

Lernsoftware-Evaluation

Tücken und Kriterien der Beurteilung von Lernsoftware

Von Cornelia Biffi

Mit dem Begriff *Lernsoftware* beschreibt Baumgartner (2002, S. 434) jene Programme, die speziell für klar bestimmbar Lernzwecke entwickelt und programmiert wurden (z.B. ein Sprachlernprogramm). Dabei sieht er im Programm bereits ein bestimmtes didaktisches Konzept realisiert (z.B. Lückentests), das sowohl einen ganz bestimmten Lerninhalt (z.B. französische Grammatik) zum Gegenstand hat und sich auf eine mehr oder weniger klar definierte Zielgruppe (z.B. Französisch-Anfänger) ausrichtet. In der Lernsoftware ist seiner Meinung nach der hauptsächliche Verwendungszweck des Programms von den Autoren bereits weitgehend festgelegt worden: allgemeine Informationsvermittlung (z.B. Software über verschiedene Vogelarten und ihre Lebensräume), Tutorials (z.B. Software, um das Kartenlesen zu lernen) oder Übungsprogramme (z.B. Software für das 1x1-Training).

Gängige Beurteilungsverfahren

Es gibt unterschiedliche Beurteilungsmöglichkeiten für Lernsoftware. Gängig sind Rezensionen, die eine Software beschreiben und einschätzen.² Der eigentliche Sinn dieser Softwarebesprechungen liegt im Verarbeiten von subjektiven Erfahrungen und Einschätzungen. Erwartet wird weder eine vollständige Inhaltsangabe noch ein «objektives» Urteil. Eine weitere Möglichkeit zur Beur-

teilung von Lernsoftware ist die Verwendung von Kriterien³. Nach Baumgartner (2002, S. 428) lassen sich verschiedene Arten von kriterienbasierten Instrumenten unterscheiden:

- Kriterienkataloge: Aufstellung von einzelnen Items oder Itemgruppen, die mit oder ohne Gewichtung versehen sind.
- Checklisten: Kataloge von Items, deren Erfüllung durch einfaches Abhaken registriert wird.
- Anforderungskataloge: Zusammenstellung von Kriterien, deren Erfüllung qualitativ zu beschreiben ist.

Lernsoftware-Evaluation

Bei der Evaluation mittels Kriterienkatalog werden zuerst die Kriterien entwickelt. Anhand dieser wird die Lernsoftware eingeschätzt und mittels Gewichtung oder Bewertung der gewonnenen Datenbasis wird die Qualitätsbeurteilung vorgenommen. Dieses Verfahren kann sowohl von Expertinnen als auch von Praktikern angewendet werden. Wie der Forschungsstand⁴ zeigt, bieten kriterienbasierte Bewertungen gewisse Vorteile, weil sie kostengünstig sind und die Beurteilung unabhängig vom realen Lerngeschehen erfolgen kann. Zudem erscheint das Verfahren durch das schrittweise Abarbeiten der Kriterien objektiv, valide und methodisch nachvollziehbar. Gleichzeitig weisen sie aber auch gravie-

rende Mängel auf, denn das Vorkommen einzelner Items sagt über die Qualität der Lernsoftware noch nicht sehr viel aus. Oft fehlen theoretisch fundierte Bewertungs- oder Gewichtungsverfahren. Zudem schwingen bei der Bewertung stets implizite und explizite Zielvorstellungen und Verwendungskontexte mit und letztlich vernachlässigt eine kontextferne Beurteilung wichtige Faktoren, die für das Lernen relevant sind.

Im Evaluationsprojekt des Pestalozzi-Anums, das im Auftrag der kantonalen Bildungsdirektion seit 1997 Lernsoftware auf ihre Unterrichtstauglichkeit hin beurteilt, wurde angestrebt, diese möglichst praxisnah zu bewerten und Lehrkräften eine «sachkundige Orientierungshilfe für den Einsatz von Lernsoftware in der Schule anzubieten» (Pfyffer, 2001). Berücksichtigt wurden neben Tutorials oder Übungsprogrammen auch Lernspiele und Spielgeschichten oder elektronische Nachschlagewerke. Die Beurteilung erfolgte kriterienbasiert mittels Beurteilungsbogen und offenen Antwortfeldern. Letztere boten die Möglichkeit, subjektive Erfahrungen und Einschätzungen zu beschreiben. Die Evaluierenden waren hauptsächlich Lehrkräfte, welche die zu beurteilende Lernsoftware im Unterricht einsetzten, was die Praxisnähe der Evaluation zusätzlich unterstreicht. Angesichts der eingeschränkten finanziellen Dotierung des

Projektes bestand keine Möglichkeit einer Schulung der Evaluator/innen. Jede Lernsoftware wurde jeweils durch eine einzige Person evaluiert, was in der Folge einen einzelnen Datensatz generierte. Insgesamt wurden 97 Datensätze aus den Bereichen Sprache (23), Mathematik (26), Mensch und Umwelt (23), Gestaltung und Musik (5) und Fremdsprachen (12) erstellt. Hauptsächlich wurden Übungsprogramme (44) und Lernprogramme/Tutorials (20) evaluiert. Weitere Programme wie Nachschlagewerke, Lernspiele und Spielgeschichten machten etwa einen Drittel der evaluierten Lernsoftware aus (30). Aus der Kategorie der Autorenprogramme oder Werkzeuge wurden nur drei Programme berücksichtigt. Simulationen oder Mikrowelten fehlen in der Evaluation.⁵

Ergebnisse der Auswertung des bestehenden Evaluationsprojektes

Mit dem Ziel, Erkenntnisse für die Weiterentwicklung des Projekts zu generieren, wurden die Ergebnisse der Evaluation – die vorliegenden 97 Datensätze – nach Vorzügen und Problemen bei der Anwendung des Kriterienkatalogs und der offenen Antwortmöglichkeiten untersucht.⁶

Positiv fällt auf, dass die Beurteilung direkt auf den Unterrichtseinsatz bezogen ist und dass die technischen Kriterien nicht überwiegen. Nachteile zeigen sich bei der Anwendung der Kriterien. Die entsprechende Checkliste sieht vor, die Erfüllung eines Items durch einfaches Ankreuzen zu registrieren. Unklar ist die Zuordnung bereits beim Angeben des Programmtyps und bei der Abgrenzung der Begriffe «lernen» und «üben». Das führt weiter dazu, dass die Kriterien je nach Programmtyp unterschiedlich interpretiert werden. Angaben zum Lehrplanbezug (besonders in Deutsch und Mathematik) fehlen, obwohl die Programme erfolgreich im Unterricht integriert wurden.

Die Ergebnisse zeigen, dass die Anwendung des Kriterienkatalogs einen grossen

Interpretationsspielraum zulässt, da keine Angaben über Indikatoren vorliegen, welche das Vorkommen eines bestimmten Sachverhaltes anzeigen. Die Evaluierenden müssen jeweils aus eigenem Ermessen beurteilen, ob und ab wann ein Kriterium zutrifft oder nicht. Neben den Anwendungsproblemen bei den Kriterien erscheint die Bewertung durch Notengebung als willkürlich, da sie nicht auf die Kriterien abgestimmt ist. Es erstaunt, dass die Notenvergabe vor allem auch für den Bereich Didaktik sehr gut ausfällt, obwohl der gewünschte Lehrmittelbezug weitgehend fehlt.⁷

Die Möglichkeit, anhand offener Antwortfelder die eigenen Erfahrungen und Einschätzungen zu beschreiben, wurde von den Evaluierenden intensiv und kompetent genutzt: von Beobachtungen bezüglich Wirkung und Einsatzmöglichkeiten bis hin zu Beschreibungen von zusammenhängenden Kriterien im Lernprozess. Die Aussagen zu Motivationseffekten, Lernaktivitäten oder Einsatzmöglichkeiten der Lernsoftware eröffnen neben der Produktebeurteilung eine zweite Ebene in der Evaluation. Es handelt sich dabei um Aussagen zur Wirkung der Software.

Überlegungen zur Umsetzung der gewonnenen Erkenntnisse

Wie die Auswertung des bestehenden Projektes aufzeigt, sind die beschriebenen Probleme nicht einfach durch Ergänzungen oder Verfeinerungen des Kriterienkatalogs zu beheben. Es braucht eine grundlegende Neukonzeption der Evaluation.

Wichtig ist die Erkenntnis, dass mittels einheitlichem Kriterienkatalog die verschiedenen Arten von Lernsoftware nicht adäquat beurteilt werden können. Das bedeutet, dass für jeden Programmtyp die entsprechenden Qualitätskriterien individuell formuliert werden müssen. Dafür muss eine klar definierte Programmtypologie entwickelt werden. Wichtig ist auch eine sorgfältige Ausarbeitung des Lehrplanbezugs sowohl in

Links zu Datenbanken mit Lernsoftware-Beurteilungen:

www.feibel.de

Die Homepage des Journalisten und Medienexperten Thomas Feibel bietet eine Datenbank mit Beurteilungen zu «Kindersoftware», Lernsoftware und Internetseiten. Die Einträge basieren auf Feibels Rezensionen aus den Rätegebern ab 1999.

www.i-CD-ROM.de

Die standardisierte Lernsoftwarebeurteilung des Staatsinstituts für Schulpädagogik und Bildungsforschung bietet Recherchiermöglichkeiten für CD-ROMs und DVDs. Die Evaluation besteht einerseits durch Offenlegung der Bewertungen als auch der zugrunde liegenden Kriterien, und bietet dadurch eine transparente Entscheidungshilfe für die Softwareauswahl.

www.infvo.ch

Die Datenbank bietet stufen- und fachspezifisch geordnete PDF-Dateien zu Lernsoftware mit Inhaltsbeschreibungen und Einsatzmöglichkeiten zum Herunterladen an. Diese Informationsquellen sind hilfreich im direkten Einsatz der Lernsoftware.

www.medienbildung.ch/lernsoftware

Die Datenbank des Pestalozzianums umfasst zur Zeit 125 evaluierte Lernprogramme mit Produktbeschreibungen, Bewertungen und Praxisberichten.

www.sodis.de

Die SODIS Datenbank ist ein Gemeinschaftsprojekt der deutschen Länder. Als Dokumentations- und Informationssystem bietet sie einen umfassenden Überblick über neue Medien und bewertet anhand ausführlicher Beschreibungen einen grossen Teil der Produkte unter Kriterien der Nutzbarkeit im Unterricht.

fachspezifischer Beziehung als auch in Richtung Schlüsselkompetenzen. Zudem müssen Auswahlkriterien für die zu evaluierende Lernsoftware vorliegen.

Nur so können auch Programme berücksichtigt werden, die im Marktangebot nicht so präsent sind, oder gezielt solche Programmtypen evaluiert werden, deren Evaluation etwas aufwändiger ist, wie z.B. Simulationen, Mikrowelten und Autorenprogramme.

Unumgänglich ist zudem eine Instruktion oder Schulung der evaluierenden Personen in der Anwendung der Evaluationsinstrumente.

Interessant ist das Ergebnis, dass die Lehrpersonen die Lernsoftware überdurchschnittlich positiv beurteilten. Da in den Kommentaren der gut bewerteten Produkte vor allem auch beobachtbare Aktivitäten und Motivationseffekte beim Lernsoftwareeinsatz im Unterricht beschrieben wurden, könnte es sein, dass die erfolgreiche Anwendung einer Lernsoftware im Unterricht die Entscheidung bezüglich Qualität des Programms bereits grundsätzlich in Richtung positiver Beurteilung beeinflusst hat. Für eine gewisse Transparenz bei der Qualitätsermittlung ist die Berücksichtigung von Leistungsstandards unumgänglich. Eine Ermittlung von Basis-kriterien (necessitata) würde eine Vorselektion der zu evaluierenden Lernsoftware ermöglichen. Die Qualität würde dann ausschliesslich über die Vorzüge (desiderata) eines Programmes ermittelt. Das Entwicklungspotenzial würde sich in den Zielvorstellungen (ideals) abzeichnen und nur dann eine sehr hohe Bewertung rechtfertigen. Wichtig ist, dass diese Leistungsstandards theoretisch fundiert sind und für jeden Programmtyp individuell formuliert werden. So wäre z.B. die Möglichkeit zur Lern-diagnose bei einem Tutorial aus konstruktivistischer Sicht als Basisfunktion aufzufassen, während dies aus behavioristischer Perspektive wohl eher als Vorzug eines Programms betrachtet werden müsste.

Möglichkeiten einer Verbindung von Objektivität und Praxisnähe

Wie ist nun im Spannungsfeld von Ob-

jektivität und Praxisnähe die Frage nach einer sinnvollen Form von Lernsoftware-Evaluation zu beantworten? Ziel der Evaluation ist die Schaffung einer Bewertungsgrundlage, die für Lehrpersonen eine geeignete Orientierungsmöglichkeit im breiten Angebot von unterschiedlichen Produkten darstellt.

Dies soll auf zwei Ebenen erfolgen. Wie in der Definition von Baumgartner bereits angedeutet ist, wird Lernsoftware speziell für Lehr- und Lernzwecke konzipiert. Dabei liegt die didaktische Komponente im Produkt selbst. Bei der Beurteilung einer Lernsoftware geht es um die Beschreibung dieser der Lernsoftware zugrunde liegenden Didaktik und der Beurteilung der Umsetzung. Die Lernsoftware ist also von dieser Intention her zu beurteilen. Die didaktischen Wunschvorstellungen können dann über die Auswahl der Produkte einfließen.

Die didaktischen Komponenten, die nicht lernsoftwareimmanent sind, z.B. Annahmen über Lernaktivitäten, Motivationseffekte oder Einsatzmöglichkeiten, müssen klar als solche ausgewiesen werden. Sie gelten vorerst nicht als Beurteilungskriterien, sondern als Hinweise für die Verwendung der Lernsoftware, die sich in der Praxis erst bewähren müssen. Die Beurteilung der Wirkungen, die der Einsatz einer Lernsoftware erzielt, könnte in einem zweiten Schritt durch leitfadengenerierte Praxisberichte erreicht werden. Praxisberichte sind methodisch sinnvoll, um die Erfahrungen beim konkreten Einsatz der Programme in die Evaluation zu integrieren.

Die Qualitätsbeurteilung des Produktes ist von der Wirkungsweise der Lernsoftware klar zu trennen. In der Trennung von lernsoftwareimmanenten und externen Faktoren und der Kombination von standardisierter Beurteilung und leitfadengenerierten Praxisberichten liegt meines Erachtens das Potenzial einer praxisnahen Evaluation von Bildungssoftware, welche die pädagogischen Anforderungen fokussiert. Damit sollte der Anspruch nach einer geeigneten Orien-

tierungsmöglichkeit für Lehrpersonen im breiten Angebot von unterschiedlichen Produkten eingelöst werden können.

- 1 Das Forschungsvorhaben wird im Rahmen der internen Projektförderung der Pädagogischen Hochschule Zürich realisiert und trägt den Arbeitstitel: «Lernsoftware-Evaluation. Theoretische Fundierung und Weiterentwicklung des Evaluationsprojekts des Pestalozzianums».
- 2 Die bekanntesten Beurteilungen anhand von Rezensionen sind die Lern- und Kindersoftware-Ratgeber von Thomas Feibel (s. auch den Link oder die Buchbesprechung in diesem Heft).
- 3 Aktuelle kriteriengestützte Lernsoftwarebeurteilungen bieten die Datenbanken SODIS oder i-CD-ROM (s. Hinweise zu den entsprechenden Links).
- 4 Zum Forschungsstand im Bereich Lernsoftwareevaluation s. Fricke, 2000 oder Tergan, 2001.
- 5 Die Datenbank ermöglicht eine Abfrage der evaluierten Lernsoftware übers Internet. Ein Angebot, das nach wie vor rege genutzt wird.
- 6 Für eine ausführliche Darstellung der Ergebnisse s. Biffi, 2002.
- 7 Wie die Evaluationsergebnisse des zürcherischen Schulversuchs «Schulprojekt 21» zum computergestützten Lernen zeigen, wird der Vorbereitungsaufwand für den Computereinsatz im Unterricht von den Lehrpersonen auch nach zwei Jahren als «hoch» angegeben. Dieser Aufwand wird vor allem auf das Fehlen von geeigneter Lernsoftware zurückgeführt. Als geeignete Lernsoftware bezeichnen die Lehrpersonen diejenigen Produkte, welche zu den Lehrmitteln passen. Hilfreich wäre aus Sicht der Lehrkräfte zudem die Möglichkeit, auf einen Fundus an Ideen für Lehr-/Lernaufträge zum computergestützten Unterricht und den Einsatz von Lernsoftware zurückgreifen zu können (vgl. Büeler et al., 2001).

Literatur

- Baumgartner, Peter. (2002). «Pädagogische Anforderungen für die Bewertung und Auswahl von Lernsoftware.» In: L. J. Issing & P. Klimsa (Hrsg.). *Information und Lernen mit Multimedia und Internet*. Lehrbuch für Studium und Praxis. 3. vollst. überarb. Aufl. Weinheim: Beltz. S. 427–442
- Biffi, Cornelia. «Evaluation von Bildungssoftware im Spannungsfeld von Objektivität und praktischer Anwendung.» *Medienpädagogik* 8.5.2002. <<http://www.medienpaed.com/02-1/biffi1.pdf>>
- Büeler, X.; Stebler, R.; Stöckli, G.; Stolz, D. (2001). Schulprojekt 21. Lernen für das 21. Jahrhundert? Externe wissenschaftliche Evaluation. Schlussbericht zuhanden der Bildungsdirektion des Kantons Zürich. S. 65–111; 242f. <<http://www.schulprojekt21.ch>>
- Fricke, R. (2000). «Qualitätsbeurteilung durch Kriterienkataloge. Auf der Suche nach validen Vorhersagemodellen.» In: P. Schenkel et al. (Hrsg.). *Qualitätsbeurteilung multimedialer Lern- und Informationssysteme*. Evaluationsmethoden auf dem Prüfstand. Nürnberg: Bildung und Wissen. S. 75–88
- Pfytter, A. (2001). *Evaluation Lernsoftware im Unterricht. Didaktische Überlegungen und Produkteempfehlungen*. Werkstatt und Projektberichte. Zürich: Pestalozzianum
- Tergan, S.-O. (2001). «Qualitätsbeurteilung von Bildungssoftware mittels Kriterienkatalogen. Problemaufriss und Perspektiven.» In: *Zeitschrift für Lernforschung*, 29. Jg., Heft 4, S. 319–341

Cornelia Biffi ist Dozentin im Bereich Medienbildung an der Pädagogischen Hochschule Zürich.