

*Annabell Preussler, Peter Baumgartner*

## **Qualitätssicherung in mediengestützten Lernprozessen – sind theoretische Konstrukte messbar?**

Preussler, A. und P. Baumgartner (2006). Qualitätssicherung in mediengestützten Lernprozessen – zur Messproblematik von theoretischen Konstrukten. In: Qualitätssicherung im E-Learning. Hrsg.: A. Sindler, C. Bremer, U. Dittler, et al. Münster, Waxmann. Medien in der Wissenschaft: 36: 73-85.

### **Zusammenfassung**

In unserem Artikel wollen wir die Problematik der Evaluation und Qualitätssicherung von Konstrukten beschreiben. So ist einerseits die Qualität selbst als ein Konstrukt anzusehen, aber auch der Lernerfolg, der ja oft als ein Indikator für Qualität gesehen wird. Eine Bewertung von Qualität wird aus dieser Sichtweise heraus schwierig, ist aber in der Praxis unumgänglich.

### **1 Einleitung<sup>1</sup>**

Mediengestütztes Lernen bzw. eLearning sind uneingeschränkt up to date. Dieser Hype wird nicht nur in der Diskussion deutlich, die in den letzten Jahren in der Bildungspolitik geführt wurde, es zeigt sich schon allein daran, wie sehr sich auch traditionell arbeitende Bereiche dieser Thematik annehmen. Als Beispiel sei die Entwicklung der FernUniversität in Hagen genannt. Gab es – und gibt es – seit Jahren und aus einer eingesessenen Tradition heraus noch das Studium mittels Studienmaterialien in Printform, die den Studierenden nach Hause geschickt werden, so zeigt sich eine Entwicklung hin zu eLearning-Angeboten. Dabei spielt nicht zuletzt der Aspekt der Wirtschaftlichkeit eine Rolle, denn diese Angebote werden von den Studierenden durchaus nachgefragt. Offenbar scheint sich alles,

---

<sup>1</sup> Bei der Angabe von Zitaten haben wir in diesem Beitrag jeweils die Vor- und Nachnamen der AutorInnen angegeben, um im Zuge von gender mainstreaming darauf hinzuweisen, ob der entsprechende Beitrag von einer Frau oder einem Mann verfasst worden ist.

was „e“ ist – also eLearning, eEducation, eInclusion, eGovernment usw. – gut zu vermarkten.

Doch wo liegen die Vorteile von eLearning? Warum soll es besser sein, als traditionelles Lernen? Wird durch eLearning eine höhere Lernqualität erzielt als mit traditionellem Lernen?

Diese Fragen werden in der Forschung in vielfältiger Weise zu beantworten versucht. Aus diesem Grund werden oftmals Vergleichsstudien herangezogen, die genau auf diesen Vergleich von Online- versus Präsenzlernen abzielen. Doch ist solch ein Vergleich sinnvoll?

## **2 Lernqualität – was ist das?**

Die Frage nach Lernqualität lässt sich nicht ohne weiteres beantworten, denn Lernqualität kann vielfältige Bedeutungen haben. Einerseits kann damit die Qualität des Lernergebnisses gemeint sein, aber auch die Qualität des Lernprozesses oder die der Behaltensleistung.

Es wird deutlich, dass es problematisch ist, eine Definition zu geben, weil es nicht DIE Lernqualität gibt. Es wird im Folgenden daher zunächst auf das Konstrukt Qualität eingegangen.

### **2.1 Das Konstrukt Qualität**

In der Wikipedia<sup>2</sup> – der größten freien Enzyklopädie im Netz – ist zur „Qualität“ zunächst verzeichnet, dass sich der Begriff von dem Lateinischen „qualitas“ ableitet, was soviel wie Beschaffenheit oder Eigenschaft bedeutet. Weiterhin wird angegeben, dass „Qualität“ in der Umgangssprache mit Wertungen verbunden ist und „die Zweckangemessenheit eines Ergebnisses menschlicher Arbeitstätigkeit bzw. der Beschaffenheit eines Produktes oder einer Dienstleistung zum Ausdruck bringt“ (Wikipedia, 2005). Am Ende des Eintrags findet sich eine sehr einfache, aber dennoch sehr einleuchtende Faustformel für die Definition von Qualität: „Qualität ist die Übereinstimmung von Ist und Soll“ (ebd).

So plausibel dies zunächst klingt – was Qualität ausmacht, lässt sich aus dieser „Regel“ nicht ableiten. Die Schwierigkeit liegt darin, dass sich Qualität nicht nur aus einer Perspektive zusammensetzt, sondern ein Konstrukt darstellt, welches aus dem Zusammenspiel vieler Faktoren entsteht (vgl. Ulf Ehlers 2004:25).

„Konstrukte [...] unterscheiden sich von Indikatoren (beobachtbare Merkmale) vor allem dadurch, daß sie nicht ohne weiteres meßbar

---

<sup>2</sup> <http://de.wikipedia.org>

sind. Konstrukte sind theoretische Größen, die innerhalb von Theoriekontexten sinnvoll sind, deren empirische Existenz jedoch nur behauptet werden kann. Für die empirische Forschung ergibt sich daraus die Notwendigkeit, Konstrukte operationalisieren zu müssen, wobei die Operationalisierungen wiederum nur einen hypothetischen Charakter haben, sich also als falsch erweisen können“ (Ute Clement & Bernd Martens 2000:106).

Zu einem Konsens über die Definition von Qualität zu gelangen, dürfte auch im wissenschaftlichen Diskurs schwierig sein. Ulf Ehlers weist darauf hin, dass es noch keine übereinstimmende Meinung darüber gibt, „wie etwa Bildungsqualität definiert ist“ (2004:51).

Dennoch bewegen wir uns nicht im definitionsleeren Raum, denn trotz dieser relativistischen Darstellung können Qualitätseigenschaften beschrieben werden.

So ist Qualität im Bildungsbereich unter anderem dadurch gekennzeichnet, dass sie erst nach Inanspruchnahme (also nach erfolgter Nutzung) sichtbar und überprüfbar ist und darüber hinaus nicht unmittelbar sichtbar ist. Außerdem ist Qualität schwer zu standardisieren (vgl. Lee Harvey & Diana Green 2000).

Des Weiteren sagt Qualität etwas über die Beschaffenheit eines Gegenstandes aus und setzt sich nicht mit dem Gegenstand an sich auseinander (vgl. ebd.). Sie lässt sich also nicht durch den Inhalt eines Produkts beschreiben, sondern entsteht aus dem Zusammenwirken verschiedener Dimensionen.

Betrachten wir die Lernqualität, so ist ein Faktor dabei auch der/die Lernende selbst. Qualität entsteht demnach erst dann, wenn die Lernenden selbst mit der Lernsituation oder dem Lernarrangement in Interaktion treten. Lernqualität entsteht erst, wenn auch gelernt wird (vgl. Gerhard M. Zimmer & Elena Psaralidis 2000:265; Ulf Ehlers 2002:7 ff.). Sie ist demzufolge auch erst nach Inanspruchnahme sichtbar und überprüfbar und schwer zu standardisieren.

Im Zuge dieser Entwicklung wird der Qualitätsbegriff erweitert, indem er auch die Rahmenbedingungen des Lernens und die Lernenden selbst einbezieht (Sigmar-Olaf Tergan 2003). Wenn Lee Harvey & Diana Green (2000:31) also „Qualität als Transformation“ beschreiben, so ist damit gemeint, dass sich die Qualität aus dem Verhältnis zwischen Lernangebot und Lernendem als „Weiterentwicklung des Lernenden durch einen Lernprozess“ konstruiert (Ulf Ehlers 2002:16).

Seit dem Aufkommen des eLearnings wurden eine Reihe von Qualitätskonzepten entwickelt. Diese zeigten jedoch keine großen Erfolge, da „ein auf alle Lerner gleichermaßen übertragbarer Zusammenhang zwischen didaktischen oder medialen Attributen eines Lernsystems [...] nach derzeitigem Stand nicht

nachgewiesen werden [konnte]“ (Ulf Ehlers 2004:24, vgl. auch Bernd Weidenmann 1997, James A. Kulik & Chen-Lin Kulik 1991).

Entscheidend für die Messung und Sicherung von (Lern-)Qualität ist Evaluation. Durch Evaluation werden sowohl Lernangebote, als auch –Ergebnisse anhand bestimmter Kriterien und im Hinblick auf die Lernziele gemessen und bewertet. Eines dieser Kriterien ist in sozialwissenschaftlichen Evaluationsstudien oftmals der Lernerfolg. Er dient sozusagen als ein Indikator für Lernqualität.

### **3 Messung von Lernqualität**

Ein beliebter Weg, Lernqualität zu belegen, ist die Messung des Lernerfolges. Es wird davon ausgegangen, dass eine Software oder ein Bildungsangebot höhere Lernerfolge erzielt, je höher die Qualität eben jener Software ist. Will man sich beispielsweise ein Lernprogramm zulegen, überprüft man es auf ein Qualitätssiegel oder eine Testnote. Gekauft wird das Programm, welches die höchste Qualität verspricht, weil ein höherer Lernerfolg erhofft wird.

Umgekehrt wird die Qualität einer Software auch darin gesehen, wie gut damit gelernt werden kann, also wie hoch der erzielte Lernerfolg ist. Wer an der Entwicklung einer Software arbeitet oder mit deren Evaluation beauftragt ist, versucht oft, den Lernerfolg zu messen um somit Aussagen zur Qualität treffen zu können. Die Qualität wird somit als Voraussetzung für den Lernerfolg gesehen (vgl. Reiner Fricke, 2000).

#### **3.1 Das Konstrukt Lernerfolg**

Was macht den Erfolg von Lernen eigentlich aus? Ist es die Effektivität des Lernprozesses, das Replizieren von Wissen oder nur die Anzahl der richtig gelösten Aufgaben in einem Test? Für die Feststellung der Qualität eines Lehr- und Lernmediums und die daraus folgende Evaluation ist die Definition dessen, was unter Lernerfolg verstanden wird, allerdings ein sehr wichtiger Punkt.

Es liegt auf der Hand, dass es nicht DEN Lernerfolg gibt, sondern dass es sich auch hierbei um ein Konstrukt handelt, welches sich aus verschiedenen Dimensionen zusammensetzt.

Michael Kerres (2001:111) sieht den Lernerfolg als „Ergebnis didaktischer Aktivitäten“, aber nicht als das bloße „Behalten von Fakten, Ereignissen oder Vorgängen“. Bei einer Reduktion des Lernerfolges auf bloße Behaltensleistung werden sowohl Persönlichkeit der Lernenden als auch ihre Denkschemata außer Acht gelassen. Dazu kommt weiterhin, dass auch der Lerntransfer nicht

ausreichend beachtet wird – also: können wir das, was wir lernen, auch praktisch umsetzen, oder: ist der Erfolg in einer Lernsituation auf den Alltag übertragbar?

Der Lernerfolg ist nach Kerres unter anderem davon abhängig, „inwieweit es gelingt, die angestrebten kognitiven und emotionalen Lernprozesse anzuregen“. Lernangebote sind demnach „so aufzubereiten, dass ihre Bearbeitung bestimmte Lernprozesse möglich machen“ (ebd:147).

Kerres stellt einige Kriterien heraus, die seines Erachtens für die Spezifizierung des Konstrukts Lernerfolg im Kontext mediengestützten Lernens geeignet sind. Diese umfassen beispielsweise die erlebte Qualität des Lernangebotes, die emotionale Reaktion und Lernmotivation, das Lernverhalten, die subjektive Zufriedenheit mit dem Lernverhalten/-ergebnis, die faktische Nutzung/Akzeptanz und „Lebensfähigkeit“ des mediengestützten Lernens im organisationalen Kontext usw., aber auch den „objektiven“ Lernerfolg in verschiedenen zeitlichen Abständen und unter unterschiedlicher Anwendungsnähe (vgl. ebd. 2001:112).

Was ist unter „objektiven“ Lernerfolg zu verstehen und kann er überhaupt gemessen werden? Auf diese Problematik möchten wir im Folgenden eingehen.

### **3.2 Lernerfolgsmessung: Fehler und Fallen der Forschung**

„Ein Schüler, der sich auf den Unterricht konzentriert, Mitschriften anfertigt, bei Unklarheiten stutzig wird und entsprechend nachfragt, der am Nachmittag liest oder mit Freunden den Unterrichtsstoff diskutiert, der sich Unklares von Eltern oder älteren Geschwistern erklären lässt, der Textstellen unterstreicht, sich durch stilles Wiederholen wichtige Informationen einprägt und der sich vergewissert, dass er auch wirklich den Stoff für die nächste Klassenarbeit beherrscht, dessen Leistung sollte einer guten Note entsprechen“. (Nadine Spörer 2003:69).

Dieses Zitat macht bereits deutlich, dass es nicht so einfach ist, Lernerfolg anhand einzelner Kriterien festzumachen. Denn eine Schülerin, die den Lernstoff bloß auswendig gelernt wiedergibt, kann unter Umständen eine bessere Note erzielen, als der im Zitat erwähnte Schüler. Die Messung des Lernerfolgs orientiert sich im Schulbereich, wie auch in Evaluationsstudien oftmals an standardisierten Wissensabfragen.

Die Verwechslung von Denkprozessen und Denkprodukten ist jedoch ein Kategorienfehler und stellt einer der schwerwiegendsten Fehler bei Evaluationen zum Lernerfolg dar (vgl. Peter Baumgartner 1999a).

Es ist nämlich keineswegs klar und eindeutig, was unter Lernerfolg zu verstehen ist. Die bisherige traditionelle, jedoch oft vorherrschende Sichtweise, dass der

Lernerfolg in der Reproduktion und/oder Anwendung von theoretischem Wissen (sei es nun Fakten- oder prozedurales Wissen) besteht, greift eindeutig zu kurz und ist noch überwiegend von den „alten“ Medien, d.h. der mehr oder weniger passiven Wissensvermittlung, geprägt. Zurückzuführen sind diese unsinnigen Evaluationen auf die Gleichsetzung von prozeduralem Wissen („know how“ bzw. Wissen, wie etwas geht) mit der tatsächlichen Fertigkeit, dem skill selbst: Wenn ich weiß, wie ein Autoreifen gewechselt wird, heißt das lange noch nicht, dass ich es auch tatsächlich kann (vgl. Peter Baumgartner 1999b).

Schulmeister et al. weisen des Weiteren darauf hin, dass bei Evaluationen auch zwischen den Variablen Lernerfolg und Lernerzufriedenheit zu unterscheiden ist, die beide erfasst werden müssen:

„Der häufig in objektivierten Maßen, etwa Klausurnoten oder Prüfungsleistungen, erfasste Lernerfolg sagt nichts über die subjektive Zufriedenheit des Nutzers mit der Lernsoftware aus [...]. Lernerfolg ist zudem ein hoch komplexes, widersprüchliches Merkmal. Niedrige Werte in „objektiven“ Lernerfolgsmaßen können mit positiven Beurteilungen der Lernsoftware durch einen Nutzer einhergehen und umgekehrt“ (Rolf Schulmeister et al., 2005).

Ute Clement & Bernd Martens verweisen in diesem Zusammenhang auch auf die Anforderungen an modernes Lernen. So werden auf der einen Seite zwar Interaktivität und ein hohes Maß an Selbststeuerung beim Lernen gefordert, der Lernerfolg soll aber dennoch kontrollierbar und möglichst vorhersagbar sein (vgl. 2000:98). Dieser Aspekt stellt eine weitere Schwierigkeit bei der Bewertung von Lernerfolg dar. Auch Kerres (2001:11) gibt zu bedenken, dass im Rahmen einer konstruktivistischen Sichtweise und dem Ansatz des situierten Lernens Wissen ständig in sozialen Kontexten konstruiert und also nicht im Kopf „abgespeichert“ wird. Es wird deutlich, dass damit Evaluationsansätze, die auf das Individuum ausgerichtet sind, wenig brauchbar sind. Ein möglicher alternativer Evaluationsansatz wäre demnach, zu prüfen, inwieweit die Lernangebote generative soziale Settings unterstützen.

In der Forschungspraxis ist ein solches Vorgehen oft schwierig umzusetzen. Es verwundert daher nicht, dass in vielen Evaluationsstudien Lernerfolg mit Behaltensleistung gleichgesetzt wird (vgl. aaO).

Werden Lernprozesse empirisch analysiert, wird in der Regel angenommen, „es bestehe ein kausaler Zusammenhang zwischen Ergebnissen von Lernprozessen und davon unabhängigen Einflussgrößen“ (Ute Clement & Bernd Martens 2000: 99). Ohne diese Annahme ergäben beispielsweise didaktische Bemühungen oder Differenzierung keinen Sinn. Der Lernerfolg stellt in diesem Modell eine abhängige Größe dar, die in einem kausalen Zusammenhang zu einer Anzahl

unabhängiger Variablen steht: „Die Forschung versucht, die Abhängigkeiten zwischen Ursache und Wirkung im empirischen Einzelfall oder verbunden mit einem allgemeinen theoretischen Anspruch zu klären“ (ebd.). In der empirischen Forschung – darauf weisen Clement & Martens ausdrücklich hin – fällt es jedoch schwer, die Gültigkeit eines gesamten Modells zu belegen. Es existieren „keine metatheoretischen Aussagen über die Gültigkeitsbereiche sozialwissenschaftlicher Methoden“ (2000:100).

Bedenkt man weiterhin die Schwierigkeit, die Lernwirksamkeit entsprechender Qualitätskriterien nachzuweisen, wird klar, dass der Ansatz der empirischen Überprüfung von Lernerfolg und Qualität als unrealistisch bezeichnet werden kann (vgl. Gerhard M. Zimmer & Elena Psaralidis, 2000; Ulf Ehlers, 2003; Sigmar-Olaf Tergan, 2003).

Weiterhin ist es ein Trugschluss zu glauben, dass man verschiedene Lernformen – wie Online- versus Präsenzlernen – problemlos miteinander vergleichen könne, obwohl dies immer wieder versucht wird. Zahlreiche Untersuchungen belegen, dass ein solches Verfahren in der Regel zu nicht-signifikanten Ergebnissen führt (vgl. z.B. Sigmar-Olaf Tergan 2003, Jane M. Carey, 2001, S. 7).

Viele Jahre lang stellten Vergleichsuntersuchungen im herkömmlichen Sinne einen methodischen Forschungsstandard dar – ein Lernsetting wurde mit einem anderen verglichen. Auf diese Weise wurde erhofft, Rückschlüsse auf den Erfolg eines Settings zulassen zu können. Nach wie vor erfreuen sich derartige Vergleichsstudien großer Beliebtheit. In vielen Bereichen mag dies auch ein sinnvolles Verfahren in der pädagogischen Forschung sein, im Bereich der Medienwirkungsforschung ist es durchaus diskussionswürdig.

Bei diesen aufwendigen experimentellen Studien wird der Lernerfolg meist implizit durch das Testverfahren definiert, wobei aus methodischen Gründen der Vergleichbarkeit das gleiche Prüfverfahren (z.B. schriftliche Tests für beide Gruppen) verlangt wird. Dieses Evaluationsdesign ist aber in zweifacher Hinsicht problematisch:

1. Wenn es sich bloß um einen Vergleich rezeptiven Lernens, wo es in erster Linie um die Präsentation (Vermittlung) von Inhalten mittels Text, Grafik oder Bewegung geht, handelt, dann werden in diesem Fall die neuen Eigenschaften und Vorteile der Interaktivität gerade nicht genutzt! Vergleichenden Evaluationen dieser Art sind daher wenig sinnvoll. Wenn mehrdimensionale Evaluanden in Hinblick auf einem bestimmten Aspekt miteinander verglichen werden, dieser Gesichtspunkt aber für einen der Evaluanden gerade *nicht* typisch ist, dann sind natürlich allgemeine Bewertungen und Aussagen zu den Evaluanden insgesamt nicht zulässig. Wenn z.B. untersucht wird ob das Lesen von Texten auf Papier oder am Bildschirm effizienter ist, und die empirischen

Ergebnisse sprechen (natürlich!) für das Buch, dann lassen sich darüber keine allgemeinen Aussagen zum Lernmedium selbst ableiten. Zulässig ist bloß der eingeschränkte Schluss, dass das Medium Buch für rezeptive Leseaufgaben besser geeignet ist.

Trotz aufwändiger und kostenintensiver Methodik unterliegen generalisierende Vergleichsstudien einem wissenschaftstheoretischen Zirkelschluss: Zuerst wird Lernen kognitivistisch eingeschränkt und anschließend behauptet, dass „Lernen“ durch interaktive Software generell nicht besser unterstützt wird, als es traditionelle Mittel (z.B. ein gutes Buch) ohnehin schon tun bzw. getan haben.

2. Oder aber umgekehrt: Wenn in diesen vergleichenden Evaluationen gezeigt wird, dass bestimmte Zusammenhänge mit interaktiven Medien besser gelernt werden können, dann ist in diesen Untersuchungen „Lernerfolg“ nicht mehr bloß rezeptiv definiert. Trotzdem ist damit das methodologische Problem einer Vergleichsstudie nicht gelöst und kann daher in einem vergleichenden Evaluationsdesign nicht adäquat eingelöst werden: Der Lernerfolg muss ja qua Evaluationsdesign mit *gleichen* Inhalten und *gleicher* Prüfmethode festgestellt werden (vgl. Peter Baumgartner 1997).

Unter Bezugnahme auf Daniel Surry & David Ensminger (2001) seien nun einige Gründe für die Kritik an Medienvergleichsforschungen genannt (nach Patrick Blumschein 2004:8):

1. Es gibt kaum signifikante Unterschiede bei Medienvergleichsstudien (vgl. Clark 1994, Kulik & Kulik 1989). Laut Blumschein deutet dies darauf hin, „dass nicht die Wahl des Mediums sondern vielmehr die angewandte Methode ausschlaggebend die Qualität des Lernens beeinflusst“ (Patrick Blumschein 2004:8)
2. Über die Medienwirkung auf Individuen ist zu wenig bekannt. „We don't know enough about the attributes of different media or the way people learn with media to design effective media comparison studies“ (Daniel Surry & David Ensminger 2001:33)<sup>3</sup>.
3. Da intervenierende Variablen in vielen Medienvergleichsstudien nicht berücksichtigt werden, sind diese Studien methodisch mangelhaft, weil sie so gut wie keine externe Validität aufweisen (vgl. Jürgen Bortz & Nicola Döring 2002; Bernhard Jacobs 1998).

---

<sup>3</sup> An dieser Stelle sei auf die ATI-Forschung hingewiesen, die vor allem durch Lee J. Cronbach bekannt wurde und die sich mit der Interaktion zwischen Vorwissen der Lernenden und des Untersuchungsverfahrens beschäftigt. Allerdings ist auch der Erfolg dieser Forschungsrichtung als gering einzuschätzen.

Wir möchten zusammenfassend mit Schulmeister feststellen, dass „aus dem Darbietungsmodus der Lerninhalte und der Unterscheidung zwischen verschiedenen Lernmedien [...] keine Unterschiede im objektiv messbaren individuellen Lernerfolg oder der von den Lernern subjektiv empfundenen Lernqualität ableitbar [sind]“ (Rolf Schulmeister et al. 2005). Oder, wie Joachim Hasebrook bereits vor 10 Jahren feststellte: „Nach dem heutigen Erkenntnisstand ist es nicht möglich, prinzipielle Aussagen über die Lernwirkungen von Multimedia zu machen“ (1995:101).

Es dürfte wohl inzwischen aus unserer Argumentation deutlich geworden sein, dass wir der Auffassung sind, dass weder klare Definitionen von Lernerfolg, noch deren empirische Erfassung problemlos durchzuführen sind. Wenn sich diese Ansicht durchsetzt, dass müssten in Zukunft konsequenterweise diese methodologisch falsch angelegten „Vergleichs“studien in Begründungsnot kommen. Das betrifft sowohl ihre Finanzierung als auch ihre Akzeptanz für Zeitschriften und Kongresse, wo sie häufig immer noch ihre fälschlicherweise verallgemeinerten Resultate ausbreiten dürfen.

Es gibt bereits eine wissenschaftlich akzeptierte und auch renommierte Diskussion nach Alternativen. Wir meinen damit den Diskurs der unter ErziehungswissenschaftlerInnen unter „Handlungsforschung“ geführt wird (vgl. John Elliott 1991, Herbert Altrichter & Peter Posch 1998 etc.) und von den PsychologInnen unter dem Titel „Design Based Research“ in letzter Zeit aufgegriffen worden ist (vgl. Gabi Reinmann 2005 etc.)

#### **4 Wohin soll die Reise gehen? – Ein Fazit**

Wie deutlich wurde ist es nicht ohne weiteres möglich, Aussagen über den Lernerfolg beim eLearning zu treffen, da Lernerfolg als ein Konstrukt zu sehen ist, von dem keine „Reinform“ existiert. So können beispielsweise Behaltensleistungen, neu erworbene Fähigkeiten und Kompetenzen oder auch nur die Anzahl richtig angekreuzter Antworten in einem Test als Lernerfolg betrachtet werden. Dies führt dazu, dass unspezifische Vergleiche – z.B. online- versus Präsenzlernen – nur bedingt geeignet sind, um Aussagen über den Erfolg eines Lernarrangements zu treffen.

Dies bedeutet aber auch, dass über die Qualität einer Lernumgebung nur wenig ausgesagt werden kann, wenn der Lernerfolg als ein Qualitätskriterium angesehen wird. Damit würde man ein Konstrukt (Qualität) unter Bezugnahme auf ein anderes Konstrukt (Lernerfolg) zu operationalisieren versuchen, was ein hoffnungsvolles Unterfangen ist.

Dennoch wollen wir keine relativistische Position vertreten. Wir vertreten damit nicht die Auffassung, dass nichts mehr gemessen werden kann, dass quantitative empirische Wissenschaft obsolet ist und keine Perspektive hat. Ganz im Gegenteil, wir meinen, dass es sich dieser Wissenschaftszweig eine Reihe von Herausforderung stellen muss und sich die Sache nicht zu einfach machen darf. Unsere (nicht relativistische) Schlussfolgerung ist vielmehr: Das (vergleichende) Untersuchungsdesign muss der Komplexität der untersuchten Wechselwirkungen entsprechen! Werden zur Reduktion der Komplexität des Untersuchungsdesign Bedingungen als unabhängige Variablen konstant gehalten, dann dürfen in der Interpretation der Ergebnisse keine (implizite) Aussagen über diese Variablen getroffen werden. Das führt natürlich zu einer deutlichen Einschränkung der Interpretation quantitativer empirischer Forschungsergebnisse und wirft die Frage auf, ob der dafür investierte Aufwand in Personal und Geld die Resultate dieser Art von Forschung rechtfertigen. Das ist aber eine wissenschaftspolitische Entscheidung und hängt zum Teil zumindest auch vom vorherrschenden Paradigma der Wissenschaft ab: Wer sitzt mit welcher Wissenschaftsauffassung in welchen Gremien und bestimmt darüber was als wissenschaftliche Forschung legitim, sozial verträglich und politisch gewünscht ist.

In der Praxis hat es wenig Sinn, auf die Frage, inwiefern man Lernplattform X gegenüber Lernplattform Y favorisiere, mitzuteilen, dass es keine hinreichende wissenschaftliche Begründung dafür gibt. Evaluationen sind für die Qualitätssicherung also nach wie vor nötig. Zwar möchten wir Gerhard M. Zimmer & Elena Psaralidis (2000:263f) zustimmen, „dass die Bewertung eines Lernangebots auf der Grundlage seiner Produktmerkmale noch nichts über die erzielbaren Wirkungen aussagt“, dennoch kommen wir in der Praxis um Bewertungen – und den dafür notwendigen komplexen Untersuchungssetting – nicht umhin.

Wie könnte also ein alternativer Ansatz zur Messung der Qualität eines eLearning-Settings aussehen? Auch hier beziehen wir uns zunächst auf Gerhard M. Zimmer & Elena Psaralidis (2000), indem wir feststellen, dass erst durch das Lernen selbst Qualität hergestellt werden kann. Lernen ist demnach in erster Linie als eine soziale Aktivität zu begreifen und nicht bloß als eine kognitive Erscheinung. Allerdings sollte man Qualität, so stellt Sigmar-Olaf Tergan (2003) fest, nicht *nur* an den Lernaktivitäten, sondern vor allem in „Hinblick auf die Bewältigung bestimmter Anforderungen von Problem- und Aufgabenstellungen“ festmachen, die in ganz unterschiedlicher Weise von den Lernenden bzw. vom Lernangebot ausgehen können. Wie wir bereits oben dargestellt haben, kann Qualität von e-Learning also aus zwei unterschiedlichen Blickrichtungen betrachtet werden: Sie kann sich einerseits auf die Qualität des angebotenen Materials und andererseits auf die Qualität des Lernergebnisses beziehen. Während es sich im ersten Fall um eine Art „Trockenprüfung“ handelt – wo das Material beispielsweise von

ExpertInnen(gruppen) durchgesehen wird –, erfordert die zweite Sichtweise die Evaluierung eines bereits stattgefundenen Lehr-/Lernvorgangs, eines dynamischen Prozesses, einer sozialen Aktivität.

Diese beiden Sichtweisen erfordern nicht nur jeweils eine unterschiedliche Methodik, sondern können auch zu völlig unterschiedlichen Ergebnissen führen. Beispielsweise kann der online Kurs XY in der „Trockenprüfung“ sehr gut abschneiden (indem er z.B. ein bestimmtes Themengebiet vollständig, inhaltlich korrekt und – nach Meinung der evaluierenden ExpertInnen – didaktisch gut aufbereitet und abdeckt), während derselbe Kurs bei der Lernerfolgskontrolle dann katastrophale Ergebnisse liefert (z.B. weil die Zielgruppe nicht adäquat war oder weil die Performance der Software im Realbetrieb eingebrochen ist). Umgekehrt können rudimentäre, schlecht aufbereitete Angebote durch ein gelungenes didaktisches Design zu sehr guten Lernerfolgen führen. (Paradoxaerweise gerade auch deswegen, weil sie rudimentär sind, d.h. viel Eigenleistung und Eigenaktivitäten der Studierenden verlangen.) Obwohl es zwischen der Qualität des Angebotes und der Qualität des Lernergebnisses natürlich einen Zusammenhang gibt, so ist doch zu betonen, dass das was letztlich zählt, natürlich die Qualität des Lernergebnisses ist.

Bei der Einschätzung des Lernergebnisses ist eine klare Vorstellung über Lernziel und damit zusammenhängend über die verwendete Lerntheorie erforderlich: Wenn beispielsweise fachliche Inhalte abstrakt abgefragt werden, ist bloß das Reproduzieren des Stoffes (Faktenwissen) gefordert; wird hingegen ein komplexes Problem präsentiert, sind unter anderem auch anderen Wissensformen wie Anwendungswissen, Auswahl und Einsatz von Methodenwissen gefordert. Während bei der Evaluierung gelernten Faktenwissens verschiedene (individuelle) Test- und Assessmentwerkzeuge verwendet werden können, die einen relativ starken Zusammenhang zur Angebotsstruktur haben, steht bei konstruktivistischen Lehrformen auch die Lehrorganisation inklusive des Betreuungsprozesses auf dem Prüfstand.

Eine der hauptsächlichen Fehler bei Evaluationen besteht gerade darin, dass die Aufmerksamkeit in erster Linie auf die Methodik der Verfahren gelegt wird und dabei die Frage nach den Lernzielen vernachlässigt wird. Das Ergebnis beschränkt sich dann häufig auf die Bewertung einfacher Formen der Wissensvermittlung (präsentative Lehrformen) und weniger auf die Evaluierung von Fähigkeiten und Kompetenzen (vgl. Peter Baumgartner & Annabell Preussler 2005:5).

An dieser Stelle könnte dann ein gänzlich anderer Ansatz zum Tragen kommen, nämlich die Idee einer kompetenzbasierten Ausbildung<sup>4</sup> und deren Bewertung.

---

<sup>4</sup> Dieses Modell wird beispielsweise im neuen Master-Studiengang „eEducation /Bildung und Medien“ an der FernUniversität in Hagen eingesetzt werden. Auch die Western Governors

Unter dem Begriff der Kompetenz verstehen wir nach John Erpenbeck & Lutz von Rosenstiel „Dispositionen selbstorganisierten Handelns“(2003:11). Dabei zeigen sich Kompetenzen vor allem in kreativen Leistungen, die im Gegensatz zu Qualifikationen eben nicht auf traditionellem Weg geprüft werden können und sollen. Sie beziehen sich eher auf selbstorganisierte subjektorientierte „Aufführungen“, die auch als Performances verstanden werden können. Die Prüfungsleistungen werden in ePortfolios gesammelt und archiviert. Diesen Ansatz – den wir auch an unserem Lehrgebiet weiter verfolgen wollen – ist unserer Ansicht nach ein wichtiger Schritt bei der Bewertung von Qualität in eLearning-Szenarien.

## **Literatur**

- Altrichter, Herbert & Posch, Peter (1998): *Lehrer erforschen ihren Unterricht. Eine Einführung in die Methoden der Aktionsforschung*. 3. durchgesehene und erweiterte Auflage, Bad Heilbrunn.
- Baumgartner, Peter (1997): *Evaluation vernetzten Lernens: 4 Thesen*. In: Wagner, Erwin & Kindt, Michael (Hrsg.) (1997): *Virtueller Campus. Forschung und Entwicklung für neues Lehren und Lernen*. Münster. S. 131-146.
- Baumgartner, Peter (1999a): *10 Todsünden in der Evaluation interaktiver Lehr – und Lernmedien*. In: K. Lehmann (1999) *Studieren 2000 - Alte Inhalte in neuen Medien? (mit CDROM)*. Münster, 199-220.
- Baumgartner, Peter (1999b): *Evaluation mediengestützten Lernens. Theorie - Logik - Modelle*. In: Kindt, Michael (Hrsg.) (1999): *Projektevaluation in der Lehre - Multimedia an Hochschulen zeigt Profile*. Münster. S. 61-97.
- Baumgartner, Peter; Häfele; Hartmut; Maier-Häfele, Kornelia (2002): *E-Learning Praxishandbuch. Auswahl von Lernplattformen : Marktübersicht - Funktionen - Fachbegriffe*. Innsbruck.
- Baumgartner, Peter (2003): *Die Theorie des Lernens*. In: Werdan, Karl; Trappe, Hans-Joachim und Zerkowski, Hans-Reinhard (2003): *Mensch und E-Learning. Beiträge zur E-Didaktik und darüber hinaus*. München-Jena. S. 957-960.
- Baumgartner, Peter & Preussler, Annabell (2005): *E-Learning-Qualität: Evaluation und Assessment*. Studienbrief für Modul 8 des Weiterbildungsstudiengangs „Technologiebasiertes Lehren und Lernen“ Virtueller Campus Rheinland-Pfalz.
- Blumschein, Patrick (2004): *Eine Metaanalyse zur Effektivität multimedialen Lernens am Beispiel der Anchored Instruction*. Dissertation in der Wirtschafts- und verhaltenswissenschaftlichen Fakultät. Freiburg, Albert-Ludwigs Universität Freiburg im Breisgau.
- Bortz, Jürgen & Döring, Nicola (1995): *Forschungsmethoden und Evaluation für Sozialwissenschaftler*. 2. vollständig überarbeitete und aktualisierte Auflage. Berlin.

---

University (WGU) hat bereits ein kompetenzbasiertes Ausbildungsmodell eingeführt (WGU 2004).

- Carey Jane M. (2001): Effective Student Outcomes: A Comparison of Online and Face-to-Face Delivery Modes. In: DEOSNEWS Vol.11 No. 9.
- Clement, Ute & Martens, Bernd (2000): Effizienter Lernen durch Multimedia? Probleme der empirischen Feststellung von Ursachen des Lernerfolgs. In: *Zeitschrift für Pädagogik*, Vol. 46, No. 1, Januar/Februar, pp. 97–112
- Ehlers, U. (2002). Qualität beim eLearning. Der Lernende als Grundkategorie der Qualitätssicherung. In: DGfE (Hrsg.). *Medienpaed.com – Onlinezeitschrift für Theorie und Praxis der Medienbildung*. Ausgabe 01/2002. [WWW-Dokument, <http://www.medienpaed.com/02-1/ehlers1.pdf> , letzter Aufruf 07.06.2005].
- Ehlers, Ulf-Daniel (2004): Qualität im E-Learning aus Lernericht. Wiesbaden.
- Elliott, John (1991): Action Research for Educational Change. Milton Keynes, Philadelphia: Open University Press.
- Erpenbeck, John und Rosenstiel, Lutz von (Hrsg.) (2003). Handbuch Kompetenzmessung. Erkennen, verstehen und bewerten von Kompetenzen in der betrieblichen, pädagogischen und psychologischen Praxis. Stuttgart, Schäffer-Poeschel.
- Fricke, Reiner (2000): Qualitätsbeurteilung durch Kriterienkataloge. Auf der Suche nach validen Vorhersagemodellen. In: Schenkel, P., Tergan, S.-O. & Lottmann, A. (Hrsg.) (2000): *Qualitätsbeurteilung multimedialer Lern- und Informationssysteme. Evaluationsmethoden auf dem Prüfstand*. Nürnberg
- Harvey, Lee & Green, Diana (2000): Qualität definieren. Fünf unterschiedliche Ansätze. In: *Zeitschrift für Pädagogik*, 41. Beiheft, 17-40.
- Hasebrook, Joachim (1995): Lernen mit Multimedia. In: *Zeitschrift für Pädagogische Psychologie* 9(2), 95-103.
- Kerres, Michael (2001): Multimediale und telemediale Lernumgebungen. Konzeption und Entwicklung. 2. Auflage, München: Oldenbourg Verlag.
- Kulik, Chen-Lin C. & Kulik, James A. (1991): Effectiveness of Computer-based Instruction: An Updated Analysis. In: *Computers in Human Behaviour*, 7, 75-94.
- Reinmann, G. (2005). „Innovation ohne Forschung? Ein Plädoyer für den Design-Based Research Ansatz in der Lehr- Lernforschung.“ In: *Unterrichtswissenschaft* 2.
- Schulmeister, Rolf (1997) Grundlagen hypermedialer Lernsysteme. Theorie - Didaktik - Design. 2. aktual. Aufl. München.
- Schulmeister, Rolf (2000): Virtuelles Lernen aus didaktischer Sicht. [WWW-Dokument, <http://www.izhd.uni-hamburg.de/pdfs/VirtLern.PDF> , letzter Aufruf April 2005).
- Schulmeister, Rolf; Vollmers, Burkhard; Gücker, Robert; Nuyken, Klaus (2005): Konzeption und Durchführung der Evaluation einer virtuellen Lernumgebung: Das Projekt Methodenlehre-Baukasten. In: Bachmair, Ben; Diepold, Peter; de Witt, Claudia: *Handbuch Medienpädagogik Bd. 5*, Opladen.
- Spörer, Nadine (2003): Strategie und Lernerfolg. Validierung eines Interviews zum selbstgesteuerten Lernen. Dissertation in der Humanwissenschaftlichen Fakultät der Universität Potsdam.
- Surry, Daniel W. & Ensminger, David (2001). What's Wrong with Media Comparison Studies? In: *Educational Technology*, 41(4), 32-35.

- Tergan, Sigmar-Olaf (2003): Realistische Qualitätsevaluation von E-Learning. In D. Meister, S.-O. Tergan & P. Zentl (Hrsg.): *Evaluation von E-Learning - Zielrichtungen, methodologische Aspekte, Zukunftsperspektiven*. Münster, S.131-154.
- Tergan, Sigmar-Olaf & Schenkel, Peter (Hrsg.). (2004). Was macht E-Learning erfolgreich? Grundlagen und Instrumente der Qualitätsbeurteilung. Berlin.
- Tergan, Sigmar-Olaf & Schenkel, Peter (2004). Was macht Lernen erfolgreich? Evaluation des Lernpotenzials von E-Learning. In Andreas Hohenstein & Karl Wilbers (Hrsg.) *Handbuch E-Learning*. Neuwied.
- Weidenmann, Bernd (1997): „Multimedia“: Mehrere Medien, mehrere Codes, mehrere Sinneskanäle? In: *Unterrichtswissenschaft*, Heft 3/1997, S. 197-206.
- Winteler, Adi (2002): Lehrqualität = Lernqualität? Über Konzepte des Lehrens und die Qualität des Lernens (Teil 1). In: *Das Hochschulwesen* 2/2002. 50. Jahrgang, S. 42-49.
- Zimmer, Gerhard M. & Psaralidis, Elena (2000): „Der Lernerfolg bestimmt die Qualität einer Lernsoftware!“ – Evaluation von Lernerfolg als logische Rekonstruktion von Handlungen. In: In: Schenkel, P., Tergan, S.-O. & Lottmann, A. (Hrsg.) (2000): *Qualitätsbeurteilung multimedialer Lern- und Informationssysteme. Evaluationsmethoden auf dem Prüfstand*. Nürnberg.

### **Web links:**

- WGU, W. G. U. (2004). What is a WGU Competency-Based Degree?  
WWW-Dokument  
[[http://www.wgu.edu/wgu/academics/comp\\_what\\_is.html](http://www.wgu.edu/wgu/academics/comp_what_is.html)] , letzter Aufruf 07.06.2005.
- Wikipedia (2005): Qualität. WWW-Dokument  
[<http://de.wikipedia.org/wiki/Qualit%C3%A4t>] , letzter Aufruf 07.06.2005.