

Martin STEIN, Münster

Online-Plattformen zum Üben im Fach Mathematik im deutsch- und englischsprachigen Raum – ein systematischer Vergleich

1. Das Eva-CBTM-Projekt

Bisher fehlt in der wissenschaftlichen Literatur zur Evaluation von online-basierter mathematischer Lern-Software ein fundiertes System, das insbesondere den Prozesscharakter mathematischen Arbeitens berücksichtigt. Das Eva-CBTM-Projekt (Evaluation of Computer Based Training Programs for Mathematics) geht systematisch daran, diese Lücke zu schließen. Hierfür werden Kataloge mit Punktwertungen für die Evaluation mathematischer Übungsprogramme entwickelt, die den spezifischen Anforderungen mathematischen Arbeitens und Übens und seiner Dynamik gerecht werden. Diese Kataloge werden in dieser Arbeit für eine Punktwertung der auf dem Markt befindlichen kommerziellen wie nicht-kommerziellen Systeme genutzt, die damit unter vielen Gesichtspunkten systematisch verglichen werden können.

2. Aspekte der Beurteilung von Software in Eva-CBTM

Schüler mit Problemen in Mathematik benötigen Rückmeldungen über ihren Bearbeitungserfolg und in den meisten Fällen Hilfe. Jede Übungs-Software benötigt somit ein

- (1) *Bewertungssystem* sowie ein *Hilfesystem*. Beide Komponenten arbeiten in vielen Fällen eng zusammen, z.B. dann, wenn das Bewertungssystem einen Fehler erkennt und dann eine Hilfe anstößt. Auch wenn Übungs-Software in Hinsicht auf die internen Bewertungsabläufe für den Anwender eine *Blackbox* darstellt und die Arbeit des Bewertungssystems nur aus den Rückmeldungen des Systems an den Nutzer erschlossen werden kann, ist es doch sinnvoll, beide Komponenten getrennt zu beschreiben.

Ferner benötigen alle Systeme einen *Aufgabenpool*, der in irgendeiner Weise *anzuordnen und zu strukturieren* ist. Arbeiten Anwender an den Aufgaben und haben Erfolg oder Misserfolg, muss das System die nächste Aufgabe oder Aufgabensequenz zur Verfügung stellen. Hierbei nutzt es die Systemstruktur. Wir haben somit

- (2) *Systemstruktur* und *Aufgaben-Auswahlssystem* als weitere zu beurteilende Charakteristika eines CBTM-Systems.

Systeme unterscheiden sich auch erheblich in der Frage, in wie weit sie alle Entscheidungen selbst übernehmen oder diese dem Nutzer überlassen. Dies wird durch die

(3) *Freiheitsgrade*

gemessen.

Schließlich ist zu unterscheiden, wie vollständig der behandelte Stoff abgedeckt wird – ein System, das nur wenige Themen abdeckt, kann offensichtlich mehr Ressourcen auf eine möglichst ausgefeilte Erstellung seiner Komponenten legen als eines, das versucht, alle inhaltlichen Facetten abzudecken. Der Extremfall wäre der eines *Prototypen*, mit dem anhand einer einzelnen Aufgabe alle Raffinessen entwickelt werden, die im Idealfall eingesetzt werden könnten. Dieser Aspekt wird in der

(4) *thematischen Vollständigkeit* gemessen.

Die *thematische Vollständigkeit* wird bei der Evaluation eine besondere Rolle spielen. Hier wird ein Prozentwert für die mehr oder weniger vollständige Abdeckung des Stoffes ermittelt. Die Anwendung dieses Prozentwertes auf die vorher ermittelten einzelnen Punktwertungen der betrachteten Systeme ergibt die abschließende Evaluation der Systeme.

Das zu entwickelnde Punktwertungssystem soll bei der Vielzahl der mittlerweile erhältlichen Angebote an Übungsprogrammen einen möglichst objektiven Vergleich ermöglichen. Die üblichen Kriterien der *usability* und die üblichen Methoden einer qualitativen Evaluation werden in dieser Studie nicht behandelt – hierzu gibt es viele Kriterienkataloge, auf deren Grundlage eine qualitative Evaluation möglich ist. Das vorgelegte System kann allerdings dazu dienen, aus der Vielzahl der Angebote eine rational begründete *Vorauswahl* zu treffen

3. Ein Prozessmodell für das Zusammenspiel von Nutzer- und Systemaktivitäten

Die bestehenden Kriterienkataloge für die Beurteilung von Lern- und Übungs-Software haben eine globale und statische Sicht auf den Arbeitsprozess gemeinsam: sie fragen: „*Passt sich das System den Nutzeraktivitäten an?*“ „*Stellt das System ausreichende Hilfen bereit?*“ usw. Dabei bleibt unbeachtet, dass die denkbaren Antworten davon abhängen, an welchem Punkt des Bearbeitungsprozesses sich Nutzer und System befinden: so muss eine Hilfe, die vom System *vor* dem Beginn der Arbeit gegeben wird, notwendigerweise anders ausfallen als eine, die nach Abschluss der Arbeit erfolgt. Wird in einem Kriterienkatalog also global danach gefragt, ob das System genügend Hilfen zur Verfügung stellt, wird vom Beurteiler verlangt, dass er mehr oder weniger intuitiv eine zusammenfassende Sicht al-

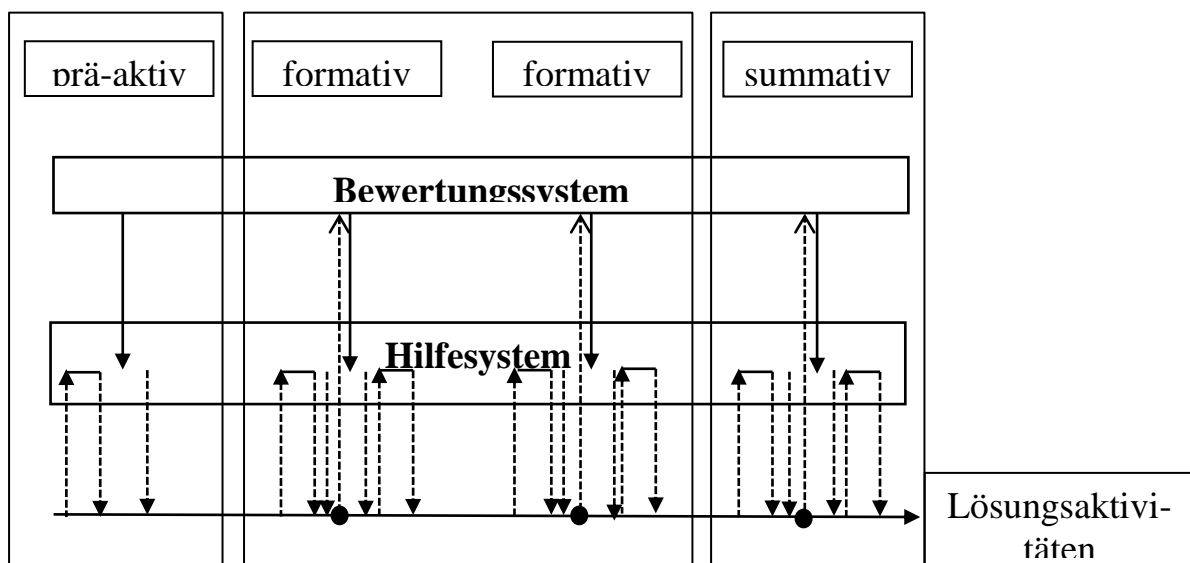
ler „Hilfeelebnisse“ erstellt. Objektivierbar bzw. zuverlässig durch andere Beurteiler überprüfbar kann eine auf dieser Grundlage erfolgte Bewertung kaum sein. Für Eva-CBTM wurde deshalb ein Prozessmodell entwickelt, das berücksichtigt, dass die Arbeit der Übenden / Lernenden ein *Prozess* ist. Dabei verläuft der Arbeitsprozess längs einer Zeitachse, deren Beginn bei der Präsentation der Aufgabe liegt und deren Ende mit dem Drücken der Eingabe-Taste nach eingetragener Lösung, eines Lösungs-Buttons etc. markiert ist.

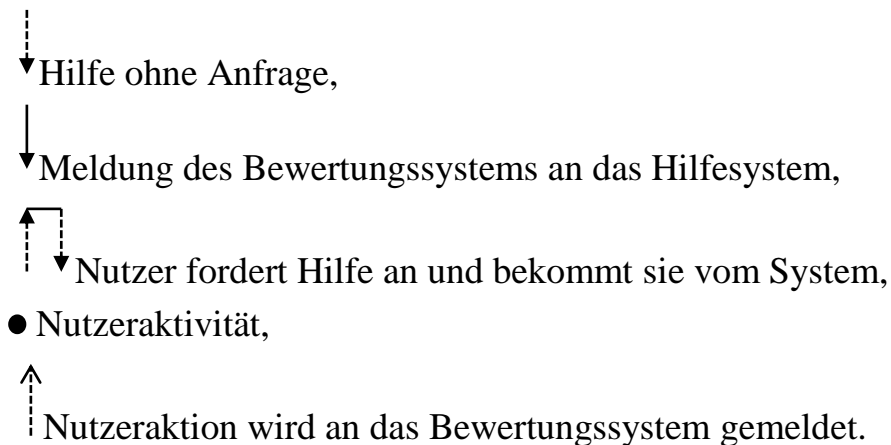
Die Bewertung der Lösungen sowie helfende Informationen können somit an drei Stellen ansetzen:

Prä-aktiv: Vor Beginn des Arbeitsprozesses – vor Beginn der Lösungsaktivitäten kann offensichtlich noch keine Nutzereingabe bewertet werden – *Hilfen* sind aber schon zu diesem Zeitpunkt möglich. Ebenso ist denkbar, dass das Bewertungssystem Informationen über frühere Nutzeraktivitäten – Erfolge und Misserfolge – bereit-stellt.

Formativ: Im Verlauf des Arbeitsprozesses – die Bewertung der Nutzeraktivitäten erfolgt auf Grund seiner Eingaben. Dieser zeitliche Verlauf unterliegt grundsätzlich jeglicher Lösungsaktivität – ob das System davon etwas „bemerkt“, ist ein zweiter Punkt. Soll das System über die Zwischenschritte einer Lösung informiert werden, muss es die Eingabe eben dieser Zwischenschritte ermöglichen. Nur in diesem Falle kann von einer formativen Bewertung / Hilfe gesprochen werden.

Summativ: Zum Abschluss des Arbeitsprozesses – dies ist die Standardform der Bewertung, die *jedes* Lehr- und Übungssystem beherrscht.





Dieses Modell kommt auf zwei Leveln zur Anwendung:

Mikro-Level: Die Bearbeitung einer einzelnen Aufgabe erfolgt grundsätzlich in mehreren Schritten. Wenn ein System entsprechend programmiert ist, kann es die einzelnen Schritte der Anwender erfassen und entsprechend reagieren.

Makro-Level: Übungssysteme stellen den Anwendern in aller Regel keine einzelnen Aufgaben, sondern Aufgabensequenzen zur Verfügung. Die Abfolge der Arbeit an diesen Aufgaben kann mit demselben Modell beschrieben werden.

Dabei ist zu berücksichtigen, dass gewisse Überschneidungen denkbar sind: so findet sich die *summative* Bewertung des Mikro-Levels als *formative* Bewertung des Makro-Levels wieder, denn der Abschluss einer einzelnen Aufgabe auf dem Mikro-Level ist ja ein einzelner „Punkt“ auf dem Makro-Level, wenn man die Aufgabe als Teil einer Serie betrachtet.

4. Ergebnisse

Aufgrund der Platzbeschränkung kann hier nur die Ergebnisse der fünf bestbewerteten Plattformen wiedergegeben werden. Eine vollständige Liste findet sich in Stein 2012 und Stein 2013.

	TV	G-BS	G-HS	G-SyS	G-AAS	G-FG	Summe
Khan Academy	0,44	15	52	13	17	26	124
IXL	0,69	21	55	21	28	3	128
Tenmarks	0,67	13	87	13	13	1	128
Mathegym	0,62	31	62	12	12	25	142
Bettermarks	0,65	65	164	26	65	46	367

Literatur

- Stein, M. (2012): Eva-CBTM. 2nd enlarged edition. Münster: WTM-Verlag
 Stein, M. (Hrsg.) (2013): Mathematik online. Münster: WTM-Verlag