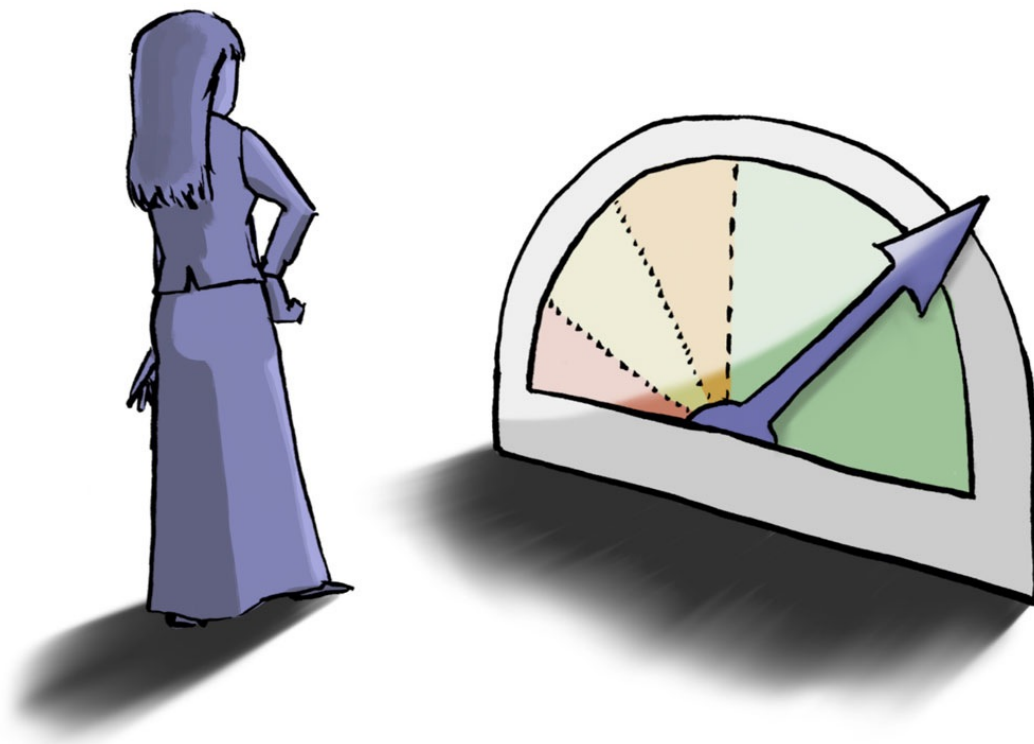


Hattie verstehen

Was John Hatties „Lernen sichtbar machen“ uns sagen will und kann



Michael Mittag

Erste Auflage

Warum Hattie lesen?	4
I. Was hat Hattie gemacht?	8
1. Hatties Arbeit ist eine Meta-Analyse	8
2. Warum Meta-Analysen zur Schulforschung notwendig sind	9
3. Hattie macht eine Meta-Analyse von Meta-Analysen zum Lernerfolg.....	13
4. Der Zwerg auf den Schultern des Riesen sieht weiter als der Riese selbst.	15
II. Was bedeutet die Effektstärke?	19
1. Eine grosse Effektstärke besagt, dass es leicht ist, Erfolge nachzuweisen....	21
2. Effektstärken bedeuten: Unterschiede verursachen Unterschiede.....	25
3. Effektstärken gelten immer „hier und jetzt“	28
4. Effektstärken sind Belege für wirksame Interventionen	32
5. Effektstärken über 0,4 zeigen „echte“ Verbesserungen.....	35
III. Fünf Effekte im Detail	38
1. Gute Lehrpersonen, guter Unterricht (Effekte um $d=0,70$)	38
Wie sieht die Gesellschaft guten Unterricht?	39
Der Mythos vom schnellen Lernen	41
Die Realität des langsamen Lernens	44
Das Schaffen von überdauernden Erinnerungen	45
2. Klassengrösse ($d=0.21$)	48
Mythos 1: Kleine Klassen steigern den Lernerfolg	49
Mythos 2: Kleine Klassen steigern das Wohlbefinden	50
Die Realität: Kleine Klassen haben viele Probleme	51
3. Video-Selbstbeobachtung (Microteaching, $d=0.88$)	52
Microteaching-Programme wirken auf vielen Ebenen	53
Microteaching ist eine wirksame Methode in der Lehrerbildung	54
4. Lob und Belohnung (negative Effektstärken)	55

Es fühlt sich gut und richtig an, zu loben und zu belohnen	55
Lob ist selten wirksam und häufig schädlich	56
Belohnungen können die Motivation zerstören	57
5. Verteiltes Lernen ($d=0.71$)	59
Exkurs: Einfach zu verstehende Faktoren erhalten mehr Aufmerksamkeit als einfach umzusetzende	60
Lernen zu verteilen ist einfach und wirksam – wenn die Umstände es zulassen	62
Literatur	63

Warum Hattie lesen?



John Hattie

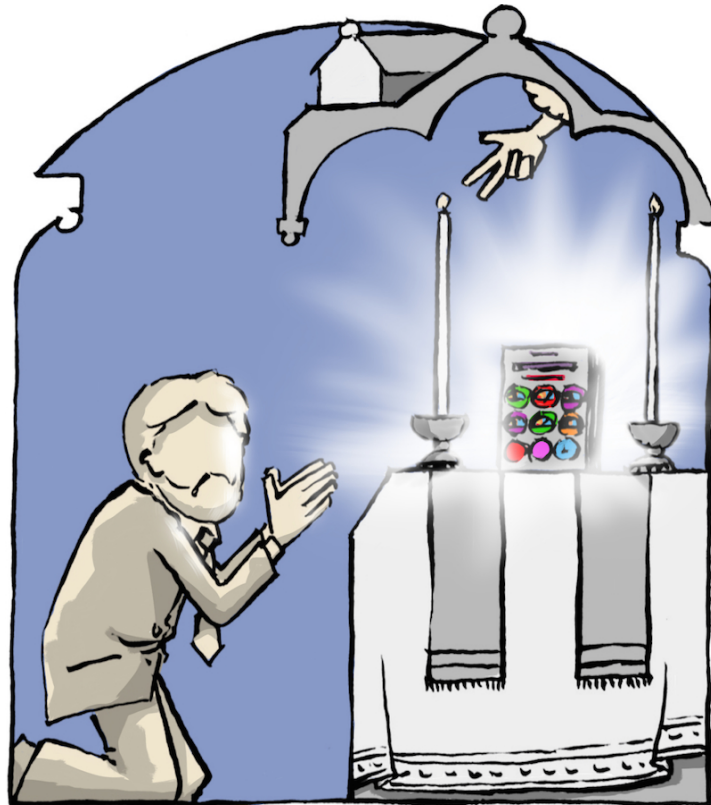
Laut New York Times ist der **heilige Gral der Erziehung** die **Antwort auf folgende Frage:**

If you could change one thing about the way our schooling system is run, what would it be?

Falls Sie eine Sache ändern könnten in Bezug darauf, wie unser Schulsystem funktioniert, was wäre es?

Warwick Mansell, Times Education Supplement, 2008

Der Artikel kommt zum Schluss, dass John Hattie mit seinem Buch „Lernen sichtbar machen“ diesen Gral endlich gefunden hat. Dass wir nach Jahrhunderten von Diskussionen jetzt endlich wissen, was wirkt.



Research reveals teaching's Holy Grail

News | Published in TES Newspaper on 21 November, 2008 | By: Warwick Mansell

Nach einer mittelalterlichen Abbildung: „Lernen sichtbar machen“ als heiliger Gral des Unterrichtens.

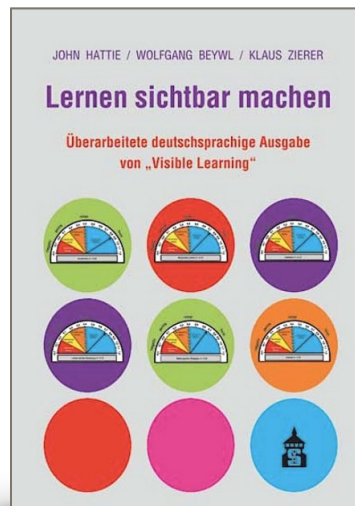
In einem gewissen Sinn ist das auch vollkommen richtig. Nur geht es nicht um eine Verkündigung, sondern um das genaue Gegenteil. **Hattie sagt uns nicht, was seiner Meinung nach wirkt, sondern er trägt zusammen, was andere herausgefunden haben.** Genauer: Was einige zehntausend Forscherinnen und Forscher an einigen zehnmillionen Versuchspersonen zum Thema Lernerfolg herausgefunden haben.

Wenn jetzt jemand sagt, dass dies möglicherweise etwas grob vereinfachend ist und es ja noch andere Dinge als Lernerfolg gibt, dann stimmt das, tut aber der Sache keinen Abbruch. Es gehört zur Wissenschaft dazu, dass man versucht, in einem begrenzten Bereich zu allgemeinen Gesetzmässigkeiten zu kommen. Ebenso, dass man sich mit den Ergebnissen kritisch auseinandersetzt, also die verwertbaren Dinge praktisch anwendet, noch offene Fragen aufklärt und vermeintlich sichere Erkenntnisse hinterfragt.

Die eigentliche **Frage ist also, was Hattie's Buch uns zu sagen hat**, wo die Grenzen sind und was sich daraus für die Praxis ableiten lässt. Das hat mit religiöser Verehrung nichts zu tun.

Wenn man die Glaubensfragen beiseite lässt, dann erweist sich Hattie als vielfältig und als zugänglich. Er spricht über (fast) alles, was Schulleistung angeht, und man muss auch nicht Mathe können, um die Ergebnisse zu verstehen.

Dieser Leitfaden ersetzt natürlich nicht die Lektüre der Bücher, die Hattie geschrieben hat:



„**Visible Learning**“ oder deutsch „**Lernen sichtbar machen**“ ist Hatties ursprüngliches Werk. Hier analysiert er **über 100 unterschiedliche Faktoren**, die das Lernen beeinflussen. Die Ergebnisse sind sehr konkret, übersichtlich und auch ohne Hintergrund in Bildungsforschung gut nachvollziehbar. **Anhang B enthält die Liste der Faktoren** nach Grösse geordnet. Dort kann man nachschauen, was wie gut wirkt. Allerdings sollte man jeweils im Buch nachlesen, ob der Faktor wirklich auch das meint, was man damit verbindet.



„**Visible Learning for Teachers**“ oder „**Lernen sichtbar machen für Lehrpersonen**“ ist Hatties Folgebuch, in dem er beschreibt, wie guter Unterricht nach wissenschaftlichen Gesichtspunkten aufgebaut ist. Es ist vor allem ein Ratgeber, wie sich

Lehrpersonen im Klassenzimmer verhalten sollen und geht nicht auf die grössere Perspektive ein – es ersetzt also nicht das erste Buch.

Beide Bücher kann man sehr gut als Nachschlagewerke im Regal stehen haben und gelegentlich hervorziehen, um sich inspirieren zu lassen oder zu schauen, was die Wissenschaft an Erkenntnissen bietet für die Dinge, die einen gerade beschäftigen. Insbesondere, wenn es Entscheidungen zu treffen gibt, welche Konsequenzen haben auf den Lernerfolg von Schülerinnen und Schülern, oder wenn man nach Inspiration sucht, wie man den eigenen Unterricht effizienter gestalten kann.

Teil 1

I. Was hat Hattie gemacht?

Hattie ist wahrscheinlich der erste berühmte Bildungsforscher. Also der erste Mensch, der mit wissenschaftlich gesicherten Erkenntnissen darüber, wie Schule funktioniert, in den Medien und in der Politik auch Beachtung fand.

Was hat Hattie also getan, das ihn als Bildungsforscher berühmt gemacht hat?

1. Hatties Arbeit ist eine Meta-Analyse

Hattie ist Schulforscher. In „Lernen sichtbar machen“ geht es allerdings nicht um seine eigene Forschung, sondern um das, was andere gemacht haben. Das ist eine so genannte Meta-Analyse, also eine **Analyse von den Analysen anderer Wissenschaftlerinnen und Wissenschaftler**. Die Idee ist, allgemeine Trends und Gesetzmässigkeiten festzustellen, die sich über viele wissenschaftliche Untersuchungen hinweg erkennen lassen.



*Eine Meta-Analyse ist der Versuch, sich über den Nebel des Alltags und der einzelnen Artikel und Meinungen zu erheben und die grossen Muster zu erkennen.
(Bild: Caspar David Friedrich, Der Wanderer über dem Nebelmeer)*

Meta-Analysen gibt es zu allen Themen, zu denen es viel Forschung gibt. Einige Beispiele sollen zeigen, wie vielfältig Zugänge und Erkenntnisse von solchen Analysen sind:

- **Klaus Grawe** hat in seiner Meta-Analyse 897 Studien zur **Wirksamkeit von Psychotherapie** zusammengefasst (Grawe, Donati & Bernauer, 1994). Er fand **fünf Faktoren, die es für eine erfolgreiche Therapie braucht**: Eine gute therapeutische

Beziehung, das Aktivieren von Ressourcen beim Patienten, das Konfrontieren mit den Problemen, das Klären der Motivation und das konkrete Bewältigen der Probleme. Grawe selbst wehrte sich gegen ein „Schulendenken“ (also dass bestimmte Therapierichtungen anderen grundsätzlich überlegen sind), geriet aber trotzdem ins Kreuzfeuer, da sich bei der Analyse zeigte, dass einige Richtungen besser durch Wirknachweise abgestützt sind als andere.

- **Irving Kirsch** hat in einer Meta-Analyse die **Effektivität von Antidepressiva** untersucht (Kirsch & Sapirstein, 1998). Er untersuchte dazu Studien der Herstellerfirmen, nachdem diese frei zugänglich gemacht wurden. Er fand in 19 grossen Doppelblind-Studien **keine Belege für eine klinische Wirksamkeit der Antidepressiva**. Wenn überhaupt ein Unterschied nachgewiesen werden konnte, dann war dieser sehr klein und beschränkte sich auf die Gruppe der sehr stark depressiven Patienten.
- **Deborah Barnes und Lisa Bero** haben untersucht, warum einige von 106 Zeitschriftenartikeln Artikel zum Schluss kommen, dass **Passivrauchen** schädlich sei, andere nicht (Barnes & Bero, 1998; Bero, 2005). Aus den vielen untersuchten Faktoren fanden sie nur einen, der die Unterschiede erklären konnte: **Autoren, welche von Tabakfirmen Honorare bekommen haben, kamen fast durchgehend zum Ergebnis, dass Passivrauchen unschädlich sei. Nicht von den Tabakfirmen bezahlte Autoren kamen überwiegend zum Schluss, dass Passivrauchen schädlich sei.**

2. Warum Meta-Analysen zur Schulforschung notwendig sind

Weltweit geschieht **sehr viel Schulforschung** und es gibt eine Unmenge von Ideen, wie man Unterricht zu gestalten hat, damit die Schülerinnen und Schüler am meisten lernen. In den grossen Fachdatenbanken finden sich mittlerweile **hunderttausende von Texten über Schule und Unterricht** (ERIC: 1.5 Mio englischsprachige Texte, FIS-Bildung: 800'000 deutschsprachige Texte).

Von der Grössenordnung her gibt es etwa so viele Forschungsartikel zur Schule wie:

- Deutschsprachige Wikipedia-Artikel (1.7 Mio)
- Lieferbare Bücher im deutschen Sprachraum (1.7 Mio, nach www.vlb.de)
- Anzahl Wörter im gesamten „Game of Thrones“-Buchzyklus (ca. 1-2 Mio, grob geschätzt)

Das mutet etwas inflationär an, aber: Jeder Forschungsartikel beantwortet eine oder mehrere pädagogische Fragen in einem bestimmten Setting, und in jedem Artikel stecken einige Monate Arbeit. **Es müsste also ausreichend pädagogisches Wissen und gesicherte Erkenntnis vorhanden sein, um die drängenden Fragen zu beantworten** (oder zumindest aus verschiedenen Richtungen auszuleuchten).

Wie wirkt sich diese Vielfalt auf die Praxis der Schule aus? – Was die Politik angeht: **sehr verhalten**. Dies soll dargestellt werden an einigen Beispielen.

Gesetzesinitiativen zur Verbesserung von Schule und Unterricht

In der Schweiz gibt es fortlaufend Bestrebungen für **Gesetzesinitiativen**, in denen das Volk bestimmt, wie die Kinder auszubilden sind. Nicht alle Initiativbegehren enden in Volksabstimmungen, und viele Abstimmungen werden verworfen, aber gelegentlich setzt das Volk relativ isolierte Aspekte der schulischen Realität fest. Die Diskussion orientiert sich üblicherweise nicht an gesicherten Erkenntnissen.



Das Niveau der Diskussion von bildungspolitischen Themen: Seifenblasen, Regenbogen, blauer Himmel. Dazwischen Angstmache und Lästern über Lehrpersonen und Bildungssysteme.

Hier einige der aktuellen oder kürzlich vergangenen Themen:

- 2014, Zürich: **Begrenzung der Klassengrösse** auf maximal 20 (in Vorbereitung)
- 2014, Wallis und St. Gallen: **Verbot von Kopftüchern** bei Schülerinnen (in Vorbereitung)
- 2012, Gesamtschweiz: **Verbot von obligatorischem Sexualekunde-Unterricht**
- 2011, Aargau: Im Kindergarten wird in **Mundart unterrichtet** (angenommen mit 54% Ja-Stimmen)

Reformbemühungen im englischen Sprachraum

Der englische Bildungsforscher (und Praktiker) **Dylan Wiliam** gibt einen Überblick über eine lange Reihe **weitgehend erfolgloser Reformprojekte** (Wiliam, 2011):

- In den USA werden **Schulen verkleinert**, um bessere Beziehung und Betreuung zu gewährleisten – bislang ohne nachweislichen Erfolg.
- In England werden wenig erfolgreiche Schulen mit erfolgreichen zu **Föderationen** (also effektiv zu grösseren Schulen) zusammengeschlossen – ebenfalls bislang ohne nachweislichen Erfolg.
- In den USA wurden so genannte **Charter Schools** eingerichtet, welche eher wie ein Wirtschaftsunternehmen als wie eine Staatsschule geführt werden. Die Effekte sind schwer einzuschätzen, wahrscheinlich führt das Bilden einer Charter School zu einer geringfügigen Verschlechterung der Schulleistungen.
- In England wurden **Schwerpunktschulen** („Specialist Schools“) gebildet, welche mit zusätzlichen Mitteln bestimmte Fächer fokussieren. Es zeigte sich, dass diese Schulen zwar etwas besser abschnitten als andere Schulen, allerdings nicht spezifisch in den fokussierten Bereichen und nicht stärker als wenn die aufgewendeten finanziellen Mittel einfach so an Schulen gegangen wären.
- 2007 wurden Schulen in London grosszügig mit **interaktiven Whiteboards** ausgestattet, allerdings ohne irgendeinen feststellbaren Effekt.



Interaktive Whiteboards erlauben es, Folien zu zeigen und mit einem Stift elektronisch darauf zu malen oder schreiben. Das ist sehr beeindruckende Technologie, trägt aber offenbar nicht entscheidend zum Lernerfolg bei.

Im **deutschen Sprachraum** gibt es eine Reihe von Schulen, welche **reformpädagogische Ansätze** in den Vordergrund stellen, welche also anstreben, Schule grundsätzlich anders anzugehen und neu zu denken. Dazu gehören demokratische Schulen, Waldorf-Schulen oder Montessori-Schulen. Diese Schulen haben in einzelnen Bereichen grosse Erfolge vorzuweisen, welche teilweise auch ins staatliche Schulwesen übernommen

werden, aber bei all dem, was ausprobiert wurde, gibt es offenbar auch nichts, was der klassischen Schule in allen Belangen klar überlegen wäre. Also auch hier: Eine komplexe Sachlage, keine einfachen Lösungen.

Privatisierung der Bildungsinnovation

Eine ähnliche Reformwelle findet unter ganz anderen pädagogischen Gesichtspunkten derzeit in den USA statt. **Viele Schulen werden von Stiftungen oder direkt von einzelnen Milliardären finanziell unterstützt**, oft in wohlthätiger Absicht und mit positiven Folgen – die Schulen können das zusätzliche Geld gut gebrauchen. **Allerdings stehen dabei häufig spezifische weltanschauliche Ziele im Vordergrund.** Der Facebook-Gründer Mark Zuckerberg wirft beispielsweise in Newark 120 Millionen Dollar auf, um dort die Qualität des Unterrichts zu verbessern. Oder genauer, ganz spezifische Bereiche der Schulqualität, wie auf der Website der Stiftung nachzulesen ist:

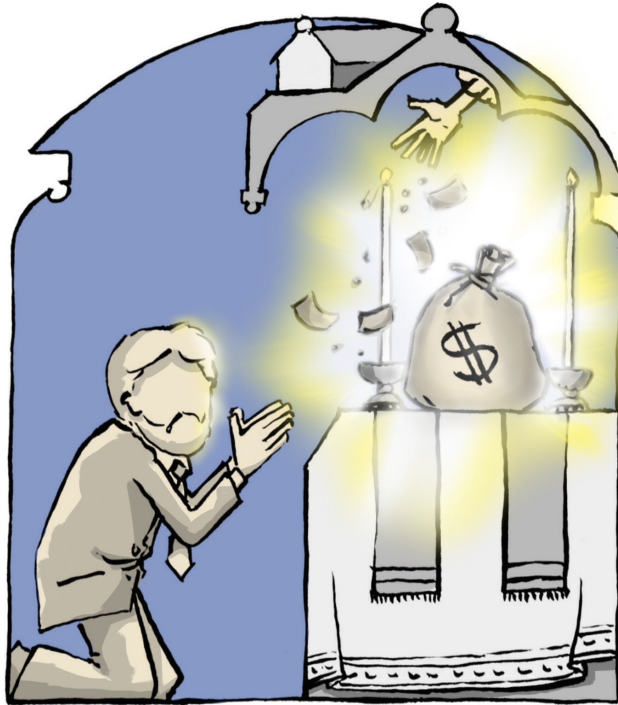
Local educators and community leaders have told us that their priorities are encouraging innovation in the classroom, helping to train a new generation of leaders, and supporting student development. Accordingly, the initial grants will go towards initiatives that help provide computers and connectivity in schools, along with teacher training and parent outreach to make these a really valuable addition to the learning experience.

Lokale Auszubildende und Community-Sprecher haben uns gesagt, ihre Prioritäten liegen darin, Neuerungen im Klassenzimmer zu ermutigen, eine neue Generation von Anführern heranzubilden und die Entwicklung der Studierenden zu fördern. Entsprechend gehen die ersten Investitionen an Initiativen, welche Computer und Vernetzung in die Schulen bringen, sowie an Lehrpersonen-Ausbildung und Elternkommunikation, um diese zu einer wertvollen Unterstützung der Lernerfahrung zu machen.

– Text von www.startupeducation.org

Zuckerberg unterstützt also Bildungsinnovation und gibt den Schulen dringend benötigte Finanzmittel, allerdings setzt er weitgehend seine eigene Vorstellung davon, wie Schule auszusehen hat: mehr Computer, mehr elektronische Vernetzung, Führungspersönlichkeiten heranbilden. Insbesondere die Forderung nach einer „neuen Generation von Anführern“ ist üblicherweise nicht das, was Lehrpersonen oder Lokalpolitiker in finanzschwachen Communities als primäre Bildungsanliegen vorbringen.

Derzeit ist die Situation so weit abgeglitten, dass man in den USA von einer **schleichenden Privatisierung der Bildungsinnovation** spricht. Gemeint ist, dass ständig irgendwelche Milliardäre ihre ganz privaten Vorstellungen davon, wie man Schule verbessert, an der Allgemeinheit ausprobieren, indem sie einfach Geld fließen lassen.



Grundsätzlich ist es positiv, wenn reiche Leute das Bildungssystem unterstützen und dringend benötigtes Geld in Schulen stecken. Häufig sind allerdings bestimmte Vorstellungen damit verbunden, wie Schule auszusehen hat.

Fazit

Offenbar **existiert also sehr viel Forschung zu Schule**, andererseits eine gewisse **Ratlosigkeit oder schlicht ein Unvermögen, diese auch wirksam anzuwenden**. Es gibt Berge von gesichertem Wissen darüber, wie Schule funktioniert, was man bereits ausprobiert hat und was dabei herausgekommen ist. Im konkreten Fall entscheidet man aber oft mehr oder weniger intuitiv und setzt mit grossem Aufwand irgendwelche Dinge um, die keine Wirkung haben. Was sehr häufig auch voraussehbar war.

3. Hattie macht eine Meta-Analyse von Meta-Analysen zum Lernerfolg

Angesichts der Fülle von Reformbemühungen einerseits und gesicherten Ergebnissen andererseits liegt es nahe, Ordnung zu schaffen. Genau das hat Hattie getan, und er hat 15 Jahre dafür aufgewendet, das bestehende Wissen zusammenzutragen. Hatties Ansatz ist wie folgt:

- Er interessiert sich nur für den **Lernerfolg**.
- Er interessiert sich nur für **Meta-Analysen** (also für bereits bestehende Zusammenfassungen der Forschung zu einem Thema).

Hattie fand **800 Meta-Analysen** zu rund **140 verschiedenen Faktoren**. Als „Faktor“ bezeichnet er eine Einflussgrösse, wie etwa die Klarheit der Lehrperson oder die finanzielle Ausstattung des Elternhauses. Hattie hat für jeden Faktor eine so genannte **Effektstärke** berechnet, welche angibt, **wie viel Einfluss der Faktor auf den Lernerfolg hat**, wenn man die gesamten zur Verfügung stehenden Ergebnisse berücksichtigt.

Natürlich ist die Meta-Analyse ein **grobes Instrument**. Hinein gehen zahlreiche Studien, in denen intelligente Menschen mit viel Energie versuchen, ein Thema anders anzugehen als die Forscherinnen und Forscher vor ihnen. Heraus kommt im Extremfall eine einzige, nüchterne Zahl.

Es versteht sich von selbst, dass die Feinheiten verloren gehen. Heraus kommt ein grobes Bild über alles, sozusagen ein **Satellitenfoto über die Bildungsforschung** – jedenfalls soweit sie sich mit Lernerfolg beschäftigt. Das dient dazu, herauszufinden, wo etwas interessantes zu finden sein könnte. Bevor man sich in wilde Spekulationen oder blinden Aktionismus stürzt, sollte man auf jeden Fall genauer hinschauen oder das vom Boden aus nachprüfen. Sprich: Bei Hattie nachlesen, worauf sich seine Aussagen tatsächlich abstützen und gegebenenfalls auch die Originalartikel beiziehen.



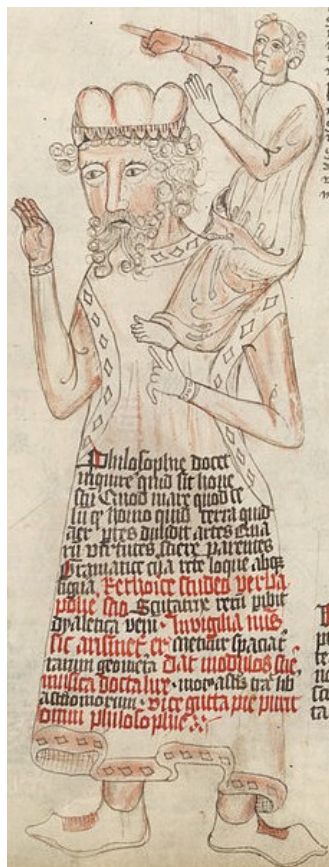
Das „Marsgesicht“ sieht von Weitem sehr spektakulär aus (links), wenn man dann genauer hinschaut (rechts), dann ist es eben doch einfach ein Hügel wie andere auch.

Also: **Hattie liefert uns eine Übersicht, und wir brauchen diese**, um zu wissen, was wir tun sollen und wo es sich lohnt, aktiv zu werden. Wenn wir als Lehrpersonen eine Stunde vorbereiten oder etwas neues ausprobieren wollen, dann können wir ja nicht jedes Mal die gesamte Forschungslage lesen, und wir wollen ja auch nicht einfach irgend etwas ausprobieren, sondern Dinge tun, deren Wirkung auch belegt ist.

4. Der Zwerg auf den Schultern des Riesen sieht weiter als der Riese selbst

Die Auseinandersetzung mit Hatties Ergebnissen gestattet uns einen **Weitblick**, den wir sonst nicht erhalten. Es geht also nicht darum, ob man Hatties Folgerungen blind übernehmen oder ebenso blind verdammen sollte, sondern darum, sie zu benutzen, um eine eigene, umfassende Perspektive auf die Bildungslandschaft zu gewinnen.

Dafür gibt es ein **mittelalterliches Bild**: Das vom **Zwerg auf den Schultern des Riesen. Der Zwerg bin ich mit meinem bodennahen Erfahrungswissen** davon, wie Schule und Unterricht funktioniert. **Der Riese ist nicht Hattie, sondern es sind diejenigen Denker und Forscher, die vor uns waren.**



Mittelalterliche Abbildung, ca. 1410 (Manuskript „Rosenwald 4“)

Pigmaei gigantum humeris impositi plusquam ipsi gigantes vident.

Zwerge, die auf den Schultern von Riesen stehen, sehen weiter als die Riesen selbst.

– Didacus Stella, 16. Jahrhundert

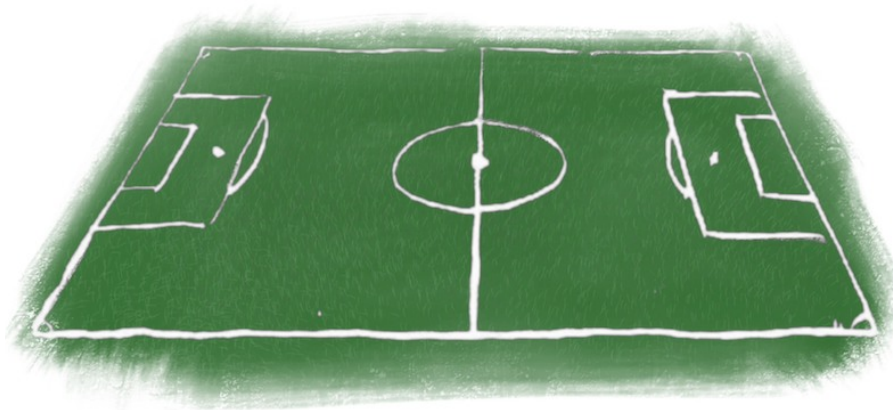
In der **mittelalterlichen Theologie und Wissenschaft**, die übrigens wesentlich weiter entwickelt war, als das manchmal dargestellt wird, beantwortete das Bild eine **grundlegende Frage: Wenn die antiken Philosophen bereits alle wesentlichen**

Bedingungen der Existenz ergründet hatten und die Bibel die ewige Wahrheit offenbart, warum forschen wir weiter? – Die Antwort wird als Gleichnis gegeben: Wir sind zwar nur Zwerge angesichts all der Riesen, die vor uns waren, aber **ein Zwerg, der auf den Schultern des Riesen steht, sieht weiter als der Riese selbst.**

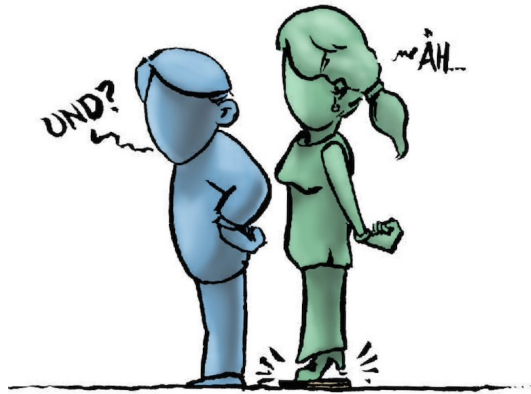
Auf unsere Verhältnisse übertragen: **Wenn eine Bachelorarbeit an ein internationales Forschungsprojekt anknüpft, dann kann die Bachelorarbeit zu Erkenntnissen gelangen, welche dem internationalen Forschungsprojekt verwehrt sind.** Wenn ich Hatties Ergebnisse bei mir im Klassenzimmer anwende, dann gelange ich zu Lernerfolgen, welche über das hinausgehen, was aus dem blinden Anwenden von irgendwelchen Ratschlägen möglich ist. Und auch weiter als das, was herauskommt, wenn ich nur mit meiner begrenzten persönlichen Einsicht handle.

Entscheidend ist in diesem Gleichnis nur, wie hoch man den Kopf am Schluss hat. Und diese Höhe wollen wir in einem weiteren (hoffentlich nicht allzu strapazierten) Gleichnis und Zahlenspiel ausloten.

Gehen wir einmal davon aus, dass es in den deutsch- und englischsprachigen Datenbanken **ungefähr 2 Millionen Texte zur Schulforschung** gibt (PEDOCS, ERIC), also ganz grob geschätzt 20 Millionen Seiten Text. Wenn man das ausdrucken und auf einem Fussballfeld verteilen würde, dann wäre das Fussballfeld etwa 60 Seiten hoch mit Forschungstexten bedeckt.

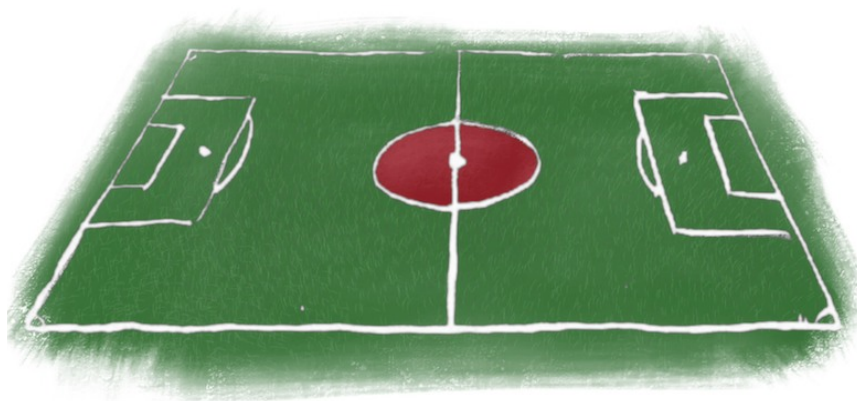


Das ist sehr viel Material, aber wenn es da einfach gleichmässig verteilt herumliegt, dann verbessert das meine Weitsicht nicht beträchtlich.



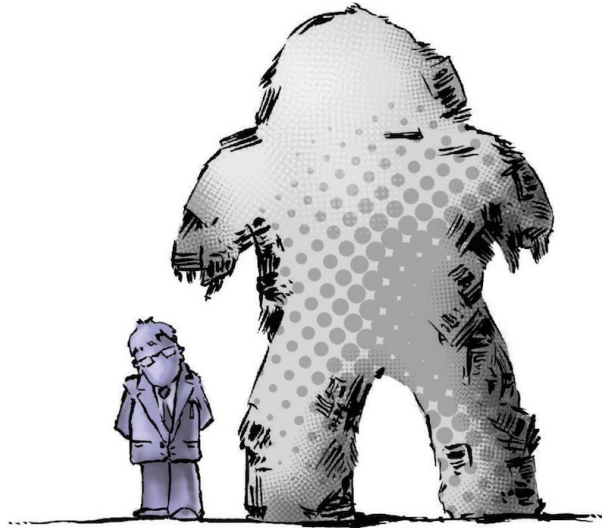
Ein Zwerg, der auf einem Forschungsartikel steht, sieht noch ungefähr genau gleich weit wie jeder andere Zwerg auch.

Was Hattie gemacht hat, ist folgendes: Er hat einen Riesen gebaut. Dazu hat er **50'000 Texte zusammengefasst und strukturiert**, so dass wir etwas haben zum Draufstehen. Wenn die gesamte Schulforschung gleichmässig über das Fussballfeld verteilt ist, dann sind das **etwa so viele Texte, wie im Mittelkreis Platz finden:**

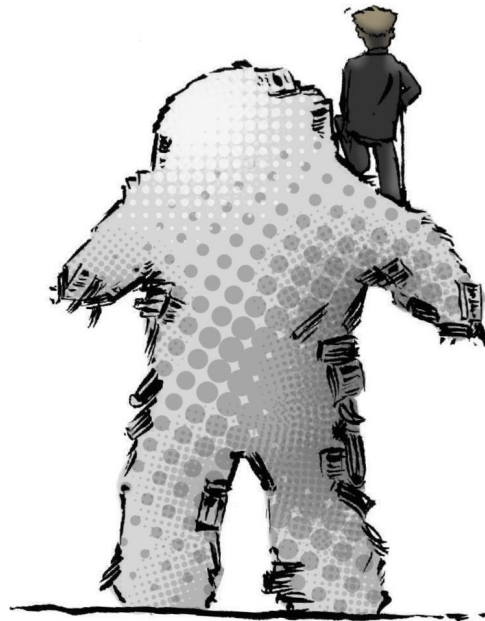


Es gibt sehr viel Forschung, die Hattie nicht abdeckt. Das sind Texte, die entweder nicht besonders spannend sind (etwa ein Bericht über ein Symposium vor 10 Jahren) oder die sehr spannend sind, aber mit anderen Themen als Lernerfolg zu tun haben. Übrig bleiben immer noch sehr viele Arbeiten, viel zu viele, um sie alle zu lesen.

Deshalb strukturiert sie Hattie, er baut sich also einen Riesen daraus. Wie gross ist jetzt dieser Riese? – Wenn man 50'000 ausgedruckte Zeitschriftenartikel zu einem ungefähr menschenförmigen Gebilde zusammenfügt (und dabei keine Luft zwischen den Blättern lässt), dann hat dieses Gebilde etwa 5-10 Meter Höhe (nach ungefähren Berechnungen aufgrund von Daten in Wikipedia und Wolfram Alpha).



Wenn wir jetzt über Bildung nachdenken, dann stehen wir auf den Schultern von 50'000 Forschungstexten und können dadurch Dinge sehen, die uns auf dem Boden verborgen blieben. Wir stehen auf den Schultern eines Giganten, den uns Hattie aus dem gesammelten Wissen der Forschung zum Thema Schulleistung zusammengebaut hat.



In den folgenden zwei Teilen werden wir zunächst besprechen, wie die Welt von den Schultern des Riesen herunter aussieht, was wir sehen, was wir nicht sehen können, und mit welchen Verzerrungen wir umgehen müssen. Dann werden wir einzelne Inhalte vertieft diskutieren, um zu sehen, wie wir eine Brücke schlagen können zwischen den Ergebnissen von Hattie und der Wirklichkeit der Schule.

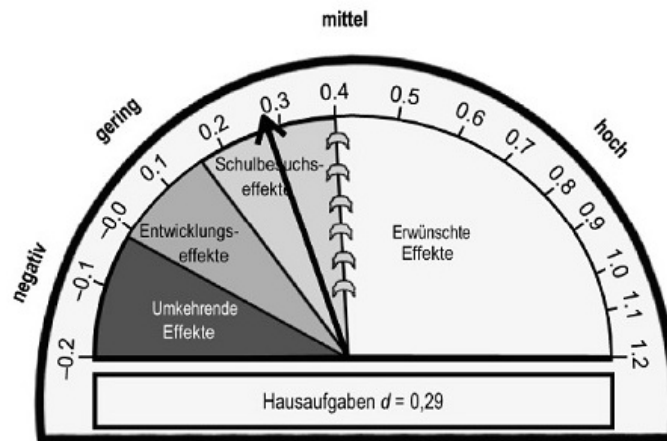
Teil 2

II. Was bedeutet die Effektstärke?



„Faktor“ kommt vom Lateinischen „facere“ (machen) und war früher die Bezeichnung von Handelsherren wie Georg Giese auf diesem Bild von Holbein. Holbein bildet hier ab, was ein Faktor macht. Hattie macht eigentlich auch nichts anderes, aber mit anderen Faktoren und anderen Abbildungen.

Ein **Faktor** ist bei Hattie eine **Variable, die mit Lernerfolg zusammenhängt**, wie zum Beispiel Klassengröße, Fachkompetenz der Lehrkraft, Motivation der Lernenden, verschiedene Fördermassnahmen oder unterschiedliche pädagogische Unterrichtskonzepte. Hattie hat zu jedem Faktor eine **Effektstärke** berechnet, die sagen soll, wie wirksam der Faktor sich in der bisherigen Forschung erwiesen hat. Also: **Wie stark sich die Lernleistung verbessert, wenn der Faktor vorhanden ist.**



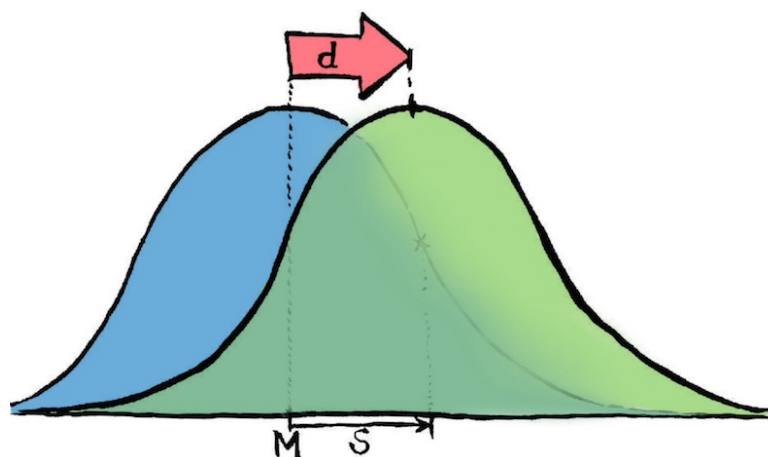
Hattie stellt Wirkung als Tacho dar: Alles, was rechts ist (über 0.4) sind ausserordentlich grosse Effekte, Effekte unter 0 bedeuten eine Verschlechterung

Die **Interpretation eines Faktors** ist zunächst ganz einfach: Eine grössere Effektstärke weist auf einen grösseren Unterschied hin. Also: **Mehr ist mehr**. Das ist auch eine der Stärken bei Hatties Forschung: Man kann sie spontan verstehen.

Konkret: Mentoring hat eine Effektstärke von $d=0.15$, Peer Tutoring $d=0.55$. Peer Tutoring hat also in der bisherigen Forschung grössere Unterschiede im Lernerfolg ausgemacht als Mentoring. Weil 0.55 eine grössere Zahl ist als 0.15.

Natürlich sollte man an dieser Stelle etwas in die Tiefe gehen. Um zuerst einmal zu verstehen, was die Effektstärke eigentlich ist, haben wir einen **Film** produziert. Er wird an geeigneter Stelle zugänglich gemacht werden und wahrscheinlich auf dem Internet leicht zu finden sein, wenn man nach „Effektstärke“ und/oder „Hattie“ sucht.

An dieser Stelle soll eine kurze Zusammenfassung genügen: Die Effektstärke ist ein **Mass, um Unterschiede auszudrücken**, und zwar unabhängig davon, wie diese Unterschiede gemessen wurden. Das braucht Hattie, da ja fast alle Untersuchungen verschiedene Methoden verwenden, um Lernerfolg zu messen. Eine Verbesserung von 10 Punkten kann also je nach Skala und Test ganz unterschiedliche Bedeutungen haben.



Kurz zusammengefasst: Die Effektstärke d sagt, wie stark sich eine Verteilung der Werte verschiebt bei einer Massnahme. Für mehr Informationen ist es am besten, den Film anzuschauen.

Die wichtigsten Punkte sind:

- Eine Effektstärke von **$d=0$** bedeutet, dass es **keinen Effekt** gibt.
- Effektstärken **können auch grösser als 1 sein**, das ist in der Praxis aber **sehr selten**. Eine Effektstärke **von 0.8 ist bereits sehr gut** und steht für einen grossen Zuwachs im Lerngewinn.
- Effektstärken kann man auch verstehen, ohne dass man die Formel weiss oder die Definition kennt: **Grössere Effektstärken bedeuten grössere Unterschiede**, also mehr Lerngewinn.

Die Effektstärke sagt also aus, wie gross der Unterschied ist, den Wissenschaftler bei einer Massnahme beobachtet haben. Können wir die Zahlen jetzt einfach nehmen und blind diejenigen Dinge umsetzen, welche hohe Effektstärken haben? – Dazu ist die Sachlage leider etwas zu komplex, weshalb wir uns in den nächsten Abschnitten überlegen wollen, was eine hohe Effektstärke wirklich bedeutet.

1. Eine grosse Effektstärke besagt, dass es leicht ist, Erfolge nachzuweisen

Die Effektstärke ist ein mittlerer Wert davon, was bei den Meta-Analysen herausgekommen ist. Dafür erntet Hattie teilweise Skepsis: **Kann man wirklich Forschung mitteln und damit zu gültigen Aussagen gelangen?** Wahrscheinlich verkörpern ja nicht alle Meta-Analysen und Einzelarbeiten zu einem Thema genau das, was wir uns inhaltlich und methodisch idealerweise wünschen würden, und bestimmt nicht alle dasselbe. Können wir uns auf Hatties Ergebnisse irgendwie verlassen?

Kleiner Exkurs: René Descartes hat in einem Zustand ähnlicher Verwirrung den Satz „Ich denke, also bin ich“ geprägt. Das heisst: Man kann eigentlich alles anzweifeln ausser die eigene Existenz. Denn wenn ich nicht existieren würde, dann könnte ich ja auch keine Dinge anzweifeln. Da ich zweifle (oder nur schon denke), muss ich auch existieren.



René Descartes (1596-1650)

Was ist bei Hattie der Punkt, wo wir uns unzweifelhaft einig sind? Die eine Aussage, auf die wir uns uneingeschränkt verlassen können? Es ist diese hier:

Eine grosse Effektstärke besagt, dass es leicht ist, Erfolge nachzuweisen.

Das heisst: Ich kann eigentlich an allem zweifeln, was den Inhalt des Faktors angeht. Grosse Effektstärken bedeuten aber, dass oft grosse Effekte gefunden werden, wenn man den Faktor untersucht. Also: Dass es einfach ist, Erfolg zu haben mit Forschung zu diesem Faktor.



Für **grosse Effektstärken** liegt es auf der Hand, was das bedeutet: Nämlich dass die **Wirkung gross, zuverlässig und offensichtlich** ist. Das heisst, dass sich jedes Mal, wenn jemand das Thema angeht, auch eine nachweisliche Verbesserung einstellt.

Bei kleinen Effektstärken müssen wir sehr viel vorsichtiger sein. Es gibt **viele Gründe**, weshalb man einen Effekt nicht findet. Der offensichtlichste Grund ist, weil der Effekt nicht existiert, also:

- Der Faktor hat nur einen **geringen Einfluss** auf die Lernleistung.

Genau so gut ist es aber **möglich, dass der Effekt existiert, dass wir ihn aber einfach nicht sehen.** Das kann mehrere Ursachen haben:

- Die Forschenden haben in ihrer Untersuchung den Faktor **schlecht umgesetzt**. Also: Der Faktor würde wirken, wenn man die Sache nur richtig angehen würde.
- Es gibt allgemeine Umstände, die eine **erfolgreiche Umsetzung erschweren**, auch wenn die Forschenden alles korrekt machen.
- Der Faktor **zieht andere Dinge nach sich**, die seine Wirkung wieder schmälern.
- Die Forschung wurde **nicht mit ausreichend Ausdauer, methodischer Kompetenz und finanziellen Möglichkeiten durchgeführt**, um Effekte nachweisen zu können.



*Salopp gesagt: Wenn ich irgendwo eine Ziege sehe, dann ist auch eine Ziege da.
Wenn ich keine Ziege sehe, dann kann das viele Ursachen haben – es bedeutet nicht, dass es in dieser Gegend grundsätzlich keine Ziegen gibt oder geben kann.*

Beispiel: Feedback ($d=0.73$) vs. offener Unterricht ($d=0.01$)

Der obige Gedankengang soll am Beispiel der beiden Faktoren „Feedback“ ($d=0,73$) und „offener Unterricht“ ($d=0,01$) illustriert werden.

Zu den Begriffen: **Feedback** ist bei Hattie ein sehr komplexer Begriff, der eine Vielzahl von Methoden umfasst, wie Lehrende und Lernende Informationen über einander austauschen (s. 206). Wichtig ist, dass gerade die gängigen Praktiken wie von der Lehrperson individuell ausgeteiltes Lob, Belohnung und Bestrafung eher nicht dazu gehören und von Hattie auch als unwirksam angesehen werden (s. 207). Es geht viel stärker darum, dass Lernende und Lehrende reichhaltige und präzise Informationen über das Lernen und den Lernstand austauschen. **Offener Unterricht** ist bei Hattie eine Sammlung pädagogischer Ansätze, die stark auf dem individuellen Lernen des Kindes fokussiert und dafür die Klassenstruktur auflöst oder zumindest in den Hintergrund stellt (s. 105).

Für Feedback bedeutet die **hohe Effektstärke** mit ziemlicher Sicherheit: **Feedback wirkt**. Die Forschenden haben offenbar sehr häufig Leistungsverbesserungen gefunden,

wenn sie Feedback untersucht haben. Wenn Feedback unwirksam wäre, dann liesse sich nur sehr schwer erklären, wo die gefundenen Unterschiede herkommen.

Für die **geringe Effektstärke des offenen Unterrichts** gibt es **verschiedene Erklärungen**, die wir in der oben bereits vorgebrachten Reihenfolge systematisch durchgehen:

- Die einfachste Erklärung ist, dass offener Unterricht **nicht wirksam** ist, um besseren Lernerfolg zu erzielen. Fairerweise muss man jedoch sagen, dass die Leistungen im offenen Unterricht auch nicht schlechter werden. Darum sollte die tiefe Effektstärke auch niemanden davon abhalten, offenen Unterricht zu machen.
- Vielleicht ist offener Unterricht sehr **schwierig umzusetzen**, und die Forschenden sind daran gescheitert. Das heisst: Richtig durchgeführt kann der offene Unterricht die Lernleistung möglicherweise steigern, aber wir wissen noch nicht, wie man offenen Unterricht richtig macht.
- Vielleicht sind es die **Umstände**: Möglicherweise sind die Lehrpersonen nicht ausreichend für offene Unterrichtsformen geschult oder motiviert, vielleicht sind die Schulzimmer dafür ungeeignet, oder die Klassen sind noch zu sehr den konventionellen Unterricht gewohnt. Also: Wenn man hinget und offenen Unterricht umzusetzen versucht, dann gibt es plötzlich Probleme, die eigentlich nichts mit dem pädagogischen Konzept zu tun haben, aber die Umsetzung erschweren.
- Möglicherweise geht offener Unterricht mit **zusätzlichen Problemen** einher, welche die Vorteile der Methode aufheben. Vielleicht ist diese Unterrichtsform aufwendiger in der Planung, oder die Kinder können leichter Arbeit vermeiden, so dass höherer Aufwand für Kontrollen und das Erhalten der Lernmotivation anfällt.
- Vielleicht liegt der offene Unterricht nicht so stark auf der Linie der Hochschulen, so dass **Studien weniger gut ausgestattet** sind und von **weniger routinierten Forschenden** durchgeführt werden.



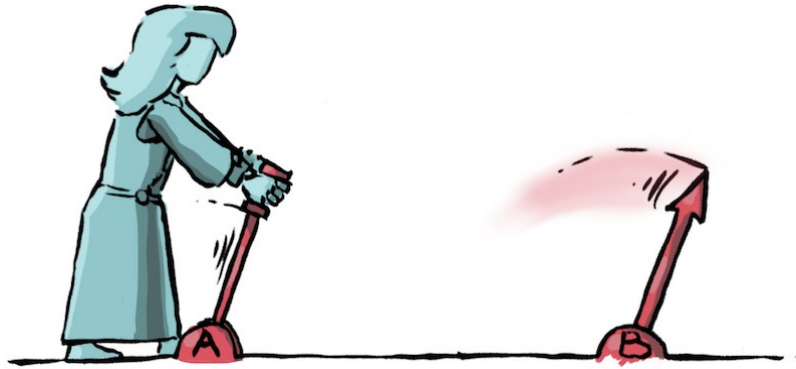
Mögliche Erklärung: Guter offener Unterricht ist sehr schwierig umzusetzen, so dass die negativen Aspekte und zusätzlichen Aufwände die potenziellen Gewinne wieder aufheben.

Fazit: **Die tiefe Effektstärke bedeutet nicht, dass wir belegt haben, dass offener Unterricht wirkungslos ist.** Das ist *eine* mögliche Erklärung. Mit der Einschränkung, dass „wirkungslos“ hier heisst: ebensogut wie anderer Unterricht auch. Es kann auch einfach sein, dass es schwierig ist, offenen Unterricht so gut zu machen, dass man bessere Lernleistungen nachweisen kann.

Beim Feedback ist die Erklärung hingegen recht geradlinig: Die hohe Effektstärke bedeutet ziemlich sicher, dass Feedback einen grossen Einfluss auf die Lernleistung hat, und dass dieser Einfluss sich auch sehr zuverlässig einstellt. Das heisst: Ich kann Feedback wahrscheinlich recht einfach auch bei mir im Klassenzimmer einsetzen, während ich beim offenen Unterricht vorsichtiger sein muss, da es offenbar gar nicht so leicht ist, damit eine Verbesserung zu erreichen.

2. Effektstärken bedeuten: Unterschiede verursachen Unterschiede

Eine Effektstärke bedeutet grundsätzlich: **Wenn sich A verändert, dann verändert sich auch B.** In „Lernen sichtbar machen“ geht es um Schulleistung, das heisst: „B“ ist jeweils Schulleistung. Die Frage ist, welche Faktoren eine Auswirkung haben auf die Schulleistung. Also: **Bei welchen Faktoren verbessert sich die Schulleistung, wenn sich der Faktor ändert?**



Idealerweise führe ich als Lehrperson eine Verbesserung in einem Faktor (A) durch, und der Lernerfolg (B) nimmt entsprechend zu.

Der Haken dabei ist: **Was bedeutet „eine Änderung“?** Wie gross ist diese, und von welchem absoluten Niveau sprechen wir dabei?

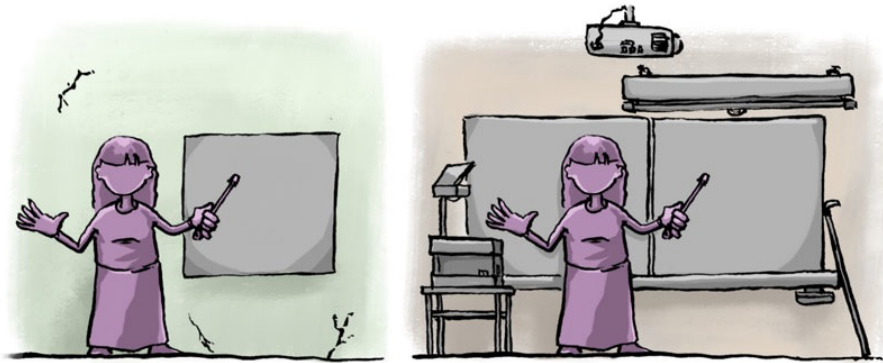
Nehmen wir hierfür ein Beispiel zur Hand: Die finanzielle Ausstattung der Schule ($d=0,23$). Also: Die finanzielle Ausstattung kann einen Unterschied von 0.23 Standardabweichungen ausmachen und hat damit einen eher geringen Einfluss auf den Lernerfolg.

Das heisst nicht, dass eine Schule ohne jegliche finanzielle Ausstattung immer noch ungefähr gleich gut funktioniert wie eine gut ausgestattete Schule. Es heisst nur, dass die vorgefundenen oder im Rahmen der Studien verursachten Unterschiede in der Finanzierung keine bedeutenden Unterschiede im Lernerfolg ausgelöst haben.



Eine tiefe Effektstärke für „finanzielle Ausstattung“ bedeutet nicht, dass man die gleichen Lernerfolge auch ganz ohne finanziellen Aufwand erreichen kann.

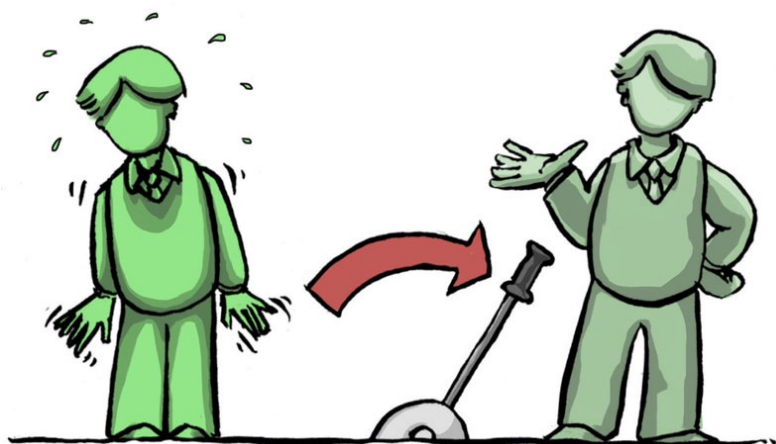
Ein sehr einfacher Grund für die kleine Effektstärke könnte sein, dass Schulen schon recht gut ausgestattet sind, so dass **zusätzliches Geld keinen grossen Einfluss** hat. Oder zumindest, dass alle Schulen einigermaßen vergleichbar ausgestattet sind, was die wirklich wichtigen Dinge angeht. Auch wenn die eine Schule etwas moderner oder grosszügiger eingerichtet ist als die andere, so können doch beide ihre Lehrpersonen bezahlen, sie haben ausreichend Räume für den Unterricht, die Kinder haben Hefte und Bücher und so weiter.



Sofern Schulen über Lehrpersonen und Unterrichtsräume verfügen, sind sie wohl einigermaßen vergleichbar, was die Lernleistung angeht. Zusätzliche Geräte und Ausstattungen können das Unterrichten zwar angenehmer machen, haben aber nur noch geringen Einfluss auf die Lernleistung.

So gesehen ist es sehr schwierig, einen Unterschied durch finanzielle Ausstattung nachzuweisen oder hervorzurufen. Nicht weil Geld grundsätzlich unwichtig ist, sondern weil das Niveau der finanziellen Ausstattung so hoch ist, dass zusätzliches Geld nicht zu deutlich besserer Lernleistung führt.

Umgekehrt zeigen einige Faktoren **sehr hohe Effektstärken, bei denen das allgemeine Niveau der Umsetzung sehr tief ist**, die also in der derzeitigen Schulrealität ein Randdasein fristen. Dazu gehören etwa das Microteaching ($d=0,88$), also das Ausbilden von Lehrpersonen mit Video-Selbstbeobachtung. Das ist derzeit nicht Standard in der Lehrpersonen-Ausbildung. Wenn wir also die Effektstärke von Microteaching anschauen, dann vergleichen wir Probanden ohne Video-Selbsterfahrung mit solchen, welche dazu ein intensives Training durchlaufen haben.



Microteaching-Programme verwandeln Personen ohne Auftrittserfahrung in sichere, routinierte und geschulte Akteure. Dieser grosse Unterschied im Faktor schlägt sich in entsprechend grossen Effektstärken nieder.

Das schmälert nicht den Effekt, es macht ihn aber leichter nachweisbar, da die untersuchten Unterschiede sehr gross sind. Aufgrund der Ergebnisse von Hattie und der Verfügbarkeit von Aufzeichnungsgeräten kann es gut sein, dass Video-Selbstbeobachtung in Zukunft stärker zum Standard in der Ausbildung wird, und wenn es einmal so ist, dann wird

die Effektstärke spezieller Microteaching-Programme auch geringer. Nicht weil das Auftreten weniger wichtig ist, sondern weil das allgemeine Niveau höher ist und zusätzliche Verbesserungen im Auftreten nicht mehr zu gleich grossen Verbesserungen der Lernleistung führen.

Bei Hattie gibt es auch einige Faktoren, wo sehr grosser Aufwand nötig ist, um etwas zu verändern. Dazu gehört zum Beispiel die Klassengrösse ($d=0,21$). Hattie differenziert hier noch weiter (s. 102 sowie Hattie, 2006): Bei einer Reduktion von 25 auf 15 Lernende pro Klasse kann man mit einer Verbesserung von ungefähr $d=0,1$ bis $d=0,2$ rechnen. Kleinere Klassen führen also zu etwas besseren Lernleistungen, allerdings ist die Verbesserung auch bei einer drastischen Reduktion der Klassengrösse nur sehr gering.



Bei einigen Faktoren bewirken auch sehr grosse Änderungen nur minimale Verbesserungen der Lernleistung.

3. Effektstärken gelten immer „hier und jetzt“

Effektstärken gelten grundsätzlich für den untersuchten Zeitraum und die untersuchte Kultur. Einige Beispiele sollen das illustrieren.

Technologie hier und in ökonomisch benachteiligten Regionen in Indien

Bei Hattie ist sind **Effektstärken für Technologie und Ausstattung sehr tief**: $d=0,16$ für technologiegestütztes Lernen zu Hause, $d=0,18$ für webbasiertes Lernen, $d=0,22$ für visuelle und audiovisuelle Medien. Klassenzimmer mit Computern, Beamern und Smartboards auszustatten führt also bei uns nicht zu gesamthaft höheren Leistungen; in früheren Studien sanken die Lernleistungen sogar bisweilen, wenn Dias oder Tonbänder eingesetzt wurden (Willet, Yamashita & Anderson, 1983).

In Indien gab es jedoch **enorm erfolgreiche Projekte** in dieser Richtung: Professor **Sugata Mitra** mauerte in Armenvierteln Computer in die Wand ein, mit denen die Bewohner Zugang zum Internet erlangten. Insbesondere Kinder nutzten dies rege und lernten selbständig, wie man Computer bedient (Mitra, 2013).



Ein so genanntes „Hole in the Wall“, also ein Loch mit einem internetfähigen Computer.

Mitra verwendete die Computer auch, um die **Kinder in Indien mit Erwachsenen aus englischsprachigen Ländern zusammenzubringen**. Diese bieten Unterstützung und Austausch, teilweise auch formalen Unterricht. Mitra setzt dabei auf selbständiges und selbstkontrolliertes Lernen, er gibt also keine Bildungsziele vor. Dadurch lässt sich der Erfolg nicht ganz so einfach messen wie in formalen Bildungssystemen, dafür erhalten die Probanden die Bildung, welche sie für wichtig erachten.



Abbildung: Ein Mitglied der so genannten „Granny Cloud“ unterhält sich mit Kindern in einem indischen Armenviertel.

Hier haben wir also zwei Faktoren, die für ökonomisch benachteiligte Kinder in Indien sehr grosse Effekte zu haben scheinen, und welche bei uns wahrscheinlich nur geringe Lerneffekte bewirken würden:

- Freier, unregelmäßiger Zugang zu internetfähigen Computern.
- Informeller Austausch mit älteren Personen aus anderen Kulturkreisen.

Natürlich können auch bei uns mit Computern und Internet Bildungsziele erreicht werden, und der Austausch mit Menschen aus anderen Kulturen ist auch hier eine

Bereicherung. Aber irgendwie scheint es wenig aussichtsreich, damit die Lernleistung unserer Schülerinnen und Schüler entscheidend zu optimieren.

Unterricht gestern, heute und morgen

Nebst kulturellen Unterschieden können auch zeitliche Entwicklungen dazu führen, dass sich die Wirkung von Faktoren ändert.

Geschlecht ($d=0,12$) hat heute **sehr geringe Effektstärken** in fast allen untersuchten Bereichen. Eine ausführliche Diskussion der untersuchten Variablen findet sich bei Hattie im Buch (s. 66); auch für klassische Geschlechtsstereotype wie Kommunikationsfähigkeit, Hilfsbereitschaft, Führungsstärke, Aggression oder Interesse für bestimmte Schulfächer liegen die Effektstärken kaum je über 0,20 (Übersicht in Hyde, 2005; ausserdem Hyde 1981; Hyde & Lynn, 1988; Freeman, 1984; Friedman, 1989; Frost, Hyde & Fennema, 1994, Hines, 1989; Hyde, Fennema & Lamon, 1990). Natürlich können wir nur spekulieren, aber wahrscheinlich hätte das vor 20, 50 oder 100 Jahren noch ganz anders ausgesehen und wir hätten sehr viel grössere Geschlechtsunterschiede festgestellt. Hier hat wohl ein Wechsel stattgefunden, so dass wir zwar die Stereotypen noch in den Köpfen haben, sie aber nur noch sehr geringe Auswirkungen auf die Lebensrealität der heutigen Kinder haben.



Um 1900 hatten Schülerinnen noch ganz andere Erziehung, gesellschaftliche Rollendefinitionen und berufliche Möglichkeiten als Schüler. Wenn es damals messbare Geschlechtsunterschiede in schulischen Variablen gab, so sind diese heute fast vollständig verschwunden (Quelle: Staatsarchiv Basel)

Klassengrösse hat ebenfalls hier und heute eine **relativ bescheidene Effektstärke** ($d=0.21$), obwohl sich in kleinen Klassen viele Aspekte des guten Unterrichts sehr viel einfacher umsetzen lassen. Jetzt wo man dies weiss, entsteht ein Interesse an der Frage, wie man mit kleinen Klassen guten Unterricht macht (s. 102). Man kann Hatties Ausführungen geradezu als Aufforderung ansehen, in diesem Bereich Forschung zu betreiben und Unterrichtsmethoden zu entwickeln. Mit etwas Glück wissen wir in einigen Jahren sehr viel besser, wie man kleine Klassen gut unterrichtet, und wenn wir das herausfinden, dann wird auch die Effektstärke grösser werden.

Ganz bedeutend ist der Effekt bei den **neuen Technologien**. Natürlich halten diese Einzug in die Schule, nur ist es eben heute so, dass wir vielfach noch nicht besonders gut wissen, wie man damit bessere Lernerfolge bei den Schülerinnen und Schülern erreicht. Das heisst: Heute messen wir relativ geringe Effektstärken, wenn wir Laptops und Tablets in die Schulzimmer bringen. Wenn wir besser wissen, wie man diese Technologien in der Schule einsetzt und wenn es mehr schulnahe Anwendungen dafür gibt, dann kann sich das ändern.



Wandtafeln gibt es seit dem Mittelalter, heute sind sie das prototypische Arbeitsgerät für Lehrpersonen. Wahrscheinlich hat es bei ihrer Einführung aber auch zunächst einmal einige Jahre gedauert, bis die Lehrpersonen damit vertraut waren und wussten, wie man damit die Schülerinnen und Schüler zu besseren Leistungen bringt. (Abbildung von Comenius, 1653)

Fazit: Effektstärken sind Momentaufnahmen



Was heisst das für Hatties Effektstärken? – Zunächst einmal, dass es Momentaufnahmen sind. **Effektstärken können sich ändern, indem Zeit verstreicht, Geld investiert wird, Technologie sich entwickelt, Menschen und Interessen sich entwickeln.** Eine tiefe Effektstärke bedeutet nicht, dass ein Faktor ganz pauschal unwirksam ist. Es bedeutet einfach, dass er hier und jetzt keine grossen Änderungen verursacht. In einem ganz anderen Schulsystem hat der Faktor vielleicht sehr entscheidende Auswirkungen, und bei uns möglicherweise in Zukunft auch.

4. Effektstärken sind Belege für wirksame Interventionen

Nebst aller Komplexität, die mit den Effektstärken verbunden sind, gibt es eine ganz einfache Interpretation. Es ist die gleiche, welche wir benutzen, um herauszufinden, **welche Medikamente wirken.**

Pharmazie ist eine sehr komplexe Wissenschaft, und wir wissen bei weitem nicht genug, um einer Substanz anzusehen, was passiert, wenn wir sie einem Menschen verabreichen. Dazu braucht es wissenschaftliche Studien, also im Grunde genommen muss man eben ausprobieren und sehen, was dabei herauskommt. Das soll nicht das massive Fachwissen schmälern, das wir haben über die biologischen und medizinischen Prozesse, was am Schluss jedoch zählt, sind die Versuchsreihen. **Medikamente kommen erst dann auf den Markt, wenn sie erfolgreich ausprobiert wurden.**

Es gibt Gründe, warum das in der Praxis nicht immer einwandfrei funktioniert – die meisten haben damit zu tun, dass **sehr viel Geld im Spiel** ist. Die englische Wikipedia hat eine Liste der grössten Betrugsfälle von Pharma-Konzernen, die Fälle selbst kann man auf der Website des amerikanischen Justizministeriums nachlesen:

http://en.wikipedia.org/wiki/List_of_largest_pharmaceutical_settlements

Ausserdem wird fast die ganze Pharma-Forschung von den Herstellern selbst finanziert, was ebenfalls sehr viele Probleme mit sich bringt. Eine Übersicht findet sich im Buch „Bad Pharma“, oder in der Zusammenfassung auf Wikipedia:

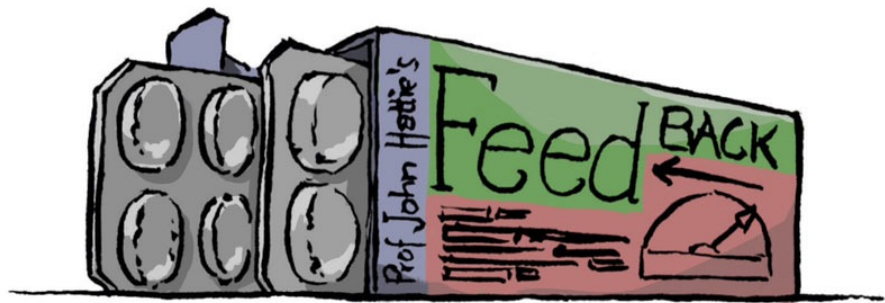
http://en.wikipedia.org/wiki/Bad_Pharma

Warum gehe ich hier auf die pharmazeutische Forschung ein? – Weil wir ihr unsere Gesundheit anvertrauen, obwohl sie viele bekannte Probleme hat und die Herstellerfirmen regelmässig dafür verurteilt werden, dass sie Patienten umbringen, um ihre Gewinne zu maximieren.

Die Schulforschung hat weitgehend die gleichen Anliegen und Forschungsmethoden, aber sie hat diese Probleme nicht. Schulforschung wird von Universitäten und Hochschulen betrieben, nicht von riesigen Konzernen, die Produkte verkaufen wollen. Überhaupt werden in der Schule etwa 50% des Geldes für Lehrpersonen ausgegeben (Bundesamt für Statistik, 2013), und die übrigen Ausgaben sind vielfach klar

vorgegeben: Hauswart, Sekretariat, Möbel, Hefte, Bücher. Und die Schulhäuser selbst, die ebenfalls viel Geld kosten.

Auch wenn es die Praxis nicht immer ist: **Die Idee hinter der pharmazeutischen Forschung ist solide und einfach:** Wir probieren aus, was funktioniert, und verwenden diejenigen Dinge, deren Wirksamkeit wir beweisen können. **Unter diesem Gesichtspunkt sind Hatties Effektstärken Wirkbelege.** Viele Dinge wurden ausprobiert und wir wissen, welche davon die Leistung von Schülerinnen und Schülern verbessert haben.



Gute Schule bedeutet, dass viele Dinge stattfinden, die wirksam sind für das Lernen.

Unter dieser Perspektive interessieren wir uns also nicht dafür, wie Schule funktioniert, sondern was in der Schule funktioniert. Konkret interessieren wir uns für Faktoren mit folgenden Eigenschaften:

- Sie sollten eine **hohe Effektstärke** haben, also eine **gut belegte Wirksamkeit**.
- Sie müssen von uns **beeinflussbar** sein. Hatties Liste enthält auch Dinge, die wir nicht beeinflussen können, wie etwa das Geburtsgewicht der Lernenden ($d=0,54$).
- Sie sollten **möglichst einfach und preiswert** sein.

Der letzte Punkt tönt eher trivial, ist aber praktisch sehr bedeutsam. Bei Hattie haben fast alle Faktoren positive Effektstärken, für sich betrachtet ist also fast jede Intervention sinnvoll. Viel schwieriger ist die Frage, welche der aufwendigen Interventionen sich rechtfertigen lassen, wenn es auch einfachere Möglichkeiten gibt, den Lernerfolg zu verbessern. Dies ist insbesondere der Fall bei **umfassenden Unterrichtsreformen** ($d=0,22$), also gross angelegten, oft tiefgreifenden Reformprojekten. Borman, Hewes, Overman und Brown (2003) haben 18 solche Reformen mit klingenden Namen wie „Roots and Wings“ oder „High Schools That Work“ untersucht und fanden Effekte zwischen $d=-0,02$ und $d=0,38$ (eine ausführlichere Liste findet sich bei Hattie auf Seite 254). Grosse Reformen sind also bestenfalls durchschnittlich wirksam.

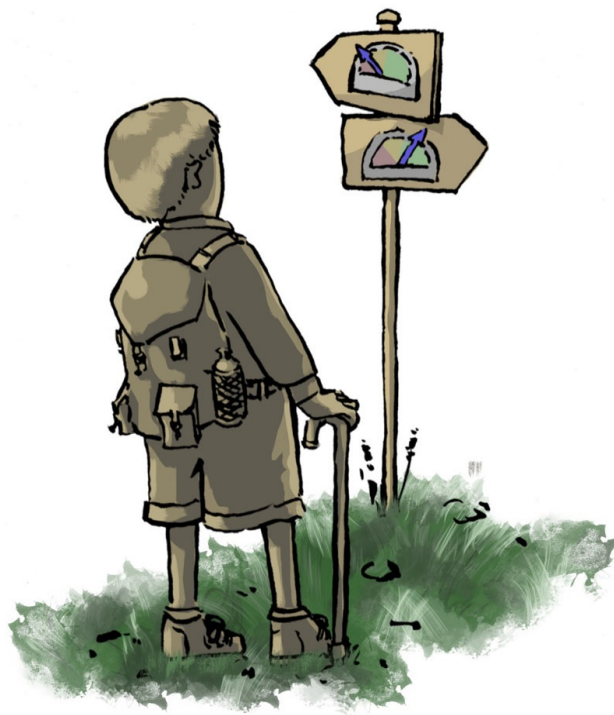
Umgekehrt gibt es eine Reihe von Faktoren, die sehr einfach umzusetzen sind und grössere Effektstärken haben, hier nur einige Beispiele:

- Definieren von **Lern- und Verhaltenszielen** („Advance Organizers“, $d=0,41$)
- Entwicklung **graphischer Darstellungen** der zu lernenden Konzepte („Concept Mapping“, $d=0,57$)

- Setzen **anspruchsvoller Ziele** ($d=0,56$)

Somit könnte man eine weitere Analogie ziehen zu Medikamenten: **Wir verfügen über eine Reihe von bewährten und preiswerten Methoden, Unterricht besser zu machen und die Lernleistung zu steigern.** Das Problem von neuen und teuren Ansätzen ist weniger, dass sie wirkungslos sind, sondern dass sie häufig nicht an die Wirksamkeit der altbekannten Methoden heranreichen. Es ist zwar verlockend, grosse Umwälzungen anzustreben, aber in der Regel sind die Erfolge geringer, als wenn man etwas einfaches und gut belegtes macht.

Diese Sichtweise von Effektstärken als Wirkbelege ist vermutlich auch der **Grund, weshalb Hattie in der Politik und bei Schulleitungen beliebt ist: Man kann daraus ableiten, welche Interventionen man durchführen soll.** Entscheidungsträgerinnen und Entscheidungsträger sind darauf angewiesen, wirksame Interventionen zur Verfügung zu haben. Hier bietet die Liste der Faktoren mit ihrem rezeptartigen Zugang grosse Vorteile. Nicht in dem Sinn, dass damit Schule umfassend und unter Würdigung aller Perspektiven und Wechselwirkungen abgebildet ist, sondern ganz einfach, weil sie hilft, Dinge zu tun, die wirken.



Effektstärken können den Weg zu sinnvollen Entscheidungen weisen.

5. Effektstärken über 0,4 zeigen „echte“ Verbesserungen

Hattie setzt eine Effektstärke über 0,4 als „**Schwellenwert**“: kleinere Effektstärken verdienen nur eingeschränkte Beachtung, grössere sind echte Belege für wirksame Effekte. Das hat folgende Gründe:

- **(Fast) alles wirkt.** In Hatties Liste gibt es nur sehr wenige Dinge, die zu einer Verschlechterung der Lernleistung führen. Somit ist es kein ausreichendes Kriterium für eine gute Intervention, dass der Effekt grösser als 0 ist – dieses Kriterium erfüllen 133 der 138 Faktoren.
- **Der Durchschnitt aller Effektstärken ist $d=0,4$.** Das heisst: Wenn ich irgend eine beliebige Intervention mache, dann kann ich im Schnitt eine Effektstärke von 0,4 erwarten.

Wie ist das zu verstehen? – Zunächst einmal ist das ein Beleg dafür, dass **Schule stabil ist** und Leistungen nicht plötzlich zusammenfallen, wenn man etwas am pädagogischen Konzept oder an den Rahmenbedingungen ändert.

Ausserdem sind **erfahrene und gut ausgebildete Lehrpersonen** am Werk, die sich Mühe geben, etwas zu verbessern, und sich bei Forschungsprogrammen meist auch besonders stark engagieren. Auch für die Lernenden ist es meist auch von Vorteil, wenn neue Dinge ausprobiert werden. Das **schafft Abwechslung und motiviert**, da sich Lehrpersonen und Forschende für ihren Lernerfolg interessieren und engagieren.

Dieser Effekt hat in der Wissenschaft einen Namen: Es ist der **Hawthorne-Effekt** (Landsberger, 1958). Er geht zurück auf Untersuchungen in den Hawthorne-Elektrowerken in den 20-er Jahren. Dort untersuchte man, welche Lichtverhältnisse optimal sind für die Angestellten (also zu höheren Produktionsleistungen führen). Nach einigen Experimenten fand man heraus, dass **jede Veränderung der Lichtverhältnisse mit einer erhöhten Produktion einhergeht** (inklusive der Rückkehr zur ursprünglichen Beleuchtung). Die Interpretation war, dass **nicht das Licht die Leistung veränderte, sondern das Wissen der Arbeitenden, dass sie Teil einer Studie waren**. Also dass man sich einerseits um sie kümmerte und andererseits auch aufzeichnete, wie viel sie leisteten.



Eine der Hawthorne-Versuchsanordnungen.

Wie können wir nun unterscheiden, ob ein Studienergebnis das Resultat des Hawthorne-Effekts ist, oder ob eine echte Verbesserung dahinter steckt? – Hier hilft etwas, das Meta-Analysen fast immer machen: Sie untersuchen nämlich neben dem Ausmass der Wirkung auch die **methodische Qualität der Studien**. Dabei geht es nicht darum, ob die Wissenschaftler gut oder schlecht gearbeitet haben, sondern wie viel Aufwand sie betrieben haben, um solche Dinge wie den Hawthorne-Effekt zu kontrollieren. Beispielsweise kann man eine **Kontrollgruppe** verwenden, welche gleich viel Aufmerksamkeit bekommt wie die Untersuchungsgruppe, aber mit der man nicht die untersuchte Intervention durchführt. Man kann auch die Studie **längerfristig** anlegen und **Nachmessungen** durchführen, um sicherzustellen, dass die positiven Effekte anhalten.

In der Regel steht den Wissenschaftlern aber nicht das Geld oder der organisatorische Rahmen zur Verfügung, um gross angelegte, kontrollierte, lange dauernde Studien zu machen. Dafür sind diese einfach **viel zu teuer**, und sie lohnen sich auch nur, wenn man schon recht viel zusammengetragen hat über ein Thema und die Erkenntnisse besser absichern möchte.



Im einfachsten Fall lassen sich Erkenntnisse aus bereits erhobenen Daten ableiten, wie etwa schulübergreifenden Vergleichstests. Daraus kann man Effektstärken berechnen, ohne vom Stuhl aufstehen zu müssen.

Bei einem echten Effekt finden methodisch bessere Studien gleiche oder grössere Effektstärken als die methodisch weniger aufwendigen. Wenn die Effekte gross sind bei den einfachen Studien und klein bei den aufwendigen Studien, dann handelt es sich eher um einen Placebo- oder Hawthorne-Effekt.

Was heisst das konkret bezogen auf den Schwellenwert von $d=0,4$? – Zunächst einmal kann man etwas pauschal davon ausgehen, dass **Effektstärken über $d=0,4$ sehr wahrscheinlich auf echte Effekte zurückzuführen sind**. In einem zweiten Schritt kann man – besonders bei Effektstärken unter 0,4 – die methodische Qualität hinzuziehen. Wenn aufwendige Studien die Effekte bestätigen oder sogar grössere Effekte finden, dann spricht das dafür, dass die Effekte wirklich existieren. Wenn die Effekte sich nur in den weniger sorgfältig abgesicherten Studien zeigen, dann kann man sich eher nicht auf sie verlassen.

Teil 3

III. Fünf Effekte im Detail

Wir wollen in diesem Teil einige ausgewählte Ergebnisse etwas genauer anschauen. Ich möchte an dieser Stelle noch einmal darauf hinweisen, dass die Faktoren und Ergebnisse sehr **zugänglich** sind, und dass man sich sehr leicht vertieft informieren kann:

- In „**Lernen sichtbar machen**“ diskutiert Hattie jeden Faktor auf ein oder zwei Seiten, gelegentlich auch etwas ausführlicher. Dabei weist er ausführlich auf die zugrundeliegenden Arbeiten und weiterführende Literatur hin.
- In „**Lernen sichtbar machen für Lehrpersonen**“ fasst er die Ergebnisse dahingehend zusammen, was sie für Lehrpersonen und Unterricht bedeuten.
- Die **zugrundeliegenden Arbeiten** sind ebenfalls leicht verfügbar. Viele sind auf dem Internet frei zugänglich, die anderen sollten aus einem Hochschulnetzwerk oder von einem Bibliothekscomputer aus elektronisch zu beschaffen sein.

In diesem Teil wollen wir einige ausgewählte Faktoren unter spezifischen Gesichtspunkten betrachten und an die hiesigen Umstände und gesellschaftlichen Diskussionen anschliessen. Zusätzlich zu allen anderen Erwägungen sind zwei Fragen zentral:

- **Wie die Faktoren aus Sicht der Lernenden zu verstehen sind.** Das ist für Hattie wichtig und kommt in pädagogischen Diskussionen meist viel zu kurz. Gemeint ist, dass wir nicht darüber reden, was unserem Bild von Kindheit, idealer Lehrperson oder guter Schule entspricht. Wir sollten versuchen, die Faktoren durch die Augen der Schülerinnen und Schüler zu sehen.
- **Was die Faktoren für Lehrpersonen und Schule bedeuten.** Müsste man angesichts der Effekte Schule anders machen als heute? Wenn wir Hatties Ergebnisse beiziehen, dann reden wir jeweils über **das Kerngeschäft: Schule soll den Kindern etwas beibringen.** Das effizient umzusetzen ist schwierig und verdient ebenfalls mehr Beachtung, als es üblicherweise bekommt.

1. Gute Lehrpersonen, guter Unterricht (Effekte um $d=0,70$)

Damit Schülerinnen und Schüler viel lernen, braucht es gute Lehrpersonen, welche guten Unterricht machen. Bei Hattie sind dies ganz zentral folgende Faktoren:

- Klarheit der Lehrperson ($d=0.75$)
- Lehrer-Schüler-Beziehung ($d=0.72$)
- Rhythmisierendes Unterrichten ($d=0.71$)
- Lehrerfort- und Weiterbildung ($d=0.62$)

- Klassenführung (d=0.52)

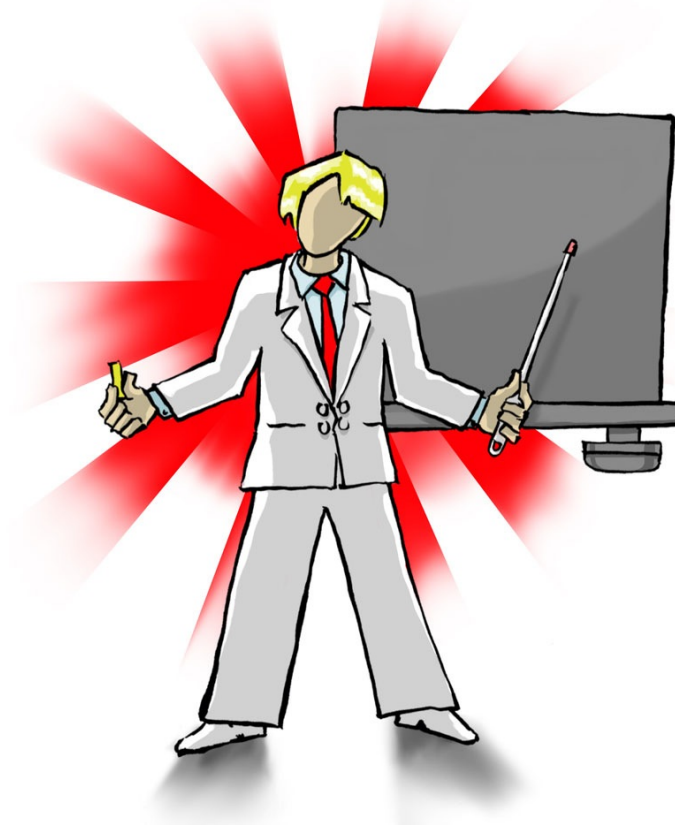
Nach Hatties Ergebnissen sind diese Faktoren grösser und konsistenter als die meisten anderen Dinge, die mit Lernerfolg zu tun haben. Das ist ein zentrales Ergebnis seiner Studie: **Für Lernerfolg braucht es gute Lehrpersonen, die guten Unterricht machen.** Stütz- und Auffangmassnahmen sind hingegen kaum effizient: Sommerschulen (d=0.23), finanzielle Ausstattung (d=0.23), Home-School-Programme (d=0.16), Nachmittags- und Sommerkurse (d=0.09). Eine Klasse zu wiederholen führt sogar zu eher schlechteren Leistungen (d=-0.16). Das heisst: Wenn der Unterricht nicht funktioniert, dann gibt es nur sehr wenig, was man ausserhalb des Unterrichts gegensteuern kann.

Wie sieht die Gesellschaft guten Unterricht?

Die gesellschaftliche Diskussion handelt nur selten davon, wie man Unterricht effizient macht. Viel **beliebtere Themen** sind zum Beispiel:

- Das Engagement der Lehrperson
- Erziehung
- Migration
- Geplante oder durchgeführte Änderungen an Lehrplan und Schulstrukturen
- Die vielen Ferien
- Die schwierigen Eltern
- Die schwierigen Kinder
- Regeln einführen / Grenzen setzen
- Welche Lehrpersonen und Fächer man hasste
- Phantasie
- Kreativität
- Strafen
- Kind sein dürfen

In den Medien sind gute Lehrpersonen jeweils charismatische Figuren, welche die Kinder packen, indem sie jedem Zwang abschwören und die Kinder nur durch reine Faszination zu fesseln vermögen.



Wie sich Hollywood eine gute Lehrperson vorstellt

Das ist nicht von Grund auf falsch, aber Zweck der Schule ist ja nicht, dass die Kinder fasziniert sind vom Stoff, sondern dass sie etwas lernen. Schlecht wäre folgende Definition von Schule:

Schools are places where kids go to watch teachers work.

Schulen sind Orte, wo Kinder hingehen, um Lehrpersonen beim Arbeiten zuzuschauen

– Quelle unbekannt, zitiert nach Dylan Wiliam (2011)

Oder mit etwas anderen Worten:

If the teachers are going home more tired than the kids at the end of the day, the wrong people are doing the work.

Wenn die Lehrpersonen abends erschöpfter nach Hause gehen als die Kinder, dann machen die falschen Leute die Arbeit.

– Dylan Wiliam (2011)

Film und Fernsehen präsentieren uns in der Regel **normale Menschen, die in besonderen Situationen über sich hinauswachsen**. Das erleichtert uns die Identifikation mit ihnen und das Verständnis der Handlung, da wir die Denkweise und Handlungsmuster der Protagonisten leicht nachvollziehen können. Etwas anders gesagt:

Filme handeln davon, wie jemand erfolgreich ist, der ungefähr die gleichen Voraussetzungen mitbringt wie das Publikum.

Bei Filmen über Lehrpersonen äussert sich das oft in Mustern, welche für professionelle Lehrpersonen beleidigend sind:

- Abgesehen vom Protagonisten sind alle Lehrpersonen unmotiviert, unfähig und feindselig.
- Die Schülerinnen und Schüler leisten praktisch unbegrenzten Einsatz, wenn ihnen die Lehrperson sympathisch ist oder kurz erklärt, weshalb das Fach spannend ist.
- Unkonventionelle Methoden führen zu sehr viel grösseren Lernerfolgen. Je weniger das, was die Schülerinnen und Schüler machen, an eine Lernsituation erinnert, desto mehr wird dabei gelernt.

Was hat das mit Hattie zu tun? – Wir haben mit „Lernen sichtbar machen“ eine umfassende Darstellung davon, was wirklich wirkt in den Schulen. Wichtig sind dabei folgende Einsichten:

- **Viele belegte Effekte tauchen in der öffentlichen Diskussion eher am Rand auf.** Es gibt kaum Diskussion über Micro-Teaching ($d=0,88$), Akzeleration ($d=0,88$), Feedback ($d=0,73$); die sehr wichtige positive Lehrer-Schüler-Beziehung ($d=0,72$) wird allenfalls als „Kuschelpädagogik“ abgewertet.
- **Viele der populären Forderungen an gute Schule sind bei weitem nicht die Wundermittel, als die sie dargestellt werden:** Kreativität ($d=0,35$), finanzielle Ausstattung der Schule ($d=0,23$), umfassende Unterrichtsreformen ($d=0,22$), unternehmerisch organisierte Schulen („Charter Schools“, $d=0,20$).
- **Lehrperson sein erfordert spezielle Fähigkeiten:** Professionelles Auftreten (Micro-Teaching, $d=0,88$), professioneller Umgang mit Störungen und Schülerverhalten ($d=0,80$ für Beeinflussen von Schülerverhalten, $d=0,34$ für das Reduzieren von Unterrichtsstörungen), Vermitteln von metakognitiven Strategien ($d=0,69$) und Lerntechniken ($d=0,59$), Wissen über Lehrstrategien ($d=0,60$).

Guter Unterricht hilft Schülerinnen und Schülern, Stoff zu lernen und sich Fähigkeiten anzueignen. Weshalb haben wir aber eine so vage Vorstellung von gutem Unterricht, wo wir doch alle jahrelang zur Schule gegangen sind? Das hat damit zu tun, wie unser Erinnerungsvermögen funktioniert. Das Lernen von Fähigkeiten ist etwas ganz anderes als das Erinnern an vergangene Zeiten und Begebenheiten. Wir wollen deshalb diese verschiedenen Arten des Gedächtnisses und Lernens gesondert betrachten.

Der Mythos vom schnellen Lernen

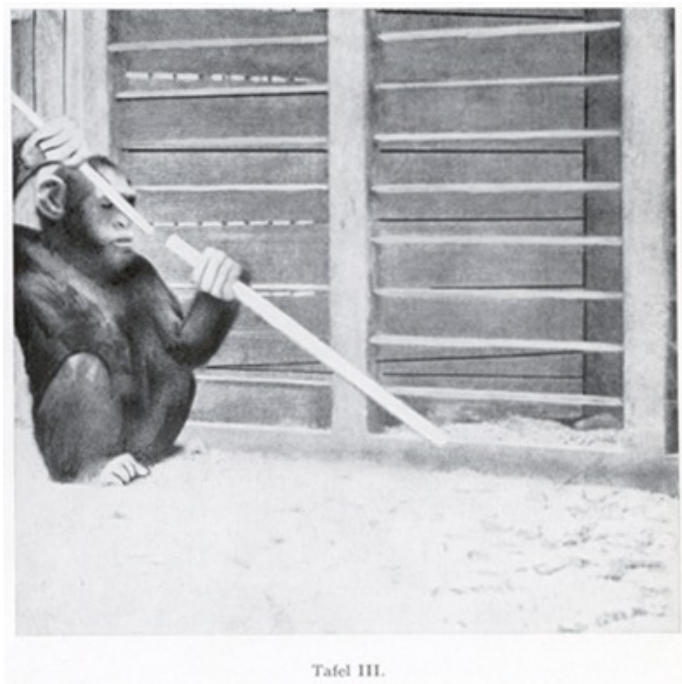
Zunächst wollen wir einen Mythos aufklären: Die Idee, dass man die mühsamen und arbeitsamen Aspekte des Lernens ablösen kann durch Einsichtsprozesse und plötzliche Erkenntnis, wenn man den Unterricht nur richtig aufbaut.

Damit das klar ist: **Nichts gegen spektakulären Unterricht und dagegen, dass Lernende plötzliche, umwälzende Einsichten haben. Nur können diese Dinge**

nicht Übung und Arbeit ersetzen. Auch wenn wir uns das wünschen würden – aus der Sicht der Lernenden wäre Schule sehr viel einfacher, wenn wir uns einfach zurücklehnen und mitreissen lassen könnten, und unser Lernen dann in plötzlichen, grossen Schritten vorwärts gehen würde.

Lernen durch plötzliche Einsicht oder spontane kreative Einfälle ist sehr gut untersucht. Zunächst einmal bei Affen (Köhler, 1917, 1925). Wolfgang Köhler schuf verschiedene Versuchsanordnungen, in denen Affen Hilfsmittel einsetzen mussten, um Bananen zu erreichen. Eine Beschreibung der einzelnen Anordnungen findet sich hier:

<https://www.uni-due.de/edit/lp/kognitiv/koehler.htm>



Ein Affe setzt zwei Rohre zusammen. Solches Verhalten zeigt sich oft im Rahmen von plötzlicher Einsicht und führt zu schnellem, dauerhaftem Lernen.

Köhler stellte fest, dass es zwei verschiedene Arten gab, wie Affen an die Bananen kamen:

- Die Tiere versuchten zunächst, die Bananen durch einfache Aktionen wie Springen zu erreichen, also durch **Versuch-und-Irrtum**.
- Wenn das nicht funktionierte, dann setzten sie sich oft zunächst einmal hin oder beschäftigten sich mit anderen Dingen. Dann setzten sie plötzlich und zielgerichtet Werkzeuge wie Kisten oder Rohre ein, um die Banane zu erreichen. **Sie scheinen aus Einsicht zu handeln.**

Die **Einsicht führte bei den Affen zu Lernerfolg**: Sie konnten sich auch später noch sehr gut an die Lösung erinnern und sie wiederholen, und schienen auch Spass daran zu haben.

Das ist uns auch als Menschen vertraut. So soll Archimedes, der wahrscheinlich grösste Wissenschaftler der Antike, nach einer plötzlichen Einsicht in seiner Badewanne nackt und „**Heureka**“ (ich habe es gefunden) schreiend durch die Strassen gerannt sein.



Warum können also nicht alle Lernprozesse so spektakulär und erfreulich ablaufen? Warum stellen wir Schülerinnen und Schüler nicht einfach vor Probleme und lassen sie dann Einsichten und die damit verbundene Freude erleben, und weshalb können Lehrpersonen nicht solche Einsichten fördern, anstatt dass sie Übungen machen und Texte schreiben lassen?

Nun, **man hat es ausprobiert. Der Ansatz heisst „Problembasiertes Lernen“** und besteht darin, dass Lernende in Kleingruppen mit echten Problemen konfrontiert werden, die sie lösen sollen. Dabei sollen sie sich selbstgesteuert die notwendigen Werkzeuge aneignen, während die Lehrperson als Tutor moderieren und anleiten kann.

Die Erfolge des problembasierten Lernen sind allerdings eher bescheiden: $d=0,15$. Und was beinahe noch schlimmer ist: Problembasiertes Lernen umfasst eine Reihe von Ansätzen, welche je für sich sehr viel höhere Effekte aufweisen:

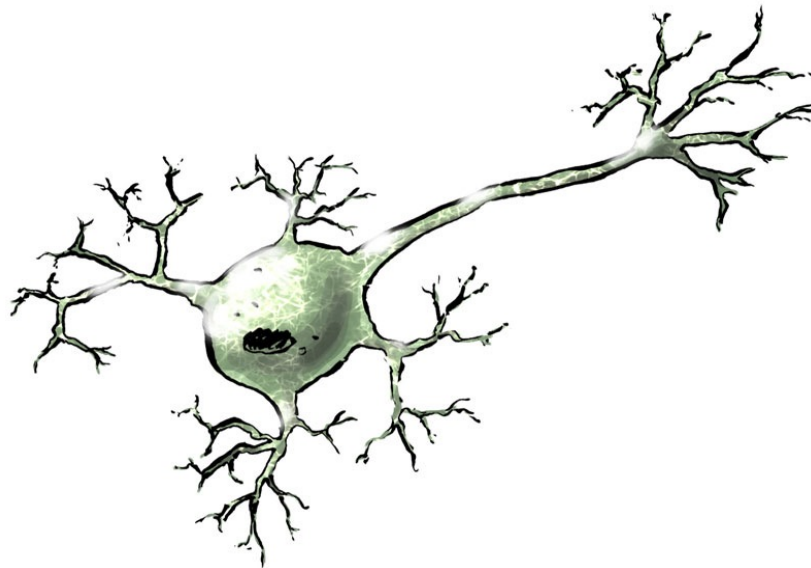
- Training von Problemlösefähigkeiten ($d=0.61$)
- Kooperatives Lernen ($d=0.41$)
- Forschendes Lernen ($d=0.31$)

Hattie kommt ausserdem zum Schluss, dass problembasiertes Lernen eher für die Anwendung von Wissen geeignet ist, und weniger für das Aneignen von neuem Wissen (s. 250). **Somit scheint die plötzliche Einsicht – bei aller Freude und Eindrücklichkeit, die sie hervorruft – nicht zu taugen als primäres Mittel, um Stoff zu vermitteln.**

Die Realität des langsamen Lernens

Das meiste Lernen geschieht allmählich, indem wir Dinge immer wieder tun, bis sie haften bleiben. Das hat damit zu tun, wie unser Gehirn funktioniert.

Wenn Dinge gelernt werden, dann **muss das Gehirn neue Verbindungen zwischen den Nervenzellen schaffen**, wie Donald Hebb bereits 1949 herausgefunden hat (Hebb, 1949). Über die Zeit entstehen komplexe Strukturen, die sich heute auch in Gehirncans nachweisen lassen. Wenn man sich komplexe Fähigkeiten aneignet, dann führt dies zu nachweisbaren Veränderungen im Aufbau des Gehirns (Jäncke, 2009).



Das Gehirn enthält ca. 20 Milliarden Nervenzellen mit total 100 Billionen Verbindungen und 5.8 Millionen Kilometern Nervenleitbahnen. Lernen bedeutet, diese Bahnen behutsam zu komplexen, aufeinander abgestimmten Strukturen zu formen.

Damit das Gehirn solche Strukturen bildet, braucht es vor allem viel Arbeit, also jede Menge Auseinandersetzung mit den eigentlichen Inhalten. Auf Ebene der Nervenzellen ist die Regel dabei (Shatz, 2007):

Fire together, wire together.

Zellen, welche gemeinsam aktiv sind, vernetzen sich.

Das ist ein relativ **behutsamer elektrochemischer Vorgang** und benötigt Zeit und Wiederholungen. Damit sich Strukturen bilden, muss das Gehirn aktiv sein und sich mit der Sache beschäftigen, die es lernen will. Wir müssen uns also mit den Lerninhalten auseinandersetzen.

Der Grad der Auseinandersetzung lässt sich nur unzureichend messen, wir können ja nicht feststellen, ob das Gehirn gerade am Strukturen aufbauen ist. Für die reine Lernzeit findet Hattie einen mittleren Effekt ($d=0.38$), hohe Effekte haben Strategien, welche das eigene Lernen optimieren ($d=0.69$ für Meta-Kognitive Strategien). Beim Lesen finden sich ebenfalls hohe Effektstärken für Ansätze, welche intensive Auseinandersetzung zur Folge

haben: $d=0.67$ für Wortschatzprogramme, $d=0.67$ für wiederholendes Lesen, $d=0.58$ für verstärktes Engagement und Verständnis beim Lesen.

Dem fortgesetzten Üben stellt sich allerdings eine weitere Eigenschaft des Gehirns entgegen: **Je mehr wir über ein Gebiet wissen, desto zurückhaltender wird unser Gehirn im Aufnehmen neuer Informationen.** Der Fachbegriff dafür ist „**proaktive Inferenz**“ (Baddeley, 1990, S.43ff), im Alltag ist das bekannt als **Unlust** bis hin zur Blockade. Es ist schwierig, sich fortgesetzt für die gleiche Sache zu interessieren.

Man kann die proaktive Interferenz senken, wenn man das **Lernen auf längere Zeiträume verteilt** (zum Beispiel: Gunter, Berry & Clifford, 1981). Man sollte sich also nicht fasziniert auf die Arbeit stürzen, da man dann unter Umständen recht schnell an dem Punkt ist, wo bei allem Elan nicht mehr viel ins System hereingeht. In der Tat ist das Verteilen von Lernen auf ausgedehnte Zeiträume **einer der effektivsten Faktoren bei Hattie ($d=0.71$, damit 12. Gesamtrang)**. Auch dieser Faktor wird oft falsch eingeschätzt – man ist versucht, in intensiven Trainingseinheiten grosse Sprünge machen zu wollen, aber in der Tat ist es viel effizienter, Lernen auf längere Zeiträume zu verteilen.

Dies sind **zwei gewichtige Bedingungen des Lernens**: Um Neues zu lernen muss ich mich **lange und intensiv damit auseinandersetzen**, damit mein Gehirn komplexe Strukturen von Nervenzellen aufbauen kann. Und je länger ich mich fortgesetzt mit einer Sache auseinandersetze, desto weniger effizient werde ich im Lernen, ich muss das Lernen also möglichst **auf einen längeren Zeitraum verteilen**, damit es effizient bleibt.

Das Schaffen von überdauernden Erinnerungen

Was Wissen und Fähigkeiten angeht, lernt unser Gehirn also eher langsam, unterbrochen allenfalls von kurzen Sprüngen und plötzlichen Einsichten. Dies ist, was in der Psychologie das **semantische Wissen** ausmacht, das Wissen von der Welt und den Dingen (Tulving, 1983).

Zusätzlich gibt es das so genannte **episodische Wissen** (ebenfalls Tulving, 1983), also die **Erinnerung an vergangene Ereignisse**. Das ist sozusagen den Film, der in unserem Kopf mitläuft und alles – zumindest bruchstückhaft – abspeichert, was uns den ganzen Tag passiert.

Auch hier werden herausragende Ereignisse bevorzugt abgespeichert. Dazu gehören zum Beispiel spektakuläre Exkursionen, wenn Unterrichtsinhalte als Rap vorgetragen werden, oder auch nur, wenn der Dozent an der Uni über ein Kabel stolpert und der Länge nach hinfällt (Tulving, 1983).



Endel Tulving, Gedächtnispsychologe. Er sagte einmal, dass er kein Problem damit hat, wenn die Studierenden die Inhalte seiner Veranstaltung vergessen. Allerdings wurmt es ihn, dass sie sich auch Jahre später noch präzise daran erinnern können, wie er einmal über ein Hellraumprojektor-Kabel gestolpert ist.

Wenn wir uns erinnern an unsere eigene Schulzeit, dann stehen diese **herausragenden Ereignisse im Vordergrund**, ganz egal, ob sie etwas mit Lernerfolgen zu tun hatten oder nicht. Ich kann mich noch an Schulausflüge erinnern, einige Chemie-Versuche (besonders die schief gegangenen), die Wertschätzung meines Deutschlehrers, die Geringschätzung meines Mathe-Lehrers und dass der Englischlehrer ein Mal unrasiert in die Stunde kam.



Ein Studienkollege von mir ging mit der Klasse Elefanten waschen im Zoo. Das ist natürlich ein absolutes Highlight, das den Kindern lange in Erinnerung bleiben wird und ihre Vorstellung davon prägt, was guter Unterricht ist. Zusätzlich muss aber noch viel anderer guter Unterricht stattfinden, damit die Kinder zum Beispiel Brüche auflösen und Texte verfassen lernen.

Dass ich heute Hochschuldozent bin, hat sicher viel mehr mit den Dingen zu tun, an die ich mich nicht mehr im Einzelnen erinnere, also mit den hunderten und tausenden Stunden Arbeit, die ich in das Erlernen meiner unterschiedlichen Fähigkeiten investiert habe. Kurz: Ich möchte sagen, dass **gute und schlechte Schule wahrscheinlich nicht das ist, was wir an positiven und negativen Erinnerungen mit uns herumtragen.**

Wenn wir darüber nachdenken, was gute Schule ist, dann sollten wir vielmehr die **Perspektive der Schülerinnen und Schüler** einnehmen und uns **überlegen, was dazu beiträgt, ihnen neue Dinge beizubringen und bestehende Fertigkeiten auszubauen.** Also wie wir daran arbeiten, dass ihre Gehirne hochkomplexe Fähigkeiten erlernen, mit denen sie später einmal ein erfolgreiches Leben führen werden. Was sie für erstrebenswert halten im Leben ist natürlich ihre Sache, aber Schule soll dazu beitragen, dass sie diese Dinge auch erreichen können.

2. Klassengrösse (d=0.21)

Klassengrösse ist ein gutes **Beispiel dafür, wo die Forschung den Erwartungen zuwider läuft**. Eigentlich müssten kleine Klassen zu besseren Lernleistungen führen, wie der Tagesanzeiger formuliert:

„Kleine Klassen haben einen positiven Einfluss auf Leistung und Wohlbefinden der Schüler, weil sich die Lehrer intensiver um die einzelnen Kinder kümmern können, weil sich die Kinder aktiver beteiligen und weil es im Unterricht weniger Störungen gibt. In kleinen Klassen erbringen die Kinder bessere Leistungen im Lesen und in der Mathematik. Die Kluft zwischen schwachen und starken Schülern ist kleiner. Es gibt weniger Kinder, die sitzen bleiben. Die Klassengrösse sollte idealerweise deutlich unter 20 Kindern liegen.“

Tages Anzeiger 23.08.2013 – zitiert ist die gekürzte Version, welche von den Befürwortern der Klassengrösse-Initiative publiziert wird, der volle Artikel enthält auch Argumente, welche gegen die Anliegen der Initiative sprechen

Klassengrösse ist ein politisch spannendes Thema, da es beispielsweise in Zürich eine Volksinitiative gibt, Klassen auf maximal 20 Lernende zu beschränken (derzeit: 25, ausnahmsweise bis 28). Das ist laut Gegnern der Initiative mit zusätzlichen Kosten von 120 Millionen Franken verbunden.

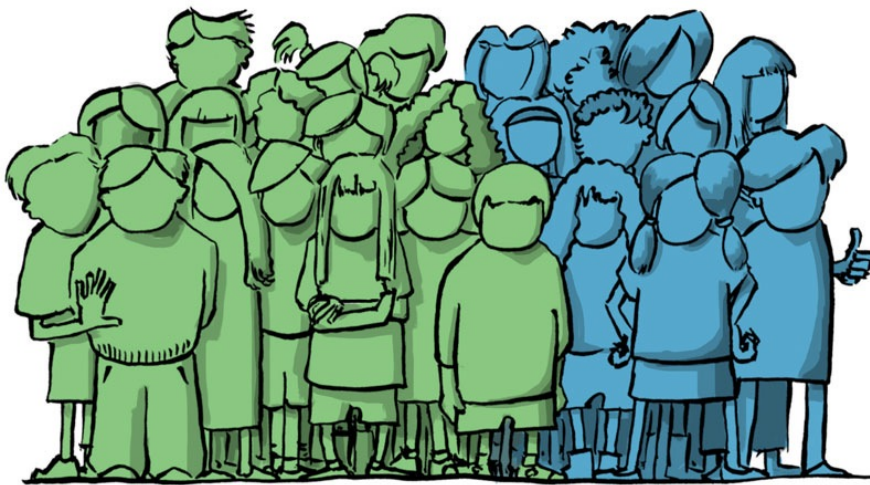
Aus **Management-Sicht** ist die Klassengrösse ein sehr verlockendes Instrument, um Schwankungen aufzufangen. **Es ist einfacher, ein paar Lernende mehr in einer Klasse aufzunehmen, als eine neue Klasse zu schaffen**. Die zusätzliche Klasse benötigt Lehrpersonen, die man unter Umständen neu anstellen muss, ausserdem Zimmer, Ausstattung und Koordinationsaufwand. Es kann auch schwierig sein, für eine kurzfristige Ausweitung (etwa: einmalig drei statt zwei Klassen in einer Stufe) eine gute Lehrperson zu finden.

Hattie selbst ist der Auffassung, dass viele der Faktoren, die wissenschaftlich belegt sind, in kleinen Klassen eigentlich besser funktionieren (s. 104). Das heisst: Kleine Klassen müssten eigentlich zu besseren Leistungen führen, als sie es tun.

Mythos 1: Kleine Klassen steigern den Lernerfolg

Hattie findet für Klassengrösse eine Effektstärke von $d=0.21$, also einen eher geringen Effekt. Schülerinnen und Schüler lernen in kleinen Klassen geringfügig besser als in grossen. Der Effekt stützt sich ab auf 3 Meta-Analysen mit total 96 Studien an 550'000 Lernenden. Die Effekte schwanken nur recht wenig und bewegen sich üblicherweise zwischen 0.1 und 0.2 (s. 103).

Entscheidend ist natürlich, was die von Hattie zitierten Studien unter einer kleineren Klassengrösse verstehen. Die Effektstärke bezieht sich auf eine **Reduktion von 25 auf 15 Lernende**. Das ist eine sehr deutliche Reduktion. Für kleinere Reduktionen dürften die Effektstärken noch weit geringer sein.



Reduktion von 25 lernenden (ganze Gruppe) auf 15 Lernende (linke, grüne Gruppe)

Brahm (2006) fand in einer Untersuchung an 89 Essener Grundschulklassen, dass die Schülerinnen und Schüler **in grösseren Klassen häufiger angaben, sich aktiv am Unterricht zu beteiligen und seltener angaben, mit den Gedanken abzuschweifen**.

In amerikanischen und englischen Studien zeigten sich positive Effekte vor allem in den ersten Schuljahren und für leistungsschwache wie auch sozial benachteiligte Lernende (Blatchford, 2003; Finn, 1998). In der englischen Studie glichen sich die Leistungen im Verlauf der weiteren Schulkarriere wieder an, in der amerikanischen blieb ein gewisser Vorsprung für die Lernenden in kleinen Klassen – allerdings nur, wenn sie auch weiterhin in kleinen Klassen unterrichtet wurden.

Dylan Wiliam (2001) sieht einen Grund in der mässigen bis fehlenden Wirkung darin, dass die Massnahmen oft die kleinen Klassen zum Ziel hatten, in der Annahme, die höhere Lernleistung stelle sich dann automatisch ein. Ihm zufolge hat die Klassengrösse allenfalls dann eine Wirkung, wenn es Vorstellungen gibt, welche Ziele man mit der geringen Grösse erreichen will und wie man die kleinere Klassengrösse einsetzt, um besseren Unterricht zu machen.

Mythos 2: Kleine Klassen steigern das Wohlbefinden

Hier gehen wir etwas über Hattie hinaus. Hattie interessiert sich ja „nur“ für den Lernerfolg, also wie man Schülerinnen und Schülern effizient etwas beibringt. Natürlich gibt es noch andere wichtige Dinge in der Schule, etwa das Wohlbefinden der Beteiligten.

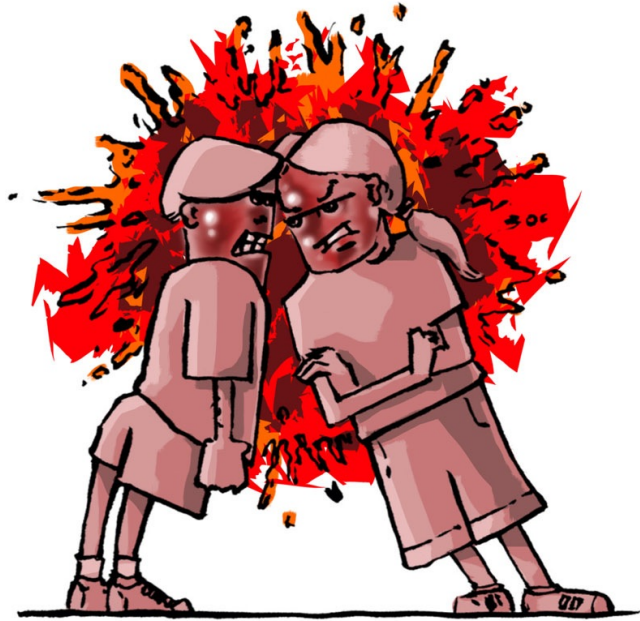


Eine Beteiligte (die Lehrperson)

Lehrpersonen sind im Allgemeinen der Ansicht, dass grosse Klassen für sie eine Belastung darstellen. In einer Umfrage des Zürcher Lehrerinnen- und Lehrerverbandes an 954 Lehrpersonen erachteten 64,7% die Klassengrösse als „grosses Problem“, 30,2% als „mässiges Problem“. Nur 5% der befragten Lehrpersonen sahen kein Problem in der Klassengrösse (ZLV, 2013). Damit ist die Klassengrösse subjektiv das gewichtigste Problem für Lehrpersonen.

Objektiv lassen sich diese Befunde nicht bestätigen. Brahm (2006, s. 19) findet bei 89 Grundschulklassen, dass die **Schüler in kleinen und grossen Klassen gleich häufig ungestört arbeiten können, der Lehrperson folgen oder nicht folgen oder den Anfang der Stunde durch Unruhe herauszögern**. Helmke und Weinert (1997, S. 246f) finden im Rahmen der Münchner Scholastik-Studie, dass **in grösseren Klassen effizientere Klassenführung, höhere Strukturierung und aktivere Unterstützung und Kontrolle durch die Lehrpersonen** stattfinden.

Verschiedene Studien haben gezeigt, dass das **soziale Klima in kleinen Klassen eher schlechter ist als in grossen**. Antisoziales und ausschliessendes Verhalten wird häufiger festgestellt (Kühn 1986; Petillion 1985; von Saldern 1985 und 1993b, Blatchford, 2003, S. 89). Petillion (1985) führt das darauf zurück, dass in kleineren Klassen sowohl ein grösserer Zusammenhalt der Gruppe als auch eine verminderte Wahlmöglichkeit der Lernenden herrschen, so dass es häufiger zu Ablehnung kommt.



In kleinen Klassen gibt es mehr Probleme zwischen den Schülerinnen und Schülern.

Die Realität: Kleine Klassen haben viele Probleme

Kleine Klassen sind also nicht besonders effektiv, um Schülerinnen und Schüler zu mehr Lernleistung zu bewegen, sie führen sie zu einem schlechteren Klima und sind teuer.

Hinzu kommen **zahlreiche praktischen Probleme**, wenn es an die konkrete Umsetzung geht:

- Es braucht **mehr Lehrpersonen** – und es ist jetzt schon sehr **schwierig, ausreichend qualifizierte und fähige Lehrpersonen zu bekommen**.
- Die Schulleitungen werden versuchen, die bestehenden und bewährten Lehrpersonen dazu zu bewegen, **zusätzliche Stunden** zu übernehmen.
- Wenn Lehrpersonen Parallelklassen unterrichten, dann müssen die Prüfungen jeweils so konstruiert und durchgeführt werden, dass sich die Klassen nicht in der Zwischenzeit über die Fragen austauschen können. Dadurch entsteht **Mehraufwand**, auch wenn pro Klasse weniger Korrekturaufwand anfällt.
- Es braucht **mehr Räume**, was dazu führen kann, dass der Unterricht in Containern stattfindet oder dass Spezialzimmer aufgehoben und als reguläre Unterrichtsräume verwendet werden.
- Mehr Klassen mit schlechterem sozialem Klima bedeutet, dass es **mehr Probleme in den Pausen und bei Schulaktivitäten** gibt und dass die Lehrpersonen häufiger mit sozialen Problemen in der Klasse konfrontiert sind.

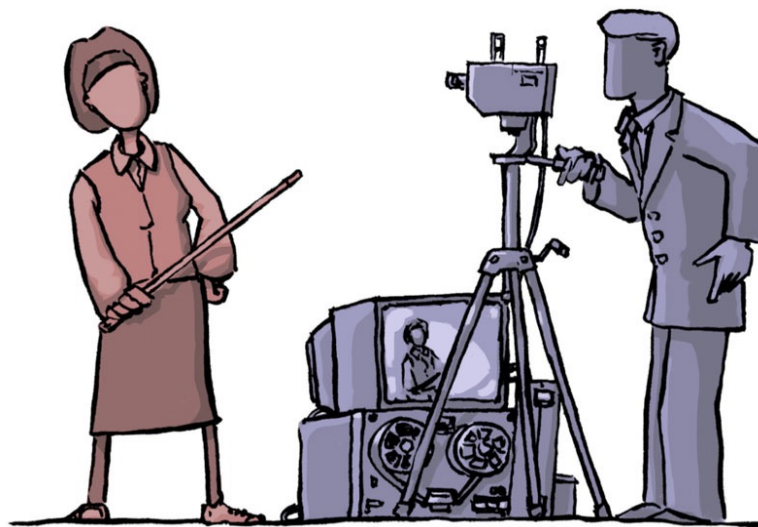
Hier geht es nicht um Schwarzmalerei. Ein zentrales Ergebnis von Hattie ist ja gerade, dass Schule sehr widerstandsfähig ist und dass es beinahe ausgeschlossen ist, dass gut gemeinte Massnahmen zu drastischen Verschlechterungen führen. Andererseits ist es **oft so, dass plausibel scheinende Massnahmen nicht die erhofften Wirkungen zeigen, weil die Schulrealität komplexer ist, als man sich das vorgestellt hatte**.

Hattie weicht mit seinen Effektstärken dieser Komplexität elegant aus. Das Ziel ist bei ihm nicht, die wechselseitigen Abhängigkeiten des Systems Schule zu begreifen. Bei der Klassengröße ist es aber so, dass die unerwartet geringe Effektstärke darauf hinweist, dass die Sachlage komplexer ist als man zunächst vermuten könnte, wodurch die offensichtlichen Vorteile der kleinen Klassen in der Realität nicht zum tragen kommen und durch negative Effekte überlagert werden.

3. Video-Selbstbeobachtung (Microteaching, $d=0.88$)

Microteaching ist ein etwas merkwürdiges Wort für eine einfache Sache, die bei näherem Hinsehen wieder sehr komplex wird. Beim Microteaching geht es darum, dass **Lehrpersonen ihr Handeln auf Video aufnehmen und dann analysieren.**

Microteaching-Programme sind üblicherweise recht umfangreich. Jo Kramis (1988) beispielsweise führte mit den Studierenden 10-12 Sitzungen à 3 Stunden durch, also rund 30 Stunden aktives Training. 30 Stunden sind viel für ein Trainingsprogramm, aber sehr wenig verglichen mit der gesamten Lehrpersonenausbildung, welche einige tausend Stunden umfasst.



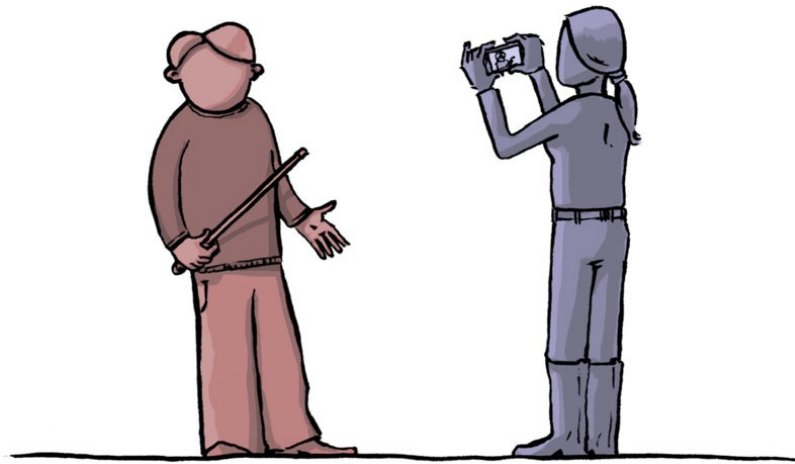
Microteaching-Programme stammen aus Zeiten, als Videoaufnahmen noch sehr aufwendig waren.

Die Videoaufnahmen geschehen je nach Programm **in einer realen Umgebung mit echten Kindern oder in simulierten Umgebungen mit Studierenden.** Microteaching wurde in den 60-er Jahren entwickelt, existiert seither in vielen Formen und wird unter verschiedenen Aspekten unterrichtet. Die Programme passen sich auch an die technischen und organisatorischen Rahmenbedingungen an:

- **Zu Beginn waren Aufnahmegeräte sehr teuer,** aber auch revolutionär und **mit grossen didaktischen Hoffnungen verbunden.** Heute trägt fast jede Person die

Technik für Videoaufnahmen mit sich herum, dafür ist es schwerer, Begeisterung dafür zu wecken.

- Der grosse technische Aufwand machte es **früher wahrscheinlich eher leichter, Programme mit echten Kindern zu organisieren**. Die teuren Aufnahme-Ausrüstungen lohnten sich nur, wenn sie optimal eingesetzt wurden.
- Die **grosse Menge an schlecht produzierten Youtube-Videos** und im Gegensatz dazu **sehr hochstehend inszenierten Mediendarstellungen in Film und Fernsehen erschweren das unvoreingenommene Anschauen der eigenen Performance auf Video**. Auch gibt es heute grosse Bedenken über die Privatsphäre, welche sich beim Thema „Video“ beinahe automatisch in den Vordergrund drängen.



Heute ist es sehr leicht, jemanden auf Video aufzunehmen. Das macht das Microteaching aber nicht nur einfacher.

Über alle Zeiten und Programmformen hinweg ist Microteaching **enorm gut belegt in der Wirksamkeit**. Jo Kramis fand eine Effektstärke von $d=0.8$, wobei er allerdings nicht Schulleistungen mass, sondern die Einschätzung der Lehrpersonen durch Experten (welche nicht wussten, welche Personen in den Programmen waren) sowie die Noten der Teilnehmenden bei den praktischen Abschlussprüfungen. Der Unterschied war auch ein Jahr später noch feststellbar, was sehr beachtlich ist. Es gibt wahrscheinlich nur wenige Dinge in der Ausbildung, deren Wirksamkeit man nach einem Jahr im Unterricht noch nachweisen kann.

Microteaching-Programme wirken auf vielen Ebenen

Die Komplexität von Microteaching-Programmen stammt daher, dass in ihnen **sehr viele Dinge geschehen**. Der Kern der Programme ist, dass angehende Lehrpersonen auf Video aufgenommen werden und sich dann (in Gruppen) über ihre Performance vor Kamera unterhalten. Aber was genau ist es, was den Effekt ausmacht? Ist es:

- Dass man **sich selbst auf Video** erlebt?
- Dass man **andere** EinsteigerInnen beim Unterrichten sieht?

- Dass man sich **darüber unterhält, welche konkreten Verhaltensweisen** in bestimmten Situationen **möglich und effektiv sind?**
- Dass man **Selbstsicherheit** und **Routine** entwickelt?
- Dass man **häufige Situationen (wie einen Unterrichtseinstieg) zu meistern lernt?**
- Dass man seine **Stärken kennen** und ausspielen lernt?
- Dass man **an seinen Schwächen arbeitet?**
- Dass man in einer sehr **verbindlichen Lernsituation** ist, deren Anforderungen man nicht ausweichen kann?

Vielleicht liegt ein Erfolgsrezept des Microteachings auch gerade darin, dass man sich mit so vielen verschiedenen Dingen intensiv auseinandersetzt. Gerade weil viele dieser Dinge in der Ausbildung nicht viel Platz finden.

Microteaching ist eine wirksame Methode in der Lehrerbildung

Eine häufig diskutierte **Frage** ist, **ob das Lehrpersonen-Studium theoretisch oder praktisch ausgerichtet sein soll**. Wir wollen diese Frage in diesem Abschnitt anhand des Microteachings erhellen. Wahrscheinlich ist die **Frage ganz einfach falsch gestellt**, und wir haben auch nicht genug Daten, um sicher sagen zu können, wie eine gute Lehrpersonen-Ausbildung aussieht.

Eines der ungemütlichen Ergebnisse bei Hattie ist, dass die **Lehrpersonen-Ausbildung gesamthaft eine eher geringe Effektstärke aufweist (d=0,11)**. Verschiedene Arten, Lehrpersonen auszubilden, scheinen nicht zu effizienteren oder weniger effizienten Lehrpersonen zu führen. Ausserdem sind neue Lehrpersonen nicht viel effizienter, wenn sie das eigene Fach unterrichten als eines, für das sie keine spezielle Ausbildung haben (d=0,09).

Belegt ist, dass man Unterrichten lernen kann. Lehrpersonen mit vollwertiger Qualifikation und mehrjähriger Berufserfahrung sind effizienter (d=0,39), insbesondere wenn sie auf dem Fach arbeiten, auf dem sie ausgebildet sind (d=0,38).

Die Effektstärke beruht auf drei Meta-Analysen und total 53 Studien. Hattie kommt hier zum Ergebnis, dass es **viel mehr Forschung braucht** (s.130), insbesondere, da das Thema sehr wichtig ist. **Microteaching ist eine der wenigen Ausbildungsmethoden, deren Wirkung auf die Lernleistung der Schülerinnen und Schüler belegt ist (d=0,88)**. Darüber hinaus wissen wir, dass einige Lehrpersonen-Faktoren sich sehr positiv auf das Lernen auswirken: Klarheit (d=0,75), Aufbau einer positiven Lehrer-Schüler-Beziehung (d=0,72), Verwendung von geeignetem Feedback (d=0,73). Allerdings ist nicht belegt, dass ein Training dieser Fähigkeiten im Studium möglich und effizient ist. Man kann aber davon ausgehen, dass es **sicher besser ist, auf wirksame Faktoren hinzuarbeiten, als auf unwirksame**. Die Unmittelbarkeit der Rückmeldung der Lehrperson ist beispielsweise etwas, das kaum zu besseren Leistungen führt (d=0,16). Ausbildung dazu – ob theoretisch oder praktisch – dürfte also wenig dazu beitragen, gute Lehrpersonen heranzubilden.

4. Lob und Belohnung (negative Effektstärken)

Lob und Belohnung ist bei Hattie **Teil des Faktors „Feedback“ (d=0.73)**. Der Faktor „Feedback“ ist bei Hattie sehr komplex – Hattie selbst scheint auch nicht immer sicher, was darunter zu verstehen ist. Es hat sich allerdings gezeigt, dass **Lob und Belohnung im Allgemeinen nicht besonders gut funktionieren und bisweilen zu massiven Leistungseinbußen führen** (bis zu $d = -0,78$ in einzelnen Studien). Hattie kommt zum Schluss, dass Lob und Belohnung möglicherweise gar nicht als Feedback verstanden werden soll, sondern als etwas ganz anderes.

An diesem Faktor kann man sehr gut sehen, warum es wichtig ist, die **Perspektive der Lernenden einzunehmen** und dem eigenen „Bauchgefühl“ gelegentlich etwas Misstrauen entgegenzubringen, oder es zumindest zu schärfen. Es ist ja auch gut, als Lehrpersonen zu wissen, dass man einen gewissen professionellen Vorsprung hat gegenüber Leuten, die das Unterrichten als reine Frage von Engagement und gesundem Menschenverstand ansehen.

Es fühlt sich gut und richtig an, zu loben und zu belohnen

Aus Sicht der Lehrperson sind Lob und Belohnung Highlights. Jemand gibt mir etwas (zum Beispiel eine richtige Antwort) und bekommt etwas dafür (meine artikuliert Wertschätzung). Das ist ein netter persönlicher Moment.

Zunächst einmal ist interessant, was denn genau gelobt (oder getadelt) wird. Bei den meisten Lehrpersonen sieht das wie folgt aus:

- **Lob** wird ausgesprochen **für gute intellektuelle Leistungen.**
- **Tadel** wird ausgesprochen **für störendes Sozialverhalten.**

Das ist ein gewisser Widerspruch. **Wenn Lob und Tadel wirksame Instrumente sind, dann sollte man auch die ruhigen Schülerinnen und Schüler für ihr vorbildliches Verhalten loben und die schlechten Schülerinnen und Schüler wegen ihrer Leistungen tadeln.** Beides tun wir üblicherweise nicht, weil wir irgendwie wissen, dass das nicht sehr wirksam wäre.

Negativ mit Lernenden umzugehen ist übrigens wenig effizient. Wenn die Lernenden positiv eingestellt sind gegenüber Lehrperson und Unterricht, dann lernen sie auch mehr, und das äussert sich auf mehreren Faktoren: Lehrer-Schüler-Beziehung ($d=0.72$), kooperatives vs. kompetitives Lernen ($d=0.54$), Angstreduktion ($d=0.40$). Hier stimmt das Bauchgefühl der meisten angehenden Lehrpersonen: **Positiver Austausch ist wichtig für Lernerfolg.**

Die Herausforderung ist zunächst, zu verstehen, wie Lob und Belohnung wirklich funktionieren und welche positiven und negativen Effekte sie auf den Lernerfolg der Schülerinnen und Schüler haben. In der Folge geht es dann darum, sich Strategien und Verhaltensmuster anzueignen, wie man positiv mit Schülerinnen und Schülern interagiert, ohne das Lernen zu beeinträchtigen.

Lob ist selten wirksam und häufig schädlich

Studien zu Lob finden üblicherweise nur sehr geringe positive Effekte. Bei Hattie ist „Lob“ kein eigener Faktor, sondern Teil des Feedbacks (s. 206), wobei er auch gleich die Frage stellt, ob man es wirklich dazuzählen soll. Zum einen ist Lob wenig wirksam ($d=0.12$ bei Wilkinson, 1980, und $d=0.09$ bei Kluger & deNisi, 1996), und zum anderen ist es genau betrachtet **kein „richtiges“ Feedback:**

- Lob gibt den Lernenden **kaum Informationen**, wie sie ihre Leistungen verbessern – es stellt einfach fest, dass sie gut sind.
- Lob fokussiert auf das zu erreichende **Ziel** und auf die **Schülerinnen und Schüler, die es erreicht haben**. Unterricht sollte sich aber vor allem mit den Lernenden befassen, welche noch unterwegs sind.
- Lob **stellt die Lehrperson ins Zentrum** und senkt die Verantwortung, welche Schülerinnen und Schüler selbst für ihr Lernen übernehmen.

Eine gute Zusammenfassung findet sich bereits 1855 bei Wilhelm Jakob Georg Curtmann und Friedrich Heinrich Christian Schwarz (leicht gekürzt):

Das Lob ist der Lohn in Worten, die Auszeichnung der Lohn in äusserlichen Zeichen. In Ihrer Anwendung muss die äusserste Behutsamkeit angewandt werden, wenn nicht mehr Schaden als Nutzen gestiftet werden soll. Lob hat schon das Bedenken gegen sich, dass es meist einen Tadel Anderer, nicht Gelobter in sich einschliesst, und dass der Gelobte die Zufriedenheit, welche ein Fremder in seinen Handlungen ausdrückt, nun auch zum Massstabe seiner Selbstschätzung macht, und über die Motive seiner Handlungen ebenso leicht hinwegsieht, als der aussen stehende; endlich weil aus dem einmaligen Lob leicht ein Anspruch auf öfteres Lob entspringt. Ein unbegründeter Tadel lässt sich meist wieder gut machen, die Erbitterung schwindet durch Anerkennung des Irrthums; ein unbegründetes Lob ist fast niemals zu verbessern, denn die Zurücknahme verletzt, ohne die wach gewordene Eitelkeit niederzudrücken.

Curtmann & Schwarz, 1855, S.104

Ein Hinweis: In diesem Abschnitt werden häufig ältere Werke zitiert. Das heisst nicht, dass die moderne Forschung wertlos ist. Gute Pädagogen und scharfe Denker gab es zu allen Zeiten, heute gibt es davon wahrscheinlich nicht mehr als vor 200 oder 2000 Jahren. Was neu ist, ist jedoch, dass wir heute Methoden haben, um zu prüfen, welche der vorgeschlagenen und überzeugend tönenden Argumente sich auch praktisch bewähren. Neu ist also wahrscheinlich nur sehr wenig von dem, was Hattie und andere an Ideen vorbringen. Neu ist, dass wir belegen können, welche davon bei den Schülerinnen und Schülern tatsächlich zu besserem Lernerfolg führen.

Wie man richtig lobt

In „Visible Learning for Teachers“ gibt Hattie einige Hinweise, wie man richtig lobt. Der wichtigste Rat ist, **Lob und Feedback zu trennen**. Wenn Lob und Feedback vermischt werden, dann kann dies sehr leicht dazu führen, dass das Feedback zu kurz kommt.

Auch hier geben Wilhelm Jakob Georg Curtmann und Friedrich Heinrich Christian Schwarz sehr präzise Anweisungen, die – im grossen Ganzen – nach wie vor Gültigkeit haben (ebenfalls leicht gekürzt):

Man spreche also das Lob wenigstens nicht sehr positiv und nicht in ungemessenen Ausdrücken aus; man prüfe die zu beurteilende Handlung sorgfältig, nicht blos nach ihrer äusseren Erscheinung, sondern nach ihren Triebfedern. Eigenschaften, besonders solche, welche nicht durch eignes Zuthun erworben sind, lobe und tadle man gar nicht, sondern immer nur das Streben, sie zu erwerben oder abzulegen. Gerade diese Art des Lobs, die giftigste unter allen, ist die gebräuchlichste bei ungebildeten Personen. Da wird Schönheit, Gewandtheit, Stärke, Fassungskraft, Gedächtnis, Verstand, Sanftmuth, Liebenswürdigkeit eines Kindes ins Angesicht gerühmt, als wäre es ein Gemälde (...). Das Lob sei nicht eine Vergleichung mit den minderen Leistungen anderer Kinder, sondern entweder eine einfache Bezeugung der Zufriedenheit des Erziehers, also ein Zeugnis, dass der Zögling seine Pflicht erfüllt habe (dabei bleibt die Voraussetzung stehen, dass noch weitere Pflichten zu erfüllen seien). (...) Diese Art des Lobes ist Aufmunterung zum Weiterstreben und besonders bei schüchternen Kindern anzuwenden. (...) Je sparsamer Lob angewandt wird, desto wirksamer ist es.

Curtmann & Schwarz, 1855, S.105

Die Resultate von Deci, Koestner und Ryan (1999) mahnen zu noch weiter gehender Zurückhaltung: auch **Mitteilungen, dass die Lernenden die von ihnen erwarteten Leistungen erbracht haben, führten oft zu schlechterer Lernleistung (d= -0,78)**. Lob ist also geeignet, wenn es die zukünftigen Anstrengungen verstärkt, und ungeeignet, um vergangene Leistungen und Anstrengungen zu würdigen.

Das heisst: **Ermuntern ist gut, loben ist schlecht**. Lob sollte nur sehr zurückhaltend eingesetzt werden, nicht auf erreichte Leistungen oder von den Lernenden nicht beeinflussbare Faktoren bezogen werden und keinesfalls als Ersatz für richtiges Feedback verstanden werden.

Belohnungen können die Motivation zerstören

Motivation ist ein sehr gut erforschtes Gebiet in der Psychologie. Edward L. Deci und Richard M. Ryan (Deci & Ryan, 1985) unterscheiden zwei Arten von Motivation:

- **Extrinsische Motivation** ist Motivation in Bezug auf ein Ziel oder eine Belohnung. Also: Niklaus macht seine Hausaufgaben, weil er nachher draussen spielen gehen darf.

- **Intrinsische Motivation** ist Motivation aus der Sache heraus. Also: Sven macht seine Hausaufgaben, weil er das Fach spannend findet und es ihm Freude macht, darin Dinge zu lernen und gut zu sein.

Es geht jetzt nicht darum, weshalb Niklaus oder Sven die Hausaufgaben machen. Entscheidend ist eine zentrale Erkenntnis der Motivationspsychologie:

Intrinsische Motivation kann durch extrinsische Belohnung zerstört werden.

Koestner, Deci und Ryan (1999) haben in einer Meta-Analyse festgestellt, dass in der bisherigen Forschung **Belohnungen generell zu schlechteren Leistungen führen** ($d = -0,34$), **insbesondere bei interessanten Aufgaben** ($d = -0,68$). Sie stellen fest, dass Belohnung zu mehr Überwachung, Bewertung und Konkurrenz führt, was wiederum das Engagement der Lernenden verringert.

Wie man richtig belohnt

Gutes Belohnen fokussiert nicht auf der Leistung und der Belohnung, sondern auf den Menschen, die sie austeilen und bekommen. Auch hier folgen wir Wilhelm Jakob Georg Curtmann und Friedrich Heinrich Christian Schwarz:

Um Belohnungen pädagogische Wirksamkeit zu verschaffen, dürfen sie erstens die Sinnlichkeit nicht stärker erregen, als absolut notwendig ist. Raffinierte Sinnreize sollte man niemals als Belohnung aussetzen. Das unverdorbene Kind freut sich schon über gar Weniges.

Curtmann & Schwarz, 1855, S.103

Also wie bei Lob: Auch bei der Belohnung ist Zurückhaltung angebracht. Grossartige Belohnungen sind unangebracht – und auch unnötig.

Zweitens sollen die Belohnungen niemals etwa Schädliches oder Verbotenes, wenngleich dem Kinde noch so erwünschtes erhalten.

Curtmann & Schwarz, 1855, S.103

Schülerinnen und Schüler werden gelegentlich versuchen, für gute Arbeit die Regeln des Schulbetriebs zu ändern, also beispielsweise einen Film schauen zu dürfen oder etwas früher gehen zu dürfen, wenn sie ihrer Ansicht nach gut gearbeitet haben. Dazu auch folgendes:

Wenn die Belohnung nicht als eine freie Gabe des Wohlwollens erscheint, wenn sie bedungen, vielleicht gar abgedungen ist, so kann dieselbe unter keinen Umständen als richtiges Erziehungsmittel gelten, sie wird vielmehr ein Reiz zum sittlichen Verderben. (...) Lohn-Diener sind schlechte Diener, Lohn-Zöglinge noch schlechtere Zöglinge.

Curtmann & Schwarz, 1855, S.103

Im konkreten Fall hilft die **Einsicht, dass wir als Lehrpersonen die intrinsische Motivation der Lernenden brauchen**, damit sich die Schülerinnen und Schüler ausgiebig und engagiert mit dem Stoff auseinandersetzen. Andererseits **können wir es uns**

auch schlicht nicht leisten, die Schülerinnen und Schüler für ihren Einsatz zu entlönnen. Womit auch? Einige Minuten zusätzliche Pausenzeit oder das Schauen eines Films in der letzten Stunde vor den Ferien entschädigen ja nicht adäquat für stundenlanges Arbeiten.

Das bedeutet nicht, dass man nie einen Film schauen oder Schülerinnen und Schüler etwas gönnen darf. Nur **eignen sich diese Dinge nicht als Hebel, um Lernen zu begünstigen.** Wenn sie als solches eingesetzt werden, dann sind sie eher schädlich.

Auch hier ist vor allem wichtig, dass Belohnung **kein Ersatz für Feedback** ist. Auch wenn die Klasse für gute Leistungen etwas bekommt, dann ersetzt dies nicht die Rückmeldung, was diese Leistungen anbelangt. Belohnungen sind dann wirksam, wenn sie zu vermehrter zukünftiger Anstrengung führen, ohne allerdings die intrinsische Motivation zu beeinträchtigen. Das kann sehr schwierig zu bewerkstelligen sein.

Eine Technik es zu erreichen ist der so genannte „**Secret Student**“ den Dylan Wiliam in der BBC-Serie „The Classroom Experiment“ vorgestellt hat (BBC, 2010). Dabei wird aus der Klasse **jeden Tag zufällig und geheim eine Schülerin oder ein Schüler ausgewählt.** Wenn das Verhalten der Schülerin oder des Schülers über den Tag hinweg zufriedenstellend war, dann bekommt die Klasse einen Punkt gutgeschrieben und der Name des „Secret Students“ wird bekannt gegeben. Wenn es keinen Punkt gibt, dann nicht. Wenn in einem bestimmten Zeitraum ausreichend Punkte erreicht werden, wird dies **mit einem Klassenausflug belohnt.** Durch den „Secret Student“ **steigt die Motivation, sich in der Schule zu benehmen, ohne dass die intrinsische Motivation für die Fächer in Gefahr ist.**

5. Verteiltes Lernen ($d=0.71$)

Lernen ist effizienter, wenn es über längere Zeiträume hinweg verteilt stattfindet (Baddeley, 1990). Als Hermann Ebbinghaus 1885 damit begann, Lernen wissenschaftlich zu untersuchen, war dies eine seiner ersten Erkenntnisse (Ebbinghaus, 1885). Wenn er Lernen gleichmässig auf zwei Tage verteilte, dann brauchte er gesamthaft weniger Zeit, um die gleiche Menge an Stoff zu lernen.

Der Effekt wurde in den folgenden Jahren ausgiebig untersucht, und 1938 konnte Robert Woodworth für sein Lehrbuch bereits eine **breite Menge von Untersuchungen** zusammentragen, die den Vorteil des verteilten Lernens für viele Aufgaben belegten, **vom Labyrinthlernen bei Tanzmäusen bis zum Bogenschiessen** (Woodworth, 1938). In einer Studie zum Lernen von **Schreibmaschine-Schreiben** lernten die Probanden mit einer Stunde Training pro Tag fast doppelt so effizient wie mit 2 Blöcken von 2 Stunden pro Tag (Baddeley & Longman, 1978). Lehrpersonen und Lernende sind hingegen üblicherweise der Ansicht, dass massiertes Lernen die besten Effekte hat (Landauer & Ross, 1977; Rothkopf, 1963). Ulrich Neisser, einer der grossen Psychologen, verfasste in der Folge ein Gedicht, um das verteilte Lernen bekannter zu machen:

You can get a good deal from rehearsal,
If it just has the proper dispersal.
You would be an ass
To do it en masse:
Your remembering would turn out much worsal.

Zu deutsch ungefähr:

Wiederholung kann dir viel erhalten,
Nur musst du verteilt sie gestalten.
Du wärest ein Pflock,
Zu tun es „en bloc“,
und würdest nur wenig behalten.

„Verteiltes Lernen“ ist übrigens der psychologische Fachbegriff, Hattie bezeichnet den Faktor als „spaced vs. massed practice“ (s. 220), in der deutschen Übersetzung „rhythmisiertes vs. geballtes Lernen“. Damit ist jeweils die gleiche Sache gemeint.

Das verteilte Lernen soll hier als Beispiel für einen einfach umzusetzenden und trotzdem wenig beachteten Faktor verwendet werden, und damit auch als Illustration, wie mit Hatties Ergebnissen umgegangen wird. Hattie selbst behandelt den Faktor recht kurz und beiläufig auf etwas mehr als einer halben Seite, obwohl es sich dabei um einen der grössten Effekte handelt, der gut an psychologische Erkenntnisse und Schulpraxis angeschlossen werden kann.

Exkurs: Einfach zu verstehende Faktoren erhalten mehr Aufmerksamkeit als einfach umzusetzende

Die Faktorenliste im Anhang von „Lernen sichtbar machen“ (s. 433) ist einerseits sehr übersichtlich, andererseits auch voll von Fachwörtern, deren Bedeutung man anhand des Namens oft nicht erfassen kann. Hinzu kommt, dass die Liste auch Faktoren enthält, die für das Unterrichten nicht relevant sind, wie etwa das Geburtsgewicht ($d=0.54$), welches ja durch Lehrpersonen und Schulstrukturen nicht beeinflusst werden kann.

Wie entscheidet sich nun, welche Faktoren eine breitere Aufmerksamkeit erhalten?

Das hängt zuerst einmal davon ab, ob man sich etwas darunter vorstellen kann, wenn man den Namen des Faktors liest. **Folgende Faktoren sind beispielsweise sehr zugänglich:**

- **Klarheit der Lehrperson** ($d=0.75$)
- **Feedback** ($d=0.73$)
- **Lehrer-Schüler-Beziehung** ($d=0.72$)

Überhaupt wird die **Wichtigkeit der Lehrperson** oft als zentrales Ergebnis der Hattie-Studie genannt:

Auf den Lehrer kommt es an

Die folgenden Faktoren haben vergleichbare Effektstärken, sind aber auch Fachpersonen oft unbekannt und erhalten sehr viel weniger Beachtung in der allgemeinen Diskussion:

- **Akzeleration** (d=0.88)
- **Reziprokes Lernen** (d=0.77)
- **Wiederholendes Lesen** (d=0.67)
- **Lautier-Methode** (d=0.60)
- **Mastery-Learning** (d=0.58)

Einige Autoren stellen auch fest, dass Hatties Ergebnisse bisweilen sehr eigenwillig interpretiert werden (Rolf, 2013). Dies ist zum Beispiel der Fall für die **direkte Instruktion** (d=0,59), welche **bisweilen als Frontalunterricht missverstanden** wird.

Bildungswesen

Frontalunterricht macht klug

Problemorientierter oder offener Unterricht - die ganze moderne Pädagogik stiftet wenig Nutzen. Am besten ist noch immer moderner Frontalunterricht, fanden Forscher heraus.

Forscher: Frontalunterricht ist besser als sein Ruf

Geschrieben von: Redaktion 13. Februar 2013 Rubrik: Praxis 4 Kommentare

MÜNCHEN. Lehrerzentrierter Frontalunterricht? Gilt nicht gerade als letzter Schrei der Pädagogik. Gleichwohl erlebt die Methode derzeit eine erstaunliche Wiederauferstehung.

Hier wird ein relativ umfassendes pädagogisches Konzept – die direkte Instruktion – auf ein Schlagwort und damit einen einzelnen Aspekt reduziert, nämlich „jemand steht vorne und redet“. In Wahrheit umfasst die direkte Instruktion (Adams & Engemann, 1996) viele Schritte: Die Lehrperson definiert die Ziele und Erfolgskriterien, demonstriert sie, überprüft sie, leitet Übungen an und fasst abschliessend die Lerngewinne zusammen. Das ist nicht das gleiche wie wenn die Lehrperson einfach vorne steht und redet bis die Pausenglocke läutet.

Was bedeutet das? – Es heisst, **wir müssen Hattie's Liste aufmerksam lesen. Die Versuchung ist gross, aus der Liste diejenigen Dinge herauszupicken, die bereits bekannt sind, oder die wir ohnehin schon umsetzen. Dabei übersehen wir diejenigen Dinge, die uns weiterbringen würden, und möglicherweise auch sehr einfach zu realisieren sind** – wie etwa das verteilte Lernen.

Lernen zu verteilen ist einfach und wirksam – wenn die Umstände es zulassen

Eigentlich sind fast alle Zugangsweisen zum verteilten Lernen sinnvoll, so sie denn umsetzbar sind. Die folgende Liste soll einige Anregungen erhalten. Es ist mir bewusst, dass viele davon nicht sinnvoll einsetzbar sind in allen Fällen, also durch Lehrpläne, Notenkonferenzen, eingeschränkte Kapazitäten oder Motivationsschwierigkeiten blockiert werden. Wenn sich aber Möglichkeiten ergeben, dann kann hier sehr einfach und effizient das Lernen verbessert werden.

- **Mit den komplizierten Dingen anfangen.** Intuitiv baut man Stoff oft so auf, dass man die einfachen Dinge an den Anfang stellt und die schwierigen an den Schluss. Wenn man mit den schwierigen Dingen beginnt, dann haben die Lernenden mehr Zeit, diese zu üben und können das Lernen über längere Zeiträume verteilen.
- **Aktive Prüfungen.** Prüfungen sind zu mehr zu gebrauchen als nur dazu, Wissen abzurufen. Multiple Choice Prüfungen (und ähnliche Formate) sind meist aktiver als Textfragen, da hier die erlernten Konzepte aktiv auf die verfügbaren Antworten angewendet werden müssen. Ausserdem sind Multiple Choice Prüfungen meist schneller durchgeführt, so dass sie im Anschluss besprochen werden können – was eine weitere Lernrunde darstellt. Auch ist eine Stunde zur Prüfungsvorbereitung oft ein guter Anlass, um Schülerinnen und Schüler zu engagiertem Lernen zu bewegen.
- **Themen längs aufteilen.** Es ist verlockend, die zur Verfügung stehende Zeit in Abschnitte zu unterteilen, an denen intensiv ein Thema nach dem anderen angegangen wird. Eigentlich wäre es besser, mehrere Themen parallel und in kürzeren, weiter verteilten Einheiten zu behandeln.
- **Zwischenprüfung.** Prüfungen sind sehr effizient, um Lernende auf einen hohen Wissensstand zu bringen. Es ist eigentlich schade, wenn man sie am Schluss des Themas macht. Dann lernen die meisten Schülerinnen und Schüler den Stoff dann nur ein Mal in Form eines massierten Lernblocks und haben auch nichts mehr davon, dass sie jetzt gut sind, da ja gleich zum nächsten Thema gewechselt wird.
- **Nachprüfung.** Bei klassischen Prüfungen erhalten die Lernenden die Note üblicherweise einige Zeit nach Abschluss des Themas. Wenn man Wiederholungsprüfungen anbietet, dann müssen die Lernenden das Thema mit etwas Abstand noch einmal aufbereiten, was die Behaltensleistung verbessert.
- **Stoff über die Ferien hinweg legen.** Was, wenn der Stoff vor den Ferien nicht mehr fertig wird? – Aus Sicht der Behaltensleistung ist es sehr vorteilhaft, wenn der Stoff über die Ferien hinweg geht. Ziel ist ja nicht, dass die Kinder in den Ferien möglichst viel zum Vergessen haben, sondern dass sie gesamthaft viel können. Wenn nach den Ferien zuerst einmal der Stoff aufbereitet, das Thema beendet und dann die Prüfung geschrieben wird, dann hilft das.
- **Wiederholungsrunden.** Es zahlt sich aus, einige Strategien und didaktische Werkzeuge zu haben, mit denen man effizient Stoff repetieren kann. Das kann zum Beispiel eine kurze Fragerunde sein, in der die Kinder sich gegenseitig die zentralen Inhalte noch einmal

erklären. Dazu gibt es viele mögliche Formen. Ich verwende häufig Ampelbecher: Die Studierenden können mit farbigen Bechern anzeigen, ob sie ein Konzept verstanden haben, und ich lasse dann gegebenenfalls das Konzept noch einmal kurz von jemandem erklären, die oder der es begriffen hat. Eine andere Möglichkeit sind Kurzvorträge (im Sinn von Vortragsübungen oder „Science Slams“) über die bereits behandelten Dinge.



Ampelbecher: Eine gute Feedbackmethode, und auch ein guter Anlass, um die Sicherheit mit vergangenen Inhalten zu überprüfen und gegebenenfalls zu verbessern. Die Becher werden jeweils übereinandergesteckt, die obere Farbe zeigt die subjektive Sicherheit mit dem Stoff (grün=hoch, gelb=mittel, rot=tief).

Das verteilte Lernen ist allerdings bisher **noch nicht hinreichend für die Schulrealität aufbereitet**. Das ist schade, da es sich um einen zuverlässigen, psychologisch und pädagogisch bewährten Faktor handelt. Die Umsetzung ist zwar organisatorisch nicht immer ganz einfach, aber in der Regel ohne grössere Kosten oder Aufwände möglich. Schliesslich werden ja die gleichen Dinge gelernt, nur eben etwas weniger massiert.

Hier wäre also noch reichlich Platz für weitere Forschung oder praktische Ratgeber.

Literatur

Adams, G. L. & Engelmann, S. (1996). *Research on direct instruction: 20 years beyond DISTAR*. Seattle, WA: Educational Achievement Systems.

Baddeley, A. D. (1990). *Human memory: Theory and practice*. East Sussex: Lawrence Erlbaum.

Baddeley, A. D. & Longman, D. J. A. (1978). The influence of length and frequency on training sessions on the rate of learning to type. *Ergonomics*, 21, 627-635.

Barnes, D.E. & Bero, L. A. (1998). Why review articles on the health effects of passive smoking reach different conclusions. *JAMA*, 279(19), 1566–1570.

BBC (2010, 11. April). The Classroom experiment (Ep. 1) [Videodatei]. Zugegriffen über www.youtube.com/watch?v=J25d9aC1GZA

BBC (2010, 11. April). The Classroom experiment (Ep. 1) [Videodatei]. Zugegriffen über www.youtube.com/watch?v=1iD6Zadhg4M

Bero, L. A. (2005). Tobacco industry manipulation of research. *Public Health Rep.* 120(2). 200–208.

Blatchford, P. (2003). *The Class Size Debate: Is Small Better?* Maidenhead: Open University Press 182pp.

Borman, G. D. , Hewes, G. M., Overman, L. T. & Brown, S. (2003). Comprehensive School Reform and Achievement: A Meta-Analysis. *Review of Educational Research*, 73(125). doi: 10.3102/00346543073002125

Brahm, G. (2006). Klassengröße: eine wichtige Variable von Schule und Unterricht? *Bildungsforschung*, 2006/3, Ausgabe 1.

Bundesamt für Statistik (2013). Bildungssystem Schweiz - Indikatoren: Investitionen und Kosten - Öffentliche Bildungsausgaben. Heruntergeladen von <http://www.bfs.admin.ch/bfs/portal/de/index/themen/15/17/blank/01.indicator.402101.4012.html>

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New Jersey: Lawrence Erlbaum Associates.

Cohen, J. (1992). A power primer. *Psychological Bulletin* 112(1). 155–159. doi: 10.1037/0033-2909.112.1.155

Curtmann, W. J. G. & Schwarz, F. H. C. (1855). *Lehrbuch der Erziehung und des Unterrichts: Die Erziehungslehre*. Heidelberg: Winter.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125(6), 627-668. doi: 10.1037/0033-2909.125.6.627

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.

Finn, J. D. (1998). *Class Size and Students at Risk. What is known? What is next?* Washington D.C: National Institute on the Education of At-Risk Students, U.S. Department of Education.

Freeman, H. E. (1984). A meta-analysis of gender differences in mathematics achievement. Unpublished Ph. D., The university of Alabama, AL.

Friedman, L. (1989). Mathematics and the gender gap: A meta-analysis of recent studies on sex differences on mathematical tasks. *Review of Educational Research*, 59(2), 185-213.

Frost, L. A., Hyde, J. S. & Fennema, E. (1994). Gender, mathematic preference, and mathematics-related attitudes and affect: A meta-analytic synthesis. *International Journal of Academic Research*, 21, 373-385.

Grawe, K. Donati, R. & Bernauer, F. (1994). *Psychotherapie im Wandel - von der Konfession zur Profession*. Göttingen: Hogrefe.

Gunter, B., Berry, C. & Clifford, B.R. (1981). Proactive interference effects with television news items: Further evidence. *Journal of Experimental Psychology: Human Learning and Memory*, 7, 480-487.

Hattie, J. A. C. (2006). The paradox of reducing class size and improved learning outcomes. *International Journal of Educational Research*, 42, 387-425.

Hattie, J. A. C., Beywl, W. & Zierer, K. (2013). *Lernen sichtbar machen*. Baltmannsweiler: Schneider.

Hattie, J. A. C., Beywl, W. & Zierer, K. (2014). *Lernen sichtbar machen für Lehrpersonen*. Baltmannsweiler: Schneider.

Hebb, Donald (1949). *The Organization of Behavior*. New York: Wiley & Sons.

Helmke, A. & Weinert, F. E. (1997). Wissenschaftliche Grundlagen, Realisierungsbedingungen und Ergebnisperspektiven des Längsschnittprojektes SCHOLASTIK. In F.E. Weinert & A. Helmke (Hrsg.), *Entwicklung im Grundschulalter* (S. 1-12) Weinheim: Psychologie Verlags Union.

Hines, H. E. (1989). *Gender-related differences in mathematics participation and achievement: A meta-analysis*. Unpublished Ed. D., University of Houston, Texas, United States.

Hyde, J. S. (1981). How large are cognitive gender differences? A meta-analysis using ω^2 and d . *American Psychologist*, 36, 892–901.

Hyde, J. S. (2005). The Gender Similarities Hypothesis. *American Psychologist*. 60 (6), 581–592 doi: 10.1037/0003-066X.60.6.581

Hyde, J. S., Fennema, E. & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, 107(2), 139-155.

Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin*, 104, 53–69.

Jäncke, L. (2009). The plastic human brain. *Restorative neurology and neuroscience* 27(5): 521-38. doi: 10.3233/RNN-2009-0519

Kirsch, I., & Sapirstein, G. (1998). Listening to Prozac but hearing placebo: A meta-analysis of antidepressant medication. *Prevention and Treatment* 1 (2). doi:10.1037/1522-3736.1.1.12a

Landauer, T. K. & Ross, B. H. (1977). Can simple instructions to use spaced practice improve ability to remember a fact? An experimental test using telephone numbers. *Bulletin of the Psychonomic Society*, 10, 215-218.

Köhler, W. (1917). *Intelligenzprüfungen an Anthropoiden*. Berlin.

Köhler, W. (1925). *The mentality of apes*. London: Kegan.

Kramis, J. (1988): Erfahrungen mit einer Kombination von Microteaching, reflective teaching und Unterrichtsbeobachtung. Eine empirische Untersuchung an 146 Lehramtskandidaten. In: *Beiträge zur Lehrerbildung* 6(3), S. 283-288.

Kühn, R. (1986). Zusammenhänge zwischen Klassenfrequenz, affektiven Persönlichkeitsmerkmalen und Schulnoten bei Schülern der vierten Klasse. In: *Psychologie der Erziehung und Unterricht* 33, 277-284.

Landsberger, H. A. (1958). *Hawthorne Revisited*, Ithaca, N.Y.: Cornell University.

- Mansell (2008, 21. November). Research reveals teaching's Holy Grail. *TES Newspaper*.
- Mitra, S. (2013). <http://www.hole-in-the-wall.com>
- Mitra, S. (2014). <http://grannycloud.wordpress.com>
- Petillon, H. (1985). Klassenfrequenz: Überlegungen zu einem systematischen Erklärungsansatz. In K. Ingenkamp, H. Petillon & M. Weiss (Hrsg.), *Klassengröße: je kleiner, desto besser? Forschungs- und Diskussionsstand zu Wirkungen der Klassenfrequenz* (S. 147-189). Weinheim: Beltz.
- Rolff, H. G. (2013). Die Hattie-Studie: Ein Rorschach-Test. *Pädagogik*, 4/13. Landsberg: Beltz.
- Rothkopf, E. Z. (1963). Some observations on predicting instructional effectiveness by simple inspection. *Journal of Programmed Instruction*, 2, 19-20.
- Saldern, M. v. (1985). Der Einfluss der Klassenfrequenz auf die subjektiv wahrgenommene Lernumwelt. In K. Ingenkamp, H. Petillon & M. Weiss (Hrsg.), *Klassengröße: je kleiner, desto besser? Forschungs- und Diskussionsstand zu Wirkungen der Klassenfrequenz* (S. 147-189). Weinheim: Beltz.
- Saldern, M. v. (1993): Klassengröße als Forschungsgegenstand. Landau: Verlag für Empirische Pädagogik.
- Shatz (2007): Zitiert nach Doidge, N. (2007). *The Brain That Changes Itself*. United States: Viking Press. p. 427.
- Tulving, 1983. *Elements of episodic memory*. Oxford: Clarendon.
- Wiliam, D. (2011). *Embedded Formative Assessment*. Cambridge: Perfect Paperback.
- Willett, J. B., Yamashita J. J. M. & Anderson, R. D. (1983). A meta-analysis of instructional systems applied in science teaching. *Journal of Research in Science Teaching*, 20(5), 405-417.
- Woodworth, R. S. (1938). *Experimental psychology*. London: Methuen.
- ZLV (2013). Umfrage zu Zufriedenheit und Belastung bei Lehrpersonen. Zugegriffen über <http://files.newsnetz.ch/upload//2/8/28353.pdf>