Explizites oder implizites Heurismentraining – was ist besser?

von

Benjamin Rott, Essen & Thomas Gawlick, Hannover

Kurzfassung: Gegenstand dieses Artikels ist eine Vortest-Nachtest-Studie zur Trainierbarkeit mathematischen Problemlösens mit Fokus auf fünf unterschiedliche heuristische Strategien (eine Replikation unter verbesserten Bedingungen) und zwei verschiedene Trainingsbedingungen (explizit und implizit). Es wurden 28 Studierende des gymnasialen Lehramts auf drei Gruppen aufgeteilt: Experimentalgruppe (Problemlöse- und Heurismentraining), Kontrollgruppe (nur Problemlösetraining) und Vergleichsgruppe (kein Training). Die Untersuchung zeigt eine signifikante Überlegenheit des kombinierten Trainings.

Abstract: This paper reports on a pretest-posttest study regarding the trainability of mathematical problem solving focusing on five different heuristic strategies (a replication with improved circumstances) and two different kinds of instruction (explicit and implicit). 28 student teachers were separated into three groups: an experimental group (problem-solving and heuristic training), a control group (only problem-solving training) and a comparison group (no training). The study shows a significant advantage of the combined training.

1 Einleitung

Eine Aufgabe wird gemeinhin als *Problem* für eine bestimmte Person bezeichnet, wenn diese Person kein Verfahren zur direkten Lösung dieser Aufgabe kennt (z. B. Dörner 1979). In diesem allgemeinen Verständnis ist Problemlösen dementsprechend ein relevantes Thema für nahezu alle Lebensbereiche (vgl. OECD 2014). Bezogen auf Mathematik(unterricht) geht es um das Fehlen entsprechender Lösungsverfahren oder Algorithmen beim Bearbeiten mathematischer Aufgaben.

Aufgrund seiner allgemeinen Bedeutung ist das Problemlösen schon lange ein wichtiges Beschäftigungsfeld nicht nur in der Psychologie, sondern auch in der Mathematikdidaktik, und es ist insbesondere auch für Schule und Unterricht von Interesse. So hat zum Beispiel die amerikanische Vereinigung der Mathematiklehrer die 80er Jahre zum Jahrzehnt des Problemlösens erhoben (NCTM 1980). Und nach Winter (1995, S. 37) gehört der Erwerb von Problemlösefähigkeiten und heuristischen Fähigkeiten zu den drei Grunderfahrungen, die den allgemeinbildenden Charakter des Mathematikunterrichts legitimieren. Diese Grunderfahrungen wurden von den KMK-Bildungsstandards (KMK 2003) und damit den Kerncurricula der Bundesländer aufgegriffen, womit das Problemlösen und damit einhergehend die Vermittlung heuristischer Techniken auch auf Basis der curricularen Vorgaben zu einem wichtigen Unterrichtsgegenstand geworden ist.

Unter den bereits erwähnten *Heurismen* bzw. *heuristischen Techniken* verstehen wir in diesem Zusammenhang „Faustregeln" („*rules of thumb*") und „allgemeine Ratschläge" („*general suggestions*") für „erfolgreiches Problemlösen" im Sinne von Schoenfeld (1985, S. 22 f.). Es handelt sich – im Gegensatz zu Algorithmen – um Tätigkeiten, die das Problemverständnis oder die Lösungssuche erleichtern können, aber keine Lösung garantieren. Beispiele für Heurismen, die insbesondere bei mathematischen Problemen hilfreich sein können, sind unter anderem das Sammeln von Daten in Tabellen, die Betrachtung von Analogien oder Systematisches Probieren.

In diesem Artikel geht es um eine Studie zur Trainierbarkeit heuristischer Techniken, genauer gesagt um die methodisch optimierte Replikation einer solchen Studie (Schoenfeld 1979; 1985, Kap. 6). Mit Schoenfeld verstehen wir den Begriff des „heuristischen Trainings" als ein Instruktionsdesign bezüglich Heurismen im Sinne eines Lernprozesses und nicht als beispielgestütztes Vor- und Nachmachen oder gar unverstandenes Auswendiglernen.[1] Der Fokus liegt dabei deutlicher auf der Art des Trainings, verglichen wird ein explizites Heurismentraining mit einem impliziten Training sowie mit einer Gruppe ohne Treatment.

2 Theoretischer Hintergrund

In der Mathematikdidaktik wird ein *Problem* – in Anlehnung an die Psychologie – verstanden als „Aufgabe, die dem Bearbeiter beim Lösen eine Barriere entgegenstellt" (Vollrath 1992, S. 127). Solche (personenspezifischen) Barrieren behindern bei der Arbeit an Problemen die Transformation eines unerwünschten Ausgangszustands in einen erwünschten Endzustand (vgl. Duncker 1935, S. 1; Dörner 1979, S. 10; Klix 1971, S. 639 f.; Zech 1992, S. 279). Probleme werden häufig explizit abgegrenzt von Routineaufgaben, für deren Bewältigung dem Bearbeiter ein Routineschema oder ein Algorithmus vorliegen, die – Rechenfehler ausgenommen – sicher zu einer Lösung führen (vgl. Pólya 1980, S. 4; Büchter & Leuders 2005, S. 28 ff.).[2]

[1] Wir verstehen das Wort „Training" im hier geschilderten Zusammenhang also in einem umfassenderen psychologischen Sinn und nicht etwa als „drill and practise" von routinemäßig erlernbaren Verfahren; vergleiche auch Hasselhorn und Gold (2009, S. 379), die „*kognitives Training*" abgrenzen von der „psychologischen *Therapie*" und insbesondere vom „pädagogischen Drill bzw. *Coaching*".

[2] Tatsächlich ist es schwer möglich, allein anhand einer Aufgabe zu entscheiden, ob ein Problem oder eine Routineaufgabe vorliegt, da dies immer vom jeweiligen Bearbeiter

Liegt kein solches Routineschema vor, sind es *Heurismen*, die wesentlich zum Erfolg der Problembearbeitung beitragen können, indem sie mögliche Lösungswege aufzeigen oder ebnen bzw. den Suchraum einschränken, in dem der Problemlöser sich bewegt (z. B. Schoenfeld 1992; Heinze 2007).

Unter Heurismen versteht man „weichere“ mathematische Tätigkeiten, die helfen können, Problemsituationen besser zu verstehen und Fortschritte auf dem Weg zur Lösung zu machen (z. B. Schoenfeld 1985, S. 22 f.; Tietze, Klika & Wolpers 2000, S. 98; Koichu, Berman & Moore 2007, S. 101). Oft werden Heurismen in Abgrenzung zu Algorithmen, die nur auf bestimmte Aufgabentypen anwendbar sind, charakterisiert (z. B. Bruder 2000, S. 72 f.). Heurismen hingegen können sehr viel breiter eingesetzt werden, ohne allerdings zwingend zum Erfolg zu führen. Ausführliche Analysen mathematischer Probleme und ihrer Bearbeitung mithilfe von Heurismen haben beispielsweise Pólya (1949) oder Schwarz (2006) durchgeführt. Beispiele für unterschiedliche Klassifikationen und Kategorisierungen von für die Mathematik wichtigen Heurismen finden sich u. a. bei Schreiber (2011) und Bruder (2000).

Abgrenzen möchten wir Heurismen an dieser Stelle von metakognitiven und selbstregulativen Tätigkeiten zur Steuerung und Regulation von (Problemlöse-) Prozessen (vgl. Schoenfeld 1985, S. 27; Konrad 2005, S. 23). Metakognitive und selbstregulative Aktivitäten sind für erfolgreiches Problemlösen zwar von großer Bedeutung (vgl. Collet 2009; Schoenfeld 1985, Kap. 4), sie liegen aber nicht im Fokus dieser Studie, da wir uns auf Bedingungen für ein erfolgreiches Heurismentraining konzentrieren möchten, ohne dies mit anderen Aspekten zu mischen.

Da Heurismen wesentlich zum Erfolg beim Problemlösen beitragen können (z. B. Rott 2013, Kap. 13), ist es für die Mathematikdidaktik von besonderem Interesse, Aussagen über die Erlern- und Trainierbarkeit von Heurismen treffen zu können.

Hinsichtlich möglicher Trainingskonzepte lassen sich grundsätzlich zwei Positionen unterscheiden. Die eine geht davon aus, dass sich heuristisches Arbeiten allein durch Üben erlernen lässt:

> „Wenn der Lehrer in seinen Schülern die Denkoperationen entwickeln will, die den Fragen und Anregungen unserer Tabelle entsprechen, so legt er diese Fragen und Anregungen den Schülern so oft vor, wie er das ungezwungen tun kann. [...]. Dank dieser Führung wird der Schüler schließlich hinter den rechten Gebrauch dieser Fragen und Anregungen kommen.“ (Pólya 1949, S. 18)

Allerdings ist fraglich, ob diese Art des Trainings und des Einsatzes der Pólya-Tabelle tatsächlich von Erfolg gekrönt ist (vgl. Schoenfeld 1985, Einleitung). Einer der Gründe gegen ein solches Training liegt für Schoenfeld darin, dass die Pólya-

abhängt und davon, ob diesem passende Verfahren zur Lösungsfindung bekannt sind (vgl. Rott 2013, Kap. 2 für eine ausführliche Diskussion zu diesem Thema).

Strategien sehr allgemein formuliert sind. Insbesondere unerfahrenen Problemlösern könne es kaum gelingen, diese in einer Problemsituation so zu konkretisieren, dass sie zur Lösungsfindung beitragen. Beispielsweise werden unter dem Heurismus „Spezialfälle betrachten“ eine Vielzahl ganz unterschiedlicher Konkretisierungen subsumiert (Schoenfeld 1992, S. 53 ff.). Allenfalls kann man sich nach dem Finden einer solchen konkreten Anwendung bewusst machen, dass diese sich als Instanz von „Spezialfälle betrachten“ auffassen lässt; und dies ist dann eine zusätzliche Anforderung, die – etwa in der Reflexionsphase – vom Lehrer zu organisieren ist. Entsprechend postuliert König, ein Heuristiker der Chemnitzer Schule, im Gegensatz zu Pólya:

> „Ausgewählte heuristische Vorgehensweisen sollten (als eine spezielle Art von Verfahrenskenntnissen) im Prozeß der Tätigkeit bewußt vermittelt werden. Das heißt, es geht um ein zielgerichtetes Aneignen und Anwenden im Unterricht und um ein explizites Abheben von methodologischen Erkenntnissen. Ein nur implizites Vermitteln etwa durch Vorbildwirkung reicht nicht aus.“ (König 1992, S. 24)

Im Anschluss an König nennen wir ein Heurismentraining *explizit*, wenn es ein solches „Abheben von methodologischen Erkenntnissen“ beinhaltet, d. h., wenn die verwendeten Heurismen während des Lernprozesses bewusst thematisiert werden. Wir sprechen von *implizitem Training*, wenn Heurismen während des Lernprozesses nicht herausgestellt werden.

Zum Einsatz und zur Erlernbarkeit von Heurismen gibt es viele Studien, die meisten von ihnen allerdings mit eher schwachen Resultaten (vgl. Schoenfeld 1992, S. 52). Lucas (1974)[3] und Kantowski (1977)[4] haben jeweils mehrwöchige Heurismentrainings durchgeführt, wobei Lucas ein – nach unserer Terminologie – explizites mit einem impliziten Training verglichen und leichte Vorteile für das explizite Training nachgewiesen hat; Kantowski hat ausschließlich explizit trainiert.

Koichu, Berman & Moore (2007) berichten über ein fünfmonatiges Feldexperiment in zwei israelischen achten Klassen, in denen das Ausmaß der Anwendungen der vermittelten heuristischen Techniken mit dem mathematischen Lernzuwachs (gemessen mithilfe einer SAT-Variante) positiv korrelierte. Es gab allerdings keine Kontrollgruppe, und die Intervention erfolgte nur indirekt über die Lehrkräfte. Die Planung der Unterrichtsstunden durch die Lehrkräfte wurde jedoch in enger Zusammenarbeit mit den Forschern durchgeführt, und es wurde dafür gesorgt, dass die SchülerInnen ein „heuristisches Vokabular“ (*heuristic vocabulary*) erwerben

---

3 Lucas hat 30 Universitätsstudierende in zwei Gruppen aufgeteilt und acht Wochen lang durch praktische Übungen im Problemlösen geschult. Eine der beiden Gruppen hat zusätzlich zu den Problemaufgaben ausführliche Informationen zu Heurismen erhalten.

4 Kantowski hat acht SchülerInnen der neunten Klasse über mehrere Monate lang geschult und getestet, wobei der Schwerpunkt auf geometrischen Beweisaufgaben lag. Problemlöseaktivitäten und Heurismen wurden dabei auch bewusst besprochen und geübt.

konnten. Es handelt sich – in unserer Terminologie – also um ein explizites Training.

Collet (2009) hat im Rahmen des Projektes PROSA ebenfalls den Erfolg von SchülerInnen (7. und 8. Klassenstufe) im Rahmen eines Heurismentrainings (sowie eines Selbstregulationstrainings, auf das hier nicht weiter eingegangen wird) untersucht, wobei dieses Training ebenfalls über zuvor geschulte Lehrkräfte erfolgte. In den Klassen mit Problemlöseintervention wurden die Heurismen explizit thematisiert. Sie konnte deutliche Zuwächse in der Anzahl der verwendeten Heurismen im Vergleich vom Vor- zum Nachtest sowie mittelstarke, signifikante Korrelationen zwischen dem Einsatz von Heurismen und dem Erfolg in Problemlöseaufgaben nachweisen.

Für einen weiteren Überblick siehe z. B. Sweller (1990); Schoenfeld (1985, S. 72 ff.); Schoenfeld (1992, S. 52 ff.); Koichu, Berman & Moore (2007, S. 103 ff.) und Rott (2013, Kap. 6).

Eine der wenigen Studien, die sich ausschließlich auf ein Heurismentraining (im Gegensatz zu zusätzlichem Training selbstregulativer Kompetenzen) konzentrieren und in denen ein relevanter Trainingserfolg berichtet wird, stammt von Schoenfeld (1979) selbst. Es handelt sich dabei um eine kleine Laborstudie zum Problemlösen lernen, in der Schoenfeld den Lernerfolg durch die Bearbeitung von Problemen vergleicht mit der Bearbeitung von Problemen und zusätzlicher, expliziter Heurismeninstruktion (s. u.). Schoenfeld weist nach, dass bestimmte Heurismen durch eine explizite Instruktion trainierbar sind und dass dieses Training (bei entsprechenden Aufgaben, die mithilfe dieser Heurismen gelöst werden können) im Vergleich zum reinen Problemlösen zu einem erhöhten Bearbeitungserfolg führt. Auf die Bedeutung dieser Erhebung und die Tatsache, dass es – zumindest bis 1990 – kaum vergleichbare Studien oder Replikationen gebe, weist Sweller (1990, S. 413) explizit hin. Die Aussagekraft von Schoenfelds Studie wird allerdings durch bestimmte Bedingungen des experimentellen Settings und eine geringe Teilnehmerzahl eingeschränkt (Näheres zu dieser Studie s. Abschnitte 2.1 und 2.2).

Bis heute gibt es zu Schoenfelds Untersuchung keine vergleichbaren quasi-experimentellen Studien, deren Aussagen nicht durch die Trainingsmethodik (zusätzlich Selbstregulation) oder die Lernumgebung (Klassenraumsituation) schwer interpretierbar werden, so dass eine Replikation angebracht erscheint. Die Studie von Schoenfeld diente uns als Vorbild für eine Studie unter leicht geänderten Ausgangsbedingungen zu Beginn des Wintersemesters 2009/10 im Rahmen einer Vorlesung zum Thema „Problemlösen“ an der Universität Hannover.

2.1 Schoenfelds „Explicit Heuristic Training“-Studie

Schoenfeld (1985, Kap. 6, eine überarbeitete Version von Schoenfeld 1979) hat sieben Studenten zufällig auf zwei Gruppen aufgeteilt – vier in die *experimental* und drei in die *control group*.[5] Alle sieben wurden einem *Vortest* unterzogen, der fünf Aufgaben enthielt, für die sie jeweils bis zu 20 Minuten Zeit hatten. In den folgenden zwei Wochen gab es für Schoenfelds Probanden fünf individuelle Trainingssitzungen mit je vier Aufgaben und zugehörigen Lösungen, die nach spätestens 15 Minuten präsentiert wurden.[6] In einer letzten Sitzung wurde schließlich ein aus fünf Aufgaben bestehender *Nachtest* bearbeitet.

Ziel der Studie war es herauszufinden, ob und unter welchen Bedingungen die Studenten in der kurzen Zeit bestimmte heuristische Strategien erlernen können. Entsprechend waren die *Test*-Aufgaben so ausgewählt, dass sie mit jeweils einer dieser Strategien gut zu lösen waren. Bei den fünf Heurismen handelte es sich um:

> „1. Draw a diagram if at all possible; 2. If there is an integer parameter, look for an inductive argument; 3. Consider arguing by contradiction; 4. Consider a similar problem with fewer variables; 5. Try to establish subgoals.“ (Schoenfeld 1985, S. 195)

Die Übungsaufgaben waren auf diese fünf Heurismen ausgelegt, die Trainingssitzungen unterschieden sich allerdings für die Gruppen: Während die Probanden der *control group* nur die Lösungen bekam (implizites Training), erhielten die Studenten der *experimental group* zusätzlich eine Liste der fünf Heurismen, und ihre Lösungen waren mit strategischen Kommentaren versehen, welche die *control group* nicht zu sehen bekam (explizites Training):

> „Although the solutions to each problem seen by both groups were identical, the students in the heuristics group saw in addition an overlay to each solution. That overlay indicated which heuristic strategy had been used to solve the problem and how it had been used.“ (Schoenfeld 1985, S. 197)

Zusätzlich wurde die Reihenfolge der 20 Trainingsaufgaben variiert: Der *control group* wurden sie durchmischt dargeboten, wohingegen die *experimental group* je Sitzung drei (von vier) Aufgaben zu nur einer Strategie bekam, um das Training zu fokussieren.

Während des *Nachtests* gab es für beide Gruppen alle fünf Minuten eine kurze Unterbrechung (*5-minute-warnings*), in denen die Probanden aufgefordert wurden, kurz Luft zu holen und ihre Vorgehensweise zu überdenken. Schoenfeld hat diese Intervention eingeführt, damit mangelnde Selbstregulation der Teilnehmer die eventuell positiven Effekte des Trainings nicht überlagert:

---

[5] Der Übersichtlichkeit halber werden in diesem Abschnitt die englischen Bezeichnungen für die Gruppen von Schoenfeld verwendet, wohingegen die deutschen Begriffe unser Forschungsprojekt beschreiben.

[6] Schoenfelds Studie ließe sich damit auch im Kontext der „worked-out examples“ (vgl. Renkl 2002; 2005; Stark et al. 2000) diskutieren. Auf diesen Aspekt seiner Studie geht er allerdings nicht ein und diskutiert auch keine alternativen Trainingsmethoden, weshalb wir diesen möglichen Forschungsrahmen hier auch ausblenden.

> „[...] the effects of bad control might well wash out any positive effects of heuristic learning. One major purpose of the experiment was to see if, and under what circumstances, students could learn to use the five heuristic strategies. It was essential to make certain that the students had the opportunity to use them." (Schoenfeld 1985, S. 208)

Zur Bewertung der *Tests* verwendete Schoenfeld zwei verschiedene Schemata[7], beide mit demselben Ergebnis: Während kein Teilnehmer im *Vortest* mehr als zwei Aufgaben lösen konnte, gab es im *Nachtest* deutliche Unterschiede zwischen der *experimental* und der *control group*. Letztere konnte sich durchschnittlich nicht verbessern, die Mitglieder der *experimental group* lösten dagegen im Schnitt zwei Aufgaben mehr als im *Vortest*.

2.2 Kritik an der Studie von Schoenfeld

Die Anzahl der Probanden in Schoenfelds Studie ist mit $n = 7$ zu gering, um den Effekt der Liste mit den Heurismen und der strategischen Kommentare bewerten, geschweige denn statistisch auswerten zu können. Auch hat Schoenfeld keine Vergleichsgruppe untersucht, die nur die beiden Tests absolviert hat, ohne an einem – irgendwie gearteten – Training teilgenommen zu haben. Hierdurch kann er nichts über die Effektivität seines Trainings aussagen und hat auch keinen Anhaltspunkt, ob die Aufgabenschwierigkeit in Vor- und Nachtest vergleichbar ist.

Hinzu kommt, dass die fünf von Schoenfeld ausgewählten heuristischen Strategien sehr unterschiedlich sind, was zum Beispiel ihr mögliches Anwendungsfeld oder ihren Abstraktionsgrad anbelangt. Auch kann nicht geklärt werden, welchen Einfluss die unterschiedliche Aufgabenreihenfolge während der Trainingssitzungen und die *5-minute-warnings* während des *Nachtests* auf die Ergebnisse hatten.[8] Schließlich hat Schoenfeld den Einsatz bzw. das Auftreten von Heurismen nicht direkt, sondern nur indirekt über das (vollständige) Lösen von Aufgaben gemessen. Es ist allerdings denkbar, dass das Training Erfolg zeigt und die Probanden im Nachtest mehr Heurismen einsetzen, ohne die Aufgaben lösen zu können.

Diese Bedingungen wurden deshalb in unserer Untersuchung verändert.

---

[7] Eine strikte Variante der Bewertung mit einem Punkt für die richtige Lösung und null Punkten sonst; sowie eine moderate Variante, bei der es auch für fast korrekte Bearbeitungen schon den einen Punkt gab.

[8] Hinweise für Promptingeffekte bei allerdings deutlich anderem Design ergeben sich aus Stark, Gruber, Renkl und Mandl (2008).

3 Unsere Studie

3.1 Fragestellungen der Studie

Die Trainierbarkeit von Heurismen besagt konkret, dass Studierende die Heurismen zur Lösung von Testaufgaben, die den zugehörigen Trainingsaufgaben ähneln, anschließend (öfter) einsetzen und dies evtl. zu einer Steigerung der Problemlöseperformanz führt. Hierzu stellen sich im Einzelnen folgende Forschungsfragen (vgl. Schoenfeld 1985, S. 189 f.):

1. Wie unterscheiden sich Studierende mit explizitem Training, mit implizitem Training und ohne Training in ihrer Problemlöseperformanz?
2. Reicht zum Erlernen und Anwenden von Heurismen das bloße Üben passender Aufgaben (implizites Training) oder bedarf es zusätzlicher Instruktionen (explizites Training)?

3.2 Das Problemlösetraining im Wintersemester 2009/10

Unsere Studie fand von Oktober bis November 2009 in den ersten fünf Übungsterminen einer Bachelor-Veranstaltung für das Lehramt Mathematik an Gymnasien und an Berufsbildenden Schulen zum Thema „Problemlösen“ statt. In der ersten und der letzten dieser Sitzungen wurde ein *Vor*- bzw. ein *Nachtest* geschrieben, die drei Termine dazwischen dienten dem Training. Beide *Tests* bestanden aus je fünf Aufgaben aus verschiedenen mathematischen Stoffgebieten (u. a. Geometrie, Algebra und Zahlentheorie), für die die Studierenden jeweils 90 Minuten Zeit hatten, die sie sich selbst einteilen konnten.

In jeder der drei *Trainingssitzungen* bekamen die Studierenden fünf Aufgaben und zugehörige Musterlösungen. Die Probanden wurden aufgefordert, alleine zu arbeiten, die Aufgaben selbstständig zu lösen und im Anschluss die Musterlösungen anzuschauen; Letzteres wurde aber nicht kontrolliert. Jeder konnte über seine Zeit und die Reihenfolge, in der er die Aufgaben und die Lösungen der jeweiligen Sitzung bearbeitete, selbst bestimmen. Nach jedem Training wurden die Unterlagen eingesammelt, um sicherzustellen, dass sich kein Proband zu Hause mit den Materialien beschäftigt und so die Trainingsergebnisse verfälscht, und um einen Austausch unter den Teilnehmern der verschiedenen Gruppen zu erschweren.

Wie in der Studie von Schoenfeld sollten fünf heuristische Strategien gezielt gelehrt und dieses Training mithilfe eines Kontroll-Experimentalgruppen-Designs überprüft werden. Die Aufgaben waren dabei so verteilt, dass es in jedem *Test* und in jeder *Trainingssitzung* je eine Aufgabe gab, die sich mit einer der Strategien besonders gut lösen lässt[9] – ohne dass den Probanden dies mitgeteilt wurde. Die fünf *Vortestaufgaben* wurden in die Trainingssitzungen eingestreut.

Die Teilnehmer wurden anhand der Ergebnisse des *Vortests* auf zwei gleich starke Gruppen aufgeteilt, die *Experimentalgruppe* (EXP) und die *Kontrollgruppe* (KON), die sich durch leicht unterschiedliche Instruktionen auszeichneten: Im Unterschied zu den Teilnehmern der Kontrollgruppe erhielten die Studenten der Experimentalgruppe eine Liste mit den fünf heuristischen Strategien und kurzen Beschreibungen dieser Strategien (insgesamt etwa eine dreiviertel DIN A4-Seite). Zusätzlich waren die Musterlösungen dieser Gruppe am Seitenrand mit Kommentaren in sogenannten „Strategiekästen" versehen, in denen auf den Einsatz der Heurismen hingewiesen wurde (s. Abb. 4 im Anhang für ein Beispiel). In diesen zwei Punkten erschöpften sich die Unterschiede zwischen den beiden Gruppen. Im Gegensatz zur Studie von Schoenfeld gab es also keine unterschiedliche Aufgabenreihenfolge für die beiden Gruppen oder *5-minute-warnings* während des *Nachtests*, um den Effekt des Treatments nicht zu konfundieren.

Die fünf heuristischen Strategien der vorliegenden Studie waren folgende (je eine Beispielaufgabe pro Heurismus findet sich im Anhang):

(Z) Zeichnung anfertigen

(S) Spezialfälle betrachten

(R) Rückwärtsarbeiten

(B) Weglassen einer Bedingung

(V) Ähnliche Aufgaben mit weniger Variablen betrachten

Man sieht, dass es sich nicht um dieselben fünf Heurismen handelt, die in Schoenfelds Studie enthalten waren, da zwei Strategien (2. und 3. aus Schoenfelds Liste: *inductive argument* und *contradiction*) als zu bereichsspezifisch und eine (5., *subgoals*) als zu allgemein eingeschätzt wurden. Die im Gegensatz zu 2. und 5. – laut Schoenfeld – gut trainierbaren Heurismen, (Z) und (V), wurden einschließlich der publizierten Aufgaben übernommen (vgl. Schoenfeld 1985, S. 210). Die übrigen Aufgaben stammen aus verschiedenen Aufgabensammlungen und mathematikdidaktischen Veröffentlichungen (u. a. Pólya 1949 und Schwarz 2006).[10]

Zusätzlich zu den zwei Gruppen, EXP und KON, haben wir die beiden *Tests* mit drei Wochen Pause dazwischen (derselbe zeitliche Abstand, der beim Training

---

[9] Natürlich kann nicht ausgeschlossen werden, dass sich die Aufgaben auch mit gänzlich anderen oder mit Kombinationen von Heurismen lösen lassen.

[10] Auch in der vorliegenden Studie streut die Reichweite der trainierten Heurismen noch stark, die von Schoenfeld übernommene Strategie (V) ist beispielsweise deutlich eingeschränkter einsetzbar als (Z) oder (R).

zwischen den Tests lag) in einer weiteren Veranstaltung erhoben. Es handelte sich dabei um eine Übung zur „Didaktik der Analysis“ mit Studierenden, die vom Ausbildungsstand in ihrem Studium mit den „Problemlöse“-Probanden vergleichbar waren; dies ist unsere Vergleichsgruppe (VGL).[11]

3.3 Auswertung

Der ursprüngliche Plan sah vor, die „Problemlöse“-Studierenden (EXP und KON) anhand des *Vortests* zu parallelisieren, um auf diese Weise zwei möglichst gleich starke Gruppen zu erhalten (vgl. Bortz 2005, S. 143 ff. bzw. Bortz & Döring 2006, S. 563). Während des Trainings schieden jedoch mehrere Teilnehmer aus, so dass (nach Entfernung der entsprechenden „Zwillinge“) nur fünf Paare für eine Auswertung übrig geblieben sind – eine schwache Datenbasis.

Aus diesem Grund wurde die Parallelisierung fallen gelassen. Stattdessen wird die Gruppe aller 19 Teilnehmer (neun Experimental- und zehn Kontrollgruppenteilnehmer) betrachtet. Die Vergleichsgruppe aus der Parallelveranstaltung bestand aus insgesamt 17 Personen, von denen 15 am *Vor*- und elf am *Nachtest* teilgenommen haben. Neun dieser Studenten haben beide Tests mitgeschrieben, und nur diese Teilgruppe wird hier ausgewertet. Die Gruppenstärken betragen also EXP = 9, KON = 10 und VGL = 9 Teilnehmer.

Ausgewertet wurden nur die Ergebnisse des *Vor*- und *Nachtests*. Die Mitschriften aus den Trainingssitzungen liegen zwar vor, wurden aber nicht analysiert (da nicht kontrolliert wurde, wann die Probanden zur Musterlösung weiter geblättert haben).

Im Gegensatz zu Schoenfeld haben wir uns gegen eine dichotome Bewertung (1/0 für richtig/falsch) der *Test*-Ergebnisse entschieden, da die unter diesen Bedingungen zu erwartenden Häufigkeitsverteilungen der Ergebnisse nur dann sinnvoll ausgewertet werden können, wenn „die Zahl der Aufgaben nicht zu klein ist“ (Lienert & Raatz 1998, S. 48) – unsere Tests bestanden aber jeweils nur aus fünf Aufgaben. Insbesondere für komplexe Aufgaben, die „eine deutlich längere Bearbeitungszeit erfordern [als einfache Testaufgaben] oder schwieriger erscheinen“ (ebd.) empfehlen Lienert und Raatz (1998), höhere Punktzahlen zu verwenden bzw. Aufgaben für ihre Bewertung in Teilaufgaben zu zerlegen. Des Weiteren gewährleistet eine dichotome Bewertung keine valide Auswertung der Tests im Sinne der Fragestellungen unserer Studie: Die Probanden könnten im Nachtest mehr Heurismen anwenden, ohne dass sich dies in einer größeren Anzahl korrekt gelöster Aufgaben widerspiegelt.

---

[11] Unter „Kontrollgruppe“ wird hier – in Anlehnung an die Terminologie Schoenfelds und abweichend zum sonstigen deutschen Sprachgebrauch – die Gruppe mit Treatment-Variante (ohne explizite, d. h. mit impliziter Heurismeninstruktion) bezeichnet, während die Kontrollgruppe i. e. S. als „Vergleichsgruppe“ fungiert.

Aus diesen Gründen haben wir für jede Aufgabe in den *Tests* eine bestimmte Anzahl an *Rohpunkten* vergeben. Ein Rohpunkt stellt – möglichst unabhängig von der jeweiligen Aufgabe – einen gewissen mathematischen Bearbeitungsschritt dar, etwa eine korrekte Umformung, einen zielführenden Lösungsansatz oder eine passende Skizze. Erfasst werden insbesondere also auch „heuristische Akte“ wie die Betrachtung sinnvoller Spezialfälle oder das Einführen passender Bezeichnungen.[12] Ziel dieser Vorgehensweise ist, die Aufgaben aufgrund ihrer tatsächlichen Schwierigkeit zu gewichten. Die Vergabe der Rohpunkte führte dazu, dass im *Vor-* und im *Nachtest* leicht unterschiedlich viele Punkte erreichbar waren (43/40) – die Tests waren allerdings so gestaltet, dass kein Proband die volle Punktzahl erreicht hat, man also von „open-ended“-Tests sprechen kann. Trotz der unterschiedlichen Maximalpunktzahl werden im Folgenden die Rohpunktzahlen und nicht etwa Prozentwerte betrachtet, denn Differenzen von Prozentzahlen sind in diesem Zusammenhang nicht sinnvoll als Leistungsveränderung interpretierbar.

Inhaltlich lassen sich die Rohpunkte – wie oben angesprochen – als „Bearbeitungsschritte“ deuten, die durch „heuristische Akte“ ausgelöst wurden. Die Punktzahl ist damit als ein Maß für den Heurismeneinsatz interpretierbar. Im Gegensatz zu Schoenfeld, der den Erfolg seines Trainings nur indirekt über die Anzahl der (vollständig) gelösten Aufgaben messen konnte, können wir einen Anstieg der Rohpunktzahl als vermehrtes heuristisches Vorgehen und damit als Zuwachs an Problemlösekompetenz interpretieren.

Für das von beiden Autoren ausgearbeitete Bewertungsschema wurden ausführliche Musterlösungen der Testaufgaben angefertigt, wobei auch unterschiedliche Herangehensweisen und Lösungswege Berücksichtigung fanden. Die tatsächliche Bewertung hat der Erstautor gemeinsam mit einer Mitarbeiterin der Arbeitsgruppe durchgeführt, wobei strittige Fälle im Sinne einer *konsensuellen Validierung* diskutiert wurden.

Aufgrund der geringen Stichprobengröße, insbesondere aber da es zur Messung von Leistungszuwächsen keinen „Goldstandard“ gibt (vgl. z. B. die Diskussion zur Verwendung von Differenzen in Bortz & Döring 2006, S. 552 ff.), präsentieren wir im Folgenden mehrere methodische Ansätze zur Auswertung und verwenden dabei insbesondere auch deskriptive Statistik.

Die Rohpunkt-Daten wurden mit dem Test von Shapiro-Wilk für kleine Stichproben auf Normalverteilung getestet; der Test liefert stets Werte größer 0,05, also

---

[12] Bei der Aufgabe B1 (siehe Anhang) gab es beispielsweise (von insgesamt 9 erreichbaren Punkten) je einen Punkt für das Anfertigen einer *Skizze der Zielkonfiguration* (Quadrat im Dreieck), den Plan zur Konstruktion eines Quadrats mit Eckpunkten auf zwei der Dreiecksseiten (*Weglassen einer Bedingung*) und das Zeichnen einer *Hilfslinie* durch die Quadrat-Eckpunkte, die noch nicht auf einer der Dreiecksseiten liegen.

keine signifikanten Abweichungen von der Normalverteilung. Dies rechtfertigt eine Auswertung mithilfe parametrischer Verfahren. Eine nicht-parametrische Auswertung der Daten[13] dieser Stichprobe ergab in Bezug auf statistische Signifikanz und Effektstärken dieselben Ergebnisse wie die Auswertung der Gesamtstichprobe (s. u.), weshalb hier nicht weiter darauf eingegangen wird.

Zur Auswertung herangezogen wurden Kovarianz- und Varianzanalysen (vgl. Rasch et al. 2006b, S. 1 ff.) sowie *t*-Tests (vgl. Rasch et al. 2006a, S. 43 ff.). Eine Übersicht über das Treatment und die Auswertung gibt Abbildung 1.

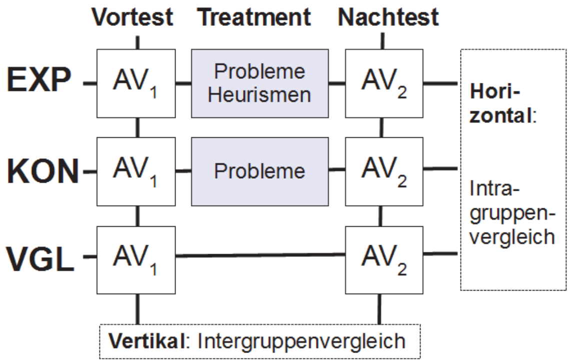


Abbildung 1: Übersicht über die Versuchsbedingungen und die Auswertungsmethoden (Kovarianzanalyse und Post-Hoc-Vergleiche); AV steht für „abhängige Variable“

4 Ergebnisse

Die *parallelisierte Stichprobe* (mit EXP-KON-„Zwillingen“) ist, wie oben erwähnt, aufgrund experimenteller Mortalität sehr klein geworden. Aus diesem Grund stellen wir im Folgenden ausschließlich die Auswertung mit Blick auf die Gesamtgruppen vor.

4.1 Auswertung der gesamten Stichprobe

Die arithmetischen Mittel und Standardabweichungen der drei Gruppen sind in Tabelle 1 dargestellt. Abbildung 2 zeigt die zugehörigen Boxplots. Auf den ersten Blick fällt der starke Anstieg bei der Experimentalgruppe auf.

[13] Mann-Whitney-U-Tests anstelle von *t*-Tests für unabhängige Stichproben, Wilcoxon-Tests anstelle von *t*-Tests für abhängige Stichproben sowie Brown-Forsythe-Tests anstelle von einfaktoriellen Varianzanalysen

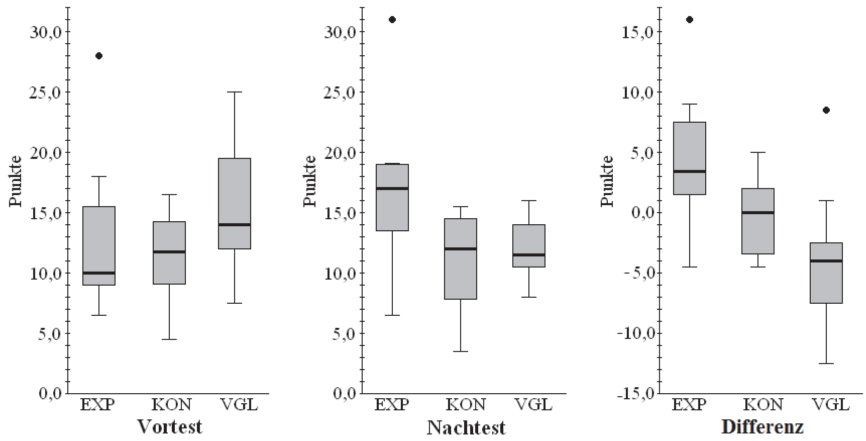


Abbildung 2: Boxplots von Vor- und Nachtest sowie für die Differenz

Gruppe	N	Vortest M (S)	Nachtest M (S)	Differenz M (S)
EXP	9	12,83 (6,84)	16,72 (6,59)	3,89 (4,44)
KON	10	11,45 (3,78)	11,05 (4,19)	–0,40 (3,24)
VGL	9	16,00 (5,88)	12,06 (2,38)	–3,94 (6,09)

Anmerkung: In Vor- und Nachtest gab es mit 43 bzw. 40 Punkten einen kleinen Unterschied in Bezug auf die maximale Punktzahl; diese Punktzahl wurde aber von keinem Teilnehmer erreicht, es handelt sich um „open-ended“-Tests.

Tabelle 1: Arithmetische Mittelwerte M und Standardabweichungen (S) der drei Gruppen

4.2 Deskriptiver Vergleich der Leistungsänderungen in den drei Gruppen

Die Zusammenschau der Vor- und Nachtestwerte (siehe die Abb. 3, nächste Seite) zeigt:

- Während sich die Experimentalgruppe im Schnitt um fast 4 Rohpunkte verbessert hat (was etwa einer zur Hälfte gelösten Aufgabe entspricht), gab es in der Kontrollgruppe insgesamt sogar eine (schwache) Abnahme der Gesamtpunktzahl und in der Vergleichsgruppe eine deutliche Abnahme von fast 4 Punkten.
- In der Experimentalgruppe gibt es sieben Probanden mit positiver und nur einen mit negativer Punktdifferenz; in der Kontrollgruppe sind es jeweils fünf mit positiver und negativer Differenz; in der Vergleichsgruppe zeigen sogar sieben Probanden negative Tendenzen und nur zwei eine positive.

- Die wirklich großen Punktzunahmen sind (bis auf je eine Ausnahme in der Kontroll- und in der Vergleichsgruppe) alle in der Experimentalgruppe zu verzeichnen.[14]
- Die Experimentalgruppe enthält einen Ausreißer; seine Entfernung aus den Daten hat keinen Einfluss auf die berichteten Signifikanzen.

4.3 Statistischer Vergleich der Leistungsänderungen in den drei Gruppen

Die „Augenschein-Auswertung“ aus den Abschnitten 4.1 und 4.2 lässt sich mithilfe traditioneller statistischer Methoden bestätigen: Eine Kovarianzanalyse mit der Punktzahl des Vortests als Kovariate ergibt einen signifikanten Effekt für die Zugehörigkeit zu einer Gruppe (EXP, KON, VGL) ($F = 6{,}593$; $p = 0{,}005$).[15]

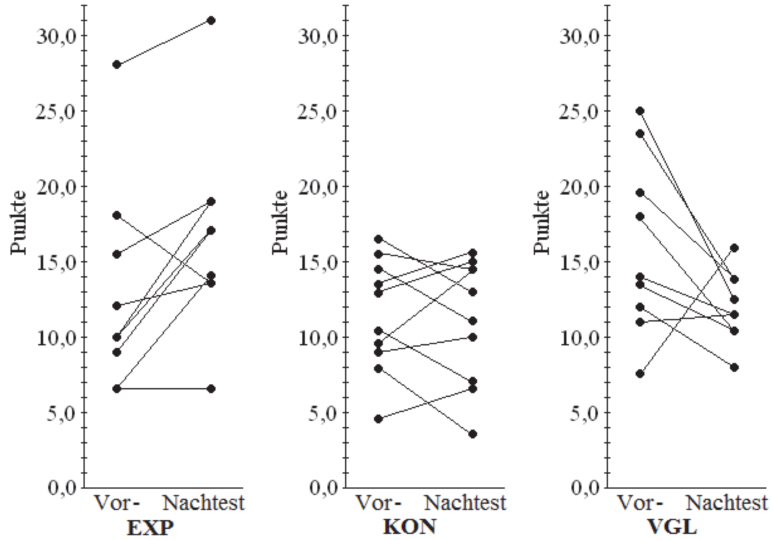


Abbildung 3: Veränderungsmuster der drei Gruppen zu den beiden Messzeitpunkten

[14] Die Unterschiede der Differenzen (Nachtest – Vortest) der Gruppen EXP und KON sind signifikant (*t*-Test für unabhängige Stichproben; 1-seitig, EXP > KON, $t = 2{,}423$, $df = 17$, $p < 0{,}05$).

[15] Eine Varianzanalyse mit Messwiederholung bestätigt einen signifikanten Interaktionseffekt von Testleistung und Gruppenzugehörigkeit ($F = 6{,}307$; $p = 0{,}006$); aufgrund des Testdesigns und der unterschiedlichen Testschwierigkeit (s. u.) hielten wir die Kovarianzanalyse an dieser Stelle für angebrachter.

Post-Hoc-Analysen in Form von einfaktoriellen Varianzanalysen und *t*-Tests decken die folgenden Zusammenhänge auf:

Intergruppenvergleich: Während es im Vortest, vor dem Training, keinen signifikanten Unterschied zwischen den drei Gruppen gab, traten dagegen im Nachtest signifikante Abweichungen auf;[16] ein Post-Hoc-Vergleich (Bonferroni-Test) ergab hier signifikante Unterschiede zwischen EXP und KON ($p = 0{,}043$).

Die korrigierte Effektstärke des Trainings[17] beträgt jeweils: $d_{\text{korr\_(EXP,KON)}} = 0{,}786$; $d_{\text{korr\_(EXP,VGL)}} = 1{,}438$; $d_{\text{korr\_(KON,VGL)}} = 0{,}640$; es handelt sich nach der Konvention von Cohen (vgl. Bortz & Döring 2006) also um mittlere bis große Effekte.

Die Betrachtung der Differenz-Werte ergab hochsignifikante Unterschiede zwischen den drei Gruppen ($F = 6{,}307$; $p = 0{,}006 < 0{,}01$); ein Post-Hoc-Vergleich zeigte hochsignifikante Unterschiede zwischen EXP und VGL auf ($p = 0{,}005$).

Intragruppenvergleich: Der Vergleich der Punktzahlen zu den zwei Messzeitpunkten (Vor- und Nachtest) innerhalb der drei Gruppen ergibt Folgendes (*t*-Tests für abhängige Stichproben, 1-seitig für EXP und KON mit der Erwartung, dass der Nachtest besser ausfällt als der Vortest, und 2-seitig für VGL, da hier kein Effekt zu erwarten ist): Einzig die Experimentalgruppe kann einen signifikanten Zuwachs verzeichnen ($t = -2{,}626$; $p = 0{,}015$);[18] die Effektgröße für abhängige Stichproben, $\delta' = \mu_D/\sigma_D$ (Bortz & Döring 2006, S. 608 f.), ihres Trainings beträgt $\delta_{EXP}' = 0{,}88$. Bei der Kontrollgruppe ($t = 0{,}391$) und bei der Vergleichsgruppe ($t = 1{,}953$) gibt es keine signifikanten Veränderungen.

4.4 Vergleich von Vor- und Nachtest-Schwierigkeit

Die trotz Trainings nur geringen Punktdifferenzen legen die Vermutung nahe, dass der *Nachtest* etwas schwieriger war als der *Vortest* – obwohl aufgrund der Wahl derselben „günstigen heuristischen Strategien", auf Basis von Musterlösungen und mithilfe vergleichbarer Anzahlen an *Rohpunkten* natürlich versucht worden war, Testaufgaben gleicher Schwierigkeit auszuwählen. Zur Überprüfung dieser Vermutung können wir die Vergleichsgruppe heranziehen. Das Ergebnis von 16 Punkten im *Vor*- und 12 Punkten im *Nachtest* (s. Tab. 1) bestätigt die Vermutung, dass die beiden Tests von unterschiedlichem Schwierigkeitsgrad waren.

---

16 Einfaktorielle ANOVA, Vortest: $F = 1{,}627$; $p = 0{,}217$. Nachtest: $F = 3{,}853$; $p = 0{,}035$

17 Berechnet mithilfe des Tools auf: http://www.psychometrica.de/effektstaerke.html [02.02.2015]

18 Da in diesen Gruppenvergleichen mehrere *t*-Tests an derselben Stichprobe durchgeführt wurden, besteht hier die Gefahr der α-Fehler-Kumulation (vgl. Bortz 2005, S. 129). Der Zuwachs ist nach der (konservativen) Bonferroni-Korrektur auch auf dem korrigierten Niveau $\alpha' = \alpha/3 = 0{,}017$ signifikant.

5 Diskussion

In der vorliegenden Studie wurden – als verbesserte Replikation einer Studie von Schoenfeld – verschiedene Arten eines Problemlöse- und Heurismentrainings verglichen: Im Rahmen einer fünf Sitzungen umfassenden, wöchentlich stattfindenden Übung haben Studierende eigenständig mathematische Probleme bearbeitet und ihre Ergebnisse anschließend mit bereitgestellten Musterlösungen verglichen. Während sich die Instruktion bei der Kontrollgruppe auf diese Musterlösungen beschränkte (implizites Training), hat die Experimentalgruppe zusätzlich Hinweise und Informationen zu jeweils passenden Heurismen erhalten (explizites Training). Eine dritte Gruppe, die Vergleichsgruppe, hat ohne Training nur den Pre- und den Posttest mitgeschrieben, den auch die anderen beiden Gruppen durchlaufen haben. Die Ergebnisse zeigen statistisch signifikant bessere Ergebnisse der Experimentalgruppe im Vergleich zu den anderen beiden Gruppen sowie der Kontrollgruppe im Vergleich zur Vergleichsgruppe.

5.1 Interpretation der Ergebnisse

Die erste Forschungsfrage zu Trainingsunterschieden in den drei Gruppen (explizites, implizites Training und Vergleichsgruppe) lässt sich mit einem Blick auf die statistische Auswertung eindeutig beantworten: Die Experimentalgruppe mit Problemlöse- und explizitem Heurismentraining hat im Vergleich von Vor- und Nachtest erwartungsgemäß besser abgeschnitten als die Kontrollgruppe, die nur anhand der(selben) Problemaufgaben und Musterlösungen geübt und die Heurismen daher nur implizit trainiert hat. Die Kontrollgruppe wiederum weist ein besseres Ergebnis im Vergleich der beiden Tests auf als die Vergleichsgruppe, die gar kein entsprechendes Training durchlaufen hat; diese Ergebnisse sind trotz der kleinen Gruppengröße alle statistisch signifikant. Die Teilnehmer der Experimentalgruppe haben im Nachtest durchschnittlich vier „heuristische Akte“ mehr gezeigt als im Vortest.[19] Aufgrund der Kürze des Trainings – lediglich drei Übungssitzungen – hätte man auch keinen größeren Effekt erwarten dürfen; andere Studien berichten von wesentlich längeren Trainingsphasen (vgl. Schoenfeld (1985, Kap. 4), der seine Probanden ein ganzes Semester lang trainiert hat; Koichu, Berman & Moore (2007) haben – genauso wie Kantowski (1977) – Schüler über fünf Monate unterwiesen). Die geringen Unterschiede zwischen der Experimental- und der Kontrollgruppe lassen sich auch damit erklären, dass in der letztgenannten Instruktionsversion nicht verhindert werden kann, dass die Lernenden selbst Verbindungen herstellen und eigenständig Abstraktionen der Vorgehensweisen schaffen.

[19] Im Vergleich zu Schoenfelds Originalstudie, in der die Probanden nach dem Training im Schnitt zwei Aufgaben mehr lösen konnten, fällt unser Ergebnis etwas schwächer aus, vier Rohpunkte entsprechen einer halben bis einer Aufgabe.

In der Fragestellung (2) ging es darum, ob es zum Erlernen bestimmter Heurismen expliziter Instruktionen bedarf oder ob bloßes Üben (implizites Training) ausreichend ist für die Aneignung der Heurismen. Die Daten legen nahe, dass zusätzliche Instruktionen und Hinweise wesentlich dazu beitragen, erlernte Heurismen erfolgreich auf andere Aufgaben zu übertragen. Die Probanden der Kontroll- und der Experimentalgruppe haben vor dem Nachtest an jeweils 15 verschiedenen Problemaufgaben, je drei pro Strategie, trainiert. Wären die expliziten Strategiehinweise überflüssig und implizites Training wäre alles, was man benötigt, hätte es keine signifikanten Unterschiede nach dem Treatment geben dürfen. Tatsächlich erhielten wir aber einen Effekt von mittlerer Stärke. Dies bestärkt Königs These, dass ein bloßes Vormachen der Heurismen weniger effektiv ist als ein Training, bei dem Heurismen in der Rückschau ausdrücklich hervorgehoben und aufgabenübergreifend thematisiert werden.

5.2 Aussagen zur Generalisierbarkeit der Ergebnisse

Auch wenn im Vergleich zur Originalstudie von Schoenfeld einige Aspekte unserer Studie optimiert werden konnten, bleibt die Aussagekraft dennoch eingeschränkt. Die Anzahl der Teilnehmer ist mit etwa 10 Studierenden pro Gruppe sehr gering und sollte in Folgestudien stark erhöht werden. Ebenso handelt es sich bei den Probanden ausschließlich um Studierende des Lehramts Mathematik. Ob sich die Ergebnisse auch auf weniger mathematik-affine Lerngruppen übertragen lassen, sollte weitergehend untersucht werden (siehe Abschnitt 5.3).

Die leicht unterschiedlichen Schwierigkeiten von *Vor*- und *Nachtest* lassen keinen Einfluss auf die Generalisierbarkeit der Ergebnisse vermuten, da diese Bedingungen für alle Teilnehmer identisch waren; sie könnten für zukünftige Trainingsmaßnahmen aber weitere Auswertungsschritte initiieren: Beispielsweise könnte man eine Schwierigkeitsanalyse der einzelnen Aufgaben in den beiden Tests durchführen und Items mit Bodeneffekten aus den Analysen herausnehmen (Aufgabe 1 aus dem Vortest wäre mit einer sehr linkssteilen Punkteverteilung ein Kandidat hierfür) um möglichst gleich schwierige Tests zu erstellen. Andererseits könnte man die Schwierigkeitsgrade dieser Erhebungen systematisch variieren, um den Transfer des Heurismentrainings näher untersuchen zu können.

Trotz der hier diskutierten Einschränkungen konnten die Ergebnisse von Schoenfeld, dass nämlich Heurismen durch Spezialisierung trainierbar sind und die Art des Trainings entscheidend ist, unter Laborbedingungen – ohne störende konfundierende Variablen wie zusätzliche Selbstregulationstrainings oder Lehrkräfte als zwischengeschaltete Trainingsinstanz – repliziert werden.

5.3 Ausblick

Bei einer weiteren Studie dieser Art erscheint es sinnvoll, neben einer Erhöhung der Probandenzahl und einer Verlängerung des Trainings, einen zusätzlichen Messzeitpunkt einzuführen, da hierdurch die Reliabilität der Ergebnisse auf ein Vielfaches erhöht werden kann (vgl. Bortz & Döring 2006, S. 554). Auch bedarf die Trainierbarkeit der in dieser Studie verwendeten Heurismen und anderer heuristischer Techniken weiterer Untersuchungen, die Liste der (schul-)relevanten Heurismen ist lang (vgl. Bruder 2000; Collet 2009, S. 125).

Zusätzlich ließe sich die Aussagekraft der Studie durch eine Erfassung der verwendeten Heurismen noch deutlich erhöhen. Durch die Kodierung der Problembearbeitungen unter diesem Aspekt ließen sich in Nachfolgestudien weitere Informationen gewinnen; der zusätzliche Einsatz von Interviews und/oder gefilmten Problemlöseprozessen könnte zu weiteren Erkenntnissen führen.[20] Zudem könnte mit umfangreicheren Vor- und Nachtests, die mehr als eine Aufgabe je trainierter heuristischer Strategie enthalten, ein differenzierteres Bild über die Aneignung von Heurismen gewonnen werden.

Wichtiger noch als zusätzliche Laborstudien erscheint uns die Übertragung ins Klassenzimmer. Eine Feldstudie zum Problemlösen, in deren Rahmen explizites und implizites Heurismentraining sowie eine Vergleichsgruppe unterschieden werden, wurde von Brockmann-Behnsen (2012, 2014) in vier 8. Klassen eines Gymnasiums durchgeführt. Erste Ergebnisse dieser Studie deuten auch hier auf einen Vorteil der expliziten Instruktion gegenüber eines impliziten Heurismentrainings sowie auf signifikante Unterschiede zwischen den Trainings- und den Vergleichsklassen hin.

Literatur

Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler*. Heidelberg: Springer, 6. Auflage.

Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. Heidelberg: Springer.

Brockmann-Behnsen, D. (2012). HeuRekAP – Erste Ergebnisse der Langzeitstudie zum Problemlösen und Beweisen am Gymnasium. In M. Ludwig & M. Kleine (Hrsg.), *Beiträge zum Mathematikunterricht 2012* (S. 149–152). Münster: WTM.

Brockmann-Behnsen, D. (2014). Explizites und implizites Heurismentraining im Unterricht. *Der Mathematikunterricht* 60 (5), S. 19–23.

Bruder, R. (2000). Akzentuierte Aufgaben und heuristische Erfahrungen – Wege zu einem anspruchsvollen Mathematikunterricht für alle. In L. Flade & W. Herget (Hrsg.),

[20] In Rott (2013, Kap. 10 und 13) wird ein Kodiermanual zur Erfassung von Heurismen in (videographierten) Prozessen sowie in (fotokopierten) Produkten elaboriert; dies könnte ohne großen Aufwand auf die in dieser Studie verwendeten Testaufgaben angepasst werden.

Mathematik lehren und lernen nach TIMSS – Anregungen für die Sekundarstufen (S. 69–78). Berlin: Volk und Wissen.

Büchter, A. & Leuders, T. (2005). *Mathematikaufgaben selbst entwickeln*. Berlin: Cornelsen Scriptor.

Collet, C. (2009). *Förderung von Problemlösekompetenzen in Verbindung mit Selbstregulation – Wirkungsanalysen von Lehrerfortbildungen*. Münster: Waxmann.

Dörner, D. (1979). *Problemlösen als Informationsverarbeitung*. Stuttgart: Kohlhammer.

Duncker, K. (1935). *Zur Psychologie des produktiven Denkens*. Berlin: Springer. Neudruck: 1974.

Hasselhorn, M. & Gold, A. (2009). *Pädagogische Psychologie – Erfolgreiches Lernen und Lehren*. Stuttgart: Kohlhammer.

Heinze, A. (2007). Problemlösen im mathematischen und außermathematischen Kontext. *Journal für Mathematik-Didaktik*, 28 (1), S. 3–30.

Kantowski, M. G. (1977). Processes involved in mathematical problem solving. *Journal for Research in Mathematics Education*, 8, S. 163–180.

Klix, F. (1971). *Information und Verhalten. Kybernetische Aspekte der organismischen Informationsverarbeitung*. Bern: Huber.

KMK (2003). *Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss*.

Koichu, B.; Berman, A. & Moore, M. (2007). Heuristic literacy development and its relation to mathematical achievements of middle school students. *Instructional Science, 35*, S. 99–139.

König, H. (1992). Einige für den Mathematikunterricht bedeutsame heuristische Vorgehensweisen. *Der Mathematikunterricht*, 38 (3), S. 24–38.

Konrad, K. (2005). *Förderung und Analyse von selbstgesteuertem Lernen in kooperativen Lernumgebungen: Bedingungen, Prozesse und Bedeutung kognitiver sowie metakognitiver Strategien für den Erwerb und Transfer konzeptuellen Wissens*. Habilitationsschrift, Universität Weingarten.

Lienert, G. A. & Raatz, U. (1998). *Testaufbau und Testanalyse*. Weinheim: Beltz.

Lucas, J. F. (1974). An exploratory study on the diagnostic teaching of heuristic problem-solving strategies in calculus. In J. G. Hervey & T. A. Romberg (Hrsg.), *Problem-Solving Studies in Mathematics.* Madison, WI: Wisconsin R&D Center Monograph Series.

NCTM – National Council of Teachers of Mathematics (1980). *An agenda for action*. Reston, VA: NCTM.

OECD (2014). *PISA 2012 Results: Creative Problem Solving: Students' Skills in Tackling Real-Life Problems (Volume V)*. PISA, OECD Publishing.

Pólya, G. (1949). *Schule des Denkens*. Tübingen: Francke.

Pólya, G. (1980). Wie lehren wir Problemlösen? *Mathematiklehrer,* 1, 3–5.

Rasch, B.; Friese, M.; Hofmann, W. & Naumann, E. (2006a). *Quantitative Methoden. Einführung in die Statistik – Band 1*. Heidelberg: Springer.

Rasch, B.; Friese, M.; Hofmann, W. & Naumann, E. (2006b). *Quantitative Methoden. Einführung in die Statistik – Band 2*. Heidelberg: Springer.

Renkl, A. (2002). Worked-out examples: instructional explanations support learning by self-explanations. *Learning and Instruction*, 12, S. 529–556.

Renkl, A. (2005). The Worked-Out Examples Principle in Multimedia Learning. In R. E. Mayer (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (S. 229–245). Cambridge: University Press.

Rott, B. (2013). *Mathematisches Problemlösen – Ergebnisse einer empirischen Studie*. Münster: WTM Verlag.

Schoenfeld, A. H. (1979). Explicit Heuristic Training as a Variable in Problem-Solving. *Journal for Research in Mathematics Education*, 10 (3), S. 173–187.

Schoenfeld, A. H. (1985). *Mathematical Problem Solving*. Orlando: Academic Press.

Schoenfeld, A. H. (1992). Learning to think mathematically: problem solving, metacognition, and sensemaking in mathematics. In D. Grouws (Hrsg.), *Handbook for Research on Mathematics Teaching and Learning* (S. 334–370). New York: MacMillan.

Schwarz, W. (2006). *Heuristische Strategien des Problemlösens*. Münster: WTM Verlag.

Schreiber, A. (2011). *Begriffsbestimmungen – Aufsätze zur Heuristik und Logik mathematischer Begriffsbildung*. Berlin: Logos Verlag.

Stark, R.; Gruber, H.; Renkl, A. & Mandl, H. (2000). Instruktionale Effekte einer kombinierten Lernmethode – Zahlt sich die Kombination von Lösungsbeispielen und Problemlöseaufgaben aus? *Zeitschrift für Pädagogische Psychologie*, 14 (4), S. 206–218.

Stark, R.; Tyroller, M.; Krause, U.-M. & Mandl, H. (2008). Effekte einer metakognitiven Promptingmaßnahme beim situierten, beispielbasierten Lernen im Bereich Korrelationsrechnung. *Zeitschrift für Pädagogische Psychologie*, 22 (1), S. 59–71.

Sweller, J. (1990). On the Limited Evidence for the Effectiveness of Teaching General Problem-Solving Strategies. *Journal for Research in Mathematics Education*, 21 (5), S. 411–415.

Tietze, U.-P.; Klika, M. & Wolpers, H. (2000). *Mathematikunterricht in der Sekundarstufe II – Band 1*. Braunschweig: Vieweg.

Vollrath, H.-J. (1992). Zur Rolle des Begriffs im Problemlöseprozeß des Beweisens. *Mathematische Semesterberichte*, 39, S. 127–136.

Winter, H. (1995). Mathematikunterricht und Allgemeinbildung. *Mitteilungen der Gesellschaft für Didaktik der Mathematik*, 61, S. 37–46.

Zech, F. (1992). *Grundkurs Mathematikdidaktik*. Weinheim und Basel: Beltz.

Anschrift der Verfasser

Jun.-Prof. Dr. Benjamin Rott
Universität Duisburg-Essen
Didaktik der Mathematik
45117 Essen
E-Mail: benjamin.rott@uni-due.de

Prof. Dr. Thomas Gawlick
Leibniz Universität Hannover
Institut für Didaktik der Mathematik und Physik
30167 Hannover
E-Mail: gawlick@idmp.uni-hannover.de

Eingang Manuskript: 03.03.2014
Eingang überarbeitetes Manuskript: 09.06.2014
Online verfügbar: 02.03.2015

Anhang

Beispiel für eine Aufgabe zum Heurismus (Z) „Zeichnung anfertigen“:

Aufgabe Z3

Gegeben sind ein Schachbrett und 32 Dominosteine. Da jeder Dominostein genau zwei der Schachbrettfelder abdeckt, kann das ganze Brett vollständig mit Dominosteinen zugelegt werden.

Nun werden zwei diagonal gegenüberliegende Ecken und ein Dominostein entfernt. Ist es möglich, die 31 Dominosteine so auf dem Brett zu platzieren, dass alle verbleibenden 62 Schachbrettfelder abgedeckt sind?

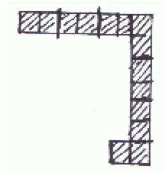
Beispiel für eine Musterlösung inkl. der strategischen Hinweise am rechten Rand:

Musterlösung:

Das Originalschachbrett kann man einfach zumauern – durch das Entfernen zweier Ecksteine geht das nicht mehr. Aber es gibt viele andere Legevarianten wo nicht klar ist wo sie hinführen. Ich mache zunächst eine Skizze

Überdecke mit

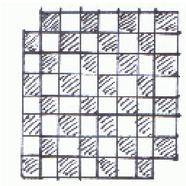
Mache eine Zeichnung!



Kann man die Steine vielleicht einfach als Schlange legen?

Trenne die Teile der Bedingung. Kannst du sie hinschreiben / einzeichnen?

Hm – mir fällt auf: jeder Stein bedeckt ein schwarzes und ein weißes Feld. Und: benachbarte Felder sind immer verschiedenfarbig. Also heißt die Aufgabe



überdecke mit

Das beschnittene Schachfeld hat also 30 weiße und 32 schwarze Felder – 31 Dominosteine können nur 31 weiße und 31 schwarze Felder belegen. Also kann man nicht alle Felder mit Dominosteinen belegen!

Abbildung 4: Musterlösung zur Aufgabe Z3 inkl. strategischer Kommentare am rechten Rand, die nur die Experimentalgruppe zu sehen bekam

Beispiel für eine Aufgabe zum Heurismus (S) „Spezialfall betrachten“:

Aufgabe S2

Gegeben sind zwei Kreise. Konstruiere eine gemeinsame Tangente.

Beispiel für eine Aufgabe zum Heurismus (R) „Rückwärtsarbeiten“:

Aufgabe R1

Beweise: In jedem Dreieck zerlegt die Winkelhalbierende eines Winkels die gegenüberliegende Seite im Verhältnis der anliegenden Seiten.

Beispiel für eine Aufgabe zum Heurismus (B) „Weglassen einer Bedingung“:

Aufgabe B1

Ein Quadrat wird so in ein Dreieck einbeschrieben, dass zwei Eckpunkte auf der Grundseite und die anderen beiden Eckpunkte auf den beiden anderen Seiten liegen. Wie lässt sich dieses Quadrat konstruieren?

Beispiel für eine Aufgabe zum Heurismus (V) „Ähnliche Aufgabe mit weniger Variablen“:

Aufgabe V3

Zeige, dass für alle reellen Zahlen $x, y \geq -1$ und jede ganze Zahl $n \geq 0$ gilt:

$(1 + x + y + xy)^n \geq 1 + n(x + y) + n^2xy$