

Michael KALLWEIT, Bochum

Der Computer als Tutor - technikbasierte Diagnostik mit Freitextaufgaben

Viele Lernplattformen ermöglichen inzwischen offene Formate für Mathematikaufgaben mit automatischer Auswertung. Die Software STACK bietet darüber hinaus intelligente Optionen für individuelles Feedback, das Lernenden und Lehrenden spezifische Rückmeldung über Fehlvorstellungen und Lernerfolge liefert. Erste Erfahrungen mit STACK-Aufgaben und ihren Feedback-Bäumen dienen zur Pilotierung einer Untersuchung über die Wirksamkeit und Nachhaltigkeit dieses Formates in verschiedenen Kontexten.

Einleitung

Technische Systeme, die für die Mathematik geeignete offene Aufgabenformate bereitstellen finden zur Zeit eine immer stärkere Verbreitung (siehe z. B. Buchsteiner, Kallweit, 2015). Das Assessmentssystem STACK (Sangwin, 2013) bietet vielfältige Möglichkeiten zur Konstruktion von digitalen Mathematikaufgaben mit automatischer Bewertung durch ein Computeralgebrasystem und individuell generiertem Feedback (siehe auch Kallweit, 2015).

Der Einsatz von STACK als Bestandteil regulärer Veranstaltungen wurde an der Ruhr-Universität Bochum (RUB) im WS 2015/2016 in zwei Szenarien erprobt und ausgewertet.

Einsatz im Eingangstest

In zwei Präsenzvorkursen für Mathematik wurde ein freiwilliger digitaler Eingangstest durchgeführt an dem insgesamt 146 Studienanfänger teilgenommen haben (48 aus dem Vorkurs für Mathematiker und Physiker, 98 aus dem Vorkurs für Ingenieure). Basis für die Aufgaben war ein Eingangstest der Westfälischen Wilhelms-Universität Münster (WWU), der im dortigen Vorkurs von 191 Teilnehmer durchgeführt wurde (vgl. Neugebauer 2013). Dieser Test bestand aus 25 MultipleChoice-Aufgaben, deren Distraktoren auf empirisch ermittelten Falschantworten und dahinterliegenden Fehlerkonzepten beruhen (Neugebauer & Winter, 2014). Pro Frage gab es mindestens 5, im Mittel 6,5 und maximal 9 verschiedene Antwortmöglichkeiten, womit sich insgesamt 145 Distraktoren ergaben. Hinzu kam zu jeder Aufgabe die Auswahlmöglichkeit „Meine Antwort ist nicht dabei“, welche von allen Testeingaben aber nur 69 (1,6%) mal ausgewählt wurde.

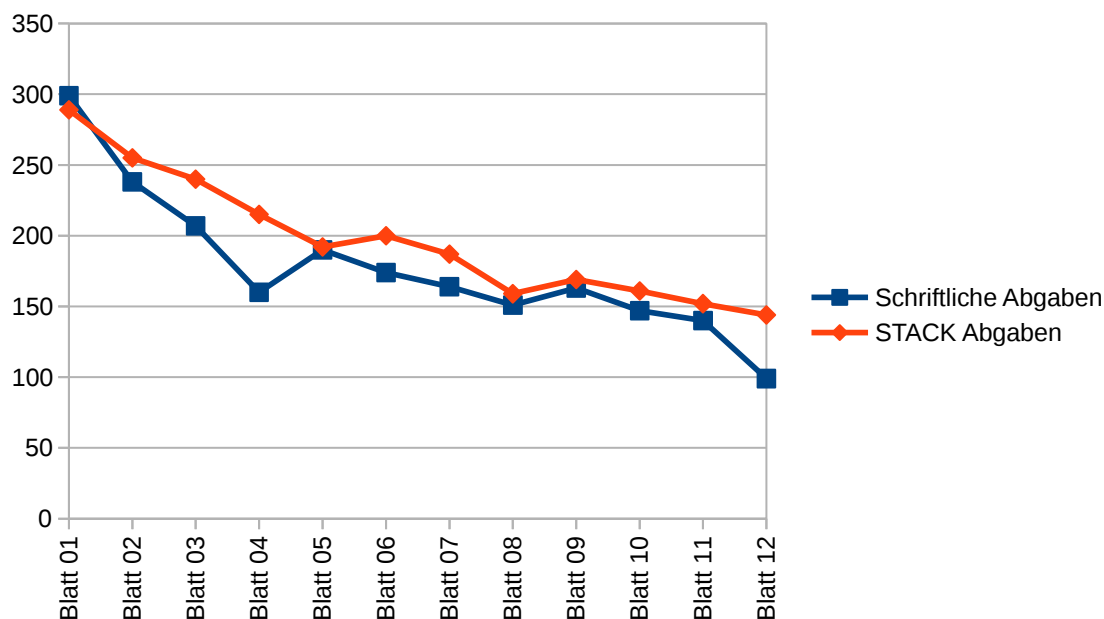
Für die Variante an der RUB wurden 23 der insgesamt 25 Aufgaben in das offene Aufgabenformat mit STACK überführt und die Auswertungs-/Feedbackbäume anhand der MultipleChoice-Distraktoren konstruiert. Durchschnittlich wurden hier 15 verschiedene Eingaben pro Aufgabe erfasst, was sich insgesamt zu 347 summierte. Ein Extrembeispiel zeigt eine Aufgabe zu Termumformungen: Hier wurden 9 verschiedene Distraktoren hinterlegt, aber 41 unterschiedliche Lösungsterme von den Teilnehmern eingegeben.

Test mit MultipleChoice (WWU)	Test mit STACK (RUB)
Ø 6 Distraktoren	Ø 15 vers. Eingaben
max. 9 Distraktoren	max. 41 vers. Eingaben
insg. 145 Distraktoren	insg. 347 vers. Eingaben
1,6% der Auswahlen sind „meine Lösung ist nicht dabei,“	>60% der Eingaben sind neue Eingaben

(Tabelle 1: Vergleich der Auswahlen bzw. Eingaben des Eingabetests)

Einsatz in den wöchentlichen Übungsaufgaben

Als zweites Szenario wurden digitale Aufgaben mit STACK als Teil der regulären wöchentlichen Übungsaufgaben der Veranstaltung Lineare Algebra I gestellt. Pro Übungsblatt gab es 3 auf Papier zu lösende Aufgaben und eine digitale Aufgabe in der eLearning-Plattform Moodle. Für die Bearbeitung beider Aufgabentypen hatten die Studierenden eine Woche Zeit. Für korrekte Lösungen gab es einen Bonus für die Modulpunkte. Der Großteil der Hörer nahm an den Übungen teil und bearbeitete Aufgaben. Die Anzahl der Abgaben sank (von 299 Studierende anfangs auf ca. 150 am Ende) über den Verlauf des Semesters (siehe Abbildung 2), wobei die elektronischen Aufgaben generell (bis auf die erste Woche) häufiger bearbeitet wurden, was die Akzeptanz dieses Aufgabenformats bei den Studierenden unterstreicht. Erfahrungsberichte zeigen, dass anfängliche Probleme mit dem technischen System, wie z. B. das Erlernen der Eingabesyntax, nach kurzer Zeit nur noch in sehr seltenen Fällen auftraten. In einem Fall musste eine Aufgabe durch Fehler bei der Konstruktion nochmals nachgebessert werden. Dies war aber ohne weitere Interaktion mit den Studierenden möglich, da STACK das nochmalige Auswerten aller Aufgaben unterstützt.



(Abbildung 2: Abgabezahlen nach Aufgabentyp)

Als Aufgaben wurden klassische Fragestellungen als Grundlage herangezogen, hierbei aber versucht die neuen Möglichkeiten des offenen Formats auszuschöpfen. So wurde u. a. die Generierung von Beispielen eines mathematischen Sachverhalts abgefragt (siehe Abbildung 3). STACK ermöglicht einen Wechsel des Auswertungsparadigmas von „Welcher Fehler wurde gemacht“ (Distraktoren bei MultipleChoice) zu „Was wurde richtig gemacht“.

<p>Sei $\varphi : (\mathbb{R}, +) \rightarrow (\mathbb{R} \setminus \{0\})$ mit $\varphi(x) = e^x$. Zeigen Sie, dass φ ein Gruppenhomomorphismus ist.</p>	↔	<p>Geben Sie einen Gruppenhomomorphismus $\varphi : (\mathbb{R}, +) \rightarrow (\mathbb{R} \setminus \{0\})$ an. $\varphi(x) =$ <input style="width: 150px; height: 20px;" type="text"/></p>
--	---	--

(Abbildung 3: Vergleich Papieraufgabe und STACK-Aufgabe)

Setzt man die erreichten Punktzahlen bei den Papier- und den STACK-Aufgaben zueinander und zum erreichten Klausurergebnis in Beziehung, ergeben sich signifikante Korrelationen (Tabelle 4).

Papier ↔ STACK	r=0,902	p<0,0001
Papier ↔ Klausur	r=0,556	p<0,0001
STACK ↔ Klausur	r=0,625	p<0,0001

(Tabelle 4: Korrelationen nach Pearson)

Fazit

Der offene mit STACK umgesetzte Aufgabentyp bietet im mathematischen Kontext eine Bereicherung der Überprüfungsmöglichkeiten von Wissen und Kompetenzen. Schwierigkeiten in der Nutzung durch Studierenden treten nur beim ersten Kontakt mit den Tools auf. Die Akzeptanz im praktischen Einsatz ist auf Seite der Studierenden und Lehrenden gleichermaßen hoch. Die Fokussierung auf die Überprüfung mathematischer Eigenschaften einer Eingabe durch den Nutzer bei der Erstellung der Aufgaben, ermöglicht zudem kürzere Entwicklungsprozesse, da nicht zwingend sämtliche möglichen Fehler im Vorfeld überlegt werden müssen. Dennoch lassen sich gerade durch das offene Format weitere Daten für eine didaktische Analyse sammeln. Der Einsatz von STACK in diagnostischen Tests in Kombination mit diesen vorher ermittelten Fehleraufklärungen ist ein vielversprechender Ansatz, der in Zukunft weiter erforscht werden soll.

Literatur

- Buchsteiner, J. , Kallweit, M. (2015): *Professionalisierung des Helpdesk Mathematik*. In H. Linneweber-Lammerskitten (Hrsg.), *Beiträge zum Mathematikunterricht 2015*, 220-223. Münster: WTM.
- Kallweit, M. (2015). *Mathematik-Kompetenzen überprüfen und fördern – Automatisiert Lehren und Lernen mit STACK*. Workshop der ASIM/GI-Fachgruppen, Argesim Report AR 50.
- Neugebauer, C., Winter, K. (2014): *Fehleranalysen bei Studienanfängern als Basis zur individuellen Förderung in Mathematik*. In: Roth, J., Ames, J. (Hrsg.) (2014): *Beiträge zum Mathematikunterricht*, Münster: WTM-Verlag, S. 851-854.
- Neugebauer, C. 2013: *Online - Test zum Self-Assessment im Themenfeld "Studierfähigkeit in Mathematik": Zur Entwicklung von Multiple-Choice-Items*. In: Hoppenbrock, A., Schreiber, S., et al. (2013): *Mathematik im Übergang Schule/Hochschule und im ersten Studienjahr*.
- Sangwin, C. J. (2013). *Computer Aided Assessment of Mathematics*. Oxford University Press