Juliane LEUDERS, Timo LEUDERS, Freiburg

Diagnostische Kompetenzen von Lehramtsstudierenden bei der Beurteilung von Schülerlösungen

Unter den vielen Aufgaben, die Lehrerinnen und Lehrern in ihrer Praxis begegnen, kommt diagnostischen Aktivitäten eine wichtige Rolle zu. Solche Aktivitäten beinhalten das Sammeln und Interpretieren von Daten, sei es aus formalen Tests, durch Beobachtung, die Analyse von Schülerprodukten oder das Führen und Auswerten Gesprächen mit Lernenden. Das Ziel von Diagnose im Unterricht ist die Erhebung von validen Informationen über Leistungen und Schwierigkeiten von einzelnen Lernenden oder auch der ganzen Klasse. Wissen, Fähigkeiten und Einstellungen der Lehrenden, die mit diesen diagnostischen Aktivitäten in Verbindung stehen, können als diagnostische Kompetenz bezeichnet werden (s. Leuders, Leuders & Philipp, 2014).

Zielsetzung

In unserer Studie untersuchen wir eine eng umrissene Facette von diagnostischer Kompetenz. Wir fokussieren auf die Analyse von Schülerlösungen zu Lernumgebungen, also auf die post-aktionale Phase, in der kein Zeitdruck herrscht. Im theoretischen Rahmen der Michigan-Group (z.B. Hill, Shilling & Ball, 2004) lässt sich dieser Aspekt als *knowledge of content and students* (KCS) fassen. Dadurch entsteht ein Instrument, das diagnostische Kompetenz in einer typischen Praxissituation erfasst. Um auch den Einfluss von fachwissenschaftlichen Fähigkeiten einzubeziehen, untersuchen wir die Auswirkung von eigenen Lösungsversuchen zu den konkreten Aufgaben und die Bedeutung des fachwissenschaftlichen Vorwissens. Unsere Fragestellungen sind: (1) Liefert das von uns entwickelte Instrument zur Erfassung diagnostischer Kompetenz reliable Ergebnisse? (2) Verbessern eigene Lösungsversuche der Studierenden ihre diagnostischen Urteile? (3) Welche weiteren Faktoren beeinflussen die diagnostische Kompetenz?

Methode

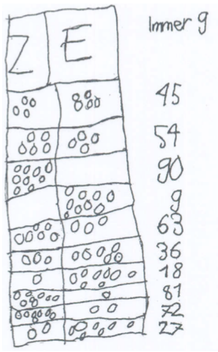
Teilnehmer: Die Hauptstudie wurde mit 110 Lehramtsstudierenden durchgeführt. Sie waren Teilnehmer einer Lehrveranstaltung zu Diagnose und Förderung im Mathematikunterricht and der Pädagogischen Hochschule Freiburg, zu der Studierende ab dem 4. Semester zugelassen sind. Der Test wurde vor Beginn der Veranstaltung durchgeführt. Die Studierenden stammen aus unterschiedlichen Studiengängen, die sich in drei Gruppen

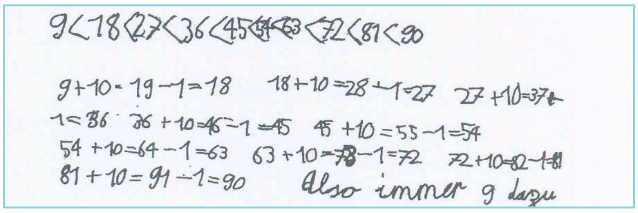
In J. Roth & J. Ames (Hrsg.), *Beiträge zum Mathematikunterricht 2014* (S. 735–738). Münster: WTM-Verlag

einordnen lassen: Primarstufe, Hauptfach Mathematik (23%); Primarstufe, fachfremd (ca. 6 SWS Mathematikdidaktik, 38%); Sekundarstufe, Haupt- oder Nebenfach Mathematik (36%).

Material: Das Instrument enthält authentische Lösungen von Zweitklässlerinnen zu zwei Lernumgebungen (Hengartner et al. 2007, Bild 1). Die Aufgaben und Lösungen sind ausreichend unterschiedlich, so dass Überschneidungen minimiert sind. Die Probanden werden gebeten, bis zu 5 kurze Aussagen zu diesen Lösungen zu notieren.

Aufgabe 1: Nimm immer 9 Plättchen. Welche Zahlen kannst du darstellen? Ordne sie nach der Größe. Wie viele verschiedene Zahlen sind es?





Instruktion für die Probanden: Tanja ist Zweitklässlerin. Beurteilen Sie Tanjas Lösung. Was fällt Ihnen auf? Formulieren Sie eine oder mehrere unterschiedliche Aussagen zu der Lösung (z.B. „Die Lösung ist...“, „Tanja hat....“, „Vielleicht wurde....“).

Bild 1: Aufgabe, Schülerlösung und Instruktion zu Aufgabe 1

Design: Die Teilnehmer wurden randomisiert auf zwei Gruppen verteilt. Gruppe A löste zunächst Aufgabe 1 selbst, analysierte dann die Schülerlösung zu Aufgabe 1 und anschließend die Schülerlösung zu Aufgabe 2, ohne diese zweite Aufgabe selbst bearbeitet zu haben. Bei Gruppe B war es umgekehrt. Zusätzlich wurde Semesterzahl, Abiturnote in Mathematik und Studiengang erfasst.

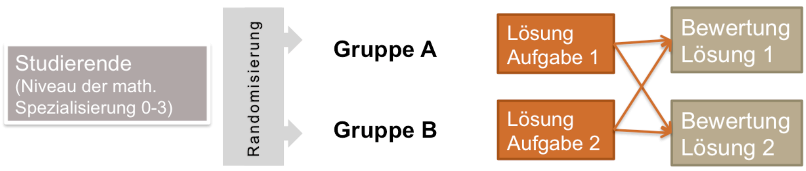


Bild 2: Studiendesign

Aus den diagnostischen Aussagen konnten induktiv ca. 20 Kategorien für jede Aufgabe entwickelt werden, die sich in 4 Typen gliedern lassen:

- spezifische Aussagen, z.B. “Ordnen nach Größe“: *„Um eine gewisse Ordnung in die Darstellung zu bekommen, hat Tanja die dargestellten Zahlen aufsteigend geordnet.“*
- unspezifische, nicht konkret auf die Lösung bezogene Aussagen, z.B. „systematisch“: *„Tanja ist sehr systematisch vorgegangen und hat alle möglichen Kombinationen gefunden.“*

- defizitorientierte Aussagen, die nur beschreiben, was in der Schülerlösung fehlt, z.B. *„Erklärung zu der Rechnung fehlt“*
- nicht-mathematische Aussagen, z.B. *„nicht mit Lineal gezeichnet“*

Ergebnisse

(1) Liefert das Instrument reliable Ergebnisse? Dies konnte bestätigt werden: Die Daten der Hauptstudie wurde durch zwei Rater kategorisiert, die im Rahmen der Pilotstudie bereits ein Ratertraining erhalten hatten. Nach dem Rating der ersten 20 Fälle ergab sich für die Interrater-Reliabilität ein Cohen's κ von 0.67. Nach einem weiteren Ratertraining konnte dies bei 23 weiteren Fällen auf Cohen's κ=0.76 gesteigert werden.

(2) Verbessern eigene Lösungsversuche die diagnostischen Urteile der Studierenden? Diagnostische Kompetenz wird hier operationalisiert als Reichhaltigkeit von diagnostischen Aussagen, gemessen als Summe von spezifischen und unspezifischen Aussagen. Aussagen, die sich nicht auf Mathematik beziehen, und defizitorientierte Aussagen, die nur beschreiben, was die Schülerlösung nicht beinhaltet, wurden nicht gewertet. Die Hypothese war, dass für die Schülerlösung zu einer selbst bearbeiteten Aufgabe reichaltigere Aussagen gemacht werden.

Für Aufgabe 1 (absteigende Linie) konnte die Hypothese bestätig werden: Die Probanden, die diese Aufgabe zuvor selbst gelöst hatten, formulierten reichhaltigere diagnostische Aussagen ($F=5,1$; $p<0,001$, partielles $\eta^2=0,12$) Für Aufgabe 2 (ansteigende Linie) zeigte sich eine ähnliche Tendenz, allerdings nicht auf signifikantem Niveau ($p=0,19$).

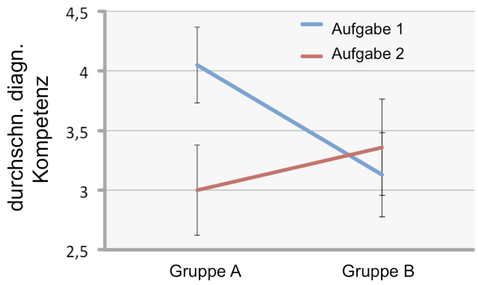


Bild 3: Effekt eigener Lösungsversuche auf die Reichhaltigkeit der diagnostischen Aussagen

(3) Welche weiteren Faktoren beeinflussen die diagnostische Kompetenz? Bezüglich des Studiengangs wäre zu erwarten, dass Primarstufenstudierende mit Hauptfach Mathematik am besten abschneiden sollten, da sie sowohl mit dem Fach als auch mit den Fähigkeiten von Zweitklässlern vertraut sind. Diese Hypothese konnte im Bezug auf Aufgabe 1 bestätigt werden: fachfremde Primarstufenstudierende schnitten schlechter ab, Sekundarstufenstudierende lagen sogar noch unter den Fachfremden. (Varianzanalyse: $F=3,56$; $p=0,017$). Es gab allerdings zwischen diesen Gruppen keine signifikante Variation des Effekts eigener Lösungen: Keine der

Gruppen profitierte mehr als die anderen von eigenen Lösungsversuchen. Bezüglich der Abiturnote zeigte sich ein marginal signifikanter Effekt in der Gruppe der fachfremden Primarstufenstudierenden: Bei einer niedrigen Abiturnote (N=17) zeigten sich auch geringere Leistungen beim Beurteilen der Schülerlösung zu einer nicht selbst bearbeiteten Aufgabe und profitierten am meisten von eigenen Lösungen (F=3,93, p=0,052).

Diskussion

Die Ergebnisse deuten darauf hin, dass sich die diagnostischen Aussagen zu den Schülerlösungen reliabel kategorisieren lassen. Für Aufgabe 1 konnte ein signifikanter Effekt eigener Lösungsversuche nachgewiesen werden. Eine genauere Analyse zeigt, dass dieser Einfluss vor allem auf einem Anstieg in den *spezifischen* Aussagen beruht. Allerdings fanden wir für Aufgabe 2 kein vergleichbares Ergebnis. Bei Betrachtung der Kategorien deutet sich hier ein Deckeneffekt an: die interessanten Aspekte der entsprechenden Schülerlösung waren leichter zu erkennen als in Aufgabe 1, auch ohne eigens Durchdringen der Aufgabe. Die Qualität von diagnostischen Urteilen zu Schülerlösungen kann also durch eigene Lösungen verbessert werden. Dies steht im Einklang mit dem Phänomen der "Dekomprimierung" bei Morris et al. (2009).

Außerdem zeigt sich, dass die Primarstufenstudierenden mit Hauptfach Mathematik auch ohne eigenes Lösen die meisten sprezifischen Aussagen formulieren, während bei fachfremden und bei Sekundarstufenstudierenden ein höherer Anteil an unspezifischen Aussagen auftritt. Da die gewählten Aufgaben relativ komplex sind, ist dies nicht überraschend. Dass die fachfremden Proimarstufenstudierenden am meisten von eigenen Lösungen profitierten (marginal signifikant) erinnert an das Ergebnis von El Mouhayar und Jurdak (2013), nach dem specialized content knowledge eine hohe Bedeutung für diagnostische Kompetenz hat.

Literatur

El Mouhayar, R.R., & Jurdak, M.E. (2013). Teachers' ability to identify and explain students' actions in near and far figural pattern generalization tasks. *Educational Studies in Mathematics*, *82*, 379–396.

Hengartner, E., Hirt, U., Wälti, B., & Primarschulteam Lupsingen (2007). *Lernumgebungen für Rechenschwache bis Hochbegabte*. Zug: Klett und Balmer.

Hill, H. C., Schilling, S. G., & Ball, D. L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *The Elementary School Journal, 105*(1), 11-30.

Morris, A. K., Hiebert, J., & Spitzer, S. M. (2009). Mathematical knowledge for teaching in planning and evaluating instruction: What can preservice teachers learn? *Journal for Research in Mathematics Education, 40*(5), 491-529.