

Hermann G. Ebner

Wissen schaffen: What you do is what you get.

I

Mit der Überschrift zu diesem Beitrag mache ich eine Anleihe bei der Formulierung des einigen Anwendungsprogrammen mitgegebenen Versprechens „What you see is what you get“ (WYSIWYG). Während Abweichungen von dem Erscheinungsbild am Monitor und dem am Papier schnell auffallen, fehlt bei der Generierung von Daten im Forschungsprozess zumeist eine synchrone Vergleichsmöglichkeit. Die Zuordnung der untersuchten Phänomene durch die Form, in der wir uns mit ihnen beschäftigen, die Überformung der Fragen, Feststellungen und Folgerungen durch die eingesetzten Untersuchungsinstrumente sind häufig deutlich weniger offenbar, denn sie vermitteln sich über das Ergebnis, sie sind ein nicht abtrennbarer Bestandteil des Untersuchungsbefundes. Wird nun die Interpretation der aus den Daten gewonnenen Information nicht strikt auf der Grundlage des Leistungspotenzials des Inputs vorgenommen, dann besteht die Gefahr, das Erhaltene falsch einzuschätzen und unzutreffende Schlüsse zu ziehen.

Ein instruktiver Fall wurde 2010 in einer Ausgabe der Frankfurter Allgemeinen Sonntagszeitung beschrieben: Auf seiner Entdeckungsreise (1520/21) „soll Magellan mit einem 700 Meter langen Seil, an dessen Ende eine Eisenkugel angebracht war, eine Messung der Meerestiefe im Pazifik angeordnet haben. Weil die Kugel keinen Grund fand, erklärte Magellan den Ozean für bodenlos“. Magellan hat alles richtig gemacht: Für das Problem hatte er die angemessene Methode gewählt (Messung), und er hatte dafür ein prinzipiell geeignetes Instrument (Seil mit Beschwerung) eingesetzt. Er erhält ein Ergebnis und begeht nun aber den entscheidenden Fehler: Das Ergebnis ist: „Es wurde kein Grund gefunden“. Magellans Folgerung, „der Ozean ist bodenlos“, wird durch das Messergebnis nicht gedeckt, sie ist logisch unzulässig. Es kann eine Reihe möglicher Gründe angeführt werden, die dazu geführt haben könnten, dass kein Grund gefunden wurde: Der offensichtlichste ist selbstverständlich, dass mit dem verwendeten Seil lediglich Tiefen bis 700 Meter festgestellt werden können. Daneben könnten aber z. B. auch starke Strömungen verhindert haben, dass die Kugel – auch bei geringeren Tiefen – den Meeresboden erreicht. Jedenfalls steht allein fest, dass kein Grund gefunden wurde, alles Weitere ist Spekulation.

Qualitätssicherung in der Forschung hat vor allem damit zu tun, dass in Bezug auf die zu untersuchenden Phänomene das Spektrum möglicher Interpretationen umschrieben wird. Entsprechend bedeutet Qualitätssicherung, dass aufgrund der Kenntnis dieser Risiken Anstrengungen unternommen werden, sie zu kontrollieren.

Vor einigen Jahrzehnten wurde der sogenannte Positivismusstreit weitgehend fern der Forschungspraxis ausgetragen, und diese Auseinandersetzungen lösten zwar einige Irritationen aus, entfalteten darüber hinaus jedoch m. E. kaum substanzielle forschungspraktische Relevanz. Heute sind – auch wenn vereinzelt noch immer Argumentationsfiguren bei Protagonisten aus jener Zeit auftreten – zentrale methodologische Diskussionen in unserem Fachgebiet eher an Fragen der Ausgestaltung des Forschungsprozesses zur theoretisch fundierten Lösung praktischer Aufgaben orientiert. Zu solch praktischen Aufgaben gehört die Gestaltung effektiver Lernumgebungen. Es ist Richard E. MAYER (2010), der in seinem Beitrag „Seeking a science of instruction“ die aktuelle Auffassung von der Wahrnehmung dieser Aufgabe und dem Beitrag der Forschung prägnant formuliert: „The design of instruction should be guided by a research-based theory of how instruction works – that is, a theory of how instructional methods affect learning processes. This is the central premise in the science of instruction.“

Damit skizziert MAYER ein Programm, dessen Relevanz nicht auf den engeren Bereich der Instruktionsforschung begrenzt ist. Würde die Berufsbildungsforschung daran orientiert, dann hieße das für sie, dass sie u. a. das Material für eine auf Forschungsergebnissen gegründete Konzept-/Maßnahmenentwicklung zu liefern hätte. Im Zusammenhang mit der Generierung dieses Wissens, dessen Aufbereitung und Bereitstellung, z. B. in Form von Prinzipien (beispielhaft dafür: MERRILL 2002), sowie dessen Verwendung stellen sich zwei Fragen: Erstens, was gilt es zu beachten, um möglichst verlässliche Forschungsergebnisse zu erhalten, und zweitens, anhand welcher Kriterien kann die Verlässlichkeit berichteter Befunde beurteilt werden. In beiden Fällen geht es um „Qualitätssicherung“: im ersten aus „Produzenten-“ und im zweiten aus „Verwertersicht“.

Im Kontext dieser Fragen besteht allerdings die Gefahr – für manche Akteure vielleicht auch die Verlockung –, den wenig fruchtbaren Disput aus früheren Jahrzehnten wieder aufzunehmen, denn das Thema „Verlässlichkeit“ scheint beinahe zwingend in die Unterscheidung zwischen „guten“ und „schlechten“ Forschungsmethoden sowie zwischen anerkannten und „unsauberen“ Forschungskonzepten zu führen: Etwas zugespitzt sehen sich wohl einige der beteiligten Fraktionen vor dem Problem, wer die „Kodifizierung von Evidenz“ dominiert. Dabei fällt auf, dass insbesondere die in den USA im letzten Jahrzehnt des 20. Jahrhunderts aufgekommene Auseinandersetzung eher von den Vertretern des „streng“ kontrollierten Experiments auszugehen schien: Aus deren Sicht liefern „Design Experiments“, „Design-based Research“ oder „Development(al) Research“ keine verlässliche Information.

Die Diskussion über Konzepte, Methoden und Instrumente bleibt aber häufig auf der formalen, inhaltsfreien Ebene. Zusammenhänge mit der Art der Fragestellung, also der Information, die gewonnen werden soll, und der intendierten Verwer-

tung werden zumeist nicht expliziert. Da m. E. die Frage nach der Verlässlichkeit der im Forschungsprozess generierten Information sinnvoll ist, wird im Folgenden zunächst skizziert, warum Verlässlichkeit für wichtig gehalten, und anschließend, wie hier die Beziehung von Verlässlichkeit, Methode und Fragestellung modelliert wird.

II

In Bezug auf die Wichtigkeit kann beispielhaft die Argumentation Robert COES (1999) aus seinem Papier zu evidenzbasierten Entscheidungen im Bildungsbereich herangezogen werden. COE weist darauf hin, dass „evidenzbasiert“ inzwischen den Status eines Modeworts habe und daher die Gefahr bestehe, dass zunehmend verloren gehe, was dessen Bedeutung in diesem Kontext sei, bzw. welches Programm damit vertreten werde. Nicht zuletzt deshalb bemüht er sich in seinem Plädoyer um die Klärung dieses Konzepts. Für COE (1999, S. 1) wird mit „Evidence-based Education“ ein Ansatz gekennzeichnet, „which argues that policy and practice should be capable of being justified in terms of sound evidence about their likely effects. Education may not be an exact science, but it is too important to allow it to be determined by unfounded opinion, whether of politicians, teachers, researchers or anyone else“. Ähnlich argumentiert auch Robert E. SLAVIN (2003, S. 27): „If one believes, as I do, that research in education has an obligation to answer the ‚what works‘ questions that educators, parents and policymakers ask, then our job is to produce answers that are well justified.“ Und COE (1999, S. 1) ergänzt: „In the UK, and no doubt elsewhere, too many policies have been imposed on schools without adequate evidence about their likely effects and costs. It is arguably a waste of public money and professionals’ time to impose policies without good evidence that they will lead to an improvement over what previously existed“. Wenn also der Nachweis nicht geführt wird, oder wenn nicht belegt werden kann, dass mit der Implementierung einer vorgeschlagenen Maßnahme die intendierten bzw. behaupteten Effekte erzielt werden, dann ist zumindest nicht auszuschließen, dass mit der Umsetzung der Maßnahme Verluste oder Belastungen entstehen, die hätten vermieden werden können und daher nicht zu rechtfertigen sind.

Um Entscheidungen evidenzbasiert treffen zu können (vgl. dazu auch SLAVIN 2008), müssen die entsprechenden Belege, die Informationen, das relevante Wissen vorliegen, und das in verlässlicher Form. An dieser Stelle greife ich die Überschrift zu diesem Beitrag wieder auf: „Wissen schaffen: What you do is what you get“ – es geht darum, zu zeigen, dass es der präzisen Entwicklung bzw. Analyse der Untersuchungsanlage bedarf, um abschätzen zu können, welcher Art die Ergebnisse sind, die man erhalten wird. Mit den folgenden Ausführungen erläutere ich anhand von Beispielen diese Aussage und den bereits oben erwähnten Zusammenhang von Fra-

gestellung, Informationsgewinnung und Verwertung bzw. Interpretation der Ergebnisse. Dabei wird unterschieden zwischen

1. der Intention, Informationen zu einem einzelnen Phänomen zu erhalten und
2. der Intention, die Wirksamkeit und die Wirkungen einer (neuen) Maßnahme zu beurteilen, eine Intervention, ein Programm zu evaluieren.

Nehmen wir – für den Fall (a) – an, in einem Landkreis komme die Frage auf, ob an einem Standort ein weiteres Berufliches Gymnasium eines im Landkreis bislang noch nicht vertretenen Typs eingerichtet werden soll. Im Kreis sind an mehreren Standorten bereits Berufliche Gymnasien aller anderen Fachrichtungen vorhanden. Wenn ein nachfrageorientiertes Konzept verfolgt wird, dann ist zu klären, ob eine hinreichende Nachfrage nach dem zusätzlichen Bildungsangebot besteht. Diese Information lässt sich gewinnen, indem die potenziellen Nachfrager, d. h. die Jugendlichen und deren Eltern, dazu befragt oder/und indem Statistiken vergleichbarer Fälle in anderen Landkreisen ausgewertet werden. Soweit eine Befragung durchgeführt werden soll, sind bei der Entwicklung des Designs der Studie einige Probleme zu lösen, die für die Verwertung der generierten Information entscheidend sind. Mit Blick auf das Interesse, verlässliche Daten zu erhalten, sind vorab u. a. folgende Fragen zu beantworten:

- Auf welche Weise werden die zu befragenden Personen ausgewählt (z. B. Vollerhebung, Ziehung einer Stichprobe, exemplarische Erhebung an einer Schule)?
- In welcher Form wird die Befragung durchgeführt (z. B. schriftlich via Postversand, online, Telefoninterview)?
- Ist sichergestellt, dass die zu befragenden Personen über den betreffenden Schultyp hinreichend Bescheid wissen (z. B. Informationsveranstaltungen, Informationsblatt, Erläuterungen auf dem Fragebogen)?

Bei diesen elementaren Dimensionen der Qualitätssicherung sind die Lösungen vergleichsweise eindeutig, wenn auch nicht immer einfach zu realisieren. So ist unstrittig, dass – in Abhängigkeit von der Größe der Grundgesamtheit – allein die Vollerhebung bzw. die Befragung einer Zufallsstichprobe die Gewinnung verlässlicher Informationen ermöglicht. Hinsichtlich des zweiten Punktes – der Form der Befragung – ist der Weg zu wählen, der unter den gegebenen Umständen die Chance auf die größtmögliche Rücklaufquote bietet. Und ebenso ist dafür Sorge zu tragen, dass die jeweiligen Befragten entsprechend informiert sind. Analog sind bei der Auswertung statistischen Materials aus anderen Landkreisen Prüfkriterien anzulegen, um vor allem sicherzustellen, dass ausreichend affine Kontextbedingungen vorliegen bzw. um Differenzen in der Ausgangslage präzise zu kennzeichnen.

Bei dieser vergleichsweise einfachen Problemstellung ist es auch nicht schwierig zu sehen, in welcher Weise die Vorgehensweise das Ergebnis beeinflusst. Abgesehen von offensichtlichen „handwerklichen“ Fehlern bei der Auswahl der zu befragenden Personen, lassen sich mit Blick auf die praktische Relevanz des Datenmaterials weitere Gefährdungspotenziale identifizieren; dazu gehört z. B. die ungenügende Berücksichtigung von Interdependenzen. In unserem Beispiel wäre u. a. die Änderung der Ausschöpfungsquote bezogen auf die fragliche Alterskohorte abzuschätzen und ebenso die möglichen „Kannibalisierungseffekte“ in Bezug auf die bestehenden Bildungsangebote. Wie wahrscheinlich bzw. wie groß diese Art von Gefährdungen der Verlässlichkeit ist, hängt ab von der Professionalität der Durchführenden bzw. von den Verfahren, die eingesetzt werden, um solche Risiken zu minimieren (wie z. B. Validierung des Indikatorenkatalogs im Rahmen von Expertengesprächen). Auf Gefährdungspotenziale, die in der interessengeleiteten Generierung, Verarbeitung und Interpretation der Daten liegen, wird hier nicht eingegangen, da mit ihnen elementare Regeln wissenschaftlichen Arbeitens verletzt werden.

Der Fall (b) – Evaluation einer Intervention oder eines Programms – ist im Vergleich zum Fall (a) mit deutlich größeren forschungsmethodologischen Herausforderungen verbunden: Die Frage ist: Unter welchen Bedingungen können im Interventionsfeld eruierte Phänomene als Effekte der Intervention gelten? Im Zusammenhang mit dieser Problemstellung sind in den letzten Jahrzehnten vielfältige Kontroversen ausgetragen worden. In Deutschland erwies sich zunächst die Konzeption der sogenannten Modellversuche bzw. Modellversuchsprogramme und der mit ihnen verbundenen „Wissenschaftlichen Begleitung“ als ein Kristallisationspunkt solcher Auseinandersetzungen (vgl. dazu die Beiträge in dem vom BUNDESINSTITUT FÜR BERUFSBILDUNG 1978 herausgegebenen Band). Die neueren Kontroversen werden vor allem zwischen den Vertretern des experimentellen Designs und jenen des Design-based Research oder der Grounded Theory ausgetragen: Es geht u. a. um die Strenge der Regeln bei der Generierung der Daten und auch um das Verständnis von „Theorie“.

Im Unterschied zu dieser vor allem auf der Konzept-Ebene geführten Diskussion steht z. B. für das NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE (2003), für Ann L. BROWN (1992) oder für Richard J. SHAVELSON u. a. (2003) die Beziehung von Untersuchungsanlage und zu gewinnender Information im Vordergrund. Informativ sind m. E. insbesondere die Ausführungen von Ann L. BROWN (1992, S. 143); sie kommt ursprünglich von der experimentellen Forschungslinie, beschäftigte sich jedoch zunehmend mit Design-based Research (sie verwendete die Bezeichnung „Design Experiments“). BROWN schreibt: „Another critical tension in our goals is that between contributing to a theory of learning, a theoretical aim that has always been a keystone of our work, and contributing to practice. This is intervention research designed to inform practice. For this to be true, we must operate al-

ways under the constraint that effective intervention should be able to migrate from our experimental classroom to average classrooms operated by and for average students and teachers, supported by realistic technological and personal support. (...) My Change in focus was a gradual evolution rather than an unpremeditated leap to instruction (...). Even though the research setting has changed dramatically, my goal remains the same: to work toward a theoretical model of learning and instruction rooted in a firm empirical base.“

Die von Brown markierten Aspekte sind beide in dem Leitfaden des National Center for Education Evaluation and Regional Assistance (NCEERA) berücksichtigt: Um die Effektivität einer Intervention oder eines Programms mit dem Gütesiegel „gut belegt“ zu versehen, erwarten die Verfasser, dass (a) die kausale Beziehung zwischen Maßnahme und Effekten zunächst im Rahmen streng kontrollierter experimenteller Studien nachgewiesen und (b) im Anschluss daran die Maßnahme mehrfach unter typischen Bedingungen des Einsatzbereichs (z. B. in Schulorganisation, Klassenzimmer) erfolgreich umgesetzt wurde. Bevor also eine Maßnahme zur Übernahme in den Regelbetrieb empfohlen werden kann, soll geprüft werden, ob es für die Wirksamkeit der Maßnahme einen hinreichenden Beleg gibt. Dieser Beleg gilt als erbracht, wenn bei der Analyse der Studien festgestellt wird, dass bei der Effektivitätsprüfung der Intervention den im Leitfaden beschriebenen Kriterien entsprechend vorgegangen wurde.

Tabelle: Die Kriterien des Leitfadens des NCEERA (2003, S. V) im Überblick

Step 1. Is the intervention backed by „strong“ evidence of effectiveness?		
Quality of studies needed to establish „strong“ evidence:	Quantity of evidence needed:	
<ul style="list-style-type: none"> • Randomized controlled trials that are well-designed and implemented 	+	Trials showing effectiveness in <ul style="list-style-type: none"> • two or more typical school settings, • including a setting similar to that of your schools/Classrooms
		= „Strong“ Evidence
Step 2. If the intervention is not backed by „strong“ evidence, is it backed by „possible“ evidence of effectiveness?		
Types of studies that can comprise „possible“ evidence:		Types of studies that do not comprise „possible“ evidence:
<ul style="list-style-type: none"> • Randomized controlled trials whose quality/quantity are good but fall short of „strong“ evidence; and/or • Comparison-group studies in which the intervention and comparison groups are very closely matched in academic achievement, demographics and other characteristics. 		<ul style="list-style-type: none"> • Pre-post studies • Comparison-group studies in which the intervention and comparison groups are not closely matched • „Meta-analyses“ that include the results of such lower-quality studies
Step 3. If the answers to both questions above are „no“, one may conclude that the intervention is not supported by meaningful evidence.		

Hilfreich sind mit Blick auf die Qualitätssicherung von Initiativen in der Berufsbildungsforschung die Hinweise auf die eingeschränkte Tauglichkeit bestimmter Untersuchungsdesigns, um Belege für die Wirksamkeit einer Intervention beibringen zu können. So werden Untersuchungsanlagen, in denen Daten zu zwei Zeitpunkten erhoben werden – vor der Intervention und danach –, zwar häufig als zweckmäßig angesehen, um die Wirksamkeit einer Intervention festzustellen. Tatsächlich erfüllen sie diese Funktion jedoch nicht oder nur bedingt. Gleiches gilt auch für Designs mit Vergleichsgruppen, bei denen die Gruppenzuteilung der Probanden nicht randomisiert erfolgt. Im ersteren Fall gelingt es nicht, etwaige Differenzen in den Pre- und Posttestergebnissen kausal der Intervention zuzuordnen, da zumindest offen bleibt, ob diese Veränderungen nicht auch unabhängig von der Teilnahme an dem Projekt eingetreten wären, ob Lerneffekte aus dem Pretest zu anderen Ergebnissen im Posttest geführt haben oder ob andere Einflüsse (z. B. die bloße Teilnahme an dem Projekt) für die Differenz der Messergebnisse verantwortlich sind. Auch im Fall der nicht randomisierten Vergleichsgruppen können Differenzen in den Gruppenwerten nicht sicher als Effekt der Intervention begriffen werden.

Die zuverlässige Prüfung der Effektivität einer Intervention oder eines Programms setzt voraus, dass in der Anlage der Untersuchung die Beziehung von Intervention und den zur Messung vorgesehenen Effekten möglichst präzise modelliert ist. Die Entscheidung für ein bestimmtes Untersuchungskonzept resultiert dabei nicht aus abstrakten Erwägungen in Bezug auf quantitative und qualitative Methoden oder hinsichtlich experimenteller, design-basierter oder an der Grounded Theory orientierter Forschungslinien. Maßgebend ist vielmehr die Art von Information, die gesucht wird, d. h. das Wissen, das generiert werden soll – diese Frage ist entscheidend für die Anlage der Untersuchung. In ihrem Beitrag zur Einschätzung des möglichen Beitrags von Design-Studien formulieren dies SHAVELSON u. a. (2003, S. 35) wie folgt: „The effort to better understand the nature and validity of design studies comes at a time when policymakers, increasingly, are defining high-quality science in terms of traditional cause and effect methods, such as randomized experiments (...). We support the use of such designs, not as a blanket template for all rigorous work, but only when they are appropriate to the research question.“

Während bei der Entwicklung des Untersuchungsdesigns die Art des angestrebten Wissens die permanent präsente Leitlinie darstellt und die Entwurfsarbeit sich als iterativer Prozess zwischen antizipiertem Wissen und technischer Ermöglichung, dieses zu generieren realisiert, lautet nach Abschluss der Datengewinnung die Frage: Was können wir jetzt wissen? Gemeint ist damit die strikte Analyse des umgesetzten Designs und dessen Folgen: Was genau ist unternommen worden, und welches Wissen kann daraus gewonnen werden?

Hierzu ein weiteres Beispiel: Seit einigen Jahren sehen wir weltweit, dass an Schulen und Hochschulen Qualitätsmanagementsysteme implementiert werden. Selbstverständlich ist mit dieser Maßnahme die Erwartung verbunden, dass dadurch in den Einrichtungen die gewünschten Effekte erzielt werden. Bislang liegen jedoch weder aus dem Bereich der Wirtschaftsunternehmen (wo das Qualitätsmanagement seit längerer Zeit eingeführt ist) noch aus dem Bildungsbereich Daten vor, die eine zuverlässige Abschätzung der Effekte erlauben würden: Insgesamt wurden erst wenige Untersuchungen durchgeführt, und diese Studien liefern in Bezug auf die Effektivität der Maßnahme keine verlässliche Information. Orientiert man sich an dem oben zitierten Leitfaden, dann lässt sich mit Blick auf diese Sachlage die flächendeckende Implementierung von Qualitätsmanagementsystemen nicht rechtfertigen. Zugleich jedoch ist es ein legitimes Interesse der aktuell politisch Verantwortlichen, Informationen über die Effektivität der eingeführten Maßnahmen zu erhalten. Daten, die als „starker Beleg“ für die Effektivität gelten könnten, lassen sich allerdings unter den gegebenen Bedingungen kaum mehr generieren. Dennoch können wichtige Informationen gewonnen werden: So ist es für das Funktionieren des Qualitätsmanagementsystems u. a. bedeutsam,

- was in den Steuerungsvorgaben geregelt wird und wie weit oder eng sie gefasst sind,
- ob die Implementierung vor Ort konzeptgemäß erfolgt,
- wie stark der Durchdringungsgrad in der Organisation ist,
- wie stark die Maßnahme von den Beteiligten akzeptiert bzw. abgelehnt wird.

Während mit diesen Analysen und Erhebungen offensichtlich lediglich einzelne Phänomene beschrieben werden, findet sich in Bezug auf Messungen in verschiedenen Implementierungsstadien die Auffassung, dass sich in unterschiedlichen Werten Effekte der Maßnahme spiegeln. Da mit den Erhebungen – gleichgültig ob Messwiederholungen in den Organisationen oder Messungen in verschiedenen Organisationen mit unterschiedlichem Implementierungsstatus – i. d. R. weder eine zufällige Gruppenzuteilung noch eine systematische Variation der Intervention vorgenommen wurde bzw. werden konnte, sind die oben in der Tabelle für Step 2 beschriebenen Bedingungen nicht erfüllt. Auch wenn das, was im Rahmen dieses Untersuchungsdesigns getan wird, nicht die Möglichkeit bietet, Aussagen zur Effektivität der Maßnahme treffen zu können, so sind solche Erhebungen – wie anhand der oben angeführten Beispiele demonstriert – dennoch sinnvoll, um z. B. die Ausprägung ausgewählter Merkmale zu ermitteln, die für die Wirksamkeit bedeutsam sein können.

III

Der Beitrag ist ein Plädoyer für die weitere Stärkung von zwei Aufmerksamkeitsrichtungen in der Berufsbildungsforschung: Die eine hat ihren Fokus in der Fragestellung, zu der Antworten gesucht werden, in dem Problem, zu dessen Lösung mit einer Studie beigetragen werden soll. Dabei kann als Orientierung die Einteilung von SHAVELSON u. a. (2002, S. 28) herangezogen werden:

- What is happening?
- Is there a systematic effect?
- Why or how is it happening?

Da die Eignung eines Forschungskonzepts oder einer Erhebungsmethode bzw. eines Sets von Methoden in Relation zu sehen ist zu der in der Studie zu bearbeitenden Fragestellung, ist deren präzise Klärung für die folgenden Entscheidungen maßgebend.

Die zweite Aufmerksamkeitsrichtung bezieht sich auf das in einer Studie realisierte Forschungskonzept und die eingesetzten Methoden. Ob Durchführende oder Verwerter der Studie, für beide Personengruppen ist eine genaue Beschäftigung mit den im Forschungsprozess eingesetzten Konzepten, Methoden und Instrumenten relevant, um zu verstehen, welches Wissen aufgrund des Designs zu gewinnen grundsätzlich möglich war, und was an neuem Wissen tatsächlich gewonnen wurde.

Literatur

- BROWN, Ann L.: Design Experiments: Theoretical and Methodological Challenges in Creating Complex Interventions in Classroom Settings. In: *The Journal of the Learning Sciences* 2 (1992) 2, S. 141–178
- BUNDESINSTITUT FÜR BERUFSBILDUNG: Modellversuche, ein Instrument zur Weiterentwicklung beruflicher Bildungspraxis: Arbeitsergebnisse e. Workshops d. Bundesinst. für Berufsbildung u. D. Wirtschafts- u. Berufspädag. Studienkreises am 28. u. 29. September 1976 in Berlin. Schriften zur Berufsbildungsforschung, Band 52. Hannover 1978
- COE, Robert: A Manifesto for Evidence-based Education. Durham University 1999. – URL: <http://www.cemcentre.org/evidence-based-education/manifesto-for-evidence-based-education> (Stand: 09.08.2011)
- MAYER, Richard E.: Seeking a science of instruction. In: *Instructional Science: An International Journal of the Learning Sciences* 38 (2010) 2, S. 143–145
- MERRILL, M. David: First Principles of Instruction. In: *Educational Technology: Research and Development* 50 (2002) 3, S. 43–59

- NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE: Identifying and Implementing Educational Practices Supported by Rigorous Evidence: A User Friendly Guide. Coalition for Evidence-based Policy. Washington, DC 2003
- SHAVELSON, Richard J. u. a.: On the Science of Education Design Studies. In: Educational Researcher 32 (2003) 1, S. 25–28
- SLAVIN, Robert E.: Education Research Can and Must Address „What Works“ Questions. In: Educational Researcher 33 (2004) 1, S. 27–28
- SLAVIN, Robert E.: Evidence-based Reform in Education: Which Evidence Counts? In: Educational Researcher 37 (2008) 1, S. 47–50

© 2013 by Bundesinstitut für Berufsbildung, Bonn
Herausgeber: Bundesinstitut für Berufsbildung, 53142 Bonn
Internet: <http://www.bibb.de/veroeffentlichungen>

aus: EBNER, Hermann G.: Wissen schaffen: What you do is what you get.
In: SEVERING, Eckart; WEISS, Reinhold (Hrsg.): Qualitätsentwicklung in der Berufsbildungsforschung.
Bonn 2013, S. 61-70



Der Inhalt dieses Werkes steht unter einer Creative Commons Lizenz
(Lizenztyp: Namensnennung – Keine kommerzielle Nutzung – Keine Bearbeitung – 3.0 Deutschland).

Das Werk wird durch das Urheberrecht und/oder einschlägige Gesetze geschützt. Jede Nutzung, die durch diese Lizenz oder Urheberrecht nicht ausdrücklich gestattet ist, ist untersagt. Weitere Informationen finden Sie im Internet auf unserer Creative Commons-Infoseite: <http://www.bibb.de/cc-lizenz>