

Experimentelle Studien in der Grundschulforschung

Andreas Hartinger

Angaben zur Veröffentlichung / Publication details:

Hartinger, Andreas. 2008. "Experimentelle Studien in der Grundschulforschung." In *Lehr-Lernforschung und Grundschulpädagogik*, edited by Frank Hellmich, 107–21. Bad Heilbrunn: Julius Klinkhardt.

Nutzungsbedingungen / Terms of use:

licgercopyright

Dieses Dokument wird unter folgenden Bedingungen zur Verfügung gestellt: / This document is made available under these conditions:

Deutsches Urheberrecht

Weitere Informationen finden Sie unter: / For more information see:

<https://www.uni-augsburg.de/de/organisation/bibliothek/publizieren-zitieren-archivieren/publiz/>



Experimentelle Studien in der Grundschulforschung

Ausgangspunkt dieses Beitrags ist eine Forderung, die Frank Hellmich und Hanna Kiper in ihrer „Einführung in die Grundschuldidaktik“ an die Grundschulforschung stellen. Sie mahnen an, dass „bislang nur wenige Unterrichtsmethoden in experimentellen Studien hinsichtlich ihrer Effektivität geprüft sind“ (2006, S. 158). Als Problem konstatieren sie in diesem Zusammenhang den Nachweis spezifischer Effekte, da die Forschung in der Grundschule bislang eher auf „einer eher groben und damit grundlagenorientierten Ebene“ (2006, S. 158) stattfände, so dass zwar allgemeine Hinweise auf Effekte gefunden werden können, es allerdings nur wenig differenzierte Befunde gibt, insbesondere auch bezüglich der Bedingungen, unter denen bestimmte Effekte auftreten.

Ziel dieses Beitrags soll nun sein zu überlegen, inwieweit es möglich und sinnvoll ist, solche „experimentellen Studien“ in der Grundschulforschung durchzuführen. Dazu soll zunächst der Begriff des Experiments (in der pädagogischen Forschung) kurz besprochen werden; im Anschluss daran folgen Überlegungen, welche Bereiche bzw. grundlegende Fragestellungen der Grundschulforschung tatsächlich einer experimentellen Überprüfung bedürfen, und inwieweit experimentelle Studien hierzu schon vorliegen. Abschließend werden die Ausführungen in einem Fazit zusammengefasst und konkretisiert.

1. Experimentelle Forschung

1.1 Allgemeine Definition

Naturgemäß ist es sinnvoll, zu Beginn eines Beitrags über „Experimentelle Studien in der Grundschulforschung“ zu klären, was unter einem „Experiment“ oder einer „experimentellen Studie“ in der pädagogisch-psychologischen Forschung verstanden wird. Ein recht bekannter Definitionsversuch hierfür stammt von Karl Josef Klauer, der den Begriff folgendermaßen fest-

legt: „Ein Experiment ist ein vom Forscher in Gang gesetztes Geschehen, das (1) nach einem bestimmten Plan abläuft, (2) grundsätzlich wiederholbar ist und bei dem (3) eine isolierende Variation der Bedingungen vorgenommen wird, um anschließend festzustellen, welche Auswirkungen die isolierende Bedingungsvariation auf mindestens eine Kriteriumsvariable hat“ (Klauer, 1973, S. 56). Damit dies erreicht werden kann, wird von einem Experiment (mit Menschen) erwartet, dass das Experiment an einer Zufallsstichprobe mit anschließender Randomisierung (also der zufälligen Verteilung der Teilnehmerinnen und Teilnehmer auf die unterschiedlichen Gruppen) durchgeführt wird. Nicht direkt erwähnt, allerdings impliziert, ist in dieser Definition, dass vor der Durchführung des Experiments eine Hypothese aufgestellt werden muss, die dann überprüft wird. Selbstverständlich gilt auch für experimentelle Untersuchungen, dass sie – genauso wie jede andere Form der wissenschaftlichen empirischen Datengewinnung – nicht ohne Theorie und ohne vorab offen gelegte Fragestellung sinnvollerweise durchgeführt werden können (vgl. z.B. schon Lay, 1905, S. 586ff.). Durch die Untersuchung von Kausalität erwartet man bei Experimenten jedoch zusätzlich, dass offene Fragestellungen bereits als Hypothesen spezifiziert worden sind (vgl. Huber, 2005, S. 69).²⁵

Ziel eines Experiments ist demnach die Überprüfung einer (oder mehrerer) Hypothesen, wobei durch die besondere Ausrichtung eines Experiments der Anspruch entsteht, Kausalität, also „Ursache-Wirkungs-Beziehungen“ nachweisen zu können. Nur durch die vorweg überlegte und dann auch gezielt durchgeführte, kontrollierte Variation genau einer Variablen (oder genau einer Variablenkombination) ist es möglich – so die wissenschaftstheoretische Logik – Wirkungen genau auf die Ursachen (eben die variierte Variable) zurückführen zu können (vgl. z.B. Huber, 2005, S. 69ff.).

1.2 Experimente in der Unterrichtsforschung

Für die pädagogische Diskussion sind die oben genannten Kriterien sicherlich zu spezifizieren; manchmal sind sie nur schwer oder auch nicht zu erfüllen – dies gilt in besonderem Maße für die Unterrichtsforschung.

a) So ist schon die Fragestellung, was *grundsätzliche Wiederholbarkeit* in pädagogischen Kontexten bedeutet, zu klären. Ein Naturwissenschaftler kann

²⁵ Als Einschränkung dieser Aussage könnte man so genannte „explorative Experimente“ bezeichnen, wie sie z.B. von Flechsig (1979, S. 212ff.) in ihrer Funktion der Hypothesengenerierung oder der Überprüfung des methodischen Inventars diskutiert werden.

z.B. Wasser und Salzwasser immer unter gleichen Bedingungen zum Kochen bringen, um dabei festzustellen, dass der Siedepunkt des Salzwassers höher liegt (bei Normaldruck und 10g Salz auf 1kg Wasser von 100,17° Celsius statt 100° Celsius). Bei Untersuchungen mit Menschen ist die Wiederholbarkeit dagegen schon aus dem Grund nur relativ, weil die meisten Experimente nicht mit den gleichen Personen wiederholbar sind. So kann man sicherlich grundsätzlich experimentell überprüfen, ob sich die phonologische Bewusstheit von Erstklasssschülerinnen und -schülern durch ein bestimmtes Förderprogramm im Vergleich zu einem anderen Unterricht überzufällig verbessert. Eine Wiederholung mit den gleichen Kindern würde hier jedoch völlig neue Bedingungen vorfinden (v.a. dann, wenn die Kinder etwas gelernt haben). Eine Wiederholung des Experiments mit anderen Kindern bedeutet aber, dass andere Individuen untersucht werden und damit keine Wiederholung im eigentlichen Sinn vorliegt (vgl. Klauer, 1973, S. 34f.).

b) In der Grundschulforschung (bzw. allgemein in der Unterrichtsforschung) ist zudem streng genommen eine *isolierende Variation der Bedingungen* ebenfalls nicht zu erreichen, insbesondere dann, wenn die Untersuchung im „Feld“ (also z.B. in den Klassenzimmern der untersuchten Klassen) und nicht in einem künstlichen Laborraum stattfindet. Schon die Tatsache, dass sich z.B. Versuchsgruppen beispielsweise durch die Ausstattung der Klassenzimmer oder durch die vorangegangene Unterrichtsstunde unterscheiden, bedeutet, dass in den unterschiedlichen Gruppen andere Variablen vorhanden sind (auch wenn man im Normalfall hier nicht jedes Mal einen Einfluss dieser Variablen annehmen muss bzw. man durch einen klaren Theoriebezug einen solchen Einfluss auch begründeter Weise ablehnen kann). Eine wirklich isolierte Variation genau einer Bedingung ist nur in einer künstlich geschaffenen Laborsituation möglich. Allerdings erschwert natürlich eine solche Situation die ökologische Validität (als Spezialfall der externen Validität, vgl. z.B. Rost, 2005, S. 92) der Ergebnisse, da nicht gesichert ist, ob sich die hier gefundenen Tatsachen auch in der „normalen“ Lebenssituation so wieder finden lassen würden (vgl. kritisch dazu z.B. auch Brügelmann, 2005, S. 319f.). So ist z.B. zu erwarten, dass Schülerinnen und Schüler in Mitschauräumen der Universität anders reagieren als in ihrem gewohnten Klassenzimmer. Diese Diskussion um Labor- oder Feldexperimente (z.T. auch mit einer anderen Terminologie geführt, wie z.B. „Technologische Experimente“ bei Flehsig, 1979, S. 211) ist weder neu, noch ist sie letztendlich lösbar. Bedenkenswert und bemerkenswert sind für die Unterrichtsforschung dabei Modelle, wie das von Flehsig (1979, S. 210), der vorschlägt, auf der Grundlage der Ergebnisse von Laborexperimenten Feldexperimente durchzuführen, deren Ergebnisse dann für konkrete Unterrichtsvorschläge dienen können.

c) Manche – für pädagogische Fragestellungen häufig wichtige – *Einflussfaktoren lassen sich nicht gezielt variieren* (zumindest nicht ethisch vertretbar), wie z.B. der soziale Status der Eltern, das Alter oder das Geschlecht der Schülerinnen und Schüler sowie ihr Vorwissen oder ihre Intelligenz. Hier ist es nur möglich, auf der Grundlage der bereits bestehenden Variation Unterschiede zu berechnen. Dieses Vorgehen wird häufig als „*Ex-post-facto-Experiment*“ (manchmal auch als „natürliches Experiment“) bezeichnet (vgl. Klauer, 1973, S. 48). Zur Frage, inwieweit dies als „echtes“ Experiment gelten kann, gibt es unterschiedliche Auffassungen (vgl. Klauer, 1973, S. 48). Sicherlich sind jedoch solche Ex-post-facto-Experimente anfälliger für Störvariablen und damit problematischer, wenn es um die Erklärung von Kausalität geht (vgl. Rost, 2005, S. 54).

d) Gerade bei Feldstudien in der Unterrichtsforschung ist es zudem häufig nicht möglich, (bei der Wahrung einer angemessenen Ökonomie) eine echte *Zufallsstichprobe* inklusive Randomisierung zu erreichen. Nur in seltenen Fällen werden die Daten von zufällig ausgewählten Schülerinnen und Schülern verschiedener Klassen erfasst; der Normalfall ist, dass die Untersuchung mit allen Schülerinnen und Schülern der Klasse durchgeführt wird, da ansonsten eine erhebliche Ausweitung der Gesamtstichprobe (und damit auch z.B. der beteiligten Lehrerinnen und Lehrer bzw. Schulen) erforderlich machen würde. Dies gilt in besonderem Maße, wenn bei einer experimentellen Studie durch eine Intervention der Unterricht o.Ä. gezielt variiert wird. Auch können viele Untersuchungen nur dann stattfinden, wenn sich Lehrerinnen und Lehrer freiwillig bereit erklären, an der Studie teilzunehmen und wenn für die Schülerinnen und Schüler eine Erlaubnis ihrer Eltern vorliegt. Dies bedeutet natürlich eine Selektion und damit eine klare Verletzung der Zufälligkeit der Stichprobe.

1.3 Zwischenfazit

Die Auflistung dieser Einschränkung macht deutlich, dass man sich zumindest die Frage stellen muss, inwieweit es in der Grundschulforschung (bzw. allgemeiner, in der Unterrichtsforschung) sinnvoll ist, von Experimenten oder von experimentellen Studien zu sprechen, wenn zentrale Bedingungen nicht erfüllt sind. In Lehrbüchern oder entsprechenden Publikationen finden sich unterschiedlich strenge Auslegungen (vgl. z.B. Wellenreuther, 2007, S. 33ff.). Die Bezeichnungen „Feldexperiment“, „Post-hoc-Experiment“ oder auch „Quasi-Experiment“ zeigen dabei Möglichkeiten auf, wie der Begriff des Experiments auch dann verwendet werden kann, wenn die strengen Bedingungen nicht durchgängig erfüllt sind – allerdings hat gerade die Bezeichnung „Quasi-Experiment“ eine eher herabsetzende Konnotation.

Sicherlich wichtiger als die strenge Einhaltung der Terminologie ist jedoch, dass wir uns in der Grundschulforschung zunächst klar machen, inwieweit die jeweilige Forschungsfrage ein echtes Experiment erfordert (oder überhaupt ermöglicht), ob Einschränkungen eher bei der internen oder der externen Validität der Studie akzeptabel sind, und inwiefern eine Zufallsstichprobe unabdingbar ist bzw. inwieweit Befunde einer selektiv gewonnenen Stichprobe aussagekräftig und/oder verallgemeinerbar sind (vgl. dazu auch Punkt 2.1 dieses Beitrags). Hierzu ist der – oben bereits angedeutete – Bezug auf Referenztheorien auch deshalb wichtig, da er es ermöglicht, Variablen in ihrer Bedeutung einzuschätzen, um so bestimmte Variablen als unwichtig auszuschließen. Damit ist dann ihre Erfassung oder Kontrolle nicht mehr erforderlich.

2. Experimentelle Studien in der Grundschulforschung

2.1 Fragestellungen der Grundschulforschung

In diesem Punkt soll überlegt werden, inwieweit zentrale Fragen, die in der Grundschulpädagogik zu beantworten sind, sinnvollerweise in einem experimentellen Zugriff zu erforschen sind.²⁶ Dabei ist es nicht leicht, die Forschungsfragen der Grundschulforschung zu systematisieren. In Anlehnung an Unterteilungen von Einsiedler (2002) bzw. von Saldern (1998) halte ich in einem heuristischen, nicht überschneidungsfreien Zugriff folgende Forschungslinien fest:

(a) Untersuchungen zur Effektivität des Unterrichts in der Grundschule

Diese Forschungstradition ist ein zentrales Feld jeglicher Lehr-Lernforschung. Für die Grundschulforschung kennzeichnend sind insbesondere Untersuchungen zum Sachunterricht als grundschulspezifisches Fach sowie Studien zum Schriftspracherwerb und zur Erstmathematik. Der Begriff der Effektivität ist breit zu fassen. Fragestellungen nach dem kognitiven Lerngewinn haben hier ebenso ihren Platz wie die Untersuchung des Selbstkonzepts oder der Lernmotivation von Schülerinnen und Schülern.

²⁶ In Anlehnung an das Thema des gesamten Bandes beziehe ich mich im Folgenden nur auf Fragen, die einen empirischen Zugriff implizieren, wobei ich in diesem Fall die historische Grundschulforschung ausklammere, da hier Experimente weder möglich noch intendiert sind. Mir ist natürlich bewusst, dass es zentrale Fragen der Grundschulpädagogik und -didaktik gibt, die sich einem empirischen Zugriff verschließen.

(b) Untersuchungen zu grundschultypischen Lehr- und Erziehungsaufgaben
In Unterscheidung zur eben genannten Forschungslinie sollen hier v.a. Fragestellungen gefasst werden, die über das eigentlich „unterrichtliche“ hinausgehen bzw. nicht direkt an den speziellen Unterricht gekoppelt sind, was z.B. bei Fragen der sozialen Entwicklung, der Gewaltprävention oder der Integration behinderter und nicht behinderter Kinder der Fall sein kann.

(c) Untersuchungen zu den Lernvoraussetzungen der Schülerinnen und Schüler

Auch hier gilt es, die oben genannte Breite aus kognitiven und nicht-kognitiven Lernvoraussetzungen zu beachten, wobei ich in diese Forschungslinie auch die elaborierte Tradition der Forschung zu außerschulischen Lebensbedingungen der Schülerinnen und Schüler – nicht selten z.B. unter dem Begriff der „veränderten Kindheit“ (vgl. z.B. Fölling-Albers, 2001) gefasst – einschließen mag.

Die Forschung zu Lernvoraussetzungen kann als eigenständige Forschungstradition verstanden werden, nicht zuletzt weil die Untersuchungen zumeist einen starken soziologischen oder psychologischen Schwerpunkt haben – und auch von den entsprechenden Kolleginnen und Kollegen bearbeitet werden. Gleichzeitig bilden diese Befunde die Grundlage für Fragestellungen der ATI (Aptitude-Treatment Interaction)-Forschung, inwieweit ein Unterricht für Schülerinnen und Schüler mit unterschiedlichen Lernvoraussetzungen in gleichem Maße geeignet ist.

(d) Untersuchungen zu Lehrerein- und -vorstellungen sowie zum Lehrwissen

Diese Forschungstradition orientiert sich nicht selten am Paradigma der Expertise-Forschung. Gefragt wird in erster Linie, welches Wissen (sowohl fachlich als auch fachdidaktisch u.Ä.) bzw. welche Vorstellungen (z.B. zum Lernen) Lehrkräfte haben, und inwieweit sich dies auf den Unterricht auswirkt. Dieser Forschungslinie sind auch Untersuchungen zuzuordnen, die sich damit beschäftigen, inwieweit solche Dispositionen oder Kognitionen durch Maßnahmen der verschiedenen Phasen der Lehrerbildung verändert werden können.

(e) Untersuchungen zur Organisationsentwicklung der Grundschule

Kern dieser Untersuchungen ist die Frage, inwieweit die Grundschule in ihrer Position im Bildungswesen funktioniert bzw. bestimmte Effekte verursacht. In Deutschland ist hier z.B. die Frage nach dem Übertritt in weiterführende Schulen zu betrachten, der (je nach Bundesland durch Noten oder Empfeh-

lungen) von der Grundschule (bzw. den Grundschullehrkräften) maßgeblich beeinflusst wird.

Zunächst ist festzuhalten, dass in allen Forschungslinien Fragen der Kausalität berührt werden bzw. berührt werden können. Beispiele dafür wären:

- Inwieweit führen bestimmte Unterrichtsmaßnahmen (insbesondere methodischer Art) zu besseren Lernergebnissen? (z.B. Unterscheiden sich die Rechtschreibleistungen von Kindern, die mit einer Fibel oder einer Anlauttabelle unterrichtet worden sind, am Ende der zweiten Jahrgangsstufe?)
- Werden behinderte und nicht behinderte Schülerinnen und Schüler angemessener in Förderklassen oder in Integrationsklassen gefördert?
- Ist das souveräne Beherrschen der deutschen Sprache ein zentraler Indikator für Lernerfolg in der Grundschule? Und als mögliche Anschlussfrage: Hilft ein Unterricht, der durch Gruppenarbeiten die Kommunikation unter den Schülerinnen und Schülern einfordert, Kindern mit unterschiedlichen Deutschkenntnissen in unterschiedlicher Art und Weise?
- Unterrichten Lehrerinnen und Lehrer, die sich (persönlich oder beruflich) für naturwissenschaftliche Themen interessieren, diese Themen anders als Lehrkräfte ohne dieses Interesse? Anschlussfragen dazu könnten sein: Lässt sich dieses berufliche Interesse durch die Teilnahme an einer Lehrerfortbildung verändern? Und: Wirkt sich das berufliche Interesse – bzw. der dadurch beeinflusste Unterricht – auf die Lernleistungen der Schülerinnen und Schüler aus?
- Führen die Selektionsmaßnahmen am Ende der vierten Jahrgangsstufe zu einem veränderten Verhältnis zwischen Lehrkräften und ihren Schülerinnen und Schülern?

Bereits an diesen Beispielen wird deutlich, dass es zwar auf der einen Seite höchst interessant wäre, Untersuchungen so zu gestalten, dass Ursache-Wirkungs-Beziehungen abgesichert werden können, dass aber „echte“ Experimente nur schwer durchzuführen sind. Einige der Fragen sind unter den gegebenen Bedingungen nur post-hoc zu beantworten (z.B. Deutschkenntnisse der Schülerinnen und Schüler, berufliches Interesse der Lehrkräfte oder Selektionsmaßnahmen). Andere Fragen bedürften zumindest einer genauen Spezifizierung. Bei der Frage nach den Rechtschreibleistungen der Schülerinnen und Schüler wäre genau zu spezifizieren, wie z.B. das freie Schreiben mithilfe der Anlauttabelle ein- und durchgeführt wurde, ob spezifische Übungen ergänzend durchgeführt wurden usw. Gleiches gilt für Integrationsmaßnahmen in der Grundschule. Auch hier wären die genauen Bedingungen bzw. die Durchführung zu spezifizieren und zu benennen.

Diese Beispiele zeigen meines Erachtens aber auch, dass eine Orientierung an der Idee des Experiments für die Beantwortung der Fragen wichtig wäre. Eine Antwort bzw. Aussage, die durch die oben gestellten Fragen nahe gelegt wird (wie z.B. Rechtschreibenlernen mit Hilfe einer Anlauttabelle ist im Vergleich zum Fibelunterricht besser/schlechter/gleichwertig) ist aus den eben genannten Gründen nicht intellektuell redlich zu treffen – es sei denn, man kann sich auf die Ergebnisse einer empirischen Studie stützen, in der dies anhand einer ausreichend großen Zufallsstichprobe überprüft wurde.²⁷

Bei den Untersuchungen, die post-hoc zu beantworten sind, ist zumindest zu bedenken und zu berücksichtigen, dass die untersuchten abhängigen Variablen evtl. auf andere zurückzuführen sind bzw. mit ihnen zusammenhängen. So ist eine Korrelation zwischen Deutschkenntnissen und allgemeiner Intelligenz zumindest möglich, und das berufliche Interesse an speziellen Themen könnte auch allgemein mit der Freude am Lehrerdasein zusammenhängen. Einfache Kausalerklärungen sind nur dann zu postulieren, wenn dies bedacht und empirisch berücksichtigt ist.

Fragen nach einfacheren methodischen Konzepten (wie z.B. Gruppenarbeit oder eine spezifisch gestaltete Lehrerfortbildung) lassen sich noch am leichtesten experimentell untersuchen – von möglichen organisatorischen Problemen einmal abgesehen.

2.2 Beispiele

Es ist nicht möglich, in einem solchen Beitrag einen umfassenden Überblick über die relevanten Studien für die Grundschulforschung zu geben.²⁸ Aus diesem Grund beschränke ich mich auf zwei Untersuchungen zur Unterrichtsqualität und damit auf das Feld, in dem experimentelle Studien durch ihr Prinzip am leichtesten und bereits am längsten durchgeführt werden bzw. worden sind. Ziel dieses Teilkapitels ist es, durch Beispiele zu illustrieren, inwieweit wichtige Fragen durch experimentell angelegte Studien bearbeitet

²⁷ Nur erwähnt werden soll, dass eine so allgemeine Aussage selbstverständlich auch nur möglich wäre, wenn verschiedene Methoden des Schreibens mit der Anlauttabelle und verschiedene Formen des fibelgebundenen Unterrichts miteinander verglichen worden wären.

²⁸ Ich mag dies durchaus grundsätzlich positiv sehen. Auch wenn die empirische Forschungssituation in der Grundschulforschung sicherlich noch weiter zu intensivieren ist, so ist sie zumindest doch dem Exotenstatus entwachsen. Obwohl nicht auf die Grundschule eingeschränkt, dafür aber auch nicht mit dem Anspruch auf Vollständigkeit, zeigt z.B. der Umfang des Bandes von Wellenreuther (2007), wie breit das Feld inzwischen ist.

werden können und bearbeitet worden sind. Da beide Studien nicht im strengen Sinne experimentell sind, kann auch illustriert werden, welche Art von Schlussfolgerungen aufgrund solcher Ergebnisse möglich sind, und welche nicht.

Am Institut für Grundschulforschung in Nürnberg wurde ein Trainingsprogramm zur Förderung der phonologischen Bewusstheit konzipiert (Forster & Martschinke 2001), dessen Effektivität durch experimentell angelegte Untersuchungen überprüft werden sollte. Neben den Auswirkungen auf die phonologische Bewusstheit selbst wurden auch Lese- und Rechtschreibleistungen bis zum Ende des zweiten Schuljahrs untersucht (vgl. z.B. Einsiedler, Frank, Kirschhock, Martschinke & Treinies 2002; Kirschhock, Martschinke, Treinies & Einsiedler, 2002).

Insgesamt wurde die Studie mit 375 Kindern aus 15 Schulklassen durchgeführt. Die Forschergruppe verglich die Wirkungen dreier Treatments. Neben einer Gruppe, die nach dem oben genannten Programm plus Fibel unterrichtet wurde, gab es eine Gruppe, in der „entwicklungsorientiert“ (Einsiedler et al., 2002, S. 199) mithilfe einer Anlauttabelle unterrichtet wurde sowie eine Gruppe, in der Fibelunterricht ohne das Trainingsprogramm stattfand. Eine Randomisierung war nicht möglich. Die Klassen konnten nicht zufällig auf die drei Gruppen verteilt werden – vermutlich, da nicht alle Lehrerinnen und Lehrer bereit gewesen wären, ihren Unterricht für die Studie entsprechend umzustellen. Dadurch war es auch nicht möglich, die Gruppen so einzuteilen, dass vor Beginn der Intervention in den Gruppen die gleichen Voraussetzungen bzgl. interessierender Variablen (wie z.B. phonologische Bewusstheit) vorhanden waren.

Zentrale Befunde der Studie waren, dass die Schülerinnen und Schüler, die das Trainingsprogramm absolviert hatten, in verschiedener Hinsicht günstigere Lernergebnisse zeigten – v.a. beim Lesen, weniger beim Schreiben. Die Leistungen der Kinder bezüglich der Lesefertigkeit, aber auch bzgl. des Leseverständnisses waren in der ersten, noch stärker jedoch in der zweiten Jahrgangsstufe in den Trainingsklassen am höchsten (vgl. Kirschhock et al., 2002). Für das erste Halbjahr konnten diese Unterschiede noch nicht gezeigt werden; allerdings stieg in diesem Zeitraum die phonologische Bewusstheit in der Trainingsgruppe am stärksten an, wobei sich diese Unterschiede v.a. für die Gruppe der (zu Beginn der Untersuchung) leistungsschwächsten Schülerinnen und Schüler zeigten (vgl. Einsiedler et al., 2002).

Naturgemäß war dieser letzte Vergleich der Entwicklungen von Kindern mit unterschiedlicher Leistungsfähigkeit nur post-hoc durchführbar. Der Befund, dass insbesondere schwächere Schülerinnen und Schüler profitieren, ist aber für die Akzeptanz des Trainingsprogramms von großer Bedeutung, da die phonologische Bewusstheit als zentrales Element eines erfolgreichen Schriftspracherwerbs gilt (vgl. z.B. Marx, 1997), und man daher hoffen kann, mit Hilfe des Programms weiterreichenden Lese-Rechtschreibschwierigkeiten erfolgreich entgegenzuwirken. Als Konsequenz der Studie kann man Lehrerinnen und Lehrern empfehlen, das untersuchte Trainingsprogramm zu verwenden, insbesondere dann, wenn Schülerinnen und Schüler im Kindergarten diesbezüglich noch nicht oder wenig gefördert worden sind. Für die Unterrichtspraxis ist diese Studie damit von großer Relevanz. Bevor die Ergebnisse vorlagen, gab es begründete Zweifel, ob eine Förderung der phonologischen Bewusstheit in der Schule nicht schon zu spät käme.

Manche – eher wissenschaftlich als unterrichtspraktisch interessierende – Fragen können durch die Studie jedoch nicht beantwortet werden, auch da die Studie nicht streng experimentell durchgeführt werden konnte.²⁹ So muss z.B. offen bleiben, inwieweit einzelne Elemente des Programms wichtiger als andere sind. Streng gesehen ist nicht einmal erwiesen, ob das Trainingsprogramm selbst zu den positiven Effekten führt oder ob z.B. durch das Programm die Aufmerksamkeit der Lehrerinnen stärker auf die phonologische Bewusstheit gelenkt wurde, sodass sie auch andere Teile des Unterrichts diesbezüglich umstrukturierten oder den Schülerinnen und Schülern andere Rückmeldungen gaben. Auch kann nur vermutet, nicht jedoch empirisch geklärt werden, weshalb die Effekte für das Lesenlernen erst gegen Ende der ersten Jahrgangsstufe deutlich werden.

Am Seminar für Didaktik des Sachunterrichts der Universität Münster wurden verschiedene Studien zum naturwissenschaftlichen Lernen in der Grundschule durchgeführt – zum Teil in enger Kooperation mit dem Forschungsbereich Erziehungswissenschaft und Bildungssysteme des Max-Planck-Instituts für Bildungsforschung in Berlin. Auch diese Studien sind experimentell angelegt, ohne im strengen Sinn experimentell zu sein.

Ziel einer der Studien war es zu überprüfen, inwieweit sich Strukturierungselemente in einem konstruktivistisch ausgerichteten Unterricht auf Lerner-

²⁹ Um nicht missverstanden zu werden: Die Beantwortung dieser Fragen wurde von der Forschergruppe nicht angezielt oder gar behauptet. Es geht hier um die Darstellung, was diese Art von Studien leisten kann und was nicht.

gebnisse der Schülerinnen und Schüler auswirken (vgl. z.B. Blumberg, Möller, Jonen & Hardy, 2003; Jonen, Möller & Hardy, 2003; Möller, Hardy, Jonen, Kleickmann & Blumberg, 2006). Verglichen wurden zwei Treatments. In beiden Gruppen erhielten die Schülerinnen und Schüler vielfältige Materialien und Anregungen, um ihre kognitiven Konzepte zum Thema „Schwimmen und Sinken“ möglichst aktiv und konstruktiv weiterzuentwickeln. Dabei wurde darauf geachtet, dass kooperative Lernprozesse möglich waren und dass das Gelernte regelmäßig reflektiert wurde. Diese Grundprinzipien waren für beide Experimentalbedingungen gleich (vgl. Jonen et al., 2003). Identisch war auch die Dauer der Intervention; beide Unterrichtseinheiten erstreckten sich über acht Doppelstunden.

Die Gruppen unterschieden sich jedoch darin, inwieweit die Lehrerin – in beiden Gruppen unterrichtete die gleiche Lehrerin – Strukturierungsmaßnahmen vornahm. In der strukturierten Gruppe wurde der Unterricht anhand verschiedener Teilfragen sequenziert, „die zwar nicht kleinschrittig und losgelöst von dem komplexen Problem erarbeitet werden sollten, aber orientiert an den Vorstellungen der Schülerinnen und Schüler und ihren Lernschwierigkeiten Teilaspekte des komplexen Problems herausgriffen“ (Jonen et al., 2003, S. 101). Des Weiteren variierten die verbalen Anweisungen und Rückmeldungen der Lehrerin, indem sie in dieser Gruppe häufiger Rückmeldungen an die Klasse gab und häufiger auf Widersprüche u.ä. hinwies sowie regelmäßig Begründungen und Zusammenfassungen einforderte (vgl. Jonen et al., 2003, S. 102). Insgesamt nahmen 190 Schülerinnen und Schüler der dritten Jahrgangsstufe an der Untersuchung teil. Jede der Gruppen bestand aus drei Klassen. Daneben wurden noch zwei Klassen als Kontrollgruppe untersucht, die während des Untersuchungszeitraumes keinen Unterricht zum untersuchten Thema „Schwimmen und Sinken“ erhalten hatte.

Die Ergebnisse ähneln in mancherlei Hinsicht der oben beschriebenen Untersuchung. Wichtige Effekte zwischen den beiden Gruppen traten nicht direkt im Anschluss an die Intervention auf, sondern erst in einem Nachtest, der ein Jahr später durchgeführt wurde (vgl. Möller et al., S. 172f.) – mit höheren Lernleistungen bzgl. des konzeptuellen Verständnisses für die Gruppe, in der mehr Strukturierungsmaßnahmen durchgeführt worden waren. Und auch hier zeigte sich, dass die Unterschiede besonders groß bei den Kindern waren, die in der Eingangsuntersuchung besonders geringes Vorwissen hatten (vgl. Möller et al., 2006, S. 173; Blumberg et al., 2003). Der gleiche Befund ergab sich auch für nicht-kognitive Variablen wie Selbstkonzept, Kompetenzerfinden und intrinsische Motivation.

Die Befunde kann man so interpretieren, dass die Strukturierungsmaßnahmen zu diesen positiven Effekten führten – wenngleich nicht festzustellen ist, ob nur bestimmte Maßnahmen oder genau die Kombination der verschiedenen Formen von Strukturierung relevant sind. Dennoch kann man auch hier eine Konsequenz für Lehrerinnen und Lehrern ziehen; in diesem Fall ist es die Anregung, naturwissenschaftlichen Unterricht in einem konstruktivistischen Sinne aufzubauen, ohne die erforderliche Strukturierung des Unterrichts zu vernachlässigen. Da es erforderlich war, für die Studie Unterricht zu planen und durchzuführen, kann die untersuchte Unterrichtssequenz als Orientierungsbeispiel dienen. Die Forschergruppe hat – auf der Grundlage der Untersuchungen – dann auch verschiedene Materialkästen für den Einsatz in der Grundschule entwickelt.

3. Fazit

Die Beispiele machen meiner Ansicht nach Verschiedenes deutlich. Zum einen sind die Befunde für die Unterrichtspraxis verwertbar. Die Überprüfung der Theorie bzw. des erstellten Unterrichtsmaterials durch eine vergleichende Untersuchung führt in beiden Fällen dazu, dass realisierbare Unterrichtsempfehlungen gegeben werden können, auch wenn manche Kausalerklärungen nicht möglich waren. Dafür ist wichtig, dass durch die Untersuchung im Feld eine hohe ökologische Validität erreicht wurde, die es ermöglicht, Unterrichtsempfehlungen zu geben, die sehr eng am untersuchten Treatment angelehnt sind. Bei einer Übersetzung von vergleichsweise künstlichen Laborsituationen wäre immer einzufordern, dass auch diese Übersetzung überprüft würde. Zudem zeigen die Befunde in beiden Untersuchungen, dass es in einer Gruppe anscheinend gelungen ist, die Schülerinnen und Schüler mit ungünstigen Lernvoraussetzungen besonders zu fördern – auch dieser Befund konnte nicht streng experimentell erhoben werden.

Damit entzogen sich zwei der für Umsetzungsempfehlungen und damit für die unterrichtspraktische Bedeutsamkeit zentrale Aspekte – hohe ökologische Validität und Betrachtung spezieller (nicht gezielt und intentional variierbarer) Schülervariablen – einer echten experimentellen Überprüfung. Aus diesem Grund bin ich der Überzeugung, dass eine *ausschließliche bzw. vorrangige* Orientierung an der strengen Vorstellung experimenteller Forschung für zentrale Ziele und Fragestellung einer – letztlich präskriptiv ausgerichteten – Unterrichtsforschung abzulehnen ist; sie würde dazu führen, dass zentrale unterrichtliche Fragestellungen nicht oder nur wenig aussagekräftig bearbeitet werden können. Die Befürchtung, dass dann eine „Innovation ohne Forschung“ (Reinmann, 2005, S. 52) geschieht, da komplexere Erneuerungen

von Unterricht nicht mehr überprüfbar sind, ist real. Der Hinweis von Valtin (2000), dass empirische Befunde so selten genutzt werden, kann unter anderem auch durch diese Kluft erklärt werden – eine „didaktische Entwicklungsforschung“ (Kahlert, 2006, S. 850), die auf empirischen Grundgedanken und Ergebnissen beruht, die jedoch offen für erforderliche Variationen ist, kann helfen, diese Kluft etwas zu überwinden (vgl. auch Kahlert, 2007).

Obwohl – oder vielleicht gerade weil – ich mich dagegen ausspreche, dass sich die Grundschulforschung vorrangig an strengen experimentellen Designs ausrichtet, scheinen mir verschiedene Hinweise angebracht:

- (a) Unabhängig von Detailfragen in der strengen Auslegung ist der experimentelle Ansatz, der theoriegeleitet verschiedene Formen des Unterrichts – wenn möglich unter Berücksichtigung von Schülerinnen und Schülern mit unterschiedlichen Lernvoraussetzungen – *empirisch vergleicht*, ein zentraler Ansatz der Grundschulforschung; für die Frage nach der Effektivität von Grundschulunterricht sicherlich sogar der entscheidende.
- (b) Zum zweiten sollten insbesondere Untersuchungen, die z.B. aufgrund einer fehlenden Randomisierung durch verschiedene Störvariablen beeinflusst sein könnten, intensiver zu *Replikationsstudien* führen – Studien, die durchaus auch in einfacheren akademischen Abschlussarbeiten durchgeführt werden könnten. Gerade in pädagogischen Kontexten, in denen auch bei einer zufällig ausgewählten und randomisierten Stichprobe nicht erwartet werden kann, dass diese z.B. die Schülerschaft in 20 Jahren repräsentiert, ist dieser Ansatz auf alle Fälle sinnvoll.
- (c) Auch wenn eine Konfundierung von Variablen – wie oben beschrieben – immer wieder erforderlich ist, um praxisrelevante Vorschläge zu überprüfen, so sollten dennoch Untersuchungen gleichgewichtig betrieben werden, die möglichst gezielt *einzelne Variablen ohne Konfundierung* isolieren. Ansonsten kann immer nur ein „Gesamtpaket“ der Innovation als überprüft gelten. Variationen und Modifikationen (die für die Passung an unterschiedliche Klassensituationen ja immer erforderlich sind) wären dann nur schwer begründbar.
- (d) Spätestens bei der Interpretation der „nicht-streng-experimentell“ gewonnenen Befunde sind *Einschränkungen deutlich zu machen* und offen zu legen. Dies betrifft insbesondere die Interpretation mit Blick auf Ursache-Wirkungs-Beziehungen. Gerade bei Felduntersuchungen müssen wir im Auge behalten, dass es schwieriger ist, die gewonnenen Ergebnisse präzise auf die Intervention zu beziehen und Störvariablen auszuschalten. Hier ist es sicherlich erforderlich, die Daten manchmal etwas vorsichtiger zu deuten, und es ist insbesondere darauf zu verzichten, aus den Befunden theoretische Modelle mit allgemeinem Gültigkeitsanspruch zu entwickeln.

(e) Auf alle Fälle sind Forscherverbünde zu unterstützen, die, wie es z.B. Flechsig (1979) bereits vorgeschlagen hat, eine *Kombination aus (echten) Laborexperimenten und – darauf aufbauend – „Quasi-Experimenten“* im Feld durchführen. Wenn die Ergebnisse in die gleiche Richtung gehen, so kann dies eine recht valide und aussagekräftige Befundlage ergeben; eine Befundlage, die sowohl auf theoretischer Ebene zur Bestätigung von Ursache-Wirkungs-Beziehungen als auch zum Erstellen von umsetzbaren Unterrichtsvorschlägen oder -empfehlungen führen kann.

Literatur

- Blumberg, E., Möller, K., Jonen, A. & Hardy, I. (2003). Multikriteriale Zielerreichung im naturwissenschaftlichen Sachunterricht der Grundschule. In D. Cech & H.J. Schwier (Hrsg.), *Lernwege und Aneignungsformen im Sachunterricht* (S. 77-92). Bad Heilbrunn: Klinkhardt.
- Brügelmann, H. (2005). *Schule verstehen und gestalten*. Lengwil: Libelle.
- Einsiedler, W. (2002). Empirische Forschung zum Sachunterricht – ein Überblick. In K. Spreckelsen, K. Möller & A. Hartinger (Hrsg.), *Ansätze und Methoden empirischer Forschung zum Sachunterricht* (S. 17-38). Bad Heilbrunn: Klinkhardt.
- Einsiedler, W., Frank, A., Kirschhock, E.-M., Martschinke, S. & Treinies, G. (2002). Der Einfluss verschiedener Unterrichtsmethoden auf die phonologische Bewusstheit sowie auf Lese- und Rechtschreibleistungen im 1. Schuljahr. *Psychologie in Erziehung und Unterricht*, 49, 194-209.
- Flechsig, K.-H. (1979). Die Funktion des Experiments in der Unterrichtsforschung. In: N. Kluge & H. Reichel (Hrsg.), *Das Experiment in der Erziehungswissenschaft* (S. 200-225). Darmstadt: Wissenschaftliche Buchgesellschaft.
- Fölling-Albers, M. (2001). Veränderte Kindheit – revisited. Konzepte und Ergebnisse sozialwissenschaftlicher Kindheitsforschung der vergangenen 20 Jahre. In M. Fölling-Albers, S. Richter, H. Brügelmann & A. Speck-Hamdan (Hrsg.), *Jahrbuch Grundschule III. Fragen der Praxis – Befunde der Forschung* (S. 10-51). Frankfurt: Grundschriftverband.
- Forster, M. & Martschinke, S. (2001). *Leichter lesen und schreiben lernen mit der Hexe Susi*. Donauwörth: Auer.
- Hellmich, F. & Kiper, H. (2006). *Einführung in die Grundschuldidaktik*. Weinheim und Basel: Beltz.
- Huber, O. (2005). *Das psychologische Experiment: Eine Einführung* (4. Aufl.). Bern: Verlag Hans Huber.
- Jonen, A., Möller, K. & Hardy, I. (2003). Lernen als Veränderung von Konzepten – am Beispiel einer Untersuchung zum naturwissenschaftlichen Lernen in der Grundschule. In D. Cech & H.J. Schwier (Hrsg.), *Lernwege und Aneignungsformen im Sachunterricht* (S. 93-108). Bad Heilbrunn: Klinkhardt.
- Kahlert, J. (2006). Zwischen den Stühlen zweier Referenzsysteme. Zum Umgang mit heterogenen Erwartungen bei der Evaluation schulnaher Disziplinen in Lehramtsstudiengängen. *Zeitschrift für Pädagogik*, 51, 840-855.
- Kahlert, J. (2007). Was kommt nach der Erkenntnis? Zum schwierigen Verhältnis pädagogischer Disziplinen zu der Erwartung, sich nützlich zu machen. In G. Reinmann & J. Kahlert (Hrsg.), *Der Nutzen wird vertagt ... Bildungswissenschaften im Spannungsfeld zwischen wissenschaftlicher Profilbildung und praktischem Mehrwert* (S. 20-45). Lengerich: Pabst.

- Kirschhock, E.-M., Martschinke, S., Treinies, G. & Einsiedler, W. (2002). Vergleich von Unterrichtsmethoden zum Schriftspracherwerb mit Ergebnissen zum Lesen und Rechtschreiben im 1. und 2. Schuljahr. *Empirische Pädagogik*, 16, 433-459.
- Klauer, K.J. (1973). *Das Experiment in der pädagogischen Forschung*. Düsseldorf: Pädagogischer Verlag Schwann.
- Lay, W.A. (1905). *Experimentelle Didaktik. Ihre Grundlegung mit besonderer Rücksicht auf Muskelsinn, Wille und Tat* (2. Aufl.). Leipzig: Otto Nemnich Verlag.
- Marx, H. (1997). Erwerb des Lesens und des Rechtschreibens: Literaturüberblick. In F.E. Weinert & A. Helmke (1997), *Entwicklung im Grundschulalter* (S. 85-111). Weinheim: Beltz/Psychologie Verlags Union.
- Möller, K., Hardy, I., Jöns, A., Kleickmann, T. & Blumberg, E. (2006). Naturwissenschaften in der Primarstufe. Zur Förderung konzeptionellen Verständnisses durch Unterricht und zur Wirksamkeit von Lehrerfortbildungen. In M. Prenzel & L. Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule* (S. 161-193). Münster: Waxmann.
- Reinmann, G. (2005). Innovation ohne Forschung? Ein Plädoyer für den Design-Based Research-Ansatz in der Lehr- Lernforschung. *Unterrichtswissenschaft*, 33, 52-69.
- Rost, D.H. (2005). *Interpretation und Bewertung pädagogisch-psychologischer Studien*. Weinheim: Beltz.
- von Saldern, M. (1998). Die Aufgabenfülle der Grundschule und ihrer Pädagogik. *Zeitschrift für Pädagogik*, 44, 907-924.
- Valtin, R. (2000). Grundschulpädagogik als empirische Forschungsdisziplin. *Zeitschrift für Pädagogik*, 46, 555-570.
- Wellenreuther, M. (2007). *Lehren und Lernen – aber wie? Empirisch-experimentelle Forschungen zum Lehren und Lernen im Unterricht* (3. Aufl.). Baltmannsweiler: Schneider Hohengehren.