

Elisa LANKEIT, Paderborn, Thomas BAUER, Marburg &
Rolf BIEHLER, Paderborn

Votingfragen in den Übungen zur Analysis – Wirkung verschiedener Einsatzszenarien

Die Methode der „Peer Instruction“ nach Eric Mazur (1999) wird seit den 1990er-Jahren in Physikvorlesungen und inzwischen auch vermehrt in Mathematikvorlesungen eingesetzt (z. B. Cline, Zullo, Duncan, Stewart & Snipes, 2013; Miller, Santana-Vega & Terrell, 2006; Pilzer, 2001). Dabei werden Votingfragen, sogenannte ConcepTests, zu einem fachlichen Konzept gestellt, über welche die Studierenden abstimmen. Anschließend diskutieren sie über ihre gewählten Antworten. Crouch und Mazur (2001) stellten deutlich bessere Leistungen bezüglich des konzeptuellen Verständnisses und der Rechenfertigkeiten bei mit Peer Instruction unterrichteten gegenüber traditionell unterrichteten Studierenden fest. Daran, dass die Methode von Mazur als „Peer Instruction“ bezeichnet wird, wird bereits deutlich, dass eine Betonung auf der Phase der Kleingruppendiskussion liegt, der eine wesentliche Bedeutung zugeschrieben wird. Auch Miller et al. (2006, S. 11) schreiben: „Data suggests that the benefit from Good Questions comes from the peer discussion“, ohne dies jedoch genauer zu belegen. Ein systematischer Vergleich des Einsatzes von ConcepTests mit und ohne Diskussionsphase fehlt bisher. An dieser Stelle setzt unsere Studie an.

Die von Mazur vorgeschlagene Methode der Peer Instruction wird seit mehreren Jahren vom zweiten Autor in den Präsenzübungen der Analysis eingesetzt (Bauer, 2019). Im Sommersemester 2018 wurde der Einsatz dieser Methode im Rahmen des Projekts WiGeMath (Wirkung und Gelingensbedingungen von Unterstützungsmaßnahmen für mathematikbezogenes Lernen in der Studieneingangsphase) evaluiert und in Hinblick auf deren Wirkung untersucht. Dazu wurden die bereits erstellten, erprobten und überarbeiteten ConcepTests in den Präsenzübungen in zwei verschiedenen Varianten (mit und ohne „Peer Discussion“) eingesetzt, welche im Folgenden vorgestellt werden.

Vorstellung der verschiedenen Einsatzszenarien (Varianten)

In jeder Übung war jeweils die erste halbe Stunde für die Auseinandersetzung mit den ConcepTests vorgesehen. Es wurden pro Übung drei verschiedene ConcepTests behandelt. Dabei folgte der Ablauf in der ersten Variante, die im Folgenden mit „P“ bezeichnet wird, im Wesentlichen dem Vorschlag von Mazur (1999): Nachdem den Studierenden der ConcepTest gezeigt wurde, hatten sie kurz Zeit, allein darüber nachzudenken. Dann fand eine

erste Abstimmung per Handkarten statt, in der jede/r Studierende für die Antwort, die sie oder er für richtig hielt, abstimmte. Anschließend fanden Diskussionen in Kleingruppen von zwei bis drei Personen („Peer Discussion“), die nicht alle gleich abgestimmt hatten, statt, in welcher das Ziel war, die anderen von der Richtigkeit der eigenen Antwort zu überzeugen. Daran schloss sich eine zweite Abstimmung an. Abschließend gab der/die Tutor/in die richtige Antwort bekannt und erklärte die Lösung kurz. In der zweiten Variante, die im Folgenden mit „T“ bezeichnet wird, wurde nach der ersten Abstimmung sofort angegeben, welche Antwort die richtige ist. Danach erklärte eine Person mit der korrekten Antwort, warum sie für diese abgestimmt hatte, und der/die Tutor/in ergänzte dies durch eine ausführliche Erklärung, warum die anderen Antworten falsch sind und welche Fehlvorstellungen diesen Distraktoren jeweils zugrunde lagen.

Forschungsfragen

In diesem Artikel befassen wir uns mit den folgenden Forschungsfragen:

F1) Gibt es Unterschiede in der studentischen Bewertung des Einsatzes der ConcepTests und der Wahrnehmung der eigenen Aktivität im Tutorium zwischen den beiden verschiedenen Einsatzszenarien (P und T)?

F2) Gibt es Unterschiede in der Wirkung auf die Klausurleistung beim Einsatz von Peer Instruction (P) vs. ConcepTests ohne Peer-Discussion-Phase (T)?

Methode der Datenerhebung

Die Vorlesung Analysis I wurde von Studierenden der Studiengänge Mathematik, Wirtschaftsmathematik, Physik und Gymnasiallehrer Mathematik besucht. Es wurde zufällig eingeteilt, welche Studierenden Präsenzübungen mit Variante P und welche Übungen mit Variante T besuchten. Personenbezogene Daten, Selbsteinschätzungen und Bewertungen wurden am Beginn und am Ende des Semesters jeweils mittels Fragebögen erhoben. Dabei wurden im WiGeMath-Projekt entwickelte Items und Skalen genutzt. Die Fragebögen waren mit individuellen Codes versehen, die eine Nachverfolgung möglich machen. Die Punktzahlen in den einzelnen Aufgaben der Klausur am Ende des Semesters konnten ebenfalls mit den Codes verknüpft werden.

Ergebnisse

F1) Es zeigt sich eine überwiegend positive Evaluation der ConcepTests. Die Studierenden ($n=58$) geben vor allem an, die Inhalte durch die ConcepTests besser verstanden zu haben ($M=4,8$, $SD=1,1$) und durch die ConcepTests ein Bewusstsein für den eigenen Kenntnisstand geschaffen zu haben ($M=4,8$,

$SD=0,8$, je auf einer 6-stufigen Likertskala). Unterschiede zwischen den beiden Varianten zeigen sich hier auf einem Signifikanzniveau von 10 % nur bei der Intensität der Kommunikation über die ConcepTests mit Kommilitonen ($p=0,029$, $d=0,67$) und dem Nutzen der entsprechenden Kommunikation ($p=0,045$, $d=0,55$), welche jeweils in Variante P höher sind. Die Werte in Variante P sind auch bei den anderen Skalen höher, diese Unterschiede sind nach dem nichtparametrischen Mann-Whitney-U-Test aber selbst auf einem Signifikanzniveau von 10 % nicht signifikant.

Itemtext: In der Übung...	M_P	SD_P	M_T	SD_T	p	d
...habe ich mich aktiv beteiligt.	3,7	1,5	3,8	1,4	0,82	0,08
...habe ich mit meinen Kommilitonen über Fragen des Tutors diskutiert.	4,0	1,4	3,2	1,6	0,02	-0,56
...habe ich mitgedacht.	4,8	0,7	4,8	0,8	0,91	-0,04

Tab.: Eigene Angaben der Studierenden am Ende des Semesters über die wahrgenommene eigene Aktivität in den Übungen (Variante P ($n=27$) vs. Variante T ($n=25$)), bei einer 6-stufigen Likertskala („trifft gar nicht zu“ bis „trifft vollständig zu“)

In der Selbsteinschätzung der eigenen Aktivität in der Übung (s. Tab.) unterscheiden sich die Studierendengruppen nur in der Frage, wie viel sie mit Kommilitonen über Fragen des Tutors diskutiert haben, wobei die Studierenden aus Variante P angaben, mehr mit Kommilitonen über Fragen des Tutors diskutiert zu haben, mit einem mittleren Effekt (Cohens $d=-0,56$).

F2) Die Studierenden aus Variante T schnitten im Mittel in der Klausur etwas besser ab als die Studierenden aus Variante P, die Unterschiede sind aber (in den einzelnen Aufgaben und in der Gesamtpunktzahl) selbst auf einem Signifikanzniveau von 10 % nicht signifikant. Es können also keine Unterschiede in der Klausurleistung zwischen den Studierendengruppen, die die verschiedenen Einsatzszenarien erlebt haben, festgestellt werden.

Diskussion

Die Studierenden bewerten den Einsatz der ConcepTests in den Analysisübungen insgesamt sehr positiv. Vor allem geben sie an, die Inhalte durch die Beschäftigung mit den ConcepTests besser verstanden und ein höheres Bewusstsein für ihren eigenen Kenntnisstand erlangt zu haben. Nicht einmal auf einem Signifikanzniveau von 10 % sind dabei Unterschiede zwischen den beiden Einsatzszenarien feststellbar. Ein Unterschied zeigt sich dabei nur in der Intensität und im wahrgenommenen Nutzen der Kommunikation mit Kommilitonen, welche in Variante P jeweils höher sind, was im Design

so angelegt ist. In der Klausurleistung und der selbst wahrgenommenen Aktivität in den Übungen konnten keine Unterschiede zwischen den Varianten festgestellt werden.

Man geht bei Peer Instruction von einer lernförderlichen Wirkung aus und betont dabei den Wert der Peer Discussion im Gegensatz zu instruktionalen Erklärungen der Lehrperson. Unsere Ergebnisse zeigen jedoch keine Überlegenheit der Variante P, also dem Szenario mit Peer Discussion, gegenüber dem Szenario mit ausführlichen tutoriellen Erklärungen, weder in Hinblick auf die Klausurleistung, noch in der Bewertung der Methode, noch in den Änderungswünschen der Studierenden, noch in der von ihnen wahrgenommenen aktiven Beteiligung an den Übungen.

Unser Befund deutet darauf hin, dass die Überlegenheit der Variante mit Diskussion nicht einfach angenommen werden sollte. Es kann sein, dass die intensive Befassung mit den Aufgaben an sich entscheidend ist. Der Unterschied zwischen Peer Discussion und Erklärung durch die Lehrperson ist möglicherweise weniger entscheidend als die Aufgaben selbst. Hervorzuheben ist, dass unsere Studie nicht die Wirkung von ConcepTests an sich untersuchen wollte, sondern die spezifische Bedeutung der Peer Discussion. Daher wurden in der Kontrollgruppe die ConcepTests nicht einfach entfernt, sondern dort die Peer Discussion durch eine andere Unterrichtsform ersetzt. Außerdem gingen wir im Gegensatz zu anderen Studien nicht davon aus, dass sich das konzeptuelle Verständnis zu den einzelnen Themen durch isomorphe Fragen im direkten Anschluss an die Übungen erfassen lässt. Stattdessen wurden die Leistungen in der Klausur am Ende des Semesters als zentrales „learning outcome“ verglichen.

Literatur

- Bauer, T. (2019). Peer Instruction als Instrument zur Aktivierung von Studierenden in mathematischen Übungsgruppen. *Mathematische Semesterberichte*, 66(2), 219–241. doi:10.1007/s00591-018-0225-8
- Cline, K., Zullo, H., Duncan, J., Stewart, A. & Snipes, M. (2013). Creating discussions with classroom voting in linear algebra. *International Journal of Mathematical Education in Science and Technology*, 44(8), 1131–1142. doi:10.1080/0020739x.2012.742152
- Crouch, C. H. & Mazur, E. (2001). Peer Instruction: Ten years of experience and results. *American Journal of Physics*, 69(9), 970–977. doi:10.1119/1.1374249
- Mazur, E. (1999). Peer instruction: A user's manual. In: AAPT.
- Miller, R. L., Santana-Vega, E. & Terrell, M. S. (2006). Can Good Questions and Peer Discussion Improve Calculus Instruction? *Primus*, 16(3), 193–203. doi:10.1080/10511970608984146
- Pilzer, S. (2001). Peer Instruction in Physics and Mathematics. *Primus*, 11(2), 185–192. doi:10.1080/10511970108965987