

Erscheint in *Lehren & Lernen*, Okt. 2018.

Kann die Reformpädagogik von der Bildungsforschung profitieren?

Georg Lind

Im Folgenden werde ich einige häufig genannte Kriterien für die Beurteilung von Bildungsforschung diskutieren, anhand derer Schulpraktiker deren Befunde kritisch bewerten und das Brauchbare vom Unbrauchbaren trennen können.

Ist der Umfang einer Studie ein Güte Merkmal?

Jeder größer die Studie, umso gültiger seien ihre Ergebnisse, sagt man. Falsch. Genau im großen Umfang liegt oft ein Grund für falsche Schlussfolgerungen. Mit wachsendem Umfang wächst nämlich rein mathematisch die Wahrscheinlichkeit, dass minimale Unterschiede statistisch "signifikant" werden, was überhaupt nichts mit "bedeutsam" (der deutschen Bedeutung von "signifikant") zu tun hat. "Signifikante" Unterschiede können winzig sein. Selten sind sie pädagogisch bedeutend.

Effektstärke und Effizienz einer Methode sind wichtig, werden aber oft übersehen

Schulpraktiker wollen von der Bildungsforschung wissen, wie groß der absolute Lernzuwachs bestimmter Lehrmethoden im Vergleich zu anderen ist und wie groß der damit verbundene zeitliche und finanzielle Aufwand im Vergleich zum Erfolg ist. Sie wollen auch wissen, welche Ausbildung nötig ist, um den gewünschten Lerngewinn zu erreichen. Leider werden Effektstärke und Effizienz in fast keiner Studie berichtet, oft selbst dann nicht, wenn die entsprechenden Daten vorliegen. Dabei macht es einen großen Unterschied, ob eine Lehrmethode nur geringe oder große Auswirkungen auf die Fähigkeitsentwicklung von SchülerInnen hat, und ob ihre Einführung und Durchführung mit hohen oder mit geringen Kosten verbunden sind. Solche Abschätzungen sollten wenn immer möglich gemacht werden. Sie können auch helfen, neue Unterrichtsmethoden einzuführen, wenn gezeigt werden kann, dass sie viel effektiver sind als die herkömmlichen, aber kaum große

Veränderungen benötigen. Wie fruchtbar Schätzungen der Effektivität und Effizienz sein können, habe ich auf dem Gebiet der moral- und demokratiepädagogischen Forschung gezeigt (Lind 2012). Lawrence Kohlberg und seine Kollegen hatten mit der Dilemma-diskussion eine sehr effektive Maßnahme zur Förderung von Moralkompetenz entwickelt. Trotzdem hat Kohlberg seine Methode für tot erklärt und von ihrer Verwendung abgeraten. Er begründete dies damit, dass nach Abschluss der Erprobungsstudien keiner der teilnehmenden Lehrpersonen die Methode weiter benutzen wollte. Ein Grund für seine Entscheidung mag auch gewesen sein, dass er das wahre Ausmaß der Effektivität seiner Methode selbst nicht kannte, obwohl alle Interventionsstudien „signifikante“ Ergebnisse zeigten. Erst meine Meta-Analyse von ca. 140 Interventions-Studien zeigte, dass die relative Effektivität seiner Dilemmamethode ($r=0.40$) weit über den mittleren Effektstärken von anderen erfolgreichen Lehr- und Therapiemethoden ($r=.30$) lag. Auch zeigte sich, dass keine der analysierten Interventionsstudien einen negativen Ausgang hatte (Lind 2012).

Ich fand, dass die Dilemmamethode viel zu wirksam war, um vergessen zu werden, und dass sie noch verbessert werden konnte. Unsere *Konstanzer Methode der Dilemma-Diskussion* (KMDD) ist tatsächlich noch effektiver und effizienter. Und anders als bei Kohlberg sind auch die Rückmeldungen der Lehrpersonen, die sie benutzen, durchweg positiv. Soweit ich weiß, benutzen alle sie auch ohne äußere Motivation weiter, wohl auch deshalb, weil die KMDD, wie sie uns sagen, nicht nur die Moralkompetenz fördert, sondern auch sofort und nachhaltig das Lernklima in der Klasse verbessert (Lind 2015).

Angemessene Forschungsdesigns sind unverzichtbar

Oft wird für Studien gefordert, dass sie repräsentativ sind, das heißt die Vielfalt einer Population abbilden. Dies ist aber nur erforderlich, wenn sie einen Zustand beschreiben sollen, wie etwa bei Wählerstudien. Für analytische Studien, die eine Hypothese über die Wirksamkeit bestimmter Lehrmethoden oder Unterrichtsorganisationen auf den Prüfstand stellen, ist es viel wichtiger, dass sie richtig angelegt sind, dass damit z.B. unerwünschte Nebenwirkungen (Verringerung der Lernmotivation), bestimmte Voraussetzungen (wie die notwendige Qualifikation der Lehrpersonen) und personenspezifische Effekte (sozialer Hintergrund der Schüler) erkannt werden können. Hier versagt die Bildungsforschung noch meist. Wenn man bei PISA statt einer bestimmten Altersgruppe zwei Schuljahrgänge erhoben

hätte, dann könnten wir abschätzen, wie groß die Effizienz unseres Bildungssystems ist. So kann jeder aus den PISA-Befunden herauslesen, was ihm gefällt.

Forschung ist die Grundlage für gute Evaluation

Ich kann heute, nach über 40 Jahren Beschäftigung mit dem Konzept der *Moral- und Demokratiekompetenz* mit Sicherheit sagen, dass sich die Investition an Zeit und Geld in die Grundlagenforschung gelohnt hat, die der pädagogischen Anwendung vorausgegangen. Ohne ein gründliches Verständnis des Gegenstands wäre es nicht möglich gewesen, ein Messverfahren zu entwickeln, das sich trotz seiner Kürze in vielen Forschungsprojekten bewährt und als valide erwiesen hat. Weil der *Moralische Kompetenz-Test* (MKT) kurz und dazu noch in vielen Altersgruppen (ab ca. zehn Jahren) eingesetzt werden kann, kann damit die Wirksamkeit verschiedener Methoden unter unterschiedlichen Bedingungen studiert werden. So können wir heute mit einiger Sicherheit sagen, dass die KMDD viel wirksamer und auch viel effizienter als viele andere Methoden ist und dass diese Wirksamkeit kaum von Merkmalen der Schüler, aber von einer guten Ausbildung der Lehrperson in der KMDD abhängt. Wir wissen dadurch auch, dass für eine merkliche Wirkung bereits ein bis zwei Veranstaltungen pro Jahr reichen, also nicht, wie bei vielen anderen Methoden, Änderungen der Stundentafel oder Einschränkungen beim Stoffplan notwendig sind (Lind 2015).

Schluss

Wenn die oben genannten Kriterien berücksichtigt würden, würde Bildungsforschung mit Sicherheit noch mehr zur Verbesserung des Lehrens und Lernens an den Schulen führen. Ich bin auch sicher, dass alternative Schul- und Lernformen stärkere Beachtung und Verbreitung finden würden, wenn nicht nur akademische Kompetenzen, sondern auch moralisch-demokratische Fähigkeiten evaluiert werden würden. Dann würde sich nämlich zeigen, dass Reformschulen, die das gemeinsame Lernen in heterogenen Lerngruppen im Fokus haben, mehr erbringen als nur vergleichbare Abschlüsse, nämlich auch eine Vergrößerung der Fähigkeit, Konflikte und Probleme auf der Grundlage universeller Moralprinzipien zu lösen, und zwar durch Denken und Diskussion, statt durch Gewalt, Betrug oder Unterwerfung unter Andere (Lind 2016).

Literatur

Lind, G. (2002). *Ist Moral lehrbar?* Berlin: Logos.

Lind, G. (2015). *Moral ist lehrbar*. Berlin: Logos.

Lind, G. (2016). Gemeinsames Lernen braucht Moralkompetenz. *Gemeinsames Lernen*, 2016, Heft 4, 42-47.

Weitere Informationen und Veranstaltungen: <https://www.uni-konstanz.de/ag-moral/>