

Die Hattie-Studie und ihre Bedeutung für den Unterricht

Ein Blick auf ausgewählte Aspekte der Lehrer-Schüler-Interaktion

Miriam Lotz & Frank Lipowsky

Lilli und Ben gehen in die vierte Klasse der Grundschule. Während Lilli in fast allen Fächern sehr gute Noten bekommt, hat Ben in vielen Fächern Schwierigkeiten. Wie lassen sich diese Unterschiede erklären? Hat Ben Schwierigkeiten, weil seine Schule in einem Problembezirk liegt oder weil er über geringe kognitive Fähigkeiten verfügt? Und ist Lilli leistungsstark, weil sie von ihren Eltern gezielt gefördert wird oder aber hat sie eine besonders »gute Lehrperson«?

Inzwischen gilt es als vielfach belegt, dass Schulleistungen *multikausal bedingt* sind. Das bedeutet, dass erst das Zusammenspiel aus individuellen Schülermerkmalen wie der Intelligenz, dem Vorwissen oder der Leistungsmotivation, aus Kontextfaktoren wie dem Elternhaus und der außerschulischen Förderung und auch schulischen sowie unterrichtlichen Faktoren einen Einfluss auf den Schulerfolg ausübt (vgl. Abbildung 1). Allein diese Erkenntnis sagt aber noch wenig darüber aus, wie bedeutsam diese einzelnen Merkmale im Vergleich zueinander sind. Welche dieser Merkmale wirken am stärksten? Diese Frage stand im Mittelpunkt der sogenannten Hattie-Studie, einer quantitativen Synthese von über 800 Metaanalysen zum Einfluss unterschiedlichster Merkmale auf den kognitiven Lernerfolg von Schülern (Hattie 2009, 2013). Die Studie resultiert unter anderem in einer Rangliste, welche die Effekte von insgesamt 138 einzelnen Merkmalen auf den Schulerfolg enthält, die den sechs in Abbildung 1 gezeigten Bereichen zugeordnet werden, und diese nach ihrer Bedeutung ordnet.

Im vorliegenden Beitrag wird die Hattie-Studie zunächst in ihren Grundzügen beschrieben und die Hauptergebnisse werden zusammengefasst. Im Vordergrund stehen dabei diejenigen Merkmale, die im direkten Einflussbereich der Lehrperson liegen, da diese auch im Unterricht veränderbar sind. Da es allerdings zur Realisierung »guten Unterrichts« nicht genügt, die Rangfolge der einzelnen Merkmale zu kennen, wird darauf aufbauend noch einmal spezifisch auf die konkrete Umsetzung einzelner Merkmale eingegangen. Dabei wird auch thematisiert, welcher Voraussetzungen und Überlegungen es bedarf, um das Potenzial dieser Merkmale möglichst gut zu nutzen. Dazu wird der Fokus auf die Unterrichtsqualität unter besonderer Berücksichtigung einer kognitiv anregenden Lehrer-Schüler-Interaktion gelegt, indem zwei Merkmale – das Stellen von Fragen und das Erteilen von Feedback – ausgewählt werden und näher darauf eingegangen wird, wie diese Aspekte zur Verbesserung der Unterrichtsqualität beitragen können.

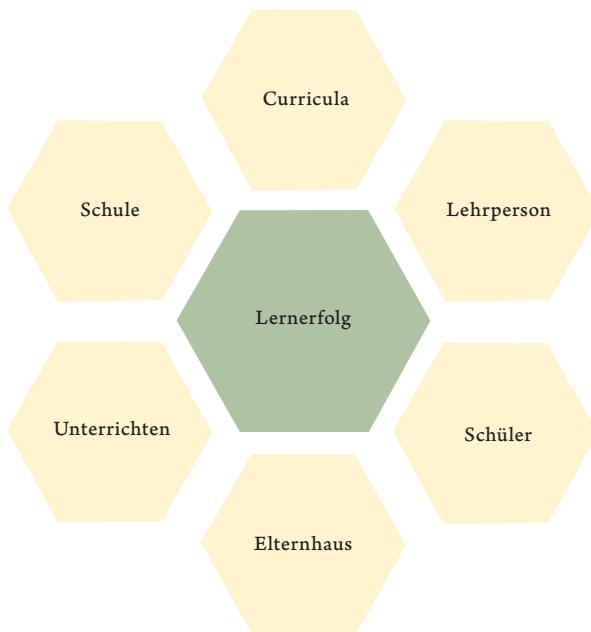


Abbildung 1: Determinanten des Lernerfolgs (vgl. Hattie 2009, 2013)

Der Boom der Hattie-Studie – Grundidee und Hauptergebnisse

Die Frage, was guten Unterricht ausmacht, beschäftigt Lehrpersonen und Wissenschaftler seit Jahrzehnten. Durch empirische Studien unterschiedlichster Art wird versucht, Erkenntnisse darüber zu gewinnen, welche Merkmale den Schulerfolg entscheidend beeinflussen. Kaum einer Studie im Bereich der Schul- und Unterrichtsforschung wurde in den letzten Jahren eine solche Aufmerksamkeit zuteil wie der sogenannten »Hattie-Studie«.

Im Jahr 2009 veröffentlichte der neuseeländische Bildungsforscher John Hattie die Monografie »Visible Learning«. Dabei geht es um die übergeordnete Frage, welche Merkmale für das schulische Lernen besonders relevant sind. Die Studie wurde im Jahr 2013 unter dem Titel »Lernen sichtbar machen« durch Beywl und Zierer übersetzt und damit auch im deutschsprachigen Raum noch einmal bekannter. Die Publikation wird als »Meilenstein in der Debatte um Voraussetzungen und Bedingungen erfolgreichen Lernens in der Schule« (Terhart 2014, S. 10) angesehen.

Hatties Vorgehen: Die Meta-Metaanalyse

Hatties Studie stellt eine *quantitative Synthese von Meta-Analysen* dar – also von Studien, die ihrerseits zu einem bestimmten Thema die Ergebnisse von Einzelstudien rechnerisch zusammenfassen. Damit kann die Hattie-Studie als Meta-Metaanalyse

bezeichnet werden (Terhart 2014). Das Besondere an der Studie ist vor allem ihre inhaltliche Breite sowie der enorme Umfang. Hattie fokussiert nicht auf einzelne mögliche Einflüsse auf das Lernen, sondern bezieht Studien aus den unterschiedlichsten Bereichen mit ein. Insgesamt fasst er die Ergebnisse von etwa 800 internationalen Metastudien zusammen, welche ihrerseits auf über 50.000 Einzelstudien beruhen und sich zusammengenommen auf eine Stichprobe von etwa 250 Millionen Lernenden beziehen.

Auf Basis der vorhandenen Metaanalysen berechnet Hattie sogenannte *Effektstärken* für jedes der insgesamt 138 einbezogenen Merkmale, die er den sechs Bereichen Lernende, Lehrperson, Elternhaus, Schule, Curriculum und Unterricht zuordnet. Dabei unterscheidet sich die Anzahl der einbezogenen Merkmale pro Bereich sowie die Anzahl der Metaanalysen, wie folgende Auflistung zeigt:

- Lernende (7 Merkmale, 35 Studien)
- Lehrperson (10 Merkmale, 31 Studien)
- Elternhaus (19 Merkmale, 139 Studien)
- Curriculum (25 Merkmale, 144 Studien)
- Schule (28 Merkmale, 101 Studien)
- Unterrichten (49 Merkmale, 365 Studien)

Was sind Effektstärken? *Effektstärken geben Auskunft darüber, wie stark der Einfluss eines Merkmals (z. B. eines Förderprogramms) auf eine Zielvariable (z. B. die Leseleistung) ist. Je höher die Effektstärke, desto bedeutsamer ist der Einfluss der jeweiligen Variable. Während negative Effektstärken anzeigen, dass ein bestimmtes Merkmal zu einer Verringerung der Leistung führt, bedeutet eine Effektstärke von Null, dass das ausgewählte Merkmal weder eine positive noch eine negative Wirkung hat. Ab einer Effektstärke von $d = 0.40$ kann eine pädagogische Maßnahme als sehr erfolgreich bezeichnet werden, da eine Effektstärke von $d = 0.40$ ungefähr dem Lernzuwachs eines Schuljahres entspricht.*

Hattie veranschaulicht die Wirkungen der 138 von ihm untersuchten Merkmale anhand eines »Effektstärke-Barometers« (vgl. Abbildung 2, adaptiert nach Hattie 2009, 2013). Hier wird noch einmal anschaulich deutlich, was die unterschiedlichen Effektstärke-Bereiche bedeuten.

Während der rot markierte Bereich negative Effekte anzeigt, folgen darauf zunächst Bereiche, die lediglich eine geringe Wirksamkeit signalisieren. Ab einer Effektstärke von 0.40 kann von deutlichen Effekten gesprochen werden. Diesen Grenzwert begründet Hattie damit, dass der mittlere Effekt aller von ihm untersuchten Merkmale

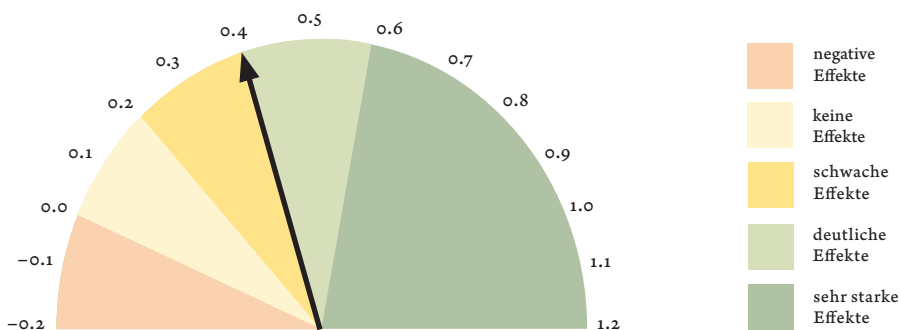


Abbildung 2: Effektstärke-Barometer nach Hattie

zusammengenommen bei 0.40 liegt. Demnach sind Effektstärken von $d \geq 0.40$ überdurchschnittlich wirksam und werden als bedeutsam beurteilt. Ab dem Bereich von 0.60 kann man dann sogar von sehr starken Effekten sprechen, da diese Wirkungen über den Lernzuwachs eines gesamten Schuljahres hinausgehen.

Zentrale Ergebnisse

Aus den zahlreichen Ergebnissen der Hattie-Studie sollen hier nur zwei herausgestellt werden: die generelle Bedeutung der Lehrperson und des Unterrichts sowie die Effektstärken ausgewählter Merkmale im Vergleich.

»Teachers make the difference«

Dies ist eines der zentralen Ergebnisse, das sehr häufig aus der Hattie-Studie herausgelesen wird. Hierbei geht es um die Frage, welcher der sechs übergeordneten Bereiche den stärksten Einfluss auf das Lernen der Schüler ausübt. Wichtig ist, dass Hattie nur den Effekt auf die Leistungen der Schüler in seine Analysen einbezogen hat, nicht beispielsweise Effekte auf die Motivation oder das Selbstkonzept der Schüler. Obwohl der vielzitierte Satz »teachers make the difference« darauf hindeutet, dass der Bereich »Lehrperson« die stärkste Wirksamkeit aufweist, sind die Merkmale der »Lernenden« selbst die bedeutungsstärksten Faktoren. Sie klären, wenn man die Studien zusammennimmt, ungefähr bis zu 50 % der Varianz in den Schulleistungen auf.¹ Das bedeutet, dass ungefähr die Hälfte der individuellen Unterschiede in den Lernleistungen auf Merkmale des Schülers – wie seine Intelligenz oder sein Vorwissen – zurückgeht.

¹ Hierbei handelt es sich um ungefähre Schätzungen, die von Fach zu Fach und auch von Kompetenz zu Kompetenz unterschiedlich ausfallen können.

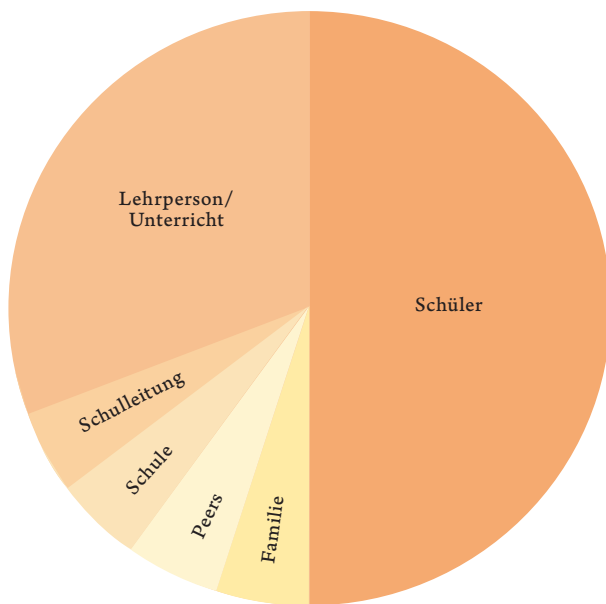


Abbildung 3: Erklärung von Leistungsunterschieden durch die sechs übergeordneten Bereiche

Auch wenn diese Lernerkmale den stärksten Einfluss ausüben, stellt Hattie (2003) heraus, dass der Beitrag der Lehrperson und des Unterrichts ebenfalls sehr bedeutsam und für bis zu 30 % der Unterschiede in den Leistungen der Schüler verantwortlich ist. Damit ist der Einfluss der Lehrperson größer als der Einfluss schulisch-struktureller Merkmale. Demnach ist es für den Lernerfolg eines Kindes entscheidender, in welche Klasse es geht als welche Schule es besucht (vgl. Abbildung 3).

Führt man sich vor Augen, dass die zentralen Lernprozesse innerhalb des Klassenraums vonstattengehen, überrascht dieses Ergebnis nicht. Es wird oftmals zum Anlass genommen, an Verbesserungen der Unterrichtsqualität anzusetzen, wenn es um eine wirksame Förderung der Schüler geht. Denn während Lehrpersonen nur wenig Einfluss auf die persönlichen Voraussetzungen der Schüler, ihre Familie und ihre Peers nehmen können und auch schulstrukturelle Maßnahmen schwieriger umsetzbar erscheinen, lässt sich der Unterricht durch Maßnahmen der Lehrperson gestalten und auch verändern. Was genau am Unterricht verändert werden könnte, um möglichst vielen Schülern hohe Lernerfolge zu ermöglichen, lässt sich aus diesem übergeordneten Ergebnis aber noch nicht ableiten. Hierzu lohnt ein Blick auf die Effektstärken der 138 untersuchten Einzelmerkmale.

Was wirkt am besten? – Effektstärken ausgewählter Merkmale im Vergleich

Tabelle 1 enthält ausgewählte Merkmale sowie ihre Effektstärken und ihre Rangplätze. Auf Platz 1 der von Hattie einbezogenen Merkmale steht die Selbsteinschätzung des eigenen Leistungsniveaus, gefolgt von der kognitiven Entwicklungsstufe. Beide Merkmale lassen sich dem Bereich des Lernenden zuordnen. Bereits auf Platz 3 und 4 folgen dann aber Merkmale, die im Einflussbereich der Lehrperson liegen: die formative Evaluation des Unterrichts (also das Einholen von Feedback zum eigenen Unterricht durch die Lehrperson) sowie das Micro-Teaching. Beim Micro-Teaching handelt es sich um eine Methode in der Lehrerbildung, bei der einzelne Merkmale lernförderlichen Lehrerverhaltens (z. B. das Stellen von Fragen) eingeübt werden, indem (angehende) Lehrpersonen eine Unterrichtssituation simulieren und zu ihrem Verhalten Rückmeldung erhalten sowie dieses reflektieren (Hattie 2013). Auch unter den weiteren vorderen Plätzen finden sich mit Merkmalen wie Feedback, der Lehrer-Schüler-Beziehung oder der inhaltlichen Klarheit einige Faktoren, welche deutlich von der Lehrperson beeinflusst werden können. Dies sind nach der Hattie-Studie also Merkmale, auf die Lehrpersonen, die ihren Unterricht verbessern möchten, einen verstärkten Blick richten sollten.

Unter den Merkmalen mit deutlichen Effekten finden sich zum einen Aspekte des Elternhauses wie das häusliche Anregungsniveau oder der sozioökonomische Status der Eltern, aber auch wieder bedeutsame unterrichtliche Faktoren, wie die Direkte Instruktion. Bei der Direkten Instruktion werden die Lernvoraussetzungen stetig geprüft, es finden Rückblicke auf vorausgegangene Lerneinheiten statt, es steht genügend Zeit zum angeleiteten Üben zur Verfügung und der Lernerfolg wird überwacht und den Schülern zurückgemeldet (Hasselhorn & Gold 2013).

Geringere Effekte konnte Hattie beispielsweise für Faktoren wie Hausaufgaben, die Klassengröße oder die Individualisierung des Unterrichts nachweisen. Weder positive noch negative Effekte haben nach den Ergebnissen von Hattie beispielsweise das Team-Teaching, jahrgangsübergreifende Klassen oder der Offene Unterricht, in dem den Schülern beispielsweise Freiheiten bei der Wahl der Lernmethode, der -materialien oder der -inhalte gewährt werden.

Diese Ergebnisse sollten allerdings nicht dahingehend missverstanden werden, dass all diese Aspekte generell überhaupt nichts »bringen«. Hier muss berücksichtigt werden, dass es sich bei den Ergebnissen um Mittelwerte handelt. So gilt für die Wirksamkeit von Hausaufgaben beispielsweise, dass insbesondere in höheren Schulstufen und bei einer guten Qualität der Aufgaben durchaus höhere Effekte erzielt werden können. Zudem muss immer auch berücksichtigt werden, mit welcher Art von Unterricht die jeweiligen Merkmale verglichen wurden. So bedeutet der Nulleffekt des Offenen Unterrichts lediglich, dass dieser keine besseren Lernergebnisse zur Folge hat als traditio-

	Merkmal	Bereich	<i>d</i>	Rang
	<i>d</i> ≥ 0.60			
Merkmale mit sehr starken Effekten	Selbsteinschätzung des eigenen Leistungsniveaus	Lernende	1.44	1
	Kognitive Entwicklungsstufe (nach Piaget)	Lernende	1.28	2
	Formative Evaluation des Unterrichts	Unterrichten	0.90	3
	Micro-Teaching	Lehrperson	0.88	4
	Klarheit der Lehrperson	Lehrperson	0.75	8
	Feedback	Unterrichten	0.73	10
	Lehrer-Schüler-Beziehung	Lehrperson	0.72	11
	Meta-kognitive Strategien	Unterrichten	0.69	13
	Vorausgehendes Leistungsniveau	Lernende	0.67	14
	0.40 ≤ <i>d</i> ≤ 0.60			
Merkmale mit deutlichen Effekten	Lerntechniken	Unterrichten	0.59	25
	Direkte Instruktion	Unterrichten	0.59	26
	Häusliches Anregungsniveau	Elternhaus	0.57	31
	Sozioökonomischer Status	Elternhaus	0.57	32
	Ziele	Unterrichten	0.56	34
	Klassenzusammenhalt	Schule	0.53	39
	Effektive Klassenführung	Schule	0.52	42
	Elternunterstützung beim Lernen	Elternhaus	0.51	45
	Motivation	Lernende	0.48	51
Fragenstellen	Lernende	0.46	53	
	0.20 ≤ <i>d</i> ≤ 0.40			
Merkmale mit geringen Effekten	Aktive Lernzeit	Unterrichten	0.38	70
	Schulleitung	Schule	0.36	74
	Reduzieren von Unterrichtsstörungen	Unterrichten	0.34	80
	Induktives Vorgehen	Unterrichten	0.33	83
	Hausaufgaben	Unterrichten	0.29	88
	Finanzielle Ausstattung	Schule	0.23	99
	Individualisierung	Unterrichten	0.23	100
	Klassengröße	Schule	0.21	106
	0.00 ≤ <i>d</i> ≤ 0.20			
Merkmale, die weder positive noch negative Effekte haben	Co-Teaching/Team-Teaching	Unterrichten	0.19	111
	Außercurriculare Aktivitäten	Curriculum	0.17	114
	Gender	Lernende	0.12	122
	Jahrgangübergreifende Klassen	Schule	0.04	131
	Freiarbeit	Unterrichten	0.04	132
Offene Klassenzimmer/Offene Lehr- u. Lernformen	Schule	0.01	133	
	<i>d</i> ≤ 0.00			
Merkmale mit negativen Effekten	Fernsehen	Elternhaus	-0.18	137
	Schulwechsel	Schule	-0.34	138

Tabelle 1: Effektstärken ausgewählter Merkmale im Vergleich

ner Unterricht. Auch bei vielen anderen Merkmalen sollte immer bedacht werden, dass es vor allem auf die jeweilige Umsetzung ankommt. Eine Studie mit einem aufgrund ihres enormen Umfangs sehr groben Auflösungsgrad kann solche differenziellen Effekte jedoch nicht abbilden.² Auch die wenigen negativen Effekte, wie beispielsweise für Schulwechsel oder Fernsehen, lassen sich plausibel erklären. Interessant ist hierbei vor allem, dass keines der untersuchten Merkmale auf Ebene der Lehrperson oder des Unterrichts negative Effekte hat. Die wenigen Merkmale mit negativen Auswirkungen auf die Schulleistung lassen sich ausschließlich den Bereichen Elternhaus oder Schule zuordnen.

Vertiefender Blick auf einzelne Merkmale der Hattie-Studie: Die Qualität der Lehrer-Schüler-Interaktion im Unterricht

Da Hattie aufzeigt, dass viele der wirkungsmächtigen Faktoren auf der Ebene des Unterrichts liegen, also von den Lehrpersonen selbst beeinflusst werden können, wird die Studie oftmals zum Anlass genommen, über Veränderungen des Unterrichts nachzudenken. Obwohl sich Hatties Studie bei der Auswahl der Merkmale nicht explizit an Theorien zur Unterrichtsqualität orientiert, lassen sich die gefundenen Merkmale nachträglich solchen Theorien und Modellen zuordnen.

Hattie und die Basisdimensionen der Unterrichtsqualität

Im deutschsprachigen Raum ist beispielsweise das *Modell der drei Basisdimensionen* von Unterrichtsqualität sehr bekannt geworden (z. B. Klieme, Lipowsky, Rakoczy & Ratzka 2006, Klieme, Schümer & Knoll 2001). Hier wird angenommen, dass zur Gewährleistung guten Unterrichts drei Bedingungen von besonderer Bedeutung sind (vgl. Abbildung 4):

- ein effektives Classroom Management,
- ein positiv ausgeprägtes Unterrichtsklima und ein unterstützendes Lehrerverhalten
- und die kognitive Aktivierung der Schüler.

Die Stärken dieses Modells liegen darin, dass die Basisdimensionen guten Unterrichts auf einer empirischen Grundlage basieren. Zudem sind in ihnen viele Einzelmerkmale guten Unterrichts enthalten, so dass sie bereits viele Einzelaspekte umfassen und in ihrer Gesamtheit bedeutsame Bereiche der Unterrichtsgestaltung beinhalten.

² Eine Diskussion, wie sich die geringe Wirksamkeit von Merkmalen wie dem Offenen Unterricht erklären lässt, findet sich bei Lotz & Lipowsky (2014) und bei Lipowsky & Lotz in diesem Band.

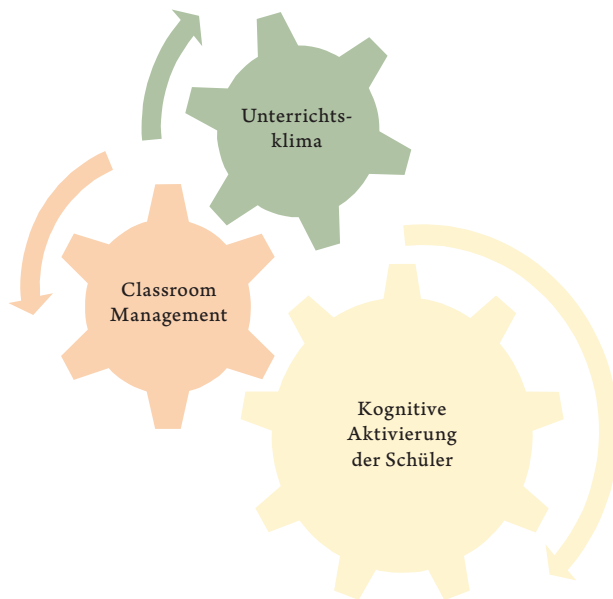


Abbildung 4: Drei Basisdimensionen der Unterrichtsqualität

Da im Modell der Basisdimensionen die einzelnen Qualitätsmerkmale faktorenanalytisch zu übergeordneten Bereichen zusammengefasst wurden, überschneiden sich die drei Basisdimensionen auch weniger stark als die Merkmale in anderen Listen zur Unterrichtsqualität, bei denen lediglich einzelne Merkmale anhand empirischer Studien abgeleitet wurden, aber nicht die Struktur der Merkmale untereinander empirisch überprüft wurde (z. B. Helmke 2009, Meyer 2010).

Eine *effektive Klassenführung* oder ein *effektives Classroom Management* beschreibt die »Fähigkeit einer Lehrperson, im Unterricht einen zügigen, gradlinigen und gut organisierten Stundenablauf zu gewährleisten, Disziplinprobleme schnell und effektiv abzuhandeln sowie die Klassenzimmerprozesse im Gesamten zu überblicken [...], um bei Schwierigkeiten und Problemen rasch und im Idealfall vorausblickend handeln zu können« (Waldis, Grob, Pauli & Reusser 2010, S. 177). Als Rahmenbedingung zur Gewährleistung eines möglichst hohen Ausmaßes an aktiver Lernzeit kommt ihr eine »Schlüsselfunktion im Unterricht« (Weinert 1998, S. 27) zu.

Auch das Unterrichtsklima oder das unterstützende Lehrerverhalten stellt eine Basisdimension von Unterrichtsqualität dar. Gruehn (2000) systematisiert Aspekte des Unterrichtsklimas, indem sie die drei Bereiche Lehrer-Schüler-Beziehung, Schüler-Schüler-Beziehung und Merkmale des Unterrichts unterscheidet. Gabriel (2014) ordnet diesen drei Bereichen folgende Einzelmerkmale zu und beschreibt ein positiv ausgeprägtes Unterrichtsklima dadurch genauer: Die Lehrer-Schüler-Beziehungen

CLASSROOM MANAGEMENT	UNTERRICHTSKLIMA	KOGNITIVE AKTIVIERUNG
<ul style="list-style-type: none"> • Klassenführung (0.52) • Aktive Lernzeit (0.38) • Reduzierung von Unterrichtsstörungen (0.34) 	<ul style="list-style-type: none"> • Lehrer-Schüler-Beziehung (0.72) • Nicht-Etikettieren von Lernenden (0.61) • Klassenzusammenhalt (0.53) 	<ul style="list-style-type: none"> • Feedback (0.73) • Meta-kognitive Strategien (0.69) • Kreativitätsförderung (0.65) • Problemlösen (0.61) • Lerntechniken (0.59) • Fragenstellen (0.46) • Forschendes Lernen (0.31)

Tabelle 2: Zuordnung ausgewählter Hattie-Merkmale zu den Basisdimensionen guten Unterrichts

lassen sich durch die Anerkennung der Schüler durch die Lehrperson charakterisieren, durch Fürsorglichkeit, Herzlichkeit und Wärme, den Einsatz von Lob, Ermutigungen, sachlich-konstruktive Rückmeldungen sowie den positiven Umgang mit Fehlern. Die Schüler-Schüler-Beziehungen sind von gegenseitiger Anerkennung und einem positiven Fehlerklima geprägt. Im Unterricht herrscht insgesamt eine humorvolle Lernatmosphäre und Wettbewerb sowie Leistungsdruck werden weitgehend vermieden.

Die kognitive Aktivierung der Schüler wird von Lipowsky (2015) als die Anregung »zum vertieften Nachdenken und zu einer elaborierten Auseinandersetzung mit dem Unterrichtsgegenstand« (S. 89 f.) beschrieben. Hier wird also gefordert, dass »alle Lernenden zur aktiven Auseinandersetzung mit den Lerninhalten auf einem für sie angemessenen Niveau angeregt werden« (Leuders & Holzäpfel 2011, S. 213), wobei sich der Begriff der Aktivität auf das Denken und nicht auf das Verhalten der Lernenden bezieht.

Analysiert man die Ergebnisse der Hattie-Studie, so finden sich darin viele Merkmale, die sich den drei Basisdimensionen guten Unterrichts zuordnen lassen (vgl. Tabelle 2). Betrachtet man die drei Basisdimensionen im Vergleich, deutet sich an, dass die kognitive Aktivierung der Schüler voraussetzungsreicher als eine effektive Klassenführung und ein unterstützendes Unterrichtsklima ist und daher das vielleicht anspruchsvollste Ziel für Lehrpersonen darstellt. Verschiedene Studien zeigen tatsächlich, dass den meisten Lehrpersonen die kognitive Aktivierung der Schüler schwerer fällt als eine effektive Klassenführung und die Schaffung eines angenehmen Unterrichtsklimas (z. B. Gabriel 2014, Klieme u. a. 2001, Lotz 2015, Pietsch 2010). Daher soll im Folgenden auf

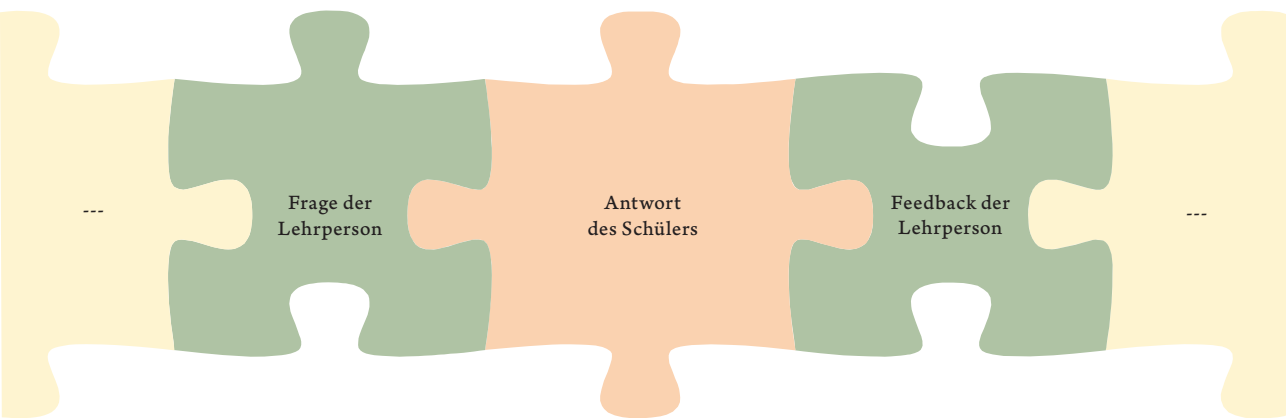


Abbildung 5: Typische Aspekte der Lehrer-Schüler-Interaktionen im Unterricht

ausgewählte Merkmale der Lehrer-Schüler-Interaktion eingegangen werden, die sich der Basisdimension der kognitiven Aktivierung zuordnen lassen.

Zentrale Elemente der Lehrer-Schüler-Interaktion

Brunner und Noack (2010) betonen, dass das Gelingen von Unterricht vor allem von der Qualität der Lehrer-Schüler-Interaktion bzw. der Kommunikation zwischen Lehrperson und Schülern abhängt, wobei davon ausgegangen wird, dass sich das Lehrer- und Schülerverhalten wechselseitig beeinflussen. Für die Kommunikation im Unterricht wurde vielfach nachgewiesen, dass hier ein bestimmtes Muster vorzufinden ist, das sich als Auffordern – Antworten – Reagieren (Bellack, Kliebard, Hyman & Smith 1974) beschreiben lässt (vgl. Abbildung 5). Mehan (1979) prägte dafür den Begriff der I-R-E-Sequenz, also der Aufeinanderfolge von Frage (»Initiation«), Antwort (»Reply«) und Bewertung (»Evaluation«). Daran wird häufig kritisiert, dass die Schüler hierbei lediglich als Stichwortgeber fungieren: »Bereits die Frage der Lehrperson ist [...] begrenzt und die Äußerung des Schülers oder der Schülerin dient nur dazu, den Redefluss der Lehrperson zu unterstützen. Die Lernenden bekommen keine Gelegenheit ihre eigenen Gedanken auszuführen oder zu erläutern« (Kobarg & Seidel 2003, S. 174).

Zentrale Bestandteile der Kommunikation im Unterricht sind daher das Stellen bedeutsamer Fragen und das Erteilen informativen und konstruktiven Feedbacks, um die Schüler im Unterricht kognitiv zu aktivieren, sie also zu einer vertieften Auseinandersetzung mit den Lerngegenständen anzuregen. Wie dies geschehen kann, wird daher in den folgenden beiden Abschnitten thematisiert.

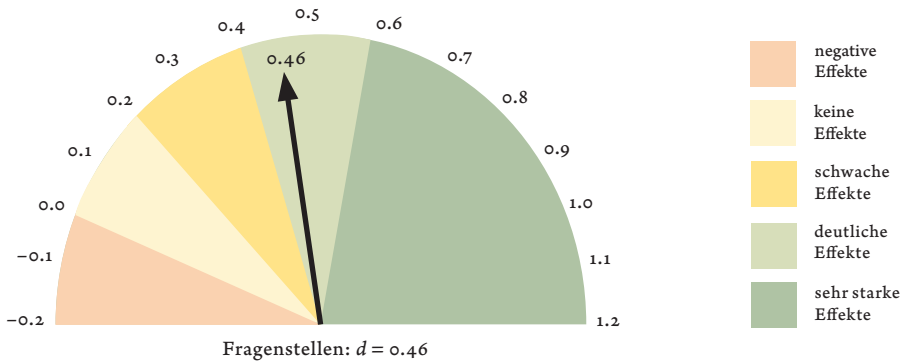


Abbildung 6: »Effektstärke-Strahl« für das Stellen von Fragen (vgl. Hattie 2009, 2013)

Was sind gute Fragen?

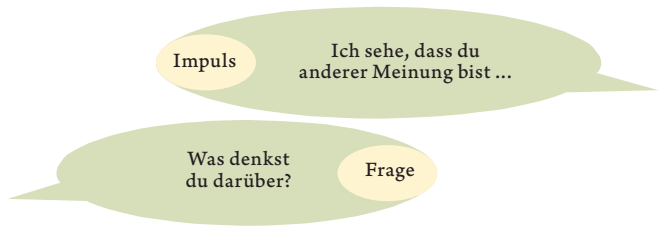
Auf Basis von sieben einbezogenen Metaanalysen, welche die Ergebnisse von insgesamt 211 Einzelstudien beinhalten, ergab sich in der Hattie-Studie eine generelle Wirksamkeit des Fragenstellens von $d = 0.46$ (vgl. Abbildung 6). Fragen befinden sich damit auf dem 53. Rang der von Hattie untersuchten Merkmale. Damit liegen sie noch im Bereich deutlicher Effekte. Dass der Wert für das Stellen von Fragen nicht höher liegt, könnte darauf zurückzuführen sein, dass vor allem die Art der Fragen entscheidend ist. Auch Hattie (2013) stellt fest, dass die Effekte von Fragen variieren, wobei die Höhe der Effekte am ehesten vom untersuchten Fragetyp abhängt.

Warum ist das Stellen von Fragen bedeutsam?

Lehrerfragen dienen dazu, den Unterricht zu strukturieren und zu steuern, die Aufmerksamkeit der Lernenden auf relevante Aspekte des Unterrichts zu lenken, das Vorwissen zu aktivieren, die Lernenden anzuregen und herauszufordern, Lernwege, (Miss-)Konzepte und (Fehl-)Vorstellungen offenzulegen, den Wissensstand der Lernenden zu ermitteln, Unterrichtsergebnisse zu sichern oder manchmal auch dazu, die Lernenden zu disziplinieren. (Lipowsky 2015, S. 80)

Fragen können von der Lehrperson somit als didaktisches Mittel eingesetzt werden, um die Schüler dazu anzuregen, über bestimmte Sachverhalte nachzudenken (Aebli 1987) und um zwischen bereits vorhandenem Wissen und neuen Lerninhalten zu vermitteln (Dubs 2008). Damit dienen Fragen also der Steuerung, Präzisierung sowie Intensivierung von Lernprozessen (zsf. Petersen & Sommer 1999) und sind eine wichtige Maßnahme zur kognitiven Aktivierung.

Studien zur Häufigkeit von Lehrerfragen ermittelten, dass Lehrpersonen durchschnittlich etwa zwei Fragen pro Minute stellen (Lotz 2015, Niegemann & Stadler



2001). Daher werden Fragen auch als »die häufigste Einzelaktivität im Unterricht, das häufigste Element des Lehrerhandelns« (Aschersleben 1999, S. 97) bezeichnet. Gleichzeitig wird aufgezeigt, dass die Komplexität der gestellten Fragen eher gering ist (z. B. Lotz 2015). Da es für die kognitive Aktivierung der Schüler jedoch sehr viel wichtiger ist, »gute« Fragen zu stellen anstatt möglichst viele Fragen zu stellen, fokussieren die folgenden Abschnitte auf die Qualität von Lehrerfragen.

Was ist besser: Fragen oder Impulse?

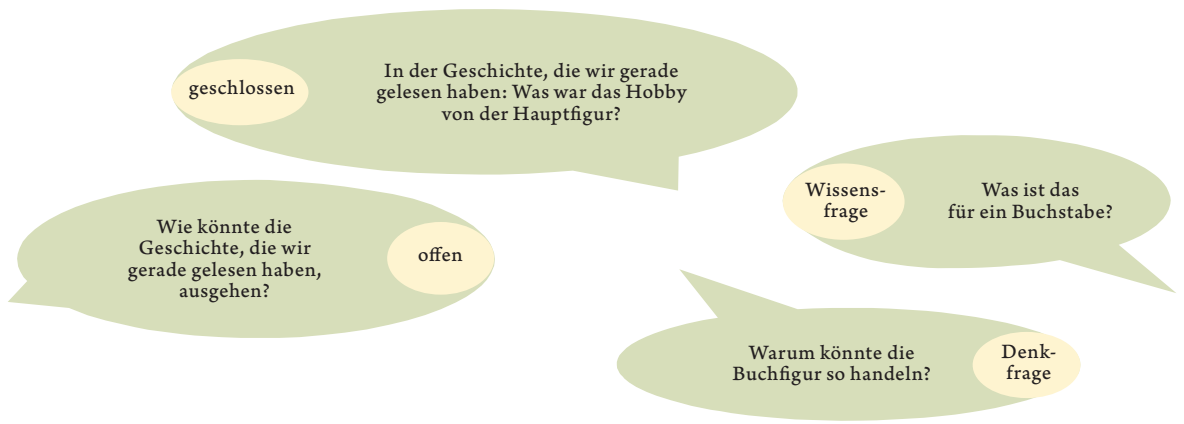
Als Fragen können sowohl alle Arten von W-Fragen (Was, Wie, Warum usw.), alle Sätze, die mit einem Fragezeichen enden als auch andere Satzformen, die den Schüler zu einer Antwort auffordern sowie nonverbale Impulse zählen (z. B. Petersen & Sommer 1999, Renkl 1991). Petersen und Sommer (1999) verstehen daher unter einer Lehrerfrage Folgendes: »Die Lehrerfrage fordert den/die Schüler zu einer Antwort auf. Sie kann grammatisch als Frage oder Aufforderung formuliert sein« (S. 21).

Aufgrund der starken Kritik an Lehrerfragen, welche das Denken der Schüler zu stark einengen würden (Gaudig 1909), wurde vielfach gefordert, Fragen durch Impulse zu ersetzen, da diese mehr Möglichkeiten zum freien Antworten bieten würden (z. B. Riedl 2010). Dass diesen Anspruch aber auch offen formulierte Fragen erfüllen können, wird inzwischen vielfach anerkannt (z. B. Aebli 1987, Aschersleben 1999, Nuding 2009, Riedl 2010, Steindorf 2000):

Lehrerfragen müssen nicht »schlecht«, Impulse nicht »gut« sein. Wie es treffend formulierte Fragen gibt, so auch leere, richtungslose und floskelhafte Lehrimpulse. Beide Lenkungsinstrumente haben ihre spezifischen Möglichkeiten und sollten daher in der Unterrichtsgestaltung nebeneinander Verwendung finden. Ihre Gemeinsamkeit liegt darin, da(ss) sie die gedankliche Tätigkeit der Schüler anzuregen und fortzubewegen vermögen. (Steindorf 2000, S. 148)

Was ist besser: geschlossene oder offene Fragen?

Während geschlossene oder konvergente Fragen nur eine richtige Antwort haben, besitzen offene oder divergente Fragen mehrere mögliche Antworten (z. B. Flanders 1970, Kobarg & Seidel 2003, Pauli 2006). Divergente Fragen können eher zu kreativem Denken anregen. Demgegenüber erfordern konvergente Fragen »regelmäßiges, eng strukturiertes, auf eine bestimmte Lösung orientiertes Denken« (Sommer 1981, S. 30).



Da beide Arten des Denkens je nach Unterrichtsziel relevant sein können, sollten Lehrpersonen reflektieren, welche Art des Denkens sie in der jeweiligen Situation fördern möchten und die Fragen dementsprechend zielorientiert auswählen. Eine generelle Überlegenheit einer der beiden Fragearten kann nicht nachgewiesen werden (Knapp 1985, Tsui, Marton, Mok & Ng 2004). Der Vorteil an offen gestellten Fragen ist allerdings, dass mehr Schüler aktiv in das Unterrichtsgespräch einbezogen werden können, indem verschiedene Antworten gesammelt und aufeinander bezogen werden können.

Wie schwierig sollten Fragen sein?

Bei der Frage nach der kognitiven Komplexität von Fragen bezieht sich die Forschung oft auf die Lernzieltaxonomie von Bloom und Kollegen (Bloom u. a. 1956, 1976), nach der man Wissens-, Verständnis-, Anwendungs-, Analyse-, Synthese- und Bewertungsfragen voneinander unterscheiden kann. Dabei werden Wissens- und Verständnisfragen meist als lower-order-questions (Fragen niedrigerer kognitiver Komplexität) bezeichnet, wohingegen ab den Analysefragen von Fragen höherer Ordnung gesprochen wird (higher-order-questions; Klinzing & Klinzing-Eurich 1982).

Zur Vereinfachung werden oft nur zwei Kategorien unterschieden, die beispielsweise als Wissens- und Denkfragen bezeichnet werden (Cecil 2008, Frey & Fisher 2007, Gayle u. a. 2006, Niegemann & Stadler 2001). Aufgrund des unterschiedlichen Vorwissens der Schüler innerhalb einer Klasse kann eine Frage, die von der Lehrperson als Wissensfrage intendiert war, für einige Schüler, die nicht über das jeweilige Vorwissen verfügen, allerdings auch als Denkfrage fungieren. Während in der Hattie-Studie lediglich der Effekt des Fragenstellens im Allgemeinen untersucht wurde, ohne dabei die Art der Frage zu berücksichtigen, konnten Gayle und Kollegen (2006) durch eine kleinere Metaanalyse zeigen, dass kognitiv anspruchsvolle Fragen zu einem höheren Niveau der Schülerantworten führen ($d = 0.41$).

Für das kognitive Niveau von Fragen wird empfohlen, dass der Unterricht sich nicht auf eine der beiden Fragearten (Wissens- oder Denkfragen) festlegen sollte, sondern dass je nach Unterrichtsziel sowie nach den Bedingungen der Klasse Wissens- und Denkfragen in einem angemessenen Verhältnis vorkommen sollten. Wichtig ist vor allem, sich als Lehrperson bewusst mit den Fragen, die man im Unterricht stellen möchte, auseinanderzusetzen und sie gezielt und reflektiert einzusetzen.

Empirische Studien zeigen aber, dass Lehrpersonen vorwiegend einfache Reproduktionsfragen stellen und dass nur sehr wenig sogenannte higher-order-questions vorkommen (Knierim 2008, Kobarg & Seidel 2007, Lotz 2015, Wilen 1982, Wragg & Brown 2001). Je nach genauer Definition der Fragearten stehen nämlich in den einzelnen Studien 4 bis 20 % Fragen höheren kognitiven Niveaus 40 bis 90 % Fragen mit niedrigerem Niveau gegenüber (Lipowsky 2015).

Wie viel Wartezeit ist nach Fragen nötig?

Gerade bei Fragen höheren kognitiven Niveaus sollte man davon ausgehen, dass Lehrpersonen Schülern genügend Wartezeit zur Verfügung stellen, bevor eine Antwort erwartet wird. Beobachtungsstudien zeigen aber, dass dies nur sehr selten der Fall ist. In einer frühen Studie zu diesem Thema zeigte Rowe (1974), dass die durchschnittliche Wartezeit zwischen dem Stellen einer Frage und dem Aufrufen des ersten Schülers lediglich 0,9 Sekunden beträgt und auch spätere Studien replizierten den Befund sehr kurzer Wartezeiten nach Lehrerfragen immer wieder (zsf. Heinze & Erhard 2006, Helmke u. a. 2008, Lotz 2015).

Obgleich die Ergebnisse zum Zusammenhang zwischen Wartezeit und Lernerfolg auch eher heterogen sind (zsf. Lipowsky 2015), wird generell eine Wartezeit von min-

Richten Sie die Art Ihrer Fragen bewusst am Unterrichtsziel aus:

- Stellen Sie offene Fragen, wenn es darum geht, verschiedene Ideen zu sammeln.
- Stellen Sie geschlossene Fragen, um den Wissensstand der Schüler zu überprüfen.

Ermöglichen Sie möglichst vielen Schülern, etwas zum Gespräch beizutragen:

- Wechseln Sie den Schwierigkeitsgrad von Fragen ab, so dass Schüler unterschiedlicher Leistungsniveaus angesprochen werden.
- Stellen Sie Fragen, die auch die leistungsstarken Schüler noch zum Denken herausfordern. Geben Sie den Schülern ausreichend Zeit zum Überlegen, bevor Sie den ersten Schüler aufrufen.

Regen Sie Ihre Schüler an, selbst Fragen zu stellen.

Abbildung 7: Praxistipps – Was sind gute Fragen?

destens drei bis fünf Sekunden empfohlen (Heinze & Erhard 2006, Rowe 1974, Tobin 1987), da dies zu elaborierteren Schülerantworten, höherer Aufmerksamkeit sowie mehr Beteiligung der Schüler führt (zsf. Gage, Berliner & Bach 1996, Heinze & Erhard 2006). Besonders wichtig ist eine etwas längere Wartezeit für schwächere Schüler, um auch ihnen die Möglichkeit zu geben, über die Frage nachzudenken.

In Abbildung 7 werden die wichtigsten Aspekte zur Qualität von Lehrerfragen zusammengefasst. Auf Lehrerfragen folgen üblicherweise Schülerantworten, welche anschließend wieder durch die Lehrperson evaluiert werden, indem diese den Schülern ein Feedback gibt. Auf die Frage, was gutes Feedback ausmacht, wird im Folgenden eingegangen.

Was ist gutes Feedback?

In Hatties Feedbackanalysen gingen insgesamt 23 Metaanalysen ein, die auf insgesamt 1287 Einzelstudien basieren. Feedback erreicht von allen in der Studie analysierten Merkmalen mit $d = 0.73$ den zehnten Rangplatz und gehört damit zu den lernwirksamsten Faktoren (vgl. Abbildung 8). Nach Hattie und Timperley (2007) stellt Feedback eine Information dar, die Aspekte der Leistung oder des Verständnisses fokussiert. Das Ziel sollte dabei – insbesondere bei noch nicht fehlerfrei ausgeführten Handlungen oder Lösungen – die Verringerung der Diskrepanz zwischen der aktuellen Leistung und dem erwünschten Lernziel sein (Hattie & Timperley 2007, Sadler 1989). Hattie (2013) weist darauf hin, dass die durch die einzelnen Metaanalysen berechneten Effektstärken stark variieren, was darauf hindeutet, dass bestimmte Arten von Feedback wirksamer als andere sind und dass die Wirksamkeit von Feedback von weiteren Bedingungen abhängig ist.

Bevor aber darauf eingegangen wird, wie Feedback effektiv gegeben werden kann, wird noch auf die generelle Bedeutsamkeit von Rückmeldungen für das Lernen eingegangen.

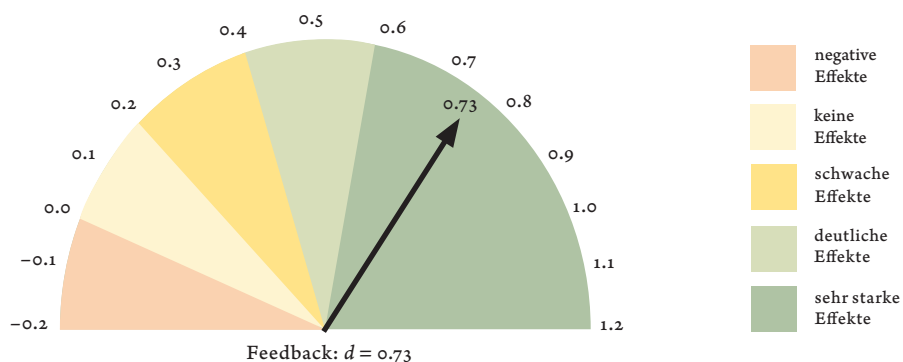


Abbildung 8: »Effektstärke-Strahl« für das Erteilen von Feedback (vgl. Hattie 2009, 2013)

Warum ist das Geben von Feedback bedeutsam?

Rückmeldungen können sich sowohl auf die Motivation als auch auf die Leistung auswirken, indem sie kognitive, motivationale und metakognitive Prozesse aktivieren (Huth 2004, Narciss 2004, Rakoczy, Klieme, Bürgermeister & Harks 2008, Vollmeyer & Rheinberg 2005). Will man Schüler langfristig zu selbstreguliertem Lernen führen (Butler & Winne 1995, van den Boom, Paas & van Merriënboer 2007), sind Rückmeldungen zentral, damit die Schüler frühzeitig lernen, eine Bewertung des eigenen Lernprozesses vorzunehmen, diesen zu reflektieren und zukünftige Lernaktivitäten effizienter zu gestalten. Denn Aufgaben oder Fragen können nur dann ihre Lernwirksamkeit entfalten, wenn sie vom Schüler zunächst aktiv bearbeitet werden und der Schüler anschließend eine sinnvolle Rückmeldung zur Aufgabenbearbeitung erhält (Jacobs 2002). Allerdings ist dazu die Art der Rückmeldung entscheidend (Bangert-Drowns u. a. 1991, Kluger & DeNisi 1996). Daher wird in den folgenden beiden Abschnitten auf zwei Aspekte der Qualität von Feedback eingegangen. Zunächst wird erläutert, welche Auswirkungen Lob hat und wie eine motivationsförderliche Rückmeldepraxis aussehen kann. Anschließend wird betrachtet, welche Informationen gutes Feedback enthalten sollte.

Wie sollte man Schüler loben?

Während Feedback als eine Rückmeldung verstanden wird, »die sich auf die Leistung oder das Verständnis des Lernenden bezieht, diesen über die Richtigkeit seiner Antwort bzw. seiner Aufgabenlösung informiert [...] oder ihm inhaltliche und/oder strategische Hilfen und Informationen zu seinem Bearbeitungsprozess zur Verfügung stellt« (Lipowsky 2015, S. 82 in Anlehnung an Mory 2004), wird Lob davon oftmals abgegrenzt, da es eher eine einfache Bekräftigung enthält (z. B. »prima«, »gut gemacht«) und nicht immer konkret auf die erbrachte Leistung bezogen ist. Wird jedoch ein sehr konkretes und differenziertes Lob erteilt, so sind die Grenzen zu Feedback fließend.

Besonders für die Motivation von Schülern ist angemessenes und sinnvolles Lob bedeutsam. Ein Lob kann das Verhalten des Schülers verstärken, also dessen zukünftige Auftretenswahrscheinlichkeit erhöhen. Dies funktioniert jedoch meist nicht automatisch: Ob das Lob die erwünschte Wirkung erzielt, ist davon abhängig, ob es spezifisch genug ist, um vom Schüler als glaubwürdig wahrgenommen zu werden. Daher empfiehlt Richert (2005): »Pädagogisch angemessenes Lob sollte nicht nur spontan, aufrichtig und glaubwürdig wirken, sondern spezifische Aspekte der erbrachten Leistung fachlich und personal bewerten« (S. 17). Ein solch spezifisches Lob, welches die Aspekte des gelobten Verhaltens explizit benennt, kommt allerdings in der Unterrichtspraxis nur selten vor (z. B. La Paro, Pianta & Stuhlman 2004).

Zu Effekten von Lob auf die Leistung von Schülern existieren uneindeutige Befunde (zsf. Henderlong & Lepper 2002, Richert 2005). Eine interessante Studie zu diesem Thema veröffentlichten Brummelman, Thomaes, de Castro, Overbeek und Bushman (2014). Sie ließen 240 Kinder zwischen acht und zwölf Jahren zunächst ihr Selbstwertgefühl einschätzen und anschließend ein berühmtes Gemälde abmalen. Anschließend wurde ihnen eine fiktive Rückmeldung erteilt, wobei drei Bedingungen realisiert wurden: Die erste Gruppe erhielt ein überschwängliches Lob (»Das ist ein unglaublich schönes Bild!«), die zweite Gruppe ein »normales« Lob (»Das ist ein schönes Bild.«) und die dritte Gruppe kein auf das Bild bezogenes Lob. Dann wurden die Kinder aufgefordert, Figuren nachzuzeichnen, wobei sie zwischen einfachen und komplexen Figuren auswählen durften. Die Wahl der Figuren wurde als Indikator für die Suche nach Herausforderungen herangezogen. Die Ergebnisse zeigen, dass überschwängliches Lob (»inflated praise«) speziell bei Kindern mit geringem Selbstwertgefühl eher zu Verunsicherung und Vermeidungsverhalten führt. Denn Kinder mit einem geringen Selbstwertgefühl, die überschwänglich gelobt wurden, wählten deutlich seltener die komplexen Figuren zum Nachzeichnen aus als Kinder, die ebenfalls ein geringes Selbstwertgefühl hatten, vorab aber nicht überschwänglich gelobt wurden. Die Autoren der Studie geben als Erklärung für diesen Befund an, dass Kinder mit einem geringen Selbstwertgefühl bei überschwänglichem Lob vor künftigen Herausforderungen möglicherweise deshalb zurückschrecken, weil sie davon ausgehen, die vorherige Leistung, welche ja schon stark gelobt wurde, kaum noch einmal erreichen zu können. Mit »Aufhören, wenn es am schönsten ist« könnte man dieses Verhalten etwas überspitzt überschreiben. Bestimmte Schüler entscheiden sich nach einem überschwänglichen Lob also für ein geringeres Risiko. Dieses Ergebnis verdeutlicht, dass Lehrpersonen sensibel auf die jeweiligen Schüler eingehen sollten, indem sie sich fragen, welche Art der Rückmeldung ein Schüler braucht, um weiterzukommen. Gleichzeitig sollten sie antizipieren, was die Art des Feedbacks beim jeweiligen Schüler auslösen könnte. Für Pädagogen, aber auch für Eltern, ist dieser Befund auch deshalb so bedeutsam, da sie Kinder mit einem geringen Selbstwertgefühl besonders motivieren und ihnen Mut zusprechen möchten und wahrscheinlich auch deshalb in solchen Fällen zu überschwänglichem Lob neigen. Nach den Ergebnissen der Studie von Brummelman und Kollegen scheint ein solch überschwängliches Loben aber die falsche Strategie zu sein, um Kinder mit einem geringen Selbstwertgefühl anzuspornen und zu motivieren.

Unter dem Stichwort »Paradoxe Effekte von Lob und Tadel« zeigt die Forschung, dass Lob unter bestimmten Bedingungen auch schädlich für das Selbstkonzept eines Schülers sein kann, dann nämlich, wenn der Schüler für Selbstverständlichkeiten gelobt wird und wenn er mitbekommt, dass Mitschüler nicht für diese Selbstverständlichkeiten gelobt werden (zsf. z. B. Rheinberg & Vollmeyer 2010). Der gelobte Schüler

kann dieses Lob als Indikator für das geringe Lehrerzutrauen ihm gegenüber interpretieren. Gelobt werden sollte ein Schüler vor allem dann, wenn er seine Leistungen verbessert hat. Darauf wird im folgenden Abschnitt eingegangen. Umgekehrt beschreibt das Phänomen des paradoxen Effekts von Tadel, dass ein Schüler, der bei einer schwierigen Aufgabe von der Lehrperson eine negative Rückmeldung bekommt, dies darauf zurückführen kann, dass die Lehrperson ihm besonders viel zutraut und viel von ihm erwartet (vgl. auch das Filmbeispiel »Rhythm Is It« im Fazit des Beitrags).

Welche Bezugsnormen sollte man anlegen?

Ein wichtiger Aspekt für das Erteilen von motivations- und selbstkonzeptförderlichem Feedback ist nämlich die Wahl einer angemessenen Bezugsnorm. Unter Bezugsnormen werden Standards verstanden, anhand derer ein Resultat verglichen wird, um es zu bewerten (z. B. Rheinberg 2006, 2008, Rheinberg & Fries 2006, 2010). Dabei können drei Bezugsnormen unterschieden werden:

- Beim Anlegen einer sozialen Bezugsnorm vergleicht die Lehrperson die Leistung eines Schülers mit den Leistungen anderer Schüler. Ein Feedback, welches eine explizite Orientierung an der sozialen Bezugsnorm enthält, wäre beispielsweise: »Sarah, du musst noch mehr üben. Schau mal, die Marie hat nur halb so viele Fehler im Diktat gemacht.«
- Bei der Verwendung einer sachlichen Bezugsnorm wird die erzielte Leistung hingegen mit einem vorab festgelegten Standard verglichen. Dies ist beispielsweise der Fall, wenn bei einer Klausur vorab festgelegt wird, wie viel Prozent der Gesamtpunkte erreicht werden müssen, um die Leistung als bestanden oder nicht bestanden zu bewerten. Aber auch ein Feedback, welches dem Schüler zurückmeldet, ob er beispielsweise einen Text richtig vorgelesen hat, enthält die Orientierung an der sachlichen Bezugsnorm (z. B. »Joshua, du hast den Text ohne Fehler vorgelesen.«). Ein weiteres Beispiel für die Verwendung sachlicher Bezugsnormen stellt die Definition der Schulnoten der Kultusministerkonferenz (KMK 1968³) dar. Als »sehr gut« (»1«) soll hier nämlich bewertet werden, »wenn die Leistung den Anforderungen in besonderem Maße entspricht«. Die Note »ausreichend« (»4«) soll beispielsweise vergeben werden, »wenn die Leistung zwar Mängel aufweist, aber im Ganzen den Anforderungen noch entspricht«. In der Unterrichtspraxis wird die Vergabe von Noten allerdings häufig auch von sozialen Bezugsnormen beeinflusst, da die Lehrperson die zu erreichenden Standards nicht generell festlegt, sondern an den Leistungsstand ihrer Klasse anpasst. So kommt es vor, dass in leistungsstarken

3 Vgl. z. B. <http://www.schule.de/22410/33,83203.htm>

	Stabile Ursachen	Variable Ursachen
	FÄHIGKEIT	ANSTRENGUNG
Internale Ursachen	»Ich bin einfach gut in Deutsch.«	»Ich habe viel geübt.«
	AUFGABENSCHWIERIGKEIT	GLÜCK/PECH/ZUFALL
Externale Ursachen	»Die Klausur war sehr schwierig«	»Diesmal hatte ich einfach Pech mit den Fragen.«

Tabelle 3: Vierfelderschema möglicher Attributionen (vgl. Weiner u. a. 1972)

Klassen faktisch bessere Leistungen nötig sind, um eine »sehr gute« Note zu erreichen als in leistungsschwächeren Klassen (Ingenkamp 1995).

- Die dritte Möglichkeit, Standards für den Vergleich erzielter Ergebnisse heranzuziehen, stellt die individuelle Bezugsnorm dar. Hier wird das aktuelle Resultat eines Schülers mit dessen vorherigen Leistungen im gleichen Bereich verglichen, indem die Lehrperson dem Schüler beispielsweise zurückmeldet: »Daniel, dieses Mal hast du schon 80 von 100 Wörtern im Diktat richtig geschrieben, das letzte Mal waren es noch 60 von 100 Wörtern. Du hast bestimmt zu Hause viel geübt, oder?«

Anhand der Beispiele dürfte bereits deutlich werden, welche Prozesse bei Schülern durch die Verwendung der unterschiedlichen Bezugsnormen ausgelöst werden können. Generell wird empfohlen, sowohl die sachliche als auch die individuelle Bezugsnorm zu verwenden, wohingegen Rückmeldungen, die sich explizit an der sozialen Bezugsnorm orientieren, vermieden werden sollten. Sie forcieren den Vergleich der Schüler untereinander, was insbesondere für schwächere Schüler meist nicht motivationsförderlich ist (z. B. Mandl & Kopp 2008). Ein Schüler wie Ben, der in den meisten Schulfächern Schwierigkeiten hat, würde durch Rückmeldungen, die sich an der sozialen Bezugsnorm orientieren, meist den Eindruck gewinnen, dass er schlechter ist als die anderen Schüler und eigene Lernfortschritte, die selbstverständlich auch von schwächeren Schülern erzielt werden, nicht bemerken.

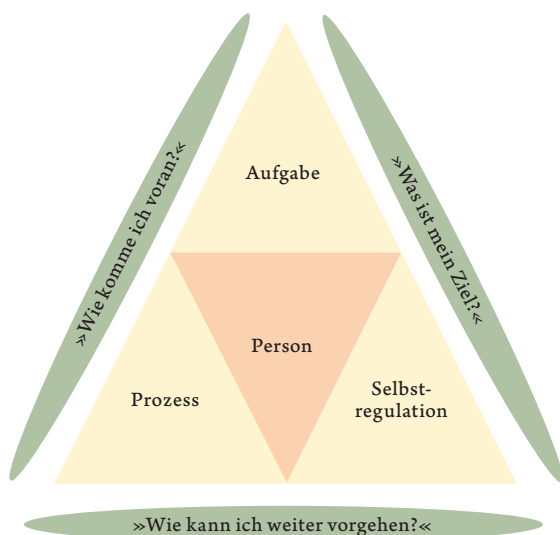


Abbildung 9: Feedback-Modell nach Hattie & Timperly (2007)

Erfahrungen bei der Bearbeitung von Aufgaben und darauf folgende Selbst- oder Fremdbewertungen können demnach eine Reihe kognitiver, motivationaler sowie emotionaler Effekte haben. Wenn Schüler eine Leistung erbringen und hierzu eine Rückmeldung bekommen, kann dies verschiedene Attributionen auslösen, also Überlegungen zu den Ursachen des Erfolgs oder Misserfolgs (z. B. Brunstein & Heckhausen 2006, Mandl & Kopp 2008). Rückmeldungen anhand der sozialen Bezugsnorm enthalten für den Schüler kaum hilfreiche Informationen und implizieren für ihn eventuell, dass seine schlechte Leistung im Vergleich zu der der Mitschüler an seinen mangelnden Fähigkeiten liegt. Hierbei handelt es sich nach dem in Tabelle 3 dargestellten Vierfelderschema möglicher Ursachenzuschreibungen um eine internale und stabile Attribution, die insbesondere bei Misserfolgen stark demotivieren kann, da die Fähigkeit dem Schüler als nicht veränderbar erscheint. Dies könnte ihn bei der zukünftigen Wahl von Aufgaben entmutigen und dazu führen, dass er Herausforderungen eher aus dem Weg geht. Dadurch wiederum wird die Chance auf zukünftige Erfolgserlebnisse minimiert (z. B. Brunstein & Heckhausen 2006).

Im Gegensatz dazu bieten Rückmeldungen, die sich an der individuellen Bezugsnorm orientieren, Schülern die Möglichkeit, eigene – wenn auch kleinere – Lernfortschritte zu erkennen. Die Schüler können dadurch bemerken, dass sich ihr Lernverhalten auszahlt. Wenn sie einen Erfolg erleben, führen sie diesen auf eine erhöhte Anstrengung zurück (internale, variable Ursachenzuschreibung), insbesondere wenn die Lehrperson durch ihr Feedback solch eine motivationsförderliche Ursachenzu-

schreibung forciert (z. B. »Du hast diesmal nur noch halb so viele Fehler gemacht wie im letzten Test. Ich freu mich, dass du so viel geübt hast.«).

Wie kann man informatives Feedback geben?

Damit Feedback seine Lernwirksamkeit entfalten kann, muss es für den Schüler relevante Informationen enthalten (vgl. Abbildung 9). Insbesondere bei fehlerhaften Antworten oder Lösungen braucht der Schüler eine Auskunft, dass (und eventuell warum) seine Lösung nicht richtig ist sowie Hinweise darüber, wie er der Lösung näher kommen kann. Differenziertes Feedback wird oft auch als elaboriert bezeichnet, wobei viele Studien zeigen, dass solches Feedback deutlich seltener vorkommt als einfa-

Feedback sollte zunächst auf drei Fragen antworten:

1. Feed-Up

»Where am I going?«/»Wo will ich eigentlich hin? Was ist mein Ziel?«

Beispiel: »Überlege noch einmal: Was möchtest du aus dem Text herausfinden?«

2. Feed-Back

»How am I going?«/»Wie komme ich voran?«

Beispiel: »Jetzt weißt du schon eine ganze Menge über die Hauptfigur.«

3. Feed-Forward

»Where to next?«/»Was sind die nächsten Schritte? Wie kann ich weiter vorgehen?«

Beispiel: »Was könntest du tun, um noch mehr über die Hauptfigur rauszufinden? Hast du dir schon bestimmte Textstellen farblich markiert?«

Jedes Feedback kann sich weiterhin auf die folgenden vier Ebenen beziehen:

1. Feedback bezogen auf die Aufgabe

Der Schüler erhält hier eine Rückmeldung über die Richtigkeit seiner Antwort oder Lösung, welche – wie im angegebenen Beispiel – auch implizit erfolgen kann.

Beispiel: »Schau dir das Ergebnis nochmal genau an!«

2. Feedback bezogen auf den Prozess der Aufgabenbearbeitung

Rückmeldungen auf dieser Ebene nehmen den Lernprozess oder den Lösungsweg in den Blick. Sie enthalten oftmals auch Hilfen zum weiteren Vorgehen.

Beispiel: »Wie bist du beim Rechnen denn vorgegangen? Erkläre mir mal deinen Rechenweg.«

3. Feedback bezogen auf die Selbstregulation

Feedback zur Selbstregulation fokussiert die Planung, Überwachung und Bewertung des Arbeitsprozesses durch den Schüler.

Beispiel: »Was könntest du denn tun, um die Lösung noch einmal zu kontrollieren?«

4. Feedback bezogen auf die Person des Schülers

Feedback mit Bezug zur Person des Schülers stellt dem Schüler kaum hilfreiche Informationen zur Verfügung.

Beispiel: »Prima!« oder »Du bist ein guter Rechner!«

Abbildung 10: Arten informativer Feedbacks nach Hattie und Timperly (2007)

ches Feedback, welches beispielsweise lediglich die Information zur Richtigkeit der Lösung enthält (z. B. Huth 2004, Lotz 2015, Mory 2004, Narciss 2006). Eine hilfreiche Unterscheidung verschiedener Arten informativen Feedbacks entwickelten Hattie und Timperley (2007; Abbildung 10).

Obwohl undifferenziertes, auf die Person des Schülers bezogenes Feedback zur Person am wenigsten effektiv ist (Hattie & Timperley 2007, Kluger & DeNisi 1998), wird in der Unterrichtspraxis gerade diese Art von Feedback häufig verwendet (Bond, Smith, Baker & Hattie 2000). Führt man die zahlreichen Anforderungen auf, die gute Rückmeldungen erfüllen sollten, überrascht dies nicht, hat die Lehrperson im Unterrichtsprozess doch nur wenig Zeit, jedem einzelnen Schüler immer solch ein differenziertes Feedback zu geben, wie dies auf der Basis der oben genannten Anforderungen sinnvoll erscheint. Hält man sich aber gleichzeitig die hohe Bedeutung von Feedback für das Lernen sowie die Motivation der Schüler vor Augen, spricht dies dafür, sich im Unterricht Freiräume zu schaffen, die es ermöglichen, immer wieder einzelnen Schülern ausführliches Feedback zu geben. Besonders wichtig ist es dabei, mit Schülerfehlern sinnvoll umzugehen. Dabei macht aber Weinert (1999) auf einen wichtigen Unterschied aufmerksam: Isolierte Flüchtigkeitsfehler werden abgegrenzt von solchen Fehlern, die durch fehlendes Regelwissen zustande kommen. Nur bei dieser zweiten Art sollten »remediale didaktische Bemühungen« (S. 103) erfolgen, während kleinere unsystematische Fehler keiner aufwändigen Korrektur bedürfen. Lediglich wenn ein Schüler vermehrt dadurch auffällt, dass er immer wieder Flüchtigkeitsfehler macht, bietet es sich an, dem nachzugehen.

Der Unterschied soll am Beispiel der schriftlichen Addition verdeutlicht werden. Abbildung 11 zeigt jeweils die Lösung dreier unterschiedlicher Schüler beim Lösen von drei Aufgaben. Betrachtet man die Aufgabenlösungen des ersten Schülers, Bastian, so fällt auf, dass er beim Übertrag Fehler macht, da er diesen allgemein nicht berücksichtigt. Da bei der zweiten Aufgabe kein Übertragen nötig ist, löst er diese Aufgabe richtig. Fehler beim Übertrag sind nach Gerster (2012) die häufigsten Fehler beim schriftlichen Addieren. Um Bastian hier sinnvolles Feedback geben zu können, muss man als Lehrperson zunächst die Fehlerursache identifizieren. Hierzu bietet es sich an, Bastian laut rechnen zu lassen, um der Fehlerursache auf die Spur zu kommen. Um Bastian dann zu verdeutlichen, dass an seinen Lösungen etwas nicht stimmen kann, könnte man ihn eine Überschlagsrechnung machen lassen. Bei der ersten Aufgabe würde er dann eventuell bereits erkennen, dass allein 300 plus 400 bereits 700 ergibt. Nimmt man nun noch 65 und 57 dazu, ist das Ergebnis 712 auf jeden Fall zu klein. Anschließend wäre es sinnvoll, Bastian das Verfahren der schriftlichen Addition noch einmal zu erklären und dabei besonders darauf einzugehen, dass die einzelnen Stellen nicht unabhängig voneinander sind. Hier kann es sinnvoll sein, die Aufgaben noch

einmal stärker zu strukturieren, indem man über die Zahlen die bekannten Abkürzungen »H«, »Z« und »E« für Hunderter, Zehner und Einer ergänzt. Zusätzlich kann ein Zurückgehen auf nicht-schriftliche Rechenverfahren sinnvoll sein, indem Bastian einige Aufgaben mit Zehnerübergängen im 100er-Zahlenraum rechnet. Darauf aufbauend kann noch einmal wiederholt werden, dass auch beim schriftlichen Rechnen – sobald die Summe zweier Ziffern mindestens 10 beträgt, die »1« in die nächste Stelle übertragen werden muss. Für die Notiz des Übertrags ist es dann – insbesondere am Anfang – sinnvoll, eine gesamte Zeile freizulassen (Gerster 2012).

Aufgabenlösungen von Bastian

$$\begin{array}{r}
 365 \\
 + 457 \\
 \hline
 712
 \end{array}
 \qquad
 \begin{array}{r}
 24 \\
 + 13 \\
 \hline
 37
 \end{array}
 \qquad
 \begin{array}{r}
 28 \\
 + 15 \\
 \hline
 33
 \end{array}$$

Bastian löst die zweite Aufgabe komplett richtig. Auch bei den beiden anderen Aufgaben addiert er die einzelnen Ziffern immer richtig. Wenn das Ergebnis größer als 9 ist, vergisst er aber den Übertrag der 1 auf die nächste Stelle. Dadurch löst er die Aufgaben 1 und 3 fehlerhaft.

Aufgabenlösungen von Daliah

$$\begin{array}{r}
 365 \\
 + 457 \\
 \hline
 822
 \end{array}
 \qquad
 \begin{array}{r}
 24 \\
 + 13 \\
 \hline
 37
 \end{array}
 \qquad
 \begin{array}{r}
 28 \\
 + 15 \\
 \hline
 42
 \end{array}$$

Daliah löst zwei Aufgaben richtig. Bei Aufgabe 3 erhält sie das Ergebnis 42 statt 43, was darauf hindeutet, dass sie sich beim Berechnen von $8 + 5$ um 1 verrechnet hat. Da dies nur einmalig vorkommt, könnte es sich um einen Flüchtigkeitsfehler handeln.

Aufgabenlösungen von Jonathan

$$\begin{array}{r}
 365 \\
 + 457 \\
 \hline
 933
 \end{array}
 \qquad
 \begin{array}{r}
 24 \\
 + 13 \\
 \hline
 47
 \end{array}
 \qquad
 \begin{array}{r}
 28 \\
 + 15 \\
 \hline
 54
 \end{array}$$

Jonathan löst keine Aufgabe richtig. Betrachtet man die einzelnen Lösungen genauer, fällt auf, dass er sich meist um 1 verrechnet. Beim Überschreiten des Zehners überträgt er zwar die 1 auf die nächste Stelle, verrechnet sich meist aber dennoch um 1. Dies deutet darauf hin, dass er noch zählend rechnet und sich hierbei ein Fehlermuster herausgebildet hat.

Abbildung 11: Beispiele für Schülerlösungen bei der schriftlichen Addition

Daliahs Aufgabenlösungen deuten darauf hin, dass sie das Verfahren grundsätzlich verstanden hat. Als Rückmeldung würde es für sie wahrscheinlich reichen, ihr zu sagen, dass zwei der drei Aufgaben richtig sind und sie auffordern, durch Nachrechnen herauszufinden, wo noch ein Fehler vorkommt. Im weiteren Verlauf könnte die Lehrperson ihr schwierigere Aufgaben geben, um zu sehen, ob sie auch in diesen Fällen alle notwendigen Verstehensschritte beherrscht. Bei der schriftlichen Addition wären dies beispielsweise Aufgaben, welche die Addition von zwei Zahlen mit einer unterschiedlichen Anzahl an Stellen verlangen (z. B. $987 + 45$). So könnte geprüft werden, ob Daliah auch der Übertrag in leere Stellen gelingt. Weitere Aufgaben könnten die Ziffer Null beinhalten (z. B. $245 + 302$), da auch dies einigen Schülern Schwierigkeiten bereitet (Gerster 2012).

An Jonathans Lösungen fällt auf, dass er bereits mit dem »Eins + Eins« Probleme hat. Das bedeutet, seine Schwierigkeiten liegen nicht, wie beispielsweise bei Bastian, am Umgang mit dem neu gelernten Verfahren der schriftlichen Addition, sondern daran, dass er das »Eins + Eins« nicht sicher beherrscht, was sich auch beim nicht-schriftlichen Rechnen zeigen dürfte. Fehler solcher Art deuten auf zählendes Rechnen hin. Meist zeigen sich hier typische Fehler in der Form, dass das vom Schüler berechnete Ergebnis um Eins größer oder kleiner ist als das richtige Ergebnis. Dies kommt häufig dadurch zustande, dass der Schüler weiterzählt und dabei beispielsweise den Fehler macht, dass er den ersten Summanden noch mitzählt. Bei der Aufgabe $7 + 4$ würde er beispielsweise ab der 7 wie folgt weiterzählen: »7, 8, 9, 10«. Dadurch ist das Ergebnis um Eins zu klein. Eine andere Möglichkeit, die in zu hohen Ergebnissen resultiert, ist, dass der Schüler die dazwischen liegenden Zahlen abzählt: Bei der Aufgabe $7 + 4$ zählt er die vier auf die 7 folgenden Zahlen ab (»8, 9, 10, 11«) und nennt dann die 12 als Ergebnis (vgl. Gerster 2012). Um herauszufinden, wie Jonathans Fehler zustande kommen, wäre es sinnvoll, ihn Aufgaben mündlich rechnen zu lassen und ihn dabei zum lauten Denken aufzufordern. Wichtig ist dabei, dass er beispielsweise nicht aus Scham ein zählendes Rechnen unter Zuhilfenahme der Finger zu verstecken versucht. Deshalb sollte ihm glaubhaft versichert werden, dass Fehler nicht schlimm sind, sondern lediglich der Lehrperson helfen, zu verstehen, wie sie ihm den Bearbeitungsweg noch besser erklären kann. Sollte sich herausstellen, dass Jonathan tatsächlich zählend rechnet, sollte abgeklärt werden, inwiefern er Zahlen als Mengen begreift (kardinaler Zahlaspekt) und nicht lediglich als Position in der Zählreihe (ordinaler Zahlaspekt). Hier können dann gezielte Übungen ansetzen, die Jonathans Zahlverständnis unterstützen (z. B. das Arbeiten mit Mengenbildern). Darauf aufbauend sollten verschiedene Möglichkeiten der Zahlzerlegung geübt werden (z. B. die verschiedenen Möglichkeiten, die Zahl 7 in die Bestandteile $1 + 6$, $2 + 5$ oder $3 + 4$ zu zerlegen). Neben dem Aufbau des Zahlverständnisses sollten langfristig dann die Aufgaben des »Eins

+ Eins« automatisiert werden, so dass die Ergebnisse aus dem Gedächtnis abgerufen werden können (Gerster 2012). Im Fall von Jonathan geht ein Feedback also weit über eine kurze Rückmeldung hinaus, sondern enthält viele instruktionale Anteile. Dabei dient das Analysieren der Schülerfehler zunächst als Ausgangspunkt für die Planung von Fördermaßnahmen, die gegebenenfalls auch nicht alle im Klassenverband stattfinden können/müssen, sondern beispielsweise in Förderkurse ausgelagert werden können. Die Bereitschaft und die Mühe der Lehrperson, sich in Jonathan hineinzusetzen und seine Schwierigkeiten nachzuvollziehen, stellt eine wichtige Grundvoraussetzung für den Erwerb handlungspraktischer diagnostischer Kompetenz dar.

Anhand der Beispiele wird deutlich, dass es zum Geben produktiver Rückmeldungen notwendig ist, den fachlichen Hintergrund der jeweiligen Lerninhalte selbst tiefgehend zu verstehen, so dass mögliche Schülerfehler antizipiert und im Falle ihres Auftretens richtig eingeordnet werden können. Zudem benötigt gutes Feedback eine genaue Diagnose, was wiederum Zeit erfordert, sich mit den Schülerlösungen über die reine »Richtig-Falsch-Bewertung« hinaus auseinanderzusetzen.

Die Abbildung 12 fasst die wichtigsten Aspekte zur Qualität von Feedback noch einmal zusammen.

Loben Sie – aber loben Sie sinnvoll.

- Loben Sie spezifisch und glaubhaft.
- Vermeiden Sie überschwängliches Lob vor allem bei unsicheren Schülern.
- Zeigen Sie Schülern ihre Lernfortschritte auf, indem Sie auch Rückmeldungen anhand einer individuellen Bezugsnorm geben.

Geben Sie informatives Feedback.

- Versuchen Sie, die Denkwege der Schüler zu verstehen, um ihnen passende Rückmeldungen und Hilfestellungen zu geben.
- Überprüfen Sie Ihre Rückmeldungen darauf, ob Sie für den Schüler wirklich relevante Informationen beinhalten und vermeiden Sie unspezifisches Feedback.

Feedback kann in verschiedene Richtungen gehen.

- Regen Sie auch ihre Schüler an, sich gegenseitig sinnvolles Feedback zu geben.
- Holen Sie sich auch selbst Rückmeldungen von Ihren Schülern zu Ihrem Unterricht ein und versuchen Sie, diese in zukünftigen Unterrichtsstunden einzubeziehen.

Abbildung 12: Praxistipps – Was ist gutes Feedback?

Was bleibt? Fazit und Würdigung der Hattie-Studie

Im Rahmen dieses Beitrags wurden ausgehend von einem Überblick über die Anlage und die wichtigsten Befunde der Hattie-Studie zwei Merkmale von Unterricht näher in den Blick genommen: Lehrerfragen und Feedback. Für beide Merkmale wurde gezeigt, dass es nicht ausreichend ist, die Rangliste von Hattie zu studieren, um zu wissen, wie guter Unterricht »geht«. Vielmehr ist es erforderlich, sich differenzierter mit den Merkmalen auseinanderzusetzen und weitere Literatur näher zu studieren, denn: Nicht jede Lehrerfrage hat positive Effekte, nicht jedes Feedback wirkt sich gleichermaßen positiv aus.

Damit ist auch schon eine Grenze der Hattie-Studie angesprochen. Die Hattie-Studie ist eine Meta-Metaanalyse, aus der man keine unmittelbaren Handlungsempfehlungen und Rezepte für die Gestaltung von Unterricht ableiten kann. Gleichwohl gibt die Studie Hinweise darauf, welche Merkmale von Unterricht eine größere Beachtung verdienen und welche unwichtiger sind.

Die Rolle der Lehrperson: Regisseur oder Moderator?

Hattie bilanziert die Ergebnisse seiner Studie, indem er größere Effektstärken für jene Merkmale identifiziert, die eine aktive, anregende und instruktive Lehrerrolle beinhalten. Er gebraucht für diese Rolle den Begriff Activator (Regisseur). Der Activator-Rolle ordnet er Merkmale wie Feedback, Direkte Instruktion und Inhaltliche Klarheit zu, alles Merkmale mit hohen Effektstärken. Demgegenüber scheinen sich unter den Merkmalen mit geringen Effekten vor allem jene Merkmale zu befinden, in denen die Lehrperson eine eher zurückhaltende Rolle einnimmt. Hattie spricht hier von der Rolle der Lehrperson als Facilitator (Moderator). Hierzu gehören z.B. der Offene Unterricht, das problembasierte Lernen oder das Merkmal Individualisierung.

Dies scheint auf den ersten Blick die häufig in der Presse verbreitete Ansicht zu stützen, Hattie wäre in seiner Studie zu dem Ergebnis gekommen, dass Frontalunterricht die beste Lehrmethode sei und dass Formen des Unterrichts, die auf zu viel Selbststeuerung der Schüler setzen, lehrergelenkten Formen des Unterrichts deutlich unterlegen seien. Diese Darstellung und Interpretation der Ergebnisse von Hattie stellt jedoch eine deutliche Verkürzung dar. Zum einen favorisiert Hattie keine bestimmte Methode oder Form des Unterrichtens. Zum zweiten ist die Direkte Instruktion keinesfalls mit einem langweiligen und eintönigen Frontalunterricht gleichzusetzen. Und zum dritten fallen bei genauer Durchsicht der Hattie-Studie sehr wohl Merkmale auf, die eine höhere Schülerelbststeuerung beinhalten und »trotzdem« hohe Effekte haben. So erreichen beispielsweise metakognitive Strategien mit $d = 0.69$ mit dem 13. Rang eine Spitzenposition und auch bei der Methode des Reciprocal Teaching ($d = 0.74$) müssen die Schüler eine sehr hohe Eigenaktivität und Selbststeuerung zeigen.

Auch weitere Metaanalysen zeigen, dass Strategien und Arbeitstechniken eine wichtige Bedeutung zukommt. Hattie bezieht sich hier auf eine Arbeit von Lavery (2008), welche die Effektstärken unterschiedlicher metakognitiver Strategien vergleicht. Dabei erwiesen sich das Erstellen eines Konzepts/Plans ($d = 0.85$), die Selbstbelohnung sowie das Erledigen der Arbeit vor angenehmen Ereignissen ($d = 0.70$) oder die Selbstbewertung des eigenen Lernprozesses ($d = 0.62$) als besonders wirksam.

Berücksichtigt man auch diese Ergebnisse, so erscheint die von Hattie vorgenommene Gegenüberstellung von effektiven Merkmalen, die der Lehrperson eine wichtige Regisseurrolle zuschreiben, und vergleichsweise wenig effektiven Merkmalen, die mit einer eher zurückhaltenden Berater- und Begleitungsfunktion der Lehrperson einhergehen, voreilig und in der Form nicht nachvollziehbar. Vielmehr dürfte es auf eine Verknüpfung beider Rollen ankommen. Dies lässt sich am Beispiel der erwähnten Strategien gut verdeutlichen: Die Forschung zeigt deutlich, dass der Erwerb dieser Lernstrategien eine aktive Lehrerrolle und eine vergleichsweise hohe Lehrerlenkung voraussetzt. Die Strategien werden den Schülern in ihrer Anwendung von der Lehrperson erklärt und demonstriert, wobei die Lehrperson in der Regel »laut« denkt, also ihre Gedanken und ihr Verhalten verbalisiert. Die ersten Anwendungen der Strategien durch die Schüler werden von der Lehrperson systematisch angeleitet und überwacht. Hinzu kommt, dass die Lehrperson den Schülern gezieltes Feedback zur Anwendung dieser Strategien gibt. Erst nach Wochen des systematischen Trainings sind die Schüler in der Regel in der Lage, diese Strategien alleine bzw. in Gruppen selbstständig anzuwenden. Die Lehrperson beobachtet, unterstützt die Schüler bei der Anwendung und fährt diese Unterstützung erst nach und nach zurück (vgl. z. B. Duffy 2002). Insofern lässt sich schlussfolgern, dass sich Lehrerlenkung und Schülerorientierung bzw. Selbststeuerung nicht ausschließen, sondern gleichermaßen wichtig sind und sich gegenseitig ergänzen (vgl. z. B. Helmke 2009, Lipowsky & Lotz in diesem Band).

Auf offene Unterrichtsformen unter Verweis auf Hatties Ergebnisse gänzlich zu verzichten, wäre demzufolge nicht angemessen. Vielmehr, und das können Hatties Ergebnisse eindrucksvoll zeigen, dürfte es darauf ankommen, auch offene Unterrichtsformen qualitativ hochwertig durchzuführen (vgl. Lipowsky & Lotz in diesem Band). Das bedeutet unter anderem, dass die Lernenden Strategien und Arbeitstechniken benötigen, um die Freiräume konstruktiv zu nutzen, dass sie auch in offenen Unterrichtsformen kognitiv herausgefordert werden müssen und konstruktives Feedback benötigen und dass gegebenenfalls die inhaltliche Komplexität reduziert werden muss, um die Lernenden kognitiv nicht zu überfordern. Guter Unterricht ist keine Frage der Unterrichtsform (vgl. auch Lipowsky 2015).

Häufige Kritikpunkte an der Hattie-Studie

Nach den euphorischen Berichten in der Presse hat sich der Hype um die Hattie-Studie etwas gelegt. Zunehmend wird auch Kritik an der Studie geübt. Im Folgenden werden einige Einwände zusammenfassend dargestellt und eingeordnet.⁴

- **Einwand 1:** *Hatties Ergebnisse beziehen sich gar nicht auf die deutsche Situation, sondern auf die Situation im angloamerikanischen Sprachraum.* Diesem Kritikpunkt ist entgegenzuhalten, dass sich Hatties Ergebnisse, wie der obere Abschnitt zu den Basisdimensionen zeigt, durchaus mit Ergebnissen der deutschen Forschung verbinden lassen. Zudem sind in den von Hattie ausgewerteten Metaanalysen auch deutsche Studien, die englischsprachig publiziert wurden, enthalten.
- **Einwand 2:** *Hatties Ergebnisse sind veraltet und deshalb nicht mehr relevant.* Diese Kritik stimmt, was den Zeitpunkt der Durchführung der Originalstudien anbelangt. Sie stimmt jedoch weniger, was die grundsätzliche Relevanz der Ergebnisse anbelangt, denn auch neuere Studien reihen sich gut in die von Hattie dargestellte Befundlage ein. Allerdings dürfte sich mit der Zeit die Anzahl der relevanten Merkmale, die von Hattie bislang noch nicht untersucht worden sind, weiter vergrößern, denn mit der Weiterentwicklung der Forschung werden auch neue Einflussfaktoren in den Fokus der Forschung geraten. So stellt Hattie z. B. fest, dass das Fachwissen der Lehrpersonen nur einen vergleichsweise geringen Einfluss auf das Lernen der Schüler hat ($d = 0,09$). Nach neueren Studien ist dieses Bild jedoch zu revidieren. Es mehren sich Befunde, dass das fachwissenschaftliche, fachdidaktische und pädagogisch-psychologische Wissen von Lehrkräften einen nicht zu unterschätzenden Einfluss darauf hat, wie gut der Unterricht ist und wie viel die Schüler lernen (Baumert u. a. 2010, Voss, Kunter, Seiz, Hoehne & Baumert 2014).
- **Einwand 3:** *Eng mit Kritikpunkt 2 hängt der folgende zusammen. Hattie skizziert ein veraltetes, nicht mehr zeitgemäßes Bild von Unterricht.* Diese Kritik ist nicht plausibel. Im Gegenteil: Im vorherigen Abschnitt wurde herausgearbeitet, dass Hatties Ergebnisse weder als Plädoyer für einen langweiligen Frontalunterricht noch für einen laissez-faire-Unterrichtsstil zu verstehen sind. Vielmehr lassen die Merkmale in der Gesamtschau einen Unterricht erkennen, der sich durch die drei oben erläuterten Basisdimensionen sowie durch eine verständnisorientierte, inhaltlich klare und interessante Gestaltung auszeichnet. Zudem wird deutlich, dass die Lernenden systematisch dazu befähigt werden müssen, ihr Lernen selbst zu steuern. Ein solcher Unterricht ist höchst anspruchsvoll zu verwirklichen.

⁴ Für Ausführungen zur Kritik an der Hattie-Studie vgl. z. B. Brügelmann 2013, Lind 2013, Pant 2014, Peschel 2014 oder Terhart 2014.

- **Einwand 4:** *Hattie legt einen einseitigen Schwerpunkt auf die Lernleistungen der Schüler.* Es stimmt zwar, dass Hattie sich auf die Analyse der schulischen Leistungen konzentriert. Allerdings beeinflusst das Lernen auch andere Merkmale der Schülerpersönlichkeit. So zeigt die Forschung, dass Leistungen mit affektiv-motivationalen Variablen (z. B. dem Selbstkonzept) in teilweise engem Zusammenhang stehen (vgl. z. B. Guay, Marsh & Boivin 2003, Valentine, DuBois & Cooper 2004), so dass mit einer positiven Beeinflussung der Leistungen auch positive Wirkungen in anderen Bereichen der Schülerpersönlichkeit zu erwarten sind. Gleichwohl dürften einige der von Hattie untersuchten Merkmale auf motivationale Aspekte andere, stärkere oder schwächere Effekte ausüben als auf die Leistung (vgl. auch Lipowsky 2015).
- **Einwand 5:** *Hattie berücksichtigt Unterschiede in den Kontextbedingungen zu wenig.* Dieser Einwand ist nicht völlig von der Hand zu weisen. Hattie schreibt zwar einleitend, dass er keine Hinweise darauf gefunden habe, dass Kontextbedingungen die Ergebnisse beeinflussen. Allerdings ist kritisch zu fragen, ob die von ihm ausgewerteten Studien Unterschiede in Kontextbedingungen (z. B. schulische Einzugsgebiete) überhaupt erfasst und dokumentiert haben.
- **Einwand 6:** *Eng mit dem Einwand 5 hängt zusammen: Hatties Ergebnisse beziehen sich vorwiegend auf den mathematisch-naturwissenschaftlichen Bereich und auf das Lesen. Andere Domänen und insbesondere musische, künstlerische und sozialwissenschaftliche Fächer werden nicht berücksichtigt.* Dieser Einwand ist insofern berechtigt, als die Forschung Lehr- und Lernprozesse vor allem in mathematisch-naturwissenschaftlichen Fächern und bedingt auch in sprachlichen Fächern untersucht hat. Demgegenüber wurde Unterricht in musisch-künstlerischen und sozialwissenschaftlichen Fächern bislang deutlich seltener untersucht, so dass diese Fächer auch in den von Hattie ausgewerteten Metaanalysen nicht bzw. kaum berücksichtigt sind. Einige fachübergreifend gültige Merkmale von Unterrichtsqualität, wie z. B. eine effektive Klassenführung oder auch Feedback, dürften aber für jegliches schulisches Lernen, also auch in musischen und sozialwissenschaftlichen Fächern, von Bedeutung sein.
- **Einwand 7:** *Die von Hattie herangezogenen Metaanalysen sind in ihrer Qualität unterschiedlich.* Dieser Einwand stimmt. Die Studien unterscheiden sich teilweise erheblich in ihrer Qualität. Einige sind sehr hochrangig publiziert, andere liegen als unveröffentlichte Schriften vor.
- **Einwand 8:** *Hattie bildet über alle Effektstärken, die sich aus den von ihm ausgewerteten Metaanalysen zu einem Merkmal ergeben, einen Mittelwert und kommt daher zu verzerrten Ergebnissen.* Tatsächlich ist die Art der statistischen Zusammenfassung seiner Ergebnisse fragwürdig. Hattie mittelt die Effektstärken aus den zu einem Merkmal gehörenden Metaanalysen, ohne zu beachten, wie viele Einzelstudien den jeweiligen Metaanalysen zugrunde liegen. Erwartbar wäre gewesen, dass er die Effektstär-

ken entsprechend der Größe der einzelnen Metaanalysen gewichtet. So wären Effektstärken aus größeren Metaanalysen mit mehr Einzelstudien mit einem größeren Gewicht in die Auswertung eingegangen als Effektstärken aus kleineren Metaanalysen. Dass eine gewichtete Berücksichtigung zu teilweise anderen Ergebnissen führt, zeigt Pant (2014) am Beispiel der Direkten Instruktion. Demnach verringert sich die Effektstärke von $d = 0.59$ auf $d = 0.39$, wenn man die unterschiedlichen Stichprobengrößen der einzelnen Metaanalysen berücksichtigt. Zudem ist die Zuordnung der von Hattie ausgewerteten Metaanalysen zu den Merkmalen nicht in allen Fällen nachvollziehbar.

Trotz dieser Einwände hat Hattie mit seiner Studie eine wertvolle empirisch fundierte Grundlage für die Diskussion um eine höhere Qualität an unseren Schulen gelegt. Seine Ergebnisse befreien Lehrkräfte, Schulleiter und Verantwortliche in der Bildungsverwaltung und Bildungspolitik jedoch nicht davon, die Studie kritisch und differenziert zu analysieren und weitere Literatur heranzuziehen.

Fazit: Worauf kommt es an?

Die Hattie-Studie verdeutlicht, dass schulisches Lernen von unterschiedlichen Einflussfaktoren abhängig ist. Geht man auf das Eingangsbeispiel von Ben und Lilli zurück, so lässt sich nicht sicher sagen, welche Faktoren für die schwachen Leistungen von Ben verantwortlich sind. Möglicherweise verfügt Ben über ungünstige Lernvoraussetzungen, die das Dazulernen erschweren. Denkbar ist auch, dass Ben (darüber hinaus) einen Unterricht erfährt, der ihn wenig fördert und anregt. Eventuell weiß die Lehrperson zu wenig über die Schwierigkeiten von Ben und wie darauf zu reagieren ist. Von allen schulischen Einflussfaktoren, das zeigt nicht nur die Hattie-Studie, ist der Unterricht und das, was im Klassenzimmer passiert, das wichtigste Faktorenbündel. Demgegenüber spielen Rahmenbedingungen der Schule eine deutlich untergeordnete Rolle.

Man kann also diesen Punkt mit »Es kommt auf das Kerngeschäft der Lehrperson, das Lehrerhandeln im Unterricht an« zusammenfassen. Dennoch können die Ergebnisse von Hattie nicht unmittelbar auf den Einzelfall angewendet werden. Die Studie hilft zwar, einen Überblick über potenzielle Ursachen von Leistungsschwierigkeiten zu gewinnen und den Blick auf vermutlich relevante Aspekte zu richten. Davon ausgehend sollte dann aber jeder Schüler mit seinen jeweiligen Stärken und Schwächen sowie den schulischen und außerschulischen Rahmenbedingungen individuell betrachtet werden.

»Visible learning« macht aber nicht nur deutlich, wie wichtig das Handeln der Lehrperson im Klassenzimmer ist, »visible learning« steht auch für entsprechende

Haltungen und Wertorientierungen der Lehrpersonen. Hierzu gehört unter anderem, dass sich die Lehrperson für die Lernprozesse ihrer Schüler interessiert und sie sichtbar macht, dass sie sich bemüht, diese zu antizipieren und nachzuvollziehen, dass sie von der Beeinflussbarkeit der Entwicklung ihrer Schüler durch Unterricht überzeugt ist und um ihre Einflussmöglichkeiten auf das Lernen der Schüler weiß. Dahinter verbirgt sich demnach eine Art handlungsbezogener Optimismus oder mit anderen Worten eine hohe Selbstwirksamkeitserwartung der Lehrkraft. Zu den wünschenswerten Haltungen und Wertorientierungen, die hinter den von Hattie identifizierten Merkmalen durchscheinen, gehören auch Zutrauen und Zumutung im Sinne anspruchsvoller Erwartungen an die Schüler.

Was man sich hierunter vorstellen kann, lässt sich gut an dem britischen Choreographen und Tanzpädagogen Royston Maldoom und seiner Arbeit in dem Projekt »Rythm Is It«⁵ illustrieren. Der Film »Rhythm Is It« zeigt, wie 250 Berliner Jugendliche in mehrwöchiger Arbeit mit Royston Maldoom Igor Strawinskis Ballett »Le sacre du printemps« einstudieren und gemeinsam mit den Berliner Philharmonikern unter der Leitung von Simon Rattle zur Aufführung bringen.

Zunächst überrascht es, dass sich Berliner Jugendliche mitten in der Pubertät überhaupt zu einem solchen klassischen Tanzprojekt und zu der Teilnahme an den vielen Proben bewegen lassen. Der Film »Rhythm Is It« vermittelt, selbst wenn einige Szenen etwas pathetisch wirken, einen guten Eindruck, wie sich einzelne Jugendliche entwickeln und im Verlaufe der Arbeit an diesem Projekt an Selbstsicherheit und Selbstwirksamkeit gewinnen. Dabei verlaufen die ersten Proben alles andere als erfolgreich. Beobachtbar ist eine Reihe von Störungen, die aus Sicht von Royston Maldoom vor allem aus einer mangelnden Selbstsicherheit, geringer Disziplin und einer unpassenden Einstellung zu dem Projekt resultieren. In einer Szene, in der die jugendlichen Tänzer sich beschwerten, dass das harte Training zu wenig »Fun mache«, verdeutlicht er den Jugendlichen, dass aus der Erfahrung, etwas Anstrengendes zu leisten und eigene Könnenserfahrungen zu sammeln, Freude erwachsen kann.

5 Zwar handelt es sich bei dem Film um einen Dokumentarfilm, doch natürlich beinhaltet auch dieser Dokumentarfilm eine Konstruktion der Wirklichkeit, d. h. bestimmte Szenen werden gezeigt, andere werden weggelassen oder herausgeschnitten. Die Ausschnitte, auf die sich die Kamera richtet, resultieren aus bestimmten Interessen und Erwartungen. Insofern ist auch eine gewisse Zurückhaltung und Distanz gegenüber diesem Film angebracht. Beispielsweise halten nicht alle Jugendlichen bis zum Ende des Projekts durch, einige brechen die Teilnahme an dem Projekt vorher ab.

Die Jugendlichen reagieren auf die Anforderungen mit Unsicherheit und glauben offenbar auch nicht daran, innerhalb weniger Wochen dieses Ballettstück einstudieren zu können. Royston Maldoom traut ihnen das zu und fordert sie entsprechend. Er schreckt auch nicht davor zurück, einzelne Schüler offen zu kritisieren und Disziplin einzufordern. Seine »harte Gangart« wird von den Lehrerinnen der teilnehmenden Schüler kritisiert.

Als Zuschauer hat man dennoch den Eindruck, dass sich hinter dieser fordernden Haltung von Maldoom neben hohen Leistungserwartungen auch Zutrauen verbirgt. Symptomatisch für diese Erwartungshaltung ist eine Szene, in der Maldoom den Jugendlichen die Angst und die Scham, sich lächerlich zu machen, zu nehmen versucht. Wieder einmal unterbricht er die Probe, weil es zu einer Störung gekommen ist. Die Jugendlichen sitzen auf dem Boden der Sporthalle, in der die Probe stattfindet.

Maldoom: Wovor habt ihr Angst? Was macht euch so nervös? Denn das Gelächter zeigt, auch wenn das euch nicht bewusst ist, dass da tief im Inneren eine Angst sitzt. Vielleicht ist es die Angst, euch lächerlich zu machen, oder es nicht zu können. Oder es ist die Angst, eure Freunde könnten euch auslachen. Wenn eure Freunde euch auslachen, wenn ihr versucht, etwas zu erreichen, dann fragt euch, ob das wirklich Freunde sind, denn ein Freund ist jemand, der da ist und euch unterstützt, wenn ihr versucht, etwas Neues zu machen. Das ist wirklich wichtig für euer ganzes Leben. Ein Freund ist jemand, der euch hilft, weiterzukommen.

Der Zuschauer gewinnt beim Betrachten dieser und anderer Szene den Eindruck: Dieser Pädagoge kritisiert die tanzenden Jugendlichen nicht, weil er sie für inkompetent oder schwach hält, sondern weil er davon überzeugt ist, dass sie ihr Können nicht abrufen und nicht an ihre Grenzen gehen und dass mehr in ihnen steckt als sie es selbst glauben. Man hat den Eindruck, dass die Jugendlichen die implizite Botschaft in den kritischen Äußerungen und Forderungen von Maldoom verstehen (siehe auch »paradoxe Effekte von Lob und Tadel«) und diese nicht als Abwertung ihrer Person, sondern als Ausdruck hoher Leistungserwartungen und eines vorhandenen Zutrauens begreifen, was sich dann auch positiv auf das Selbstkonzept und das Wirksamkeitserleben der Jugendlichen auswirken kann. Übrigens: Nach sechs Wochen Training endete das Projekt mit dem umjubelten Auftritt der Jugendlichen in der Berlin Arena vor 3.000 Zuschauern.

Wie die oben genannten Basisdimensionen von Unterrichtsqualität und hierfür erforderliche Orientierungen und Haltungen von Lehrpersonen zusammenhängen, wird auch am Nachwort in Hatties Studie deutlich. Hier zitiert er seinen Kollegen Paul Brock, der in seiner Autobiographie 2004, konfrontiert mit der Diagnose einer unheilbaren Krankheit, den künftigen Lehrpersonen seiner Töchter Folgendes auf den Weg gibt:

Fördern Sie die intellektuellen Fähigkeiten und die Fantasie meiner Töchter und fordern Sie sie zu neuen Horizonten heraus [...]. Bevormunden Sie sie nicht durch [den, M. L. & F. L.] Pudding des kleinsten gemeinsamen Nenners, der als Wissen und Lernen verkleidet ist, und zerstören Sie ihre Liebe für das Lernen nicht durch eine Pädagogik der Langeweile [...]. Kümmern Sie sich um Sophie und Millie mit Menschlichkeit und Einfühlungsvermögen und sehen Sie sie als Menschen, die es wert sind, mit echtem Respekt [...] unterrichtet zu werden. (Brock 2004, zitiert nach Hattie 2013, S. 308)

Literatur

- Aebli, Hans (1987): Grundlagen des Lehrens. Eine Allgemeine Didaktik auf psychologischer Grundlage. Stuttgart.
- Aschersleben, Karl (1999): Frontalunterricht – klassisch und modern. Eine Einführung. Neuwied.
- Bangert-Drowns, Robert L.; Kulik, Chen-Lin C.; Kulik, James A. & Morgan, Mary Teresa (1991): The instructional effect of feedback in test-like events. Review of Educational Research 61 (2), S. 213–238.
- Baumert, Jürgen; Kunter, Mareike; Blum, Werner; Brunner, Martin; Voss, Thamar; Jordan, Alexander; Klusmann, Uta; Krauss, Stefan; Neubrand, Michael & Yi-Miau, Tsai (2010): Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. American Educational Research Journal 47 (1), S. 133–180.
- Bellack, Arno; Kliebard, Herbert M.; Hyman, Ronald T. & Smith, Frank L. (1974): Die Sprache im Klassenzimmer. Düsseldorf.
- Bloom, Benjamin S.; Engelhart, Max D.; Furst, Edward J.; Hill, Walker H. & Krathwohl, David R. (1956): Taxonomy of Educational Objectives. The Classification of Educational Goals. New York.
- Bloom, Benjamin S.; Engelhart, Max D.; Furst, Edward J.; Hill, Walker H. & Krathwohl, David R. (1976): Taxonomie von Lernzielen im kognitiven Bereich. Weinheim.

- Bond, Lloyd; Smith, Tracy W.; Baker, Wanda K. & Hattie, John A. (2000): The Certification System of the National Board for Professional Teaching Standards. A Construct and Consequential Validity Study. Greensboro.
- Brügelmann, Hans (2013): Die Hattie-Studie. Der heilige Gral der Didaktik? Grundschule aktuell (121), S. 25–26.
- Brummelman, Eddie; Thomaes, Sander; Orobio de Castro, Bram; Overbeek, Geertjan & Bushman, Brad J. (2014): »That's not just beautiful – That's incredibly beautiful!« The adverse impact of inflated praise on children with low self-esteem. Psychological Science 25 (3), S. 728–735.
- Brunner, Ewald Johannes & Noack, Peter (2006): Lehrer-Schüler-Interaktion. In: Rost, Detlef H. (Hrsg.): Handwörterbuch pädagogische Psychologie. Weinheim, S. 421–430.
- Brunstein, Joachim C. & Heckhausen, Heinz (2006): Leistungsmotivation. In: Heckhausen, Jutta & Heckhausen, Heinz (Hrsg.): Motivation und Handeln. Heidelberg, S. 143–187.
- Butler, Deborah L. & Winne, Philip H. (1995): Feedback and self-regulated learning. A theoretical synthesis. Review of Educational Research 65 (3), S. 245–281.
- Cecil, Nancy L. (2008): Mit guten Fragen lernt man besser. Die besten Fragetechniken für den Unterricht. Mülheim an der Ruhr.
- Dubs, Rolf (2008): Das Lehrgespräch im Klassenunterricht. Nicht mehr zeitgemäß? In: (Hrsg.): Bundesarbeitskreis der Seminar- und Fachleiter/innen e. V. (BAK). Unterrichtsqualität. Baltmannsweiler, S. 7–16.
- Duffy, Gerald G. (2002): The case for direct explanation of strategies. In: Block, Cathy Collins & Pressley, Michael (Hrsg.): Comprehension Instruction. Research-Based Best Practices. Solving Problems in the Teaching of Literacy. New York, S. 28–41.
- Flanders, Ned A. (1970): Analyzing Teaching Behavior. New York.
- Frey, Nancy & Fisher, Douglas (2007): Reading for Information in Elementary School. Content Literacy Strategies to Build Comprehension. Upper Saddle River.
- Gabriel, Katrin (2014): Videobasierte Erfassung von Unterrichtsqualität im Anfangsunterricht der Grundschule. Klassenführung und Unterrichtsklima in Deutsch und Mathematik. Kassel.
- Gage, Nathaniel L.; Berliner, David C. & Bach, Gerhard (1996): Pädagogische Psychologie. Weinheim.
- Gaudig, Hugo (1909): Didaktische Präludien. Leipzig.
- Gayle, Barbara M.; Preiss, Raymond W. & Allen, Mike (2006): How effective are teacher-initiated classroom questions in enhancing student learning? In: Gayle, Barbara M.; Preiss, Raymond W.; Burrell, Nancy & Allen, Mike (Hrsg.): Classroom Communication and Instructional Processes. Advances Through Meta-Analysis. Mahwah/NJ, S. 279–293.

- Gerster, Hans-Dieter* (2012): Schülerfehler bei schriftlichen Rechenverfahren. Diagnose und Therapie. Münster.
- Gruehn, Sabine* (2000): Unterricht und schulisches Lernen. Schüler als Quellen der Unterrichtsbeschreibung. Münster.
- Guay, Frédéric; Marsh, Herbert W. & Boivin, Michel* (2003): Academic self-concept and academic achievement: Developmental perspectives on their causal ordering. *Journal of Educational Psychology* 95 (1), S. 124–136.
- Hasselhorn, Marcus & Gold, Andreas* (2013): Pädagogische Psychologie. Erfolgreiches Lehren und Lernen. Stuttgart.
- Hattie, John* (2003): Teachers make a difference. What is the research evidence? <http://www.educationalleaders.govt.nz/Pedagogy-and-assessment/Building-effective-learning-environments/Teachers-Make-a-Difference-What-is-the-Research-Evidence>.
- Hattie, John* (2009): Visible Learning. A Synthesis of Over 800 Meta-Analyses Relating to Achievement. London.
- Hattie, John* (2013): Lernen sichtbar machen. Baltmannsweiler.
- Hattie, John & Timperley, Helen* (2007): The power of feedback. *Review of Educational Research* 77 (1), S. 81–112.
- Heinze, Aiso & Erhard, Markus* (2006): How much time do students have to think about teacher questions? An investigation of the quick succession of teacher questions and student responses in the German mathematics classroom. *ZDM (International Journal on Mathematics Education)* 38 (5), S. 388–398.
- Helmke, Andreas* (2009): Unter richtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts. Seelze-Velber.
- Helmke, Andreas; Helmke, Tuyet; Schrader, Friedrich-Wilhelm; Wagner, Wolfgang; Nold, Günter & Schröder, Konrad* (2008): Alltagspraxis des Englischunterrichts. In: Klieme, Eckhard (Hrsg.): Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie. Weinheim, S. 371–381.
- Henderlong, Jennifer & Lepper, Mark R.* (2002): The effects of praise on children's intrinsic motivation. A review and synthesis. *Psychological Bulletin* 128 (5), S. 774–795.
- Huth, Katja* (2004): Entwicklung und Evaluation von fehlerspezifischem informativem tutoriellem Feedback (ITF) für die schriftliche Subtraktion. <http://www.qucosa.de/fileadmin/data/qucosa/documents/1243/1105354057406-4715.pdf>
- Ingenkamp, Karlheinz* (Hrsg.) (1995): Die Fragwürdigkeit der Zensurengebung. Texte und Untersuchungsberichte. Weinheim.
- Jacobs, Bernhard* (2002): Aufgaben stellen und Feedback geben. <http://psydok.sulb.uni-saarland.de/voll-texte/2004/438/pdf/feedback.pdf>

- Klieme, Eckhard; Lipowsky, Frank; Rakoczy, Katrin & Ratzka, Nadja* (2006): Qualitätsdimensionen und Wirksamkeit von Mathematikunterricht. Theoretische Grundlagen und ausgewählte Ergebnisse des Projekts »Pythagoras«. In: Prenzel, Manfred & Allolio-Näcke, Lars (Hrsg.): Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms BIQUA. Münster, S. 127–146.
- Klieme, Eckhard; Schümer, Gundel & Knoll, Steffen* (2001): Mathematikunterricht in der Sekundarstufe I. »Aufgabenkultur« und Unterrichtsgestaltung. In: Klieme, Eckhard & Baumert, Jürgen (Hrsg.): TIMSS – Impulse für Schule und Unterricht. Forschungsbefunde, Reforminitiativen, Praxisberichte und Video-Dokumente. Bonn, S. 43–57.
- Klinzing, Hans G. & Klinzing-Eurich, Gisela* (1982): Die Klarheit der Lehrerfrage. Auswirkungen eines Trainings von Fragen höherer Ordnung. Unterrichtswissenschaft 10 (4), S. 313–328.
- Kluger, Avraham N. & DeNisi, Angelo* (1996): The effects of feedback interventions on performance. A historical review, a meta-analysis, and a preliminary feedback intervention theory. Psychological Bulletin 119 (2), S. 254–284.
- Knapp, Angela* (1985): Die Lehrerfrage unter dem Aspekt des Lehrfertigkeitsstrainings. Regensburg.
- Knierim, Birte* (2008): Lerngelegenheiten anbieten – Lernangebote nutzen. Eine Videostudie im Schweizer Physikunterricht. Hamburg.
- Kobarg, Mareike & Seidel, Tina* (2003): Prozessorientierte Lernbegleitung im Physikunterricht. In: Seidel, Tina; Prenzel, Manfred; Duit, Reinders & Lehrke, Manfred (Hrsg.): Technischer Bericht zur Videostudie »Lehr-Lern-Prozesse im Physikunterricht«. Kiel, S. 151–200.
- Kobarg, Mareike & Seidel, Tina* (2007): Prozessorientierte Lernbegleitung – Videoanalysen im Physikunterricht der Sekundarstufe I. Unterrichtswissenschaft 35 (2), S. 148–168.
- La Paro, Karen M.; Pianta, Robert C. & Stuhlman, Megan* (2004): The classroom assessment scoring system: Findings from the prekindergarten year. The Elementary School Journal 104 (5), S. 409–426.
- Leuders, Timo & Holzäpfel, Lars* (2011): Kognitive Aktivierung im Mathematikunterricht. Unterrichtswissenschaft 39 (3), S. 213–230.
- Lind, Georg* (2013): Meta-Analysen als Wegweiser? Zur Rezeption der Studie von Hattie in der Politik. http://www.uni-konstanz.de/ag-moral/pdf/Lind-2013_meta-analysen-als-wegweiser.pdf
- Lipowsky, Frank* (2006): Auf den Lehrer kommt es an. Empirische Evidenzen für Zusammenhänge zwischen Lehrerkompetenzen, Lehrerhandeln und dem Lernen der Schüler. In: Allemann-Ghionda, Christina & Terhart, Ewald (Hrsg.): Kompetenz

- und Kompetenzentwicklung von Lehrerinnen und Lehrern. Beiheft der Zeitschrift für Pädagogik, S. 47–70.
- Lipowsky, Frank* (2015): Unterricht. In: Wild, Elke & Möller, Jens (Hrsg.): Pädagogische Psychologie. Heidelberg, S. 69–105.
- Lotz, Miriam* (2015): Kognitive Aktivierung im Leseunterricht der Grundschule. Eine Videostudie zur Gestaltung und Qualität von Leseübungen im ersten Schuljahr. Wiesbaden.
- Lotz, Miriam & Lipowsky, Frank* (2014): Was wirklich wichtig ist. Folgerungen aus der Hattie-Studie für den Grundschulunterricht. Die Grundschulzeitschrift (272/273), S. 6–11.
- Mandl, Heinz & Kopp, Brigitta* (2008): Ursachenzuschreibung aus Sicht der Attributonstheorie. In: Zumbach, Jörg & Mandl, Heinz (Hrsg.): Pädagogische Psychologie in Theorie und Praxis. Ein fallbasiertes Lehrbuch. Göttingen, S. 177–183.
- Mehan, Hugh* (1979): Learning Lessons. Social Organization in the Classroom. Cambridge.
- Meyer, Hugh* (2010): Merkmale guten Unterrichts. Ein Kriterienmix. In: Jürgens, Eiko & Standop, Jutta (Hrsg.): Was ist »guter« Unterricht? Namhafte Expertinnen und Experten geben Antwort. Bad Heilbrunn, S. 159–174.
- Mory, Edna Holland* (2004): Feedback research revisited. In: David H. Jonassen (Hrsg.): Handbook of Research on Educational Communications and Technology. Mahwah, S. 745–783.
- Narciss, Susanne* (2004): The impact of informative tutoring feedback and self-efficacy on motivation and achievement in concept learning. *Experimental Psychology* 51 (3), S. 214–228.
- Narciss, Susanne* (2006): Informatives tutorielles Feedback. Entwicklungs- und Evaluationsprinzipien auf der Basis instruktionspsychologischer Erkenntnisse. Münster.
- Niegemann, Helmut & Stadler, Silke* (2001): Hat noch jemand eine Frage? Systematische Unterrichtsbeobachtung zu Häufigkeit und kognitivem Niveau von Fragen im Unterricht. *Unterrichtswissenschaft* 29 (2), S. 171–192.
- Nuding, Anton* (2009): Lehrerfrage oder Impuls? Ein ewig altes Thema doch aktuell. *Pädagogik* 61 (1), S. 11–13.
- Pant, Hans Anand* (2014): Visible Evidence? Eine methodisch orientierte Auseinandersetzung mit John Hatties Meta-Metaanalysen. In: Terhart, Ewald (Hrsg.): Die Hattie-Studie in der Diskussion. Probleme sichtbar machen. Seelze, S. 134–146.
- Pauli, Christine* (2006): Klassengespräch. In: Hugener, Isabelle; Klieme, Eckhard & Pauli, Christine (Hrsg.): Videoanalysen. Frankfurt/M., S. 124–147.
- Peschel, Falko* (2014): Mr. Hattie und der Offene Unterricht. Interpretation eines Einzelfaktors. Die Grundschulzeitschrift 28 (272/273), S. 6–11.

- Petersen, Jörg & Sommer, Hartmut* (1999): Die Lehrerfrage im Unterricht. Ein praxisorientiertes Studien- und Arbeitsbuch mit Lernsoftware. Donauwörth.
- Pietsch, Marcus* (2010): Evaluation von Unterrichtsstandards. *Zeitschrift für Erziehungswissenschaft* 13 (1), S. 121–148.
- Rakoczy, Katrin; Klieme, Eckhard; Bürgermeister, Anika & Harks, Birgit* (2008): The interplay between student evaluation and instruction. Grading and feedback in mathematics classrooms. *Zeitschrift für Psychologie* 216 (2), S. 111–124.
- Renkl, Alexander* (1991): Die Bedeutung der Aufgaben- und Rückmeldungsgestaltung für die Leistungsentwicklung im Fach Mathematik (Dissertation Universität Heidelberg). Heidelberg.
- Rheinberg, Falko* (2006): Bezugsnormorientierung. In: Arnold, Karl-Heinz; Sandfuchs, Uwe & Wiechmann, Jürgen (Hrsg.): *Handbuch Unterricht*. Bad Heilbrunn, S. 643–648.
- Rheinberg, Falko* (2008): Bezugsnormen und die Beurteilung von Lernleistung. In: Schneider, Wolfgang & Hasselhorn, Marcus (Hrsg.): *Handbuch der pädagogischen Psychologie*. Göttingen, S. 178–186.
- Rheinberg, Falko & Fries, Stefan* (2006): Bezugsnormorientierung. In: Rost, Detlef H. (Hrsg.): *Handwörterbuch Pädagogische Psychologie*. Weinheim, S. 61–68.
- Rheinberg, Falko & Fries, Stefan* (2010): Bezugsnormorientierung. In: Rost, Detlef H. (Hrsg.): *Handwörterbuch Pädagogische Psychologie*. 4. Aufl. Weinheim, S. 61–66.
- Rheinberg, Falko & Vollmeyer, Regina* (2010): Paradoxe Effekte von Lob und Tadel. In: Rost, Detlef H. (Hrsg.): *Handwörterbuch pädagogische Psychologie*. 4. Aufl. Weinheim, S. 635–641.
- Richert, Peggy* (2005): Typische Sprachmuster der Lehrer-Schüler-Interaktion. Empirische Untersuchung zur Feedbackkomponente in der unterrichtlichen Interaktion. Bad Heilbrunn.
- Riedl, Alfred* (2010): *Grundlagen der Didaktik*. Stuttgart.
- Rowe, Mary Budd* (1974): Reflections on wait-time. Some methodological questions. *Journal of Research in Science Teaching* 11 (3), S. 263–279.
- Sadler, D. Royce* (1989): Formative assessment and the design of instructional systems. *Instructional Science* 18, S. 119–144.
- Sommer, Hartmut* (1981): *Grundkurs Lehrerfrage*. Ein handlungsorientiertes einführendes Arbeitsbuch für Lehrer. Weinheim.
- Steindorf, Gerhard* (2000): *Grundbegriffe des Lehrens und Lernens*. Bad Heilbrunn.
- Terhart, Ewald* (2014): Der Heilige Gral der Schul- und Unterrichtsforschung – Gefunden? Eine Auseinandersetzung mit Visible Learning. In: Terhart, Ewald (Hrsg.): *Die Hattie-Studie in der Diskussion*. Probleme sichtbar machen. Seelze, S. 10–23.

- Terhart, Ewald* (2014): Die Hattie-Studie in der Diskussion. Probleme sichtbar machen. Seelze.
- Tobin, Kenneth* (1987): The role of wait time in higher cognitive level learning. *Review of Educational Research* 57 (1), S. 69–95.
- Tsui, Amy B. M.; Marton, Ference; Mok, Ida Ah Chee & Ng, Dorothy F. P.* (2004): Questions and the space of learning. In: Marton, Ference & Tsui, Amy B. M. (Hrsg.): *Classroom discourse and the space of learning*. Mahwah, S. 113–117.
- Valentine, Jeffrey C.; DuBois, David L. & Cooper, Harris* (2004): The relation between self-beliefs and academic achievement. A metaanalytic review. *Educational Psychologist* 39, S. 111–133.
- Van den Boom, Gerard; Paas, Fred & van Merriënboer, Jeroen* (2007): Effects of elicited reflections combined with tutor or peer feedback on self-regulated learning and learning outcomes. *Learning and Instruction* 17 (5), S. 532–548.
- Vollmeyer, Regina & Rheinberg, Falko* (2005): A surprising effect of feedback on learning. *Learning and Instruction* 15 (6), S. 589–602.
- Voss, Tamar; Kunter, Mareike; Seiz, Johanna; Hoehne, Verena & Baumert, Jürgen* (2014): Die Bedeutung des pädagogisch-psychologischen Wissens von angehenden Lehrkräften für die Unterrichtsqualität. *Zeitschrift für Pädagogik* 60 (2), S. 184–201.
- Waldis, Monika; Grob, Urs; Pauli, Christine & Reusser, Kurt* (2010): Der Einfluss der Unterrichtsgestaltung auf Fachinteresse und Mathematikleistung. In: Reusser, Kurt; Pauli, Christine & Waldis, Monika (Hrsg.): *Unterrichtsgestaltung und Unterrichtsqualität. Ergebnisse einer internationalen und schweizerischen Videostudie zum Mathematikunterricht*. Münster, S. 209–252.
- Weiner, Bernhard; Frieze, Irene; Kukla, Andy; Reed, Linda; Rest, Stanley & Rosenbaum, Robert M.* (1972): Perceiving the causes of success and failure. In: Jones, Edward E. (Hrsg.): *Attribution. Perceiving the Causes of Behavior*. Morristown, S. 95–120.
- Weinert, Franz E.* (1998): Lehrerkompetenz als Schlüssel der inneren Schulreform. *schulreport* 2, S. 24–27.
- Weinert, Franz E.* (1999): Aus Fehlern lernen und Fehler vermeiden lernen. In: Althof, Wolfgang (Hrsg.): *Fehlerwelten. Vom Fehlermachen und Lernen aus Fehlern. Beiträge und Nachträge zu einem interdisziplinären Symposium aus Anlass des 60. Geburtstags von Fritz Oser*. Opladen, S. 101–109.
- Wilén, William W.* (1982): Questioning skills, for teachers. What research says to the teacher. Washington: National Education Association. <http://files.eric.ed.gov/full-text/ED222488.pdf>
- Wragg, Edward Conrad & Brown, George* (2001): *Questioning in the Primary School*. London.