch Metaanalysen zur „Open education“ berücksichtigt. Bei diesem Konzept werden mehrere Lerngruppen bzw. Klassen in einem Großraum unterrichtet, wie es in Deutschland wohl nur noch an der Laborschule Bielefeld üblich ist. In der deutschen Übersetzung wurde deshalb der Faktor mit „Offene Klassenzimmer / Lehr- und Lernformen“ bezeichnet. Um das geschilderte Problem zu entschärfen, besteht in Verbindung mit den deutschsprachigen Hattie-Übersetzungen ein Glossar auf der Homepage www.lernensichtbarmachen.ch. Dieses Portal bietet zahlreiche Hilfen für eine Auseinandersetzung mit Hatties Forschungsarbeiten und für entsprechende praktische Konsequenzen an.¹¹

Sehr umstritten in Hatties Forschungsbilanz ist seine Eingrenzung der Metaanalysen auf solche, in denen es um die *Leistungen von Schülerinnen und Schülern im fachlichen Bereich* geht. Hans Günter Rolff kritisiert an der Hattie-Studie u.a. diese einseitige Ausrichtung auf den Wissenserwerb in einigen Hauptfächern. Sehr wichtige Felder, beispielsweise der Aufbau von Sozialkompetenzen und moralischen Orientierungen, geraten dabei nicht in den Blick, so Rolff. Und so ergibt sich ein Gesamtbild, das Hans Günter Rolff zu der zugespitzten Aussage führt: „Bildung ist kein Thema, bestenfalls ein sehr einseitiges Verständnis davon“ (Rolff 2013, S. 49). Und das ist, gerade in Deutschland – anders als im angelsächsischen Raum –, sicher ein gravierender Einwand gegen die von John Hattie praktizierte Form einer *Bildungsforschung*.

11 Dies betrifft insbesondere eine Darstellung der untersuchten Einflussgrößen und ihre jeweiligen Definitionen, ein Glossar zu zentralen Fachbegriffen, die Möglichkeit, Fragen im Zusammenhang der Hattie-Studie zu stellen und beantwortet zu bekommen, eine Sammlung an Antworten auf relevante und häufig gestellte Fragen, eine Auseinandersetzung mit herausgehobenen Themen- und Problemstellungen zu Hatties Forschungsarbeiten. Darüber hinaus werden von dem Autorenteam seit 2012 jährlich mehrere „Newsletter“ herausgegeben, die sich mit zentralen Aspekten der Hattieschen Konzeption und ihrer Realisierung befassen. Von besonderer Relevanz in dem Portal sind die „Praxisberichte“ von Lehrpersonen über ihre Erfahrungen mit praktischen Maßnahmen im Anschluss an Hattie.

Das Besondere an Metaanalysen und Hatties Umgang damit

Eine Forschungsbilanz der geschilderten Größenordnung ist, wie bereits ausgeführt wurde, nicht in Form einer inhaltsanalytischen Auswertung möglich, sondern nur mit Hilfe eines statistischen Verfahrens, bei dem sich über verschiedene Studien hinweg die Daten erfassen und vergleichen lassen. In den betreffenden Metaanalysen werden Ergebnisse zu einem Einflussfaktor, z.B. zur Leseverständnisförderung oder zum Sitzenbleiben, über viele diesbezügliche Studien hinweg zu einem einzigen durchschnittlichen Gesamtwert ‚verdichtet‘. Auf diese Weise bieten Metaanalysen die Möglichkeit, *auf einen Blick* zu erfahren, was sich über viele Studien hinweg immer wieder als wirksam herausgestellt hat und für die große Masse Durchschlagskraft besitzt. Sie zeigen damit die großen Linien bzw. Trends über viele einzelne Studien hinweg. Hans Brügelmann schreibt dazu: „Die Stärke von statistischen Metaanalysen ist zugleich ihre Schwäche – wie beim Einsatz des Zoomens in der Fotografie, also der Entscheidung für Wald oder Bäume: Entweder gewinnt man einen guten Überblick, vernachlässigt dafür aber bedeutsame Details – oder es gelingt, einzelne Aspekte genauer in den Blick zu nehmen.“ (Brügelmann, 2014, S. 40) Und an anderer Stelle schreibt er: „Hatties Überblick ist ähnlich hilfreich wie die Weitwinkelaufnahme einer Stadt aus der Luft – für Planer, die die Infrastruktur einer Stadt gestalten wollen. Wer dagegen ein einzelnes Haus bauen oder kaufen will, braucht Detailinformationen, die in der Vogelperspektive verlorengehen.“ (Brügelmann, 2014, S. 48) Dieses Bild beschreibt sehr schön die Grenzen und Möglichkeiten von Metaanalysen.

Die Ermittlung von Durchschnittswerten über viele Studien hinweg führt allerdings zu einer *Nivellierung von Ergebnisdifferenzen*. Gerade aber solche Differenzen könnten von besonderem Forschungsinteresse sein. So würde man im vorliegenden Beispiel der Leseförderung danach fragen, was Programme auszeichnet, die besonders hohe bzw. besonders geringe Effektmaße erzielen. Gerade für die Bildungsplanung und Schulentwicklung dürfte es darum gehen, in Erfahrung zu bringen, welche Programme besonders wirksam sind und welche Programmkomponenten im Einzelnen dazu besonders beitragen.

Diese fehlende notwendige Differenzierung der Befunde kritisiert auch Andreas Gold an der Hattie-Studie: „Wo es vornehmlich auf öffentlichkeitswirksame Tops und Flops und auf das Erreichen des magischen Umschlagpunkts ($d > 0.40$) einer Effektstärke ankommt, wird nämlich glattgebügelt, was wir über Bedingungen (Moderatoren) der Wirksamkeit von Einflussfaktoren wissen [...]“ (Gold, 2014, S. 333) Und er führt weiter aus: Dabei können wir aus den Studien von Andreas Helmke, Eckhard Klieme, Jürgen Baumert und Mareike Kunter sehr viel mehr über guten Unterricht erfahren als aus Hatties Rangliste.“ (S. 332). Diese Liste würde mehr verdecken, als sie enthüllt. So vermisst er auch „einen seriösen Umgang mit uneinheitlichen Studienergebnissen“ (S. 333) und merkt an, „dass Aussagen über die Wirksamkeit von Einzelfaktoren stets etwas ‚Künstliches‘ anhaftet, weil es in der Realität stets eines Zusammenspiels der wichtigsten Qualitätsdimensionen des Unterrichts bedarf, um die intendierten Lernprozesse auszulösen und zu befördern“ (Gold 2015, S. 165).

Nach all den Ausführungen ist offensichtlich, dass Metaanalysen – neben ihren Möglichkeiten – auch mit einigen methodenbedingten Einschränkungen behaftet sind (vgl. dazu auch Terhart, 2011; Brügelmann, 2014; Rolff, 2014; für eine ausführlichere Methodenkritik vgl. Pant, 2014; Schulmeister & Loviscach, 2014 und Steffens & Höfer, 2016, Kap. 6).

Neben diesen allgemeinen Einschränkungen bezüglich Metaanalysen muss auch auf Besonderheiten in *Hatties Umgang mit seinen Auswertungen von Metaanalysen* hingewiesen werden. So nennt Hattie keine Auswahlkriterien bzw. Qualitätsstandards bei den von ihm berücksichtigten Metaanalysen – was bei Auswertungen im Zusammenhang mit Metaanalysen unüblich ist. Ferner bewertet er in seinen Bilanzierungen alle Metaanalysen gleichgewichtig, unbeschadet der Tatsache, wie viele Primärstudien einer Metaanalyse zugrunde liegen. Auch unterscheidet er in seinen Ergebniszusammenstellungen nicht zwischen unterschiedlichen Adressaten von Bildungsmaßnahmen; vielmehr werden die Ergebnisse von Kleinkindern, Jugendlichen und Erwachsenen undifferenziert zu einem Gesamtwert verrechnet. Darüber hinaus ordnet er die Ergebnisse für die 138 Einflussfaktoren nach sechs Untersuchungsbereichen (wie weiter oben benannt) und bildet für diese Bereiche nochmals Mittelwerte der Mittelwerte von Metaanalysen und interpretiert diese Ergebnisse.

Unterrichts- und schultheoretischer Stellenwert in Hatties Forschungsbilanz

Andreas Gold thematisiert diese Art und Weise der Ergebnisdarstellung und kritisiert dabei das „ungeordnete Nebeneinander von Lehrmethoden, unterrichtlichen Sozialformen, inhaltlichen Förderprogrammen, didaktischen Prinzipien und persönlichen Merkmalen der Lehrenden“ (Gold, 2014, S. 334). Hierbei drängt sich der Eindruck einer *fehlenden analytischen Ordnungsstruktur* auf. In diese Richtung geht auch Hilbert Meyers Kritik an John Hattie: So würdigt er zwar Hatties empirische Basis einerseits (s.o.), kritisiert aber andererseits die fehlende theoretische Unterfütterung: „In meinen Augen ist er ein pragmatisch orientierter Reformpädagoge, dessen Stärke darin besteht, dass er das, was er normativ einklagt, empirisch absichern kann, und dessen Schwäche darin besteht, dass sein bildungstheoretischer Überbau nur angedeutet wird.“ (Meyer, 2014, S. 118) Insofern ist es konsequent, wenn Meyer an anderer Stelle kritisiert: „Es fehlt an einer allgemein- und ebenso an einer fachdidaktischen Rahmung [...]“ (S. 128). Und er fährt fort: „Mit ein wenig ‚mehr Klafki‘ hätte Hattie dieses Defizit beheben können. Dann hätte der im ‚Visible Learning for Teachers‘ (2012, S. 35ff.) enthaltene und einseitig auf die Zielanalyse fixierte ‚Leitfaden‘ zur Vorbereitung, Durchführung und Auswertung von Unterrichtsstunden wohl deutlich anders ausgesehen. Selbst die schlichten Fragen aus Klafkis Didaktischer Analyse von 1959 hätten gut in seinen Fragenkatalog hineingepasst: Was ist die Struktur des vorgesehenen Unterrichtsinhalts? Wofür steht er exemplarisch? Hat er für die Schülerinnen und Schüler einen Gegenwarts- und Zukunftsbezug?“ (Meyer, 2014, S. 128). Meyer würdigt folglich Hattie als Empiriker, aber bemängelt ihn als Theoretiker: „Jeder Versuch,

eine Didaktik nur auf Grundlage der Analyse von Unterrichtseffekten zu zimmern, muss scheitern. Das sieht man an den Defiziten von Visible Learning.“ (S. 128)

Insofern ist auch Hatties Ordnungsstruktur der eingangs dargelegten sechs Untersuchungsbereiche (Elternhaus, Lernende, Schule, Curriculum, Lehrende, Unterricht) zu hinterfragen. Ein diesbezügliches theoretisches Referenzsystem ist dabei nicht zu erkennen. Solange hierfür kein Begründungszusammenhang geliefert wird, besteht der Verdacht eines vorrangig empiristischen Vorgehens. Insofern ist zu konstatieren, dass Hatties Forschungsbilanz statt einer pragmatischen Ordnungsstruktur mehr Theorie der Schule verdient gehabt hätte. Im gleichen Maße, wie sich Hilbert Meyer bei Hattie „mehr Klafki“ gewünscht hätte (s.o.), wäre im vorliegenden Zusammenhang auch „mehr Fend“ wünschenswert gewesen.

Hatties pragmatisches Vorgehen kommt auch in der Untergliederung des Untersuchungsbereichs „das Unterrichten“ zum Ausdruck, den er in zwei Teile aufteilt, und zwar in „Die Beiträge des Unterrichtens – Teil I“ und in „Die Beiträge des Unterrichtens – Teil II“. Dazu erläutert er lapidar: „Um die Behandlung der verschiedenen Unterrichtsansätze in einem vernünftigen Umfang zu halten, werden diese Beiträge, etwas willkürlich, auf zwei Kapitel aufgeteilt.“ (Hattie 2013, S. 192) Den Umfang als Grund anzugeben, ist insofern nicht nachvollziehbar, weil es im Bereich Unterrichten insgesamt dennoch bei derselben Anzahl an Einflussfaktoren und dem gleichen Textumfang bleibt. Im Übrigen führt Hattie bei seinen Bilanzierungen beide Teile zusammen und vergleicht den Bereich Unterrichten mit den anderen Untersuchungsbereichen. Dabei ermittelt er Durchschnittswerte für die Anzahl an Metaanalysen und berücksichtigten Studien, für die Anzahl der Probanden und ermittelten Effekte sowie für die Effektmaße des Untersuchungsbereichs Unterrichten. Insofern macht die von ihm vorgenommene pragmatische Aufteilung in zwei Teile keinen Sinn. Auch hier wird deshalb ein theorieorientiertes Vorgehen vermisst.¹²

So gesehen, scheint Hatties Werk vorrangig aus zwei Herangehensweisen zu bestehen – einer empirischen Forschungsbilanz zum einen und einer Art pragmatischen „Erziehungslehre“ zum anderen. Dementsprechend monieren Frank Nix und Jens Wollmann: „Hattie geht von einem didaktischen Modell aus, das er mit seiner Studie zu belegen versucht. Die Ergebnisse werden zielgerichtet ausgewählt und präsentiert, um dieses Modell zu stützen. Dies muss bei der Interpretation der Ergebnisse berücksichtigt werden.“ (Nix & Wollmann, 2015, S. 139) Insofern ist Hatties Werk „Visible Learning“ in seiner Darstellung zwiespältig angelegt:

Einerseits ist es gekennzeichnet durch eine recht objektive, geradezu ‚nüchterne‘ Beschreibung der Ergebnisse zu den 138 untersuchten Einflussgrößen, so dass es dem Leser bzw. der Leserin nicht immer leicht fällt, dem ‚trockenen‘ Text zu fol-

12 Inzwischen hat John Hattie seine sechs Untersuchungsbereiche („Domänen“ genannt) auf neun ausgeweitet. Und zwar wurden die klassenbezogenen Faktoren aus dem Untersuchungsbereich „Schule“ herausgenommen und werden jetzt in einem eigenen Bereich „Klasse“ geführt. Eine Auffächerung erfuhr auch der Untersuchungsbereich „Unterricht“. Die darin versammelten Faktoren werden nunmehr unterteilt in „Lernstrategien“, „Lehrstrategien“ und „Lehr-Lernarrangements“. (Zu den Aktualisierungen vgl. Beywl & Strasser, 2017) Trotz dieser Fortschritte in Hatties Forschungsbilanzierungen schreiben Beywl & Strasser (2017): „Die in zahlreichen Veröffentlichungen und teils schon in den Vorworten zu den deutschsprachigen Ausgaben genannten Schwächen bestehen wohl zumindest teilweise fort.“

gen. So wird ein Forschungsdetail ans nächste geknüpft und noch eine Metaanalyse zitiert und auf ein weiteres Ergebnismuster verwiesen, so dass man den Eindruck gewinnt, die Ergebnisse aus über 50.000 Studien bzw. über 700 Metaanalysen so objektiv wie möglich referiert zu bekommen. Andererseits sind Hatties Ausführungen immer dann, wenn es um praktische Konsequenzen geht (insbesondere in Kapitel 11, S. 279ff.), geprägt durch klare Botschaften und durch eine emphatisch vorgetragene normative pädagogische Ausrichtung. Dies wirkt eindringlich und überzeugend, wenngleich der Eindruck entsteht, dass Hattie, sobald es um seine Erziehungslehre geht, sich teilweise mehr durch sein eigenes Anliegen leiten lässt als durch seine zusammengetragenen empirischen Befunde. Insbesondere lässt sich das an seiner Kritik an konstruktivistischen Ansätzen zu Lehren und Lernen (siehe dazu z.B. Hattie, 2013, S. 31f.) sowie an seinem Plädoyer für eine Lehrperson als „activator“ in Abgrenzung zu einer Lehrperson als „facilitator“ ablesen (siehe dazu v.a. Hattie, 2013, S. 286ff.).

„Direkte Instruktion“ und Offener Unterricht im Widerstreit?

John Hattie macht in seinen Publikationen unmissverständlich klar, dass er *die aktiv steuernde Lehrperson* dem weitgehend zurückhaltend agierenden, auf Selbststeuerung der Lernenden setzenden „Moderators“ vorzieht. So kritisiert er deutlich ein zurückgenommenes Lehrerhandeln, das im Wesentlichen der Eigeninitiative der Lernenden Raum gibt. Ein solches Vorgehen verbindet er mit radikal-konstruktivistischen Konzeptionen, die er für nicht zielführend hält: „Konstruktivismus wird zu oft im Sinne eines schülerzentrierten, forschenden problem- und aufgabenbasierten Lernens gesehen und dann in der Fachwelt mit Begriffen wie ‚authentisches‘, ‚entdeckendes‘ und ‚intrinsisch motiviertes Lernen‘ belegt. Es wird gesagt, dass die Rolle der konstruktivistischen Lehrperson vor allem darin liege, Gelegenheiten für einzelne Lernende zu schaffen, dass diese durch eigene Aktivität und durch Diskussion, Reflexion und Austausch von Einfällen mit anderen Lernenden Wissen erwerben und Bedeutung konstruieren können, und dass all dies mit minimaler korrigierender Intervention verbunden ist. [...] Solche Aussagen sind aber fast das genaue Gegenteil eines erfolgreichen Rezepts für Lehren und Lernen [...]“ (Hattie, 2013, S. 32) Mit dem „Rezept“ meint Hattie das Konzept der „Direkten Instruktion“, die „fälschlicherweise einen schlechten Ruf (hat), insbesondere dann, wenn es mit Frontalunterricht verwechselt wird“ (Hattie, 2013, S. 242).

Aus der oben abgebildeten „Tabelle 21 Effekte der Lehrperson als Regisseur und Moderator“ (Hattie, 2013, S. 287) ist der Gegenüberstellung ein klares Muster zu entnehmen, dass zunächst einmal für Hatties Argumentation spricht, namentlich die „Direkte Instruktion“ und damit verbundene Konzepte. Auffallend sind demgegenüber die recht geringen Effektmaße für die *im Vorliegenden* aufgelisteten Variablen zur „Lehrperson als Moderator“, die einer reformpädagogisch ausgerichteten Konzeption zuzurechnen sind und häufig im Zusammenhang mit der Förderung selbstständigen und sozialen Lernens praktiziert werden.

Wie Hatties Forschungsbilanz zu entnehmen ist, besteht offenbar bei diesen Konzepten die Gefahr, dass Schülerinnen und Schülern hierbei zu wenig Ordnungsstrukturen und Orientierungen in Lernprozessen zur Verfügung gestellt werden und sie deshalb neues Wissen nicht effektiv verarbeiten und angemessen anwenden können. Dies gilt vor allem dann, wenn erforderliches Vorwissen fehlt, was insbesondere bei der Einführung in ein neues Themenfeld der Fall ist. Gerade schwächere Schülerinnen und Schüler kommen mit offeneren Lernkontexten weniger gut klar, weil ihnen dazu die „kognitiven Landkarten“ zur Selbstorganisation der Lernprozesse fehlen. Für sie ist eine engere Führung mit kürzeren Anleitungsintervallen umso wichtiger. Deshalb gelten offene Lernarrangements, in denen die Lernenden auf sich gestellt sind, als voraussetzungsreich und können sicherlich nicht als „Königsweg“ unterrichtlicher Prozesse angesehen werden.

Insofern sollte der Hattie-Befund auf offene Lernkontexte hinweisen und zumindest eine Warnung sein, die Voraussetzungen eines offenen Unterrichts nicht zu unterschätzen. So ist sich Andreas Gold mit Hattie darüber einig, dass „die Wirksamkeit der offenen, problemorientierten Methoden teilweise deutlich überschätzt (wird)“ (Gold, 2015, S. 149f.). „Wahrscheinlich muss man sich dieses ‚Zerrbild‘ über die relative Wirksamkeit der Unterrichtsmethoden vor Augen halten, um John Hatties vehementes Plädoyer für ein stärker lehrergelenktes Vorgehen besser verstehen zu können. Hattie sieht sich hier als ‚Aufklärer‘“ (Gold 2015, S. 150).

Ist deshalb davon auszugehen, dass „offene“ Lernarrangements wirklich wirkungslos sind? Um Missverständnisse zu vermeiden, ist auf Folgendes hinzuweisen: Auch wenn die Effektmaße dieser Herangehensweisen nahe null liegen, so folgt daraus nicht, dass diese nicht wirksam oder schlechter als der traditionelle Unterricht sind, sondern nur, dass das Effektmaß im Vergleich zum vorherrschenden üblichen Unterricht keinen Unterschied anzeigt – denn damit wurden die reformpädagogischen Konzepte verglichen. Dabei gilt es zu betonen, dass das Leistungskriterium sich auf fachliches Lernen bezieht. Das ist auch ein Grund für Andreas Helmke, dass die Konzepte einer Lernbegleitung nach wie vor ihre Berechtigung haben, nämlich mit Blick auf eine Förderung überfachlicher Kompetenzen. Im Übrigen geht es für ihn nicht um eine „Verabsolutierung eines Unterrichtsstils“ (vgl. Helmke, 2013, S. 12). Zur Frage offener Unterrichtsformen macht Helmke deutlich, dass diese genauso lernwirksam sein können wie traditioneller Unterricht (vgl. S. 13): „Offener Unterricht kann – wenn gekoppelt mit fachübergreifenden Qualitätskriterien wie kognitiver Aktivierung, herausfordernden und passenden Aufgaben, lernförderlichem fehlerfreundlichen Klima und (mit) vielem Feedback – sehr lernwirksam sein“ (S. 13). Für wichtig hält er dabei, dass die Kinder und Jugendlichen dem selbstständigen Lernen gewachsen sein müssen und dass Lehrpersonen die Phasen offenen Unterrichts „gründlich vorbereiten und über seinen Verlauf wachen, anstatt sich zurückzuziehen“ (S. 13). Dabei betont er, dass Hatties Prinzipien für Lernwirksamkeit „unabhängig von der jeweils gewählten Methode gelten“ (S. 13).

Dennoch gilt es, die geringen Effekte zur Lehrperson als Moderator genauer in Augenschein zu nehmen: Wofür sind die referierten Daten ein Beleg? Was bringen sie zum Ausdruck? So erscheint Hatties Zuordnung der Konzepte nicht zwingend, im

Einzelnen sogar willkürlich, zumal diese nicht näher begründet wird und eine theoretische Konzeptualisierung zu fehlen scheint. Warum gerade diese Auswahl? Wieso wird „Freiarbeit“ ($d = .04$) nicht berücksichtigt, da sie doch eine recht ‚klassische‘ Vorgehensweise offenen Unterrichts darzustellen scheint. Warum hat Hattie beispielsweise „Peer-Tutoring“ ($d = .55$) und „reziprokes Lehren“ ($d = .74$) nicht dem Typus der moderierenden Lehrperson zugerechnet? Einige der wirkungsmächtigen Faktoren zu „Lehrperson als Regisseur“ könnten genauso gut den offenen Lernformen zugeschlagen werden. Beispielsweise enthalten die Konzepte „Reziprokes Lehren“ ($d = .74$), „Metakognitive Strategien“ ($d = .69$) und „Mastery-Learning“ ($d = .58$) auch typische Komponenten offener Lernformen. In der Originalausgabe von 2009 wurden dem Moderator-Typus (dort: „Teacher as facilitator“) gar noch Faktoren zugerechnet, für die es keinen erkennbaren inhaltlichen Bezug gibt, nämlich „Smaller class sizes“ ($d = .21$) und „Different teaching for boys and girls“ ($d = .12$). Bei dieser Gegenüberstellung wird deutlich, dass eine pauschale Bilanzierung zu Fehldeutungen führen kann und dass eine differenziertere Betrachtungsweise erforderlich ist (vgl. dazu ausführlicher Steffens & Höfer, 2016, Kap. 5.2 und 6.2).

Schaut man sich die Variablen, die Hattie dem Moderatoren-Typus zurechnet, näher an, dann fällt auf, dass sie überwiegend die Sichtstrukturen von Unterricht erfassen und weniger die Tiefendimensionen von Lehr- und Lernprozessen betreffen. Insofern macht es Sinn, die reformpädagogischen Ansätze dementsprechend zu unterscheiden. In der folgenden Tabelle sind nun solche Konzepte aufgelistet, die etwas mit ‚konstruktiver Unterstützung‘ zu tun haben.¹³ Für sie zeigt sich ein gänzlich anderes Ergebnismuster.

Reformpädagogische Konzepte mit aktivierenden Vorgehensweisen	Effekt maß d
Concept Mapping	.57
Mastery learning (zielerreichendes Lernen)	.58
Problemlösen (problemsolving teaching)	.61
Lautes Denken (self-verbalisation / self-questioning)	.64
Metakognitive Strategien	.69
Durchschnitt	.62

¹³ Gold merkt an, dass eine entscheidende Qualitätsdimension, die konstruktive Unterstützung, in Hatties Forschungsbilanz gar nicht vorkommt. Man muss sie aufgrund des ‚Fehlens einer ordnenden Struktur‘ in ihren einzelnen Facetten herausfiltern: „Für einige der einflussreichsten Hattie-Faktoren, wie z.B. die ‚Reaktion auf Interventionen‘ (Rangplatz 3), die ‚formative Evaluation‘ (Rangplatz 4), die ‚Interventionen bei Lernenden mit besonderem Förderbedarf‘ (Rangplatz 8), die ‚Klarheit der Lehrpersonen‘ (Rangplatz 9) oder das ‚reziproke Lehren‘ (Rangplatz 11) ist es naheliegend zu vermuten, dass eine schülerorientiert-unterstützende Grundhaltung eine Gelingensbedingung ihrer Wirksamkeit ist“ (Gold, 2015, S. 91).

Synthese von direktivem und selbstständigkeitsförderndem Unterricht

Offenbar können die erwähnten offenen Konzepte nicht ‚im Selbstlauf‘ wirksame Lernprozesse auslösen, vor allem dann nicht, wenn sie in einem reformpädagogisch naiven „Laissez-faire“-Modus praktiziert werden und nicht die Synthese direkter und selbstständigkeitsfördernder Unterrichtsmomente betonen (vgl. z.B. den Beitrag von Messner & Blum im vorliegenden Band). Um die erforderlichen Tiefenstrukturen zu erreichen, benötigen sie darüber hinaus Lehr- und Lernstrategien, die mit kognitiven Aktivierungen einhergehen (vgl. dazu auch Messners Beitrag „Bausteine eines kognitiv aktivierenden Fachunterrichts“ in diesem Band). Die Scheidelinie zwischen wirksamen und weniger wirksamen Faktoren liegt demzufolge nicht so sehr in der vermeintlich richtigen oder falschen Unterrichtsmethode, sondern vielmehr in mehr oder weniger kognitiv herausfordernden Lehr-Lernsequenzen. Dementsprechend kommt es Andreas Gold darauf an, „auf die eigentlichen Dimensionen der Unterrichtsqualität zu schauen, statt auf die leicht sichtbaren Methoden und Sozialformen des Unterrichts“. „Wenn John Hattie vom ‚aktivierenden‘ Lehrer spricht, scheint eine solche Tiefendimension der Unterrichtsqualität immerhin durch.“ (Gold, 2015, S. 165)

Bei den gegenüberstellenden Vergleichen entsteht häufig der Eindruck, als ob es um ein Entweder-oder verschiedener Unterrichtsmethoden gehe. Ein solcher abschließender Vergleich ist allenfalls von allgemeinem Erkenntnisinteresse, aber letztendlich ohne praktische Konsequenzen, weil angeleitete und offene Unterrichtsarrangements nie in Reinform den Unterrichtsalltag bestimmen dürften. Insofern geht es nicht darum, die verschiedenen Unterrichtsformen gegeneinander auszuspielen, weil sie – je nach Zielsetzung – ihre jeweilige Berechtigung haben. Worauf es vielmehr ankommt, ist ihr angemessenes Verhältnis zueinander. Deshalb geht es bei der Steigerung von Unterrichtsqualität nicht um ein Entweder-oder der verschiedenen Methoden, sondern um ein Sowohl-als auch und um eine angemessene Balance. So schreibt auch Hattie: „Das Modell des sichtbaren Lehrens und Lernens kombiniert *lehrerzentriertes Lehren und schülerzentriertes Lernen*, statt beide gegeneinander auszuspielen“ (Hattie, 2013, S. 31; Hervorhebung: US). Und noch etwas dürfte in der vorstehenden Auseinandersetzung mit Hatties Befunden deutlich geworden sein: „Das Ranking der Methoden hilft nicht wirklich. Im Prinzip lässt sich mit allen Unterrichtsmethoden guter Unterricht machen, soweit es den Lehrern dabei gelingt, die Schüler zum Denken herauszufordern und ihre individuellen Lernprozesse konstruktiv zu unterstützen“ (Gold, 2015, S. 148).

Literatur

- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9, 469–520.
- Baumert, J. & Kunter, M. (2011). Das mathematikspezifische Wissen von Lehrkräften, kognitive Aktivierung im Unterricht und Lernfortschritte von Schülerinnen und Schülern. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand

- (Hrsg.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (S. 163–192). Münster: Waxmann.
- Beywl, W. (2015). Den Unterricht wirksam gestalten. Anregungen aus der Hattie-Studie. *Bündner Schulblatt*, Ausgabe Mai, 14–17.
- Beywl, W. & Strasser, J. (2017). Lernen sichtbar machen – neu: mehr als 250 Einflussfaktoren auf die Lernleistungen. *Newsletter Lernen sichtbar machen*, 20, Oktober. Verfügbar unter: https://www.lernensichtbarmachen.ch/wpcontent/uploads/2017/10/Newsletter20_Mail_final.pdf [20. April 2018].
- Beywl, W. & Zierer, K. (2013). Lernen sichtbar machen. Zur deutschsprachigen Ausgabe von „Visible Learning“. In J. Hattie, *Lernen sichtbar machen. Überarbeitete deutschsprachige Ausgabe von „Visible Learning“ besorgt von Wolfgang Beywl und Klaus Zierer* (S. VI–XXVI). Baltmannsweiler: Schneider Verlag Hohengehren.
- Beywl, W. & Zierer, K. (2014). „Visible Learning“ wird zu „Lernen sichtbar machen“. Ein Kommentar zur Übersetzung und Überarbeitung der Hattie-Studie (S. 147–162). In E. Terhart (Hrsg.), *Die Hattie-Studie in der Diskussion. Probleme sichtbar machen*. Seelze: Kallmeyer.
- Black, P. & Wiliam, D. (2001). *Innerhalb der „Black Box“: Verbesserung der Bildungsstandards durch die Bewertung des gesamten Unterrichtsgeschehens*. London.
- Brügelmann, H. (2014). Gilt nach Hattie: Je häufiger, desto besser? Zur Bedeutung von „Evidenzbasierung“ für pädagogisches Handeln vor Ort (S. 38–50). In E. Terhart (Hrsg.), *Die Hattie-Studie in der Diskussion. Probleme sichtbar machen*. Seelze: Kallmeyer.
- Fuchs, L.S. & Fuchs, D. (1986). Effects of systematic formative evaluation: A metaanalysis. *Exceptional Children*, 53 (3), 199–208.
- Gold, A. (2014). Vergesst Hattie! Warum John Hatties MetaMetaanalyse für die Lehrerbildung wenig brauchbar ist. *Schulverwaltung, Ausgabe Hessen/RheinlandPfalz*, 19 (12), 332–336.
- Gold, A. (2015). *Guter Unterricht. Was wir wirklich darüber wissen*. Göttingen: Vandenhoeck & Ruprecht.
- Hattie, J. (2009). *Visible Learning – A synthesis of over 800 metaanalyses relating to achievement*. London, New York: Routledge.
- Hattie, J. (2013). *Lernen sichtbar machen. Überarbeitete deutschsprachige Ausgabe von „Visible Learning“ besorgt von Wolfgang Beywl und Klaus Zierer*. Baltmannsweiler: Schneider Verlag Hohengehren.
- Hattie, J. (2014). *Lernen sichtbar machen für Lehrpersonen. Überarbeitete deutschsprachige Ausgabe von „Visible Learning for Teachers“ besorgt von Wolfgang Beywl und Klaus Zierer*. Baltmannsweiler: Schneider Verlag Hohengehren.
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77 (1), 81–112.
- Helmke, A. (2013). Interview mit Prof. Dr. Andreas Helmke zur Hattie-Studie, interviewt von Prof. Dr. Volker Reinhardt. *Lehren & Lernen. Zeitschrift für Schule und Innovation aus BadenWürttemberg*, 39 (7), 9–16.
- Klieme, E., Lipowsky, F., Rakoczy, K. & Ratzka, N. (2006). Qualitätsdimensionen und Wirksamkeit von Mathematikunterricht. Theoretische Grundlagen und ausgewählte Ergebnisse des Projekts „Pythagoras“. In M. Prenzel & L. Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFGSchwerpunktprogramms* (S. 127–146). Münster: Waxmann.
- Köller, O., Möller, J. & Möller, J. (2013). *Was wirkt wirklich? Einschätzungen von Determinanten schulischen Lernens* (SchulmanagementHandbuch, Band 45). München: Oldenbourg Schulbuchverlag.

- lernensichtbarmachen.ch (2018). *Formative Evaluation des Unterrichts*. Verfügbar unter: https://web.fhnw.ch/plattformen/hattiewiki/begriffe/Formative_Evaluation_des_Unterrichts [20. April 2018].
- Maier, U. (2010). Formative Assessment – Ein erfolgversprechendes Konzept zur Reform von Unterricht und Leistungsmessung? *Zeitschrift für Erziehungswissenschaft*, 13 (2), 293–308.
- Meyer, H. (2014). Auf den Unterricht kommt es an! Hatties Daten deuten lernen. In E. Terhart (Hrsg.), *Die Hattie-Studie in der Diskussion. Probleme sichtbar machen* (S. 117–133). Seelze: Kallmeyer.
- Nix, F. & Wollmann, J. (2015). *Hattie und die Folgen. Empirische Befunde und didaktische Konsequenzen zum erfolgreichen Unterrichten*. Berlin: Cornelsen Schulverlage.
- Pant, H.A. (2014). Aufbereitung von Evidenz für bildungspolitische und pädagogische Entscheidungen: Metaanalysen in der Bildungsforschung. *Zeitschrift für Erziehungswissenschaft*, 17, 79–99.
- Rolff, H.G. (2013). Die Hattie-Studie. Ein RorschachTest. *Pädagogik*, 65 (3), 46–49.
- Rolff, H.G. (2014). Sind schulische Strukturaktoren wirklich unwichtig? Hattie und das deutsche Schulsystem. In E. Terhart (Hrsg.), *Die Hattie-Studie in der Diskussion. Probleme sichtbar machen* (S. 67–77). Seelze: Kallmeyer.
- Schulmeister, R. & Loviscach, J. (2014). Kritische Anmerkungen zur Studie „Lernen sichtbar machen“ (Visible Learning) von John Hattie. *Seminar*, 20 (2), 121–130.
- Steffens, U. & Höfer, D. (2016). *Lernen nach Hattie. Wie gelingt guter Unterricht?* Weinheim & Basel: Beltz.
- Terhart, E. (2011). Hat John Hattie tatsächlich den Heiligen Gral der Schul und Unterrichtsforschung gefunden? Eine Auseinandersetzung mit Visible Learning (S. 277–292). In E. Keiner et al. (Hrsg.), *Metamorphosen der Bildung. Historie – Empirie – Theorie. Festschrift für Heinz-Elmar Tenorth*. Bad Heilbrunn: Klinkhardt.
- Terhart, E. (2014). Der Heilige Gral der Schul und Unterrichtsforschung – gefunden? Eine Auseinandersetzung mit Visible Learning. In E. Terhart (Hrsg.), *Die Hattie-Studie in der Diskussion. Probleme sichtbar machen* (S. 10–23). Seelze: Kallmeyer.