Universität Duisburg-Essen

Fakultät Bildungswissenschaften

Lehrstuhl für Lehr-Lernpsychologie

Analytisches Problemlösen

Labor- und feldexperimentelle Untersuchung von Aspekten der kognitiven Potenzialausschöpfungshypothese

Dissertation

zur Erlangung des Grades Dr. phil.

vorgelegt von Dipl.-Psych. Florian Buchwald

geboren am 26.12.1984 in Bonn

Erstgutachter: Prof. Dr. Detlev Leutner, Universität Duisburg-Essen

Zweitgutachter: Prof. Dr. Joachim Wirth, Ruhr-Universität Bochum

Tag der mündlichen Prüfung: 16. März 2015

Danksagung

Mein herzlicher Dank gilt allen beteiligten Personen und Institutionen, die diese Dissertation ermöglicht haben.

Ich danke Prof. Dr. Dr. h. c. Detlev Leutner für die lehrreiche Zeit am Lehrstuhl für Lehr-Lernpsychologie sowie die Betreuung und Begutachtung dieser Dissertation. Prof. Dr. Joachim Wirth möchte ich für seine kritischen Rückmeldungen und die Begutachtung danken.

Jens Fleischer danke ich für sein hilfreiches Feedback, den wertvollen Erfahrungsaustausch und die Zusammenarbeit im DFG-Projekt „Kompetenzstrukturen fächerübergreifenden und fachlichen Problemlösens“.

Für konstruktive Diskussionen und ihr Feedback im Laufe der letzten Jahre möchte ich mich herzlich bei Prof. Dr. Stefan Rumann, Prof. Dr. Joachim Funke, Dr. Thilo Kleickmann, Prof. Dr. Günter Daniel Rey, Markus Thielgen, Andreas Fischer und Julia Kobbe bedanken. Mein Dank gilt ferner Christian Spoden, der mich motiviert hat, meine Kenntnisse der Programmiersprache R weiter zu vertiefen.

Derya Bayar, Kirsten Breuer, Björn Panek, Christian Lubsczyk, Moritz Rathjen, Ralf Ricken, Cepideh Saadat, Denise Schulz, Romina Skupin, und Volker Zischka haben als studentische Hilfskräfte diese Arbeit durch viele Fleißarbeiten unterstützt, vor allem durch die Eingabe etlicher Testhefte. Kirsten Breuer und Derya Bayar danke ich zudem für die Mithilfe bei der teils mühevollen Telefonakquise. Ich danke zudem allen Beta-Testern der Laborexperimente. Thomas Kiefer, Tobias Ottenheym und Alexander Robitzsch gaben mir hilfreiche Tipps zur Benutzung verschiedener R-Pakete für die Datenauswertung.

Derya Bayar, Helene Becker, Kirsten Breuer, Jennifer Chmielnik, Jan Dworatzek, Christian Führer, Jana Goertzen, Anne Hommen, Michael Kalkowski, Julia Kobbe, Kinga Oblonczyk, Ralf Ricken, Sönke Schröder, Jana Wächter, Sabrina Windhövel und Kristina Wolferts haben als Trainer und Testleiter in der feldexperimentellen Trainingsstudie mitgearbeitet.[1] Vielen Dank dafür! Dr. Theresa Dicke hat dankenswerterweise ihre erprobten Classroom-Management-Schulungsmaterialien zur Vorbereitung der Trainer zur Verfügung gestellt. Bei Dr. Viktoria

[1] Sofern nicht anders gekennzeichnet wird aus Gründen der besseren Lesbarkeit in dieser Arbeit das generische Maskulinum verwendet.

Arling bedanke ich mich für die Nutzung ihres Planungskompetenztrainings und die angenehme Kooperation. Benjamin Klein und Dr. Ferdinand Stebner danke ich sehr für den hilfreichen Erfahrungsaustausch zu zeitaufwändigen Trainings im Schulkontext.

Herzlich danken möchte ich den teilnehmenden Schulen mit allen beteiligten Personen, ohne die der empirische Teil dieser Arbeit nicht möglich gewesen wäre.

Abschließend möchte ich mich für das sorgfältige Korrekturlesen bei Derya Bayar, Kirsten Breuer, Klaus Buchwald, Monika Buchwald, Andreas Hoffmann, Tobias Ottenheym und Markus Thielgen bedanken.

Diese Dissertation wurde ermöglicht durch eine Sachbeihilfe der Deutschen Forschungsgemeinschaft (Kennz.: LE 645/12-2, LE 645/12-3, RU 1437/4-3) im Schwerpunktprogramm „Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen“ (SPP 1293).

Inhaltsverzeichnis

Zusammenfassung

Schülerinnen und Schüler in Deutschland erzielten beim „Programme for International Student Assessment“ (PISA) im Jahr 2003 überdurchschnittliche Ergebnisse in der Domäne (fächerübergreifendes) Problemlösen. Die Leistungen in den fachlichen Domänen blieben im Vergleich dazu jedoch hinter den Erwartungen zurück. Diese Diskrepanz wird insbesondere für die Domänen Mathematik und Naturwissenschaften als Zeichen mangelnder kognitiver Potenzialausschöpfung interpretiert (Leutner, Klieme, Meyer & Wirth, 2004; OECD, 2004). In einer Serie von drei Experimenten werden Aspekte dieser kognitiven Potenzialausschöpfungshypothese für die mathematische Domäne labor- und feldexperimentell untersucht. Dabei steht die Frage möglicher Trainings- und Transfereffekten (von Komponenten) fächerübergreifenden Problemlösens auf mathematisches Problemlösen bzw. Modellieren im Fokus. Die Ergebnisse der beiden Laborexperimente zeigten teilweise Trainingseffekte, jedoch keine Transfereffekte auf die mathematische Domäne. Das Feldexperiment zeigt v. a. für leistungsschwache SuS kleine Trainingseffekte und kleine Transfereffekte auf die mathematische Domäne. Die Experimente werden in ihren Limitationen und Konsequenzen für zukünftige Forschung anschließend kritisch diskutiert.

Abstract

Students in Germany achieved above-average results in the domain of cross-curricular problem solving at "The Programme for International Student Assessment" (PISA). The results in the subject-specific domains lag behind the expectations based on the problem solving results. This discrepancy is interpreted especially for the domains mathematics and science as an indication that there exists a lack of exploitation of cognitive potential (Leutner, Klieme, Meyer & Wirth, 2004; OECD, 2004). In a series of three experiments aspects of this cognitive exploitation hypothesis for the domain of mathematics are analyzed by means of labor and field experiments. Thereby, the question of possible training and transfer effects from (components of) cross-curricular problem solving to mathematical problem solving is particularly focused. The results of both labor experiments show in part effects of training, but no effects of transfer to the mathematical domain. The field experiment shows small effects of training and transfer effects to the mathematical domain for low-achieving students. The results are discussed with respect to limitations and consequences for future research.

Tabellenverzeichnis

Abbildungsverzeichnis

1 Einleitung und Struktur der Arbeit

Um die Anforderungen sich schnell ändernder Gesellschaften erfolgreich bewältigen zu können, gewinnen breit definierte Schlüsselqualifikationen wie Problemlösekompetenz an Bedeutung (Klieme, 2004). Die Wichtigkeit des Problemlösens als *fachspezifische* Kompetenz wird vor allem in der Mathematik (Blum, Drüke-Noe, Hartung & Köller, 2006; KMK, 2004a, 2004b, 2012; NCTM, 2000) und den Naturwissenschaften (AAAS, 1993; KMK, 2005a, 2005b, 2005c) hervorgehoben. Problemlösen wird zudem als *fächerübergreifende* Kompetenz angesehen, die wichtig ist für erfolgreiches Lernen innerhalb der Schule sowie für die Bewältigung beruflicher und privater Herausforderungen (Levy & Murnane, 2005; OECD, 2004b, 2013). Angesichts der Bedeutung des Problemlösens als fächerübergreifende und fachspezifische Kompetenz wurde Problemlösekompetenz im Rahmen internationaler Bildungsvergleichsstudien erfasst (z. B. beim *Programme for International Student Assessment* (PISA), OECD, 2004a, 2004b, 2014).

Bei PISA 2003 erzielten Schülerinnen und Schüler (SuS) in Deutschland im internationalen Vergleich überdurchschnittliche Ergebnisse im Bereich fächerübergreifenden Problemlösens (OECD, 2004a, 2004b). Bei den fachlichen Tests in den Bereichen Lesen, Mathematik und Naturwissenschaften erzielten die SuS jedoch trotz konzeptuell ähnlicher Prozesse und Anforderungen nur durchschnittliche Ergebnisse. Dieses bessere Abschneiden im Bereich Problemlösen relativ zu fachlichen Kompetenz, z. B. in der Mathematik, wird im PISA-Kontext als ungenutztes Potenzial interpretiert (z. B. Walter, Ramm, Zimmer, Heidemeier & Prenzel, 2006) und führte post hoc zur *kognitiven Potenzialausschöpfungshypothese* (Leutner, Klieme, Meyer & Wirth, 2004; OECD, 2004b). Diese Hypothese postuliert kognitives Potenzial von SuS in Deutschland, das sich bislang nicht in den fachlichen Leistungen z. B. im Schulfach Mathematik zeige. Das vorhandene kognitive Potenzial, das im Test zum Problemlösen sichtbar wird, würde nicht hinreichend zum Aufbau fachbezogener Kompetenz genutzt. Diese mangelnde Potenzialausschöpfung wird insbesondere für das Fach Mathematik in Deutschland betont (Fleischer, Wirth, Rumann & Leutner, 2010; Leutner, Fleischer & Wirth, 2006; Leutner et al., 2004; Leutner, Klieme, Meyer & Wirth, 2005), da u. a. das ungenutzte Potenzial unter allen 29 Teilnehmerländern nur in Ungarn und Japan größer war (Leutner et al., 2004; OECD, 2004b). Zudem zeigte sich das Potenzial vor allem bei leistungsschwachen SuS (Leutner et al., 2004, 2005) und auch in den Naturwissenschaften (Leutner et al., 2006; Rumann, Fleischer, Stawitz, Wirth & Leutner, 2010).

Ziel dieser Dissertation ist die experimentelle Untersuchung von Aspekten dieser kognitiven Potenzialausschöpfungshypothese für die Domäne Mathematik (Leutner et al., 2004; OECD, 2004b). Dazu wurde untersucht, ob das Training weniger, zentraler Komponenten der fächerübergreifenden Problemlösekompetenz (konditionales Wissen, Planungsfähigkeit, Handlungswissen) diese Komponenten der fachbezogenen Problemlösekompetenz im Fach Mathematik und die mathematische (Problemlöse-)Kompetenz insgesamt steigert. Zusätzlich wurden differentielle Effekte analysiert, um der Frage nachzugehen, ob leistungsschwächere SuS und SuS mit hohem kognitivem Potenzial im Sinne der kognitiven Potenzialausschöpfungshypothese besonders profitieren.

Im theoretischen Hintergrund (Kapitel 2) werden Arbeitsdefinitionen gegeben und ausgewählte Theorien des Problemlösens skizziert (Kapitel 2.1). Im Anschluss werden die Operationalisierung der fächerübergreifenden und mathematischen Problemlösekompetenz bei PISA 2003 (Kapitel 2.2), empirische Befunde zur fächerübergreifenden und mathematischen Problemlösekompetenz bei PISA 2003 (Kapitel 2.3) sowie die kognitive Potenzialausschöpfungshypothese und alternative Erklärungen inklusive der jeweiligen Befundlage dargestellt (Kapitel 2.4). Anschließend werden Theorien und Ergebnisse der Trainings- und Transferforschung skizziert (Kapitel 2.5) und die Hauptfragestellungen der Arbeit vorgestellt (Kapitel 2.6).

Im empirischen Teil werden Aspekte dieser kognitiven Potenzialausschöpfungshypothese in einer aus drei Experimenten bestehenden Studie untersucht (Kapitel 3). Experiment 1 ist ein Pilotexperiment unter Laborbedingungen. In Experiment 2 wird Experiment 1 in modifizierter Form unter Laborbedingungen repliziert. Die Ergebnisse führten zu einem mehrwöchigem Feldexperiment im Schulkontext, das die kognitive Potenzialausschöpfungshypothese direkter untersuchte.

In der zusammenfassenden Diskussion (Kapitel 4) werden die wichtigsten Ergebnisse zusammengefasst (Kapitel 4.1), die theoretischen und praktischen Erträge der Arbeit dargestellt (Kapitel 4.2), die Grenzen der Arbeit beschrieben und ein Ausblick auf potentielle zukünftige Forschung gegeben (Kapitel 4.3).

2 Theoretischer Hintergrund

2.1 Probleme und Problemlösen

Ein Problem besteht aus drei Elementen: Einer Problemsituation, auch Istzustand oder Ausgangszustand genannt, einem Soll- oder Zielzustand und einem nicht unmittelbar einsichtigen Lösungsweg (Funke, 2003; Mayer, 1992), d. h. die Zielerreichung ist durch eine „Barriere“ (Dörner, 1976) blockiert. Beispielsweise weiß der Problemlösende zu Beginn womöglich nicht, wie er bekannte Operatoren (Schritte, Mittel) neu kombinieren bzw. in sinnvoller Reihenfolge anwenden muss, um einen angestrebten bekannten Zielzustand zu erreichen (z. B. beim Schach oder Routenplanen; *Interpolationsproblem* nach Dörner, 1976). Auch kann das Ziel zwar bekannt sein, aber dem Problemlöser sind die notwendigen Operatoren zumindest teilweise unbekannt bzw. werden für irrelevant gehalten (z. B. bei Denksportaufgaben wie Dunckers Kerzenproblem oder beim Neun-Punkte-Problem; *Syntheseproblem* nach Dörner, 1976).

Alle drei Elemente können mehr oder weniger gut definiert sein. Nimmt man vereinfachend an, dass Ausgangszustand, Operatoren, Zielzustand entweder bekannt oder unbekannt sind, so ergeben sich insgesamt $2^3 = 8$ Klassen von Problemen (z. B. Bruder, 2003). Ob insbesondere die Problemklassen ohne bekannten Zielzustand als Probleme im engeren Sinne zählen, ist umstritten, da bei derartigen Situationen kein Problembewusstsein vorhanden ist (Funke, 2003). In der Praxis ist teilweise unbekannt, ob überhaupt eine Lösung für ein bestimmtes Problem existiert (z. B. Finden der physikalischen Weltformel) oder ob die Lösung jemals realisiert wird (z. B. Weltfrieden, Quantencomputer). In der psychologischen Forschung und im Schulunterricht (z. B. im Fach Mathematik) hingegen haben es Probanden bzw. SuS traditionell meist mit gut definierten Problemen (*well-defined problems*) zu tun, d. h. mit Problemen, bei denen zumeist Ausgangszustand, zulässige Operatoren und Ziele klar definiert sind, und eher selten mit offeneren, schlecht definierten, Problemen (Rourke & Sweller, 2009).

2.1.1 Aufgaben und Probleme

Zur Präzisierung des Problembegriffs ist es hilfreich, die Begriffe *Problem* und *Aufgabe* abzugrenzen. Probleme sind als nicht-routinemäßige Anforderungen (produktives Verhalten) und Aufgaben als routinemäßige Anforderungen (reproduktives Verhalten) beschreibbar (Funke, 2003). In der englischsprachigen Mathematikdidaktik unterscheidet man analog zwischen *problems* und *excercices* (Schoenfeld, 1983b).

Wenn nicht zwischen Problemen und Aufgaben unterschieden werden soll, wird in dieser Arbeit der Begriff *Item* verwendet, der – im Gegensatz zum kybernetischen Systembegriff von Vollmeyer und Funke (1999) – unabhängig von spezifischen theoretischen Vorstellungen und Modellen verwendbar ist. *Item* ist der Oberbegriff für „eine Aufgabe, Frage oder Feststellung“ (Kubinger, Rasch & Yanagida, 2011, S. 555). Ein Item ist strukturell charakterisiert durch einen Itemstamm, der den Ausgangszustand beschreibt, eine (kognitive) Handlungsaufforderung und ein bestimmtes Antwortformat (z. B. offene Aufgabe, *multiple choice*).

Ob ein Item *i* für eine konkrete Person *P* eine Aufgabe oder ein Problem darstellt, ist nicht mit Sicherheit feststellbar, *bevor* die Person mit dem Item konfrontiert worden ist.[2] Dasselbe Item kann also für eine Person ein Problem darstellen und für eine andere Person mit entsprechendem Vorwissen bzw. Routinen eine Aufgabe. Im Prinzip ist aber eine Prognose möglich, ob ein Item eine Aufgabe oder ein Problem darstellt. Denn was als routinemäßige Anforderung gilt und was nicht, kann beispielsweise im Vorfeld für eine bestimmte Zielgruppe (z. B. 15-jährige SuS bei PISA) bestimmt werden. Möglichkeiten sind durch Expertenurteile, auf Grundlage vorhergehenden Unterrichts, durch empirische Vorstudien, Selbsteinschätzung durch die Probanden vor, während oder nach der Bearbeitung, durch Bezug auf Bildungsstandards (z. B. KMK, 2003) oder Kombinationen verschiedener Vorgehensweisen (z. B. OECD, 2003) gegeben.

Wenn eine Person mit einem Item konfrontiert wird, entscheidet sich, ob es für sie ein Problem darstellt. Verfügt sie über routinierte Handlungsweisen etwa in Form psychologischer *Skripte* (Schank & Abelson, 1977), d. h. im Langzeitgedächtnis gespeichertes Wissen über den typischen Ablauf einer spezifischen Handlungsepisode (z. B. Restaurantbesuch, Benzin-Tanken, Online-Shopping), oder *Schemata,* d. h. kognitive Strukturen, die die Zuordnung von Problemen zu bestimmten Problemkategorien erlauben (Kalyuga, 2009; Paas & van Merriënboer, 1994), oder mathematische *Algorithmen* zur Lösung, so handelt es sich um eine Aufgabe.

[2] Selbstverständlich ist es möglich, SuS vor Handlungsaufforderungen zu stellen, die für sie mit Sicherheit Probleme darstellen, wenn man z. B. eindeutig nicht altersgerechte Probleme präsentiert, die hohes Vorwissen erfordern. Wer kein Latein kann, kann ohne Hilfe auch kein lateinisches Gedicht übersetzen. Jemanden in eine solche Situation zu bringen, ist jedoch im Allgemeinen weder pädagogisch sinnvoll noch psychometrisch informativ.

Algorithmen beschreiben Regeln (z. B. zur schriftlichen Division), mit denen eine bestimmte Klasse von Problemen bei korrekter Ausführung sicher lösbar ist. Bei Problemen hingegen kennt der Problemlösende die nötigen Schemata bzw. Lösungsalgorithmen (zunächst) nicht oder kann sich nicht unmittelbar daran erinnern. Womöglich gibt es auch keinen Algorithmus. Theoretisch sind dabei mehrere Fälle denkbar, warum es keinen Algorithmus für ein Problem gibt. Das Problem kann *unlösbar* sein, d. h. es kann kein Algorithmus zur Lösung existieren. Das Problem kann *lösbar* sein, ohne dass ein formalisierter Algorithmus verfügbar ist, oder das Problem ist lösbar, jedoch nicht mittels eines Algorithmus‘. Insbesondere im letzten Fall spielen *Heuristiken* (Chu, Li, Su & Pizlo, 2010; Pólya, 1945) und *Strategien* als metakognitive Elemente eine wichtige Rolle (Carson, 2007; Gobet, 2005; Schoenfeld, 1987, 1992; Schraw, 1998). Heuristiken sind einfache Methoden, die helfen sollen, schnell eine Problemlösung zu finden, jedoch im Gegensatz zu Algorithmen weder das Finden einer Lösung noch ihre Optimalität garantieren (Dunbar, 1998; Hertwig, 2006). Beispiele für Heuristiken sind die Zerlegung in Teilziele oder die Suche nach Analogien. Ist eine Heuristik in vielen unterschiedlichen Problemsituationen einsetzbar, spricht man auch von einer Strategie (Hertwig, 2006). Ob Algorithmen oder Heuristiken bessere Denkmodelle sind, lässt sich nicht generell sagen. Krulik und Rudnick (1980), zwei Verfechter von Problemlöseheuristiken, betonen, dass Algorithmen ohne tiefes konzeptuelles Verständnis des Problems und des Algorithmus’ angewandt werden können und bei neuen Problemarten scheitern (Carson, 2007).

Ob jemand ein Item löst oder nicht, sagt im strengen Sinne nichts darüber aus, ob das Item für eine Person ein Problem oder lediglich eine Aufgabe darstellt, solange Prozess-Informationen darüber fehlen, warum bzw. wie jemand ein Item gelöst hat. Die folgenden Beispiele skizzieren diesen Gedanken, der mit der Idee einer probabilistischen Item-Lösungswahrscheinlichkeit verknüpft ist (vgl. *item response theory*, z. B. Embretson & Reise, 2000). Wenn eine Person über die notwendigen Routinen verfügt, sollte ihre Chance steigen, ein Item (hier: eine Aufgabe) zu lösen. Es besteht aber aus verschiedenen Gründen (z. B. Zeitknappheit, Flüchtigkeitsfehler, Motivationsmangel, *faking bad*) die Möglichkeit, dass das Item nicht gelöst wird. Wenn eine Person nicht über die notwendigen Routinen verfügt, sollte ihre Chance sinken ein Item (hier: ein Problem) zu lösen. Es besteht aber aus verschiedenen Gründen (z. B. Rateeffekte, Betrug, Lerneffekte während der Aufgabenbearbeitung, Anwendung bestehenden Wissens) die Möglichkeit, dass das Item doch gelöst wird. Die Schwierigkeit eines Items ist also ebenfalls kein *sicherer* Indikator dafür, ob ein Item für eine Person ein Problem oder lediglich eine Aufgabe ist.

In dieser Arbeit werden zur Diagnostik des Problemlösens hauptsächlich Items aus *large-scale assessments* verwendet, die für 15-jährige SuS als valide Indikatoren der fächerübergreifenden bzw. mathematischen Problemlösekompetenz gelten (Leutner, Fleischer, Spoden & Wirth, 2007, OECD, 2004a, 2004b).

2.1.2 Analytische und komplexe Probleme

Es gibt in der Geschichte der Psychologie zahlreiche Versuche, Probleme zu klassifizieren (siehe z. B. Funke, 2003; Jonassen, 2012; Robertson, 2001, für Taxonomien). Die gewählte Klassifikation der Probleme hängt von der theoretischen Position bzw. der diagnostischen Intention ab. Eine allgemeine erkenntnistheoretische Überlegenheit einer Klassifikation ist nicht ersichtlich.

Insbesondere im deutschen Sprachraum trennt man zwischen analytischen und komplexen Problemen bzw. analytischem und komplexem Problemlösen (Funke, 2003, 2006a; Wirth & Klieme, 2003). Nachfolgend wird nur diese Klassifikation vorgestellt, da beide Arten von Problemen im Rahmen der PISA-Studien eingesetzt wurden (Leutner et al., 2005; OECD, 2004b, 2014) und das zu untersuchende Befundmuster aus PISA 2003 stammt.

Bei analytischen Problemen sind alle lösungsrelevanten Informationen in der Problemsituation enthalten oder durch schlussfolgerndes Denken und Vorwissen erschließbar. Ein PISA-Item zum analytischen Problemlösen besteht z. B. in der Aufteilung von Kindern und Betreuern auf Schlafzimmer während einer Ferienfreizeit unter der Berücksichtigung von Nebenbedingungen wie Bettenanzahl, Geschlecht der Kinder und Geschlecht der Betreuer im Zimmer (OECD, 2004b). Diese Art von Problemen ist statisch. Es kommt also nicht zu dynamischen Veränderungen des Problemzustandes (vgl. Dynamik bei komplexen Problemen). Analytische Probleme erfordern oftmals die Strukturierung, Repräsentation und Integration verschiedener Informationen. Analytisches Problemlösen umfasst Problemsituationen, in denen Informationsverarbeitung und schlussfolgerndes Denken gefordert sind, da kognitive Anforderungen wie „Suchen, Erfassen, Systematisieren, Ordnen, Evaluieren, schlussfolgernde[s] Verarbeiten und Kombinieren von Informationen und […] Planen“ (Klieme, Artelt & Stanat, 2001, S. 206) gestellt werden.

Beim komplexen Problemlösen ist eine Exploration der Problemsituation durch Interaktion mit dem System notwendig (z. B. Bedienung eines unbekannten Fahrkartenautomaten; OECD, 2014). Die Problemsituation komplexer Probleme kann sich durch die Interaktion des Problemlösers mit dem System ändern (z. B. neues Untermenü, Wechsel der Sprache der Menüführung).

Komplexe Probleme sind zumeist durch fünf Merkmale gekennzeichnet, wobei nicht immer alle fünf Merkmale gegeben sein müssen (Funke, 2003, 2006a):

1. Komplexe Probleme weisen eine hohe Anzahl an Variablen auf, die eine gewisse *Komplexität (complexity)* erzeugen. Bei der Tailorshop-Simulation (Danner et al., 2011) etwa müssen 24 Variablen (u. a. Einkaufs- und Verkaufspreise, Warenbestand, Reparaturkosten, Angestelltengehälter) berücksichtigt werden, um erfolgreich ein Geschäft zu führen.

2. Komplexe Probleme sind durch *Vernetztheit* (*interconnectedness)* gekennzeichnet, d. h. es gibt Querverbindungen und Zusammenhänge (z. B. bei Problemen der Logistik), die zudem auch nicht-linear sein können.

3. Ein komplexes Problem weist eine gewisse *Dynamik* (*dynamics*) auf, d. h. die Problemsituation kann sich im zeitlichen Verlauf ändern. Gründe für die veränderte Problemsituation können Handlungen des Problemlösers oder anderer Beteiligter sein, zufällig zustande kommen oder durch eigendynamische Prozesse entstehen (z. B. Wetteränderung).

4. Komplexe Probleme sind daher durch partielle *Intransparenz* (*intransparency*) kennzeichnet, d. h. zumindest anfangs stehen nicht alle zur Lösung des Problems erforderlichen Informationen zur Verfügung. Diese Undurchsichtigkeit der Zusammenhänge erschwert Prognosen (z. B. über den optimalen Fahrpreis).

5. Komplexe Probleme zeichnen sich durch *Polytelie* (*polytely)* aus, d. h. es kann mehr als ein Ziel geben. Diese Ziele können teilweise in Konflikt zueinander stehen (z. B. Gewinnmaximierung und umweltfreundliche Produktion).

Komplexe Probleme werden auch als *dynamische Probleme* oder *interaktive Probleme* bezeichnet, wobei die Verwendung nicht konsistent ist (vgl. Greiff & Fischer, 2013). Die Unterscheidung in analytische und komplexe Probleme darf nicht zu dem Fehlschluss führen, komplexe Probleme seien per se schwierig, wohingegen analytische Probleme einfach zu lösen seien, auch wenn sie manchmal in Abgrenzung zu komplexen Problemen als *einfache Probleme* bezeichnet werden (Öllinger & Knoblich, 2006). Die Schwierigkeit von Aufgaben bzw. Problemen lässt sich zwar theoretisch beurteilen, aber nur auf der Grundlage empirischer Ergebnisse berechnen (vgl. Kapitel 2.1.1).

Die in dieser Arbeit verwendeten Testitems sind ausnahmslos analytische Probleme.[3] Die meisten stammen aus dem Itempool von PISA 2003 (OECD, 2003, 2004b). Es handelt sich um eine spezielle Klasse von Problemen: Ausgangssituation und Ziel sind relativ klar definiert (*well-defined problems*). Diese Probleme sind – wie meistens im Mathematikunterricht und mathematikdidaktischer Forschung – *lösbar* und das binnen weniger Minuten (Sieprinska, 2004). Zudem ist das zur Lösung nötige Vorwissen aus Expertensicht überschaubar, da als Zielgruppe der Untersuchung, wie bei PISA, SuS ausgewählt worden sind.

2.1.3 Der Prozess des Problemlösens

Nach der einleitenden Analyse des Problembegriffs und der Abgrenzung von analytischen zu komplexen Problemen wird in diesem Unterkapitel der Prozess des Problemlösens dargestellt.

Aus Perspektive der Graphentheorie ist eine Problemlösung formal definiert als *Pfad*, d. h. eine mathematische Folge von Kanten, vom *Ausgangsknoten* zum *Zielknoten*, wobei der *Problemraum* als Graph gegeben ist (vgl. auch Newell & Simon, 1972). Knoten repräsentieren alle möglichen Zustände des Problems, wobei zwischen zwei Knoten a und b des Graphs genau dann eine Kante $k(a,b)$ existiert, wenn die Problemzustände a und b ineinander durch eine zulässige Operatoranwendung transformierbar sind (Schmid, 2008). Problemlösen ist also in diesem informationstheoretischem Paradigma als die Suche im Problemraum nach einem Pfad vom Ausgangszustand zum Zielstand beschreibbar (Schmid, 2008).

[3] Wenn nicht explizit anders gekennzeichnet meint Problemlösen im Folgenden analytisches Problemlösen. Komplexes Problemlösen ist nicht Gegenstand der vorliegenden Arbeit.

Man beachte, dass diese Definition keine Aussagen über die *Qualität* der Problemlösung und die Anzahl möglicher Lösungen macht.

Zahlreiche Autoren (z. B. Frensch & Funke, 1995; Wirth & Klieme, 2003) betonen, dass sowohl die Definition des Problemlösens und die Konzeption assoziierter Konstrukte in der Literatur sehr unterschiedlich ausfällt. Das überrascht wenig, da psychologische Prozesse und Konstrukte aus unterschiedlichen theoretischen Perspektiven untersucht werden können (Dörner, 1984). Als weitgehend gemeinsame Kernelemente des Problemlösens lassen sich dennoch Absichten, Pläne, Barrieren und mangelnde Routinen festhalten (Funke, 2003). Der Problemlöser versucht, die Barriere (Lücken im Handlungsplan, Zielblockierung, Ziel-Mittel-Inkongruenz) in einer nicht routinemaßigen Situation zu überwinden, um sein Ziel zu erreichen. Dabei spielen das Verstehen der Problemsituation, der Aufbau einer gedanklichen Repräsentation und deren schrittweise Veränderung, die durch Planen und Schlussfolgern erleichtert wird, eine zentrale Rolle (Deutsches PISA Konsortium, 2003; Klieme, Funke, Leutner, Reimann & Wirth, 2001).

Der Prozess des Problemlösens von analytischen fächerübergreifenden und mathematischen Problemen lässt sich als eine Abfolge prototypischer Phasen beschreiben (Carlson & Bloom, 2005; Dewey, 1933; Krulik & Rudnick, 1987; Pólya, 1945; Young, 1924; vgl. Tabelle 1). Vor einem Versuch der Problemlösung muss das Problem zunächst als Problem mit seinen relevanten Bedingungen erkannt, verstanden und mental repräsentiert werden (z. B. Mayer, 1992). Dann erfolgt die Planung eines Lösungsweges, wobei diese Planung gedanklich, mündlich oder schriftlich geschehen kann. Die Ausführung dieses Planes, die durch metakognitive Prozesse überwacht werden sollte, führt zu einem Ergebnis, das mit dem Ausgangszustand des Problems verglichen wird. Gegebenenfalls ist eine erneute Planung, Ausführung und Bewertung nötig. Abschließend ist das Ergebnis angemessen zu kommunizieren (vgl. Leutner et al., 2004).

Diese Prozesse werden nicht nur für analytisches Problemlösen postuliert. Auch komplexes Problemlösen erfordert konzeptionell die beschriebenen Phasen des Problemlösens, stellt aber typischerweise erhöhte Anforderungen an selbstreguliertes Lernen und bedarf der feedbackgestützten adaptiven Anpassungen des eigenen Handelns an die sich verändernde Problemsituation (*learning by doing*; Wirth & Klieme, 2003).

Die Problemlösung kann je nach Problem auf verschiedene Art und Weise erfolgen. Edelmann und Wittmann (2012) unterscheiden fünf verschiedene Ansätze zur Lösung von Problemen:

1. *Problemlösen durch Versuch und Irrtum*, insbesondere bei intransparenten Problemen (siehe auch Middleton, 2002).
2. *Problemlösen durch Umstrukturieren* der Problemsituation (vgl. Aha-Erlebnis der Gestaltpsychologie, Situationsanalyse, Mittelaktualisierung; siehe z. B. Knoblich & Öllinger, 2006 für einen Überblick), auch mit Hilfe von Repräsentationswechseln, z. B. mittels Skizzen, Tabellen oder Formeln (Bruder, 2002).
3. *Problemlösen durch Anwendung von Strategien* (z. B. Vorwärtsarbeiten, Mittel-Ziel-Analyse, Analogien, Repräsentationswechsel; King, 1991).
4. *Problemlösen durch Kreativität*, d. h. die Anwendung neuer bzw. ungewöhnlicher und nützlicher Lösungen (Mayer, 1999).
5. *Problemlösen durch Systemdenken* (v.a. bei komplexen Problemen, vgl. Funke, 2006b).

Die verschiedenen Vorgehensweisen beim Problemlösen schließen sich im Laufe des Problemlöseprozesses nicht aus, d. h. sie sind je nach Problem kombinierbar.

2.1.4 Mathematisches Problemlösen

Die getrennte Darstellung von Theorien bzw. Modellen des Problemlösens nach Fachdisziplinen (Psychologie, Mathematik) ist didaktischer Natur. Sie soll nicht verschleiern, dass es inhaltlichen Austausch und wechselseitigen Einfluss gegeben hat und gibt.

Die Schul-Mathematik behandelt keine bislang ungelösten mathematischen Probleme, wie etwa die sogenannten Millennium-Probleme der Mathematik (Carlson, Jaffe & Wiles, 2006), sondern stellt SuS typischerweise vor mathematische Probleme die, fachwissenschaftlich betrachtet, lösbar und historisch längst gelöst sind. Auch wenn die Bedeutung des Problemlösens für die Mathematik unbestritten ist, so sind die Semantik des Problemlösebegriffs und die Bedeutung von Problemlösen für den Mathematikunterricht vielfältig (Arcavi & Friedlander, 2007; Freudenthal, 1981; Leuders, 2010). Lerntheoretisch kann man fast jedes Lernen als Problemlöseprozess auffassen (Leuders, 2010). Dieses Begriffsverständnis ist für mathematisches Problemlösen jedoch zu breit gefasst. Demnach sind nicht alle Formen mathematischer Beschäftigung als Problemlöseprozess aufzufassen, wobei insbesondere eindeutig definierte Textaufgaben (mit klarem Ausgangspunkt, Lösungsprozess und Ziel) aus mathematikdidaktischer Sicht teilweise nicht als Problemlöseaufgaben verstanden werden (Leuders, 2010). Mathematisches Problemlösen im engeren Sinne umfasst unbekannte Ausgangssituationen bzw. Ziele, deren Lösung bzw. Erreichung über die gelernte Anwendung bekannter Algorithmen und Regeln hinausgeht (Reiss & Törner, 2007).

Spätestens seit Georg Pólyas Buch *How to solve it* (Pólya, 1945) steht der *Prozess* des Problemlösens auch im Fokus mathematikdidaktischen Interesses. Pólya beschrieb Problemlösen als Prozess mit vier Phasen: *understanding the problem*, *devising a plan*, *carrying out the plan*, und *looking back.* Ähnliche Ideen finden sich auch schon früher, etwa bei Young (1924), der mathematisches Arbeiten von SuS als mehrschrittigen Prozess beschrieben hat: 1. Problem verstehen (Was ist bekannt? Was ist gesucht?), 2. Plan entwickeln, 3. Plan ausführen und gegebenenfalls neuen Plan ausdenken, 4. Ergebnisüberprüfung (siehe Higgins, 1997). Auch spätere Prozess-Modelle des Problemlösens (Tabelle 1) und Darstellungen des mathematischen Modellierungskreislaufes (Blum & Niss, 1989; Borromeo Ferri, Greefrath & Kaiser, 2013; Greefrath, 2010; Schupp, 1988) sind diesen frühen Modellen sehr ähnlich. Der Modellierungskreislauf stellt eine wesentliche theoretische Basis der PISA-Mathematiktests dar (OECD, 2003), die hauptsächlich mathematische Probleme beinhalten (vgl. Kapitel 2.2.3).

Tabelle 1. Vergleich der Schritte verschiedener Modelle mathematischen Problemlösens

Young (1924) [a]	*Dewey (1933)* [b]	*Pólya (1945)* [b]	*Krulik und Rudnick (1987)* [b]	*Schoenfeld (1983a)*	*Carlson und Bloom (2005)*
1. Problem begreifen	1. Problem begegnen	1. Problem verstehen	1. Lesen	1. Lesen	
	2. Diagnose, Definition des Problems		2. Explorieren	2. Exploration / Analyse	1. Orientierung
2. Planung	3. Bestandsaufnahme mehrerer Lösungen	2. Plan erstellen	3. Strategie auswählen	3. Planung	2. Planung
3. Ausführung	4. Vermute Konsequenzen der Lösung	3. Plan ausführen	4. Problem lösen	4. Durchführung	3. Ausführung
4. Ergebnisse testen	5. Konsequenzen testen	4. Rückschau	5. Rückschau und Erweiterung	5. Verifikation	4. Kontrolle

Anmerkung.
[a] Nach Higgins (1997). [b] Nach Carson (2007).

2.1.5 Problemlösen in den deutschen Bildungsstandards und Kernlehrplänen für das Fach Mathematik

Problemlösen ist als Maßnahme zur Begabtenförderung und -diagnostik, etwa in Form mathematischer Wettbewerbe, spätestens seit Ende des 19. Jahrhunderts Teil institutionalisierter mathematischer Bildung (siehe z. B. Freudenthal, 1969, für eine Historie). Heute gilt Problemlösen in Deutschland längst als wichtiger Unterrichtsbestandteil für alle SuS im Fach Mathematik (Bruder, 2000; Lindquist, 1989) von der Grundschule (Cobb & Merkel, 1989; KMK, 2005d; Rathmell & Huinker, 1989) bis zur gymnasialen Oberstufe (KMK, 2012), wozu auch die internationalen Vergleichsstudien wie PISA beigetragen haben. Im Rahmen des schulischen Mathematikunterrichts ist Problemlösen Ziel (vgl. *teaching for problem solving*), Unterrichtsgegenstand (vgl. *teaching via problem solving*) und Mittel (vgl. *teaching about problem solving*) zugleich (Schroeder & Lester, Jr., 1989).

In den deutschen Bildungsstandards für das Fach Mathematik (KMK, 2003, 2004a) und den daraus resultierenden Kernlehrplänen für das Fach Mathematik hat Problemlösen einen hohen Stellenwert. Die Bildungsstandards, die mathematische Aufgaben theoretisch entlang dreier Dimensionen (Leitideen, Kompetenzbereiche, Anforderungsbereiche) unterscheiden, betrachten Problemlösen als einen von sechs Kompetenzbereichen. Die Kompetenzanforderung im Bereich mathematischen Problemlösens war dabei für alle weiterführenden Schulformen zeitweise gleich formuliert z. B. in Nordrhein-Westfalen (Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2007, Ministerium für Schule, Jugend und Kinder des Landes Nordrhein-Westfalen, 2004a, 2004b, 2004c). 2011 wurde der Kernlehrplan Mathematik der Hauptschule für Nordrhein-Westfalen überarbeitet (Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2011). Vergleicht man den alten mit dem neuen Kernlehrplan der Hauptschule, so fällt auf, dass einige Punkte nicht mehr explizit genannt werden:

1. Planung und Beschreibung der Vorgehensweise zur Lösung eines Problems,
2. Nutzung von Repräsentationswechseln,
3. Plausibilitätsüberlegungen und Skizzen.

Das erscheint unter mathematikdidaktischen Gesichtspunkten (Kuntze & Prediger, 2005; Pugalee, 2004) problematisch, speziell für leistungsschwächere SuS (Pugalee, 2001), da Schreiben – u. a. durch Kompensation der begrenzten Arbeitsgedächtniskapazität – den Prozess des Problemlösens unterstützen kann (Kießwetter, 1983).[4] Und auch neuere BMBF-Vorhaben (Schneider et al., 2012) fordern explizit eine Sprachförderung im (mathematischen) Fachunterricht.

Problemlösen spielt auch in anderen Schulfächern eine wichtige Rolle, z. B. bei Diskussionen moralischer Dilemmata im Fach Philosophie (Lind, 1996, 2002, 2009) oder beim Experimentieren als (komplexes) Problemlösen in den Naturwissenschaften (Hammann, Hoi Phan & Bayrhuber, 2007; Rieß, Wirtz, Barzel & Schulz, 2012; Scherer, 2012; Scherer & Tiemann, 2012). Dies ist jedoch nicht Gegenstand der vorliegenden Arbeit.

2.2 Problemlösen und Mathematik als Testdomänen bei PISA 2003

Dieses Kapitel skizziert zunächst, was unter fächerübergreifenden Kompetenzen (*cross-curricular competencies*) verstanden wird, bevor die Konzeption der Tests und die für die Untersuchung der kognitiven Potenzialausschöpfungshypothese relevanten Ergebnisse zum fächerübergreifenden Problemlösen und zur Mathematik bei PISA 2003 beschrieben werden.

2.2.1 Fächerübergreifende Kompetenzen

Fächerübergreifende Kompetenzen spielen als Idee eine wichtige Rolle in Diskussionen von Bildungswissenschaft und -politik, mit der jedoch verschiedene theoretische, praktische und normative Assoziationen verbunden sind (Klieme, Artelt et al., 2001). Der Kompetenzbegriff wird zudem uneinheitlich (Tenorth, 2011) und bisweilen inflationär verwendet. Es sind mindestens sechs Bedeutungen des Kompetenzbegriffs gebräuchlich (Weinert, 1999).

Das Deutsche PISA-Konsortium (2003) definiert eine *fächerübergreifende* Kompetenz (*cross-curricular competency*) als kognitive Fähigkeit mit folgenden Merkmalen: *Situations- und Inhaltsunabhängigkeit*, *Voraussetzung* für unterschiedliche Schulfächer bzw. Domänen, *Förderbarkeit* in unterschiedlichen Schulfächern bzw. Domänen, *Relevanz* für die „Bewältigung komplexer, ganzheitlicher Anforderungen“ (S. 2) in realistischen Kontexten (vgl. OECD, 2003) und die *Transferierbarkeit* auf Aufgaben außerhalb des jeweiligen expliziten Curriculums (vgl. die Idee der Schlüsselqualifikation).

[4] Praktisch dürfte der Wegfall der Formulierungen m. E. wenig an der Unterrichtspraxis ändern.

Im Kontext des DFG-Schwerpunktprogramms 1293 „Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen“ (Klieme & Leutner, 2006) werden Kompetenzen definiert als „Systeme aus spezifischen, prinzipiell erlernbaren Fertigkeiten, Kenntnissen und metakognitivem Wissen, die es erlauben, eine Klasse von Anforderungen in bestimmten Alltags-, Schul- oder Arbeitsumgebungen zu bewältigen“ (Klieme, Funke et al., 2001, S. 182). Mit Bezug auf Weinert (1999) werden dabei vier charakterisierende Merkmale der Definition des Kompetenzbegriffs herausgestellt:

1. Es handelt sich um eine *funktionale Definition*, da als Indikator einer Kompetenz (z. B. im Fach Mathematik) betrachtet wird, wie erfolgreich eine Person mit bestimmten beschriebenen Anforderungen umgeht.
2. Kompetenzen sind *kognitive Konstrukte*, d. h. der Kompetenzbegriff wird auf kognitive Fähigkeiten und Fertigkeiten beschränkt. Affektive Konstrukte wie Anstrengungsbereitschaft oder Leistungsmotivation (Rheinberg, Vollmeyer & Burns, 2001) sollen separat definiert und gemessen werden.
3. Kompetenzen beinhalten die Annahme einer *Bereichsspezifität* (z. B. Gültigkeit für die Domäne Mathematik).
4. Kompetenzen sind als *Dispositionen* konzipiert, d. h. im engen Rahmen verallgemeinerbar. Sie meinen mehr als die konkrete Leistungsmessung (Performanz).

Zusammenfassend werden Kompetenzen im DFG-Schwerpunktprogramm 1293 (Klieme & Leutner, 2006; Koeppen, Hartig, Klieme & Leutner, 2008; Wilhelm & Nickolaus, 2013) und in dieser Arbeit als *funktional definierte, bereichsspezifische kognitive Dispositionen* betrachtet.

2.2.2 Fächerübergreifende Problemlösekompetenz

Fächerübergreifendes Problemlösen (vgl. Leutner et al., 2004) wird bei PISA 2003 definiert als

> „an individual's capacity to use cognitive processes to confront and resolve real, cross-disciplinary situations where the solution path is not immediately obvious and where the literacy domains or curricular areas that might be applicable are not with a single domain of mathematics, science or reading” (OECD, 2003, S. 156).

Das Attribut fächerübergreifend (*cross-disciplinary*, *cross-curricular*) ist in diesem Zusammenhang erklärungsbedürftig. Anders als beim im pädagogischen Kontext verbreiteten Begriff des *fächerübergreifenden Unterrichts* (Peterßen, 2000) besteht das fächerübergreifende Element bei Problemlöseaufgaben aus dem PISA-2003-Test zumeist nicht darin, Probleme zu lösen, die mindestens zwei Domänen (bzw. Schulfächer) betreffen, obwohl dies intendiert ist (OECD, 2004b), sondern Probleme, die meist kein spezifisches Fachwissen aus den anderen Testdomänen, sondern nur ein gewisses Maß an Weltwissen (Funke, 2003) oder geringe mathematische Kenntnisse erfordern (Fleischer et al., 2010). Im Rahmen von PISA 2003 (OECD, 2003, 2004a) wird fächerübergreifendes Problemlösen negativ definiert durch Abwesenheit der Domänen Deutsch, Mathematik und Naturwissenschaften.

Problemlösen kann, wie jedes theoretische Konstrukt, empirisch auf verschiedene Weisen untersuchen werden (Arbinger, 1997). Die naheliegende Möglichkeit, das Konstrukt "fächerübergreifende Problemlösekompetenz" zu messen, besteht darin eine Verhaltensstichprobe zu erheben, d. h. Personen fächerübergreifende Probleme vorzugeben, die sie lösen sollen. Dieser Testansatz wird auch bei PISA realisiert (OECD, 2004b). Die bei PISA 2003 eingesetzten Probleme werden in mehrfacher Weise klassifiziert. Einerseits werden Probleme je nach theoretischen Anforderungen in drei Problemtypen eingeteilt: Entscheidungen treffen (*decision making*), Systeme analysieren und entwerfen (*system analysis and design*) sowie Fehler suchen (*trouble shooting*). Ausführliche Beispielaufgaben und Lösungen findet man bei OECD (2003; 2004b). Anderseits unterscheiden sich die Aufgaben-Kontexte, in die das Problem eingebettet ist: Privatleben (*personal life*), Arbeit und Freizeit (*work and leisure*) sowie Gesellschaft (*community and society*). Die Ergebnisse zum Problemlösen werden bei PISA jedoch unidimensional berichtet (OECD, 2004b) ohne Berücksichtigung der theoretischen Klassifikation. Reanalysen der PISA-2003-Daten (Leutner, Fleischer, Wirth, Greiff & Funke, 2012) zeigten, dass sich auf psychometrischer Ebene eine unidimensionale Modellierung über alle Problemtypen und eine dreidimensionale Modellierung mit Berücksichtigung der Problemtypen kaum voneinander unterscheiden.

Zur Veränderung des diagnostischen Schwerpunkts vom analytischen Problemlösen bei PISA 2003 über das komplexe/interaktive Problemlösen bei PISA 2012 bis zum kooperativen Problemlösen bei PISA 2015 siehe Greiff, Holt und Funke (2013).

2.2.3 Mathematische Kompetenz

Der Mathematiktest bei PISA 2003 basiert auf der Idee mathematischer Grundbildung (*mathematical literacy*; OECD, 2003), d. h. der Annahme einer kognitiven Fähigkeit, die es ermöglicht, „die Rolle zu erkennen und zu verstehen, die Mathematik in der Welt spielt, fundierte mathematische Urteile abzugeben und Mathematik in einer Weise zu verwenden, die den Anforderungen dieser Person als konstruktivem, engagiertem und reflektiertem Bürger entspricht" (OECD, 2003, zitiert nach Neubrand et al., 2005, S. 51).

Mathematische Kompetenz wird bei PISA hauptsächlich in Form von diskontinuierlichen Textaufgaben zu mehr oder weniger authentischen bzw. realen Problemstellungen (*real world situations*) erfasst, zu deren Lösung mehr als die bloße Anwendung gelernter Regeln nötig ist (OECD, 2004a). Die verwendeten Aufgaben (OECD, 2004a) lassen sich größtenteils als in der Tradition Freudenthals stehende Modellierungsprobleme beschreiben (Baumert, Brunner, Lüdtke & Trautwein, 2007), die einen oder mehrere Aspekte des Regelkreislaufs von „Strukturieren, Mathematisieren, Verarbeiten, Interpretieren und Validieren" (Neubrand et al., 2005, S. 52) erfassen (vgl. auch Kapitel 2.4.3.5). Modellieren und Problemlösen stehen in einer engen Beziehung. Sie lassen sich als „zwei Seiten der gleichen Medaille" (Greefrath, 2010, S. 44) auffassen. Modellieren ist der Versuch, anwendungsorientierte Probleme durch mathematische Modelle zu lösen, wobei mehrfache Übersetzungsprozesse zwischen Realsituation und mathematischem Modell stattfinden. Mathematisches Problemlösen im engeren Sinne wird beim Lösen der innermathematischen Problemstellungen betrieben und ist damit Teil des Modellierungsprozesses (Blum, Galbraith, Henn & Niss, 2007; OECD, 2004b).

Für detaillierte Informationen und Beispielaufgaben aus den vier mathematischen Inhaltsbereichen (Quantität, Raum und Form, Veränderung und Beziehungen, Unsicherheit), den drei Kompetenzclustern (Reproduktion, Verbindungen, Reflexion) und den Kompetenzstufen bei PISA 2003 siehe z. B. OECD (2003, 2004a) oder Neubrand et al. (2005).

2.3 Ausgewählte Befunde der PISA-Studie 2003 zum Problemlösen und zur Mathematik

Dieses Kapitel beschreibt ausgewählte Befunde zum fächerübergreifenden Problemlösen und zur Mathematik, die zentral für die nachfolgenden empirischen Studien zur kognitiven Potenzialausschöpfungshypothese sind.

Vergleicht man die internationalen PISA-Ergebnisse 2003 zum fächerübergreifenden Problemlösen (PL) mit denen zur Mathematik (M), so erkennt man kein einheitliches Muster (Leutner et al., 2004; OECD, 2004b):

Fall 1 (M > PL): In 13 Ländern schnitten die SuS statistisch signifikant besser in Mathematik ab als beim Problemlösen. In den Niederlanden ist diese Differenz zugunsten der Mathematik am größten.

Fall 2 (M = PL): In 15 Ländern gab es keinen statistisch signifikanten Unterschied zwischen der Leistung beim Problemlösen und der Leistung in Mathematik. Angesichts der Stichprobengröße bei PISA lassen sich diese Befunde nicht durch eine zu geringe Teststärke erklären.

Fall 3 (M < PL): In 12 Ländern schnitten die SuS statistisch signifikant besser beim Problemlösen ab als in Mathematik. In Japan, Ungarn und Deutschland ist diese Differenz zugunsten des Problemlösens besonders groß.

Man beachte, dass alle internationalen Tests bei PISA 2003 auf einen Mittelwert von 500 ($SD = 100$) skaliert worden sind. Die Schwierigkeiten des Mathematik- und Problemlösetests sind im internationalen Vergleich also psychometrisch per Konstruktion gleich.

Deutschlands Ergebnisse bei PISA 2003 (Leutner et al., 2004; OECD, 2004b) lassen sich in Fall 3 einordnen: SuS in Deutschland erzielten durchschnittliche Ergebnisse in den Domänen Mathematik ($M = 503$, $SD = 103$), Naturwissenschaften ($M = 502$, $SD = 111$) und Lesen ($M = 491$, $SD = 109$). Beim Problemlösen lagen die Ergebnisse ($M = 513$, $SD = 95$) jedoch über dem Durchschnitt (verglichen mit dem OECD-Mittelwert von 500; $SD = 100$).

Vergleicht man die deutschen Bundesländer hinsichtlich der mittleren Kompetenz der SuS in den Domänen Mathematik und Problemlösen (Leutner et al., 2005), so zeigt sich, dass vor allem in Bundesländern, in denen im Vergleich niedrigere mathematische Kompetenzen erreicht werden (z. B. in Nordrhein-Westfalen), die

Problemlösekompetenz deutlich größer ist als die mathematische Kompetenz. Über alle Bundesländer betrachtet ist das Muster wie im internationalen Vergleich uneinheitlich.

2.4 Erklärungen der Diskrepanz der Problemlöse- und Mathematikkompetenz bei PISA 2003: Die kognitive Potenzialausschöpfungshypothese und alternative Erklärungen

Es gibt verschiedene Erklärungen für die Diskrepanz-Befunde in den Domänen fächerübergreifendes Problemlösen und Mathematik bei PISA 2003. Das Kapitel beginnt mit der kognitiven Potenzialausschöpfungshypothese (Kapitel 2.4.1), die kognitives Potenzial postuliert, das zum Kompetenzerwerb im Fach Mathematik unzureichend genutzt wird, und der zugehörigen Befundlage (Kapitel 2.4.2). Anschließend werden alternative Erklärungen diskutiert (Kapitel 2.4.3). Man beachte, dass sich die Hypothesen zur Erklärung der bei PISA 2003 gefundenen Differenz zwischen Mathematik- und Problemlösekompetenz von SuS in Deutschland nicht in allen Fällen gegenseitig ausschließen.

2.4.1 Kognitive Potenzialausschöpfungshypothese

Zur Erklärung der Diskrepanz zwischen den – bei PISA 2003 – relativ hohen Lösungsquoten von SuS im Bereich des fächerübergreifenden Problemlösens und dem relativ niedrigen Abschneiden im Bereich der Mathematik wurde eine *kognitive Potenzialausschöpfungshypothese* postuliert (Leutner et al., 2004; OECD, 2004b): Sind die Leistungen – wie in Deutschland – im fachlichen Problemlösen wesentlich geringer als die Leistungen beim fächerübergreifenden Problemlösen (Fall 3, vgl. Kapitel 2.3), wird dies im PISA-Kontext als *ungenutztes Potenzial* interpretiert (Leutner et al., 2004; Walter et al., 2006; Zimmer, Burba & Rost, 2004), d. h. das vorhandene kognitive Potenzial, das beim fächerübergreifenden Problemlösen sichtbar wird, zeigt sich nicht beim fachlichen Problemlösen.

Bezogen auf die Domäne Mathematik zeigt der Test fächerübergreifender Problemlösekompetenz „generic skills that may not be fully exploited by the mathematics curriculum“ (OECD, 2004b, S. 56). Gemäß der kognitiven Potenzialausschöpfungshypothese besitzen SuS im Jahr 2003 in Deutschland also ein kognitives Potenzial, das sich beim fächerübergreifenden Problemlösen zeigt, das jedoch nicht optimal zum Aufbau fachlicher Problemlösekompetenz in der Mathematik genutzt wird (Leutner et al., 2004; Walter et al., 2006). Diese Differenz zwischen der Problemlösekompetenz und den fachspezifischen Kompetenzen, v. a. in der Mathematik, wird besonders in Deutschland betont (Leutner et al., 2004):

> „Im Vergleich zu anderen OECD-Staaten wird in Deutschlands Schulen offensichtlich eine durchschnittlich geringere mathematische Kompetenz erreicht als aufgrund der durchschnittlichen Problemlösekompetenz der Schülerinnen und Schüler zu erwarten wäre“ (Leutner et al., 2004, S. 170).

In Deutschland waren die Differenzen zugunsten des Problemlösens bei PISA 2003 besonders groß. Unter allen 29 OECD-Teilnehmerländern zeigten elf Länder Differenzen zugunsten des Problemlösens, und nur in Ungarn und Japan waren die Differenzen zugunsten des Problemlösens größer als in Deutschland (OECD, 2004b). Diese große Differenz ist besonders überraschend, da zwischen den Domänen Problemlösen und Mathematik eine hohe latente Korrelation gefunden wurde (r = .89; OECD, 2005). In Deutschland wird dieses ungenutzte kognitive Potenzial vor allem für mathematisch leistungsschwächere SuS bzw. SuS im untersten Kompetenzbereich betont (Leutner et al., 2004, 2005; vgl. Abbildung 1).

Die Ergebnisse des Vergleichs der PISA-Ergebnisse zwischen den deutschen Bundesländern interpretieren auch Prenzel et al. (2005) so, dass in den meisten Bundesländern – z. B. in Nordrhein-Westfalen – das im Test zum Problemlösen „erkennbare kognitive Potential noch nicht angemessen in mathematische Kompetenz umgesetzt wird“ (S. 17). Querschnittliche Vergleichsstudien können jedoch kaum Erklärungen schwacher Leistungen und Interventionsideen zur Verbesserung der getesteten Kompetenzen liefern (Prenzel & Allolio-Näcke, 2006).

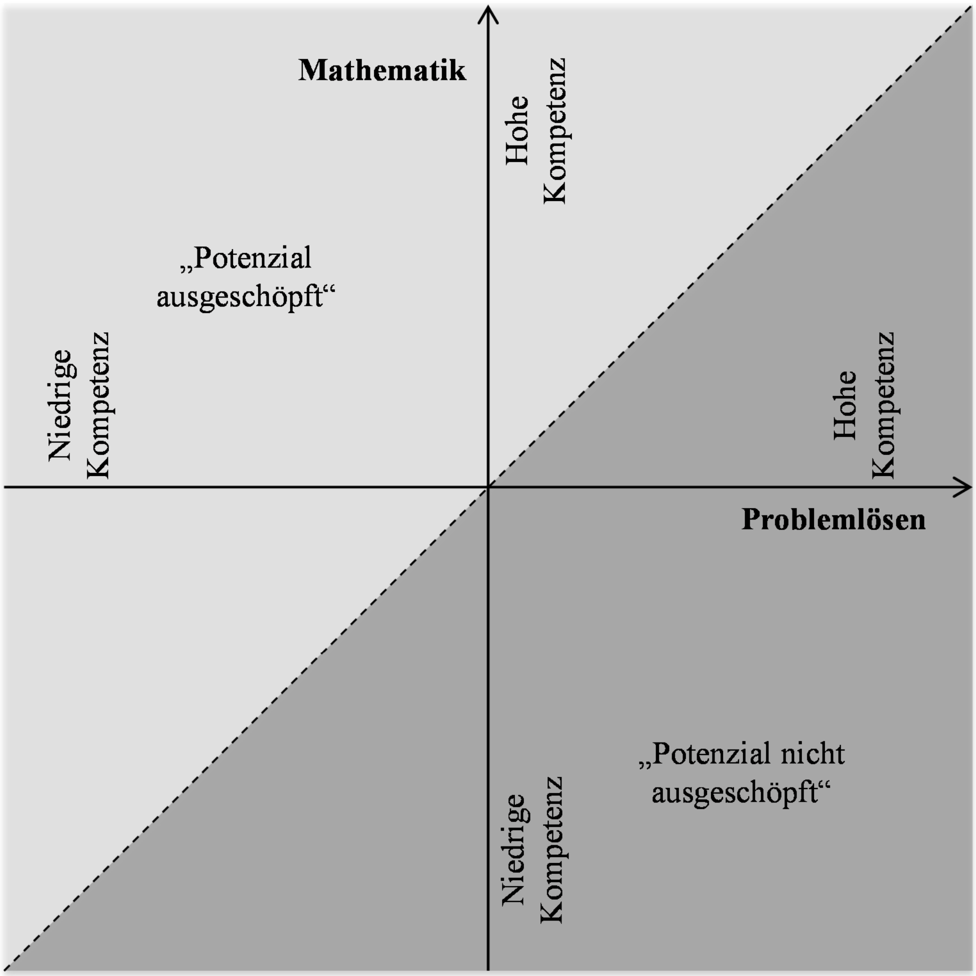


Abbildung 1. Schema zur kognitiven Potenzialausschöpfungshypothese

Anmerkung. Da der erwartete Zusammenhang zwischen Problemlösen und Mathematik nicht perfekt ist und zudem keine messfehlerfreien Tests existieren, ist mit einer gewissen Unschärfe zu rechnen.

2.4.2 Theoretische und empirische Befundlage zur kognitiven Potenzialausschöpfungshypothese

Die kognitive Potenzialausschöpfungshypothese (Leutner et al., 2004; OECD, 2004b) entstand als weitgehend theoriefreie Post-hoc-Hypothese.[5] Daher werden in diesem Unterkapitel theoretische und empirische Argumente für die kognitive Potenzialausschöpfungshypothese gesammelt. Anschließend werden alternative Erklärungen beschrieben (Kapitel 2.4.3).

[5] Mit der mathematischen Potential-Theorie (siehe z. B. Helms, 2009), die sich mit harmonischen Funktionen beschäftigt und z. B. zur Beschreibung von Gravitationsfeldern eingesetzt wird, und den physikalischen Begriffen Potential und Potentialdifferenz hat die kognitive Potenzialausschöpfungshypothese nichts zu tun.

2.4.2.1 Kognitive Anforderungen und Prozesse

Vergleicht man die theoretischen kognitiven Anforderungen und Prozesse fächerübergreifenden und mathematischen Problemlösen, so stellt man große Ähnlichkeiten fest (OECD, 2003). Die Aufgaben beider Domänen erfordern nur eine geringe Lesekompetenz und geringes naturwissenschaftliches Wissen und stellen hohe Anforderungen an schlussfolgerndes Denken (OECD, 2003). Während die mathematischen Aufgaben erwartungsgemäß hohe Anforderungen an mathematisches Wissen stellen, ist zur Lösung der Problemlöseaufgaben höchstens geringes mathematisches Wissen nötig (Tabelle 2). Die theoretischen kognitiven Prozessmodelle der beiden Domänen sind vergleichbar konzipiert (OECD, 2003, vgl. Greefrath, 2010): Am Anfang erfolgreichen fächerübergreifenden und mathematischen Problemlösens stehen das Verständnis des Problems (Situationsmodell) und das Erkennen relevanter Bedingungen (*constraints*, Realmodell), die zu einer Repräsentation des Problems führen. Anschließend wird das Problem gelöst, interpretiert und kommuniziert (Tabelle 3). Diese theoretischen Überschneidungen zwischem (fächerübergreifendem) Problemlösen und Mathematik bei PISA 2003 werden empirisch durch eine Anforderungsanalyse gestärkt, die Ähnlichkeiten und Unterschiede der Items der beiden Domänen untersuchte (siehe Fleischer et al., 2010).

Tabelle 2. Vergleich der kognitiven Anforderungen der Domänen Problemlösen und Mathematik (vgl. OECD, 2003)

Kognitive Anforderungen	*Problemlösen*	*Mathematik*
- Lesekompetenz	gering	gering
- Naturwissenschaftliches Wissen	gering	gering
- Schlussfolgerndes Denken	hoch	hoch
- Mathematisches Wissen	gering	hoch

Tabelle 3. Vergleich der kognitiven Prozesse der Domänen Problemlösen und Mathematik (vgl. OECD, 2003)

Problemlösen	*Mathematik (Modellieren)*
1. Verstehen des Problems	1. Verstehen des Problems (Situationsmodell)
2. Erkennen relevanter Bedingungen	2. Erkennen relevanter Bedingungen (Realmodell)
3. Mentale Repräsentation	3. Mathematisieren
4. Lösen des Problems	4. Lösen des Problems
5. Interpretieren	5. Prüfen und Bewerten der Problemlösung
6. Validieren	6. Kommunizieren der Problemlösung

2.4.2.2 *PISA-I-Plus*

Die Daten der Messwiederholungsstudie PISA-I-Plus zur Kompetenzentwicklung in Mathematik und Naturwissenschaften in der Jahrgangsstufe 10 geben einen ersten empirischen Beleg für die kognitive Potenzialausschöpfungshypothese (Leutner et al., 2006): Pfadanalytische Auswertungen zur Prognose der Mathematikleistung in der Jahrgangsstufe 10 zeigten sowohl einen eigenständigen (direkten) Beitrag der Problemlöseleistung in der Jahrgangsstufe 9 als auch einen indirekten Beitrag der Problemlöseleistung vermittelt über die Mathematikleistung in der Jahrgansstufe 9 (Abbildung 2). Postuliert wird – zumindest implizit – eine (direkte bzw. indirekte) Wirkung von analytischem Problemlösen auf Mathematik. Zusätzlich zeigte eine Kommunalitätenanalyse derselben Daten (Leutner et al., 2006), dass der gemeinsame Varianzanteil analytischen Problemlösens und Mathematik (R^2 = .127) größer war als der gemeinsame Varianzanteil von Intelligenz und Mathematik (R^2 = .042), während der einzelne Varianzanteil für Intelligenz (R^2 = .006) und Problemlösen (R^2 = .005) eher unbedeutend war. Diese Ergebnisse werden so interpretiert, dass Problemlösen und Mathematik aus verschiedenen, teilweise überlappenden Komponenten bestehen, die unterschiedlich zum Erwerb zukünftiger mathematischer Kompetenz beitragen. Sie sprechen für die Bedeutung fächerübergreifender Problemlösekompetenz beim Aufbau fachlicher Kompetenzen in der Mathematik und geben damit Hinweise für die Gültigkeit der kognitiven Potenzialausschöpfungshypothese.

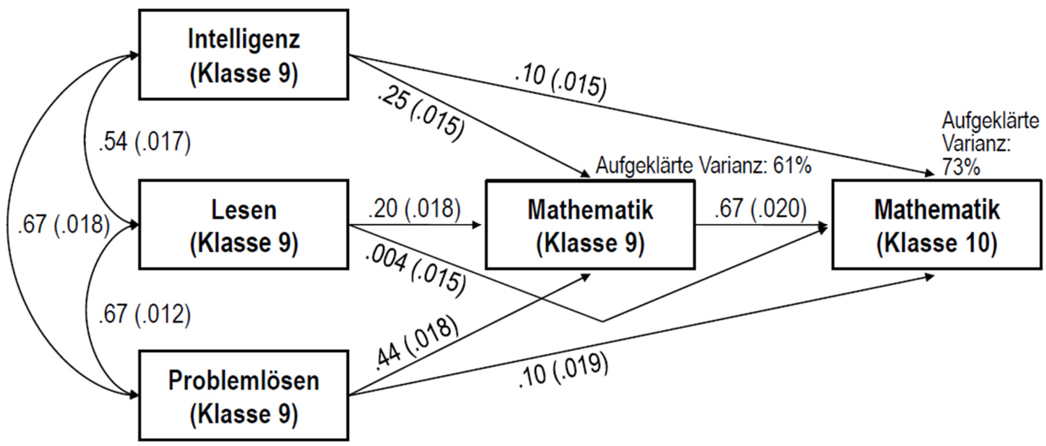


Abbildung 2. Pfadanalyse zur Prognose der Mathematikleistung in der Jahrgangsstufe 10 (Fleischer, Wirth & Leutner, 2009)

Anmerkungen. Angegeben sind standardisierte Pfadkoeffizienten (SE in Klammern). Abdruck mit freundlicher Genehmigung der Autoren.

2.4.2.3 Komponenten fächerübergreifender Problemlösekompetenz

Wichtige Komponenten der Problemlösekompetenz sind unter anderem inhaltsspezifisches Wissen (z. B. Chi, Glaser & Rees, 1982), Sachwissen und Handlungswissen (Süß, 1996), konditionales Wissen („knowing when and why“), das die Umstände der Anwendung von Operationen beschreibt (Paris, Lipson & Wixson, 1983), die Verfügbarkeit und Anwendung allgemeiner Problemlösestrategien und Heuristiken (vgl. Gick, 1986) sowie die Fähigkeit zur Selbstregulation, um Prozesse des Problemlösens zu planen, zu überwachen, zu bewerten und gegebenenfalls zu modifizieren (vgl. Davidson, Deuser & Sternberg, 1994). Mit dieser Liste wichtiger Komponenten der Problemlösekompetenz wird kein Anspruch auf Vollständigkeit erhoben. Zur Lösung von Problemen, vor allem bei subjektiv schwierigen Problemen, spielen weitere Faktoren wie Motivation, Fähigkeitsselbstkonzept (Marsh, 1990b), Selbstwirksamkeitserwartungen (Zimmerman, 1995, 2008) und weitere Voraussetzungen wie Interesse, Ausdauer, Organisation und kommunikative Fähigkeiten (Stacey, 2005) eine wichtige Rolle (siehe Mayer, 1998, für eine ausführlichere Diskussion kognitiver, metakognitiver und motivationaler Fakotren beim Problemlösen). Ergänzend sei bemerkt, dass keine einheitliche Klassifikation von Wissens- und Metakognitionsarten existiert (Hasselhorn, 2010).

Inhaltsspezifisches Wissen: Sach- und Handlungswissen

Die Berücksichtigung inhaltsspezifischen Wissens in der Problemlöseforschung ist wichtig, da aus einer vermeintlichen Problemstellung für eine Person eine Aufgabe wird, wenn sie über Wissen und (algorithmische) Prozeduren zur routinemäßigen Lösung verfügt (vgl. Kapitel 2.1.1). Zudem betont die psychologische Expertise-Forschung (z. B. Chi et al., 1982) die Rolle der domänenspezifischen Wissensbasis (*knowledge base*), z. B. beim Schach (de Groot, 1978). Vorwissen (Sachwissen, Fachwissen, Domänen-spezifisches Wissen) wurde in der Geschichte der Erforschung höherer kognitiver Fähigkeiten (wie Intelligenz, Arbeitsgedächtnis) jedoch lange vernachlässigt (Hambrick, 2005). Die meisten Probleme, denen wir begegnen, benötigen jedoch zu ihrer Lösung die kognitive Verfügbarkeit und Anwendung *inhaltsspezifischen Wissens* (Chi et al., 1982; Greeno, Collins & Resnick, 1996). Süß (1996) gliedert dieses inhaltsspezifische Wissen in *Sach- und Handlungswissen*, das beides im Gedächtnis deklarativ (als Propositionen) und prozedural (als Produktionssysteme) repräsentiert sein kann. Sachwissen, das zumeist deklarativ ist, beinhaltet im Kontext des Problemlösens das Wissen über Objekte, Situationen und Zustände, das hilft, die Problemsituation und den Zielzustand mental zu repräsentieren und zu erfassen (Fleischer et al., 2010, S. 241). Handlungswissen (auch prozedurales Wissen, *knowing how*) beinhaltet das Wissen über Mittel bzw. Operatoren, die durch Anwendung (Transformation) zur Veränderung der Problemsituation führen sowie die Fähigkeit, eine kognitive Operation in die Tat umzusetzen (vgl. Fleischer et al., 2010; Süß, 1996). Nach dieser Definition ist Handlungswissen damit kein eindimensionales Konstrukt. Die Repräsentation prozeduralen Wissens wird in Form von Produktionen bzw. Produktionssystemen, d. h. als Regeln bzw. Wenn-Dann-Beziehungen, angenommen (Arbinger, 1997) oder konkreter in Form von Heuristiken und Strategien (Schraw, 1998). Mit Blick auf die PISA-2003-Aufgaben wird prozedurales Wissen u. a. dazu benötigt, Informationen aus einer Tabelle entnehmen zu können oder ein Entscheidungsdiagramm zu benutzen (vgl. Anhang W).

Metakognition

Metakognition bezieht sich auf das Wissen über eigene kognitive Prozesse und Produkte (Flavell, 1976) und spielt insbesondere beim Problemlösen (Berardi-Coletta, Buyer, Dominowski & Rellinger, 1995; Davidson et al., 1994; Davidson & Sternberg, 1998; Flavell, 1976) und beim selbstregulierten Lernen eine wichtige Rolle (Boekaerts, 1997; Opfermann, Azevedo & Leutner, 2012; Schütte, Wirth & Leutner, 2012). Metakognition gilt als mehrdimensional, domänenunspezifisch und förder- bzw. lernbar (Schraw, 1998). Metakognition bezeichnet kein einheitliches Konzept, sondern einen „Sammelbegriff für eine Reihe von Phänomenen, Aktivitäten und Erfahrungen, die mit dem *Wissen* und der *Kontrolle* über eigene kognitive Funktionen […] zu tun haben“ (Hasselhorn, 2010, S. 541). Man beachte, dass die systematische Unterscheidung der Komponenten der Metakognition (vgl. Abbildung 3) konzeptuell nützlich ist, die Komponenten aber in Beziehung stehen (Schraw, 1998). Metakognitives Wissen umfasst deklaratives Wissen, prozedurales Wissen und konditionales Wissen (Schraw & Moshman, 1995). Selbstregulation, auch metakognitive Fähigkeiten, exekutive Kontrollprozesse, beinhaltet die metakognitiven Fähigkeiten Planung, Evaluation und Monitoring (Schraw, 1998).[6] Schon Dörner, Kreuzig, Reither und Stäudel (1983) betonen die Bedeutung der Selbstregulation für erfolgreiches (komplexes) Problemlösen. Im Kontext der vorliegenden Arbeit spielen insbesondere die metakognitiven Aspekte konditionales Wissen und Planung (bzw. Planungsfähigkeit) eine Rolle.

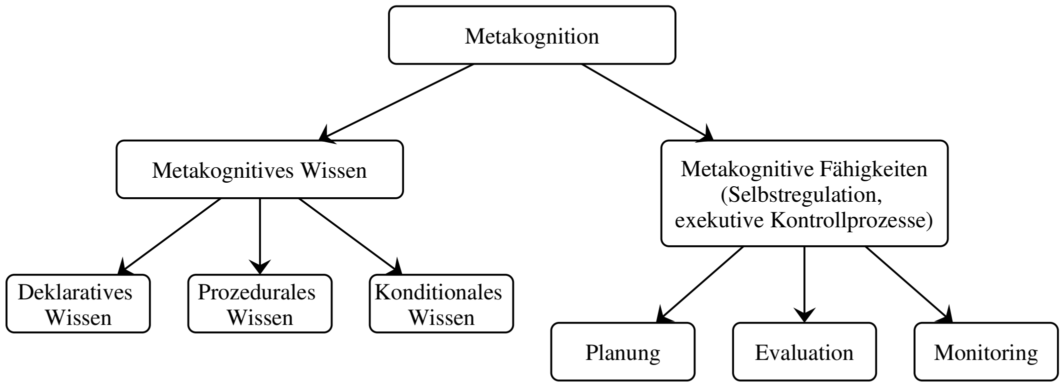


Abbildung 3. Komponenten der Metakognition (Eigene Visualisierung des Modells der Metakognition von Schraw, 1998)

Konditionales Wissen (knowing when and why)

Probleme haben typischerweise nicht genau einen Lösungsweg. Es gibt oftmals zahlreiche Wege zum Ziel. Doch welche der Optionen und Vorgehensweisen sind möglich bzw. unmöglich, zielführend, hinderlich oder irrelevant? Konditionales Wissen, das „knowing when" und „knowing why" (Paris et al., 1983; Schraw, 1998), umfasst dieses Wissen über die „Umstände der Anwendung von Operationen" (Fleischer et al., 2010; Paris et al., 1983). Es ist wichtig in Situationen, in denen – wie beim Problemlösen – mehr als eine Handlungsoption verfügbar ist.

Die Entscheidungen, die dazu führen, dass jemand eine Handlung zu einem bestimmten Zeitpunkt aus gegebenen Gründen ausführt, kann man aus didaktischen Gründen als Entscheidungsbaum visualisieren oder als Menge deterministischer bzw. probabilistischer If-then-else-Aussagen beschreiben (zum konditionalen Wissen in der Logik bzw. künstlichen Intelligenz siehe Eiter & Lukasiewicz, 2000; Lehmann & Magidor, 1992).

Die konzeptuelle Nähe konditionalen Wissens zum Konzept der Schemata ist offensichtlich. Ein kognitives Schema erlaubt, Probleme in Kategorien einzuordnen,

[6] Selbstregulation und selbst-reguliertes Lernen ist im DFG-Schwerpunktprogramm 1293 Gegenstand eines eigenen Projekts (für Ergebnisse siehe Schütte, 2012; Schütte, Wirth und Leutner, 2010).

die bestimmte Lösungsmethoden nahe legen bzw. erfordern (Chi, Feltovich & Glaser, 1981; Fuchs et al., 2006; Paas & van Merriënboer, 1994).

Konditionales Wissen ermöglicht beispielsweise einen effektiven Umgang mit kognitiven Ressourcen und Strategien durch die flexible Auswahl und Anpassung an die Erfordernisse der jeweiligen (Lern-)Aufgabe (Bouffard-Bouchard, 1994; Gick, 1986; Schraw, 1998). Obwohl schon Kinder aus verschiedenen ihnen bekannten Strategien auswählen (Sodian, 1998), sind die genauen Bedingungen des Strategieeinsatzes weitgehend unbekannt (Fritz & Hussy, 2001). Es existieren zahlreiche mögliche Gründe, die mangelnde Strategienutzung erklären können, z. B. unzureichendes Monitoring, mangelnde Routinen und Vorwissen, ungünstige Attributionen, Lernziele, ausbleibender Transfer (Garner, 1990).

Die in dieser Arbeit verwendeten Testaufgaben zum konditionalen Wissen bestehen aus einem kurzen Szenario, zu dem verschiedene Handlungsmöglichkeiten jeweils auf einer 6-stufigen Likertskala eingeschätzt werden sollen, die die deutschen Schulnoten (von 1 = *sehr gut* bis 6 = *ungenügend*) repräsentieren (siehe Anhang E und Anhang X). Dieses Format kommt in zahlreichen Strategiewissenstests in gleicher (Artelt, Beinicke, Schlagmüller & Schneider, 2009; Neuenhaus, Artelt, Lingel & Schneider, 2010; Schlagmüller & Schneider, 2007; Thillmann, 2007) oder sehr ähnlicher Form (Scherer & Tiemann, 2012; Shahat, Ohle, Treagust & Fischer, 2013) zum Einsatz. Strategien können als Sammlung von Heuristiken (Ohlsson, 2012) und Strategiewissen kann als Teilmenge konditionalen Wissens aufgefasst werden (Stacey, 2005). Strategiewissenstests erfassen somit einen spezifischen Teil konditionalen Wissens, d. h. meist spezifische Strategien in einem umgrenzten Kontext (z. B. Turm von Hanoi, Welsh & Huizinga, 2005; Experimentierstrategien, Marschner, 2011; Thillmann, 2007). Marewski und Schooler (2011) postulieren für jede Strategie eine kognitive Nische (*cognitive niche*), d. h. einen begrenzten Anwendungsbereich, in dem die jeweilige Strategie erfolgreich nutzbar ist. Die Strategiekenntnis impliziert dabei im Allgemeinen noch nicht die adäquate situationsspezifische Auswahl von Strategien und eine gute Strategie-Nutzung (Marschner, Thillmann, Wirth & Leutner, 2012; vgl. auch die Forschung zu *simple heuristics*, z. B. Todd & Gigerenzer, 2000). Begründungswissen („knowing why“) ist in diesen Strategiewissenstests zwar theoretisch notwendig, um eine Strategie beurteilen zu können, wird aber selten explizit getestet.

Planungsfähigkeit

Planungsfähigkeit als Aspekt allgemeiner Problemlösefähigkeit (Davidson et al., 1994) ist die Fähigkeit, in einer strukturierten und geplanten Art und Weise zu handeln. Ihr wird eine entscheidende Bedeutung in Schule, Beruf und Freizeit über die Lebensspanne zugesprochen (Dreher & Oerter, 1987; Lezak, 1995). Planen kann definiert werden als „any hierarchical process in the organism that can control the order in which a sequence of operations is to be performed" (Miller, Galanter & Pribram, 1960, S. 16). Geplant werden also zeitliche Abfolgen von Handlungen. Planen ist dabei ein mehrschrittiger Prozess (Fritz & Funke, 2003; Funke & Fritz, 1995b): Der Prozess des Planens beginnt mit einer Antizipation eines Ziels, gefolgt von einer Aufgabenanalyse, die hilft, den Problemraum adäquat zu repräsentieren. Danach folgt die Ausarbeitung des Handlungsplans bzw. von Teilplänen unter Berücksichtigung relevanter Beschränkungen (*constraints*). Dessen Ausführung wird überwacht und kontrolliert, indem Teilpläne koordiniert werden und Erfahrungen beim Problemlösen zur Evaluation des Problemraums, der Teilziele oder des Gesamtplans genutzt werden können (Fritz & Funke, 2003). Der Prozess des Planens kann dabei in zwei (nicht zwingend disjunkte) Phasen unterschieden werden, *Planerstellung* und *Planausführung*, die jeweils eine Reihe von Teilkompetenzen erfordern (Funke & Glodowski, 1990; Abbildung 4).

Je nach theoretischem Rahmen und diagnostischer Intention (vgl. Fritz & Funke, 2003) wird Planen a) als eigenständiges Konstrukt (Arling, 2006; Arling & Spijkers, 2012; Funke & Glodowski, 1990), b) als gleichbedeutend mit Problemlösen oder als wesentliche Komponente des Problemlösens betrachtet (Dörner, 1976; Halpern, 2003) oder c) zusammen mit Monitoring und Evaluation als wichtige Komponente metakognitiver selbstregulatorischer Fähigkeiten (Flavell, 1976; Schraw, 1998) bzw. als exekutiver Kontrollprozess und damit als Teil der Metakognition klassifiziert (Hasselhorn, 1992).

Diagnostische Möglichkeiten, Planungsfähigkeit zu erfassen, gibt es zahlreiche (Funke & Fritz, 1995a), z. B. durch Spiele (Zoo-Spiel; Fritz & Funke, 2003), Tagesplanungsaufgaben (Arling, 2006; Funke, Krüger & Schulze, 1999) oder die Vorgabe von Teilschritten eines Handlungsplanes, die in eine chronologische Reihenfolge zu bringen sind (Pascha, Schöppe & Hacker, 2001; Schütte, 2012; Anhang E).

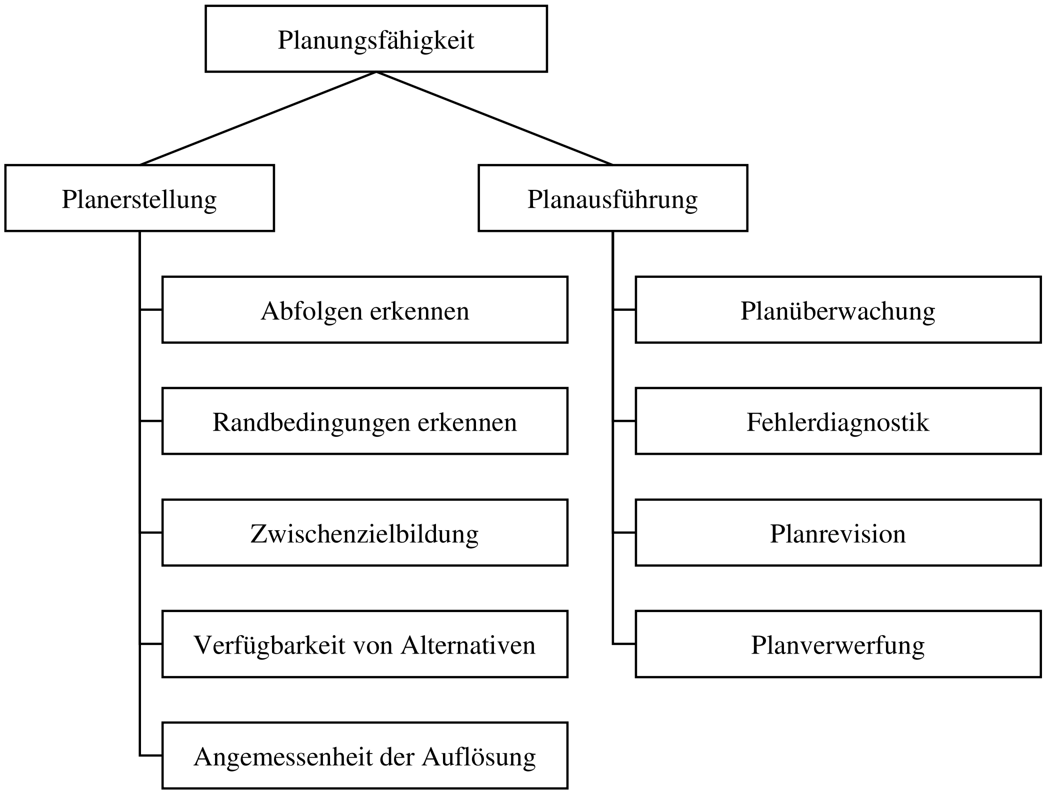


Abbildung 4. Teilkompetenzen der Planungsfähigkeit (Eigene Visualisierung des Modells von Funke und Glodowski, 1990)

2.4.2.4 Komponenten mathematischer Problemlösekompetenz

Problemlösen und Mathematik stehen in einem engen Verhältnis (vgl. Kapitel 2.2.1). Entsprechend analog lassen sich Teilkomponenten mathematischer Problemlösekompetenz identifizieren (vgl. Stacey, 2005).

Zum Lösen mathematischer Probleme benötigen SuS mathematisches *Sachwissen* etwa in Form gelernter Fachbegriffe (z. B. mathematische Vokabeln wie Ziffer, Zahl, Laplace-Versuch und mathematische Sätze wie der Satz des Pythagoras) und Rechenregeln (z. B. Kommutativgesetz, Klammerregeln, Äquivalenzumformungen) sowie *Handlungswissen*, d. h. die Kenntnis mathematischer Schemata, Rechenregeln und Prozeduren (z. B. Kurvendiskussion) und ihrer Anwendungen, die in der Schule oftmals Kochrezepten gleichen (Kießwetter, 1983).

Wie beim nicht-fachlichen Problemlösen spielen auch beim mathematischen Problemlösen *metakognitive selbstregulatorische Fähigkeiten* (d. h. v. a. Planung, Monitoring, Evaluation) eine wichtige Rolle (Schneider & Artelt, 2010). *Planungsprozesse bzw. –phasen*, die sich durch geeignete Fragen einleiten lassen (etwa: Worin besteht das Problem? Wie fange ich an? Was ist gegeben? Was ist gesucht? Hilft mir eine Skizze? Habe ich die Aufgabe mehrfach gelesen?), sind expliziter Bestandteil vieler mathematischer Modelle des Problemlösens (z. B. Carlson & Bloom, 2005; Krulik & Rudnick, 1987; Pólya, 1945; Schoenfeld, 1983a, vgl. Tabelle 1).

Bei der Auswahl der Vorgehensweise bzw. Operatoren spielt *konditionales Wissen* eine Rolle. Mathematisch-fachdidaktische Studien zu starrer Anwendung (Übergeneralisierung) von Schlüsselwortstrategien (Schoenfeld, 1982) oder zu Kapitänsaufgaben (Baruk, 1989), d. h. mathematisch sinnlosen unlösbaren Aufgaben (Stern, 1992), weisen eindrucksvoll auf Fallstricke für SuS beim Lösen mathematischer Aufgaben hin (etwa: Überlesen, starre Regelanwendung, Überlernen, Glaube an die Lösbarkeit), auch wenn die SuS dabei teilweise rational vorgehen (Selter, 1994). Diese Studien lassen sich als weitere empirische Belege für die Bedeutung von Metakognition und konditionalem Wissen in der Mathematik auffassen (vgl. auch die Studien von Swanson, 1990).

Die Bedeutung der theoretischen Komponenten des Problemlösens für die Items aus PISA 2003 wird *empirisch gestützt durch eine Aufgabenanalyse* von Fleischer et al. (2010), die unter anderem Planungsfähigkeit, prozedurales und konditionales Wissen als wichtige gemeinsame Komponenten fächerübergreifender und mathematischer Problemlösekompetenz aufzeigt. Die Aufgabenanalyse zeit aber auch Unterschiede auf: Problemlöse-Items stellen anspruchsvollere Anforderungen bezüglich systematischen und strategischen Vorgehens, des Umgangs mit einschränkenden Bedingungen sowie prozeduralen Wissens. Mathematik-Items anderseits sind formalisierter und stellen (theoriekonform) höhere Anforderungen an mathematisches Sachwissen.

Zur allgemeinen Bedeutung der einzelnen Komponenten mathematischen Problemlösens werden unterschiedliche Positionen vertreten. Auf der einen Seite werden Selbstregulation und Heuristiken propagiert (Bielaczyc, Pirolli & Brown, 1995; Bruder, 2000, 2002; Bruder, Gürtler, Schmitz & Perels, 2002; Collet & Bruder, 2008; Gürtler, Perels, Schmitz & Bruder, 2004; Komorek, Bruder, Collet & Schmitz, 2006; Otto, Perels, Schmitz & Bruder, 2006; Perels, Schmitz & Bruder, 2003; Schoenfeld, 1987). Auf der anderen Seite wird die Rolle der Wissensbasis für Problemlösen und Transfer hervorgehoben (Carson, 2007; Perkins & Salomon, 1989; Schoenfeld, 1985). Die Frage, ob Wissen oder Selbstregulation und Heuristiken entscheidender für erfolgreiches Problemlösen seien (vgl. Carson, 2007), ist m. E. nicht zielführend. Zur Förderung mathematischer Problemlösekompetenz spielen alle Komponenten eine wichtige Rolle.[7] Daher werden kombinierte Ansätze für den Mathematikunterricht empfohlen, die sowohl domänenspezifisches Wissen als auch generelle Strategien und Metakognition beinhalten (Bransford, Sherwood, Vye & Rieser, 1986; Kießwetter, 1983; Perkins & Salomon, 1989).

2.4.3 Alternative Erklärungen

Allgemeine und sozialpsychologische Gründe schwacher Mathematikleistungen, die in Kapitel 2.4.3.1 und 2.4.3.2 vorgestellt werden, reichen nicht aus, um die Diskrepanz von fächerübergreifendem und mathematischen Problemlösen bei PISA 2003 zu erklären. Es gibt jedoch neben der kognitiven Potenzialausschöpfungshypothese (Leutner et al., 2004; OECD, 2004b) eine Reihe weiterer möglicher Erklärungen des Befundmusters aus PISA 2003, die anschließend vorgestellt und diskutiert werden (Kapitel 2.4.3.3 bis 2.4.3.9).

2.4.3.1 Allgemeine Gründe schwacher Mathematikleistungen

Auch wenn die individuellen Gründe für schwache mathematische Leistungen nicht sicher diagnostizierbar sind, da eine Monokausalitätsannahme unplausibel ist, so sind auf Gruppenebene Faktoren identifizierbar, die mit einer schwachen mathematischen Leistung einhergehen (Denvir, Brown & Stolz, 1982). Wie für andere Schulfächer gibt es im Fach Mathematik zahlreiche mögliche Gründe für schwache Leistungen von SuS (Barnes, 2005; Denvir et al., 1982), z. B.

– *physiologische* Gründe (sensorische Defekte und Wahrnehmungsprobleme, Probleme des Gedächtnisses).

– *kognitive* Gründe (mangelnde Sprach- und Lesekompetenz, mangelndes strategisches und prozedurales Wissen; Cardelle-Elawar, 1992).

– *sozial-emotionale* Gründe wie emotionale Probleme, Depression, Mathematikangst, Stereotypbedrohungen (Johnson, Barnard-Brak, Saxon &

[7] Weder Schulsysteme noch Lehrer müssen sich zwischen Sachwissen und Selbstregulation / Heuristiken entscheiden. Schüler verbringen Jahre in der Schule. Es geht m. E. nicht um die Frage, ob mathematisches Sachwissen oder Selbstregulation und Heuristiken im Unterricht behandelt werden sollen. Beides ist psychologisch nützlich, um Probleme zu lösen, und politisch in den Bildungsstandards verankert, sondern um die Frage, wie Problemlösekompetenzen bestmöglich gefördert werden können.

Johnson, 2012), Strebervorwürfe (Pelkner & Boehnke, 2003; Pelkner, Günther & Boehnke, 2004), geringe Wertschätzung, die Mathematik in Deutschland entgegengebracht wird (Dambeck, 2013; Enzensberger, n.d.; Hesse, 2014).

- *Lehrermerkmale* wie unzureichendes Fachwissen, z. B. über Heuristiken (Törner & Zielinski, 1992), pädagogisches Wissen, Motivierung, Organisation.
- *Unterrichtsmerkmale* wie Materialien und Lernumgebungen.
- *Systembedingungen* wie häufige Lehrerwechsel, Zeitdruck, mangelnde Qualitätskontrolle.

Ferner sind Interaktionen verschiedener Gründe denkbar, etwa Absentismus (Schulschwänzen) als Konsequenz des Zusammenspiel von Mathematikangst der SuS und autoritärer Persönlichkeit des unterichtenden Lehrers.

Diese Gründe können jedoch nicht die PISA-Befundmuster zur Diskrepanz fächerübergreifender und mathematischer Kompetenz erklären, da dazu Ergebnisse international vergleichender mathematischer Fachdidaktik z. B. zu fächerübergreifendem und mathematischen Problemlösen benötigt werden.

2.4.3.2 Sozialpsychologische Gründe schwacher Mathematikleistung

Aus sozialpsychologischer Perspektive kann Leistung u. a. durch Strebervorwürfe und Stereotypbedrohungen negativ beeinflusst werden. In einer kulturvergleichenden Studie anhand einer Stichprobe aus Deutschland, Kanada und Israel konnte im Streber-Projekt (Pelkner & Boehnke, 2003; Pelkner et al., 2004) gezeigt werden, dass *Strebervorwürfe* vor allem Mädchen in Deutschland hindern, ihre mathematischen Fähigkeiten auszuschöpfen. Dass sich jedoch Schüler ganzer Länder im Mittel Strebervorwürfen ausgesetzt fühlten und daher weniger Leistung zeigten, erscheint unplausibel, zumal das nicht ausgeschöpfte kognitive Potenzial nach Leutner et al. (2005) vor allem im *unteren* Leistungsbereich vorhanden sei.

Leistungsbezogene Stereotype (z. B. „Frauen sind schlechter in Mathematik.“) sind soziale Stigmata kognitiver Unterlegenheit (Aronson et al., 1999). Sie können bei ihrer Aktivierung, der *Stereotypbedrohung,* die Testleistung schwieriger Aufgaben in relevanten Domänen senken (Spencer, Steele & Quinn, 1999). Ein aktuelles Experiment (Johnson et al., 2012) zeigte, dass (a) schon ein Stereotyp aktivierender Satz, dass Frauen in diesem Mathematiktest schlechter abschneiden, die Leistung von Frauen senken kann, und (b) Männer unter der Bedingung Stereotypbedrohung (d. h. wenn sie glaubten, an einem Mathematiktest teilzunehmen, bei dem Männer schlechter abschneiden) sogar bessere Ergebnisse erzielten, (c) in der Kontrollgruppe keine Geschlechtsunterschiede auftraten. Zur Erklärung der PISA-Befunde eignet sich das Phänomen der Stereotypbedrohung jedoch nicht, da die Testhefte keine Hinweise erwarteter Geschlechtseffekte enthalten und desweiteren das Stereotyp selektiv in bestimmten Ländern gelten müsste, um das teilweise unterschiedliche Abschneiden in den Domänen Mathematik und Problemlösen zu erklären (vgl. Fall 3; siehe Kapitel 2.3).

Zusammenfassend sind die beiden sozial-kognitiven Phänomene Strebervorwürfe und Stereotypbedrohung mögliche Erklärungen für *differentielle Effekte* und Anlass für pädagogische Interventionen. Sie können jedoch nicht die Differenz zwischen Problemlösen und Mathematik erklären, der in Deutschland bei PISA 2003 im internationalen Vergleich deutlich geworden ist.

2.4.3.3 Intra- und interindividuelle Kontext-Effekte

Kontexteffekte spielen beim Lernen und während des Abrufs von Informationen eine wichtige Rolle (vgl. *realistic mathematics education*, Oberflächenmerkmale beim analogen Problemlösen, *cued recall*). Der Aufgaben-Kontext kann z. B. abstrakt-mathematisch („Löse $x^2 - 5 = -1$“) oder konkret-realistisch sein („2 Kinder wollen 6 Bonbons gerecht unter sich aufteilen. Wie viele Bonbons erhält jedes Kind?“) oder ein eingekleidetes Pseudo-Problem darstellen. Beispiele für Letztere, d. h. inhaltlich praktisch bedeutungslose Textaufgaben, sind die Aufgaben bei Bassok, Wu und Olseth (1995), bei denen u. a. nach der Wahrscheinlichkeit einer Zuordnung von bestimmten Computer-Seriennummern zur Erfahrung der Sekretärinnen gefragt wird.

Intraindividuelle Vergleiche der Lösungsquoten *strukturgleicher* Aufgaben mit innermathematischem und außermathematischem Kontext (Zöttl, Heinze & Reiss, 2007) zeigten, dass bei einer mittleren Korrelation der Leistung zwischen den Kontexten ($r = .38$, $p < .01$) Aufgaben mit außermathematischem Kontext deutlich leichter sind. Als Begründungen werden die Kontextgebundenheit der Strategien und nicht näher bestimmte „individuelle und kontextspezifische Präferenzen“ (S. 229) vermutet. Warum manche Kontexte erleichternd bzw. erschwerend wirken, bleibt offen. Fehlendes mathematisches Fachwissen kann hier als Begründung weitgehend ausgeschlossen werden, da die benötigten Definitionen (das „Begriffswissen“) angegeben waren. Diese Studie zeigt damit erneut, dass strukturelle Gemeinsamkeiten alleine nicht ausreichen, um spontanen Transfer zu erreichen (vgl. Stein, 1986).

Realistische Kontexte müssen jedoch nicht zwingend zu besseren Leistungen führen. Beispielsweise stellten Yoshida, Verschaffelt und De Corte (1997) Fünftklässlern in Belgien und Japan Standardprobleme (z. B. „Kuniko hat 5 Bretter von je 2 m gekauft. Wie viele Bretter mit je 1 m Länge kann sie aus diesen Brettern sägen?“ [Übersetzung des Autors]) und sogenannte parallele problematische Probleme (z. B. „Kuniko hat 4 Bretter von je 2.5 m gekauft. Wie viele Bretter mit je 1 m Länge kann sie aus diesen Brettern sägen?“ [Übersetzung des Autors]). Die Ergebnisse zeigten schon für Fünftklässler die Tendenz, den gesunden Menschenverstand und realistische Überlegungen beim Bearbeiten mathematischer Aufgaben zu ignorieren.

Interindividuell erzeugte Kontext-Effekte in einem Quasi-Experiment, in dem pro Schulklasse durch Deckblätter „Mathematik“ bzw. „Problemlösen“ der Testkontext salient gemacht wurde, sprechen für die Bedeutung der Selbstwirksamkeit und Mathematikangst beim fächerübergreifenden Problemlösen (Fleischer, Wirth & Leutner, 2014) und damit für eine Hypothese mangelnder Potenzial*nutzung*. Allerdings beruhen die Ergebnisse nur auf einer deutschen Ad-hoc-Stichprobe. Ein internationaler Vergleich ist somit nicht möglich. Und die praktische Konsequenz kann auch nicht ein Umtaufen des Mathematikunterrichts in Problemlöseunterricht sein.

2.4.3.4 Testmotivation

War bei PISA 2003 womöglich die Testmotivation der SuS in einigen Ländern bei der Bearbeitung des Mathematiktests geringer als bei der Bearbeitung des Problemlösetests? Waren die Problemlöseaufgaben interessanter als die Mathematikaufgaben (vgl. Kontexteffekte)? Testmotivation, als Teil der Leistungsmotivation, ist ein augenscheinvalides, theoretisch plausibles Konstrukt. Testmotivation wird bei *low stakes testing*, d. h. bei Testungen, die für die Getesteten relativ frei von Risiken und negativen Folgen im Falle des Scheiterns sind, für die Getesteten oft statistisch ignoriert (Asseburg & Frey, 2013). Insbesondere bei leistungsschwächeren SuS ist fraglich, ob sie die theoretisch erwartete maximale Testleistung zeigten (Asseburg & Frey, 2013), wenn die Testung für sie, wie bei PISA 2003, nicht persönlich bedeutsam ist. Experimentelle Befunde im Kontext von PISA 2000 (Max-Planck-Institut für Bildungsforschung, Berlin, n.d.) zeigten zumindest für einen Mathematiktest, dass die Testleistung und die berichtete Anstrengungsbereitschaft unabhängig davon waren, ob die SuS glaubten,

- an einer internationalen Bildungsstudie teilzunehmen,
- an einem Mathematiktest teilzunehmen, zu dem sie eine Rückmeldung über die Anzahl gelöster Aufgaben von ihrem Lehrer bekommen,
- der Mathematiktest würde von ihrem Mathematiklehrer benotet,
- sie würden für 10 Deutsche Mark an einem Mathematiktest teilnehmen.

Diese Befunde können jedoch nicht die Fragen beantworten, ob sich die SuS *maximal* angestrengt haben, und ob es bei Problemlösetests andere Ergebnisse bezüglich der Anstrengungsbereitschaft bzw. Motivation gibt als bei Mathematiktests (vgl. Kontexteffekte).

Empirische Belege für differentielle motivationale Effekte für die Subtests Mathematik bzw. Problemlösen bei PISA 2003 lassen sich aus den bei PISA 2003 erhobenen Daten nicht gewinnen, da affektiv-motivationale Variablen nicht pro Testdomäne erfasst worden sind. Um das internationale Muster im Bereich Mathematik und Problemlösen, das zur kognitiven Potenzialausschöpfungshypothese führte, durch Testmotivation zumindest partiell zu erklären, müsste es zwischen den teilnehmenden Ländern differentielle motivationale Effekte der teilnehmenden SuS hinsichtlich der beiden Domänen geben. Diese These ist mit den PISA-2003-Daten nicht zu beantworten. Eine eigene explorative Umfrage an einer kleinen Stichprobe in Deutschland ergab keine domänenspezifischen motivationalen Unterschiede hinsichtlich der Einschätzung von Problemlöse- und Mathematikaufgaben aus PISA 2003 (siehe Anhang A).

2.4.3.5 Realistic mathematics education und mangelnde Testfairness?

Die Konzeption der Mathematikaufgaben bei PISA ist durch das niederländische Konzept der *realistic mathematics education* (RME) von Hans Freudenthal (1968; 1991) und dem *Freudenthal Institute for Science and Mathematics Education* geprägt (Baumert et al., 2007; Neubrand et al., 2001; Neubrand, 2003). Wichtige

Merkmale dieses mathematikdidaktischen Ansatzes sind die Betonung eines Realitätsbezuges durch die Verwendung von Lernaufgaben mit bedeutsamem realem Kontext (anknüpfend an Erfahrungen der Lernenden), die Mathematisierung realer Phänomene (horizontale Mathematisierung; vgl. Modellierung) und innermathematischer Problemlösungen (vertikale Mathematisierung) sowie sozial-konstruktivistische Komponenten wie Prozesse des angeleitenden Wieder-Erfindens mathematischer Ideen (*guided reinvention*) durch aktive und interagierende SuS (van den Heuvel-Panhuizen, 2000, 2003).

Zeigt sich bei PISA 2003 in den Niederlanden das zu Deutschland inverse Muster, d. h. ein besseres Abschneiden in Mathematik als beim Problemlösen (OECD, 2004b), womöglich weil die PISA-Mathematikaufgaben und die niederländische Schulmathematik durch RME geprägt sind? Prägnant formuliert: Wurden einige Teilnehmerländer durch die Art (bzw. Auswahl) der Mathematik-Aufgaben bevorteilt bzw. benachteiligt? Wurden die SuS in Deutschland durch die Art (Auswahl) der Mathematikaufgaben benachteiligt? Es stellt sich die Frage nach Konstruktvalidität bzw. Testfairness des PISA-Mathematiktests. Gegen diese Testfairness-These für die Ergebnisse in Deutschland spricht der starke Befund, dass Mathematik-Schülerleistungen aus PISA 2003 und Leistungen in einem für Deutschland curricularem Mathematiktest latent mit .91 korrelieren (Prenzel & Drechsel, 2004). Man beachte ferner, dass das gute Abschneiden niederländischer SuS bei PISA 2003 in Mathematik womöglich zum Teil durch die Testkonzeption bei PISA erklärbar ist. Das erklärt jedoch nicht das relativ schwächere Abschneiden niederländischer SuS im Bereich Problemlösen, da die *realistic mathematics education* stark auf realistische Problemlöse- und Modellierungsaufgaben setzt.

Insgesamt erscheint diese These mangelnder Testfairness wenig plausibel, um das im Vergleich zum Problemlösen relativ schwächere Abschneiden von SuS in Deutschland in der Domäne Mathematik zu erklären.

2.4.3.6 Identitätshypothese

Mit Blick auf die sehr hohen latenten Korrelationen zwischen Problemlösen und Mathematik bei PISA 2003 folgern Reiss und Törner (2007), dass Problemlösen und Mathematik bei PISA mehr oder weniger dasselbe messen. Auch wenn Reiss und Törner vorsichtig formulieren, so formulieren sie doch eine *Identitätshypothese*. Die vermeintlich verschiedenen Konstrukte Problemlösen und Mathematik seien gar nicht verschieden. Die These, beide Tests, d. h. der PISA-Mathematiktest und der PISA-Problemlösetest, würden dasselbe Konstrukt messen, wirft Fragen auf: Was ist das erfasste Konstrukt *X*, das mit den Tests mit den Labels *Mathematik* und *Problemlösen* erfasst wird? Ist es Mathematik wie Reiss und Törner (2007) meinen? Ist es Problemlösen? Wie erklärt man dann, dass es in einigen Ländern (z. B. in Deutschland, Niederlande) deutliche Unterschiede beim Abschneiden in den beiden Tests gibt, wenn sie doch dasselbe Konstrukt erfassen? Wie plausibel ist diese Identitätshypothese? Grundlage dieser Identitätshypothese sind u. a. die hohe latente Korrelation ($r = .89$) und die theoretischen Ähnlichkeiten der Domänen Problemlösen und Mathematik. Einschränkend sei bemerkt, dass die Korrelation zwar sehr hoch, aber auch nicht perfekt ist. 20.79 Prozent der Varianz werden nicht aufgeklärt. Zudem sind auch die latenten Korrelationen der Lesekompetenz mit der Problemlösekompetenz ($r = .82$), mit der mathematischen Kompetenz ($r = .77$) und mit der naturwissenschaftlichen Kompetenz ($r = .83$) ähnlich hoch (OECD, 2004b).

Aus Perspektive der Identitätshypothese stellt sich die Frage nach ungenutztem Potenzial im Sinne der kognitiven Potenzialausschöpfungshypothese (Leutner et al., 2004; OECD, 2004b) nicht: Die Differenz zweier Tests die dasselbe Konstrukt erfassen, ist weniger auf inhaltlicher Seite als auf methodischer Seite (z. B. unterschiedliche Testmerkmale und -anforderungen) zu suchen.

2.4.3.7 Mathematik und Problemlösen als kognitive Grundfertigkeiten?

Eine gewisse Fortführung der Identitätshypothese greift ebenfalls die hohen latenten Skaleninterkorrelationen bei PISA 2003 auf. Erfassen alle *vier* PISA-Tests (Problemlösen, Mathematik, Naturwissenschaften, Lesen) womöglich dasselbe Konstrukt? Misst PISA vielleicht nur allgemeine psychometrische Intelligenz (*g-factor*) bzw. kognitive Grundfertigkeiten (Rindermann, 2006)? Diese Argumentation wird aus Perspektive der empirischen Bildungsforschung nicht geteilt (siehe Prenzel, Walter & Frey, 2007), zumal es empirische Argumente für die fachspezifische Trennbarkeit der Kompetenzen gibt (Baumert et al., 2007). Mathematische Kompetenzen und kognitive Grundfertigkeiten sind empirisch trennbar wie verschiedene Validitätsstudien anhand von *Nested-factor*-Modellen zeigten, die sowohl auf die Bedeutung kognitiver Grundfertigkeiten wie Schlussfolgern als auch auf fachspezifische mathematische Komponenten für das Lösen mathematischer

Tests hinweisen (z. B. Köller & Saß, 2013; Winkelmann, Robitzsch, Stanat & Köller, 2012).

2.4.3.8 Auswertungsfehler (technisches Argument)

Die Potenzialdiskrepanz wird von Wuttke (2006) verneint. Analysiere man die Daten von PISA 2003 „richtig“, sei der Unterschied zwischen der Leistung beim fächerübergreifenden Problemlösen und Mathematik bei PISA 2003 in Deutschland nicht statistisch signifikant (Wuttke, 2006). Damit wird die Potenzialausschöpfungshypothese gemäß dieser Position gegenstandslos. Wuttkes Einwände gegen die PISA-Auswertung, die sich noch auf weitere Aspekte beziehen (Wuttke, 2006), wurden in der Stellungnahme von Köller (2006) weitgehend für unberechtigt erklärt oder ignoriert (Meyerhöfer, 2006). Eine Überprüfung der statistischen Einwände Wuttkes durch neutrale Dritte, die eine Bewertung der Einwände zulässen würde, fehlt bislang (vgl. Kapitel 4.3.2).

2.4.3.9 Bewertungsproblem

Die PISA-2003-Ergebnisse zeigten, dass insbesondere SuS in Deutschland beim Problemlösetest im Durchschnitt eine höhere Punktzahl erreichten als beim Mathematiktest (Leutner et al., 2004). Das könnte man statt der Interpretation der erwartungswidrig niedrigeren Mathematikkompetenz auch invers interpretieren: SuS in Deutschland schneiden beim Problemlösen besser ab, als auf Grundlage der Mathematikergebnisse erwartet wurde – interessanterweise ist genau diese Umkehrung von Prädiktor und Kriterium bei PISA 2012 zu finden (OECD, 2014). Prägnanter formuliert: Das System Schule hilft oder verhindert zumindest nicht, dass SuS lernen, Probleme mit Alltagsbezug zu lösen. SuS lernen das Lösen von Problemen nicht nur im Mathematikunterricht. Die SuS in Deutschland erzielen bei PISA 2003 hohe Werte beim Problemlösen trotz relativ geringerer Mathematikkompetenz. SuS werden also trotz vermeintlich schlechten Mathematikkompetenzerwerbs mit alltagsnäheren Problemen fertig. Das ist insofern plausibel, da zu deren Lösung nicht nur Mathematik benötigt wird. Aus dieser Perspektive erübrigt sich die Untersuchung der kognitiven Potenzialausschöpfungshypothese in Deutschland, da das Vorhandensein ungenutzten Potenzials im Sinne der kognitiven Potenzialausschöpfungshypothese verneint wird.

Wie interpretationsbedürftig die Diskrepanz-Befunde zu den Domänen Problemlösen und Mathematik bei PISA 2003 sind, zeigt sich auch am Beispiel der niederländischen Ergebnisse. In den Niederlanden beträgt die Differenz 18 Punkte zugunsten der Mathematikleistung und ist so groß wie in keinem anderen OECD-Teilnehmerland bei PISA 2003 (OECD, 2004b). Im Vergleich zu Deutschland zeigte sich in den Niederlanden ein inverses Befundmuster (Fall 1, vgl. Kapitel 2.3). Die OECD-Interpretation eines erfolgreichen niederländischen Mathematikunterrichts (OECD, 2004a) wird von Doorman et al. (2007) um die These ergänzt, „problem-solving skills are lagging behind" (S. 408), die als Begründung der Forderung nach *Veränderungen* des niederländischen Mathematikunterrichts verwendet wird.

2.5 Training und Transfer

In Kapitel 2.4 wurden verschiedene Erklärungen der Diskrepanz von Problemlösen und Mathematik bei PISA 2003 in Deutschland dargestellt. In dieser Arbeit soll hauptsächlich die kognitive Potenzialausschöpfungshypothese (Leutner et al., 2004; OECD, 2004b) untersucht werden. Dazu sind Trainingsstudien notwendig, die untersuchen, welche geförderten Teilkomponenten der fächerübergreifenden Problemlösekompetenz möglicherweise zu verbesserten Leistungen im Fach Mathematik führen (siehe Kapitel 3). Deshalb werden in diesem Kapitel Ergebnisse der Trainings- und Transferforschung vorgestellt. Es werden zunächst verschiedene Trainingsansätze skizziert. Anschließend werden der Begriff Transfer erläutert und Transfertheorien beschrieben.

2.5.1 Trainingsansätze

Klauer (2001) unterscheidet fünf Ansätze zur Entwicklung von (kognitiven) Trainingsprogrammen, d. h. wiederholte Handlungen, die auf die Verbesserung kognitiver Tätigkeiten bzw. Leistungs*potenziale* (Hasselhorn & Gold, 2006) abzielen:

1. *Testaufgabenansatz*: Grundidee des Testaufgabenansatzes ist die plausible Annahme, dass zur Übung Aufgaben verwendet werden, die in ähnlicher Form als Testaufgaben zum Einsatz kommen. „Problemlösen lernt man demnach durch Problemlösen" (Klauer, 2001, S. 6). Dieses Trainingsprinzip ist jedoch bei komplexen Trainingszielen – wie Problemlösen – meist zu einfach, etwa wenn die Trainierenden zur Problemlösung notwendige Teilleistungen noch nicht beherrschen (Klauer, 2001).

2. *Kognitive Korrelate*: Ausgehend von einem angenommenen Kausalzusammenhang zweier Leistungen A und B, der sich empirisch zumindest in ei-

ner positiven Korrelation zeigen sollte, kann das Training von A auch zur Leistungssteigerung bei B führen, weil ein gemeinsamer Varianzanteil existiert, der z. B. auf ähnliche Anforderungen oder Prozessschritte zurückgeht (Klauer, 2001). So fördert beispielsweise Klauers Denktraining bei lernbehinderten Sonderschülern auch die Leistung bei mathematischen Textaufgaben (Sonntag, 2004).

3. *Trainingsmethoden:* Bei diesen Trainings sind die Methoden selbst Gegenstand des Trainings, z. B. beim Erlernen metakognitiver Fragen (King, 1991), Selbstregulationsstrategien (Otto, Perels & Schmitz, 2008) oder beim Methoden-Training (Klippert, 2002).
4. *Prozessanalyse:* Ausgehend von Analysen der interessierenden Prozesse und Komponenten, z. B. von Strategien und metakognitiven Prozessen bei erfolreichen Problemlösern, werden Trainings konzipiert (Klauer, 2001). Ein zusätzlicher Vergleich erfolgreicher Problemlöser mit dem Vorgehen von Novizen ist hilfreich (Klauer, 2001).
5. *Instruktionspsychologie:* Der instruktionspsychologische Ansatz basiert auf einem Vergleich des Soll-Zustands, gewonnen aus theoretischen Überlegungen, bildungspolitischen Vorgaben oder Prozessanalysen bei Experten, mit dem Ist-Zustand der Lernenden. Die resultierende Differenz, der Lehrstoff, wird im Lehrplan (Curriculum) in Form von definierten Teilzielen festgehalten und mit lernpsychologisch angemessenen Methoden trainiert (Klauer, 2001).

Es ist zu beachten, dass die konkrete Durchführung kognitiver Trainings immer auch metakognitive Elemente enthält (z. B. werden Übungen erklärt und Fehler verbalisiert und korrigiert), und die verschiedenen Ansätze sind kombinierbar (Klauer, 2001). Die Experimente in Kapitel - nutzen die Ansätze 1 bis 4 in unterschiedlicher Gewichtung.

Die mögliche Bandbreite zusätzlicher positiver Trainingseffekte (z. B. Anwärmeeffekt, Hawthorne-Effekt, Novitätseffekt, Zuwendungseffekt, *testwiseness*, Abbau störender Ängste und Kognitionen, Tempomotivation, Einübung analytischen und systematischen Arbeitsstils) und negativer Trainingseffekte (z. B. Zeitverlust, psychische Sättigung, Demotivation durch Langeweile, Trainingsstress, mangelnde Automatisierung neuer Verhaltensweisen) ist groß (Klauer, 2001). Die intendierten und die zusätzlichen Trainingseffekte können dabei differentiell wirken: Bei Trainings, von denen v. a. leistungsschwächere Personen profitieren, spricht man von *kompensatorischen Trainingseffekten*. Profitieren hingegen anfänglich leistungsstärkere Personen besonders, spricht man vom *Matthäuseffekt* („Wer hat, dem wir gegeben“; Klauer, 2001) oder *Schereneffekt* (Hasselhorn & Gold, 2006). Matthäuseffekte sind insbesondere in Schulfächern zu erwarten und zu finden, deren Inhalte und Fertigkeiten stark aufeinander aufbauen (z. B. im Fach Mathematik; Murayama, Pekrun, Lichtenfeld & Vom Hofe, 2013).

2.5.2 Lerntransfer

Da im Rahmen konkreter Lehr-Lernprozesse Inhalte fast immer nur *exemplarisch* gelehrt bzw. erarbeitet werden (können), ist die Frage nach Generalisierbarkeit des Gelernten oder Lerntransfer (kurz Transfer) zentral für pädagogisch-psychologische Arbeit. Die Verwendung des Begriffs Transfer ist historisch uneinheitlich, wodurch kaum Konsens zu Bedingungen des Transfers besteht (Mähler & Stern, 2010). Allgemein lässt sich Transfer als Effekt früheren Lernens auf neues Lernen und Problemlösen definieren (Mayer, 2008). Detterman (1993, nach Klauer, 2011, S. 14) kritisiert, dass der Transfer-Begriff inflationär gebraucht werde und die bloße „Anwendung des Gelernten in neuen Situationen“ nicht als Transfer zu werten sei. Transfer wird spezifischer als ein *nichttrivialer Lerneffekt* definiert, d. h. als „Lerneffekt bei Aufgaben, die in dem fraglichen [Lern-]Prozess weder gelernt noch geübt wurden“ (Klauer, 2011, S. 17). Den Grad der Ähnlichkeit von Übungs- und Transferaufgaben bezeichnet man dabei als *Transferdistanz* (Klauer, 1989, 2011; Laker, 1990). Eine psychometrische Quantifizierung der Transferdistanz ist im Allgemeinen kaum möglich (für Spezialfälle siehe Klauer, 1989). Es sind daher die Unterscheidungen *naher* (oder *proximaler*) versus *weiter* (oder *ferner* oder *distaler*) Transfer auf der Grundlage von Experteneinschätzungen üblich (Mähler & Stern, 2010; Tomic, 1995). Mandl, Prenzel und Gräsel (1992, s.a. Schaper, 2004) klassifizieren anhand des Erfolgs bei nicht trainierten Aufgaben zudem die Qualität des Transfers. Liegt Trainingserfolg vor, spricht man von *positivem Transfer*, liegt kein Trainingserfolg vor, spricht man von *Null-Transfer*. Führt das Training zu schlechteren Leistungen, weil beispielsweise die gelernten Konzepte zu rigide angewendet werden, ohne die spezifischen Aufgabenanforderungen zu beachten (vgl. Übergeneralisierung), so spricht man von *negativem Transfer* (vgl. Anhang BB für ein Beispiel aus Experiment 3).

Da Instrumente fehlen, die allgemeinen Transfer von Lernleistungen erfassen, wird Transfer häufig über die Leistung in fachspezifischen Tests operationalisiert (Klauer, 2011). So geschieht es auch in dieser Arbeit.

2.5.3 Transfertheorien

Wie kann man das Auftreten bzw. Ausbleiben von Transfer theoretisch erklären? Im Laufe der Psychologie-Geschichte sind einige Theorien des Transfers formuliert worden (Edelmann & Wittmann, 2012; Klauer, 2011):

1. Die *Theorie der formalen Bildung* basiert auf der Annahme einer unspezifischen Generalisierung von Fähigkeiten und Fertigkeiten, die etwa durch die Beschäftigung mit Griechisch, Latein, oder Mathematik erworben wurden. Sie gilt im strengen Sinne als widerlegt.

2. Gemäß der *Theorie der identischen Elemente* (Thorndike & Woodworth, 1901) und ihrer Weiterentwicklungen (Singley & Anderson, 1989) findet Transfer nur statt, wenn Quellproblem (*source*) und Zielproblem (*target*) partiell identische Wissenselemente aufweisen.

3. Die *strukturalistische Transfertheorie* postuliert, dass erkannte „strukturelle Zusammenhänge, Prinzipien und Gesetzmäßigkeiten auf neue Inhalte übertragen werden“ (Edelmann & Wittmann, 2012, S. 227). Dies ist z. B. der Fall beim Problemlösen durch Anwendung von Analogien, dem *analogen Problemlösen* (vgl. auch *case-based reasoning*; Norman & Schmidt, 1992).

4. Beim *Transfer von Strategien* steht das Erlernen und Üben von Strategien im Fokus. Strategien können anhand ihrer *Spezifität* (spezifisch bis unspezifisch) klassifiziert werden, d. h. danach, ob sie eher einen engen Anwendungsbereich (z. B. vollständige Induktion) oder eher einen weiten Anwendungsbereich haben (z. B. Bildung von Teilzielen, Vorwärtsarbeiten, Rückwärtsarbeiten, indirekter Beweis). Eine weitere Klassifikation besteht entlang einer vereinfachten Kognition-Metakognitions-Dimension, d. h. danach, ob die Strategien eher kognitiv sind (z. B. vollständige Induktion) oder eher metakognitiv (Depaepe, De Corte & Verschaffel, 2010) wie bei Selbsterklärungsprompts (Thillmann, 2007) und Heuristiken (Perels, Gürtler & Schmitz, 2005).

Die Untersuchung von Problemlösen durch Analogien, bei dem erlernte Prinzipien oder Schemata in ähnlicher Art und Weise bei Übungsaufgaben (Quellproblem) und Testaufgaben (Zielproblem) anzuwenden sind, hat eine lange (experimental-)psychologische Forschungstradition (Anderson, 2013; Bassok, 2003; Gick & Holyoak, 1980, 1983; Hummel & Holyoak, 1997). Zentrale Ergebnisse sind unter anderem, dass die Ähnlichkeit der Oberflächen- und Tiefenstruktur von Quell- und Zielproblem beeinflusst, ob positiver bzw. negativer Transfer auftritt (Stein, 1986), und dass spontaner Transfer selten ist (siehe Chen, 2002; Schmid, Wirth & Polkehn, 2003, für Details und Theorien). Transfer kann zudem durch kognitive und metakognitive Strategien des Problemlösers mediiert werden (Mayer & Wittrock, 1996; vgl. Kap. 2.4.2.3).

Als weitere Voraussetzungen des Transfers von Problemlösefähigkeiten gelten Sachwissen (Carson, 2007), Schema-Erwerb und die Automatisierung von Regeln (Cooper & Sweller, 1987) sowie Anleitung, z. B. durch *priming*, *cuing*, *guiding* (Perkins & Salomon, 1989).

2.5.4 Training und Transfer von Problemlösen

Ist fächerübergreifende Problemlösekompetenz veränderbar und förderbar? Schon Tiere sind in der Lage, Probleme zu lösen (Köhler, 1963; Perler & Wild, 2005), die Menschheitsgeschichte ist voll von Problemlösungen und technischen Innovationen, und Kinder lernen im Laufe ihrer Entwicklung immer wieder neue Probleme zu lösen (vgl. Ontogenese; Lernen am Modell). Problemlösen findet statt und ist lernbar. Wie weit Problemlösen jedoch lehrbar ist, ist umstritten (vgl. z. B. Tricot & Sweller, 2014). Da Problemlösen im Gegensatz zu Intelligenz bei PISA als Kompetenz konzeptualisiert wird (OECD, 2003, 2004b), besteht die Annahme, Problemlösen lehren und lernen zu können. Denn bei PISA sind Kompetenzen nicht als stabile *traits*, sondern als Menge prinzipiell erlernbarer Fertigkeiten, Kenntnisse und als erlernbares metakognitives Wissen definiert (Klieme, Funke et al., 2001), d. h. Kompetenzen – wie die Problemlösekompetenz – gelten *per definitionem* als veränderbar und förderbar. Hingegen fehlen e*mpirische Belege der Förderbarkeit* der Schlüsselqualifikation Problemlösekompetenz weitgehend (Klieme, Artelt et al., 2001). Das gilt auch für die sich auf PISA 2003 beziehenden Modelle relevanter Teilkompetenzen der Problemlösekompetenz von Fleischer et al. (2010). Der Transfer von Schlüsselqualifikationen bzw. fächerübergreifenden Kompetenzen „auf neue Situationen wird häufig überschätzt oder einfach ungeprüft unterstellt“ (Klieme, Artelt et al., 2001, S. 205). Vor allem ist „unklar, ob ein Transfer von außermathematisch erworbenen Problemlösestrategien für innermathematische Problemstellungen stattfindet“ (Zöttl et al., 2007, S. 223). Die Experimente dieser Arbeit (Kapitel -) leisten im Rahmen labor- bzw. feldexperimenteller Möglichkeiten auch einen Beitrag zur Prüfung der Trainierbarkeitsannahme und der Transfereffekte der Teilkompetenzen des Problemlösens (Fleischer et al., 2010) und zur Frage, ob fächerübergreifende Problemlösekompetenz auf mathematische Problemlösekompetenz transferiert.

2.6 Forschungsfragen und Hypothesen

Zum Abschluss des theoretischen Hintergrunds werden die zentralen Forschungsfragen dargestellt.

Die Ergebnisse von PISA 2003 (OECD, 2004b) führten zur kognitiven Potenzialausschöpfungshypothese (Leutner et al., 2004; OECD, 2004b), die u. a. durch Ergebnisse der Messwiederholungsstudie PISA-I-Plus (Leutner et al., 2006) und einer Anforderungsanalyse der PISA-Items (Fleischer et al., 2010) weiter gestützt wird. Zur Bedeutung fächerübergreifenden Problemlösens für mathematisches Problemlösen liegen quer- und längsschnittliche korrelative Zusammenhänge vor (Fleischer et al., 2009; Leutner et al., 2004; 2006; OECD, 2004b). Randomisierte Kontrollgruppenexperimente zur kognitiven Potenzialausschöpfungshypothese fehlen bislang. Einer ersten experimentellen Untersuchung der kognitiven Potenzialausschöpfungshypothese soll im Rahmen der vorliegenden Arbeit in drei Experimenten nachgegangen werden.

Die methodische Idee der Experimente ist, die quer- und längsschnittliche Korrelation von fächerübergreifendem und mathematischem Problemlösen experimentell zu untersuchen, indem die beobachtete Variabilität von Komponenten der fächerübergreifenden Problemlösens durch experimentell zu manipulierende Variabilität ersetzt wird, um mögliche Effekte eines Trainings auf fächerübergreifendes und mathematisches Problemlösen kausal interpretieren zu können (Leutner, Rumann & Wirth, 2009).

In Experiment 1 und 2 werden zentrale Komponenten fächerübergreifender Problemlösekompetenz (konditionales Wissen, Planungsfähigkeit, Handlungswissen) in einem computerbasierten Experiment mit 45-minütiger Trainingsdauer trainiert (vgl. Leutner et al., 2009). Das Ziel ist die Untersuchung potentieller Transfereffekte von fächerübergreifender Problemlösekompetenz auf die mathematische Domäne, die als Evidenz für die kognitive Potenzialausschöpfungshypothese (Leutner et al., 2004; OECD, 2004b) zu werten wären. Kann eine systematische Verbesserung der fächerübergreifenden Problemlösekompetenz helfen, die Problemlösekompetenz im Schulfach Mathematik zu verbessern?

Um zu prüfen, ob die experimentelle Manipulation, d. h. das Training der Komponenten fächerübergreifender Problemlösekompetenz, erfolgreich ist, wird ein Treatment-Check vorgenommen:

> *Fragestellung 1* (*Treatment-Check*): Lassen sich zentrale Komponenten der fächerübergreifenden Problemlösekompetenz (konditionales Wissen, Planungsfähigkeit, Handlungswissen) mittels laborexperimenteller Lehr-Lernprogramme messbar fördern?

Dieser Treatment-Check dient zugleich als ein Schritt zur „experimentelle[n] Validierung der Kompetenzstruktur fächerübergreifenden Problemlösens“ (Leutner et al., 2009, S. 14). Zusätzlich werden potentielle Effekte des Trainings der Komponenten fächerübergreifenden Problemlösens auf fächerübergreifendes Problemlösen insgesamt, operationalisiert durch Items aus PISA 2003 (OECD, 2004b), untersucht:

> *Fragestellung 2* (*Transfer Problemlösen*): Steigert das Training weniger, zentraler Komponenten der fächerübergreifenden Problemlösekompetenz (konditionales Wissen, Planungsfähigkeit, Handlungswissen) die fächerübergreifende Problemlösekompetenz insgesamt?

Der Transfertest zur Beantwortung von Fragestellung 2 kann auch als zusätzlicher Treatment-Check aufgefasst werden, da die kognitive Potenzialausschöpfungshypothese auf Basis der Diskrepanz der Ergebnisse für fächerübergreifendes Problemlösen und Mathematik (Leutner et al., 2004) und nicht auf Grundlage der *Komponenten* fächerübergreifender bzw. mathematischer Problemlösekompetenz postuliert wurde (T. Kleickmann, persönl. Mitteilung, 7.10.2011). In Experiment 1 werden zur Beantwortung der Fragestellungen zum fächerübergreifenden Problemlösen (Fragestellungen 1 und 2) Effizienzmaße verwendet (siehe Kapitel 3.2.2.6), da es unerwartet zu unterschiedlichen Bearbeitungszeiten der Gruppen im Problemlösetest gekommen ist. In Experiment 2, einer Modifikation von Experiment 1, werden Leistungsmaße analysiert.

Ein erfolgreiches Training fächerübergreifenden Problemlösens ist eine theoretische Voraussetzung für den Existenznachweis potentieller Transfereffekte auf die mathematische Domäne. In Experiment 1 und 2 werden potentielle Effekte auf die mathematische Domäne durch zwei Fragestellungen unterschiedlicher Transferdistanz untersucht (Leutner et al., 2009):

> *Fragestellung 3* (*Naher Transfer Mathematik*): Steigert ein Training weniger, zentraler Komponenten der fächerübergreifenden Problemlösekompetenz (konditionales Wissen, Planungsfähigkeit, Handlungswissen) ebenfalls diese Komponenten der fachbezogenen Problemlösekompetenz im Fach Mathematik?

> *Fragestellung 4* (*Ferner Transfer Mathematik*): Steigert ein Training weniger, zentraler Komponenten der fächerübergreifenden Problemlösekompetenz (konditionales Wissen, Planungsfähigkeit, Handlungswissen) die fachbezogene Problemlösekompetenz im Fach Mathematik insgesamt?

In Experiment 1 und 2 werden zudem mögliche differentielle Effekte untersucht, da das ungenutzte kognitive Potenzial vor allem im unteren Leistungsbereich vorhanden sein sollte (Leutner et al., 2004, 2005). Empirische Studien im Schulbereich (Murayama et al., 2013) hingegen zeigen oft Trainingseffekte für leistungsstärkere SuS bzw. SuS mit höherem Vorwissen (Matthäuseffekt; Klauer, 2011):

> *Fragestellung 5* (*kompensatorischer Trainingseffekt*): Zeigen sich differentielle Trainingseffekte hinsichtlich der Schulform, d. h. profitieren eher leistungsschwächere SuS (Hauptschüler) vom Training?

In Experiment 3 erfolgte im Rahmen eines Feldexperiments im Schulkontext ein mehrwöchiges Training fächerübergreifender Problemlösekompetenz, das hinsichtlich des Trainingsansatzes deutlich breiter und ökologisch valider angelegt war als in den Laborexperimenten. Die Grundidee blieb dabei erhalten, d. h. es erfolgte ein Training fächerübergreifender Problemlösekompetenz, ein Treatment-Check und der Test möglicher Transfereffekte auf die mathematische Domäne. Im Gegensatz zu Experiment 1 und 2 wurde jedoch von den drei Teilkompetenzen aus Zeitgründen nur Planungsfähigkeit getestet:

> *Fragestellung 6 (Treatment-Check)*[8]: Lässt sich mit dem ursprünglich für Erwachsene konzipierten Aachener Planungskompetenztraining (Arling & Spijkers, 2012) die Planungskompetenz (als eine Teilkomponente des Problemlösens) in der Jahrgangsstufe 9 trainieren?

Wie in Experiment 1 und 2 werden in Experiment 3 Effekte des Trainings auf fächerübergreifende Problemlösekompetenz getestet. Die Fragestellungen 7 und 2 sind inhaltlich gleich, beziehen sich jedoch auf verschiedene Trainings:

> *Fragestellung 7 (Transfer Problemlösen)*: Steigert das fächerübergreifende Problemlösekompetenztraining mit Fokus auf spezifischen Komponenten (Planungsfähigkeit, konditionales Wissen, Handlungswissen, Heuristiken) die allgemeine fächerübergreifende Problemlösefähigkeit wie sie bei PISA 2003 erfasst wird?

Die Erfassung potentieller Transfereffekte auf die Domäne Mathematik wird in Experiment 3 aus Zeitgründen auf die Problemlösekompetenz im Fach Mathematik insgesamt beschränkt, d. h. es erfolgt keine Erfassung der Komponenten mathematischer Problemlösekompetenz:

> *Fragestellung 8 (Transfer Mathematik)*: Steigert das fächerübergreifende Problemlösekompetenztraining mit Fokus auf spezifischen Komponenten (Planungsfähigkeit, konditionales Wissen, Handlungswissen, Heuristiken) die Problemlösekompetenz im Fach Mathematik?

Werden Transfereffekte auf die Domäne Mathematik gefunden, so spricht dies für die Rolle fächerübergreifenden Problemlösens beim Aufbau mathematischer Kompetenz und damit für die kognitive Potenzialausschöpfungshypothese. Zusätzlich wird in Experiment 3 durch die Verwendung eines Prätests die Möglichkeit geschaffen, das vorhandene kognitive Potenzial im Sinne der kognitiven Potenzialausschöpfungshypothese (Leutner et al., 2004, 2005) vor Trainingsbeginn zu messen. Damit können differentielle Effekte hinsichtlich des kognitiven Potenzials untersucht werden, die eine direktere Prüfung der kognitiven Potenzialausschöpfungshypothese erlauben:

- *Fragestellung 9a (differentielle Effekte und kognitives Potenzial):* Profitieren im Sinne der kognitiven Potenzialausschöpfungshypothese (Leutner et

[8] Die Fragestellungen werden fortlaufend nummeriert, um eine eindeutige Referenzierbarkeit zu erlauben.

al., 2004; OECD, 2004b) insbesondere SuS vom Training, die im Prätest fächerübergreifenden Problemlösens eher hoch und im Prätest Mathematik eher niedrig abgeschnitten haben?

- *Fragestellung 9b (kognitive Potenzialausschöpfung):* Profitieren insbesondere SuS vom Training, bei denen zum Zeitpunkt des Prätests kognitives Potenzial im Sinne der kognitiven Potenzialausschöpfungshypothese (Leutner et al., 2004; OECD, 2004b) vorhanden ist?

3 Studien zur kognitiven Potenzialausschöpfungshypothese[9]

Basierend auf dem in Kapitel 2 dargestellten theoretischen Hintergrund, der zunächst kurz zusammengefasst wird, werden in diesem Kapitel drei Experimente berichtet, die Aspekte der kognitiven Potenzialausschöpfungshypothese (Leutner et al., 2004; OECD, 2004b) für das Fach Mathematik untersuchen.[10]

3.1 Einleitung

Ausgangspunkt der Experimente waren unerwartete Ergebnisse von PISA 2003 (Leutner et al., 2004; OECD, 2004b): Während die Ergebnisse beim Problemlösen (M = 513, SD = 95) über dem Durchschnitt (im Vergleich zum OECD-Mittelwert von 500; SD = 100) lagen, erzielten die SuS in Deutschland „nur“ durchschnittliche Ergebnisse in den Domänen Mathematik (M = 503, SD = 103), Naturwissenschaften (M = 502, SD = 111) und Lesen (M = 491, SD = 109).

Die Differenz zwischen fächerübergreifender Problemlösekompetenz und den fachspezifischen Kompetenzen, z. B. in Mathematik, wird besonders in Deutschland betont (Leutner et al., 2004), da nur in Ungarn und Japan die Differenzen zugunsten des Problemlösens größer waren als in Deutschland (OECD, 2004b). Diese große Differenz war besonders überraschend, da zwischen Problemlösen und Mathematik bei PISA 2003 eine hohe latente Korrelation gefunden wurde (r = .89; OECD, 2005). In Deutschland wird das Vorhandensein dieses ungenutzten Potenzials zudem für leistungsschwächere SuS betont (Leutner et al., 2004, 2005). Diese Differenz bzw. Diskrepanz zwischen fächerübergreifender Problemlösekompetenz und der Kompetenz in Mathematik kann durch die *kognitive Potenzialausschöpfungshypothese* erklärt werden. Das bessere Abschneiden im Problemlösetest relativ zum Mathematiktest zeigte, dass SuS über ungenutztes kognitives Potenzial verfügen, das im Fach Mathematik möglicherweise nicht voll zum Aufbau fachlicher Problemlösekompetenz ausgeschöpft wurde (OECD,

[9] Teile der vorläufigen Ergebnisse von Experiment 1 und 3 wurden bereits auf wissenschaftlichen Tagungen und Konferenzen präsentiert: Buchwald, Fleischer und Leutner (2012, Oktober), Buchwald, Fleischer und Leutner (2013, März), Buchwald, Arling, Fleischer und Leutner (2013), Buchwald, Fleischer, Arling und Leutner (2014, February), Leutner (2013a), Leutner (2013b). Entsprechende Publikationen befinden sich in Vorbereitung: Buchwald, Fleischer, Rumann, Wirth und Leutner (in preparation), Buchwald, Fleischer und Leutner (in preparation), Fleischer, Buchwald, Wirth, Rumann und Leutner (in preparation).

[10] Möglicherweise ungenutztes kognitives Potenzial für naturwissenschaftlichen Unterricht wird in einem anderen Projekt untersucht (DFG-Kennzeichen RU 1437/4-3).

2004b; Leutner et al., 2004). Für diese kognitive Potenzialausschöpfungshypothese gibt es einige Argumente:

1. Das Lösen fächerübergreifender und mathematischer Probleme erfordert ähnliche *theoretische Prozesse* bzw. Schritte (vgl. Pólya, 1945, vgl. auch die zyklischen Modelle der mathematischen Modellierung, Carlson & Bloom, 2005).
2. In den Domänen Problemlösen und Mathemaik *werden ähnliche kognitive Anforderungen* gestellt (eher niedrige Anforderungen an Lesekompetenz und naturwissenschaftliche Kompetenz, hoher Grad an schlussfolgerndem Denken (OECD, 2003; vgl. Fleischer, Buchwald, Wirth, Rumann & Leutner, in preparation).
3. Analysen der deutschen Messwiederholungsstudie PISA-I-Plus zur Kompetenzentwicklung in Mathematik von Jahrgangsstufe 9 zu Jahrgangsstufe 10 zeigen die *Bedeutung fächerübergreifender Problemlösekompetenz beim Aufbau fachlicher Kompetenzen in der Mathematik* (Leutner et al., 2006).
4. Die Ergebnisse einer Kommunalitätenanalyse derselben Daten (Leutner et al., 2006) implizieren, dass fächerübergreifendes *Problemlösen und Mathematik aus verschiedenen – teilweise überlappenden – Komponenten bestehen, die unterschiedlich zum Erwerb zukünftiger Kompetenz im Fach Mathematik beitragen.*
5. Theoretisch lassen sich *verschiedene Komponenten der Problemlösekompetenz* unterscheiden (z. B. Sachwissen, Handlungswissen, konditionales Wissen, allgemeine Problemlösestrategien, selbstregulatorische Fähigkeiten wie Planung, Monitoring, Evaluation), die in beiden Domänen, Problemlösen und Mathematik, wichtig sind (Kapitel 2.4.2.3 und 2.4.2.4).
6. Eine Aufgabenanalyse von Fleischer et al. (2010) liefert *empirische Argumente für die Bedeutung der theoretischen Komponenten des Problemlösens* für die PISA-2003-Items. Sie zeigt u. a. Planungsfähigkeit als Aspekt der Selbstregulation, prozedurales und konditionales Wissen als wichtige *gemeinsame Komponenten* fächerübergreifender und mathematischer Problemlösekompetenz auf.

Zusammenfassend gibt es sowohl auf einer theoretischen als auch auf einer empirischen Ebene Evidenz für eine starke Überschneidung zwischen fächerübergreifender und mathematischer Problemlösekompetenz. Daher sollte eine Verbesserung (von Komponenten) fächerübergreifender Problemlösekompetenz auch zu Transfereffekten auf die mathematische Domäne führen. Um diese Idee zu untersuchen, sind experimentelle Studien nötig.

3.2 Experiment 1

3.2.1 Fragestellungen

Die bisherigen korrelativen Zusammenhänge fächerübergreifenden und mathematischen Problemlösens (Fleischer et al., 2009; Leutner et al., 2004; 2006; OECD, 2004b) sollen in Experimenten kausal überprüft werden. Führt erhöhte fächerübergreifende Problemlösekompetenz zu erhöhter mathematischer Problemlösekompetenz? Als Trainingsmethode werden gemeinsame Komponenten fächerübergreifenden und mathematischen Problemlösens (Fleischer et al., 2010) genutzt. Führen verbesserte Teilkompetenzen fächerübergreifenden Problemlösens zu erhöhter fächerübergreifender und mathematischer Problemlösekompetenz? Um dies zu untersuchen, wird die beobachtete Variabilität der Komponenten bzw. Teilkompetenzen fächerübergreifender Problemlösekompetenz ersetzt durch experimentell zu verursachende Variabilität (Leutner et al., 2009). In Experiment 1 werden dazu Komponenten fächerübergreifender Problemlösekompetenz trainiert und Transfereffekte auf die fächerübergreifende Problemlösekompetenz insgesamt, auf Komponenten mathematischer Problemlösekompetenz und die mathematische Problemlösekompetenz insgesamt getestet. Durch die Analyse möglicher Transfereffekte auf die mathematische Problemlösekompetenz wird ein Aspekt der kognitiven Potenzialausschöpfungshypothese (Leutner et al., 2004; OECD, 2004b) geprüft (Leutner et al., 2009). Im Detail werden mit Experiment 1 folgende Fragestellungen untersucht (vgl. Leutner et al., 2009):

- *Fragestellung 1* (*Treatment-Check*): Lassen sich zentrale Komponenten der fächerübergreifenden Problemlösekompetenz (konditionales Wissen, Planungsfähigkeit, Handlungswissen) mittels laborexperimenteller Lehr-Lernprogramme messbar fördern?
- *Fragestellung 2* (*Transfer Problemlösen*): Steigert dieses Training weniger, zentraler Komponenten der fächerübergreifenden Problemlösekompetenz (konditionales Wissen, Planungsfähigkeit, Handlungswissen) die fächerübergreifende Problemlösekompetenz insgesamt?

- *Fragestellung 3* (*Naher Transfer Mathematik*): Steigert dieses Training weniger, zentraler Komponenten der fächerübergreifenden Problemlösekompetenz (konditionales Wissen, Planungsfähigkeit, Handlungswissen) ebenfalls diese Komponenten der fachbezogenen Problemlösekompetenz im Fach Mathematik?
- *Fragestellung 4* (*Ferner Transfer Mathematik*): Steigert dieses Training weniger, zentraler Komponenten der fächerübergreifenden Problemlösekompetenz (konditionales Wissen, Planungsfähigkeit, Handlungswissen) die fachbezogene Problemlösekompetenz im Fach Mathematik insgesamt?
- *Fragestellung 5* (*kompensatorischer Trainingseffekt*): Gemäß der PISA-2003-Befunde ist die Potenzialdifferenz im unteren Leistungsbereich besonders groß (Leutner et al., 2004). Das ungenutzte kognitive Potenzial sollte vor allem im unteren Leistungsbereich vorhanden sein (Leutner et al., 2004). Daher wird untersucht, ob sich *differentielle* Trainingseffekte hinsichtlich der Schulform zeigen. Profitieren also insbesondere leistungsschwächere SuS (Hauptschüler) vom Training (kompensatorischer Trainingseffekt) oder zeigt sich der häufig beobachte Matthäuseffekt („Wer hat, dem wir gegeben", Klauer, 2001)?

3.2.2 Methode

3.2.2.1 Stichprobe

Es nahmen 143 SuS, 79 Jungen (55.2 %), 63 Mädchen (44.1 %), aus Nordrhein-Westfalen am Trainingsexperiment teil. Eine Person machte keine Angabe zu ihrem Geschlecht. 59 SuS besuchten Hauptschulen und 84 SuS Gymnasien. Das Durchschnittsalter beträgt M = 15.04 Jahre (SD = 0.837). 65 % der SuS geben als Familiensprache Deutsch an, 25 % Deutsch und eine andere Sprache, 7 % Deutsch und mehrere andere Sprachen. 3 % machten keine Angaben. Der realisierte Stichprobenumfang liegt zwischen dem optimalen Stichprobenumfang von 76 für den Treatment-Check und 212 für den Transfer.[11]

[11] Die benötigte Stichprobengröße (berechnet mit G*Power 3.1.9 von Faul, Erdfelder, Lang und Buchner (2007), um bei einer 2 (Bedingungen) x 2 (Schulform) MANOVA mit drei abhängigen Variablen (je eine pro trainierter Teilkomponente) mindestens einen mittelgroßen Effekt (beim Treatment-Check) von f^2 = .15 mit großer Sicherheit ($\alpha = \beta = .05$) zu detektieren, ist eine Stichprobengröße von N = 76 Personen erforderlich. Um bei einer 2 (Bedingung) x 2 (Schulform) A-NOVA mit einer abhängigen Variablen (Problemlösen) mindestens einen mittelgroßen Interakti-

Die teilnehmenden Klassen erhielten 50 Euro für ihre Klassenkasse. Die Lehrer erhielten zusätzlich die Möglichkeit, Rückmeldung über die aggregierten Ergebnisse ihrer Klasse zu erhalten. Von dieser Möglichkeit machten alle Schulen Gebrauch.

3.2.2.2 Design

Es wurde ein Proaktionsplan (Klauer, 2011) in Form eines 2 x 2 Experiments mit den beiden Faktoren Treatment (Problemlösetraining in der Experimentalgruppe, GeoGebra-GUI-Tutorial als Kontrolltraining) und Schulform (Hauptschule, Gymnasium) verwendet. Die SuS wurden innerhalb jeder Klasse randomisiert einer Bedingung zugewiesen. Die Treatments werden im Abschnitt *Instrumente und Materialien* beschrieben.

3.2.2.3 Instrumente und Materialien

Die Durchführung der computerbasierten Teile des Experiments (Treatment, Tests zum fächerübergreifenden Problemlösen, demographische Angaben) erfolgte mit der Umfrage-Software EFS Survey (QuestBack AG, 2012). Die Lernmaterialien wurden vor der Haupterhebung in informellen *cognitive lab sessions* (siehe z. B. Zucker, Sassman & Case, 2004) mit vier Probanden überprüft und auf Grundlage der Ergebnisse modifiziert. Eine ausführliche Screenshot-Dokumentation der Materialien ist in Anhang E bis Anhang G zu finden. Daher werden die Materialien im Folgenden nur kurz beschrieben.

Experimentalgruppentreatment

Die Experimentalgruppe bearbeitete während des Treatments eine multimediale Einführung in das Thema Problemlösen (Anhang E), die u. a. die verschiedenen Phasen des schematisierten Problemlöseprozesses (Pólya, 1973) enthielt: 1. Problem lesen und verstehen, 2. Pläne oder Lösungsideen entwickeln, 3. Plan oder Strategie auswählen, 4. Plan durchführen, Problem lösen, 5. Lösung überprüfen.

ons-Effekt von f = .25 mit großer Sicherheit ($\alpha = \beta = .05$) zu detektieren, ist eine Stichprobengröße von N = 212 Personen erforderlich. Empirische Ergebnisse, an denen man die Effektstärken orientieren kann, liegen für das konzipierte Material nicht vor. Große Effektstärken (f^2 = .35) sind angesichts der Kürze des Trainings eher nicht zu erwarten, und um kleine Effektstärken (f^2 = .02) mit akzeptablen Irrtumswahrscheinlichkeiten zu finden, werden Stichprobenumfänge benötigt, die im Rahmen des Projekts nicht realisierbar sind.

Anschließend bearbeiten die SuS Übungsaufgaben zum Sachwissen, Handlungswissen, konditionalem Wissen und Planen (Anhang E). Dabei erhielten die SuS zu den Aufgaben Feedback und im Falle unvollständiger oder falscher Antworten einen zweiten Lösungsversuch. Gelang es den SuS nicht, eine Aufgabe im zweiten Versuch zu lösen, wurde eine Musterlösung präsentiert. Lösten die SuS eine Aufgabe korrekt, so wurde ein lachender Smiley mit einem entsprechenden Hinweistext präsentiert.

Die Aufgaben zum Sachwissen bestehen aus Fragen zu Konzepten und Begrifflichkeiten, deren deklarative Kenntnis für einige der weiteren Aufgaben relevant sind (Anhang E, z. B. „Was versteht man in einer Bibliothek unter überfälligen Ausleihen?“).

Die Aufgaben zum Handlungswissen beziehen sich auf prozedurale Fähigkeiten, die beim Umgang mit diskontinuierlichen Texten benötigt werden, z. B. das Entnehmen von Informationen aus einer Tabelle oder das Lesen einer abstrakten Karte (Anhang E).

Die Aufgaben zum konditionalen Wissen bestehen aus einem kurzen Szenario, d. h. einer Problemsituation, zu der verschiedenen Handlungsoptionen vorgegeben werden, die auf einer sechsstufigen Schulnoten-Skala (von 1 = *sehr gut* bis 6 = *ungenügend*) zu bewerten sind. Derartige, auch als Strategiewissenstests oder als Tests zur Erfassung metakognitiven Wissens bezeichneten Testformate, werden inzwischen in zahlreichen Domänen verwendet – z. B. Lesestrategien (Schlagmüller & Schneider, 2007), Textverstehen (Artelt et al., 2009; Neuenhaus et al., 2010), Deutsch und Englisch (Artelt, Neuenhaus, Lingel & Schneider, 2012), Mathematik (Lingel, Götz, Artelt & Schneider, 2014), Naturwissenschaften (Scherer & Tiemann, 2012; Shahat et al., 2013; Thillmann, 2007).

Die Aufgaben zur Planungsfähigkeit (vgl. Schütte, 2012, Anhang E) bestehen aus einem kurzen Szenario, d. h. einer Problemsituation, dessen vorgegebene Lösungsschritte ungeordnet in Form einer Mindmap präsentiert werden. Die Aufgabe der Probanden besteht darin, die Lösungsschritte in eine sinnvolle (chronologische) Reihenfolge zu bringen. Die eingesetzten Aufgaben zur Planungsfähigkeit ähneln zudem den Abfolgeplanungsaufgaben von Pascha, Schöppe und Hacker (2001), die jedoch rein textbasiert präsentiert und dichotom ausgewertet werden.

Kontrollgruppentreatment

Das Kontrollgruppentreatment besteht aus einer multimedialen Einführung in die Mathematik-Software GeoGebra (Hohenwarter, 2006), die als *iframe* in EFS Survey (QuestBack AG, 2012) eingebunden worden ist (Anhang F). Dabei lernen die SuS die graphische Benutzeroberfläche (GUI) kennen, die zum Umgang mit elementarer Geometrie (Zeichnen von Punkten, Linien, Polyedern) notwendig ist. Die Kontrollgruppe bearbeitete keine mathematischen Problemlöseaufgaben in GeoGebra (siehe Anhang F).

Posttest

Der Posttest besteht sowohl für fächerübergreifendes Problemlösen als auch für mathematisches Problemlösen aus Skalen zu den Komponenten Planungsfähigkeit, konditionales Wissen, Handlungswissen sowie Problemlösen (siehe Anhang G und Anhang H). Die fächerübergreifenden und mathematischen Problemlöseaufgaben stammen aus PISA 2003 (OECD, 2004a, 2004b). Skalenmittelwerte, Streuungen und Reliabilitäten der Tests sind Tabelle 4 zu entnehmen.

Tabelle 4. Reliabilität der Leistungsskalen (N = 143)

Domäne	*Skala*	*Items*	*M*	*SD*	*Cronbachs α*
Fächer-übergreifendes Problemlösen	Planungsfähigkeit	69[a]	45.89	15.07	.959
	Konditionales Wissen	29[a]	14.26	6.15	.861
	Handlungswissen	17	15.13	6.20	.868
	Problemlösen	7	2.72	2.09	.654
Mathematisches Problemlösen	Planungsfähigkeit	54[a]	33.69	16.69	.974
	Konditionales Wissen	17[a]	10.06	3.51	.825
	Handlungswissen	9	4.52	2.43	.797
	Problemlösen	13	5.05	2.78	.731

Anmerkung.
[a] Angegeben ist die Anzahl der Paarvergleiche, die in die Auswertung einfließt (siehe Anhang D für Details).

Drittvariablen

Als Kovariaten und für potenzielle zukünftige Analysen wurden Intelligenz- und Motivationsindikatoren erhoben. Zur Erfassung schlussfolgernden Denkens als Intelligenzindikator wurde die Skala *Figurenanalogien* (*N2*) des KFT 4-12+R (Heller & Perleth, 2000) administriert (M = 14.84, SD = 6.81, Cronbachs α = .92, 25 Items), wobei es sich um korrigierte Testhefte ohne die beiden fehlerhaften

Items (Möller, Bonerad & Pohlmann, 2006; Segerer, Marx & Marx, 2012) handelte.

Zur Erfassung motivationaler Aspekte wurde die Skala *Anstrengungsbereitschaft* des Schülerfragebogens aus PISA 2003 eingesetzt (Boekaerts & Otten, 1993; Deutsches PISA-Konsortium, 2006), die aus drei Fragen (zu Sorgfalt, Konzentration und Mühe) besteht, die auf einer vierstufigen Antwortskala (1 = *trifft zu* bis 4 = *trifft nicht zu*) einzuschätzen sind. Diese Skala (Wertebereich 4 bis 12) wurde vor dem Treatment ($M = 4.28$, $SD = 1.39$, Cronbachs $\alpha = .77$), nach dem Treatment ($M = 5.20$, $SD = 1.98$, Cronbachs $\alpha = .80$) sowie vor dem Posttest ($M = 5.02$, $SD = 1.99$, Cronbachs $\alpha = .87$) eingesetzt. Die SuS berichteten im Mittel also eine hohe Anstrengungsbereitschaft.

3.2.2.4 Durchführung

Die 135-minütige Durchführung des Experiments erfolgte im Computerraum der jeweiligen Schule. Jeder Schüler hatte einen PC-Arbeitsplatz zur Verfügung. Die SuS wurden mittels EFS Survey (QuestBack AG, 2012) randomisiert auf die Experimental- und die Kontrollgruppe verteilt. Zunächst erhielten die SuS allgemeine Instruktionen und einen kurzen demographischen Fragebogen, der Alter, Geschlecht, Muttersprache, Schulnoten sowie Ort und Häufigkeit der Computernutzung (Deutsches PISA-Konsortium, 2006) erfragte sowie drei Fragen zur Anstrengungsbereitschaft (Boekaerts & Otten, 1993) umfasste, die auch bei PISA 2003 zum Einsatz kamen (Deutsches PISA-Konsortium, 2006). Anschließend bearbeiteten die SuS das jeweilige Lernprogramm (Anhang E bzw. Anhang F). Im Anschluss wurden die retrospektiv formulierte Skala zur Anstrengungsbereitschaft, der Posttest und Bewertungsmöglichkeiten des jeweiligen Lernprogramms (siehe Anhang G) administriert. Die Bearbeitungszeiten wurden mittels EFS Survey geloggt. Nach diesem computerbasierten Teil bearbeiteten die SuS die Skala *Figurenanalogien* (*N2*) des KFT 4-12+R (Heller & Perleth, 2000) und die Skalen zum mathematischen Problemlösen als Paper-pencil-Tests. Um Reihenfolge- und Müdigkeitseffekte zu kontrollieren, wurden die mathematischen Skalen in einem lateinischen Quadrat angeordnet. Abbildung 5 visualisiert zusammenfassend den Ablauf des Experiments. Die Durchführung und Dateneingabe erfolgte durch geschulte studentische Hilfskräfte, die nicht über die Ziele und Hypothesen des Experiments informiert waren.

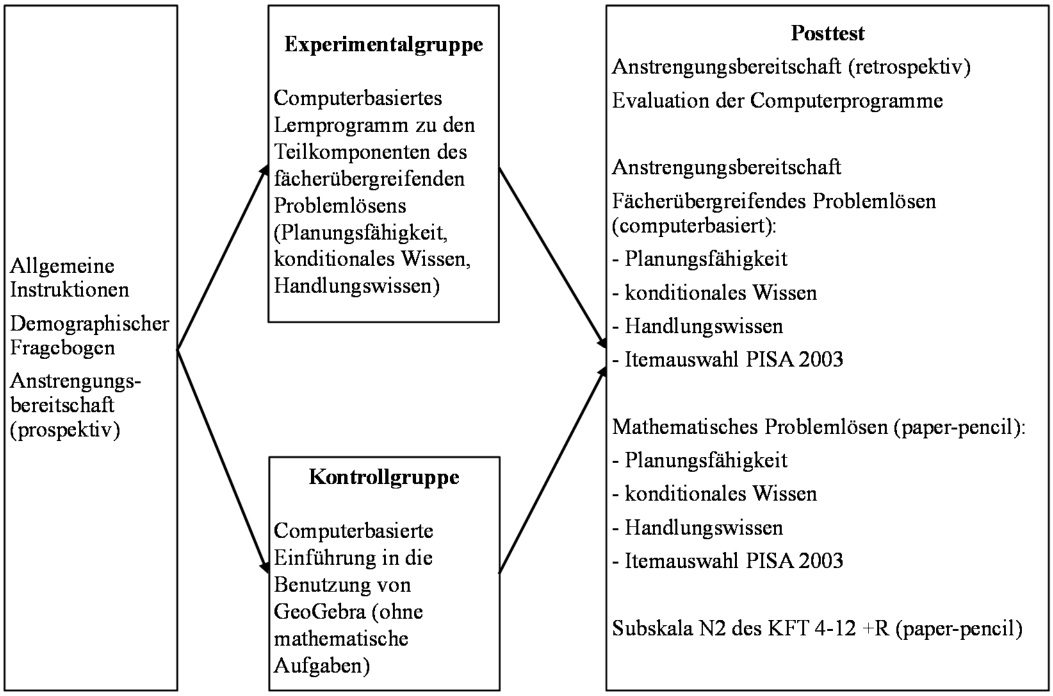


Abbildung 5. Ablaufdiagramm zum Experiment 1

3.2.2.5 Scoring

Die Aufgaben zum Sachwissen, Handlungswissen und die PISA-Items wurden trichotom ausgewertet (*vollständig richtig* = 2 Punkte, *teilweise richtig* = 1 Punkt, 0 = *sonst*). Bei den PISA-Items wurden die Musterlösungen (OECD, 2004b) verwendet. Für die selbst entwickelten Aufgaben zum fächerübergreifenden konditionalen Wissen wurde im Vorfeld des Experimentes mit Hilfe eines Expertenratings geprüft, ob eine Handlungsmöglichkeit H_i besser bewertet wird als eine Handlungsmöglichkeit H_j (vgl. Neuenhaus et al., 2010). Die SuS erhielten je einen Punkt auf der Skala, wenn sie in Übereinstimmung mit dem Expertenrating urteilten, wobei nur die Größer-Gleich-Relation und nicht die Höhe der Benotung eine Rolle spielt (vgl. Anhang *I*). Man beachte, dass die SuS bei Gleichheit der Beurteilung der Antwortoptionen zum fächerübergreifenden konditionalen Wissen keinen halben Punkt erhielten, da die 36 von 143 Personen (22 aus der EG, 14 aus der KG), die bei mindestens einer Aufgabe alle Antwortoptionen gleich bewertet haben (Spaltenkreuzen), von dieser 0.5-Punkte-Scoringregel profitieren würden: Spaltenkreuzen (1 = *ja*, 0 = *nein*) korreliert signifikant (r = .258, p = .002) mit dem Effizienzmaß, das 0.5 Punkte bei Gleichheit vorsieht, nicht jedoch mit dem Effizienzmaß ohne diese 0.5-Punkte-Scoringregel (r = -.112, p = .181). Dieses Spaltenkreuzen ist jedoch kein sicherer Indikator für allgemeine Verweigerung (vgl. Anhang *I*). Beim *mathematischen konditionalen Wissen* wurden drei Szenarien verwendet (Deutsches PISA-Konsortium, 2006), die inzwischen in einer weiter entwickelten Version publiziert sind (Lingel et al., 2014).

Pro *Planungsaufgabe* wurde im Vorfeld des Experimentes mit Hilfe eines Expertenratings geprüft, ob je zwei Teilschritte in eine chronologische Abfolge zu bringen sind. Die unter den Experten übereinstimmenden Paarvergleiche wurden zum Scoring verwendet: Falls Teilschritt A vor Teilschritt B auszuführen ist, erhielten die SuS einen Punkt, wenn sie Teilschritt A vor Teilschritt B angaben. Dabei spielte es keine Rolle, ob Teilschritt B unmittelbar oder mittelbar nach Teilschritt A angegeben wurde.

3.2.2.6 Datenanalyse

Der Reihenfolge der Hypothesen folgend werden zunächst die Ergebnisse zum fächerübergreifenden Problemlösen und dann zum mathematischen Problemlösen berichtet. Zur Analyse werden zweifaktorielle Kovarianzanalysen mit Bedingung (Experimentalgruppe vs. Kontrollgruppe) und Schulform (Hauptschule vs. Gymnasium) als unabhängige Variablen, Effizienz- bzw. Leistungsvariablen als abhängige Variablen und der Skala *Figurenanalogien* (N2) des KFT 4-12+R (Heller & Perleth, 2000) als Kovariate berechnet. Tabellen mit Mittelwerten, Standardabweichungen und Fallzahlen sind Anhang A zu entnehmen.

Die Bearbeitungszeit der Tests zum fächerübergreifenden Problemlösen unterscheidet sich zwischen den experimentellen Bedingungen. Dieses Problem konnte auftreten, da das Computerprogramm *self-paced* konzipiert war, d. h. es gab einen fließenden Übergang zwischen Lern- und Testphase (Blanchard & Thacker, 2010). Die Kontrollgruppe verwendete, obwohl bei vorhergehenden Zeitpilotierungen die Bearbeitungszeit in beiden Gruppen vergleichbar war, weniger Zeit mit dem Treatment und mehr Zeit auf den Posttest als die Experimentalgruppe (Abbildung 6). Die unterschiedliche Bearbeitungszeit führt dazu, dass die Kontrollgruppe beim fächerübergreifenden Problemlösen bessere Leistungen erzielt als die Experimentalgruppe (Abbildung 7). Um die unterschiedliche Bearbeitungszeit bei der Auswertung zu berücksichtigen, werden im Ergebnisteil für die Skalen zum fächerübergreifenden Problemlösen *Effizienzmaße* betrachtet, die die pro Skala erreichte Rohpunktsumme an der Bearbeitungszeit relativieren (Effizienz = $\frac{\text{erreichte Punktzahl}}{\text{benötigte Zeit [min]}}$; für andere Effizienzmaße siehe z. B. Paas & van

Merriënboer, 1993; Stevens & Thadani, 2007). Die Skalen zum mathematischen Problemlösen wurden mit gleicher Bearbeitungszeit für alle SuS in Papierform bearbeitet. Daher können für das mathematische Problemlösen Leistungsskalen analysiert werden. Die Leistungs- und Effizienzskalen wurden z-standardisiert ($M = 0$, $SD = 1$).

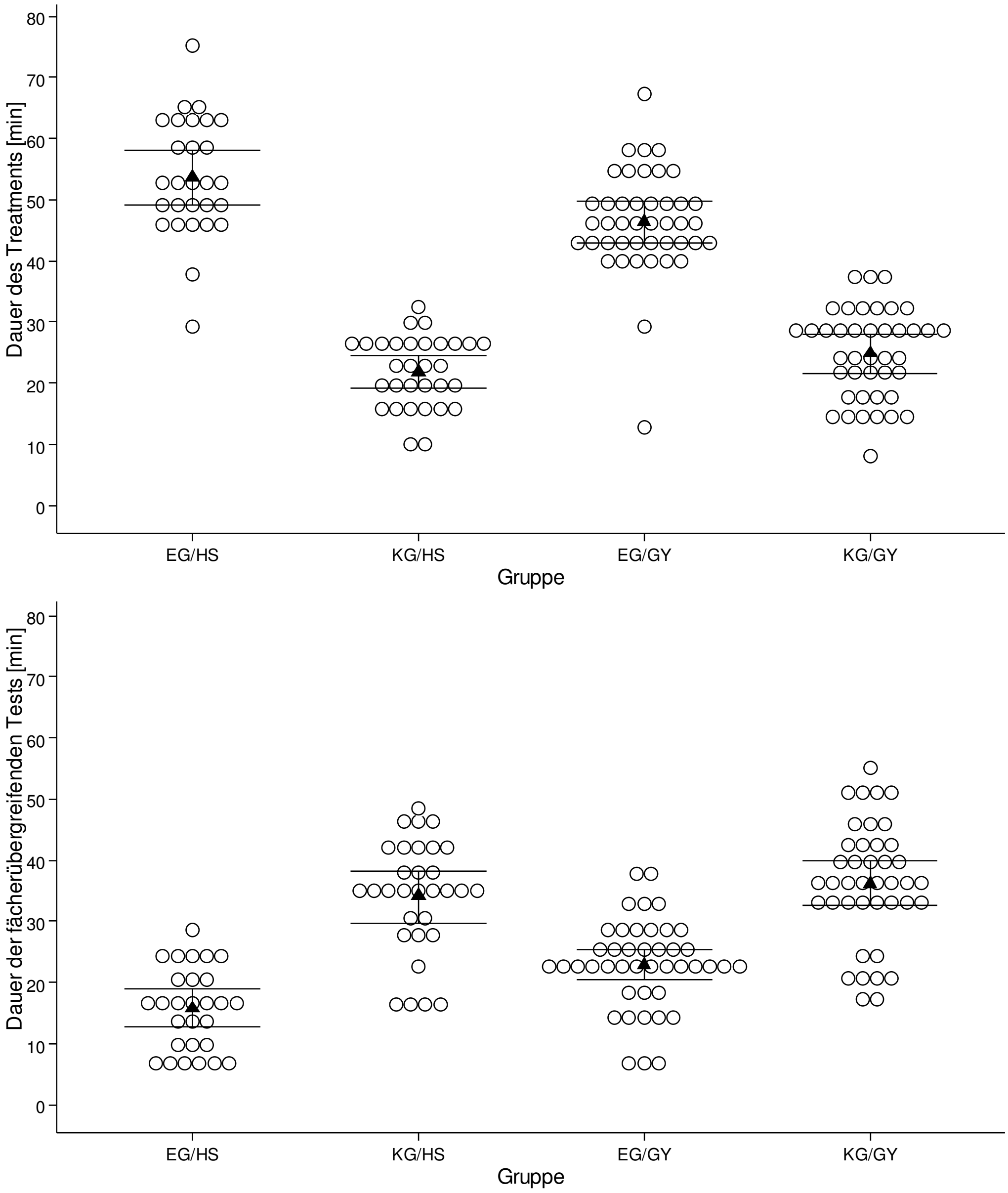


Abbildung 6. Dauer der Treatments und Tests zum fächerübergreifenden Problemlösen für Experimentalgruppe (EG) und Kontrollgruppe (KG) getrennt nach Schulform (HS = Hauptschule, GY = Gymnasium)

Anmerkung. Das schwarze Dreieck kennzeichnet den Mittelwert. Fehlerbalken repräsentieren das 99%-Bootstrap-Konfidenzintervall des Mittelwerts (mit 1000 Replikationen). Die Abbildung wurde mit dem R-Paket *ggplot 2* (Wickham, 2009) erzeugt.

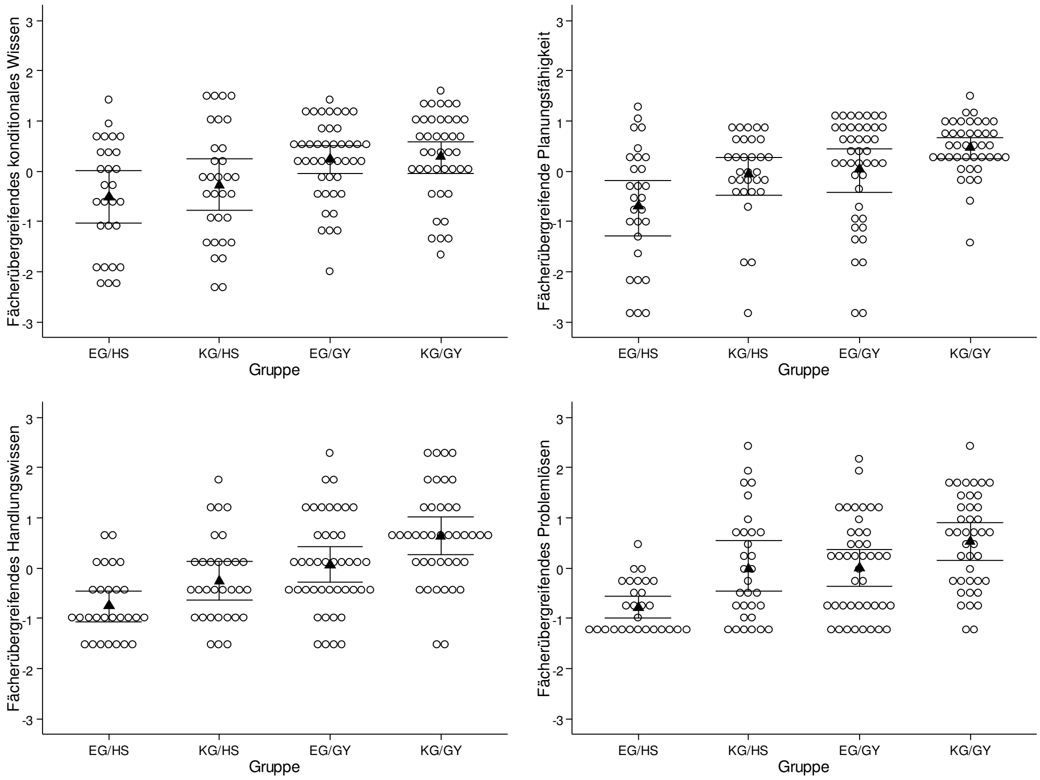


Abbildung 7. Leistung der fächerübergreifenden Skalen (z-standardisiert) im Posttest für Experimentalgruppe (EG) und Kontrollgruppe (KG) getrennt nach Schulform (HS = Hauptschule, GY = Gymnasium)

Anmerkung. Das schwarze Dreieck kennzeichnet den Mittelwert. Fehlerbalken repräsentieren das 99%-Bootstrap-Konfidenzintervall des Mittelwerts (mit 1000 Replikationen). Die Abbildung wurde mit dem R-Paket *ggplot 2* (Wickham, 2009) erzeugt.

3.2.3 Ergebnisse

3.2.3.1 Treatment-Check: Effizienz bei den Teilkomponenten fächerübergreifenden Problemlösens

Aufgrund des Bearbeitungszeitproblems (siehe Methode) wurde Fragestellung 1 folgendermaßen reformuliert: Arbeitet die Experimentalgruppe bei den Tests der Komponenten der fächerübergreifenden Problemlösekompetenz (Planungsfähigkeit, konditionales Wissen, Handlungswissen) *effizienter* als die Kontrollgruppe (Fragestellung 1*)?

Die Experimentalgruppe arbeitete deskriptiv im Mittel bei den Skalen Planungsfähigkeit und konditionales Wissen *effizienter* als die Kontrollgruppe. Beim Handlungswissen war deskriptiv kein Unterschied der experimentellen Bedingungen zu erkennen (Abbildung 8, Anhang I).

Um Fragestellung 1* zu beantworten, wurde eine 2 (Treatment: EG vs. KG) x 2 (Schulform: Hauptschule vs. Gymnasium) MANCOVA mit den z-standardisierten *Effizienzmaßen* der Skalen zum konditionalen Wissen, zur Planungsfähigkeit und zum Handlungswissen als abhängigen Variablen und der z-standardisierten Skala *Figurenanalogien* (*N2*) des KFT 4-12+R (Heller & Perleth, 2000) als Kovariate berechnet.

Die multivariaten Ergebnisse zeigten, dass die Kovariate keinen signifikanten Effekt hatte (Wilks $\Lambda = .978$, $F(3, 135) = 1.026$, p = .383, $\eta^2 = .022$).[12] Die multivariate Interaktion von Schulform und Bedingung war nicht signifikant (Wilks $\Lambda = .982$, $F < 1$). Die beiden multivariaten Haupteffekte waren signifikant: Die Schulform hatte einen signifikanten multivariaten Effekt (Wilks $\Lambda = .936$, $F(3, 135) = 3.070$, $p = .030$, $\eta^2 = .064$), ebenso die Bedingung (Wilks $\Lambda = .785$, $F(3, 135) = 12.334$, $p < .001$, $\eta^2 = .215$). Follow-up-ANCOVAs (Tabelle 5) zeigten, dass bei den Handlungswissensaufgaben Gymnasiasten effizienter arbeiteten als Hauptschüler ($\eta_p^2 = .048$). Bei den Aufgaben zum konditionalen Wissen war der Schulformunterschied einseitig signifikant ($\eta_p^2 = .020$). Bei den Planungsaufgaben war der Schulformunterschied hingegen nicht signifikant ($\eta_p^2 = .011$). Die Experimentalgruppe arbeitete bei den Aufgaben zur Planungsfähigkeit ($\eta_p^2 = .073$) und zum konditionalen Wissen ($\eta_p^2 = .200$) signifikant effizienter als die Kontrollgruppe. Bei den Aufgaben zum Handlungswissen war der Unterschied zwischen Experimental- und Kontrollgruppe einseitig signifikant ($\eta_p^2 = .020$).

Zwischenfazit: Das Treatment kann somit hinsichtlich der *Effizienz der Aufgabenbearbeitung* in den trainierten Komponenten des Problemlösens als erfolgreich angesehen werden, wobei sich die Effektstärken zwischen den Skalen deutlich unterscheiden.

[12] Berichtet wird Wilks Λ. Man beachte, dass die Ergebnisse für alle von SPSS 20 berichteten Teststatistiken gleich ausfallen.

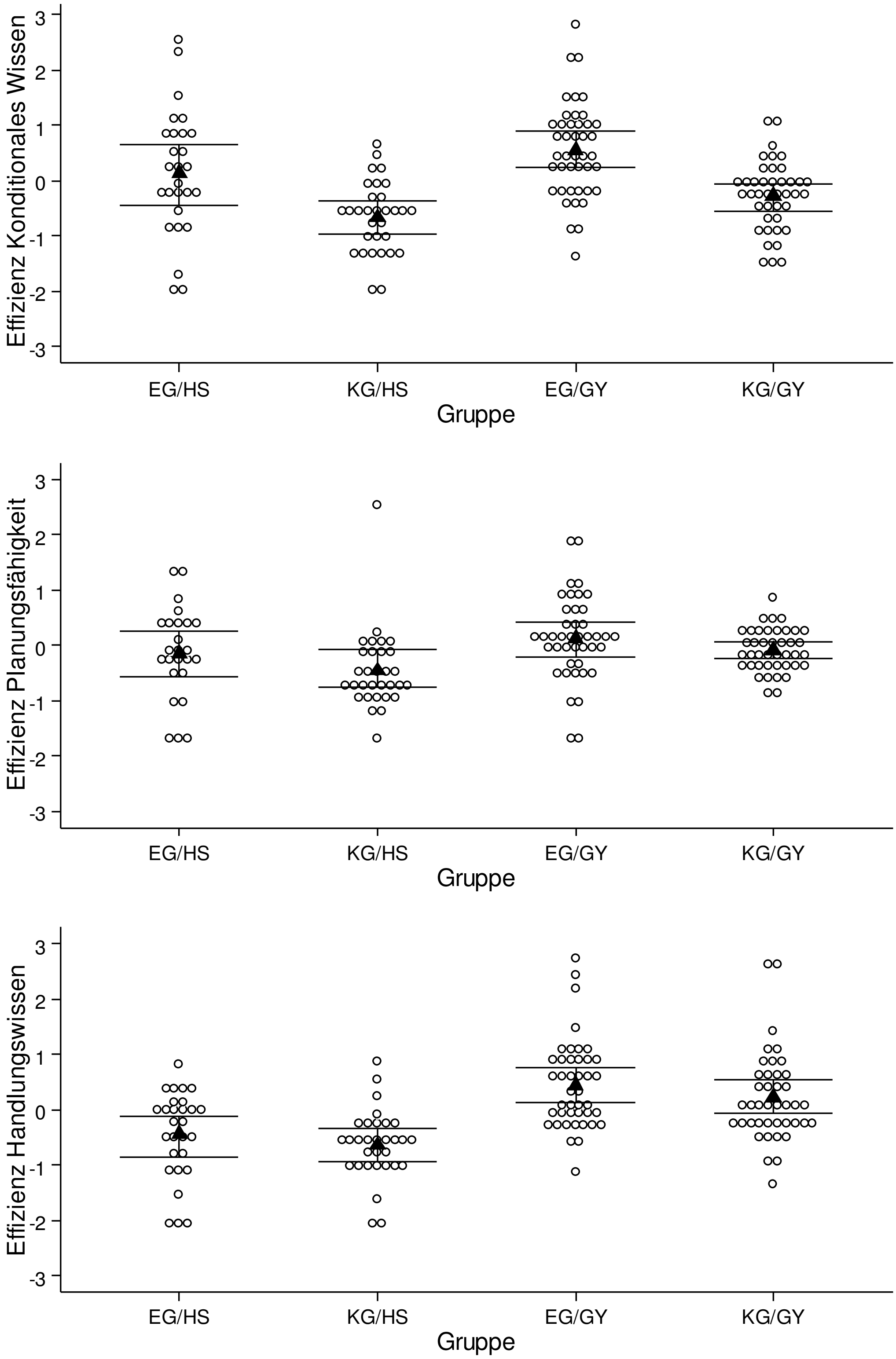


Abbildung 8. Effizienz der drei Komponenten fächerübergreifenden Problemlösens (z-standardisiert) für Experimentalgruppe (EG) und Kontrollgruppe (KG) getrennt nach Schulform (HS = Hauptschule, GY = Gymnasium)

Anmerkung. Das schwarze Dreieck kennzeichnet den Mittelwert. Fehlerbalken repräsentieren das 99%-Bootstrap-Konfidenzintervall des Mittelwerts (mit 1000 Replikationen). Die Abbildung wurde mit dem R-Paket *ggplot 2* (Wickham, 2009) erzeugt.

Tabelle 5. Ergebnisse der follow-up ANCOVAs zu den drei Komponenten Handlungswissen (HW), Planungsfähigkeit (PLA) und konditionales Wissen (KW)

Quelle der Variation	*Abhängig Variable*	*SS*	*df*	*MS*	*F*	*Sig.*	η_p^2
KFT, Skala N2	PLA[a]	0.827	1	0.827	0.877	.351	.006
	KW[b]	0.045	1	0.045	0.056	.813	.000
	HW[c]	1.218	1	1.218	1.559	.214	.011
Bedingung	PLA	1.226	1	1.226	1.839	.001	.073
	KW	27.358	1	27.358	34.223	.000	.200
	HW	2.218	1	2.218	2.838	.094	.020
Schulform	PLA	1.467	1	1.467	1.555	.215	.011
	KW	2.286	1	2.286	2.860	.093	.020
	HW	9.992	1	9.992	12.784	.000	.085
Bedingung x Schulform	PLA	2.133	1	2.133	2.261	.135	.016
	KW	0.157	1	0.157	0.197	.658	.001
	HW	0.251	1	0.251	0.321	.572	.002
Fehler	PLA	129.249	137	0.943			
	KW	109.517	137	0.799			
	HW	107.079	137	0.782			
Gesamt	PLA	141.529	142				
	KW	141.401	142				
	HW	14.813	142				

Anmerkungen.
[a] R^2 = .240 (korrigiertes R^2 = .217). [b] R^2 = .087 (korrigiertes R^2 = .060).
[c] R^2 = .225 (korrigiertes R^2 = .203).

3.2.3.2 *Ferner Transfer: Effizienz beim Problemlösen*

Steigert ein Training weniger, zentraler Komponenten der fächerübergreifenden Problemlösekompetenz die fächerübergreifende Problemlösekompetenz? Präziser formuliert: Schneidet die Experimentalgruppe *effizienter* beim Test der fächerübergreifenden Problemlösekompetenz (bestehend aus PISA 2003 Items) ab als die Kontrollgruppe (Fragestellung 2*)?

Die Mittelwerte der Effizienz der Experimentalgruppe beim fächerübergreifenden Problemlösen lagen für beide Schulformen (Hauptschule und Gymnasium) deskriptiv leicht über denen der Kontrollgruppe (Abbildung 9).

Um Fragestellung 2* zu prüfen, wurde eine ANCOVA berechnet mit der Bedingung (Experimentalgruppe vs. Kontrollgruppe) und Schulform (Gymnasium vs. Hauptschule) als unabhängigen Variablen und dem z-standardisierten *Effizienzmaß* der Skala zum fächerübergreifenden Problemlösen aus PISA 2003 als abhängiger Variable sowie der z-standardisierten Skala *Figurenanalogien* (*N2*) des KFT

4-12+R (Heller & Perleth, 2000) als Kovariate. Der Effekt der Kovariate war nicht signifikant ($F < 1$). Die Interaktion von Bedingung und Schulform hatte keinen signifikanten Effekt ($F < 1$). Hauptschüler und Gymnasiasten unterschieden sich nicht in der Effizienz ($F < 1$). Die Experimentalgruppe ($M = 0.16$, $SD = 1.31$) schnitt signifikant effizienter als die Kontrollgruppe ($M = -0.16$, $SD = 0.52$) ab ($F(1, 137) = 3.805$, $p = .03$, $\eta_p^2 = .027$, gerichtete Testung).

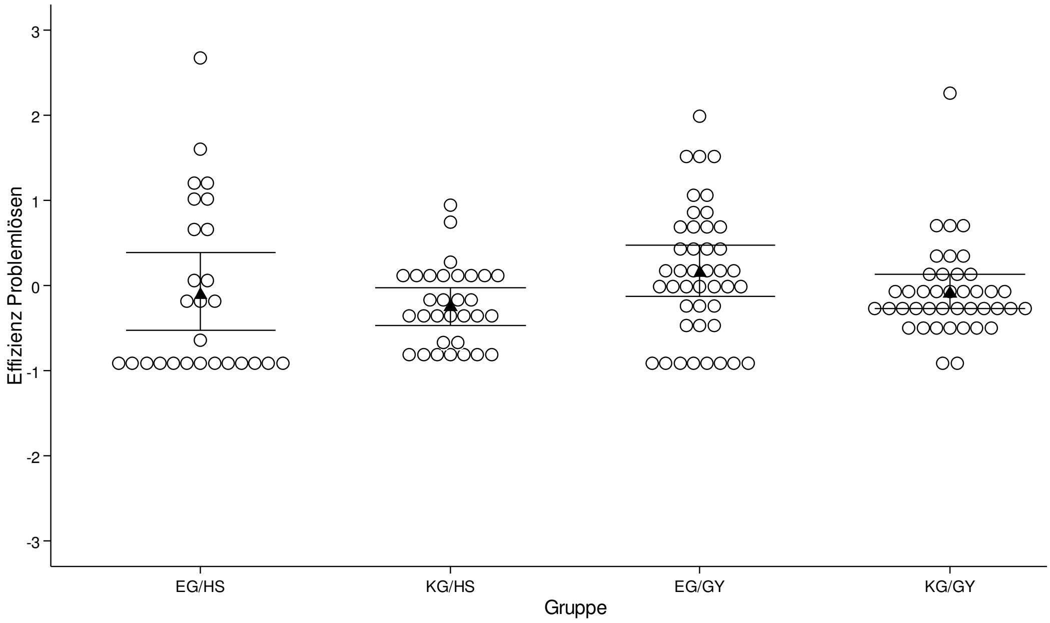


Abbildung 9. Effizienz des fächerübergreifenden Problemlösens im Posttest (z-standardisiert) für Experimentalgruppe (EG) und Kontrollgruppe (KG) getrennt nach Schulform (HS = Hauptschule, GY = Gymnasium)

Anmerkung. Das schwarze Dreieck kennzeichnet den Mittelwert. Fehlerbalken repräsentieren das 99%-Bootstrap-Konfidenzintervall des Mittelwerts (mit 1000 Replikationen). Die Abbildung wurde mit dem R-Paket *ggplot 2* (Wickham, 2009) erzeugt.

3.2.3.3 Naher Transfer Mathematik: Komponenten mathematischer Problemlösekompetenz

Steigert ein Training weniger, zentraler Komponenten der fächerübergreifenden Problemlösekompetenz ebenfalls diese Komponenten der fachbezogenen Problemlösekompetenz in der Mathematik? Präziser formuliert: Schneidet die Experimentalgruppe besser bei den Tests der Komponenten mathematischer Problemlösekompetenz ab als die Kontrollgruppe (Fragestellung 3)?

Die Deskriptivstatistik zum mathematischen Problemlösen (Anhang I) zeigte, dass die Gruppenmittelwerte und Streuungen in der Experimental- und der Kontrollgruppe über alle vier Skalen sehr ähnlich waren. Es zeigten sich Schulformeffekte zugunsten des Gymnasiums beim mathematischen Handlungswissen (Abbildung 10). Ferner zeigten die Aufgaben zur mathematischen Planungsfähigkeit bei Gymnasiasten leichte Deckeneffekte (Abbildung 10).

Um Fragestellung 3 zu prüfen, wurde eine 2 (Treatment: EG vs. KG) x 2 (Schulform: Hauptschule vs. Gymnasium) MANCOVA mit den z-standardisierten mathematischen Skalen zum konditionalen Wissen, zur Planungsfähigkeit und zum Handlungswissen als abhängige Variablen und der z-standardisierten Skala *Figurenanalogien* (*N2*) des KFT 4-12+R (Heller & Perleth, 2000) als Kovariate berechnet. Die Kovariate KFT hatte einen multivariaten Effekt (Wilks $\Lambda = .860$, $F(3,\ 135) = 7.315$, $p\ <\ .001$, $\eta^2\ = .140$). Die Interaktion von Bedingung und Schulform war nicht signifikant (Wilks $\Lambda = .984$, $F < 1$), ebenso der Effekt der Bedingung (Wilks $\Lambda = .997$, $F < 1$). Die Schulform hatte einen signifikanten Effekt (Wilks $\Lambda = .864$, $F(3,\ 135) = 7.106$, $p < .001$, $\eta^2 = .136$). Follow-up-Analysen der signifikanten multivariaten Effekte zeigten, dass die Gymnasiasten bei den Skalen zum konditionalen Wissen ($F(1,\ 137) = 5.364$, $p = .014$, $\eta^2 = .043$) und zum Handlungswissen ($F(1,\ 137) = 8.613$, $p < .001$, $\eta^2 = .100$) besser abschnitten als die Hauptschüler. Bei den Planungsaufgaben gab es keine Schulformunterschiede ($F < 1$).

Zwischenfazit: Bei den Teilkomponenten mathematischen Problemlösens zeigten sich keine Effekte nahen Transfers. Bei zwei Skalen (konditionales Wissen und Handlungswissen) gab es Schulformeffekte zugunsten des Gymnasiums.

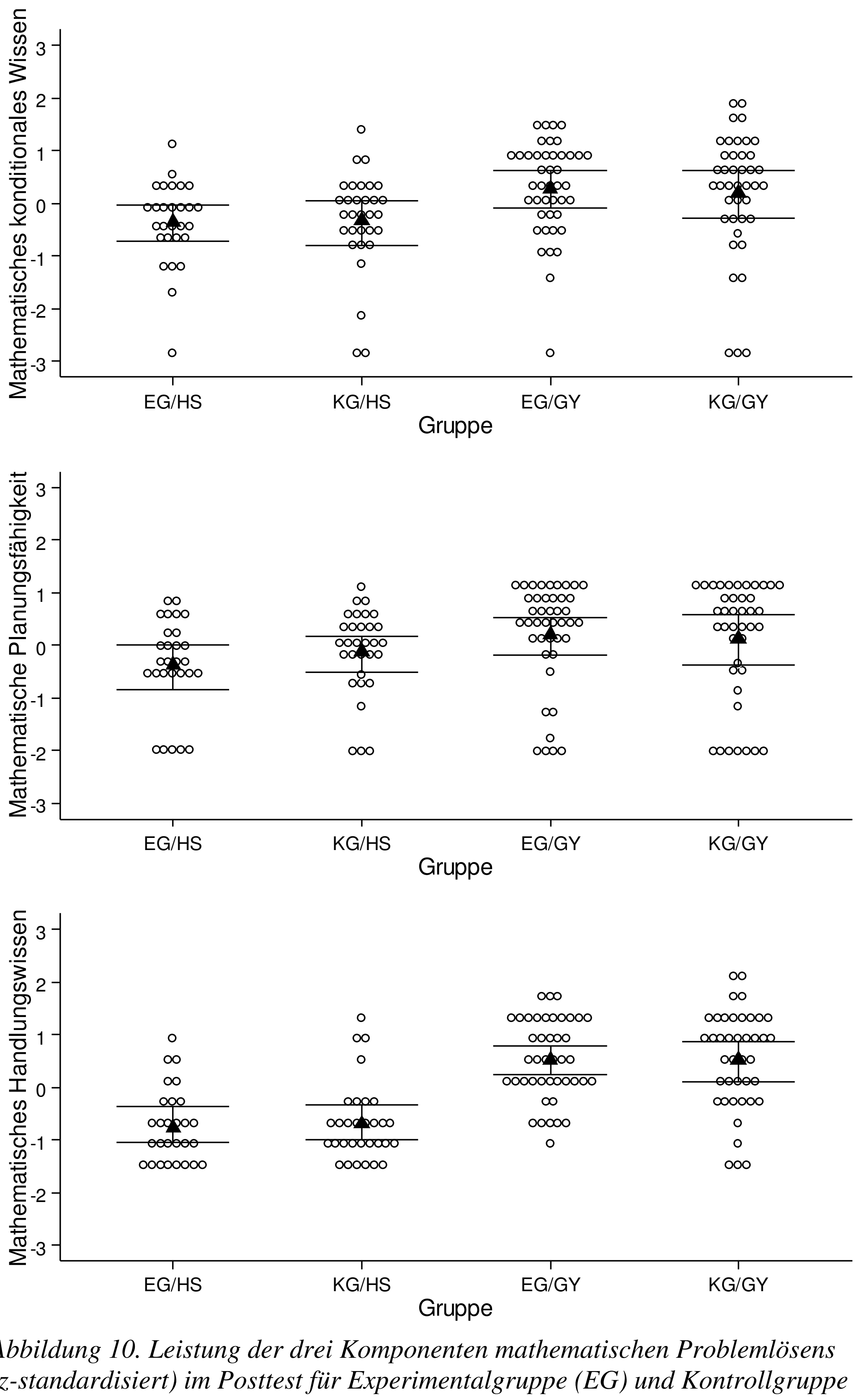


Abbildung 10. Leistung der drei Komponenten mathematischen Problemlösens (z-standardisiert) im Posttest für Experimentalgruppe (EG) und Kontrollgruppe (KG) getrennt nach Schulform (HS = Hauptschule, GY = Gymnasium)

Anmerkung. Das schwarze Dreieck kennzeichnet den Mittelwert. Fehlerbalken repräsentieren das 99%-Bootstrap-Konfidenzintervall des Mittelwerts (mit 1000 Replikationen). Die Abbildung wurde mit dem R-Paket *ggplot 2* (Wickham, 2009) erzeugt.

3.2.3.4 Ferner Transfer Mathematik: Mathematische Problemlösekompetenz

Steigert ein Training weniger, zentraler Komponenten der fächerübergreifenden Problemlösekompetenz die fachbezogene Problemlösekompetenz in der Mathematik? Präziser formuliert: Schneidet die Experimentalgruppe besser beim Test der mathematischen Problemlösekompetenz (bestehend aus PISA 2003 Items) ab als die Kontrollgruppe (Fragestellung 4)?

Deskriptiv zeigten sich beim mathematischen Problemlösen deutliche Schulformeffekte zugunsten des Gymnasiums und im Mittel keine Unterschiede zwischen den experimentellen Bedingungen (Abbildung 11).

Um Fragestellung 4 zu prüfen, wurde eine ANCOVA berechnet mit der Bedingung (Experimentalgruppe vs. Kontrollgruppe) und der Schulform (Gymnasium vs. Hauptschule) als unabhängigen Variablen und der z-standardisierten Skala zum mathematischen Problemlösen aus PISA 2003 als abhängiger Variable sowie der z-standardisierten Skala *Figurenanalogien* (*N2*) des KFT 4-12+R (Heller & Perleth, 2000) als Kovariate. Der Effekt der Kovariate KFT war signifikant ($F(1, 137) = 11.095$, $p < .001$, $\eta^2 = .149$). Die Interaktion von Bedingung und Schulform war nicht signifikant ($F < 1$). Experimentalgruppe ($M = -0.00$, $SD = 0.97$) und Kontrollgruppe ($M = -0.01$, $SD = 1.03$) unterschieden sich nicht signifikant ($F < 1$). Gymnasiasten ($M = 0.57$, $SD = 0.79$) schnitten signifikant besser ab ($F(1, 137) = 24.458$, $p < .001$, $\eta^2 = .172$) als Hauptschüler ($M = -0.82$, $SD = 0.64$).

Zwischenfazit: Transfer auf mathematisches Problemlösen ließ sich in Experiment 1 nicht nachweisen. Es zeigten sich wie beim mathematischen Handlungswissen auch beim mathematischen Problemlösen Schulformeffekte zugunsten des Gymnasiums.

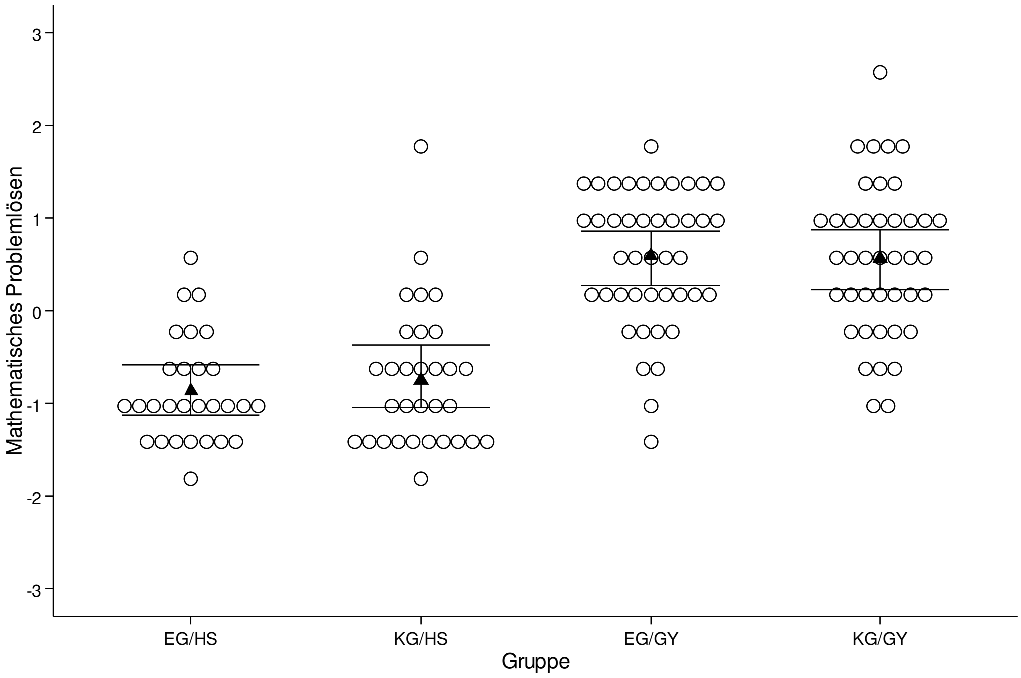


Abbildung 11. Leistung mathematischen Problemlösens (z-standardisiert) im Posttest für Experimentalgruppe (EG) und Kontrollgruppe (KG) getrennt nach Schulform (HS = Hauptschule, GY = Gymnasium)

Anmerkung. Das schwarze Dreieck kennzeichnet den Mittelwert. Fehlerbalken repräsentieren das 99%-Bootstrap-Konfidenzintervall des Mittelwerts (mit 1000 Replikationen). Die Abbildung wurde mit dem R-Paket *ggplot 2* (Wickham, 2009) erzeugt.

3.2.3.5 Weitere Befunde

Die vor dem Treatment eingesetzte Skala *Anstrengungsbereitschaft* (M = 4.28, SD = 1.39, Cronbachs α = .765, 3 Items) war rechtsschief verteilt. Die vierte Kategorie (*trifft nicht zu*) wurde nicht verwendet. Die selbst berichtete Anstrengungsbereitschaft der SuS konnte somit als erfreulich hoch angesehen werden, wobei Effekte der sozialen Erwünschtheit (Bortz & Döring, 2006) nicht auszuschließen sind, da negative Äußerungen zur Anstrengungsbereitschaft in der Schule üblicherweise sanktioniert werden. Zusätzliche Kontrollanalysen zu den Fragestellungen mit der Skala *Anstrengungsbereitschaft* als Kovariate ergaben keine Effekte für die Anstrengungsbereitschaft.

Ein Blick auf die Interkorrelation der acht Skalen (Tabelle 6) zeigte, dass die Skala *Problemlösen* und die neu konstruierten Skalen der Teilkomponenten des Problemlösens (Planungsfähigkeit, konditionales Wissen und Handlungswissen) pro

Domäne positiv miteinander korrelieren, wobei sich die Stärke der Korrelationen deutlich unterschied.

Tabelle 6. Interkorrelation der Skalen (untere Diagonalmatrix) und Partialkorrelation der Skalen bei Berücksichtung der Skala Figurenanalogien (N2) des KFT-4-12+R (obere Diagonalmatrix) in Experiment 1

			Fächerübergreifende Skalen (Effizienz)				*Mathematische Skalen (Leistung)*			
			PLA	KW	HW	PISA	PLA	KW	HW	PISA
Fächerübergreifende Skalen (Effizienz)	Planungs-fähigkeit (PLA)	*r*		.380	.163	.547	.054	.028	.110	.061
		p		.000	.027	.000	.263	.373	.098	.238
	Konditionales Wissen (KW)	*r*	.381		.217	-.028	.050	.199	.182	-.028
		p	.000		.005	.371	.280	.009	.015	.371
	Handlungs-Wissen (HW)	*r*	.162	.252		.130	.004	.203	.290	.130
		p	.027	.001		.063	.479	.008	.000	.063
	Problemlösen (PISA)	*r*	.548	.148	.130		.055	-.026	.000	.155
		p	.000	.039	.061		.260	.378	.499	.033
Mathematische Skalen (Leistung)	Planungs-fähigkeit	*r*	.046	.000	.210	.060		.128	.334	-.011
		p	.293	.498	.006	.239		.066	.000	.448
	Konditionales Wissen	*r*	.037	.094	.140	-.005	.198		.156	.171
		p	.332	.132	.047	.475	.009		.033	.021
	Handlungs-Wissen	*r*	.101	.237	.390	.033	.416	.328		.300
		p	.115	.002	.000	.349	.000	.000		.000
	Problemlösen	*r*	.062	.224	.462	.151	.173	.351	.593	
		p	.230	.004	.000	.036	.019	.000	.000	

Anmerkungen. Produkt-Moment-Korrelation, einseitige Signifikanztestung, $N = 143$. Signifikante Korrelationen ($p < .05$) sind grau hinterlegt. *P*-Werte sind auf drei Nachkommastellen gerundet. Anhang J enthält eine erweiterte Version dieser Tabelle, die auch die Korrelationen der Bearbeitungszeit der fächerübergreifenden Skalen und die Leistungsmaße der fächerübergreifenden Skalen enthält.

3.2.4 Diskussion

Das Problemlöse-Treatment kann hinsichtlich der *Effizienz der Aufgabenbearbeitung* in den trainierten Komponenten des Problemlösens als erfolgreich angesehen werden, wobei sich die Effektstärken zwischen den Skalen deutlich unterscheiden (konditionales Wissen: $\eta_p^2 = .200$, Planungsfähigkeit: $\eta_p^2 = .073$, Handlungswissen: $\eta_p^2 = .020$). Auch bei den PISA-Aufgaben zum fächerübergreifenden Problemlösen (OECD, 2004b) arbeitet die Experimentalgruppe *effizienter* als die Kontrollgruppe ($\eta_p^2 = .028$), wobei die Lösungsquote bzw. Effizienz in beiden Gruppen eher niedrig ist. Es finden sich keine nahen und fernen Transfereffekte auf den mathematischen Bereich, jedoch erwartungsgemäß Schulformunterschiede und ein Effekt der kognitiven Grundfertigkeit. Die fächerübergreifenden und mathematischen Skalen zu den Teilkomponenten des Problemlösens korrelieren pro Domäne alle positiv miteinander. Das kann als Validitätsbeleg der teilweise neu konstruierten Skalen interpretiert werden. Auffällig ist beim Vergleich zwischen den Domänen jedoch, dass die Effizienz beim fächerübergreifenden Problemlösen mit allen mathematischen Leistungsskalen unkorreliert ist. Da es sich beim fächerübergreifenden Problemlösen um Effizienzskalen und bei den mathematischen Skalen um Leistungsmaße handelt, sollten die Ergebnisse zurückhaltend interpretiert werden. Die allgemein niedrige Effizienz bei den PISA-Aufgaben zum fächerübergreifenden Problemlösen, die möglicherweise dem Umfang der Aufgaben, mit der Position der PISA-Aufgaben am Ende der Computersession, oder mit der geringen Lösungsquote erklärbar sind.

Durch das Ausbleiben des Transfers auf mathematische Skalen, liefert Experiment 1 keinen Beleg für die Gültigkeit der kognitiven Potenzialausschöpfungshypothese von Leutner et al. (2004). Aufgrund der unerwartet ungleich ausgefallenen Treatment-Zeit in Experimental- und Kontrollgruppe, die dazu führte, dass die Kontrollgruppe mehr Zeit mit dem Posttest verbrachte, ist eine Betrachtung der Leistungsmaße für die Tests zum fächerübergreifenden Problemlösen nicht sinnvoll. Das Effizienzmaß scheint in Kombination mit geschlossenen Aufgabenformaten zudem anfällig für Rateeffeke (blindes Durchklicken) zu sein. Experiment 1 daher in modifizierter Form repliziert (Kapitel 3.3). Dabei wird unter anderem das *self-pacing* zwischen Treatment und Posttest durch starre Zeitraster ersetzt. Auf diese Weise können wieder klassische Leistungsmaße ohne Relativierung an der Bearbeitungszeit verwendet werden.

3.3 Experiment 2

Experiment 2 ist eine modifizierte Replikation des Experiments 1. Daher werden vor allem die Änderungen gegenüber Experiment 1 beschrieben. Um mehr Lernzeit zu erreichen, wurde das experimentelle Design so geändert, dass pro Experimentalgruppe jeweils zwei der drei Komponenten fächerübergreifenden Problemlösens (Planungsfähigkeit, konditionales Wissen, Handlungswissen) aus Experiment 1 trainiert werden. Zusätzlich wurde das *self-pacing* zwischen Treatment und Posttest zum fächerübergreifenden Problemlösens aufgegeben, um gleiche Testzeiten zwischen den Gruppen zu gewährleisten.

3.3.1 Fragestellungen

Zur Herleitung der Fragestellungen siehe Experiment 1 (Kapitel 3.2.1). Da in Experiment 2 ein modifiziertes Design verwendet wird, bei dem jede Experimentalgruppe nur zwei der drei Komponenten fächerübergreifenden Problemlösens aus Experiment 1 (Planungsfähigkeit, konditionales Wissen, Handlungswissen) trainiert, werden die Fragestellungen nachfolgend entsprechend präziert.

Fragestellung 1 (*Treatment-Check*): Lassen sich zentrale Komponenten der fächerübergreifenden Problemlösekompetenz in lernpsychologischen Laborexperimenten messbar steigern? Präziser formuliert werden Trainingseffekte in den beiden jeweils trainierten Komponenten erwartet (Fragestellung 1, vgl. Leutner et al., 2009): Schneiden die Experimentalgruppen *besser* bei den Tests der Komponenten der fächerübergreifenden Problemlösekompetenz ab, zu denen sie das Lernprogramm absolviert haben, als die Kontrollgruppen? Als Kontrollgruppen dienen jeweils die Gruppe, die jeweilige Komponente nicht trainiert hat sowie die Gruppe, die wie in Experiment 1 das GeoGebra-GUI-Tutorial absolvierte.

Fragestellung 2 (*Transfer Problemlösen*): Steigert das Training weniger, zentraler Komponenten der fächerübergreifenden Problemlösekompetenz die fächerübergreifende Problemlösekompetenz insgesamt (Fragestellung 2a, vgl. Leutner et al., 2009)? Lassen sich die Ergebnisse von Experiment 1 zur Effizienz des fächerübergreifenden Problemlösens in Experiment 2 mit Leistungsmaßen replizieren? Präziser: Schneiden die drei Experimentalgruppen *besser* beim Test der fächerübergreifenden Problemlösekompetenz (bestehend aus PISA 2003 Items) ab als die Kontrollgruppe? Unterscheiden sich die drei Experimentalgruppen hinsichtlich des Transfers, d. h. ist eine bestimmte Kombination zweier trainierter Komponenten am besten geeignet, Transfer beim Problemlösen zu erzielen (Fragestellung 2b)?

Fragestellung 3 (*Naher Transfer Mathematik*): Steigert ein Training zentraler Komponenten der fächerübergreifenden Problemlösekompetenz ebenfalls diese Komponenten der fachbezogenen Problemlösekompetenz im Fach Mathematik (Fragestellung 3a)? Präziser formuliert: Schneiden die Experimentalgruppen *besser* ab bei den Tests der Komponenten der mathematischen Problemlösekompetenz, zu denen Sie das Lernprogramm absolviert haben, als die Kontrollgruppen? Kontrollgruppen sind jeweils die Gruppe, die zwei andere Komponenten trainiert hat, und die GeoGebra-Gruppe. Unterscheiden sich die drei Experimentalgruppen hinsichtlich des Transfers, d. h. ist eine bestimmte Kombination zweier trainierter Komponenten am besten geeignet, Transfer auf die mathematischen Teilkomponenten zu erzielen (Fragestellung 3b)?

Fragestellung 4 (*Ferner Transfer Mathematik*): Steigert ein Training weniger, zentraler Komponenten der fächerübergreifenden Problemlösekompetenz die fachbezogene Problemlösekompetenz im Fach Mathematik (Fragestellung 4a)? Präziser formuliert: Schneiden die drei Experimentalgruppen besser beim Test der mathematischen Problemlösekompetenz (bestehend aus PISA 2003 Items) ab als die Kontrollgruppe? Unterscheiden sich die drei Experimentalgruppen hinsichtlich des Transfers, d. h. ist eine bestimmte Kombination zweier trainierter Komponenten am besten geeignet, Transfer beim Problemlösen zu erzielen (Fragestellung 4b)?

Fragestellung 5 (*kompensatorischer Trainingseffekt*): Gemäß der PISA-2003-Befunde ist die Potenzialdifferenz im unteren Leistungsbereich besonders groß (Leutner et al., 2004). Das ungenutzte kognitive Potenzial sollte vor allem im unteren Leistungsbereich vorhanden sein (Leutner et al., 2004). Daher wird untersucht, ob sich *differentielle* Trainingseffekte hinsichtlich der Schulform zeigen. Profitieren also insbesondere leistungsschwächere SuS (Hauptschüler, Realschüler) vom Training (kompensatorischer Trainingseffekt) oder zeigt sich im Gegenteil der häufig beobachte Matthäuseffekt („Wer hat, dem wir gegeben“, Klauer, 2001)?

3.3.2 Methode

3.3.2.1 Stichprobe

Die Durchführung des Experiments 2 erfolgte im Sommer 2012. Es nahmen 203 SuS verschiedener Hauptschulen (52 SuS), Realschulen (115 SuS) und Gymnasien (36 SuS) aus Nordrhein-Westfalen am Experiment teil.[13] Davon sind 93 SuS

[13] Um bei einer 4 (Bedingung) x 3 (Schulform) MANOVA mit drei abhängigen Variablen (je eine pro Teilkomponente) mindestens einen mittelgroßen Effekt (beim Treatment-Check) von $f^2 = .15$ mit großer Sicherheit ($\alpha = \beta = .05$) zu detektieren, wurde mit G*Power 3.1.9 von Faul et al. (2007) eine benötigte Stichprobengröße von $N = 74$ Personen berechnet. Um einen mittelgroßen Effekt (beim Problemlösen) von $f = .25$ mit üblichen Irrtumswahrscheinlichkeiten ($\alpha = .05$, $\beta = .80$) zu finden, ist eine Stichprobengröße von $N = 225$ Personen notwendig.

(46.0 %) weiblich. Eine Person machte keine Angabe zu ihrem Geschlecht. Das Durchschnittsalter der SuS beträgt 15.09 Jahre (*SD* = 1.26).

Die teilnehmenden Klassen erhielten 50 Euro für ihre Klassenkasse. Die Lehrer erhielten zusätzlich die Möglichkeit, Rückmeldung über die aggregierten Ergebnisse ihrer Klasse zu erhalten.

3.3.2.2 Instrumente und Materialien

Die Instrumente und Materialien entsprechen den in Experiment 1 beschriebenen mit folgenden Änderungen (vgl. Anhang L bis Anhang N):

1. *Lernmaterialien der Experimentalgruppe:* Den Blöcken zur Planungsfähigkeit und zum konditionalen Wissen wurde ein *worked example* hinzugefügt. Es wurden zusätzliche Hinweise beim konditionalen Wissen im Falle unbeantworteter Items eingebaut. Der Testblock Sachwissen wurde gestrichen, da er zu leicht war. Kleinere sprachliche Verbesserungen wurden vorgenommen. Den Schülerwünschen aus Experiment 1 entsprechend wurde das Lernmaterial etwas farbiger gestaltet (hellblau statt grau) und um eine schwerere Übungsaufgabe (Sitzordnung) ergänzt.

2. *Lernmaterialien der Kontrollgruppe:* Das Treatment der Kontrollgruppe wurde mit Untergrenzen der Bearbeitungszeit pro Seite versehen, da *logfile*-Analysen von Experiment 1 gezeigt haben, dass viele Schüler der Kontrollgruppe nur wenige Sekunden auf einzelnen HTML-Seiten verbracht haben.

3. *Fixe Testzeiten*: Um eine gleiche Posttestzeit in den Experimentalgruppen und der Kontrollgruppe für den computerbasierten Teil des Experiments zu erhalten, wurden fixe Zeitfenster geschaffen, d. h. 45 Minuten Treatment und 45 Minuten Posttest fächerübergreifendes Problemlösen. Eine Kontrolle zeigt, dass die vier Gruppen sich nicht in den Bearbeitungszeiten beim Treatment und Posttest unterscheiden (Anhang O). Wie in Experiment 1 dauerte der mathematische Posttest in Papierform 45 Minuten.

4. *Änderung des fächerübergreifenden Posttests*: Zu leichte Aufgaben zum Handlungswissen (*Fehlstunden* und *Einzelverbindungsnachweis*) wurden gestrichen. Die Posttest-Blöcke wurden zufällig rotiert, ebenso die Aufga-

ben innerhalb eines Blocks. Die selbst entwickelte Aufgabe *Sitzordnung* wurde vom Posttest in das Experimentalgruppentreatment verschoben.

Wie in Experiment 1 wurden zur Erfassung von Drittvariablen folgende Instrumente eingesetzt:

1. die Skala *Figurenanalogien* (*N2*) des KFT 4-12+R (Heller & Perleth, 2000) (M = 13.87, SD = 5.88, Cronbachs α = .881, 25 Items), wobei es sich um korrigierte Testhefte ohne die beiden fehlerhaften Items (Möller et al., 2006; Segerer et al., 2012) handelt,
2. die Skala *Anstrengungsbereitschaft* des Schülerfragebogens aus PISA 2003 (Boekaerts & Otten, 1993; Deutsches PISA-Konsortium, 2006) vor dem Treatment (M = 4.51, SD = 1.86, Cronbachs α = .870, 3 Items), nach dem Treatment (M = 5.68, SD = 2.13, Cronbachs α = .850, 3 Items) sowie vor dem Posttest (M = 5.16, SD = 2.27, Cronbachs α = .918, 3 Items). Die SuS berichteten im Mittel also eine hohe Anstrengungsbereitschaft.

3.3.2.3 Design

Das experimentelle Design wurde im Vergleich zu Experiment 1 verändert. Um die Lernzeit zu erhöhen, werden, statt alle drei Komponenten zu trainieren, pro Experimentalgruppe jeweils zwei der drei Komponenten trainiert. Es handelt sich somit um ein einfaktorielles Trainingsexperiment mit vier Gruppen (Proaktionsplan; Klauer, 2011). Die drei Experimentalgruppen trainierten jeweils zwei Komponenten fächerübergreifender Problemlösekompetenz (siehe Materialien):

- Experimentalgruppe 1: Planungsfähigkeit und konditionales Wissen,
- Experimentalgruppe 2: Planungsfähigkeit und Handlungswissen,
- Experimentalgruppe 3: Handlungswissen und konditionales Wissen.

Die Kontrollgruppe absolvierte wie in Experiment 1 ein Tutorial zur Verwendung der graphischen Benutzeroberfläche von GeoGebra (Hohenwarter, 2006). Die SuS wurden innerhalb jeder Klasse randomisiert einer Bedingung zugewiesen.

3.3.2.4 Durchführung

Die 135-minütige Durchführung erfolgte im Computerraum der jeweiligen Schule. Jeder Schüler hatte einen PC-Arbeitsplatz zur Verfügung. Die SuS wurden computerbasiert randomisiert einer von vier Gruppen zugeteilt (siehe Design).

Nach allgemeinen Instruktionen, einem kurzen demographischen Fragebogen und drei Fragen zur Anstrengungsbereitschaft, bearbeiteten die SuS das jeweilige Lernprogramm. Im Anschluss wurden die retrospektiv formulierten Skalen zur Anstrengungsbereitschaft und Bewertungsmöglichkeiten des jeweiligen Lernprogramms administriert. Anschließend wurde der Posttest administriert (Anhang N). Abbildung 12 visualisiert zusammenfassend den Ablauf des Experiments 2.

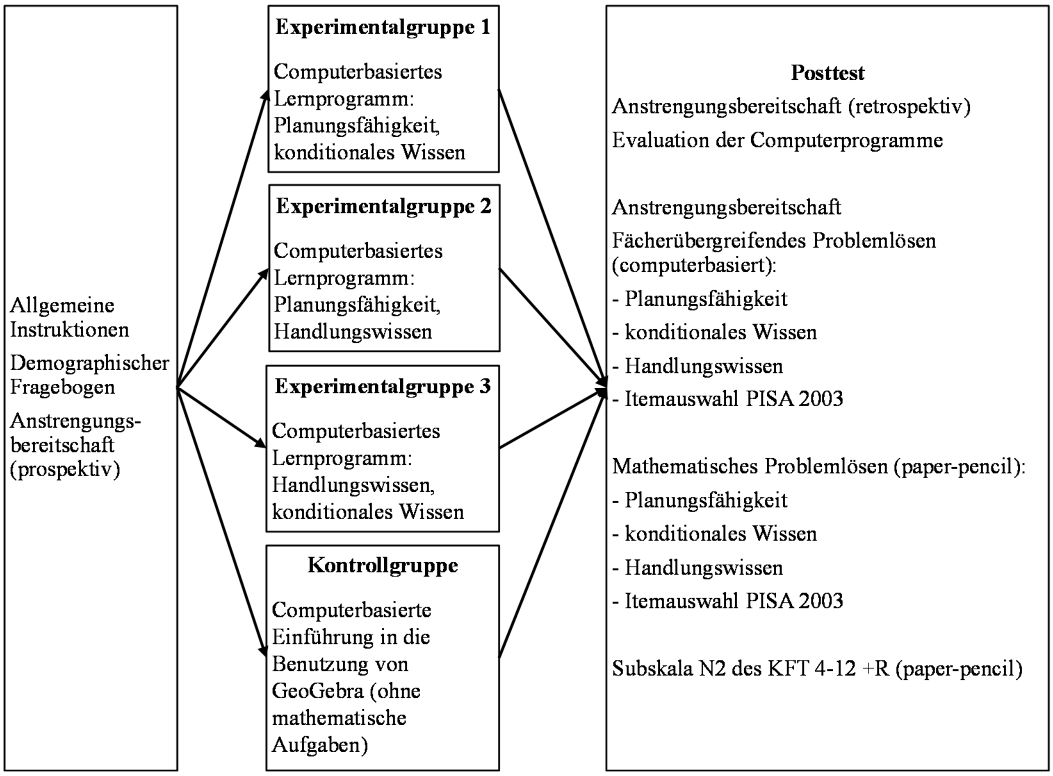


Abbildung 12. Ablaufdiagramm zum Experiment 2

3.3.2.5 Scoring

Das Scoring erfolgte wie in Experiment 1 (siehe Kapitel 3.2.2.5).

3.3.2.6 Datenanalyse

Da im Gegensatz zu Experiment 1 jeweils 45 Minuten für Treatment und Posttest zur Verfügung standen, können statt Effizienzmaßen Leistungsmaße analysiert werden. Eine Betrachtung der Bearbeitungszeiten pro experimenteller Bedingung zeigt, dass sich die Bearbeitungszeiten sowohl für das Treatment als auch den Posttest in allen Bedingungen nicht unterscheiden (Anhang O).

Zur Hypothesenprüfung werden Kovarianzanalysen mit anschließenden Bonferoni-korrigierten paarweisen Vergleichen betrachtet (Field, 2009).

3.3.3 Ergebnisse

Im Folgenden wird zunächst der Treatment-Check für die drei Komponenten fächerübergreifenden Problemlösens berichtet. Anschließend werden verschiedene Transfer-Maße untersucht: fächerübergreifendes Problemlösen mit PISA-2003-Items, drei Komponenten mathematischen Problemlösens, und mathematisches Problemlösen mit PISA-2003-Items.

3.3.3.1 Treatment-Check: Komponenten fächerübergreifenden Problemlösens

Deskriptiv unterschieden sich die z-Werte des Handlungswissens der trainierten Gruppen und der Kontrollgruppe kaum (Abbildung 13, Tabelle P-1). Deskriptiv erzielten die beiden im konditionalen Wissen trainierten Gruppen mehr Punkte im konditionalen Wissen als die Kontrollgruppen (Tabelle P-2). Bei der Skala Planungsfähigkeit erzielte die Kontrollgruppe deskriptiv die meisten Punkte. Schulformeffekte waren nicht zu beobachten (Tabelle P-3).

Für den Treatment-Check wurde eine 4 (Treatments) x 3 (Schulform: Hauptschule vs. Realschule vs. Gymnasium) MANCOVA mit den z-standardisierten Skalen zum konditionalen Wissen, zur Planungsfähigkeit und zum Handlungswissen als abhängigen Variablen und der z-standardisierten Skala *Figurenanalogien* (*N2*) des KFT 4-12+R (Heller & Perleth, 2000) als Kovariate berechnet. Da in einer Gymnasialklasse aus Zeitgründen die Skala *N2* nicht erhoben werden konnte, reduzierte sich die Stichprobengröße für die folgenden Analysen auf 182. Die multivariaten Ergebnisse zeigten, dass die Kovariate einen signifikanten Effekt hat (Wilks $\Lambda = .792$, F(3, 167) = 14.649, p < .001, $\eta^2 = .208$). Die multivariate Interaktion von Schulform und Bedingung war nicht signifikant (Wilks $\Lambda = .864$, $F(18, 472.833) = 1.396$, $p = .128$, $\eta^2 = .048$). Die Schulform hatte keinen signifikanten multivariaten Effekt (Wilks $\Lambda = .974$, $F < 1$). Die Bedingung hatte einen signifikanten multivariaten Effekt (Wilks $\Lambda = .898$, $F(9, 406.585) = 2.047, p = .033$, $\eta^2 = .035$). Follow-up-ANCOVAs (Tabelle P-4) zeigten, dass sich die Bedingungen nur hinsichtlich des konditionalen Wissens unterschieden ($F(3, 169) = 5.095$, $p = .002$, $\eta^2 = .083$). Bonferoni-korrigierte paarweise Vergleiche zeigten, dass Gruppe 1, die konditionales Wissen und Planungsfähigkeit trainierte, signifikant höhere Werte beim konditionalen Wissen erzielte als die drei anderen

Gruppen.[14] Alle anderen Gruppen unterschieden sich nicht signifikant (siehe Tabelle P-5). Zwischenfazit: Das Treatment kann somit nur für Gruppe 1 hinsichtlich des konditionalen Wissens als erfolgreich betrachtet werden.

[14] In Abbildung 13 erscheint der Gruppenunterschied nicht signifikant, da Abbildung 13 keine Kovariate berücksichtigt.

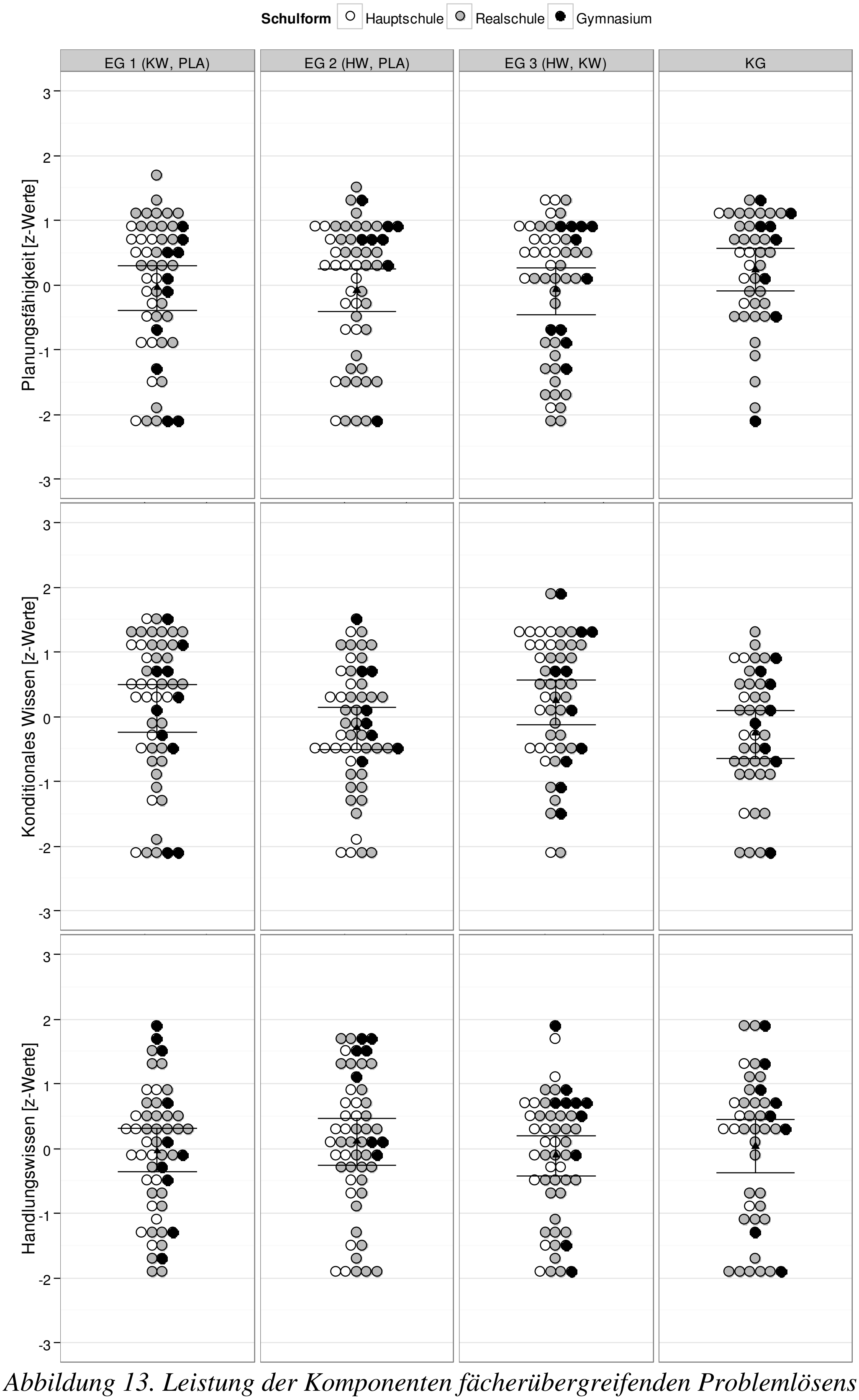


Abbildung 13. Leistung der Komponenten fächerübergreifenden Problemlösens

Anmerkungen. EG = Experimentalgruppe, KW = Konditionales Wissen, HW = Handlungswissen, PLA = Planungsfähigkeit, KG = Kontrollgruppe. Das schwarze Dreieck kennzeichnet den Mittelwert. Fehlerbalken repräsentieren das 99%-Bootstrap-Konfidenzintervall des Mittelwerts (mit 1000 Replikationen). Die Abbildung wurde mit dem R-Paket *ggplot 2* (Wickham, 2009) erzeugt.

3.3.3.2 Transfer Problemlösen

Deskriptiv zeigte sich bei der Skala zum fächerübergreifenden Problemlösen, dass sich die Gruppen kaum unterscheiden (Abbildung 14, Tabelle P-6). Um den Transfer auf fächerübergreifendes Problemlösen inferenzstatistisch zu prüfen, wurde eine ANCOVA berechnet mit Bedingung und Schulform als unabhängigen Variablen, der Skala *Figurenanalogien* (*N2*) des KFT 4-12+R (Heller & Perleth, 2000) als Kovariate und der *Leistung* beim fächerübergreifenden Problemlösetest als abhängiger Variable. Die ANCOVA-Ergebnisse zeigten, dass die Kovariate einen signifikanten Effekt hatte ($F(1, 169) = 27.385, p < .001$, $\eta^2 = .139$). Die Interaktion von Schulform und Bedingung war nicht signifikant ($F(6, 169) = 1.550, p = .165$, $\eta^2 = .052$). Die Schulform hatte einen signifikanten multivariaten Effekt ($F(2, 169) = 5.707, p = .004$, $\eta^2 = .067$). Die Bedingung hatte keinen signifikanten Effekt ($F(3, 169) = 2.057, p = .108$, $\eta^2 = .035$). Bonferoni-korrigierte paarweise Vergleiche zeigten, dass Gymnasiasten signifikant höhere Werte beim Problemlösen erzielten als Haupt- und Realschüler, die sich nicht in ihrer Problemlöseleistung unterschieden (Tabelle P-7).

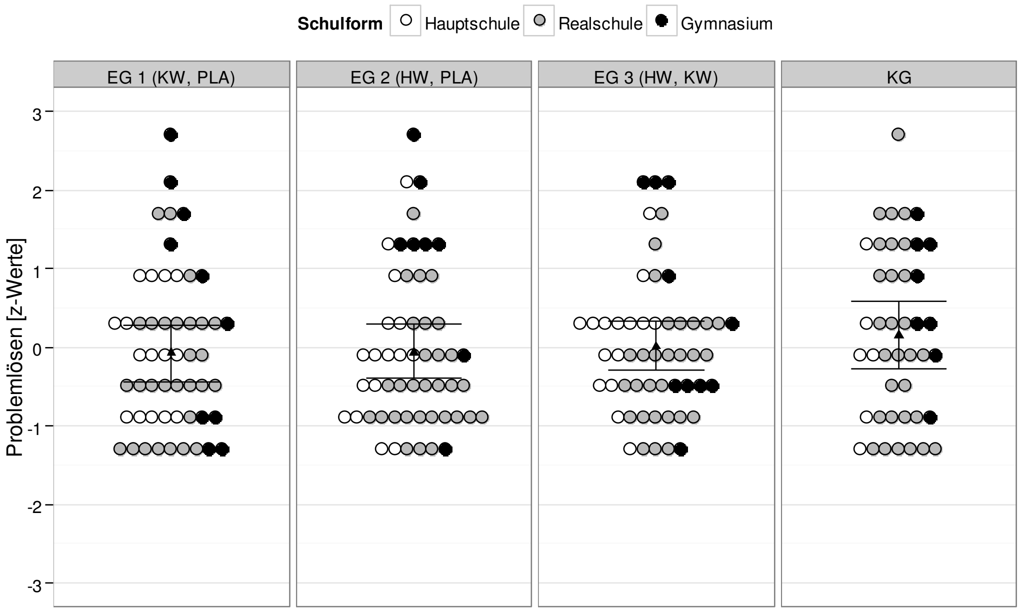


Abbildung 14. Leistung im fächerübergreifenden Problemlösen

Anmerkungen. EG = Experimentalgruppe, KW = Konditionales Wissen, HW = Handlungswissen, PLA = Planungsfähigkeit, KG = Kontrollgruppe mit GeoGebra-Treatment. Das schwarze Dreieck kennzeichnet den Mittelwert. Fehlerbalken repräsentieren das 99%-Bootstrap-Konfidenzintervall des Mittelwerts (mit 1000 Replikationen). Die Abbildung wurde mit dem R-Paket *ggplot 2* (Wickham, 2009) erzeugt.

3.3.3.3 Naher Transfer Mathematik: Komponenten mathematischen Problemlösens

Die Dotplots der mathematischen Skalen (Abbildung 15) sprachen gegen Gruppenunterschiede (vgl. auch Tabelle P-8 bis P-10). Zur Hypothesenprüfung wurde eine 4 (Treatments) x 3 (Schulform: Hauptschule vs. Realschule vs. Gymnasium) MANCOVA mit den z-standardisierten Skalen zum konditionalen Wissen, zur Planungsfähigkeit und zum Handlungswissen als abhängigen Variablen und der z-standardisierten Skala *Figurenanalogien* (*N2*) des KFT 4-12+R (Heller & Perleth, 2000) als Kovariate berechnet. Die multivariaten Ergebnisse zeigten, dass die Kovariate einen signifikanten Effekt hatte (Wilks $\Lambda = .859$, $F(3, 167) = 14.649$, $p < .001, \eta^2 = .141$). Die multivariate Interaktion von Schulform und Bedingung war nicht signifikant (Wilks $\Lambda = .851$, $F(18, 472.833), p = .072$, $\eta^2 = .052$). Die Schulform hatte einen signifikanten multivariaten Effekt (Wilks $\Lambda = .667$, $F(6, 334), p < .001$, $\eta^2 = .184$). Die Bedingung hatte keinen signifikanten multivariaten Effekt (Wilks $\Lambda = .945$, $F(9, 406.585) = 1.055, p = .395$, $\eta^2 = .019$). Follow-up-ANCOVAs (Tabelle P-11) zeigten, dass sich die Schulformen hinsichtlich des konditionalen Wissens ($F(2, 169) = 4.705, p = .010$, $\eta^2 = .053$) und des Handlungswissens ($F(2, 169) = 48.588$, $p < .001$, $\eta^2 = .313$) unterschieden und bezüglich der Planungsfähigkeit ($F(2, 169) = 1.613$, $p = .202$, $\eta^2 = .019$) nicht unterschieden. Bonferoni-korrigierte paarweise Vergleiche (Tabelle P-12) zeigten, dass beim konditionalen Wissen Gymnasiasten signifikant höhere Werte erzielten als Haupt- und Realschüler, die sich nicht unterschieden. Beim Handlungswissen erzielten Gymnasiasten höhere Werte als Haupt- und Realschüler und Realschüler höhere Werte als Hauptschüler. Zwischenfazit: Es fanden sich keine Transfereffekte auf Komponenten mathematischen Problemlösens.

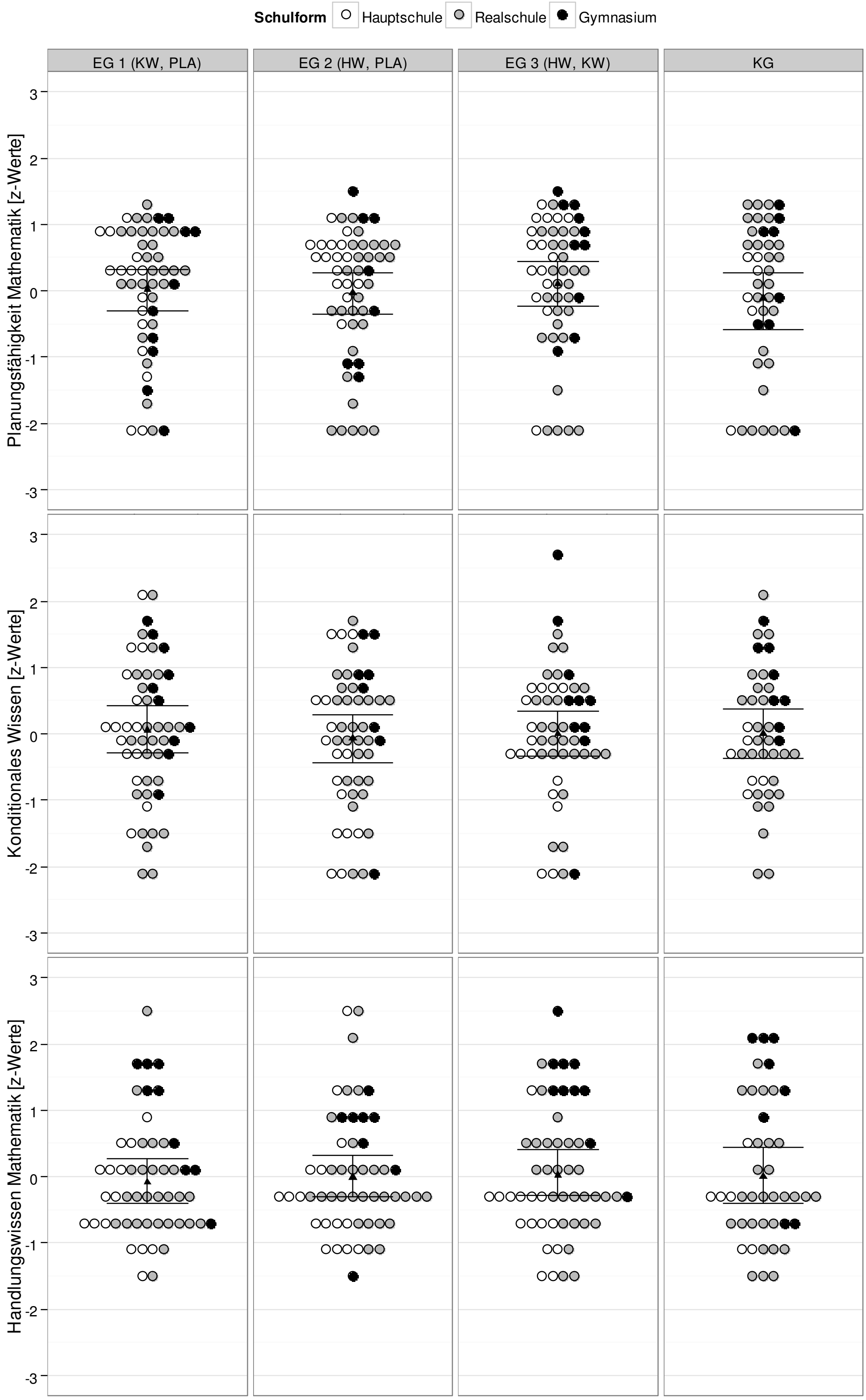


Abbildung 15. Leistung der drei Komponenten mathematischen Problemlösens

Anmerkungen. EG = Experimentalgruppe, KW = Konditionales Wissen, HW = Handlungswissen, PLA = Planungsfähigkeit, KG = Kontrollgruppe. Das schwarze Dreieck kennzeichnet den Mittelwert. Fehlerbalken repräsentieren das 99%-Bootstrap-Konfidenzintervall des Mittelwerts (mit 1000 Replikationen). Die Abbildung wurde mit dem R-Paket *ggplot 2* (Wickham, 2009) erzeugt.

3.3.3.4 Ferner Transfer Mathematik: Mathematische Problemlösekompetenz

Die Deskriptivstatistik der Skala *mathematisches Problemlösen* (Tabelle P-13) und die Dotplots (Abbildung 16) fielen in allen Gruppen sehr ähnlich aus. Um den Transfer auf mathematisches Problemlösen inferenzstatistisch zu prüfen, wurde eine ANCOVA berechnet mit Bedingung und Schulform als unabhängigen Variablen, der Skala *Figurenanalogien* (*N2*) des KFT 4-12+R (Heller & Perleth, 2000) als Kovariate und der *Leistung* beim mathematischen als abhängiger Variable. Die ANCOVA-Ergebnisse zeigten, dass die Kovariate einen signifikanten Effekt hatte ($F(1, 169) = 28.760, p < .001,\ \eta^2 = .145$). Die Interaktion von Schulform und Bedingung war nicht signifikant ($F(6, 169) = 1.031, p = .407,\ \eta^2 = .035$). Die Bedingung hatte keinen signifikanten Effekt ($F(3, 169) = 1.062, p = .367,\ \eta^2 = .019$), d. h. es gibt keinen Hinweis auf Transfereffekte von fächerübergreifendem auf mathematisches Problemlösen. Die Schulform hatte einen signifikanten Effekt ($F(2, 169) = 5.003, p = .008,\ \eta^2 = .056$). Bonferoni-korrigierte paarweise Vergleiche zeigten, dass Gymnasiasten signifikant höhere Werte beim mathematischen Problemlösen erzielten als Haupt- und Realschüler, die sich nicht in ihrer Problemlöseleistung unterschieden (siehe Tabelle P-14).

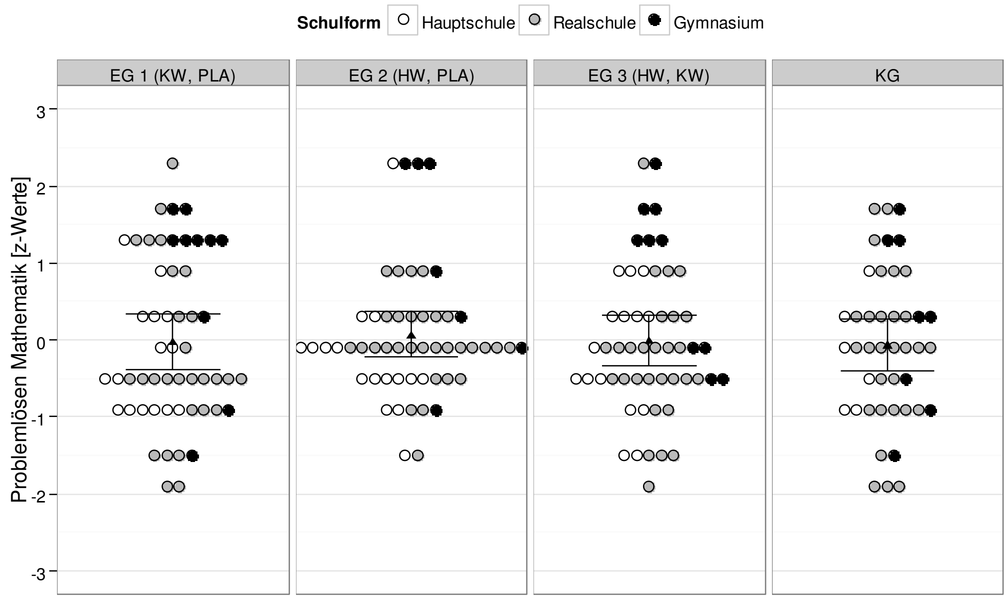


Abbildung 16. Leistung im mathematischen Problemlösen

Anmerkung. Das schwarze Dreieck kennzeichnet den Mittelwert. Fehlerbalken repräsentieren das 99%-Bootstrap-Konfidenzintervall des Mittelwerts (mit 1000 Replikationen). Die Abbildung wurde mit dem R-Paket *ggplot 2* (Wickham, 2009) erzeugt.

3.3.3.5 Weitere Befunde

Wie in Experiment 1 zeigten sich zumeist Zusammenhänge mittlerer Größe zwischen den abhängigen Variablen (Tabelle 7). Im Unterschied zu Experiment 1 korrelierten in Experiment 2 alle fächerübergreifenden Skalen mit allen mathematischen Skalen. Allerdings korrelierten im Gegensatz zu Experiment 1 nicht alle mathematischen Skalen untereinander (Tabelle 7).

Tabelle 7. Interkorrelation der Skalen (untere Diagonalmatrix, N = 203) und Partialkorrelation der Skalen bei Berücksichtung der Skala Figurenanalogien (N2) des KFT-4-12+R (obere Diagonalmatrix, N = 179) in Experiment 2

			Fächerübergreifende Skalen (Leistung)				*Mathematische Skalen (Leistung)*			
			PLA	KW	HW	PISA	PLA	KW	HW	PISA
Fächerübergreifende Skalen (Leistung)	Planungsfähigkeit (PLA)	*r*		.407	.384	.410	.140	.155	.063	.169
		p		.000	.000	.000	.030	.019	.201	.011
	Konditionales Wissen (KW)	*r*	.473		.336	.280	.171	.211	.127	.099
		p	.000		.000	.000	.011	.002	.045	.091
	Handlungs-Wissen (HW)	*r*	.477	.414		.390	.167	.142	.221	.205
		p	.000	.000		.000	.012	.028	.001	.003
	Problemlösen (PISA)	*r*	.502	.379	.483		.149	.215	.387	.458
		p	.000	.000	.000		.022	.002	.000	.000
Mathematische Skalen (Leistung)	Planungsfähigkeit	*r*	.197	.202	.192	.157		.080	.270	-.018
		p	.002	.002	.003	.013		.141	.000	.405
	Konditionales Wissen	*r*	.153	.239	.164	.247	.103		.301	.206
		p	.015	.000	.010	.000	.071		.000	.003
	Handlungs-Wissen	*r*	.113	.197	.304	.386	.306	.336		.284
		p	.054	.002	.000	.000	.000	.000		.000
	Problemlösen	*r*	.257	.187	.344	.526	.010	.300	.369	
		p	.000	.004	.000	.000	.446	.000	.000	

Anmerkungen. Signifikante Korrelationen (mindestens $p < .05$) sind grau hinterlegt. *P*-Werte sind auf drei Nachkommastellen gerundet.

Da die multivariaten Interaktionseffekte von Bedingung und Schulform in Experiment 2 alle nicht signifikant waren und die Teilnehmerzahlen aus den einzelnen Schulformen unterschiedlich waren (siehe Stichprobe), werden in Abbildung 17 die Ergebnisse der acht Leistungsskalen ohne Berücksichtigung der Schulform zusammenfassend dargestellt.

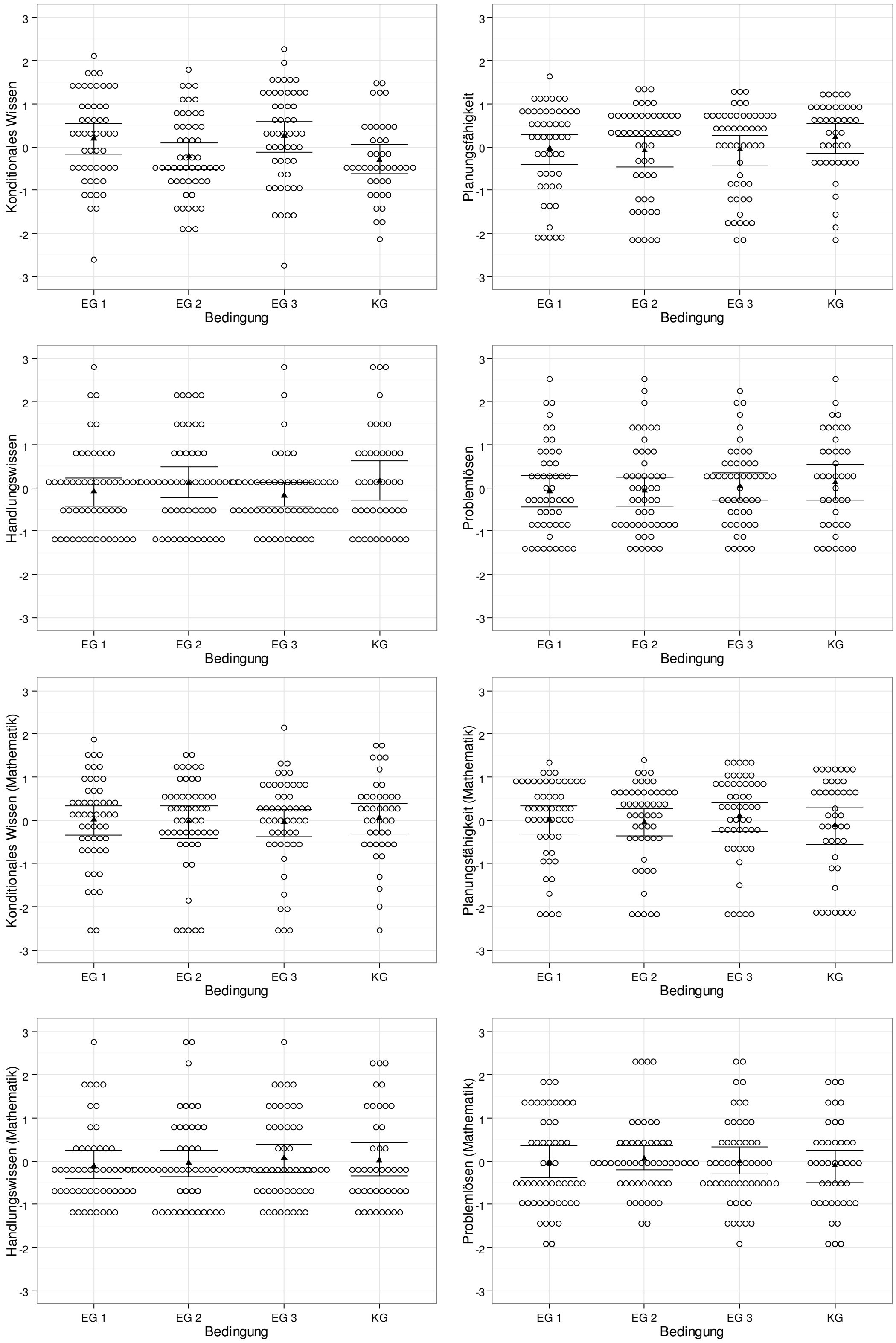


Abbildung 17. Leistung der vier Skalen zum fächerübergreifenden Problemlösen (obere Hälfte) und der vier Skalen zum mathematischen Problemlösen (untere Hälfte).

Anmerkung. Das schwarze Dreieck kennzeichnet den Mittelwert. Fehlerbalken repräsentieren das 99%-Bootstrap-Konfidenzintervall des Mittelwerts (mit 1000 Replikationen). Die Abbildung wurde mit dem R-Paket *ggplot 2* (Wickham, 2009) erzeugt.

3.3.4 Diskussion

Der Treatment-Check der Komponenten fächerübergreifenden Problemlösens in Experiment 2 war nur für konditionales Wissen erfolgreich. Bezüglich der Skalen Handlungswissen und Planungsfähigkeit unterscheiden sich die Gruppen nicht. Die Gruppen unterscheiden sich auch nicht beim Problemlösetest bestehend aus Items aus PISA 2003 (OECD, 2004b), d. h. es gibt keinen Hinweis auf nahen Transfer innerhalb des fächerübergreifenden Problemlösens, wobei die Lösungsquote in allen Gruppen eher niedrig ist. Das spricht einerseits dafür, dass die Aufgaben valide Problemlöseaufgaben sind. Andererseits scheinen sie für die untersuchten Stichproben etwas zu schwer zu sein. Wie in Experiment 1 unterscheiden sich die Gruppen in Experiment 2 nicht bezüglich der Leistung in den vier Tests der mathematischen Domäne (Handlungswissen, konditionales Wissen, Planungsfähigkeit, Problemlösen). Es gibt somit keinen Hinweis auf fernen Transfer, der als Argument für die kognitive Potenzialausschöpfungshypothese von Leutner et al. (2004) gewertet werden könnte.

Anders als in Experiment 1 wurden in Experiment 2 pro Experimentalgruppe jeweils zwei von drei Teilkomponenten fächerübergreifenden Problemlösens trainiert. Es stellt sich die Frage, wie viele Komponenten trainiert werden müssen. Wenn man die Reihenfolge der trainierten Komponenten außer Acht lässt, so ergeben sich bei drei Komponenten (Handlungswissen, konditionales Wissen, Planungsfähigkeit) folgende mögliche Trainingskombinationen: ein Training aller drei Komponenten (wie in Experiment 1), drei Trainings je zweier Komponenten (wie in Experiment 2), drei Trainings je einer Komponente. Kombinierte Trainings müssen jedoch nicht zwingend zu positiven summativen Effekten führen. Es kann zu Interferenzen kommen, die sich in Nulleffekten oder sogar negativem Transfer niederschlagen (Klauer, 2010). Um diese unerwünschten Effekte ausschließen zu können, müssten die Komponenten einzeln trainiert und analysiert werden. Alternativ ist die Verwendung der Subtraktionsmethode (Klauer, 2001) denkbar, bei der die Experimentalgruppen unterschiedlich viele Komponenten trainieren. Ferner berücksichtigen die computerbasierten Trainings nicht, dass manche SuS womöglich über hinreichende Ausprägungen auf einzelnen Teilkomponenten des Problemlösens verfügen. Binnendifferenzierung oder stärker adaptive Trainings bieten hier entsprechende Ansatzpunkte. Dazu ist eine entsprechende Diagnostik vor bzw. während des Trainings notwendig. Für diese ist im fächerübergreifenden Bereich die Entwicklung eines größeren Itempools wünschenswert.

Die verwendeten Trainings- und Testmaterialien enthalten Problemlöseaufgaben, die sich auf je eine Komponente der Problemlösefähigkeit beziehen. Ob jedoch einzelne Aufgaben ausschließlich eine Komponente der Problemlösekompetenz steigern, ist fraglich. Dies arbeitet jedoch nicht gegen die Hypothesen. Im Gegenteil, Problemlösekompetenz soll möglichst umfassend gesteigert werden. Ergänzend muss zu den Planungsaufgaben angemerkt werden, dass sie nicht alle Aspekte der Selbstregulation trainieren bzw. erfassen. Die Korrelation der mathematischen Planungsaufgaben mit mathematischen Problemlösen war in Experiment 1 gering und in Experiment 2 nicht signifikant. Umfangreichere Strukturgleichungsmodellierungen zur Struktur der Tests der Komponenten fächerübergreifender und mathematischer Problemlösekompetenz sind in Vorbereitung (J. Fleischer, persönl. Mitteilung, 25.05.2012).

Zudem weisen Strategiewissenstests, wie sie in Experiment 1 und 2 eingesetzt worden sind, trotz zufriedenstellender psychometrischer Kennwerte eine Reihe von potentiell problematischen Schwachpunkten auf (z. B. Wahl der Szenarien, Qualität und Anzahl der Expertenurteile, Kriterien der Expertenübereinstimmung, potentielle Testlet-Effekte, Konsequenzen bestimmter Scoringregeln), die in Anhang I ausführlicher dargestellt werden. Weitere Studien mit zusätzlichen Szenarien erscheinen wünschenswert, um z. B. zu prüfen, welche Konsequenzen sich aus verschiedenen Scorings ergeben (McDaniel & Nguyen, 2001) oder wie situationsspezifisch bzw. allgemein das erfasste Strategiewissen ist. Strategiewissen ist zudem nur ein Teil konditionalen Wissens. Zukünftige Studien sollten daher gegebenenfalls einen breiteren Ansatz zum Training und zur Diagnostik konditionalen Wissens verwenden.

Experiment 1 und 2 untersuchen, ob eine experimentelle Variation der fächerübergreifenden Problemlösekompetenz zu veränderter mathematischer Problemlösekompetenz führt (Leutner et al., 2009). Diese Prüfung möglicher (kausaler) Transfereffekte von fächerübergreifendem Problemlösen auf mathematisches Problemlösen untersucht einen Aspekt der kognitiven Potenzialausschöpfungshypothese. Insgesamt finden sich in beiden computerbasierten Experimenten 1 und 2 keine Transfereffekte auf die mathematische Domäne. Damit liefern die beiden Laborexperimente keine Evidenz für die Gültigkeit der kognitiven Potenzialausschöpfungshypothese von Leutner et al. (2004). Einschränkend sei bemerkt, dass der Nulltransfer auf die mathematischen PISA-Aufgaben auch positiv als Validitätsindikator für die verwendeten Aufgaben interpretierbar ist. Ein 45-minütiges Training reicht eben nicht aus, damit SuS im Mittel höhere Leistungen bei diesen Aufgaben erzielen. Womöglich waren die Aufgaben also nicht sensitiv genug, um Trainingseffekte kurzzeitiger Trainings zu detektieren.

Die Experimente 1 und 2 enthielten aus Zeitgründen jedoch keinen Prätest, da die Experimente jeweils insgesamt bereits drei Schulstunden dauerten. Dadurch konnte das mathematikspezifische Vorwissen, das eine große Rolle für mathematisches Problemlösen spielt (Carson, 2007; Renkl & Stern, 1994), nicht berücksichtigt werden. Ein Prätest fächerübergreifenden und mathematischen Problemlösens ist zusätzlich sinnvoll, da er z. B. ermöglicht zu prüfen, ob bzw. wie viele SuS eingangs besser beim fächerübergreifenden Problemlösen abschneiden als beim mathematischen Problemlösen. Die Erhebung eines Prätests erlaubt daher eine Messung des Potenzials im Sinne der kognitiven Potenzialausschöpfungshypothese (Leutner et al., 2004) und damit eine direktere Prüfung der kognitiven Potenzialausschöpfungshypothese, indem differentielle Trainings- und Transfereffekte untersucht werden können (siehe Experiment 3).

Die zeitlichen Grenzen des Laborexperiments führen zu einer wichtigen Frage: Was ist von kurzen Laborexperimenten bezüglich der kognitiven Potenzialausschöpfungshypothese (Leutner et al., 2004) erwartbar? Mit anderen Worten: Sind längerfristige Interventionen zur Hypothesenprüfung notwendig? Die kritische Frage, ob die kurzen Trainingsmodule überhaupt langfristige Effekte haben können, wobei bewusst keine Zeitangabe für Langfristigkeit gegeben wird, ist pädagogisch berechtigt, zur Beantwortung der vorliegenden psychologischen Forschungsfragen jedoch eher nachrangig. Ziel der Experimente 1 und 2 ist nicht primär die Entwicklung einer praxis-tauglichen Problemlöseförderungsmaßnahme, sondern eine experimental-psychologische Theorieprüfung, die mit der Frage der prinzipiellen Förderbarkeit der Problemlösekompetenz im Sinne der Operationalisierung bei PISA 2003 (OECD, 2003) einhergeht. Das Experiment 3 setzt die Untersuchung von Aspekten der kognitiven Potenzialausschöpfungshypothese dennoch in einem längerfristigen Trainingskontext in der Schule fort. Das trägt einerseits der psychologischen Forschungstradition stärker Rechnung, die zeigt, dass spontaner Transfer selten ist (siehe z. B. Barnett & Ceci, 2002; Perkins & Salomon, 1989). Anderseits wird eine mathematikdidaktische Perspektive stärker berücksichtigt, die den Erwerb von mathematischer Problemlösekompetenz als einen längerfristigen Prozess betrachtet, der kontinuierlich in den (Mathematik-)Unterricht integriert werden muss, damit Problemlösen tatsächlich zum „Lernangebot für alle" (Bruder, 2000, S. 2) wird, wie es die deutschen Bildungsstandards vorsehen (KMK, 2004b, 2005d).

3.4 Experiment 3

Da die laborexperimentellen Befunde keine Evidenz für die kognitive Potenzialausschöpfungshypothese erbracht haben, wurde ein veränderter methodischer Zugang gewählt, um die Potenzialausschöpfungshypothese erneut empirisch zu prüfen. Im Rahmen einer mehrwöchigen feldexperimentellen Trainingsstudie im Schulkontext wird die kognitive Potenzialausschöpfungshypothese ökologisch valider und direkter untersucht. Dazu wird u. a. die Trainingszeit deutlich verlängert und ein Prätest erhoben, d. h. in Experiment 3 wird das Vorwissen der SuS erfasst, da Vorwissen ein guter Prädiktor für domänenspezifisches Problemlösen ist (Clarke, Ayres & Sweller, 2005; Fyfe, Rittle-Johnson & DeCaro, 2012; Kalyuga, 2005; Mayer, 1992; Renkl & Stern, 1994; Rittle-Johnson, Star & Durkin, 2009). Ein Prätest ermöglicht es zudem, eine eventuelle Diskrepanz fächerübergreifenden und mathematischen Problemlösens zu Trainingsbeginn zu quantifizieren. Dieser Prätest ermöglicht also das kognitive Potenzial im Sinne der kognitiven Potenzialausschöpfungshypothese (Leutner et al., 2004) zu erfassen. Um Zeit für diese im Vergleich zu Experiment 1 und 2 umfangreichere Testung fächerübergreifenden und mathematischen Problemlösens zu gewinnen, wurde auf die Testung von Handlungswissen und konditionalem Wissen verzichtet. Die Komponente Planungskompetenz wurde mit dem Aachener Planungskompetenztraining (Arling & Spijkers, 2012) trainiert, das auch einen Prä- und Posttest enthält.

3.4.1 Fragestellungen

Wie in Experiment 1 und 2 wurden potentielle Transfereffekte eines fächerübergreifenden Problemlösetrainings auf die Domäne Mathematik untersucht. Dazu wurden in Experiment 3 relevante Teilkomponenten des fächerübergreifenden (und mathematischen) Problemlösens (Fleischer et al., 2010) kombiniert mit Heuristiken als Transferhilfen im regulären Unterrichtskontext über einen Zeitraum von mehreren Wochen trainiert. Experiment 3 untersuchte im Einzelnen folgende Fragestellungen:[15]

- *Fragestellung 6 (Treatment-Check)*: Lässt sich mit dem ursprünglich für Erwachsene konzipierten Aachener Planungskompetenztraining (Arling & Spijkers, 2012) Planungskompetenz (als eine Teilkomponente des Problemlösens) in der Jahrgangsstufe 9 trainieren?
- *Fragestellung 7 (Transfer Problemlösen)*: Steigert das fächerübergreifende Problemlösekompetenztraining mit Fokus auf spezifischen Komponenten (Planungsfähigkeit, konditionales Wissen, Handlungswissen, Heuristiken) die allgemeine fächerübergreifende Problemlösefähigkeit?
- Fragestellung *8 (Transfer Mathematik)*: Steigert das fächerübergreifende Problemlösekompetenztraining mit Fokus auf spezifischen Komponenten (Planungsfähigkeit, konditionales Wissen, Handlungswissen, Heuristiken) die Problemlösekompetenz im Fach Mathematik?

Werden Transfereffekte auf die Domäne Mathematik gefunden, so spricht dies für die Rolle fächerübergreifenden Problemlösens beim Aufbau mathematischer Kompetenz und damit für die kognitive Potenzialausschöpfungshypothese.

In Experiment 3 ist durch die Verwendung eines Prätests der Domänen Problemlösen und Mathematik die Möglichkeit gegeben, das vorhandene kognitive Potenzial im Sinne der kognitiven Potenzialausschöpfungshypothese (Leutner et al., 2004, 2005) vor Trainingsbeginn zu untersuchen. Damit können differentielle Effekte untersucht werden:

- *Fragestellung 9a (differentielle Effekte und kognitives Potenzial):* Profitieren im Sinne der kognitiven Potenzialausschöpfungshypothese (Leutner et al., 2004; OECD, 2004b) insbesondere SuS vom Training, die im Prätest fächerübergreifenden Problemlösens eher hoch und im Prätest Mathematik eher niedrig abgeschnitten haben?

Als direktere Untersuchung der kognitiven Potenzialausschöpfung wird geprüft, ob sich Effekte der experimentell zu manipulierenden fächerübergreifenden Problemlösekompetenz beim Lösen mathematischer Probleme insbesondere für SuS zei-

[15] Die Fragestellungen werden über die Experimente hinweg fortlaufend nummeriert, um eine eindeutige Referenzierbarkeit zu erlauben.

gen, die zu Beginn des Trainings schlechter in Mathematik waren als auf Grundlage ihrer fächerübergreifenden Problemlösekompetenz zu erwarten war:

- *Fragestellung 9b (kognitive Potenzialausschöpfung):* Profitieren insbesondere SuS vom Training, bei denen zum Zeitpunkt des Prätests kognitives Potenzial im Sinne der kognitiven Potenzialausschöpfungshypothese (Leutner et al., 2004; OECD, 2004b) vorhanden ist?

Geprüft wird ob SuS, die zu Beginn des Trainings niedriger in Mathematik abgeschnitten hatten, als es auf Grundlage ihrer fächerübergreifenden Problemlösekompetenz zu erwarten war, im Posttest ihr kognitives Potenzial ausschöpfen.

Als ergänzende Indikatoren des Trainingserfolgs werden ferner die *Bearbeitungsquote und strategische Herangehensweisen* der SuS untersucht: Mathematikdidaktische und psychologische Studien zu mathematischen Textaufgaben zeigen, dass viele SuS erst gar nicht versuchen, ein Problem zu verstehen und die Aufgabe auslassen oder planlos rechnen (siehe Kajamies, Vauras & Kinnunen, 2010, für Literaturangaben).

> „Es besteht kein Mut zu Unternehmungen mit ungewissem Ausgang. Vielmehr existiert […] die Erwartung von Schemata, schön ordentlich auf eine Reihe von Kästen verteilt, so daß man nur noch den richtigen herausfinden muß“ (Kießwetter, 1983, S. 74).

Selbst gute SuS zeigten die Tendenz, bei schwierigen Aufgaben nichts zu schreiben (Doorman et al., 2007). In einer Studie von van den Heuvel-Panhuizen und Bodin-Baarends (2004) beispielsweise machten sich selbst unter mathematisch leistungsstarken Grundschülern (Top 20 %) nur ein Drittel Notizen beim Bearbeiten mathematischer Probleme. Ein erfolgreiches Problemlösetraining sollte daher auch eine veränderte Herangehensweise an Probleme bewirken (vgl. Bruder, 2002) in dem Sinne, dass SuS vermehrt versuchen, Probleme zu lösen:

Fragestellung 10a (Bearbeitungsquote): Fördert das das fächerübergreifende Problemlösekompetenztraining mit Fokus auf spezifische Komponenten (u. a. Planungsfähigkeit, konditionales Wissen) den Willen zur Aufgabenbearbeitung in Form einer höheren Bearbeitungsquote?

Dabei können mehr Notizen bzw. eine höhere Bearbeitungsquote im Vergleich zur Kontrollgruppe als Indikator für den Willen zur Aufgabenbearbeitung (*task persistence*) gelten (Ho & Hedberg, 2005). Zusätzlich sollten ein erfolgreiches Problemlösetraining bewirken, dass SuS strategischer vorgehen und damit verstärkt systematische Problemlöseansätze wählen und weniger unsystematisches Problemlösen (vgl. Bertling, 2012):

> *Fragestellung 10b (Herangehensweise an Probleme):* Fördert das das fächerübergreifende Problemlösekompetenztraining mit Fokus auf spezifische Komponenten (u. a. Planungsfähigkeit, konditionales Wissen) die Wahl systematischer Problemlöseansätze?

3.4.2 Methode

3.4.2.1 Stichprobe

Die Stichprobenumfangsplanung basierte auf einer Kovarianzanalyse mit der Bedingung als Faktor und 5 Kovariaten (Prätest Mathematik, Prätest Problemlösen, Prätest Mathematik x Bedingung, Prätest Problemlösen x Bedingung, Prätest Mathematik x Prätest Problemlösen x Bedingung). Eine Berechnung mit G*Power 3.1.9 (Faul, Erdfelder, Buchner & Lang, 2009) unter Annahme eines mittelgroßen Effekts von $f = .25$, eines Signifikanzniveaus von $\alpha = .05$ und einer Teststärke von .90 ergab eine benötigte Stichprobengröße von 171.

Die Durchführung des Feldexperiments erfolgte im 2. Schulhalbjahr 2012/2013 an einer städtischen Gesamtschule in Nordrhein-Westfalen. Die Teilnahme war für alle 173 SuS der Jahrgangsstufe 9 (70 Jungen, 103 Mädchen) als Bestandteil des regulären Unterrichts verpflichtend. Das Durchschnittsalter betrug zu Beginn des Trainings 14.79 Jahre ($SD = 0.676$). Der Anteil der SuS, die Zuhause hauptsächlich Deutsch sprechen, beträgt 59 Prozent (102 von 152). 21 SuS machten keine Angabe. Nach Aussage der Schule strebten die meisten SuS die gymnasiale Oberstufe an.

3.4.2.2 Design

Die einfaktorielle feldexperimentelle Trainingsstudie mit zwei Gruppen (Experimentalgruppe, EG, und Kontrollgruppe, KG) wurde mit einem Prä-Posttestdesign durchgeführt (Retroaktionsplan; Klauer, 2011), wobei ein unvollständiges balanciertes Multi-Matrix-Testheftdesign verwendet wurde (*incomplete balanced matrix booklet*; Gonzales & Rutkowski, 2010, siehe Instrumente). Der Prätest erlaubt Vorwissenseffekte, mögliche *Aptitude-Treatment-Interactions* (ATI-Effekte) und das kognitive Potenzial im Sinne der kognitiven Potenzialausschöpfungshypothese zu erfassen. Das EG-Treatment besteht aus einer Unterrichtseinheit zum Thema „Planen und Problemlösen“, das KG-Treatment aus einer Unterrichtseinheit zum „Präsentieren“ (siehe Materialien und Durchführung). Die SuS wurden innerhalb jeder der sechs Klassen computerbasiert randomisiert einer der beiden Bedingungen zugewiesen.

3.4.2.3 Instrumente

Planen

Zur Erfassung der Teilkomponente Planungsfähigkeit wurden als Prätest der Tour-Planer (Cronbachs $\alpha = .91$, 36 Items) und als Posttest der Routen-Planer (Cronbachs $\alpha = .92$, 36 Items) des Aachener Planungskompetenztrainings (Arling & Spijkers, 2012) eingesetzt (siehe Durchführung und Materialien).

Problemlösen

Leistungstests. Der Prä- und Posttest besteht aus einem Test zum fächerübergreifenden Problemlösen und einem Test zum mathematischen Problemlösen. Die Aufgaben zum fächerübergreifenden Problemlösen entstammen der PISA-Studie 2003 (OECD, 2004b). Die Aufgaben zur Mathematik entstammen der PISA-Studie 2003 (OECD, 2004b), den zentralen Lernstandserhebungen des Landes Nordrhein-Westfalen 2004/2005 im Fach Mathematik (Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2005b, 2005a; vgl. Leutner et al., 2007) und Pugalee (2001). Die Aufgaben wurden in einem balancierten unvollständigen Testheftdesign (Gonzales & Rutkowski, 2010) mit vier Clustern pro Domäne (Problemlösen, Mathematik) administriert.[16] Denn balancierte Testheftdesigns führen im Vergleich zu unbalancierten zu weniger Verzerrungen bei der Schätzungen von Populationsparametern (Frey & Bernhardt, 2012). Zur Vermeidung von Reihenfolgeeffekten und des Abschreibens wurde die Reihenfolge der Itemcluster variiert. Durch Rotation der Domänen und Rotation der Cluster innerhalb der Domäne entstanden insgesamt 12 Testheftversionen pro Testzeitpunkt (Tabelle 8). Das Linking der Testhefte erfolgt über Cluster 4 beider Domänen (Tabelle 8). Jedes Item kommt nur in einem Cluster vor (Anhang Y). Wer im Prätest Heft i erhielt, erhielt auch im Posttest Heft i, um Gedächtniseffekte auszuschließen ($i = 1, ..., 12$).

[16] Jens Fleischer, Christian Spoden und Prof. Dr. Martin Brunner danke ich für die Anmerkungen und Hilfen bei der Wahl des Testheftdesigns.

Tabelle 8. Testheftdesign: Verteilung der 4 Itemcluster des Problemlösens (PL) und der Mathematik (M) auf die 12 Testheftversionen beim Prä- und Posttest

Heftversionen Prätest											
Heft 1	*Heft 2*	*Heft 3*	*Heft 4*	*Heft 5*	*Heft 6*	*Heft 7*	*Heft 8*	*Heft 9*	*Heft 10*	*Heft 11*	*Heft 12*
PL1	M 1	PL2	M 2	PL3	M 3	PL1	M 1	PL2	M 2	PL3	M 3
PL2	M 2	PL1	M 1	PL1	M 1	PL3	M 3	PL3	M 3	PL2	M 2
M 1	PL1	M 2	PL2	M 3	PL3	M 1	PL1	M 2	PL2	M 3	PL3
M 2	PL2	M 1	PL1	M 1	PL1	M 3	PL3	M 3	PL3	M 2	PL2

Heftversionen Posttest											
Heft 1	*Heft 2*	*Heft 3*	*Heft 4*	*Heft 5*	*Heft 6*	*Heft 7*	*Heft 8*	*Heft 9*	*Heft 10*	*Heft 11*	*Heft 12*
PL3	M 3	PL4	M 4	PL2	M 2	PL4	M 4	PL1	M 1	PL4	M 4
PL4	M 4	PL3	M 3	PL4	M 4	PL2	M 2	PL4	M 4	PL1	M 1
M 3	PL3	M 4	PL4	M 2	PL2	M 4	PL4	M 1	PL1	M 4	PL4
M 4	PL4	M 3	PL3	M 4	PL4	M 2	PL2	M 4	PL4	M 1	PL1

Als weiterer Test wurde direkt nach dem Posttest des Planungskompetenztrainings ein kleiner Test zum fächerübergreifenden Problemlösen administriert (siehe Tabelle 10), um mögliche Transfereffekte zu testen. Dieser Papier-Bleistift-Test besteht aus der Beispielaufgabe *Birthday Party* der PISA-2012-Erhebung (Australian Council for Educational Research, n.d.), bei der eine Tischordnung unter Beachtung zulässiger Sitznachbarn geplant werden muss, und zwei selbst entwickelten Items. Letztere verlangen die Planung einer Speisenauswahl unter Berücksichtigung von Allergien bzw. Unverträglichkeiten der Gäste (Anhang L). Die Items waren jedoch unkorreliert (siehe Ergebnisse).

Strategiemaße. Um neben den Leistungsmaßen auch Auswirkungen auf die allgemein eingesetzten Strategien beim Problemlösen bzw. die Herangehensweise an Probleme erfassen zu können, wurde der *problem-solving situational judgement test* (PS-SJT; Bertling, 2012) eingesetzt. Der PS-SJT besteht aus drei Szenarien, zu denen je vier Handlungsmöglichkeiten (Items) gegeben sind, die bezüglich ihrer Eignung in der jeweiligen Problemsituation beurteilt werden müssen. Beim PS-SJT werden drei Arten von Herangehensweisen an Probleme unterschieden, die faktorenanalytisch gewonnen wurden (Bertling, 2012): *systematisches Problemlösen* (typisch für Kompetenzstufen 2 und 3 der Problemlösekompetenz bei PISA 2003), *Hilfesuchen* (typisch für Kompetenzstufe 1 der Problemlösekompetenz), *unsystematisches Problemlösen* (typisch für Problemlösen unterhalb von Kompetenzstufe 1). Die in Experiment 3 niedrigen Reliabiltäten der Skalen des PS-SJT betragen $\alpha = .61$ für *systematisches Problemlösen*, $\alpha = .53$ für *unsystematisches Problemlösen* und $\alpha = .56$ für *Hilfesuchen*. Auch in der PISA-

2012-Erhebung wies der PS-SJT große Reliabilitätsprobleme auf (D. Leutner, persönl. Mitteilung, 20.11.2014).

Drittvariablen

Zur Verwendung in der multiplen Imputation, d. h. zur Schätzung von fehlenden Werten, und für spätere Analysen wurden im Laufe des Schulhalbjahrs eine Reihe von weiteren Variablen erhoben, die im Zusammenhang mit mathematischer Leistung stehen. Tabelle 9 listet diese Drittvariablen zusammenfassend auf.

Tabelle 9. Erhobene Drittvariablen in Experiment 3

Konstrukt	*Skalen bzw. Fragebogen*
Schlussfolgern	*Figurenanalogien* (N2, Form B) und *Wortanalogien* (V3) des KFT4-12+R (Heller & Perleth, 2000)
Lern- und Leistungsmotivation	*Fragebogen zur Erfassung aktueller Lern- und Leistungsmotivation* (FAM; Rheinberg et al., 2001) in der Kurzversion von Freund, Kuhn und Holling (2011), *Skalen zur Erfassung der Lern- und Leistungsmotivation* (SELLMO; Spinath, Pelster-Stiensmeier, Schöne & Dickhäuser, 2002)
Anstrengungsbereitschaft	*Anstrengungsbereitschaft* (Deutsches PISA-Konsortium, 2006)
Freude an Denkaufgaben	*Need for Cognition* (Bless, Wänke, Bohner, Fellhauer & Schwarz, 1994)
Ängstlichkeit in Mathematik	*Ängstlichkeit in Mathematik* (Deutsches PISA-Konsortium, 2006)
Mathematikbezogene Selbstwirksamkeit	*Mathematikbezogene Selbstwirksamkeit* (Deutsches PISA-Konsortium, 2006)
Fachspezifisches Selbstkonzept	*DISK-Gitter* (Rost & Sparfeldt, 2002) für die Fächer Mathematik und Deutsch

Schlussfolgern. Um kognitive Voraussetzungen berücksichtigen zu können, wurden die Skalen *Figurenanalogien* (Skala *N2*; M = 14.97, SD = 6.35, Cronbachs α = .942, N = 148, 25 Items) und *verbale Analogien* (Skala V3; M = 9.17, SD = 3.63, Cronbachs α = .953, N = 132, 20 Items) des KFT 4-12+R (Heller & Perleth, 2000) eingesetzt, die schlussfolgerndes Denken als Aspekt der Intelligenz testen. Man beachte, dass es sich um neue Testhefte handelt, die nicht die beiden fehlerhaften Items der *N2*-Skala beinhalten, die Möller et al. (2006) und Segerer et al. (2012) aufgezeigt haben.

Lern- und Leistungsmotivation. Die Lern- und Leistungsmotivation der SuS ($N = 132$) wurde zu Beginn des Trainings mit den vier Skalen *Lernziele* ($M = 3.31$, $SD = 0.27$, Cronbachs $\alpha = .78$, 8 Items), *Vermeidungsleistungsziele* ($M = 2.34$, $SD = 0.31$, Cronbachs $\alpha = .86$, 8 Items), *Annäherungsleistungsziele* ($M = 2.83$, $SD = 0.43$, Cronbachs $\alpha = .75$, 7 Items), *Arbeitsvermeidung* ($M = 2.07$, $SD = 0.21$, Cronbachs $\alpha = .81$, 8 Items) der *Skalen zur Erfassung der Lern- und Leistungsmotivation* (SELLMO; Spinath et al., 2002) erhoben. Zusätzlich wurde im Posttest der *Fragebogen zur Erfassung aktueller Lern- und Leistungsmotivation* (FAM; Rheinberg et al., 2001) in der Kurzversion von Freund et al. (2011) erhoben ($M = 3.77$, $SD = 0.98$ Cronbachs $\alpha = .771$, 12 Items, $N = 136$). Zusätzlich wurde die *Anstrengungsbereitschaft* (Boekaerts & Otten, 1993; Deutsches PISA-Konsortium, 2006) erfragt ($M = 3.19$, $SD = 0.69$, Cronbachs $\alpha = .91$, 3 Items, $N = 136$).

Freude an Denkaufgaben. Zur Erfassung des Engagements und der Freude an Denkaufgaben wurde die Kurzversion des Fragebogens *Need for Cognition* von Bless et al. (1994) verwendet (Cronbachs $\alpha = .778$, 19 Items, $N = 114$). Die Skala besteht aus 19 Items, die auf einer 7-stufigen Antwortskala (von -3 = *völlig unzutreffend* bis 3 = *trifft ganz genau zu*) einzuschätzen sind ($M = 0.40$, $SD = 0.70$).

Ängstlichkeit und Selbstwirksamkeit im Fach Mathematik. Zur Erfassung der Mathematikangst wurde die Skala *Ängstlichkeit in Mathematik* (Deutsches PISA-Konsortium, 2006; Cronbachs $\alpha = .88$, 5 Items, $N = 148$) verwendet. Zur Erfassung mathematischer Selbstwirksamkeit wurde die Skala *Selbstwirksamkeit Mathematik* (Deutsches PISA-Konsortium, 2006) verwendet (Cronbachs $\alpha = .80$, 8 Items, $N = 145$).

Selbstkonzept. Zur Diagnostik des fachspezifischen Selbstkonzepts für die Fächer Mathematik (Cronbachs $\alpha = .93$, 8 Items, $N = 142$) und Deutsch (Cronbachs $\alpha = .93$, 8 Items, $N = 138$) wurde die jeweilige Skala des DISK-Gitters (Rost & Sparfeldt, 2002) eingesetzt.

3.4.2.4 Durchführung und Materialien

Ablauf. Die Trainingsstudie wurde im Schuljahr 2012/2013 an einer städtischen Gesamtschule in Nordrhein-Westfalen durchgeführt. Die Teilnahme war für die SuS als regulärer Bestandteil des Unterrichts am Vormittag verpflichtend, jedoch nicht versetzungsrelevant. Das Training fand einmal pro Woche für 90 Minuten statt. In Woche 1 erfolgte ein thematischer Überblick und ein demographischer Fragebogen wurde administriert (Tabelle 10). Den SuS wurde in Woche 1 mitgeteilt, dass alle Gruppen im Laufe des Schulhalbjahres dieselben Inhalte behandeln und dass sich aus organisatorischen Gründen (Nutzung des Computerraums) lediglich die Reihenfolge unterscheidet. Der Unterricht in Woche 2 entfiel aus schulinternen Gründen (Karnevalsferien). In Woche 3 bearbeiteten die SuS den Prätest zum fächerübergreifenden und mathematischen Problemlösen. Im Anschluss absolvierten die SuS zur Auflockerung eine Übung zum Präsentieren. In Woche 4 erfolgte der Prätest des Planungskompetenztrainings (Arling & Spijkers, 2012) und ein kleiner Transfertest zum Problemlösen. Die Experimentalgruppe bearbeitete im Anschluss Reflexionsübungen zum Planungskompetenztraining (Hoffmann, 2012), die Kontrollgruppe begann mit der mehrwöchigen Unterrichtseinheit zum Thema Präsentieren. In den Wochen 5 und 6 absolvierte die Experimentalgruppe das Planungskompetenztraining. In Absprache mit Dr. Arling wurde das Planungskompetenztraining in Partnerarbeit statt in Einzelarbeit durchgeführt und um Reflexionsübungen (Hoffmann, 2012) ergänzt. In Woche 7 bearbeiteten beide Gruppen den Posttest des Planungskompetenztrainings. Nach den Osterferien (Woche 8 und 9) wurden die Treatments zum Problemlösen (Experimentalgruppe) und Präsentieren (Kontrollgruppe) fünf Wochen lang fortgeführt. In Woche 15 erfolgte der Posttest zum fächerübergreifenden und mathematischen Problemlösen. Nach dem Posttest in Woche 15 absolvierte die Experimentalgruppe aus Gründen der schulischen Gleichbehandlung die Unterrichtseinheit zum Thema Präsentieren und die Kontrollgruppe die Unterrichtseinheit zum Thema Problemlösen. Am Schuljahresende wurde eine kurze summative Evaluation des gesamten Trainings durchgeführt, deren Ergebnisse bei Goertzen (2014) berichtet werden. Anhang Q enthält eine ausführlichere Übersicht zum Ablauf von Experiment 3.

Tabelle 10. Überblick über den Studienablauf von Experiment 3

Woche	*Experimentalgruppe*	*Kontrollgruppe*
1	Einführung und demographischer Fragebogen	
2	schulfrei	
3	Prätest fächerübergreifendes und mathematisches Problemlösen	
4	Prätest Planungskompetenz	
	Treatment: Reflexionsübung	Treatment Präsentieren
5-6	Treatment: Planungskompetenz-training	Treatment Präsentieren
7	Posttest Planungskompetenz, kleiner Transfertest Problemlösen	
8-9	schulfrei	
10-14	Treatment Problemlösen	Treatment Präsentieren
15	Posttest fächerübergreifendes und mathematisches Problemlösen	

Lehrpersonen. Die Unterrichtseinheiten wurden von bezahlten Lehramtsstudierenden durchgeführt, die über ein abgeschlossenes erziehungswissenschaftliches Grundstudium, Pflichtpraktika und zusätzliche pädagogische Vorerfahrungen (wie Jugendarbeit, Nachhilfe, Vertretungslehrer, Leitung von Arbeitsgemeinschaften, Testleitertätigkeiten) verfügten. Es erfolgte kein Wechsel der Lehrpersonen während des Schulhalbjahres, um stabile Lehrer-Schüler-Beziehungen zu ermöglichen. Pro Trainingsgruppe (13-15 SuS) wurden zwei Trainer eingesetzt – mit Ausnahme einer Klasse (zweier Trainingsgruppen). Dort wurde jeweils ein Trainer mit Erfahrungen im selbstständigen Unterrichten eingesetzt. Im Vorfeld der Durchführung wurden die Trainer geschult. Die Themen umfassten eine Testleiterschulung, das inhaltliche Konzept des jeweiligen Trainings sowie eine Einführung in *classroom management*.[17] Während der Durchführung erhielten die Lehrpersonen wöchentliche Verlaufspläne sowie regelmäßige Schulungen und Supervision.

Materialien. Zur Förderung der Teilkomponente Planungsfähigkeit wurden der Einkaufs-Planer und der Urlaubs-Planer des Aachener Planungskompetenztrainings (Arling & Spijkers, 2012) eingesetzt, die im Gegensatz zum Manual und nach Absprache mit Dr. Arling in Partnerarbeit statt in Stillarbeit bearbeitet wurden, um die Attraktivität des (ursprünglich für erwachsene Rehabilitanden konzipierten) Verfahrens für SuS zu erhöhen. Die Trainingssitzungen, die aus

[17] Ich danke Dr. Theresa Dicke für ihre erprobten Schulungsmaterialien zum *classroom management*!

Tagungsplanungsaufgaben bestehen (z. B. Planung verschiedener Einkäufe unter Berücksichtigung von Nebenbedingung wie Öffnungszeiten, Pausen und Geld) wurden zudem um eine Reflexionsübung nach dem Prätest (Tour-Planer, Arling, 2006) und eine Transferübung nach dem Urlaubs-Planer ergänzt (Hoffmann, 2012, modifiziert). Was wird beim Planungskompetenztraining (Arling & Spijkers, 2012) trainiert, das einen Transfer auf Problemlöseaufgaben erwarten lässt? Auf einer Makroebene ist Planen ein wesentlicher Bestandteil des Problemlöseprozesses (vgl. Davidson et al., 1994; Pólya, 1945, 1973). Daher sollte bessere Planung im Prinzip zu besseren Problemlösungen führen. Theoretisch lassen sich die Aufgaben des Planungskompetenztrainings auch als umfangreiche nichtfachliche Problemlöseszenarien auffassen, dessen Ziel in der Planung eines Tagesablaufs unter Berücksichtigung von Nebenbedingungen liegt (vgl. das *Plan-a-Day*-Szenario; Funke & Krüger, 1995, bzw. Rodewald et al., 2011). Auf einer Mikroebene werden beim Planungskompetenztraining von Arling und Spijkers (2012) folgende Anforderungen gestellt und geübt, die beim Problemlösen ebenfalls wichtig sein können:

- *Motivation und Volition.* Um Probleme zu lösen, darf man nicht sofort aufgeben, sondern muss sich auch herausfordernden Problemen bzw. Arbeitsanforderungen stellen und diese über einen längeren Zeitraum verfolgen (Ausdauer, Persistenz).
- *Anwenden von Arbeitstechniken und Haltungen* (vgl. Leuders, 2010) bzw. strukturiertes Arbeiten im Umgang mit anspruchsvollem Arbeitsmaterial wie das Markieren von Texten (eher implizit), das explizit empfohlene Verwenden von Tabellen (was SuS selten selbstständig tun; vgl. Törner & Zielinski, 1992) und das explizit empfohlene Anfertigen von Notizen.
- *Zwischenzielsetzung* (z. B. durch Selbstinstruktion: „Erstelle erst die Planung bis zur Mittagspause und plane dann den weiteren Tag!“).
- *Berücksichtigung zahlreicher Nebenbedingungen* (*constraints*), die *Planen, Planüberwachung* und *Kontrollprozesse* erfordern (Arling & Spijkers, 2012).
- *Metakognitive Reflexion des Bearbeitungsprozesses* durch Übungen, bei denen die SuS in Partnerarbeit erarbeiten sollen, was ihnen bei der Bearbeitung leicht und was schwer gefallen ist, welche Strategien sie eingesetzt

haben, welche Lösungsideen sie haben, welche Strategien sie im Alltag (Hoffmann, 2012) und welche sie beim Lösen mathematisch-naturwissenschaftlicher Aufgaben einsetzen können.

Zur Förderung weiterer Komponenten der Problemlösekompetenz wurde ein breiter Ansatz gewählt (siehe Tabelle 11 für Details). Er umfasst sowohl kognitive Elemente (Übungsaufgaben zum Planen und Problemlösen) als auch metakognitive Elemente wie Strategien, Heuristiken, metakognitive Fragen (Blum et al., 2006; Clement & Konold, 1989; King, 1991). Kognitive Trainings, die Strategien und Heuristiken vermitteln und als Transfervehikel nutzen, gelten als sehr effektiv (Hasselhorn & Gold, 2006). Wie in Experiment 1 und 2 wurden teilweise Aufgaben verwendet, die auf die Teilkomponenten Planungsfähigkeit (Arling & Spijkers, 2012; Hoffmann, 2012), Handlungswissen (Buchwald, 2012a) und konditionales Wissen (Buchwald, 2012b, 2012c; Keller, 2005) abzielen.

Das Kontrollgruppentreatment besteht aus rhetorischen Übungen und Arbeitsblättern zu Lampenfieber (Klippert, 2008; Nöllke & Schmettkamp, 2011), Körpersprache (Brüning, 2011; Vlcek, 2000) und Kriterien guten Präsentierens (Berthold, Diehl & Kühne, 2010; Klippert, 2008; Mattes, 2011) mit einer anschließenden Einführung in Microsoft PowerPoint inklusive dem Vortragen der erarbeiteten Präsentationen.

Im Anschluss an den Posttest absolvierte jede Gruppe aus schulischen Gründen der Gleichbehandlung die jeweils andere Unterrichtseinheit in leicht gekürzter Form. Um motivationale Effekte der Gruppeneinteilung zu minimieren, wurde den SuS am Beginn des Schuljahres mitgeteilt, dass sich die Trainingsgruppen aus organisatorischen Gründen (Kapazität des Computerraums) nur in ihrer Reihenfolge unterscheiden. Am Ende des gesamten Trainings wurde eine kurze summative Evaluation durchgeführt (siehe Goertzen, 2014).

Tabelle 11. Auflistung der Trainingsinhalte zum Problemlösen

Trainingsinhalte zum Problemlösen
- Merkblätter und Tipps zum Problemlösen (Clement & Konold, 1989, King, 1991, vgl. auch Teong, 2003).
- Übung *Eierflugmodell* (Anhang S; vgl. Gilsdorf & Kistner, 1998; Weidenmann, 2008), die Planen, Problemlösen und Präsentieren integriert. Bei dieser Aufgabe gab es eine explizite Planungsphase, um an das Planungskompetenztraining (Arling & Spijkers, 2012) anzuknüpfen, und eine Problemlösephase.
- Komponentenspezifische Übungen zum konditionalen Wissen, z. B. alternative Nutzungsweisen generieren (Keller, 2005, vgl. Bruder, 2002, 2003, Januar, auch um die Nützlichkeit schriftlichem Vorgehens zu verdeutlichen), Beurteilung von Handlungsoptionen (Buchwald, 2012c; Anhang W), Vergleich von Lösungen und Strategien (Buchwald, 2012b; Anhang R) sowie Handlungswissen (u. a. Umgang mit Entscheidungs-diagrammen; Buchwald, 2012a).
- Wasserumfüllaufgabe (Luchins, 1942), um Nutzen und Grenzen von Strategien zu zeigen (vgl. den Darmstädter Ansatz; Bruder et al., 2002; Gürtler, Perels, Schmitz & Bruder, 2002).
- Komplexes Problemlösen (Systematisches Probieren und VOTAT-Strategie) mit dem Computerprogramm *Genetics Lab* (Sonnleitner et al., 2012), das nur zur Übung und nicht als Test eingesetzt worden ist.
- Materialien und Übungen zu den in den Kernlehrplänen des Landes Nordrhein-Westfalen (Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2007, 2011, Ministerium für Schule, Jugend und Kinder des Landes Nordrhein-Westfalen, 2004a, 2004b, 2004c) geforderten Problemlöseheuristiken im Fach Mathematik als *Transfervehikel*:
- Übersicht mathematisch-heuristischer Hilfsmittel (Analogieprinzip, Zerlegungsprinzip, Veranschaulichen, Vorwärtsarbeiten, systematisches Probieren, Rückwärtsarbeiten; Blum et al., 2006, Anhang T),
- systematisches Probieren mit Übungen (Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2010a, Anhang U)
- Arbeitsblatt zum Vorwärtsarbeiten[a] (Arbeitsgruppe der Fachdidaktik der Mathematik der TU Darmstadt, n.d.; Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2010b)
- Ergänzungsaufgaben zum Problemlösen[b]
- Hobbits-und-Orks-Problem (Thomas, 1974)
- „Was ist in welcher Flasche?“ (Gärtner & Scharf, 2001, modifiziert)

Anmerkungen.
[a] Didaktischer Kommentar: Auch wenn dieses Arbeitsblatt für Klasse 5 vorgesehen ist, stellten die Aufgabe die SuS der Klasse 9 nach Aussage der Trainer teilweise vor Herausforderungen. [b] Die Ergänzungsaufgaben waren optional am Stundenende und kamen aus Zeitgründen nicht in allen Gruppen zum Einsatz.

3.4.3 Datenanalyse

3.4.3.1 Skalierung des Prä- und Posttests der Domänen Problemlösen und Mathematik

Die Problemlöse- und Mathematikitems aus PISA 2003 und der Lernstandserhebung 2004/2005 in Nordrhein-Westfalen wurden entsprechend der jeweiligen Musterlösungen ausgewertet (Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2005d, 2005c; OECD, 2004b). Anschließend erfolgte eine IRT-Skalierung der Multi-Matrix-Testheftdaten mit R (R Core Team, 2014) und dem R-Paket *TAM* (Kiefer, Robitzsch & Wu, 2014), einer freien Alternative zu *Conquest* (Wu, Adams, Wilson & Haldane, 2007). Dabei wird durch Linking sichergestellt, dass sich die Personenparameter von Prätest und Posttest pro Domäne auf derselben Metrik befinden (Davier, 2011; De Ayala, 2009). Für Details dieser Skalierung siehe Anhang CC. Als Schätzer der jeweiligen Personenfähigkeit für die weiteren Analysen werden *Weighted-Likelihood*-Personenparameterschätzungen (WLE; Warm, 1989) verwendet.

3.4.3.2 Umgang mit fehlenden Werten

Bei Längsschnittstudien kommt es fast immer zu fehlenden Werten, da Personen nicht an allen Testterminen anwesend sind. Diese fehlenden Werte sollten durch Schätzungen ersetzt werden (Lüdtke, Robitzsch, Trautwein & Köller, 2007). Tabelle 12 gibt eine Übersicht über die Anzahl der anwesenden SuS beim Prä- und Posttest der Domänen Problemlösen und Mathematik.

Tabelle 12. Anwesenheit der SuS beim Prä- und Posttest der Domänen Problemlösen und Mathematik

Anwesenheit beim Prätest Problemlösen und Mathematik	*Anwesenheit beim Posttest Problemlösen und Mathematik*	*Anzahl SuS*
Ja	Ja	114
Ja	Nein	23
Nein	Ja	22
Nein	Nein	14

Da 14 SuS beim Prä- und Posttest nicht anwesend waren und zudem fehlende Werte auf Drittvariablen auftraten, wurde anschließend eine multiple Imputation der Skalenwerte vorgenommen, um fehlende Werte zu ersetzen (Graham, 2012).

Um fehlende Werte zu imputieren, wurden hypothesenrelevante Variablen und korrelierende Drittvariablen verwendet. Die multiple Imputation (MI) wurde mit R (R Core Team, 2014) und dem R-Paket *mice* (van Buuren & Groothuis-Oudshoorn, 2011) nach dem Vorgehen von van Buuren und Groothuis-Oudshoorn (2011) durchgeführt (siehe Anhang DD für Details). Dabei wurden die fehlenden Werte der interessierenden Leistungsvariablen graphisch mit dem R-Paket *VIM* (Templ, Alfons, Kowarik & Prantner, 2013) überprüft und abgeleitete Variablen wie Differenzvariablen und Interaktionsterme passiv imputiert. Es wurden 20 Datensätze imputiert, da die verbreitete Anzahl von 5 Imputationen nicht mehr empfohlen wird (Graham, Olchowski & Gilreath, 2007; Robitzsch, 2014b). *Trace line plots* sprechen für eine gute Qualität der Konvergenz des MI-Algorithmus (siehe Anhang DD).

3.4.3.3 Problemlösen und Mathematik

Zur Beantwortung der Fragestellungen 1b, 2 und 3 wurden mit R (R Core Team, 2014) und dem R-Paket *mice* (van Buuren & Groothuis-Oudshoorn, 2011) hierarchische Regressionen pro imputierten Datensatz berechnet und die Ergebnisse der Analysen aggregiert (gepoolt). Die Gruppenzugehörigkeit war dummy-kodiert (0 = *Kontrollgruppe*, 1 = *Experimentalgruppe*). Zusätzlich werden aufgrund der zu erwartenden Korrelation der Prädiktoren Problemlösen und Mathematik eine Kommunalitätenanalyse mit dem R-Paket *yhat* (Nimon, Lewis, Kane & Haynes, 2008) und eine Analyse der relativen Wichtigkeit mit dem R-Paket *relaimpo* (Grömping, 2006) berechnet (Böhme, Robitzsch & Busé, 2010; Kraha, Turner, Nimon, Zientek & Henson, 2012).

3.4.3.4 Wille zur Aufgabenbearbeitung (Bearbeitungsquote)

Eine geringe Anzahl ausgelassener Aufgaben bzw. fehlender Werte stellt einen Indikator für Anstrengung und Trainingserfolg dar (Ho & Hedberg, 2005). Es gibt jedoch verschiedene Gründe fehlender Werte (z. B. ungültige Antworten, ausgelassene Items, nicht-erreichte Items, missing by design; Duchhardt & Gerdes, 2013, April). Für die Analyse bedeutet dies, dass folgendermaßen festgelegt wird, wann eine Aufgabe als nicht-bearbeitet zählt: Da genügend Bearbeitungszeit für die Testaufgaben zur Verfügung stand – die meisten SuS gaben ein paar Minuten vor Ende der Bearbeitungszeit ihr Testheft ab – werden sowohl nicht-erreichte Items als auch ausgelassene Items als nichtbearbeitet gezählt. Ungültige Antworten stellen in der Regel einen Bearbeitungsversuch dar und werden daher als bearbeitet gezählt. Von der Analyse ausgeschlossen werden fehlende Werte, die dem Testheftdesign geschuldet sind (*missing by design*) und fehlende Werte von Personen, die nicht am Test teilgenommen haben. Als invers-gepolter Indikator für Trainingserfolg wird daher pro Aufgabe im Posttest die Anzahl an SuS betrachtet, die eine Aufgabe nicht bearbeitet haben, hier definiert als Summe der SuS mit der Kodierung *nicht-erreicht* und *ausgelassen*. Man beachte, dass dieses Maß eher streng ist, da ausschließlich Aufgeschriebenes Grundlage der Analyse ist und es SuS geben kann, die zwar über eine Aufgabe nachgedacht haben, aber nichts aufgeschrieben haben. Somit wird die reale Anzahl an Nicht-Bearbeitern wahrscheinlich überschätzt.

3.4.3.5 Problem-Solving Situational Judgement Test

Die Bildung der drei Skalen *systematisches Problemlösen*, *unsystematisches Problemlösen* und *Hilfesuchen* des Problem-Solving Situational Judgement Tests (PS-SJT; Bertling, 2012) erfolgte als Summenscore über die jeweils vier Items, die beim PS-SJT auf einem Faktor laden. Die Reliabilität der Skalen war nicht zufriedenstellend (*systematisches Problemlösen*: $\alpha = .61$, *unsystematisches Problemlösen*: $\alpha = .53$ und *Hilfesuchen*: $\alpha = .56$). Angesichts der Kürze der Skalen konnte die interne Konsistenz durch Itemausschluss nicht weiter erhöht werden.

3.4.4 Ergebnisse

3.4.4.1 Fragestellung 6: Treatment-Check Planungskompetenz

Lässt sich mit dem ursprünglich für Erwachsene konzipierten Aachener Planungskompetenztraining (Arling & Spijkers, 2012) die Planungskompetenz (als eine Teilkomponente des Problemlösens) in der Jahrgangsstufe 9 steigern? Aufgrund der Berichte der Trainer wurde die Durchführung des Planungskompetenztrainings in zwei Klassen als *nicht manualkonform* eingestuft: SuS einer Klasse brachen den Prätest frühzeitig ab[18], wodurch sich niedrige Prätestwerte (Tabelle 13) und damit artifiziell höhere Punktgewinne von Prä- zu Posttest des Planungskompetenztrainings ergeben hätten (Abbildung 18). In einer Experimentalgruppe wurde das Planungskompetenztraining nicht mit allen Reflexionsübungen und

[18] Die SuS berichteten, ihre Klassenlehrerin hätte Ihnen gesagt, die SuS müssten nicht mitmachen, wenn sie keine Lust hätten.

Tipps durchgeführt. Aus Gründen guter statistischer Praxis (Simmons, Nelson & Simonsohn, 2011) werden die Ergebnisse aller Gruppen berichtet.

Tabelle 13. Erreichte Punkte im Prä- und Posttest des Planungskompetenztrainings (Maximum: 38)

	Prätest			*Posttest*		
Gruppe	*M*	*SD*	*n*	*M*	*SD*	*n*
EG (normal)	12.31	8.37	45	17.62	8.69	37
KG (normal)	12.02	9.16	45	13.32	10.38	41
EG (nicht manualkonform)	5.48	4.81	25	12.83	7.80	23
KG (nicht manualkonform)	5.54	6.15	26	15.88	10.56	16
Gesamtsumme	9.76	8.33	141	14.93	9.53	117

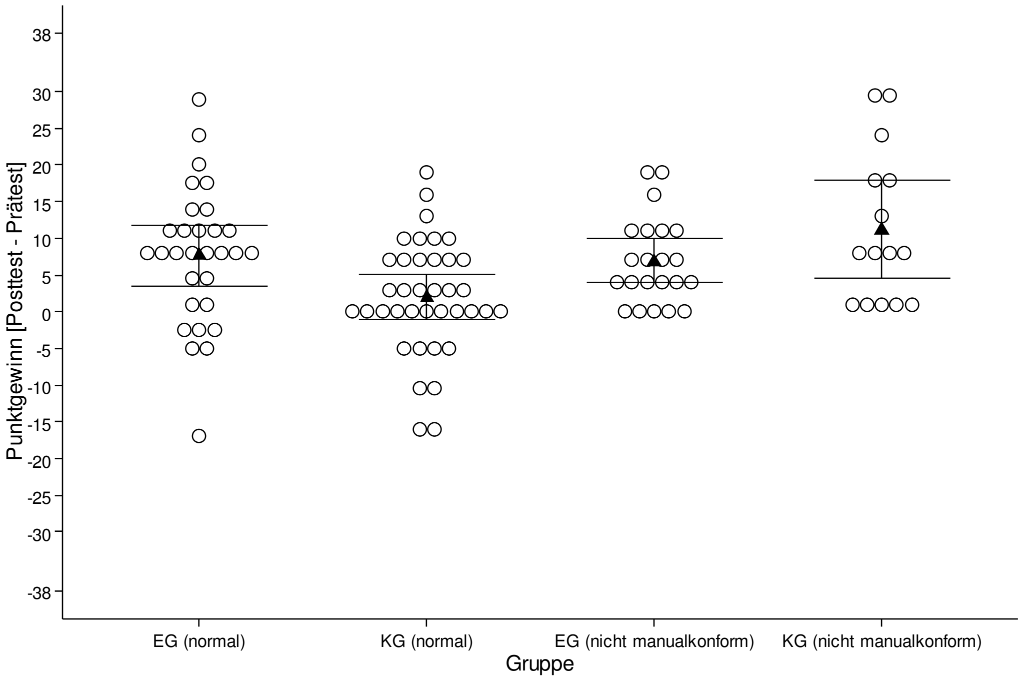


Abbildung 18. Punktgewinn im Planungskompetenztraining der Experimentalgruppe (EG) und Kontrollgruppe (KG)

Anmerkung. Das schwarze Dreieck kennzeichnet den Mittelwert. Fehlerbalken repräsentieren das 99%-Bootstrap-Konfidenzintervall des Mittelwerts (mit 1000 Replikationen). Die Abbildung wurde mit dem R-Paket *ggplot 2* (Wickham, 2009) erzeugt.

Deskriptiv (vgl. Abbildung 19) erzielte die Experimentalgruppe, die gemäß Manual trainierte, beim Posttest des Planungskompetenztrainings (Routen-Planer) mehr Punkte (M = 17.62, SD = 8.69, n = 37) als die Kontrollgruppe (M = 13.32, *SD* = 10.38, *n* = 41).[19] Die Experimentalgruppe, die nicht manualkonform trainierte, schnitt deskriptiv schlechter ab (*M* = 12.83, *SD* = 7.82, *n* = 23) als die Kontrollgruppe, die nicht manualkonform trainierte (*M* = 15.88, *SD* = 10.58, *n* = 16). Zur Prüfung der Fragestellung 6 wurde eine ANCOVA mit der Bedingung (Experimentalgruppe vs. Kontrollgruppe) und der Manualkonformität (normal vs. abweichend) als Faktoren, der Leistung im Posttest des Planungskompetenztrainings (Routen-Planer) als abhängiger Variable und der Leistung im Prätest des Planungskompetenztrainings (Tour-Planer) als Kovariate berechnet. Die Kovariate war signifikant ($F(1, 98) = 36.58$, $p < .001$, $\eta^2 = .272$), ebenso die Interaktion der Faktoren ($F(1, 98) = 8.11$, $p = .005$, $\eta^2 = .076$). Die Haupteffekte der Bedingung ($F(1, 98) = .63$, $p = .43$, $\eta^2 = .006$) und der Manualkonformität ($F(1, 98) = .17$, $p = .68$, $\eta^2 = .002$) waren nicht signifikant. Die Ergebnisse zeigten, dass in den Klassen, die wie im Manual vorgesehen trainiert hatten, die Experimentalgruppe besser beim Posttest des Planungskompetenztrainings (Routen-Planer) abschnitt als die Kontrollgruppe (Cohens $d = 0.45$).

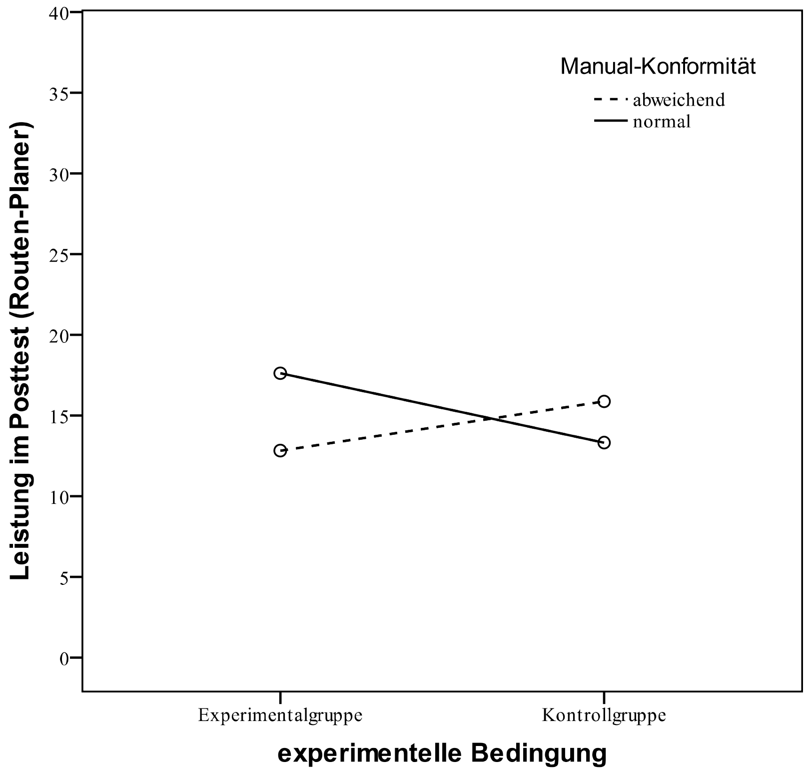


Abbildung 19. Leistung im Posttest des Planungskompetenztrainings

Anmerkung. Als Kovariate wurde die Prätestleistung (Tour-Planer) berücksichtigt (*M* = 9.72).

---

[19] Vielen Dank an Kirsten Breuer, die einige Eingabe- und Übertragungsfehler der Trainer gefunden und beseitigt hat!

Weitere Befunde zum Planungskompetenztraining

Als kleiner Transfertest wurden nach dem Posttest des Planungskompetenztrainings drei Items zum fächerübergreifenden Problemlösen präsentiert. Maximal waren in diesem Transfertest 5 Punkte erreichbar. Die Experimentalgruppe ($M = 2.28$, $SD = 0.98$) erreichte deskriptiv etwas mehr Punkte bei geringerer Streuung als die Kontrollgruppe ($M = 2.24$, $SD = 1.16$). Der Unterschied war nicht signifikant ($t(1, 108) = .17$, $p = .862$, Cohens $d = 0.03$).[20] Mit einer Reliabilität von Cronbachs $\alpha = .083$ lässt sich allerdings nicht von einer Skala sprechen. Korrelationsanalysen zeigten, dass die drei Items unkorreliert waren ($r_{12}= .06$, $p_{12}= .29$; $r_{13}= -.02$, $p_{13}=.43$; $r_{23}= .09$, $p_{23}= .18$; jeweils einseitige Signifikanztestung). Daher wurde zusätzlich die Leistung pro Item betrachtet (Tabelle 14). Die Experimentalgruppe löste Item 1 schlechter als die Kontrollgruppe, Item 2 besser und bei Item 3 bestanden keine Unterschiede. Es war also kein Muster erkennbar. Auf weitere Signifikanztestungen wurde daher verzichtet.

Tabelle 14. Deskriptivstatistik des kleinen Transfertests nach dem Planungskompetenztraining getrennt für Experimentalgruppe (EG) und Kontrollgruppe (KG)

		Item 1		*Item 2*		*Item 3*	
		EG	*KG*	*EG*	*KG*	*EG*	*KG*
Erreichte Punktzahl	*0*	40	30	4	9	36	29
	1	10	19	11	12	17	24
	2	nicht möglich		42	32	4	0
	M	0.20	0.38	1.68	1.51	0.44	0.49
	SD	0.41	0.49	0.59	0.74	0.64	0.51

Verglich man die Leistung der SuS beim Prätest des Planungskompetenztrainings in Experiment 3 mit den Leistungen einer von Arling und Spijkers (2012) untersuchten erwachsenen Rehabilitanden-Stichprobe ($M = 21.9$, $SD = 7.3$, $N = 775$), zeigte sich, dass die SuS im Mittel weniger Punkte erzielten ($M = 9.76$) und die Streuung ($SD = 8.33$) bei ihnen etwas größer war (Tabelle 13; Buchwald, Arling, Fleischer & Leutner, 2013). Bodeneffekte beim Planungskompetenztraining waren in Experiment 3 – Teilnahmemotivation vorausgesetzt – nicht zu erkennen. Für eine Analyse der Bearbeitungsspuren des Planungskompetenztrainings im Rahmen von Experiment 3 siehe Breuer (2014).

[20] Das Berücksichtigen der Manualkonformität als zusätzlichem Faktor ändert nichts an den Ergebnissen.

3.4.4.2 Fragestellung 7: Transfer Problemlösen

Steigert das fächerübergreifende Problemlösekompetenztraining mit Fokus auf spezifische Komponenten (u. a. Planungsfähigkeit, konditionales Wissen) die allgemeine fächerübergreifende Problemlösefähigkeit? Zur Beantwortung der Fragestellung wurden multiple Regressionen, eine Kommunalitätenanalyse mit dem R-Paket *yhat* (Nimon et al., 2008) und eine Analyse der relativen Wichtigkeit der Prädiktoren (Grömping, 2006) berechnet, da multiple Regressionen alleine nicht ausreichen, um die Korreliertheit der Prädiktoren (Tabelle 15) zu berücksichtigen.

Multiple Regression

Um den Einfluss der Prätestleistung der Domänen Problemlösen (PL.T1) und Mathematik (M.T1) sowie der Gruppenzugehörigkeit (Gruppe; 0 = Kontrollgruppe, 1 = Experimentalgruppe) und mögliche Interaktionseffekte (PL.T1 x M.T1, PL.T1 x Gruppe, M.T1 x Gruppe, PL.T1 x M.T1 x Gruppe) zu berücksichtigen, wurde mit dem R-Paket *mice* (van Buuren & Groothuis-Oudshoorn, 2011) eine gepoolte hierarchische multiple Regression mit der Leistung im Posttest Problemlösen (PL.T2) als Kriterium berechnet (Tabelle 16). Es wurden zunächst schrittweise die Prätestvariablen für Problemlösen und Mathematik sowie die Gruppenzugehörigkeit (0 = Kontrollgruppe, 1 = Experimentalgruppe) in die Analyse aufgenommen. Anschließend wurden in den Modellen 4 bis 7 mögliche Interaktionseffekte geprüft.

In Modell 1 wurde die Leistung im Posttest Problemlösen (PL.T2) vorhergesagt durch die Leistung im Prätest Problemlösen (PL.T1). Die Varianzaufklärung betrug 10.15 % und war signifikant, $F(1, 115.1308) = 12.659, p < .001$. Die Hinzunahme der Leistung im Prätest Mathematik (M.T1) in Modell 2 erklärte zusätzlich 4.10 % der Varianz, und diese Veränderung in R^2 war signifikant, $F(1,136.9765) = 5.258, p = .023$. Model 1 und 2 weisen erneut auf die Bedeutung des Vorwissens hin. Die Hinzunahme der Gruppenzugehörigkeit (KG = 0, EG = 1) in Modell 3 erklärte zusätzlich 1.62 % der Varianz, und diese Veränderung in R^2 war nicht signifikant, $F(1, 99.31026) = 1.728, p = .192$. Die Hinzunahme der Interaktion der Prätestleistung der beiden Domänen in Modell 4 erklärte zusätzlich 0.26 % der Varianz, und diese Veränderung in R^2 war nicht signifikant, $F < 1, p = .919$. Die Hinzunahme der Interaktion der Prätestleistung Problemlösen und Gruppenzugehörigkeit in Modell 5 erklärte zusätzlich 1.82 % der

Varianz, und diese Veränderung in R^2 war nicht signifikant, $F(1, 82.6795) = 1.875, p = .175$. Die Hinzunahme der Interaktion der Prätestleistung Mathematik und Gruppenzugehörigkeit in Modell 6 erklärte zusätzlich 0.58 % der Varianz, und diese Veränderung in R^2 war nicht signifikant, $F < 1, p = .480$. Die Hinzunahme der Dreifachinteraktion von Prätestleistung Problemlösen, Prätestleistung Mathematik und Gruppenzugehörigkeit in Modell 6 erklärte zusätzlich 0.47 % der Varianz, und diese Veränderung in R^2 war nicht signifikant, $F < 1, p = .806$. Insgesamt klärten die 7 Prädiktoren 19.00 % der Varianz im Posttest Problemlösen auf.

Tabelle 15. Korrelation wichtiger Variablen in Experiment 3

Nr.	Variable	1	2	3	4	5	6	7	8
1	Gruppe (0 = KG, 1 = EG)								
2	Prätest Problemlösen (PL.T1)	-.216**							
3	Prätest Mathematik (M.T1)	.015	.383***						
4	Posttest Problemlösen (PL.T2)	.066	.319***	.304**					
5	Posttest Mathematik (M.T2)	-.027	.320**	.404**	.336***				
6	PL.T1 x M.T1	-.206*	.544***	-.568***	.007	-.083			
7	Gruppe x PL.T1	.004	.766***	.222**	.175*	.122	.481***		
8	Gruppe x M.T1	.116	.231**	.659***	.131	.210*	-.393***	.334***	
9	Gruppe x PL.T1 x M.T1	-.091	.515***	-.331***	.055	-.058	.760***	.647***	-.505***

Anmerkungen. Dargestellt sind mit dem R-Paket *miceadds* (Robitzsch, 2014a) gemittelte Korrelationen über die imputierten Datensätze.
* $p < .05$, ** $p < .01$, *** $p < .001$

Tabelle 16. Gepoolte hierarchische multiple Regression über die 20 imputierten Datensätze mit der Problemlöseleistung im Posttest als Kriterium

						95 % CI				
	B	*SE*	*t*	*df*	*p*	UG	OG	NA	FMI	λ
Modell 1[a]										
P	.344	.097	3.558	66.033	.001	.151	.537	36	.361	.342
Modell 2[b]										
P	.256	.101	2.535	7.015	.013	.055	.458	36	.342	.323
M	.225	.098	2.293	72.719	.025	.029	.421	36	.330	.312
Modell 3[c]										
P	.290	.107	2.709	6.612	.009	.076	.504	36	.386	.366
M	.211	.099	2.130	7.286	.037	.013	.408	36	.339	.321
G	.307	.234	1.315	59.77	.194	-.160	.775	0	.390	.370
Modell 4[d]										
P	.290	.108	2.692	6.618	.009	.074	.505	36	.385	.365
M	.208	.107	1.935	7.899	.057	-.006	.422	36	.335	.317
G	.307	.234	1.313	6.840	.194	-.160	.774	0	.384	.364
P x M	.009	.084	.102	65.759	.919	-.160	.177	36	.359	.340
Modell 5[e]										
P	.482	.188	2.568	43.943	.014	.104	.860	36	.488	.466
M	.193	.108	1.789	69.117	.078	-.022	.408	36	.342	.324
G	.405	.257	1.577	48.146	.121	-.111	.921	0	.458	.436
P x M	-.007	.083	-.080	72.966	.937	-.171	.158	36	.325	.307
P x G	-.292	.213	-1.369	52.393	.177	-.720	.136	36	.431	.409
Modell 6[f]										
P	.444	.192	2.309	47.367	.025	.057	.831	36	.463	.441
M	.270	.145	1.856	86.310	.067	-.019	.559	36	.271	.254
G	.406	.256	1.587	48.401	.119	-.109	.922	0	.456	.434
P x M	-.022	.084	-.262	79.389	.794	-.189	.145	36	.298	.280
P x G	-.232	.230	-1.010	53.927	.317	-.693	.229	36	.420	.399
M x G	-.146	.206	-.709	71.187	.481	-.556	.264	36	.332	.313
Modell 7[g]										
P	.443	.196	2.263	45.493	.028	.049	.837	36	.475	.453
M	.255	.166	1.538	71.087	.128	-.076	.587	36	.331	.313
G	.430	.257	1.676	6.026	.099	-.083	.944	0	.385	.365
P x M	.001	.126	.006	65.025	.995	-.251	.252	36	.359	.340
P x G	-.229	.233	-.981	52.136	.331	-.697	.239	36	.430	.409
M x G	-.127	.231	-.549	59.116	.585	-.589	.336	36	.390	.370
P x M x G	-.047	.189	-.247	49.749	.806	-.426	.333	36	.446	.424

Anmerkungen. P = Prätest Problemlösen, M = Prätest Mathematik, G = Gruppe, CI = Konfidenzintervall, UG = Untere Grenze. OG = Obere Grenze, NA = Anzahl fehlender Werte, FMI = fraction of missing information[21], λ = Proportion of the variation attributable to the missing data. Varianzaufklärung [%]:
[a] R^2 = 10.15, ΔR^2 = 10.15***. [b] R^2 = 14.25, ΔR^2 = 4.10*.
[c] R^2 = 15.87, ΔR^2 = 1.62. [d] R^2 = 16.12, ΔR^2 = 0.26. [e] R^2 = 17.79, ΔR^2 = 1.82.
[f] R^2 = 18.53, ΔR^2 = 0.58. [g] R^2 = 19.00, ΔR^2 = 0.47.
* $p < .05$, *** $p < .001$.

[21] FMI entspricht im Allgemeinen nicht dem prozentualen Anteil fehlender Daten (Rubin, 1987, nach Schafer & Olsen, 1998). Es ist ein Maß für die Unsicherheit der Schätzung der imputierten Daten und erlaubt den Vergleich verschiedener Imputationsmodelle (Wagner, 2010).

Kommunalitätenanalyse

Eine Kommunalitätenanalyse mit dem R-Paket *yhat* (Nimon et al., 2008) auf Basis der imputierten Daten zeigte, dass die Vorwissensvariablen (Problemlösen, Mathematik, Problemlösen x Mathematik) den größten Anteil der aufgeklärten Gesamtvarianz erklären. Die gemeinsamen Varianzanteile sind dabei etwas größer als die uniquen Varianzanteile (Tabelle 17). Der unique Varianzanteil der Gruppenzugehörigkeit liegt mit 2.40 % deskriptiv über dem uniquen Varianzanteil der Mathematikleistung (1.91 %) und ist jeweils als klein zu beurteilen ($f^2 = .03$). Der negative Anteil aufgeklärter Varianz weist auf einen kleinen Supressoreffekt der Gruppenvariable hin, der zustande kommt, da ohne Gruppenvariable die ATI-Effekte nicht berücksichtigt werden können. Der totale Anteil der aufgeklärten Varianz der Interaktion von Problemlösen und Treatment ist klein (3.27 %, $f^2 = .03$).

Tabelle 17. Ergebnisse der Kommunalitätenanalyse mit der Posttestleistung Problemlösen als Kriterium

Prädiktor	*Unique [%]*	*Common [%]*	*Total [%]*
Problemlösen T1	4.81	5.52	10.33
Mathematik T1	1.91	7.52	9.44
Gruppe	2.40	-1.72	0.68
Problemlösen T1 x Mathematik T1	0.28	2.20	2.48
Problemlösen T1 x Gruppe	1.07	2.20	3.27
Mathematik T1 x Gruppe	0.51	1.46	1.96
Problemlösen T1 x Mathematik T1 x Gruppe	0.45	-0.10	0.35

Anmerkungen. T1 = Prätest. Die Kommunalitätenanalyse wurde mit dem R-Paket *yhat* (Nimon et al., 2008) berechnet. Man beachte, dass sich die Summe der Spalte „Total“ nicht zu R^2 summiert, da es gemeinsame Varianzanteile der einzelnen Prädiktoren gibt (Kraha et al., 2012).

Analyse der relativen Wichtigkeit

Ergänzend wurde eine Analyse der relativen Wichtigkeit der Prädiktoren vorgenommen (Grömping, 2006, 2007). Da das R-Paket *relaimpo* (Grömping, 2006) in der aktuellen Version 2.2-2 noch keine Dreifachinteraktionen berücksichtigen kann, wird zumindest für Modell 6 (siehe Tabelle 16) eine LMG-Analyse (Lindeman, Merenda & Gold, 1980) berechnet, die die über alle zulässigen Prädiktorreihenfolgen gemittelte Varianzaufklärung einzelner Prädiktoren bestimmt. Die aufgeklärte Gesamtvarianz von 18.48 % im Posttest Problemlösen wird zu

44.26 % (R^2 = 8.24 %) durch die anfängliche Problemlöseleistung, zu 35.05 % (R^2 = 6.40 %) durch die anfängliche Mathematikleistung, zu 7.76 % (R^2 = 1.21 %) durch die Interaktion von Problemlösen und Gruppenzugehörigkeit, zu 6.50 % (R^2 = 1.47 %) durch die Gruppenzugehörigkeit, zu 4.95 % (R^2 = 0.89 %) durch die Interaktion von Mathematik und Gruppenzugehörigkeit und zu 1.38 % (R^2 = 0.27 %) durch die Interaktion von Problemlösen und Mathematik erklärt (Abbildung 20).

Der unique Effekt des Trainings (f^2 = .01) ist als klein zu beurteilen (Konvention nach Cohen, 1988). Der Interaktionseffekt von Problemlösen und Gruppenzugehörigkeit (f^2 = .01) ist ebenfalls als klein zu beurteilen. Durch die Korrelation der Prädiktoren waren diese Effekte in der multiplen Regression nicht signifikant. Das Training zeigte damit beim Posttest Problemlösen einen kleinen ATI-Effekt zugunsten schwächerer Problemlöser. Diese Interpretation wird unterstützt durch die Signifikanz des entsprechenden ATI-Effekts beim Posttest Mathematik (siehe Kapitel 3.4.4.3), der einen ATI-Effekt beim Posttest Problemlösen voraussetzt.

3.4.4.3 Fragestellung 8: Transfer Mathematik

Steigert das fächerübergreifende Problemlösekompetenztraining mit Fokus auf spezifische Komponenten (u. a. Planungsfähigkeit, konditionales Wissen) die Problemlösekompetenz im Fach Mathematik? Zur Beantwortung der Fragestellung wurden multiple Regressionen, eine Kommunalitätenanalyse mit dem R-Paket *yhat* (Nimon et al., 2008) und eine Analyse der relativen Wichtigkeit der Prädiktoren (Grömping, 2006) berechnet, da multiple Regressionen alleine nicht ausreichen, um die Korreliertheit der Prädiktoren (Tabelle 15) zu berücksichtigen.

Multiple Regression

Um den Einfluss der Prätestleistung der Domänen Mathematik (M.T1) und Problemlösen (PL.T1) sowie der Gruppenzugehörigkeit (Gruppe; 0 = Kontrollgruppe, 1 = Experimentalgruppe) und mögliche Interaktionseffekte (PL.T1 x M.T1, PL.T1 x Gruppe, M.T1 x Gruppe, PL.T1 x M.T1 x Gruppe) zu berücksichtigen, wurde mit dem R-Paket *mice* (van Buuren & Groothuis-Oudshoorn, 2011) eine gepoolte hierarchische multiple Regression mit der Leistung im Posttest Mathematik (M.T2) als Kriterium berechnet (Tabelle 18). Es wurden zunächst schrittweise die Prätestvariablen für Mathematik und Problemlösen sowie die Gruppenzugehörigkeit (0 = Kontrollgruppe, 1 = Experimentalgruppe) in die Analyse aufgenommen. Anschließend wurden in den Modellen 4 bis 7 mögliche Interaktionseffekte geprüft.

In Modell 1 wurde die Leistung im Posttest Mathematik (M.T2) vorhergesagt durch die Leistung im Prätest Mathematik (M.T1). Die Varianzaufklärung betrug 16.35 % und war signifikant, $F(1{,}175.4328) = 24.226, p < .001$. Die Hinzunahme der Leistung im Prätest Problemlösen (PL.T1) in Modell 2 erklärte zusätzlich 3.46 % der Varianz, und diese Veränderung in R^2 war einseitig signifikant, $F(1, 75.99154) = 3.827, p = .054$. Model 1 und 2 weisen erneut auf die Bedeutung des Vorwissens für mathematisches Problemlösen hin. Die Hinzunahme der Gruppenzugehörigkeit (KG = 0, EG = 1) in Modell 3 erklärte zusätzlich 0.14 % der Varianz, und diese Veränderung in R^2 war nicht signifikant, $F < 1, p = .903$. Die Hinzunahme der Interaktion der Prätestleistung der beiden Domänen in Modell 4 erklärte zusätzlich 0.26 % der Varianz, und diese Veränderung in R^2 war nicht signifikant, $F < 1, p = .933$. Die Hinzunahme der Interaktion der Prätestleistung Mathematik und Gruppenzugehörigkeit in Modell 5 erklärte zusätzlich 2.97 % der Varianz, und diese Veränderung in R^2 war einseitig signifikant, $F(1, 107.8456) = 3.829, p = .053$. Die Hinzunahme der Interaktion der Prätestleistung Problemlösen und Gruppenzugehörigkeit in Modell 6 erklärte zusätzlich 2.92 % der Varianz, und diese Veränderung in R^2 war einseitig signifikant, $F(1, 95.31898) = 3.000, p = .087$, $f^2 = .03$. Die Hinzunahme der Dreifachinteraktion von Prätestleistung Problemlösen, Prätestleistung Mathematik und Gruppenzugehörigkeit in Modell 6 erklärte zusätzlich 0.42 % der Varianz, und diese Veränderung in R^2 war nicht signifikant, $F < 1, p = .828$. Insgesamt klärten die 7 Prädiktoren 23.71 % der Varianz im Posttest Mathematik auf.

Tabelle 18. Gepoolte hierarchische multiple Regression über die 20 imputierten Datensätze mit der Mathematikleistung im Posttest als Kriterium

						95 % CI				
	B	*SE*	*t*	*df*	*p*	UG	OG	NA	FMI	λ
Modell 1[a]										
M	.421	.086	4.922	82.872	< .001	.251	.592	36	.290	.273
Modell 2[b]										
M	.343	.092	3.731	78.086	< .001	.160	.527	36	.307	.290
P	.205	.105	1.956	49.725	.056	-.005	.415	36	.451	.429
Modell 3[c]										
M	.343	.093	3.676	76.464	.000	.157	.528	36	.313	.295
P	.208	.108	1.918	49.505	.061	-.010	.425	36	.451	.429
G	.025	.204	.122	92.735	.903	-.380	.429	0	.252	.236
Modell 4[d]										
M	.340	.104	3.250	66.848	.002	.131	.548	36	.354	.335
P	.207	.108	1.924	51.666	.060	-.009	.423	36	.436	.415
G	.024	.205	.118	9.546	.906	-.384	.432	0	.259	.242
P x M	.007	.083	.085	58.471	.933	-.159	.173	36	.396	.376
Modell 5[e]										
M	.418	.134	3.115	8.599	.003	.151	.685	36	.294	.277
P	.209	.108	1.930	5.448	.059	-.008	.426	36	.443	.421
G	.048	.205	.235	94.178	.815	-.359	.455	0	.245	.229
P x M	-.012	.085	-.143	64.231	.886	-.181	.157	36	.366	.346
P x G	-.156	.166	-.941	117.271	.349	-.484	.172	36	.173	.159
Modell 6[f]										
M	.327	.146	2.249	69.079	.028	.037	.618	36	.341	.323
P	.452	.191	2.366	41.429	.023	.066	.838	36	.507	.484
G	.151	.212	.712	87.425	.478	-.271	.573	0	.267	.251
P x M	-.014	.082	-.174	7.018	.863	-.178	.150	36	.337	.318
P x G	-.015	.181	-.084	105.722	.933	-.374	.343	36	.206	.191
M x G	-.373	.216	-1.732	57.587	.089	-.805	.058	36	.399	.379
Modell 7[g]										
M	.317	.164	1.928	6.370	.059	-.012	.646	36	.383	.363
P	.451	.190	2.376	42.830	.022	.068	.835	36	.495	.472
G	.171	.235	.727	71.626	.470	-.298	.640	0	.329	.310
P x M	.002	.121	.019	61.704	.985	-.240	.245	36	.376	.356
P x G	.000	.210	-.002	72.068	.999	-.420	.419	36	.327	.308
M x G	-.372	.212	-1.755	63.259	.084	-.795	.052	36	.368	.349
P x M x G	-.040	.184	-.217	45.625	.829	-.412	.331	36	.474	.452

Anmerkungen. P = Prätest Problemlösen, M = Prätest Mathematik, G = Gruppe, CI = Konfidenzintervall, UG = Untere Grenze. OG = Obere Grenze, NA = Anzahl fehlender Werte, fmi = fraction of missing information, λ = Proportion of the variation attributable to the missing data. Varianzaufklärung [%]:
[a] R^2 = 16.35, ΔR^2 = 16.35***. [b] R^2 = 19.81, ΔR^2 = 3.46* (einseitig).
[c] R^2 = 19.96, ΔR^2 = 0.14. [d] R^2 = 20.22, ΔR^2 = 0.26. [e] R^2 = 20.80, ΔR^2 = 0.58.
[f] R^2 = 23.29, ΔR^2 = 2.92* (einseitig). [g] R^2 = 23.71, ΔR^2 = 0.42.
* p < .05, *** p < .001.

Kommunalitätenanalyse

Eine Kommunalitätenanalyse mit dem R-Paket *yhat* (Nimon et al., 2008) zeigte, dass der unique Varianzaufklärungsanteil von Problemlösen (5.26 %) höher ist als von Mathematik (2.91 %). Letzterer ist ähnlich groß wie der unique Varianzaufklärungsanteil durch die Interaktion von Problemlösen und Gruppenzugehörigkeit (2.40 %, $f^2 = .02$). Die vier Prädiktoren Gruppenzugehörigkeit, Problemlösen x Mathematik, Mathematik x Gruppenzugehörigkeit sowie die Dreifachinteraktion klären unique jeweils weniger als ein Prozent der Varianz auf. Mit Ausnahme des ATI-Effekts Mathematik x Gruppenzugehörigkeit (4.32 %, $f^2 = .05$) liegt auch ihr gemeinsamer Varianzaufklärungsanteil in diesem Bereich.

Tabelle 19. Ergebnisse der Kommunalitätenanalyse mit der Posttestleistung Mathematik (M.T2) als Kriterium

Prädiktor	*Unique [%]*	*Common [%]*	*Total [%]*
Mathematik T1	2.91	13.52	16.43
Problemlösen T1	5.26	5.24	10.50
Gruppe	0.54	-0.35	0.19
Problemlösen T1 x Mathematik T1	0.25	3.87	4.12
Mathematik T1x Gruppe	0.19	4.32	4.51
Problemlösen T1 x Gruppe	2.40	-0.68	1.72
Problemlösen T1 x Mathematik T1 x Gruppe	0.42	-0.09	0.33

Anmerkungen. T1 = Prätest. Die Kommunalitätenanalyse wurde mit dem R-Paket *yhat* (Nimon et al., 2008) berechnet. Man beachte, dass sich die Summe der Spalte „Total" nicht zu R^2 summiert, da es gemeinsame Varianzanteile der einzelnen Prädiktoren gibt (Kraha et al., 2012).

Analyse der relativen Wichtigkeit

Ergänzend wurde eine Analyse der relativen Wichtigkeit der Prädiktoren vorgenommen (Grömping, 2006, 2007). Da das R-Paket *relaimpo* (Grömping, 2006) in der aktuellen Version 2.2-2 noch keine Dreifachinteraktionen berücksichtigen kann, wird zumindest für Modell 6 (siehe Tabelle 18) eine LMG-Analyse (Lindeman et al., 1980) berechnet, die die über alle zulässigen Prädiktorreihenfolgen gemittelte Varianzaufklärung einzelner Prädiktoren bestimmt. Die aufgeklärte Gesamtvarianz von 23.48 % im Posttest Mathematik wird zu 55.13 % ($R^2 = 12.70$ %) durch die anfängliche Mathematikleistung, zu 29.08 % ($R^2 = 7.01$ %) durch die anfängliche Problemlöseleistung, zu 12.35 % ($R^2 = 2.96$ %) durch die Interaktion von Problemlösen und Gruppenzugehörigkeit, zu 1.50 % ($R^2 = 0.35$ %) durch die Interaktion von Mathematik und Gruppenzugehörigkeit, zu 1.09 % ($R^2 = 0.25$ %) durch die Interaktion von Problemlösen und Mathematik und zu 0.85 % ($R^2 = 0.20$ %) durch die Gruppenzugehörigkeit erklärt (Abbildung 20). Verglich man die relative Wichtigkeit der Prädiktoren in den beiden multiplen Regressionen für die Kriterien Problemlösen und Mathematik, so erwiesen sich das domänenspezifische Vorwissen in Form der Prätestleistung, gefolgt von der Prätestleistung der jeweils anderen Domäne sowie die Interaktion von Problemlösen und Gruppenzugehörigkeit als beste Prädiktoren der Posttestleistung (Abbildung 20).

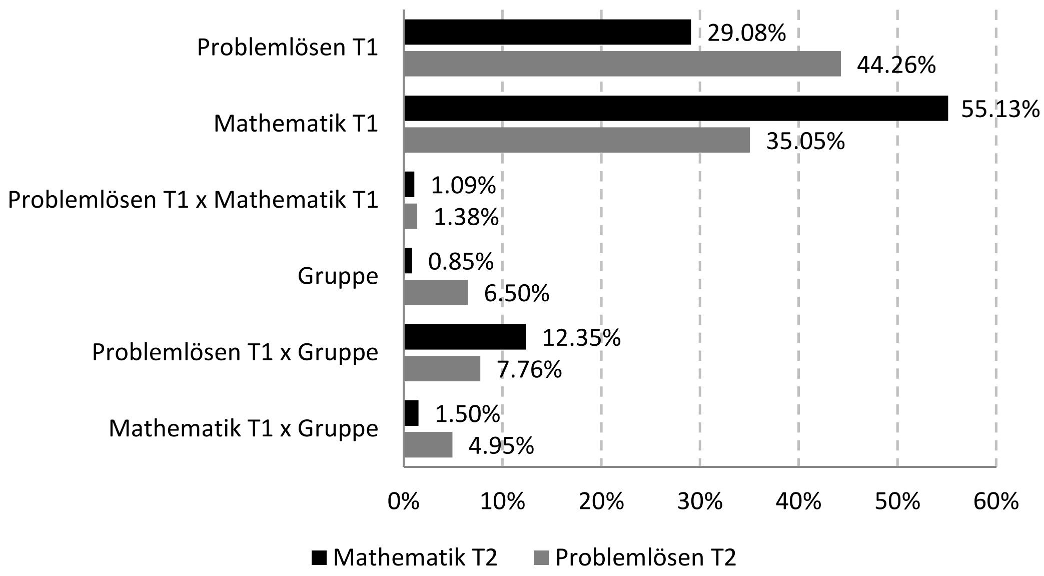


Abbildung 20. Analyse der relativen Wichtigkeit: Prozentualer Anteil der Varianzaufklärung durch die verschiedenen Prädiktoren an der jeweiligen aufgeklärten Varianz von Problemlösen ($R^2 = 18.48$ %) und Mathematik ($R^2 = 23.48$ %) im Posttest (T2)

Insgesamt ist der Haupteffekt des Trainings ($f^2 < .001$) zur Vorhersage der Mathematikleistung nicht bedeutsam. Der Interaktionseffekt von Problemlösen und Gruppenzugehörigkeit ($f^2 = .03$) gilt als kleiner Effekt (Konvention nach Cohen, 1988). Das Training zeigte damit beim Posttest Mathematik einen kleinen ATI-Effekt zugunsten schwächerer Problemlöser (vgl. Buchwald, Fleischer, Rumann, Wirth & Leutner, in preparation).

3.4.4.4 *Fragestellung 9a: Differentielle Trainingseffekte und kognitives Potenzial*

Profitieren im Sinne der kognitiven Potenzialausschöpfungshypothese (Leutner et al., 2004; OECD, 2004b) insbesondere SuS vom Training, die im Prätest fächerübergreifenden Problemlösens eher hoch und im Prätest Mathematik eher niedrig abgeschnitten haben? Die multiplen Regressionen zum Problemlösen (Kapitel 3.4.4.2) und zur Mathematik (Kapitel 3.4.4.3) zeigten in beiden Domänen keine signifikanten Effekte der Dreifachinteraktion von Treatment und den Prätestleistungen Mathematik und Problemlösen. Die Ergebnisse der Kommunalitätenanalysen (Tabelle 17 und 17) zeigten, dass der unique und gemeinsame Varianzanteil der Dreifachinteraktion von Gruppe x Problemlösen x Mathematik zur Vorhersage beider Kriterien unbedeutend ist. Auch die Strukturkoeffizienten d. h. die Korrelation von Prädiktorvariable und geschätzter Kriteriumsvariable, sind für die Produktvariable der Dreifachinteraktion und die Kriterien Posttest Problemlösen (.031) und Posttest Mathematik (.016) nahe Null. Insgesamt gibt es in Experiment 3 damit keine Hinweise auf differentielle Trainingseffekte zugunsten von SuS, die zu Beginn im Problemlösen eher gut und in Mathematik eher schlecht abgeschnitten haben.

3.4.4.5 *Fragestellung 9b: Kognitive Potenzialausschöpfung*

Profitieren insbesondere SuS vom Training, bei denen zum Zeitpunkt des Prätests kognitives Potenzial im Sinne der kognitiven Potenzialausschöpfungshypothese (Leutner et al., 2004; OECD, 2004b) vorhanden ist? Zur Beantwortung der Fragestellung wurde für Kontroll- und Experimentalgruppe getrennt ein McNemar-χ^2-Test für abhängige Stichproben (Bortz, Lienert & Boehnke, 2008) pro imputiertem Datensatz berechnet, um zu prüfen, ob SuS, die zu Beginn des Trainings niedriger in Mathematik abgeschnitten hatten, als es auf Grundlage ihrer fächerübergreifenden Problemlösekompetenz zu erwarten war, im Posttest ihr kognitives Potenzial ausschöpfen. Vorbereitend wurde für Prä- und Posttest eine Indikatorvariable für das Vorhandensein kognitiven Potenzials angelegt (1 = *Potenzial vorhanden*, d. h. die Mathematikleistung war deskriptiv geringer als auf Basis der Problemlöseleistung vorhergesagt; 0 = *Potenzial ausgeschöpft*, d. h. die Mathematikleistung war deskriptiv mindestens so groß, wie auf Basis der Problemlöseleistung vorhergesagt; vgl. Abbildung 1 und Abbildung 21). Da die McNemar-χ^2-

Tests sämtlich nicht signifikant waren ($X^2_{df=1} < .001$, $p = 1.00$), war eine Aggregation über die 20 Analysen (*pooling*) nicht nötig. Es zeigten sich keine differentielle Trainingseffekte zugunsten von SuS, die im Prätest schlechter in Mathematik abgeschnitten haben, als es auf Grundlage ihrer fächerübergreifenden Problemlösekompetenz zu erwarten war.

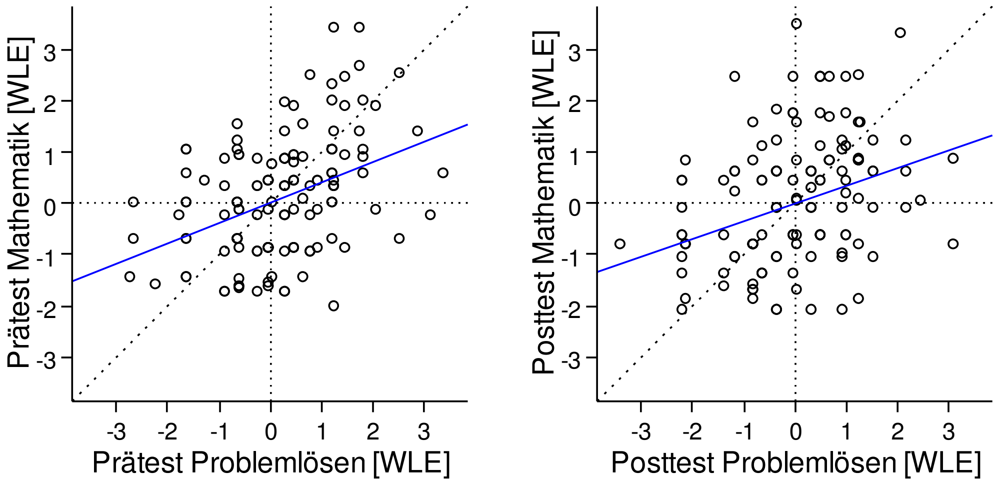


Abbildung 21. Streudiagramme der Leistungen in den Domänen Problemlösen und Mathematik für Prä- und Posttest auf Basis der nicht-imputierten Daten

Anmerkung. Der Zusammenhang der Prätestleistungen der Domänen Problemlösen und Mathematik ist deutlich weniger eng ($r = .38$, $p < .001$) als in der PISA-2003-Stichprobe ($r = .89$, $p < .001$), was nur teilweise durch die dortigen latenten Korrelationen zu erklären ist.

3.4.4.6 Fragestellung 10a: Bearbeitungsquote

Die Trainer in Experiment 3 berichteten von Schüleräußerungen, die vermuten lassen, dass zumindest einige SuS nicht versucht haben subjektiv, schwierig erscheinende Aufgaben zu bearbeiten („Wer soll denn den ganzen Text lesen?“, „Denken ist echt nix für mich!“, „Ich sehe, dass ich das nicht kann!“). Diese anekdotische Evidenz wird im Folgenden auf quantitativer Basis untersucht.

Nur 24 von 136 Personen hatten alle Aufgaben ihres Testhefts bearbeitet. Abbildung 22 zeigt, dass die meisten Personen eher wenige fehlende Werte aufweisen. Zusätzlich gibt es keine Totalverweigerer unter den Getesteten. Der durchschnittliche Anteil der fehlenden Werte lag zwischen 19 und 22 Prozent (Tabelle 20), was in etwa 2 Items entspricht. Der durchschnittliche Anteil der fehlenden Werte unterschied sich zwischen den Domänen und den experimentellen Bedingungen kaum, wobei die Mittelwerte der Kontrollgruppe etwas niedriger waren (Tabelle 20). Angesichts dessen und der großen Standardabweichungen zwischen 22 und 23 Prozent wurde auf eine inferenzstatistische Prüfung von Gruppenunterschieden verzichtet. Es sind keine Gruppenunterschiede hinsichtlich der Anzahl nicht-bearbeiteter Aufgaben zu erkennen, und die Domänen Problemlösen und Mathematik unterscheiden sich nicht im Muster fehlender Werte. Eine Übersicht nicht-geplanter fehlender Werte pro Item ist in Anhang Z zu finden.

Tabelle 20. Prozent nicht-geplanter fehlender Werte beim Posttest

Domäne	*experimentelle Bedingung*	*M*	*SD*	*n*
Problemlösen	Experimentalgruppe	19.33	22.60	70
	Kontrollgruppe	19.09	21.89	66
	Gesamt	19.22	22.18	136
Mathematik	Experimentalgruppe	21.71	22.67	70
	Kontrollgruppe	18.92	23.12	66
	Gesamt	20.35	22.85	136

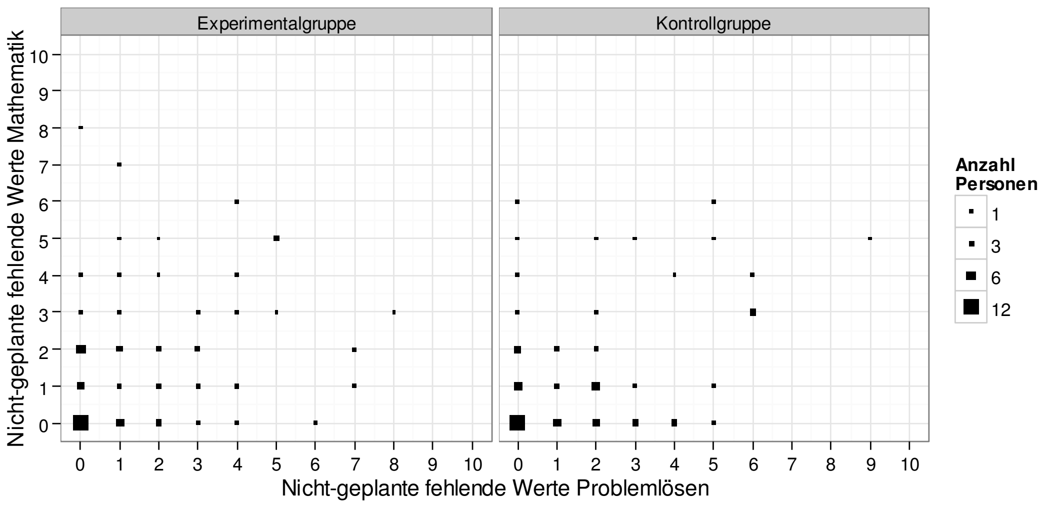


Abbildung 22. Anzahl nicht-geplanter fehlender Werte beim Posttest Problemlösen und Mathematik (N = 136)

3.4.4.7 Fragestellung 10b: Strategische Herangehensweise

Deskriptiv zeigte die Experimentalgruppe im Vergleich zur Kontrollgruppe leicht niedrigere Mittelwerte und leicht geringere Streuungen bei allen drei Skalen *systematisches Problemlösen*, *unsystematisches Problemlösen* und *Hilfesuchen* des PS-SJT (Bertling, 2012; Tabelle 21). Dotplots und Boxplots (Abbildung 23) zeigten, dass sich Experimentalgruppe und Kontrollgruppe hinsichtlich der Skalen *systematisches Problemlösen*, *unsystematisches Problemlösen* und *Hilfesuchen* in der Form der Verteilungen optisch leicht unterscheiden, dass jedoch nicht mit signifikanten Gruppenunterschieden zu rechnen ist. Zur inferenzstatistischen Prüfung eventueller Gruppenunterschiede wurde eine multivariate Varianzanalyse mit der experimentellen Bedingung (EG vs. KG) als unabhängiger Variable und den drei Skalen des PS-SJT als abhängigen Variablen berechnet. Die Voraussetzung der Varianzhomogenität wurde mit Boxs *M*-Test geprüft ($M(6, 122445.28) = 12.734$, $p = .053$). Der multivariate Haupteffekt war nicht signifikant (Wilks $\Lambda = .986$, $F < 1$). Die mit G*Power 3.1.5 (Faul, Erdfelder, Lang & Buchner, 2007) berechnete Teststärke liegt bei .97 (unter Annahme eines mittelgroßen Effekts von $f^2 = .15$ und $\alpha = .05$, $N = 132$, 2 Gruppen, 3 abhängige Variablen). Es fanden sich also keine Hinweise darauf, dass sich Experimental- und Kontrollgruppe bezüglich der drei Skalen *systematisches Problemlösen*, *unsystematisches Problemlösen* und *Hilfesuchen* bedeutsam unterscheiden.

Tabelle 21. Deskriptivstatistik der Skalen des Problem-Solving Situational Judgement Tests (Bertling, 2012)

Skala	*experimentelle Bedingung*	*M*	*SD*	*n*
systematisches Problemlösen	Experimentalgruppe	12.26	2.12	66
	Kontrollgruppe	12.44	2.81	66
	Gesamt	12.35	2.48	132
unsystematisches Problemlösen	Experimentalgruppe	8.15	2.53	66
	Kontrollgruppe	8.58	2.82	66
	Gesamt	8.36	2.68	132
Hilfesuchen	Experimentalgruppe	10.80	2.66	66
	Kontrollgruppe	10.45	2.76	66
	Gesamt	10.63	2.71	132

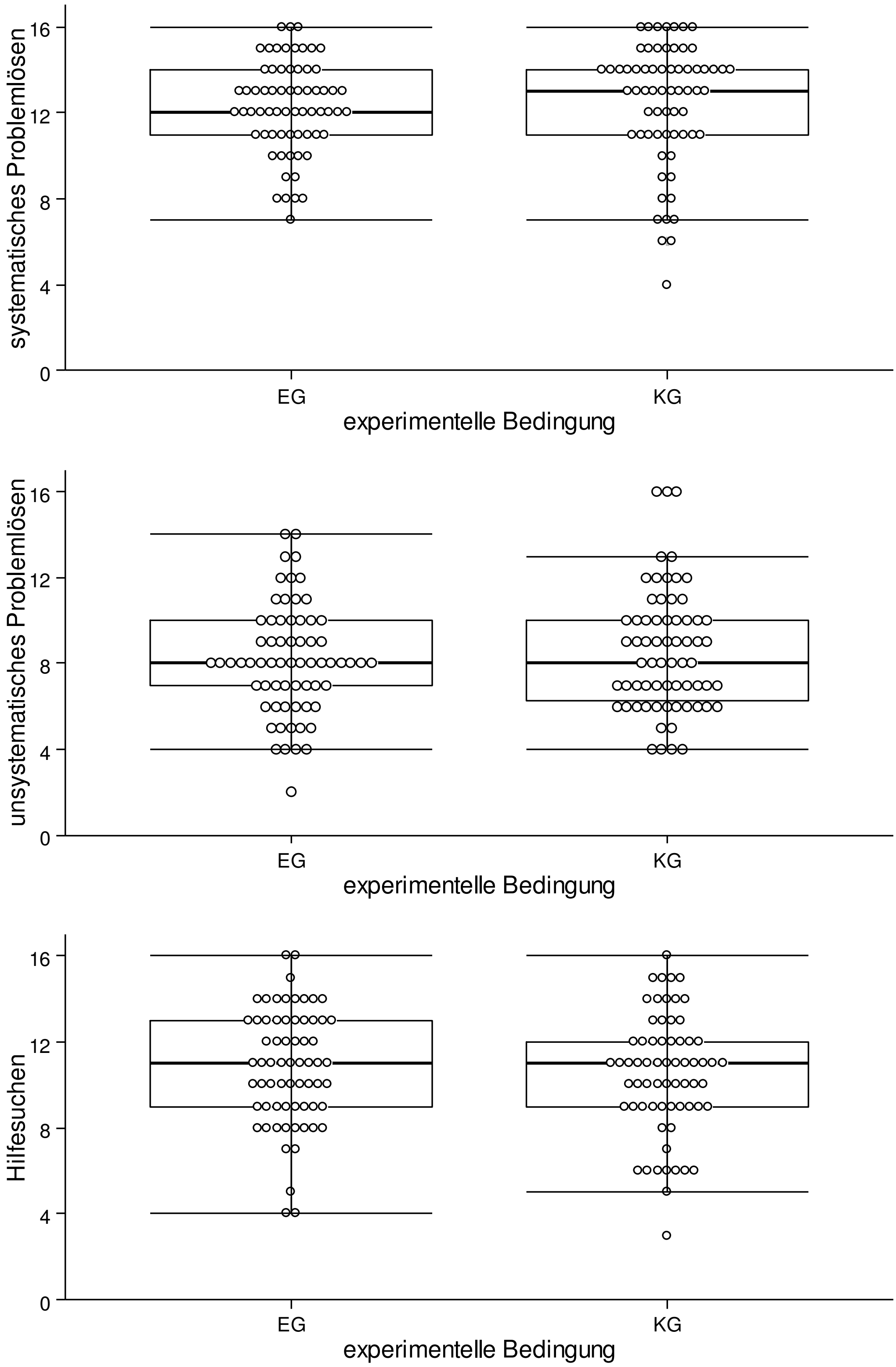


Abbildung 23. Kombination von Dotplot und Boxplot für die drei Skalen des Problem-Solving Situational Judgement Tests (Bertling, 2012) gruppiert nach Experimentalgruppe (EG) und Kontrollgruppe (KG)

3.4.4.8 Weitere Befunde

Anstrengungsbereitschaft

Die Experimentalgruppe (M = 3.20, SD = 0.82, n = 68) und Kontrollgruppe (M = 3.22, SD = 0.76, n = 66) unterschieden sich im Posttest nicht hinsichtlich der berichteten Anstrengungsbereitschaft (t(132) = -.211, p = .834). Das Niveau der selbst berichteten Anstrengung war in beiden Gruppen eher hoch.

Summative Evaluation

Am Ende des Schulhalbjahres wurden die SuS um eine kurze Evaluation des Trainingsprogramms gebeten. Eine erste Auswertung der Evaluation ist bei Goertzen (2014) zu finden. Die summativen Ergebnisse zeigten unter anderem, dass den SuS das Programm mittelmäßig gefallen hatte (mit großer Streuung), sie sich an einzelne Trainingsinhalte namentlich nicht mehr erinnern konnten, und dass viele SuS die zahlreichen Tests monierten. Die mittlere Selbsteinschätzung der SuS auf einer sechsstufigen Schulnotenskala hinsichtlich Motivation, Anstrengung, Mitarbeit und Lernerfolg lag jeweils im *guten* bis *befriedigenden* Bereich (Goertzen, 2014).

3.4.5 Diskussion

Der Reihenfolge der Hypothesen folgend werden zunächst die Ergebnisse der einzelnen Fragestellungen in Experiment 3 kurz zusammengefasst und kritisch diskutiert.

3.4.5.1 Planungskompetenztraining

Wenn das Planungskompetenztraining (Arling & Spijkers, 2012) ergänzt um Reflexionsaufgaben (Hoffmann, 2012) gemäß dem Manual durchgeführt wurde, wurden diese SuS besser bei den im Planungskompeteztraining eingesetzten Planspiel-Aufgaben (vgl. Fragestellung 6). Transfereffekte in einem kleinen Transfertest zum Problemlösen, der unmittelbar nach dem Posttest des Planungskompetenztrainings durchgeführt wurde, konnten nicht aufgezeigt werden. Gründe für diesen Null-Transfer zum Problemlösen kann es viele geben. Vor allem die mangelnde Reliabilität des Transfertests, die auf die Unkorreliertheit der drei Items zurückgeht, war überraschend. Möglicherweise ist das Planungskompetenztraining auch erst bei zeitlich umfangreicheren Aufgaben wie dem 45-minütigen

Posttest des Planungskompetenztrainings wirksam. Einschränkend muss zudem betont werden, dass zumindest für die untersuchte Gesamtschulstichprobe das Testformat des Planungskompetenztrainings neu und damit ungewohnt war. Die 45-minütige Stillarbeitsaufgabe brachte einzelne SuS an ihre motivationalen und kognitiven Grenzen. Nicht aufzugeben fiel einzelnen SuS schwer, wie die Trainer berichteten. Ein mögliches Argument für das gute Abschneiden einer Kontrollgruppe, die im Prätest Verweigerungen zeigte, im Posttest des Planungskompetenztrainings ist, dass diese Gruppe von den Trainern besonders motiviert wurde, um eine gute Mitarbeit während des restlichen Trainings zu erreichen.

3.4.5.2 Problemlösen und Mathematik

Die Ergebnisse der Domäne Problemlösen zeigten einen kleinen Trainingseffekt und einen kleinen ATI-Effekt für anfänglich schwächere Problemlöser (vgl. Fragestellung 7). Beide Effekte waren in der multiplen Regression nicht signifikant. Insbesondere der ATI-Effekt war jedoch in der Kommunalitätenanalyse sichtbar. Die Ergebnisse der Domäne Mathematik zeigten ebenfalls keinen Effekt des Trainings, jedoch einen kleinen signifikanten ATI-Effekt für anfänglich schwache Problemlöser (vgl. Fragestellung 8; Buchwald et al., in preparation). Sowohl für die Domäne Problemlösen als auch für die Domäne Mathematik zeigten sich weder Effekte für die Interaktion von Training und Prätestleistung Mathematik noch für die Dreichfachinteraktion von Training, Prätestleistung Problemlösen und Prätestleistung Mathematik (vgl. Fragestellung 9b). Die Ergebnisse des McNemar-χ^2-Tests zeigten, dass sich die Anzahl der SuS, die ihr im Prätest vorhandenes kognitives Potenzial im Posttest ausschöpften, nicht verbessert hat. Angesichts der potenziellen Fehlerhaftigkeit der Dichotomisierung des Potenzials, die durch die mittlere Korrelation der Domänen Problemlösen und Mathematik und die verbesserungswürdige Reliabilität der Tests entsteht, sollten die Ergebnisse des McNemar-χ^2-Test jedoch vorsichtig interpretiert werden. Der Versuch der Replikation der Befunde erscheint wünschenswert.

Mit Blick auf Zöttl et al. (2007), die den Transfer außermathematisch erworbener Problemlösekompetenzen auf mathematisches Problemlösen skeptisch sehen, konnte Experiment 3 zumindest für schwache Problemlöser einen kleinen Transfereffekt des fächerübergreifenden Trainings auf die mathematische Domäne aufzeigen. Die mathematikdidaktisch interessante Frage, ob Problemlösestrategien im Training vermittelt wurden oder ob nur ihre Verwendung bewirkt wurde (A. Heinze, persönl. Mitteilung, 10.10.2013), ist auf Basis von Experiment 3 nicht entscheidbar, da das Training eher breit angelegt war und keine spezifischen Prätests stattgefunden haben, die die vorherige Kenntnisse spezischer Problemlösestrategien erfassten.

3.4.5.3 Bearbeitungsquote

Die Anzahl nicht-geplanter fehlender Werte als Indikator des Willens zur Aufgabenbearbeitung unterschied sich nicht signifikant zwischen Experimental- und Kontrollgruppe (vgl. Fragestellung 10a). Eine Betrachtung der Testhefte der SuS zeigt, dass diese im Allgemeinen wenig aufgeschrieben haben, ein typischer Befund der Mathematikdidaktik (z. B. Doorman et al., 2007; Törner & Zielinski, 1992; van den Heuvel-Panhuizen & Bodin-Baarends, 2004). Dazu passt die anekdotische Schüleräußerung „Brauchen wir heute schon wieder einen Stift?“. SuS dazu zu motivieren, vermehrt ihre Lösungsideen, Gedanken und Unklarheiten beim Problemlösen in Testsituationen aufzuschreiben, ist im Rahmen des Trainingsexperiments nicht gelungen. Es bleibt eine längerfristige Aufgabe des Mathematikunterrichts (Bruder, 2000; Collet, 2009; Törner & Zielinski, 1992).

3.4.5.4 Strategische Herangehensweise

Es wurden keine Gruppenunterschiede zwischen Experimental- und Kontrollgruppe hinsichtlich der Skalen *systematisches Problemlösen*, *unsystematisches Problemlösen* und *Hilfesuchen* des PS-SJT (Bertling, 2012) gefunden (vgl. Fragestellung 10b). Angesichts der mangelnden Reliabilität der Skalen kann die Frage nach Trainingseffekten bzgl. des systematischeren Vorgehens der SuS der Experimentalgruppe mit dem PS-SJT in der vorliegenden Stichprobe nicht adäquat beantwortet werden. Ein mehrmaliger Einsatz eines reliablen Instruments zur Strategiediagnostik ist für zukünftige Studien wünschenswert, um eventuelle Unterschiede vor Interventionsbeginn berücksichtigen zu können und den zeitlichen Verlauf, z. B. mit latenten Wachstumskurvenmodellen (McArdle, 2009), adäquat zu modellieren. Ergänzend sei bemerkt, dass Bertlings Zuordnung der Strategien (Faktoren) zu den Kompetenzstufen des Problemlösens zumindest diskussionswürdig ist. Hilfesuchen als erste Reaktion auf eine schwierige Problemsituation beispielsweise kann auch Ausdruck kompetenten Problemlösens sein (vgl. z. B. das Verhalten der Firma Tepco bei der Nuklearkatastrophe in Fokushima im Jahr

2011 als Negativbeispiel). Dafür sprechen auch die von Bertling (2012) berichteten Faktorinterkorrelationen, die für *systematisches Problemlösen* und *Hilfesuchen* höher sind (.579 bis .770 je nach Strukturgleichungsmodell) als für *unsystematisches Problemlösen* und *Hilfesuchen* (.224 bis .250 je nach Strukturgleichungsmodell) und ebenfalls höher als für *systematisches Problemlösen* und *unsystematisches Problemlösen* (.019 bis .053 je nach Strukturgleichungsmodell).

3.4.5.5 Limitationen und zukünftige Forschung

In diesem Unterkapitel werden Grenzen von Experiment 3 beschrieben und Ausblicke auf weitere Forschungsmöglichkeiten gegeben. Dabei wird zunächst die Durchführung kritisch diskutiert, und Empfehlungen für eine mögliche Replikationsstudie werden gegeben. Anschließend werden der Bedarf reliablerer Instrumente, der vor allem bei potentieller Individualdiagnostik nötig ist, Überlegungen zu differenzierteren Auswertungsschemata sowie weitere Operationalisierungsmöglichkeiten kognitiven Potenzials diskutiert. Abschließend wird ein Blick auf die Ergebnisse der Domänen Problemlösen und Mathematik bei PISA 2012 geworfen.

Anmerkungen zur Durchführung und Empfehlungen für Replikationen

Die Testzusammenstellung und das Training wurde aus Zeitgründen an einer Stichprobe realisiert. Die Unterrichtsmaterialien konnten aus Zeitgründen in ihrer Kombination nicht im Vorfeld erprobt werden. Die folgenden Erfahrungen sollten bei der Planung einer möglichen Replikationsstudie beachtet werden:

- In einzelnen Trainingsgruppen kam es aus schulinternen Gründen (Elternsprechtag, Lehrerausflug) zu unerwartetem Stundenausfall, der nicht nachgeholt werden konnte. Die Möglichkeit, ausgefallene Einheiten nachholen zu können, ist erstrebenswert.
- Eine Hausaufgabenkultur existierte in den teilnehmenden Klassen kaum. Daher wurde die ursprüngliche Idee wiederholender Übungsaufgaben, die die SuS selbstständig Zuhause bearbeiten sollten, im Laufe des Projekts verworfen.
- Der Trainingsumfang war durch die schulischen Rahmenbedingungen festgelegt. Der optimale Trainingsumfang kann durch entsprechende experimentelle Variation der Trainingszeit untersucht werden.

- Sowohl die computerbasierten Experimente 1 und 2 als auch die feldexperimentelle Trainingsstudie waren offen als Veranstaltung der Universität Duisburg-Essen ersichtlich. Die Vorteile externer geschulter studentischer Hilfskräfte als Testleiter bzw. Trainer (z. B. hohes Interesse und *commitment*, bezahlbarer Stundenlohn) in Experiment 3 geht mit möglichen Nachteilen einher (z. B. geringere pädagogische Erfahrung und Autorität, kaum Notendruck), da es sich bei ihnen um pädagogische Novizen handelt, die als Externe mit der jeweiligen Schulkultur und den SuS zu Beginn nicht vertraut sind. Die Trainer berichten, dass die Übungsaufgabe „Was ist in welcher Flasche?“ (Gärtner & Scharf, 2001) viel zu schwer für die SuS war. Keine Gruppe konnte die Aufgabe lösen. Die Items von Pugalee (2001) im Posttest waren für die vorliegende Stichprobe ebenfalls eher zu schwer. Anhang AA dokumentiert typische Schülerfehler bei den Testaufgaben.
- Im Gegensatz zu Experiment 1 und 2 gab es in Experiment 3 aus Zeitgründen keine Tests konditionalen Wissens und Handlungswissens pro Domäne. In zukünftigen Studien zur kognitiven Potenzialausschöpfungshypothese sollten diese wieder aufgenommen werden.
- Eine kurze summative Evaluation der Schülerwahrnehmung des gesamten Trainings am Ende des Schulhalbjahres zeigt u. a., dass es den SuS im Mittel durchschnittlich gefallen hat, die Anzahl und Länge der Tests aber für die SuS eher wenig attraktiv waren (Goertzen, 2014).

Bei einer möglichen Neuauflage der feldexperimentellen Trainingsstudie sollte am Ende der Intervention geprüft werden, ob die SuS die Trainingsinhalte (z. B. Heuristiken, metakognitive Fragen) behalten haben. Denn ohne Behalten bzw. Möglichkeit des deklarativen Abruf scheint eine Anwendung unwahrscheinlich. Und selbst wenn z. B. Heuristiken gelernt sind, erfolgt ihr Einsatz nicht automatisch (Stacey, 1991) und muss nicht zu besseren Testleistungen führen (Depaepe et al., 2010).

Teile der Trainingseffekte können zudem auf die Vermittlung von Arbeitsweisen zurückgehen (vgl. Leuders, 2010). Das ist etwa in einer Studie prüfbar, in der eine zusätzliche Kontrollgruppe aufgenommen wird, die nur Arbeitstechniken übt.

Ferner sind für Aussagen über Entwicklung und nicht nur über Veränderung der Problemlösekompetenz entsprechende längsschnittliche Studiendesigns mit mindestens drei Messzeitpunkten und Auswertungsmethoden wie Wachstumskurvenmodellen nötig (McArdle, 2009).

Reliablere Individualdiagnostik

Obwohl bewährte Aufgaben aus Large-scale-Assessments (Leutner et al., 2007, OECD, 2004a, 2004b) eingesetzt worden sind, ist die Reliabilität der eingesetzten Instrumente in Experiment 3 verbesserungsfähig. Zur möglichen mathematikdidaktisch nutzbaren Individualdiagnostik des kognitiven Potenzials, d. h. der Quantifizierung der Diskrepanz fächerübergreifenden und mathematischen Problemlösens, sind reliablere und zeitökonomischere Tests notwendig.

Verwendung differenzierterer Auswertungsschemata und änderungssensitiver Aufgaben?

Problemlösen ist ein mehrschrittiger Prozess. Das gilt auch für mathematische Aufgaben, die zahlreiche Zwischenschritte, Berechnungen und Argumentationen erfordern können (Blum et al., 2006; KMK, 2004b). Die eingesetzten PISA-Aufgaben zum (mathematischen und fächerübergreifenden) Problemlösen sind überwiegend als geschlossene Aufgaben formuliert, deren Items dichotom oder trichotom ausgewertet werden, auch wenn sie mehrschrittige Lösungsprozesse erfordern (OECD, 2004a, 2004b). Das Scoring bei PISA (OECD, 2003) berücksichtigt zudem nicht die Unterschiedlichkeit möglicher Lösungswege und die Qualität falscher Antworten. Ob jemand beispielsweise nichts geschrieben hat, einen oder mehrere nicht-erfolgreiche Lösungsversuche unternommen hat oder sein richtiges Ergebnis durchgestrichen hat, vermag das strenge und ökonomische Scoring nicht zu differenzieren. Das ist für *large-scale*- Bildungsmonitoring im Allgemeinen adäquat. Für mathematikdidaktische und instruktionspsychologische Forschungszwecke sind Aufgaben und Auswertungen wünschenswert, die Informationen ermöglichen, die über ein „banales Richtig-falsch-Schema“ (Wernicke, 2013) hinausgehen, indem sie z. B. zusätzliche diagnostische Informationen liefern über fehlendes bzw. verfügbares Vorwissen, Bearbeitungsprozesse, die Strategienutzung bzw. Lösungswege oder typische Fehler und Fehlvorstellungen. Will man beispielsweise praxisnahe Lernfortschrittsdiagnostik (Souvignier, Förster & Salaschek, 2014) betreiben, sind fein differenziertere theoretische Modelle, Erwartungshorizonte, änderungssensitive Testaufgaben und entsprechende Scoringregeln wichtig, um Prozessfortschritte abzubilden und Lernprozesse in der Schule adäquat begleiten zu können. Wie die Analysen der Längsschnittstudie PISA-I-Plus (PISA-Konsortium Deutschland, 2006) zeigten, sind die Mathematikitems zumindest sensitiv genug, um die Veränderung mathematischer Kompetenz im Verlaufe eines Schuljahres abzubilden.

Die Grundfrage, ob die Verwendung differenzierterer Auswertungsschemata und womöglich änderungssensitiver Aufgaben, zu Vorteilen gegenüber der Verwendung dicho- oder trichotom kodierter Items führt, ist für jeden Einsatzbereich allerdings nur empirisch zu beantworten und setzt eine klare Definition der Ziele und einen empirischen Vergleich voraus. Ein solcher Vergleich war nicht Gegenstand von Experiment 3. Die Daten aus Experiment 3 können aber prinzipiell mittels neu zu entwickelnder Kategoriensysteme gescort werden, was einen Vergleich verschiedener Scorings erlauben würde.

Alternativen zur Operationalisierung des kognitiven Potenzials

Neben der verwendeten Interaktionsterm-Operationalisierung und dem Anlegen einer Indikatorvariable für kognitives Potenzial durch Dichotomisierung (vgl. Kapitel 3.4.4.4) sind weitere Möglichkeiten denkbar, kognitives Potenzial im Sinne der kognitiven Potenzialausschöpfungshypothese statistisch zu operationalisieren:

1. Durch *einfache Differenzbildung* der Testergebnisse x_{PL} in der Domäne Problemlösen und x_M in der Domäne Mathematik kann eine Differenzvariable erzeugt werden ($D := x_{PL} - x_M$). Diese berücksichtigt allerdings weder die Standardabweichungen der Tests, was z. B. durch z-Standardisierung lösbar ist, noch die Reliabilität der Tests. Zur Individualdiagnostik wird dieses stark messfehlerbehaftete Vorgehen nicht empfohlen (Walter et al., 2006).
2. Um zu prüfen, ob sich bei einer Person die Testergebnisse x_{PL} in der Domäne Problemlösen und x_M in der Domäne Mathematik unterscheiden, ist also ein entsprechendes Kriterium wünschenswert. Eine Möglichkeit besteht in der Angabe einer *kritischen Differenz* der Testwerte $D_{Krit}(x_{PL}, x_M)$, ab der sich die Testergebnisse statistisch signifikant voneinander unterscheiden. D_{Krit} kann definiert werden als aus der klassi-

schen Testtheorie stammende *kritische Differenz* zweier Testscores (Bühner, 2010): $D_{Krit} = z \cdot s_x \cdot \sqrt{2 - (r_{tt1} + r_{tt2})}$, wobei z der zum gewünschten Sicherheitsbereich gehörige z-Wert (z. B. 1.96 für 95 % zweiseitig), s_x die Streuung der Tests und r_{tti} die Reliabilität der Tests (i = 1, 2) ist. Das Problem dieser Variante für Experiment 3 besteht in der geringen Reliabilität der verwendeten Tests, die zudem auf IRT-Basis konstruiert sind (Anhang CC). Damit ergäbe sich für Experiment 3 eine kritische Differenz von fast zwei Standardabweichungen ($D_{Krit} = 1.90$). Zwischenfazit: Für präzise Individualdiagnostik sind die Tests nicht reliabel genug. Allerdings ist dieses D_{Krit} ein strenges Kriterium. Selbst bei zwei z-standardisierten Tests mit sehr guter Reliabilität von .90 beträgt die kritische Differenz $D_{Krit} = 1.96 \cdot \sqrt{2 - (1.8)} = .88$ Einheiten.

3. *Regressionsanalytische Methoden*. Per Definition ist das Residuum die Differenz des Kriteriumswert y und des vorhergesagten Kriteriumswerts $\hat{y}$ ($Residuum\ =\ y - \hat{y}$), d. h. es erfasst also Abweichungen von erwarteten Werten. Als Voraussetzung der linearen Regression sollte das Residuum standardnormalverteilt sein (Field, Miles & Field, 2012; Jarque & Bera, 1987). Wie kann für Personen entschieden werden, ob ihre Messwerte signifikant von der Regressionsgeraden abweichen, d. h. wie kann für eine Person entschieden werden, ob ihr Residuum signifikant von Null abweicht? Eine Möglichkeit besteht darin, Personen zu klassifizieren, die eine bestimmte Anzahl von Standardabweichungen über bzw. unter dem Mittelwert der Residualverteilung liegen (vgl. die Diagnostik von Underachievement). Das führt bei kleinen Stichproben (z. B. 200 Personen) und einseitiger Signifikanztestung bei 2 *SD* als Kriterium jedoch zu einer Fallauswahl von 2.5 Prozent der Personen (z. B. 5 von 200 Personen). Eine weniger strenge Möglichkeit basiert auf dem Standardschätzfehler der Regression (Bühner, 2010; Schmidt-Atzert & Amelang, 2012): $D_{Krit} = z \cdot s_x\sqrt{1 - r_{12}^2}$, wobei r_{12} die Korrelation der beiden Tests ist und s_x die Standardabweichung des Kriteriums. Durch die nur mittlere Korrelation der beiden Tests in der vorliegenden Stichprobe von r = .39, ergibt sich $D_{Krit} = 1.53$. Statt der individualdiagnostischen Prüfung der von Null abweichenden Residuen und der Selektion von Personen, kann bei Grup-

penstatistiken das Residuum als Prädiktor aufgenommen werden. Das hat den Vorteil, dass die Stichprobengröße gleich bleibt.

Zukünftige Studien zur kognitiven Potenzialausschöpfungshypothese an einer größeren Stichprobe mit reliableren Instrumenten können die verschiedenen Operationalisierungsmöglichkeiten kognitiven Potenzials vergleichen.

3.5 Gesamtdiskussion

Zur Erklärung der Diskrepanz zwischen der bei PISA 2003 relativ hohen Kompetenz im Bereich des Problemlösens und der im internationalen Vergleich relativ niedrigeren Kompetenzen im Bereich der Mathematik und der Naturwissenschaften von SuS in Deutschland postulierten Leutner et al. (2004) die kognitive Potenzialausschöpfungshypothese (siehe auch OECD, 2004b). Sie besagt, dass SuS in Deutschland über kognitves Potenzial verfügen, das zum Aufbau mathematischer Kompetenz nicht hinreichend genutzt wird. In einer Serie von drei Experimenten wurden Aspekte der kognitiven Potenzialausschöpfungshypothese (Leutner et al., 2004) für die Domäne Mathematik untersucht. Die Grundidee der Experimente 1 bis 3 war, die quer- und längsschnittliche Korrelation von fächerübergreifendem und mathematischem Problemlösen experimentell zu untersuchen. Dabei soll die beobachtete Variabilität von Komponenten der fächerübergreifenden Problemlösens durch experimentell manipulierte Variabilität ersetzt werden, um mögliche Transfereffekte auf mathematisches Problemlösen kausal interpretieren zu können (Leutner et al., 2009).

Zunächst wurde in zwei Laborexperimenten versucht, experimentell interindividuelle Variabilität wichtiger Komponenten fächerübergreifenden Problemlösens zu schaffen und postulierte Veränderungen auf fächerübergreifenden Problemlösekompetenz sowie auf mathematischer Problemlösekompetenz in Form von Transfereffekten zu testen. Dazu wurde ein computerbasiertes Trainingsprogramm eingesetzt, das auf wichtige Teilkomponenten der fächerübergreifenden Problemlösekompetenz abzielt (Planungsfähigkeit, konditionales Wissen, Handlungswissen), die auch für mathematisches Problemlösen wichtig sind (Fleischer et al., 2010; Leutner et al., 2009).

In Experiment 1 war das Training hinsichtlich der Effizienz der trainierten Komponenten fächerübergreifenden Problemlösens (konditionales Wissen, Planungsfähigkeit, prozedurales Wissen) und der Effizienz fächerübergreifenden Problemlösens erfolgreich. Das Niveau beim fächerübergreifenden Problemlösen war jedoch eher niedrig. Experiment 1 zeigte keine Transfereffekte bei den Tests der Domäne Mathematik. Der Treatment-Check in Experiment 2 war nur für eine Experimentalgruppe, die Planungsfähigkeit und konditionales Wissen trainierte, hinsichtlich einer Komponente (konditionales Wissen) erfolgreich. Die experimentelle Manipulation fächerübergreifender Problemlösekompetenz war nicht erfolgreich. Das Ausbleiben positiven Transfers in den Tests der Domäne Mathematik überrascht daher nicht. Insgesamt zeigten die Experimente 1 und 2 keine Evidenz für die kognitive Potenzialausschöpfungshypothese von Leutner et al. (2004). Sie sprechen angesichts der kaum erfolgreichen experimentellen Manipulation aber auch nicht dagegen. Es gelang in den Experimenten kaum die beobachtete Variabilität der Komponenten der fächerübergreifenden Problemlösens durch experimentell induzierte Variabilität zu ersetzen (Leutner et al., 2009).

In der feldexperimentellen Trainingsstudie (Experiment 3) zeigte sich im Posttest Mathematik ein kleiner Transfereffekt für anfänglich eher schwächere Problemlöser. Dieser kleine ATI-Effekt für die Domäne Mathematik steht in Einklang mit der Betonung des ungenutzten kognitiven Potenzials im unteren Leistungsbereich (Leutner et al., 2004), der in seiner Größe für ein relativ kurzes Training einer breit definierten Kompetenz erwartbar ist (Hasselhorn & Gold, 2006; vgl. auch Lester, Jr., 1994). Für SuS mit kognitiven Potenzial im Sinne der kognitiven Potenzialausschöpfungshypothese konnte Experiment 3 zudem keine Transfereffekte des fächerübergreifenden Trainings auf die Domäne Mathematik aufzeigen. Insgesamt liefert die Studie also nur schwache Evidenz für die kognitive Potenzialausschöpfungshypothese.

Um abschließend über die kognitive Potenzialausschöpfungshypothese zu urteilen, ist weitere Forschung, insbesondere zur Einordnung der Befunde von PISA 2012, nötig. Bei PISA 2012 schnitten SuS in Deutschland beim fächerübergreifenden Problemlösen (Kriterium) leicht besser ab, als es auf Grundlage ihrer Mathematikkompetenz (Prädiktor) erwartet wurde (OECD, 2014). Können die PISA-2012-Ergebnisse so interpretiert werden, dass das kognitive Potenzial im Sinne der kognitiven Potenzialausschöpfungshypothese (Leutner et al., 2004; OECD, 2004b) von SuS in Deutschland inzwischen besser zum Aufbau mathematischer Kompetenzen genutzt wird? Da PISA 2012 vor allem komplexes Problemlösen computerbasiert erfasst (OECD, 2014) und die kognitive Potenzialausschöpfungshypothese auf Basis der Ergebnisse zum analytischen Problemlösen des Papier-Bleistift-Tests bei PISA 2003 basiert (Leutner et al., 2004; OECD, 2004b), ist zur Beantwortung dieser Frage weitere Forschung nötig. Im Vergleich zu PISA 2003 (Leutner et al., 2004; OECD, 2004b), wo fächerübergreifendes Problemlösen als Prädiktor für das Kriterium Mathematik verwendet wurde, sind zudem die Rollen von Prädiktor und Kriterium bei PISA 2012 (OECD, 2014) teilweise vertauscht (vgl. den Ausblick in Kapitel 4.3.8), was auf unterschiedliche Kausalmodelle hindeutet. Die beiden Kausalitätsrichtungen schließen im zeitlichen Verlauf jedoch nicht aus. Einerseits lernen Menschen auch außerhalb des Mathematikunterrichts, Probleme zu lösen, und anderseits sollen im Unterricht erworbene mathematische Problemlösekompetenzen im Alltag eingesetzt werden (KMK, 2004b). Theoretisch sind also wechselseitigen Rückkopplungsprozessen bzw. reziproken Kausaleffekten zwischen fächerübergreifender und mathematischer Problemlösekompetenz denkbar (Leutner et al., 2005), deren Untersuchung und Modellierung im zeitlichen Verlauf wünschenswert erscheint.

4 Zusammenfassende Diskussion

Dieses Kapitel beginnt mit einer zusammenfassenden Übersicht der zentralen empirischen Ergebnisse der Arbeit. Anschließend werden der theoretische und praktische Ertrag der Arbeit kritisch diskutiert. Abschließend werden wichtige Grenzen der Studien und Ideen für zukünftige Forschung beschrieben.

4.1 Zentrale Ergebnisse

Die empirischen Ergebnisse werden in der Reihenfolge der zentralen Fragestellungen berichtet.

4.1.1 Fragestellung 1 (Treatment-Check): Lassen sich zentrale Komponenten der fächerübergreifenden Problemlösekompetenz (konditionales Wissen, Planungsfähigkeit, Handlungswissen) mittels laborexperimenteller Lehr-Lernprogramme messbar fördern?

In Experiment 1 war das Training hinsichtlich der *Effizienz* der trainierten Komponenten fächerübergreifenden Problemlösens (konditionales Wissen, Planungsfähigkeit, prozedurales Wissen) erfolgreich. In Experiment 2 war das Training nur in einer Gruppe, die Planungsfähigkeit und konditionales Wissen trainierte, und nur für die Komponente konditionalen Wissens erfolgreich. Insgesamt bewertet konnte das computerbasierte Trainingsprogramm in Experiment 1 und 2 zentrale Komponenten der fächerübergreifenden Problemlösekompetenz (konditionales Wissen, Planungsfähigkeit, Handlungswissen) nicht ausreichend messbar fördern.

4.1.2 Fragestellung 2 (Transfer Problemlösen): Steigert das laborexperimentelle Training weniger, zentraler Komponenten der fächerübergreifenden Problemlösekompetenz (konditionales Wissen, Planungsfähigkeit, Handlungswissen) die fächerübergreifende Problemlösekompetenz insgesamt?

In Experiment 1 war das Training weniger, zentraler Komponenten der fächerübergreifenden Problemlösekompetenz (konditionales Wissen, Planungsfähigkeit, Handlungswissen) hinsichtlich der *Effizienz* des fächerübergreifenden Problemlösens erfolgreich. In Experiment 2, in der jede Experimentalgruppe jeweils zwei der drei Komponenten aus Experiment 1 trainierte, war das Training hinsichtlich der *Leistung* des fächerübergreifenden Problemlösens nicht erfolgreich. Insgesamt bewertet konnte das computerbasierte Trainingsprogramm in Experiment 1 und 2 die fächerübergreifende Problemlösekompetenz nicht praktisch bedeutsam fördern.

4.1.3 Fragestellung 3 (Naher Transfer Mathematik): Steigert ein laborexperimentelles Training weniger, zentraler Komponenten der fächerübergreifenden Problemlösekompetenz (konditionales Wissen, Planungsfähigkeit, Handlungswissen) ebenfalls diese Komponenten der fachbezogenen Problemlösekompetenz im Fach Mathematik?

In Experiment 1 und 2 wurden keine Transfereffekte bei den Tests der Komponenten konditionales Wissen, Planungsfähigkeit und Handlungswissen der fachbezogenen Problemlösekompetenz im Fach Mathematik gefunden.

4.1.4 Fragestellung 4 (Ferner Transfer Mathematik): Steigert ein laborexperimentelles Training weniger, zentraler Komponenten der fächerübergreifenden Problemlösekompetenz (konditionales Wissen, Planungsfähigkeit, Handlungswissen) die fachbezogene Problemlösekompetenz im Fach Mathematik insgesamt?

In Experiment 1 und 2 zeigten sich keine fernen Transfereffekte bezüglich der fachbezogenen Problemlösekompetenz im Fach Mathematik. Da auch keine Transfereffekte bezüglich einzelner Komponenten der fachbezogenen Problemlösekompetenz im Fach Mathematik auftraten (vgl. Kapitel 4.1.3) ist dieses Ergebnis erwartbar und konsistent.

4.1.5 Fragestellung 5 (kompensatorischer Trainingseffekt): Zeigen sich differentielle Trainingseffekte hinsichtlich der Schulform, d. h. profitieren eher leistungsschwächere SuS (Hauptschüler) vom laborexperimentellen Training?

Die Interaktion von Schulform und experimenteller Bedingung in Experiment 1 und 2 war für keine abhängige Variable signifikant. Damit zeigten die beiden Experimente keine differentiellen Trainingseffekte hinsichtlich der Schulform.

Insgesamt zeigten die Experimente 1 und 2 damit keine Evidenz für die kognitive Potenzialausschöpfungshypothese (Leutner et al., 2004). Da die experimentelle Manipulation jedoch nur teilweise und eingeschränkt erfolgreich war, sprechen die ausbleibenden Transfereffekte auf die mathematische Domäne in Experiment 1 und 2 jedoch auch kaum gegen die kognitive Potenzialausschöpfungshypothese.

4.1.6 Fragestellung 6 (Treatment-Check): Lässt sich mit dem ursprünglich für Erwachsene konzipierten Aachener Planungskompetenztraining (Arling & Spijkers, 2012) die Planungskompetenz (als eine Teilkomponente des Problemlösens) in der Jahrgangsstufe 9 trainieren?

Die Ergebnisse des Experiments 3 zeigten, dass das Planungskompetenztraining (Arling & Spijkers, 2012) auch in der Jahrgangsstufe 9 erfolgreich eingesetzt werden kann. Die Effektgröße lag im mittleren Bereich (Cohens $d = 0.45$). Beim Einsatz des zukünftigen Einsatz des Planungskompetenztrainings in Schulen ist stark auf eine manualkonforme Durchführung und die Motivation der SuS zu achten, um Trainingserfolge zu erzielen.

4.1.7 Fragestellung 7 (Transfer Problemlösen): Steigert das fächerübergreifende Problemlösekompetenztraining mit Fokus auf spezifischen Komponenten (Planungsfähigkeit, konditionales Wissen, Handlungswissen, Heuristiken) die allgemeine fächerübergreifende Problemlösefähigkeit?

Der Haupteffekt des Treatments in Experiment 3 war nicht signifikant. Die Ergebnisse des Experiments 3 zeigten allerdings für anfänglich schwache Problemlöser Trainingseffekte für fächerübergreifendes Problemlösen (vgl. Buchwald et al., in preparation). Dieser kleine ATI-Effekt war in der multiplen Regression nicht signifikant. Die Kommunalitätenanalyse zeigte hingegen einen kleinen uniquen Erklärungsbeitrag. Man beachte, dass der entsprechende ATI-Effekt für die abhängige Variable Mathematik signifikant war, was ohne Trainingseffekte schwierig zu erklären ist.

4.1.8 Fragestellung 8 (Transfer Mathematik): Steigert das fächerübergreifende Problemlösekompetenztraining mit Fokus auf spezifischen Komponenten (Planungsfähigkeit, konditionales Wissen, Handlungswissen, Heuristiken) die Problemlösekompetenz im Fach Mathematik?

Die Ergebnisse des Experiments 3 zeigten für anfänglich schwache Problemlöser Transfereffekte des Problemlösetrainings auf die Domäne Mathematik. Anfänglich schwächeren Problemlösern gelang es nach dem Training mathematische Probleme besser zu lösen (vgl. Buchwald et al., in preparation). Dieser kleine ATI-Effekt steht in Einklang mit der Betonung des kognitiven Potenzials im Sinne der kognitiven Potenzialausschöpfungshypothese im unteren Leistungsbereich (Leutner et al., 2004, 2005). Mit großen Effekten ist bei einem breiten Trainingsansatz in einem kurzen Training einer breit definierten Kompetenz ohnehin nicht zu rechnen (Hasselhorn & Gold, 2006; vgl. Lester, 1994).

4.1.9 Fragestellung 9a (differentielle Effekte und kognitives Potenzial): Profitieren im Sinne der kognitiven Potenzialausschöpfungshypothese (Leutner et al., 2004; OECD, 2004b) insbesondere SuS vom Training, die im Prätest fächerübergreifenden Problemlösens eher hoch und im Prätest Mathematik eher niedrig abgeschnitten haben?

Die Interaktion von Gruppenzugehörigkeit x Prätestleistung Problemlösen x Prätestleistung Mathematik in Experiment 3 war für die Kriterien Problemlösen und Mathematik nicht signifikant. Es finden sich in Experiment 3 somit für die Kriterien Problemlösen und Mathematik keine differentiellen Trainings- und Transfereffekte für SuS, die im Prätest fächerübergreifenden Problemlösens eher hoch und im Prätest Mathematik eher niedrig abschnitten.

4.1.10 Fragestellung 9b (kognitive Potenzialausschöpfung): Profitieren insbesondere SuS vom Training, bei denen zum Zeitpunkt des Prätests kognitives Potenzial im Sinne der kognitiven Potenzialausschöpfungshypothese (Leutner et al., 2004; OECD, 2004b) vorhanden ist?

Der McNemar-χ^2-Test für abhängige Stichproben (Bortz et al., 2008) zur Prüfung, ob SuS, die im Prätest vorhandenes kognitives Potenzial zeigten, dieses kognitives Potenzial im Posttest ausschöpften, war nicht signifikant. Es finden sich in Experiment 3 somit keine Hinweise, dass SuS, die zu Beginn des Trainings niedriger in Mathematik abgeschnitten hatten, als es auf Grundlage ihrer fächerübergreifenden Problemlösekompetenz zu erwarten war, im Posttest ihr kognitives Potenzial ausschöpfen.

4.2 Theoretischer und praktischer Ertrag

In diesem Unterkapitel wird der theoretische und praktische Ertrag dieser Dissertation kurz zusammengefasst und an einigen Stellen ergänzt.

4.2.1 Theorie und Empirie zur kognitiven Potenzialausschöpfungshypothese (Leutner et al., 2004; OECD, 2004b)

In Kapitel 2 wurden verschiedene mögliche Erklärungen gesammelt und diskutiert, die (in Deutschland) zur Diskrepanz der Kompetenz in den Domänen Problemlösen und Mathematik bei PISA 2003 (Leutner et al., 2004; OECD, 2004b) geführt haben könnten. Eine Übersicht dieser Erklärungen (u. a. mangelnde Potenzialausschöpfung, mangelnde Potenzialnutzung, Kontexteffekte, Auswertungsfehler) fehlte bislang.

Wie steht es um die kognitive Potenzialausschöpfungshypothese? Theoretische Ähnlichkeiten der Prozesse und gemeinsame Komponenten fächerübergreifenden und mathematischen Problemlösens bzw. Modellierens (OECD, 2003), querschnittliche Korrelationen bei PISA 2003 (OECD, 2004b), längsschnittliche Korrelationen in einer deutschen PISA-Stichprobe (Leutner et al., 2006) und Kommunalitätsanalysen dieser Daten (Leutner et al., 2006), die sowohl einen hohen gemeinsamen Varianzanteil fächerübergreifenden Problemlösens und Mathematik für die Prognose mathematischer Kompetenz aufzeigen, sprechen für die kognitive Potenzialausschöpfungshypothese (Leutner et al., 2004). Zur Untersuchung der kognitiven Potenzialausschöpfungshypothese lagen vor dieser Arbeit jedoch m. E. keine experimentellen Untersuchungen vor, die den Einfluss von (Komponenten) fächerübergreifenden Problemlösens auf mathematisches Problemlösen untersucht haben.

Die Experimente 1 bis 3 untersuchten Aspekte der kognitiven Potenzialausschöpfungshypothese, die Diskrepanz fächerübergreifenden und mathematischen Problemlösens als Hinweis auf kognitives Potenzial sieht, das in Deutschland nicht optimal zum Aufbau fachlicher Leistung genutzt wurde (Leutner et al., 2004).

Die Experimente 1 und 2 konnten, bedingt durch den weitgehend ausbleibenden Nachweis der erfolgreichen experimentellen Manipulation der Komponenten der fächerübergreifenden Problemlösekompetenz, keine Transfereffekte auf die mathematische Domäne aufzeigen. Das mehrwöchige Feldexperiment (Experiment

3) ergab in Übereinstimmung mit den Erwartungen der kognitiven Potenzialausschöpfungshypothese (Leutner et al., 2004) für schwächere Problemlöser Transfereffekte des weitgehend fächerübergreifenden Trainings auf die mathematische Domäne. Ein zusätzlicher McNemar-χ^2-Test in Experiment 3 zur Analyse, ob SuS, die im Prätest vorhandenes kognitives Potenzial zeigten, dieses kognitives Potenzial im Posttest ausschöpften, war nicht signifikant. Die Ergebnisse dieser Analyse sprechen angesichts der potenziellen Fehlerhaftigkeit der Dichotomisierung des Potenzials, die durch die mittlere Korrelation der Domänen Problemlösen und Mathematik und die suboptimale Reliabilität der Tests zu vermuten ist, jedoch nur in begrenztem Maße kognitive gegen die kognitive Potenzialausschöpfungshypothese.

Aus Sicht der Fachdidaktik ist die kognitive Potenzialausschöpfungshypothese *kühn*, da die fachspezifische Wissensbasis weitgehend ignoriert wird (vgl. Carson, 2007). In der Tat liegt es nahe, zu postulieren, dass man, um bessere mathematische Kompetenzen zu erreichen, Mathematik üben müsse. Diese Perspektive ist jedoch unvollständig. Um mathematische Probleme erfolgreich zu lösen, bedarf es neben mathematischer Kompetenz u. a. auch Lesekompetenz und schlussfolgernden Denkens (vgl. Baumert et al., 2007), Problemlösekompetenzen, Weltwissen (Funke, 2003), metakognitiver Fähigkeiten (vgl. Stillman & Mevarech, 2010) und (Leistungs-)Motivation (Schiefele & Csikszentmihalyi, 1995). Aus dieser Perspektive erklärt sich, die verwendete Untersuchung potentieller Transfereffekte von (Komponenten fächerübergreifender) Problemlösekompetenz auf die mathematische Domäne als Ansatz zur Untersuchung der kognitiven Potenzialausschöpfungshypothese. Die kausalen Zusammenhänge zwischen fachlichem und fächerübergreifendem Problemlösen im Laufe der Entwicklung sind jedoch keineswegs völlig erforscht. Die umgekehrte Kausalrichtung, d. h. eine positive Wirkung mathematischen Problemlösens auf allgemeines Problemlösen, wird auch theoretisch angenommen, z. B. in der Mathematikdidaktik (Schwartz & Yerushalmy, 1995; Törner & Zielinski, 1992; Zöttl et al., 2007) und den deutschen Bildungsstandards (KMK, 2003, 2004a). Dort wird die Auffassung vertreten, Mathematik bilde auch allgemeine Problemlösefähigkeiten (vgl. auch die Idee formaler Bildung). Leuders (2010, S. 123) hält vor allem die durch die intensive Beschäftigung mit mathematischen Problemlösen „entwickelten allgemeinen Haltungen“ (im Umgang mit neuen Situationen, Arbeitstechniken) und weniger mathematische Beweistechniken für Transfervehikel.

Während bei PISA 2003 die Leistung in der Domäne Mathematik durch die Leistung in der Domäne Problemlösen vorhergesagt wurde, wurde bei PISA 2012 umgekehrt die Leistung in der Domäne Problemlösen durch die Leistung in der Domäne Mathematik vorhergesagt (siehe Kapitel 4.3.8). Die beiden – mal mehr, mal weniger explizit angenommen – Kausalitätsrichtungen schließen sich im Laufe der individuellen Lern- und Bildungsgeschichte jedoch nicht aus. Menschen lernen Problemlösen auch außerhalb des Mathematikunterrichts und fachlich erworbene mathematische Problemlösekompetenzen sollen in alltäglichen Situationen angewandt werden (KMK, 2004b). Es kann also zu wechselseitigen Rückkopplungsprozessen bzw. reziproken Kausaleffekten kommen (Leutner et al., 2005). Die beiden Hypothesen zur postulierten Kausalrichtung zwischen allgemeiner fächerübergreifender und mathematischer Problemlösekompetenz sind in ihrer allgemeinen Formulierung also miteinander kompatibel, bedürfen zu ihrer spezifischen Prüfung etwaiger Rückkopplungsprozesse jedoch entsprechender längerfristiger längsschnittlicher Forschungsdesigns (vgl. die Forschung zu *skill development* und *skill enhancement*; Helmke & Aken, 1995; Marsh, 1990a), die im Rahmen dieser Arbeit nicht realisierbar waren.

4.2.2 Entwicklung von Aufgaben, Tests, Lernmaterialien

Zur Vorbereitung auf Experiment 1 erfolgte eine Entwicklung von Aufgaben und Tests zum konditionalen Wissen, zur Planungsfähigkeit und zum Handlungswissen (Anhang E, Anhang G). Für das Feldexperiment wurden zusätzliche Aufgaben und Materialien zum Planen und Problemlösen entwickelt (Anhang W, Anhang X). Die Aufgaben sind im Anhang dokumentiert und können für weitere Forschung und Anwendung verwendet werden. Die Trainingsmaterialien stehen den kooperierenden Fachdidaktikern der Naturwissenschaften zur Nutzung und Weiterentwicklung zur Verfügung, die potentielle Transfereffekte analytischen Problemlösens im Bereich der Naturwissenschaften untersuchen möchten (Leutner, Rumann & Wirth, 2011, Januar). Die im Rahmen der Experimente entwickelten Trainingsmaterialien können in leicht modifizierter Form auch im schulischen Kontext eingesetzt werden. Dabei sind mehrere Szenarien denkbar. SuS können selbstständig und freiwillig Zuhause (online) mit den Materialien üben.

Das Computerprogramm aus Experiment 2 kann als verpflichtende Hausaufgabe durchlaufen werden. Lehrerinnen und Lehrer können die Problemlöseaufgaben im Fachunterricht nutzen, indem sie z. B. Aufgaben im Computerraum (mit Internetzugang) bearbeiten lassen (*blended learning*) oder gedruckte Materialien des Trainingsexperiments verwenden. Trotz großer Beliebtheit computerbasierter Lernprogramme im Bereich multimedialer Forschung (Mayer, 2005) sind computerbasierte Lehr-Lernprogramme unserer Projekterfahrung nach allerdings noch nicht in der Breite angekommen, was teilweise an der ungenügenden technischen Ausstattung mancher Schulen liegt.

4.2.3 Umfragesoftware als Experimentalsoftware

Mit der Beschreibung der Möglichkeiten der Nutzung von Umfragesoftware für Lehr-Lernpsychologische Zwecke in Forschung und Praxis (Buchwald, Spoden, Fleischer & Leutner, 2013) ist diese technische Möglichkeit einer breiteren Fachöffentlichkeit vorgestellt worden (Anhang B). Am Beispiel der Software EFS Survey (Globalpark AG, 2011; QuestBack GmbH, 2013), die in Experiment 1 und 2 eingesetzt wurde, wurden unterschiedliche Anwendungsszenarien skizziert, wie Online-Umfragesoftware für verschiedene Formen des verzweigten Lernens und Testens im Bereich der pädagogischen Psychologie und Diagnostik einsetzbar sind.

4.2.4 Strategiewissenstests: Kritik und Forschungsbedarf

In der Diskussion zu Experiment 2 und in Anhang I wurden die Konstruktion und Auswertung von Strategiewissenstests kritisch diskutiert. Wichtige Aspekte und mögliche Fragen für zukünftige Forschung seien kurz zusammengefasst:

1. *Szenario-Konstruktion*: Die potentielle Gleichwertigkeit von Strategien wird bislang kaum berücksichtigt. Das ist insbesondere im Bereich des Problemlösens bedenklich, da eine Einschränkung der Validität zu befürchten ist, wenn nur Probleme ausgewählt werden, bei denen es z. B. genau einen korrekten Lösungsweg gibt oder deren angegebene Handlungsoptionen qualitativ deutlich unterscheidbar sind.

2. *Experten und die Auswahl der Paarvergleiche*: Wie hoch müssen die Experten in ihren Ratings übereinstimmen, damit ein Paarvergleich in das Scoring aufgenommen wird? Wer gilt als Experte? Wie viele Experten werden befragt? Wünschenswert sind einerseits klare Beschreibungen der

Experten und anderseits systematischere Untersuchungen beispielsweise zum Umfang des Expertenkonsens mittels Delphi-Technik (Hasson, Keeney & McKenna, 2000), zu Konsequenzen unterschiedlichen Expertenkonsens‘oder zur Anzahl der befragten Experten.

3. *Konsequenzen des Scorings*: Die typischerweise gewählten Scoringregeln (Neuenhaus et al., 2010; Thillmann, 2007) vergleichen die Handlungsoptionen nur über Größer-kleiner-Relationen und betrachten keine Differenzen. Dadurch gehen potentiell Informationen verloren. Das beginnt beim Expertenrating und betrifft auch mögliche Niveauabweichung zwischen Experten- und Probandenrating sowie eine Differenzierung der Lösungsgüte zwischen Probanden. Zusätzlich macht die teilweise verwendete *0*.5-Punkte-Regel für gleich bewertete Handlungsoptionen die Rohpunktzahl anfällig für Verweigerer, die bei mindestens einem Szenario nur eine Spalte durchkreuzen. Wer z. B. alle Alternativen mit *mangelhaft* bewertet, erhält die Hälfte der maximal möglichen Punkte des Szenarios. Ebenso beträgt die Ratewahrscheinlichkeit pro Paarvergleich 50 Prozent (Schlagmüller & Schneider, 2007; Anhang I). Außerdem werden bei der Verwendung von Rohpunktsummen die Szenarien eines Tests unterschiedlich gewichtet, wenn sie unterschiedlich viele Paarvergleiche enthalten. Vergleiche verschiedener Scoringvarianten für Strategiewissenstests fehlen bislang, erscheinen aber wünschenswert.
4. *Testlet-Effekte*: Tests zum konditionalen Wissen bestehen aus verschiedenen Szenarien. Wie groß der Effekt des Kontexts der Szenarien bzw. die Generalisierbarkeit der Strategien ist, bleibt oftmals unklar. Dazu sind Studien mit größeren Stichproben und Strukturgleichungsmodelle als Auswertungsmethode nötig, die entsprechende Methodenfaktoren (siehe z. B. Brown, 2006) berücksichtigen.

4.2.5 Planungskompetenztraining in der Jahrgangsstufe 9

Als Teil von Experiment 3 erfolgte die Untersuchung des Einsatzes des Aachener Planungskompetenztrainings (Arling & Spijkers, 2012) in der Jahrgangsstufe 9, das bei einer so jungen Stichprobe noch nicht eingesetzt worden war. Das ursprünglich für Erwachsene entwickelte Planungskompetenztraining wurde zwar schon in der Jahrgangsstufe 11 an Gesamtschulen eingesetzt, bislang jedoch noch nicht in der Jahrgangsstufe 9. Die Ergebnisse des Treatment-Checks aus Experiment 3 zeigten, dass das Planungskompetenztraining erfolgreich in der Jahrgangsstufe 9 eingesetzt werden kann. Dabei ist stark auf eine manualkonforme Durchführung und Motivation der SuS zu achten, um Trainingserfolge zu erzielen. Aufgrund der im Vergleich zu erwachsenen Rehabilitanden (Arling & Spijkers, 2012) relativ niedrigen Leistungen der SuS im Prätest sollte gegebenenfalls vor dem Training eine längere Instruktionsphase erfolgen, in der die Anforderungen und der Ablauf stärker verdeutlicht werden. Aufgrund der Textlastigkeit des Verfahrens empfiehlt sich ferner nach Möglichkeit im Forschungskontext die Erhebung von Lesekompetenz als zusätzlicher Kovariate (z. B. mit der Lesetestbatterie für die Klassenstufen 8-9; Bäuerlein, Lenhard & Schneider, 2012).

4.3 Limitationen und zukünftige Forschung

Konkrete empirische Forschung hat fast immer Schwächen und impliziert weitergehende Forschung. Daher werden abschließend einige Grenzen der beschriebenen Experimente und Ausblicke auf weitere Forschung diskutiert.

4.3.1 Notwendigkeit reliablerer Instrumente

Die eingesetzten Tests und Fragebögen weisen teilweise geringe Reliabilitäten auf. Insbesondere der PS-SJT (Bertling, 2012) und der kleine Transfertest zum Problemlösen in Experiment 3, der im Anschluss an das Aachener Planungskompetenztraining (Arling & Spijkers, 2012) eingesetzt wurde, sollten vor einem erneuten Einsatz überarbeitet werden.

4.3.2 PISA-Aufgaben und -Metrik

Der Aufgabenpool zum fächerübergreifenden Problemlösen bei PISA 2003 (OECD, 2004b) ist eher klein. Zukünftige Forschung sollte diesen Aufgabenpool für Trainings und (adaptive) Testzwecke gegebenenfalls erweitern. Dabei erscheinen kognitive Diagnosemodelle (Huff & Goodmann, 2007; Kunina-Habenicht, Wilhelm, Matthes & Rupp, 2010), die auf theoretischer Basis (a priori) Anforderungen und Prozesse von Aufgaben beschreiben, für eine theoriegeleitete Item-Generierung und Testung vielversprechend (vgl. Leuders & Sodian, 2013).

Zur Einordnung der Ergebnisse der vorliegenden Arbeit ist zu beachten, dass die Experimente 1 bis 3 *nicht* auf der PISA-Metrik basieren, die die Leistungen der

SuS in Deutschland im *internationalen Vergleich* sieht. Ein internationaler Vergleich war nicht Ziel der Experimente.

Ferner sollten die Vorwürfe von Wuttke (2006), dass die Kompetenzschätzungen technische Fehler aufweisen (vgl. Kapitel 2.4.3.1), nicht nur wie bei Köller (2006) rhetorisch zurückgewiesen, sondern mathematisch geprüft werden. Der praktische Nutzen einen möglichen Auswertungsfehler von PISA 2003 mehr als 11 Jahre und drei PISA-Studien später aufzuzeigen, dürfte allerdings eher gering sein. Zusätzlich könnte man untersuchen, ob die Diskrepanz der Kompetenzen in den Domänen Mathematik und Problemlösen bestehen bleiben, wenn man *random item effects models* betrachtet (Fox, 2010; Fox & Verhagen, 2010), um z. B. länderspezifische Itemparameter als zufällige Schwankungen der Itemparameter anzunehmen ohne die länderübergreifende Messinvarianz zu verletzen.

4.3.3 Kognitive, metakognitive und motivationale Voraussetzungen des Problemlösens

Erfolgreiches Problemlösen bedarf seitens des Problemlösers oftmals einer *triple alliance* (Short & Weissberg-Benchell, 1989) von Kognition, Metakognition und Motivation (vgl. Hasselhorn, 1992; Mayer, 1998). Beschränkungen hinsichtlich dieser drei Aspekte werden nachfolgend mit Blick auf die empirischen Studien kurz diskutiert.

Problemlösen ist eine der kognitiv anspruchsvollsten Tätigkeiten des Menschen (Funke, 2003). Neben kognitiven Fertigkeiten, die für den Problemlöseprozess relevant sind (vgl. Kapitel 2.1.3), bedarf es einer Reihe von basalen Grundfertigkeiten, um die vorliegenden Problemlöseaufgaben zu bearbeiten und zu lösen. Zunächst muss man über die notwendigen Sprach- und Lesekenntnisse verfügen, um die Materialien in der zur Verfügung stehenden Zeit überhaupt lesen zu können. Einzelnen Berichten der durchführenden Trainern zufolge gab es SuS, die Probleme beim Lesen der Aufgabentexte hatten. Wenn einzelne SuS in der neunten Jahrgangsstufe noch nicht über die Basisfertigkeit des flüssigen Lesens verfügen, ist es zumindest fraglich, ob sie textbasierte Problemlöseaufgaben adäquat verstehen können. Zukünftige Studien zum analytischen Problemlösen sollten zusätzlich zu anderen Kovariaten, wenn möglich, auch die Lesekompetenz erfassen. Das gilt v.a. bei textlastigen Aufgaben, wie sie etwa bei PISA 2003 (OECD, 2004b) häufig vorkommen.

Probleme sind immer Probleme für jemanden (Funke, 2003). In der Schule müssen SuS die gestellten Probleme zumindest temporär zu ihren eigenen Problemen machen. Problemlösen kann durch die Konfrontation mit einer nicht-routinemäßig bearbeitbaren Aufgabe physiologisches Arousal erzeugen, dessen Interpretation zu positiven oder negativen Emotionen führen kann (McLeod, 1988). Die berichtete Beobachtung der Trainer in Experiment 3, dass einige SuS Probleme erst gar nicht zu lösen versuchten, wenn ihnen nicht unmittelbar ein Lösungsweg einfiel (teilweise begleitet von Äußerungen wie „Das haben wir nicht geübt.“, „Ich sehe, dass ich das nicht kann.“, „Denken ist einfach nix für mich.“), zeigt die Notwendigkeit, sich explizit mit Problemen und Problemlösen in der Schule zu beschäftigen. SuS dazu zu bewegen, sich mit Aufgaben auseinanderzusetzen, die man beim ersten Lesen nicht versteht oder glaubt nicht lösen zu können, ist eine pädagogische Herausforderung für Lehrer, nicht nur im Mathematikunterricht.

> “Man darf nicht vergessen, daß die Problemlösetätigkeit eine harte Arbeit für die Schüler bedeutet, die oft viel Willenskraft und Ausdauer erfordert“ (Ambrus, 1991, S. 92).

Obwohl die Bedeutung affektiver Faktoren (z. B. Frustration, Panik, Befriedigung, Freude, Einstellungen) beim Problemlösen oft betont wird (Pehkonen, 1991), wird sie kaum systematisch untersucht (McLeod, 1988). Zukünftige Kompetenz-Forschung sollte daher Lern- und Testmotivation stärker berücksichtigen (Asseburg & Frey, 2013; Jansen, Schroeders & Stanat, 2013).

4.3.4 Adaptivität

Die in den Experimenten 1 bis 3 verwendeten Materialien verwenden Binnendifferenzierung nur in geringem Maße. Dadurch besteht die Möglichkeit, leistungsschwächere SuS zu überfordern und leistungsstärkere SuS zu unterfordern. Zukünftige Forschung zur kognitiven Potenzialausschöpfungshypothese sollte stärker aus differentieller Perspektive betrieben werden, wie sie in der mathematischen Fachdidaktik bereits zum Einsatz kommen (z. B. mittels *Blütenaufgaben*, einer offenen Aufgabenform, die selbst-differenzierende Teilaufgaben enthält, Bruder, Reibold & Weise, 2014; von der Groeben & Kaiser, 2011). Dabei können stärker adaptive Verfahren zum Einsatz kommen, um eine bessere Aufgabenauswahl beim Lernen (Leutner, 1992a) und computerbasiertem Testen (Frey & Seitz, 2009; van der Linden & Glas, 2000, 2010) zu ermöglichen.

4.3.5 Alternative Operationalisierungen und Foci

Empirische Forschung erlaubt fast immer verschiedene Vorgehensweisen. Die in den Experimenten 1 bis 3 verwendeten Operationalisierungen sind nicht die einzig denkbaren Möglichkeiten, Aspekte der kognitiven Potenzialausschöpfungshypothese zu untersuchen.

Bei Beibehaltung des Fokus auf kognitive Aspekte sind auch andere Möglichkeiten denkbar, Kompetenzerwerb- und entwicklung zu untersuchen. Kompetenzentwicklung muss theoretisch nicht eindimensional sein, wie es durch eindimensionale Skalierung bei PISA 2003 (OECD, 2004b) angenommen wird. Intraindividuelle Kompetenzentwicklung kann bedeuten, (a) schwierigere Probleme zu lösen (Problemkomplexität), (b) ähnliche Probleme qualitativ besser zu lösen (bessere Lösungswege, schnellere Bearbeitung), (c) Probleme eigenständiger zu lösen, (d) Probleme über verschiedene Zeitpunkte hinweg stabil zu lösen (Gut-Glanzmann, 2012, 2014). Geeignete Instrumente, die entsprechend zu entwickeln wären, erlauben also möglicherweise, Kompetenzentwicklung im Bereich fachlichen und fächerübergreifenden Problemlösens jeweils umfassender bzw. multidimensional zu erfassen.

Alternative Forschungszugänge und zusätzliche Forschungsfragen stellen sich beispielsweise aus *ländervergleichender Perspektive*: Wie kommen die Unterschiede zwischen den Kompetenzen im Bereich Problemlösen und Mathematik bei PISA 2003 bzw. 2012 in verschiedenen Ländern zustande? Welche Gemeinsamkeiten haben die einzelnen Ländergruppen, die beim Problemlösen höhere / gleiche / niedrigere Ergebnisse erzielen als in der Mathematik? Was unterscheidet diese drei Ländergruppen? Wie unterscheidet sich der Mathematikunterricht in diesen Ländern? Zur Beantwortung dieser Fragen ist Forschung aus einer international vergleichenden fachdidaktischen Perspektive nötig. Ähnliche Fragen lassen sich jedoch auch für den innerdeutschen Bundesländervergleich (Leutner et al., 2005) stellen.

Darüber hinaus ist der gewählte *Untersuchungsfokus* änderbar. So beschränkt sich der gewählte Forschungs- und Analysezugang weitgehend auf aggregierte kognitive Schülerleistungen. Im Rahmen des DFG-Schwerpunktprogramms „Kompetenzmodelle" liegt diese selbstgewählte forschungsmethodische Beschränkung auf kognitive Aspekte verbreitet vor (Wilhelm & Nickolaus, 2013). Insbesondere für das komplexe Problemlösen (Dörner, 1984) und mathematische Problemlösen (McLeod, 1988) liegen jedoch längst Arbeiten vor, die die Rolle affektiver Variablen betonen. Die stärkere Einbeziehung solcher Variablen und zusätzlicher kognitiver und metakognitiver Prozessmaße sind weitere Möglichkeiten die Untersuchungsperspektive je nach spezifischer Fragestellung zu erweitern. Denkbar sind beispielsweise die Methode lauten Denkens (van Someren, Barnard & Sandberg, 1994), Lerntagebücher (Ehmann, 2008; Perels et al., 2003; Perels, Schmitz & Bruder, 2005), Logfile-Analysen (de Jong, 2005) Videoanalysen des Verhaltens beim Problemlösen (de Jong, 2005) oder retrospektives Debriefing (Taylor & Dionne, 2000).

4.3.6 Mathematisches Problemlösen lehren und lernen

Mathematisches Problemlösen ist nicht nur Gegenstand von Begabtenförderungsmaßnahmen (Holling, Preckel, Vock & Schulze Willbrennig, 2004). In einem breiteren Verständnis des Problembegriffs (z. B. Bruder, 2000; Schoenfeld, 1992) wird Problemlösen in Einklang mit den Kernlehrplänen (KMK, 2003, 2004a) als Schlüsselqualifikation bzw. Kernkompetenz betrachtet. Problemlösen soll gemäß dem politischen Willen ein expliziterer Gegenstand im Mathematikunterricht sein. Nicht zuletzt der Umgang mit schwierigen Aufgaben, das kreative Finden neuer Lösungen sowie eine neue *Kultur des Fragens (Bruder, 2002; Vollmer, 1992)* und des *Voran-Scheiterns* bzw. *Empor-Irrens* (vgl. *Lesch, 2011; Popper, 2005; Vollmer, 1995*), in der produktiv mit Schülerfehlern umgegangen wird (Katzenbach, 2004), halten zunehmend Einzug in den Mathematikunterricht. Dabei sollte auch der Aspekt des *problem posings* (Cai, 1998; Chang, Wu, Weng & Sung, 2012; Knapp, Pfaff & Werner, 2010; Silver, 1994; Silver & Cai, 1996) berücksichtigt werden, d. h. SuS sollen lernen, selbst mathematische Fragen und Aufgaben zu stellen.

Für amerikanische Grundschulen (Krulik & Rudnick, 1998) und Highschools (Krulik & Rudnick, 1996) gibt es seit langem Ansätze, die Schlussfolgern und Problemlösen explizit zum Gegenstand haben (Fuchs et al., 2006). In den letzten Jahren sind vermehrt deutschsprachige Unterrichtsmaterialien erschienen, die Mathematik im Alltag (Bettner, 2011; Christensen & König, 2011), Modellieren (z. B. Maaß & Siller, 2014) und Problemlösen (z. B. Bruder & Collet, 2011; siehe Reiss & Törner, 2007 für weitere Quellen) fokussieren. In Deutschland werden solche Materialien inzwischen zunehmend auch für Mathematik in der Grundschule entwickelt (Selter, 2009). Dies trägt dem Gedanken Rechnung, dass Problemlösenlernen ein langfristiges Anliegen ist, das im Mathematikunterricht frühzeitig begonnen werden sollte (KMK, 2005d). Auch für den Vorschulbereich gibt es einzelne metakognitive Förderansätze mathematischen Problemlösens (z. B. Annevirta & Vauras, 2006). Doch gibt es bislang weder breit verwendete Ansätze noch entsprechende umfassende empirische Evaluationen der auf diesen Materialien und Ansätzen beruhenden Unterrichtseinheiten.

Um fächerübergreifendes und mathematisches Problemlösen (bzw. Modellieren) in der Schule zu fördern, reichen die Existenz von Unterichtsmaterialien und die Kenntnis einer positiven Korrelation von fächerübergreifendem Problemlösen und Mathematik nicht aus, um zu bestimmen, wann und wie fächerübergreifende Problemlösekompetenz im schulischen Mathematikunterricht konkret nutzbar gemacht werden kann (Renkl, 2012). Man muss auch auf der Ebene der Lehrerausbildung und der Lehrerweiterbildung ansetzen (Brunner et al., 2006) und diese Maßnahmen evaluieren (z. B. Collet, 2009).

4.3.7 Problemlösen in den Naturwissenschaften und weitere Konstruktvalidierungen

Die kognitive Potenzialausschöpfungshypothese postuliert ungenutztes Potenzial nicht nur für die mathematische Domäne, sondern auch für die naturwissenschaftliche Domäne (Leutner et al., 2006; Rumann et al., 2010; Stawitz, Rumann, Fleischer & Wirth, 2008). Die Experimente 1 bis 3 untersuchten die Hypothese nur bezüglich der mathematischen Domäne. Sie lassen daher keine Verallgemeinerung auf andere Domänen zu. Fragen potentieller Trainings- und Transfereffekte in diesen Bereichen werden im Rahmen weiterer Forschung Berücksichtigung finden (RU 1437/4-3). Aus Sicht der naturwissenschaftlichen Fachdidaktiken erscheint insbesondere beim Experimentieren als naturwissenschaftlicher Methode des Problemlösens weiterer Forschungsbedarf.[22] Beispielsweise könnten mögliche Effekte eines modifizierten in den Fachunterricht integrierten Trainings kognitiver und metakognitiver Strategien untersucht werden, da zahlreiche Studien den Zusammenhang metakognitiver Variablen mit naturwis-

[22] Vielen Dank an die Teilnehmer der „Denkfabrik“ im Rahmen der Jahrestagung des DFG-Schwerpunktprogramms „Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen" (SPP 1293) am 7.10.2013 für ihre Anregungen und ihr Feedback zu möglichen Anschlussprojekten.

senschaftlicher Leistung bzw. naturwissenschaftlichem Problemlösen belegen (Scherer & Tiemann, 2012; Shahat et al., 2013). Hierzu bedarf es weiterer Anforderungsanalysen und empirischer Studien. Da Experimentieren oftmals komplexes Problemlösen ist, stellen sich in diesem Zusammenhang auch Fragen, welche analytischen Aspekte beim domänenspezifischen komplexen Problemlösen besonders wichtig und trainierbar sind. In diesem Rahmen erscheint auch eine weitergehende Konstruktvalidierung nötig (vgl. Leutner et al., 2012), die z. B. die Struktur der Teilkomponenten des fächerübergreifenden Problemlösens und ihre Beziehung zu weiteren etablierten Konstrukten wie Arbeitsgedächtnis (vgl. Bühner, Kröner & Ziegler, 2008; Hoffman & Schraw, 2009; Keeler & Swanson, 2001; Swanson, 2011; Swanson & Cooney, 1993) und Intelligenz (vgl. Humphreys, 1976; Wilhelm & Nickolaus, 2013; Wüstenberg, Greiff & Funke, 2012) analysiert und die Struktur bzw. Zusammenhänge der Teilkomponenten des Problemlösens in verschiedenen Domänen untersucht.

4.3.8 Ist das Befundmuster der PISA-2003-Studie noch aktuell? Ein Blick auf PISA 2012

Das Schulsystem in Deutschland ist vom Föderalismus geprägt und ständigen Reformen unterworfen. Auch nach den Ergebnissen der PISA-Studien ist das deutsche Schulsystem kritisiert und verändert worden (einheitliche Bildungsstandards, Kompetenzorientierung, achtjähriges Gymnasium (G8), teilweise Abschaffung der Hauptschule, mehr Ganztagsschulen, teilweise Wiederabschaffung von G8, usw.). Wie haben sich die mathematischen Kompetenzen von SuS in Deutschland seit 2003 verändert? Welche Konsequenzen hat dies für die kognitive Potenzialausschöpfungshypothese? Bei PISA 2009 beispielsweise liegt die Leistung der SuS in Deutschland mit $M = 513$ über dem OECD-Durchschnitt von $M = 496$ und 10 Punkte über dem Ergebnis von PISA 2003, wobei auch die Streuung mit $SD = 98$ über dem OECD-Durchschnitt von $SD = 92$ liegt (Frey, Heinze, Mildner, Hochweber & Asseburg, 2010). Inhaltlich wurde dies als Kompetenzzuwachs fünfzehnjähriger SuS in Höhe eines Drittel Schuljahres interpretiert. Der Zuwachs im Vergleich zu PISA 2003 kam insbesondere dadurch zustande, dass der Anteil von SuS mit Leistungen unterhalb von Kompetenzstufe 1 gesunken war (Frey et al., 2010). Allerdings wurde Problemlösen im Rahmen der PISA-Studien erst wieder bei PISA 2012 gemessen. Mit Blick auf die kognitive Potenzialausschöpfungshypothese ergeben sich einige spannende Fragen: Existiert die in der PISA-Studie 2003 gefundene Diskrepanz zwischen mathematischer und fächerübergreifender Problemlösekompetenz neun Jahre später bei PISA 2012 noch? Wird das kognitive Potenzial der SuS in Deutschland inzwischen besser zum Aufbau fachlicher Fähigkeiten genutzt? Inwieweit schöpft der Fachunterricht in Deutschland inzwischen das kognitive Potenzial der SuS besser aus? Schaffen es Länder wie Deutschland, die bei PISA 2003 eine Diskrepanz zugunsten des Problemlösens aufwiesen, besser in Mathematik abzuschneiden? Die letzte Frage kann für Deutschland bejaht werden: Bei PISA 2012 verbesserte sich Deutschland im Mathematikranking deutlich, und der Anteil an SuS auf den untersten Kompetenzstufen ging zurück (Prenzel, Sälzer, Klieme & Köller, 2013a). Heutige Neuntklässler weisen im Mittel ungefähr einen Leistungsvorsprung von einem halben Schuljahr im Vergleich zu den Neuntklässlern bei PISA 2003 auf (Prenzel, Sälzer, Klieme & Köller, 2013b). Bedeutet das, dass inzwischen das vorhandene kognitive Potenzial besser im schulischen Mathematikunterricht genutzt wird? Diese Frage ist schwer zu beantworten. Die Gründe für die verbesserten Mathematiktestergebnisse können vielfältig sein. Bei PISA 2012 erzielen SuS in Deutschland beim Problemlösen im Mittel überdurchschnittliche Werte, die etwas niedriger sind, als durch die fachliche Leistung in Mathematik zu erwarten wäre (OECD, 2014). Man beachte, dass das der kognitiven Potenzialausschöpfungshypothese zugrunde liegende Argumentationsmuster von PISA 2003 bei PISA 2012 invers verwendet wird. Die Rolle von Prädiktor und Kriterium wurde getauscht. Man beachte ferner, dass sich der Assessmentschwerpunkt innerhalb der Domäne Problemlösen von PISA 2003 zu PISA 2012 verschoben hat (Greiff et al., 2013). Beim internationalen Teil von PISA 2003 ist mit fächerübergreifendem Problemlösen analytisches Problemlösen gemeint, das als Papier-Bleistift-Test erhoben worden ist. Bei PISA 2012 liegt der Fokus auf komplexem Problemlösen, das computerbasiert erhoben wurde. Es ergeben sich also für die Vergleichbarkeit der Ergebnisse von PISA 2003 und PISA 2012 hinsichtlich des Problemlösens und der kognitiven Potenzialausschöpfungshypothese Einschränkungen, die in der Veränderung des Testschwerpunkts und der Erhebungsmethode begründet liegen. Es ist somit weitere Forschung nötig, um die Frage zu beantworten, ob das kognitiven Potenzial von SuS zum Erwerb fachlicher Problemlösekompetenz im Fach Mathematik mehr als 10 Jahre nach PISA 2003 besser ausgeschöpft wird.

5 Literaturverzeichnis

AAAS (American Association for the Advancement of Science). (1993). *Benchmarks for science literacy*. New York, NY: Oxford University Press.

Ambrus, A. (1991). Problemlösen im ungarischen Mathematikunterricht. *ZDM Mathematics Education, 23,* 84–92.

Anderson, J. R. (2013). *Kognitive Psychologie* (7. Aufl.). Berlin: Springer VS.

Annevirta, T. & Vauras, M. (2006). Developmental changes of metacognitive skill in elementary school children. *Journal of Experimental Education, 74,* 197–225.

Arbeitsgruppe der Fachdidaktik der Mathematik der TU Darmstadt. (n.d.). *Arbeitsblatt rund ums Vorwärtsarbeiten.* Zugriff am 03.02.2013. Verfügbar unter http://www.problemloesenlernen.dvlp.de/files/material/klasse5/Laengerfristige_Aufgaben.pdf

Arbinger, R. (1997). *Psychologie des Problemlösens*. Darmstadt: Primus.

Arcavi, A. & Friedlander, A. (2007). Curriculum developers and problem solving: the case of Israeli elementary school projects. *ZDM Mathematics Education, 39,* 355–364.

Arling, V. (2006). *Entwicklung und Validierung eines Verfahrens zur Erfassung von Planungskompetenz in der beruflichen Rehabilitation. Der "Tour-Planer"*. Berlin: Logos.

Arling, V. & Spijkers, W. (2012). *Planungskompetenztraining [PKT]*. Aachen: RWTH-Aachen, Lehr- und Forschungsgebiet Berufliche Rehabilitation.

Aronson, J., Lustina, M. J., Good, C., Keough, K., Steele, C. M. & Brown, J. (1999). When white men can't do math. Necessary and sufficient factors in stereotype threat. *Journal of General Internal Medicine, 35,* 29–46.

Artelt, C., Beinicke, A., Schlagmüller, M. & Schneider, W. (2009). Diagnose von Strategiewissen beim Textverstehen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 41,* 96–103.

Artelt, C., Neuenhaus, N., Lingel, K. & Schneider, W. (2012). Entwicklung und wechselseitige Effekte von metakognitiven und bereichsspezifischen Wissenskomponenten in der Sekundarstufe. *Psychologische Rundschau, 63,* 18–25.

Asseburg, R. & Frey, A. (2013). Too hard, too easy, or just right? The relationship between effort or boredom and ability-difficulty fit. *Psychological Test and Assessment Modeling, 55,* 92–104.

Australian Council for Educational Research. (n.d.). *PISA: Examples of computer-based items.* Zugriff am 22.12.2013. Verfügbar unter http://cbasq.acer.edu.au/index.php?cmd=cbaItemPreview&unitVersionId=824

Barnes, H. (2005). The theory of Realistic Mathematics Education as a theoretical framework for teaching low attainers in mathematics. *Pythagoras, 61,* 42–57.

Barnett, S. M. & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128,* 612–637.

Baruk, S. (1989). *Wie alt ist der Kapitän? Über den Irrtum in der Mathematik*. Basel: Birkhäuser.

Bassok, M. (2003). Analogical transfer in problem solving. In J. E. Davidson & R. J. Sternberg (Hrsg.), *The Psychology of Problem Solving* (S. 343–372). Cambridge: University Press.

Bassok, M., Wu, L. L. & Olseth, K. L. (1995). Judging a book by its cover. Interpretative effects of content on problem-solving transfer. *Memory & Cognition, 23,* 354–367.

Bäuerlein, K., Lenhard, W. & Schneider, W. (2012). *Lesetestbatterie für die Klassenstufen 8 -9. Verfahren zur Erfassung der basalen Lesekompetenz und des Textverständnisses*. Göttingen: Hogrefe.

Baumert, J., Brunner, M., Lüdtke, O. & Trautwein, U. (2007). Was messen internationale Schulleistungsstudien? - Resultate kumulativer Wissenserwerbsprozesse. *Psychologische Rundschau, 58,* 118–128.

Berardi-Coletta, B., Buyer, L. S., Dominowski, R. L. & Rellinger, E. R. (1995). Metacognition and problem solving. A process-oriented approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 205–223.

Berthold, S., Diehl, R. & Kühne, J. (2010). *Methodentraining: Präsentationstechniken. Handlungsorientierte Materialien zum mündlichen und mediengestützten Vortrag*. Buxtehude: Persen.

Bertling, J. (2012). *Student questionnaire for PISA 2012 initial main survey analyses - students' approaches to problem solving*. Heidelberg, Germany: Questionnaire Expert Group Meeting & Problem Solving Expert Group Meeting.

Bettner, M. (2011). *Mathematik im Alltag. Kopiervorlagen für die Sekundarstufe I* (6. Aufl.). Buxtehude: Persen.

Bielaczyc, K., Pirolli, P. L. & Brown, A. L. (1995). Training in self-explanation and self-regulation strategies. Investigating the effects of knowledge acquisition activities on problem solving. *Cognition and Instruction, 13,* 221–252.

Blanchard, P. N. & Thacker, J. W. (2010). *Effective training. Systems, strategies, and practices* (4. Aufl.). Upper Saddle River, N.J.: Pearson/Prentice Hall.

Bless, H., Wänke, M., Bohner, G., Fellhauer, R. F. & Schwarz, N. (1994). Need for cognition. Eine Skala zur Erfassung von Engagement und Freude bei Denkaufgaben. *Zeitschrift für Sozialpsychologie, 25,* 147–154.

Blum, W., Drüke-Noe, C., Hartung, R. & Köller, O. (2006). *Bildungsstandards Mathematik: Konkret*. Berlin: Cornelsen Scriptor.

Blum, W. & Niss, M. (1989). Mathematical problem solving, modelling, applications, and links to other subjects. State, trends and issues in mathematics instruction. In W. Blum (Hrsg.), *Modelling, applications and applied problem solving* (S. 1–21). Chichester: Ellis Horwood Limited.

Blum, W., Galbraith, P. L., Henn, H. W. & Niss, M. (Hrsg.). (2007). *Modelling and applications in mathematics education. The 14th ICMI study* (New ICMI study series, Bd. 10). New York, NY: Springer.

Boekaerts, M. (1997). Self-regulated learning. A new concept embraced by researchers, policy makers, educators, teachers, and students. *Learning and Instruction, 7,* 161-186.

Boekaerts, M. & Otten, R. (1993). Handlungskontrolle und Lernanstrengung im Schulunterricht. *Zeitschrift für Pädagogische Psychologie, 7,* 109–116.

Böhme, K., Robitzsch, A. & Busé, A.-K. (2010). Zur Abgrenzung des Hörverstehens gegenüber dem Leseverstehen mit Hilfe schwierigkeitsbestimmender Merkmale bei der Entwicklung von Testaufgaben. In V. Bernius & M. Imhof

(Hrsg.), *Zuhörkompetenz in Unterricht und Schule. Beiträge aus Wissenschaft und Praxis* (S. 81–104). Göttingen: Vandenhoeck & Ruprecht.

Borromeo Ferri, R., Greefrath, G. & Kaiser, G. (Hrsg.). (2013). *Mathematisches Modellieren für Schule und Hochschule. Theoretische und didaktische Hintergründe* (Realitätsbezüge im Mathematikunterricht, Bd. 1). Wiesbaden: Springer.

Bortz, J. & Döring, N. (Hrsg.). (2006). *Forschungsmethoden und Evaluation* (4. Aufl.). Heidelberg: Springer.

Bortz, J., Lienert, G. A. & Boehnke, K. (2008). *Verteilungsfreie Methoden in der Biostatistik* (2. Aufl.). Heidelberg: Springer.

Bouffard-Bouchard, T. (1994). Effect of activating conditional knowledge on self-efficacy and comprehension monitoring. *International Journal of Behavioral Development, 17,* 577–592.

Bransford, J. D., Sherwood, R., Vye, N. & Rieser, J. (1986). Teaching thinking and problem solving. Reserach foundations. *American Psychologist, 41,* 1078–1089.

Breuer, K. (2014). *Exploration potentieller Faktoren des Trainingserfolgs eines Planungskompetenztrainings*. Unveröffentlichte schriftliche Hausarbeit im Rahmen der Ersten Staatsprüfung für das Lehramt an Berufskollegs, Universität Duisburg-Essen. Essen.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research* (Methodology in the social sciences). New York, NY: Guilford Press.

Bruder, R. (2000). Problemlösen im Mathematikunterricht - ein Lernangebot für alle? *Mathematische Unterrichtspraxis. Zeitschrift für den Mathematikunterricht, 21* (1), 2–11.

Bruder, R. (2002). Lernen, geeignete Fragen zu stellen. Heuristik im Unterricht. *mathematik lehren* (115), 4–9.

Bruder, R. (2003). Konstruieren - auswählen - begleiten. Über den Umgang mit Aufgaben. *Friedrich Jahresheft "Aufgaben. Lernen fördern - Selbstständigkeit entwickeln",* 12–15.

Bruder, R. (2003, Januar). *Methoden und Techniken des Problemlösenlernens*. SINUS-Materialien. Zugriff am 27.12.2013. Verfügbar unter http://nibis.ni.schule.de/~as-lg/Mathe2/Dokumente/probleme%20loesen.pdf

Bruder, R. & Collet, C. (2011). *Problemlösen lernen im Mathematikunterricht* (Scriptor Praxis: Mathematik). Berlin: Cornelsen Scriptor.

Bruder, R., Gürtler, T., Schmitz, B. & Perels, F. (2002). Problemlösenlernen in Verbindung mit Selbstregulation. *mathematik lehren* (115), 59–62.

Bruder, R., Reibold, J. & Weise, T. (2014). *Binnendifferenziertes Aufgabenmaterial für den Mathematikunterricht der Sek I* (MABIKOM - Mathematische binnendifferenzierende Kompetenzentwicklung im Mathematikunterricht). Braunschweig: Schroedel.

Brüning, L. (2011). *Methodentraining: Vortragen, Präsentieren, Referieren. Praktische Unterrichtsmaterialien für die Sekundarstufe* (3. Aufl.). Donauwörth: Auer.

Brunner, M., Kunter, M., Krauss, S., Baumert, J., Blum, W., Dubberke, T. et al. (2006). Welche Zusammenhänge bestehen zwischen dem fachspezifischen Professionswissen von Mathematiklehrkräften und ihrer Ausbildung sowie beruflichen Fortbildung. *Zeitschrift für Erziehungswissenschaft, 9,* 521–544.

Buchwald, F. (2012a). *Arbeitsblatt "Handynutzertypen"*. Unveröffentlichtes Manuskript, Essen: Lehrstuhl für Lehr-Lernpsychologie, Universität Duisburg-Essen.

Buchwald, F. (2012b). *Arbeitsblatt "Kuchenstücke"*. Unveröffentlichtes Manuskript, Essen: Lehrstuhl für Lehr-Lernpsychologie, Universität Duisburg-Essen.

Buchwald, F. (2012c). *Arbeitsblatt "Nachdenken über Alternativen"*. Version 1.1. Unveröffentlichtes Manuskript, Essen: Lehrstuhl für Lehr-Lernpsychologie, Universität Duisburg-Essen.

Buchwald, F. (2014). *Problematic aspects of strategy knowledge tests*. Internal Report. Essen: Department of Instructional Psychology, University Duisburg-Essen.

Buchwald, F., Arling, V., Fleischer, J. & Leutner, D. (2013). Das Aachener Planungskompetenztraining. Analyse eines Trainingsexperiments in der Jahrgangsstufe 9 [Abstract]. In A. Beuter, G. Bornemann, C. Jörns & C. Mähler (Hrsg.), *"gemeinsam verschieden" - 14. Fachgruppentagung Pädagogische Psychologie* (S. 84–85). Hildesheim. Zugriff am 29.11.2013. Verfügbar unter http://www.paeps-hildesheim-2013.de/images/Tagungsprogramm/130914_Abstractband.pdf.

Buchwald, F., Fleischer, J., Arling, V. & Leutner, D. (2014, February). *Planning skills and problem solving competence. The question of transfer*. Paper presented at the 1st Duisburg-Essener Workshop on Problem Solving and Dynamic Decision Making (EPSDDM), Essen, Germany.

Buchwald, F., Fleischer, J. & Leutner, D. (in preparation). A field experimental study of analytical problem solving competence. Investigating effects of training and transfer. *Thinking Skills and Creativity*.

Buchwald, F., Fleischer, J. & Leutner, D. (2012, Oktober). *Erste labor-experimentelle Befunde zur Potentialausschöpfungshypothese*. Roundtable-Discussion während des Jahreskolloquium des DFG-Schwerpunktprogramms „Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen“, München.

Buchwald, F., Fleischer, J. & Leutner, D. (2013, März). *Training zentraler Komponenten fächerübergreifender Problemlösekompetenz. Ein Experiment zur kognitiven Potentialausschöpfungshypothese*. Poster während der 1. Tagung der Gesellschaft für empirische Bildungsforschung (GEBF), Kiel.

Buchwald, F., Fleischer, J., Rumann, S., Wirth, J. & Leutner, D. (in preparation). Training of components of problem solving competence – An experimental study on the cognitive potential exploitation hypothesis. In D. Leutner, J. Fleischer, J. Grünkorn & E. Klieme (Hrsg.), *Competence assessment in education - Research, models and instruments.* Berlin: Springer.

Buchwald, F., Spoden, C., Fleischer, J. & Leutner, D. (2013). Verzweigte Lernumgebungen und Tests mit EFS Survey 8. *Diagnostica, 59,* 113–117.

Bühner, M. (2010). *Einführung in die Test- und Fragebogenkonstruktion* (3. Aufl.). München: Pearson.

Bühner, M., Kröner, S. & Ziegler, M. (2008). Working memory, visual–spatial-intelligence and their relationship to problem-solving. *Intelligence, 36,* 672–680.

Cai, J. (1998). An investigation of U.S. an chinese students' mathematical problem posing and problem solving. *Mathematics Education Research Journal, 10,* 37–50.

Cardelle-Elawar, M. (1992). Effects of teaching metacognitive skills to students with low mathematical ability. *Teaching and Teacher Education, 8* (2), 109–121.

Carlson, J., Jaffe, A. & Wiles, A. (Hrsg.). (2006). *The millennium prize problems*. Providence, R.I, Cambridge, MA: American Mathematical Society, Clay Mathematics Institute.

Carlson, M. P. & Bloom, I. (2005). The cyclic nature of problem solving. An emergent multidimensional problem-solving framework. *Educational Studies in Mathematics, 58,* 45–75.

Carson, J. (2007). A problem with problem solving. Teaching thinking without teaching knowledge. *Mathematics Educator, 17* (2), 7–14.

Chang, K.-E., Wu, L.-J., Weng, S.-E. & Sung, Y.-T. (2012). Embedding gabe-based problem-solving phase into problem-posing system for mathematics learning. *Computers & Education, 58,* 775–786.

Chen, Z. (2002). Analogical problem solving. A hierachical analysis of procedural similarity. *Journal of Experimental Psychology, 28,* 81–98.

Chi, M. T. H., Feltovich, P. J. & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5,* 121–152.

Chi, M. T. H., Glaser, T. & Rees, E. (1982). Expertise in problem solving. In R. J. Sternberg (Hrsg.), *Advances in the psychology of human intelligence* (S. 7–75). London: Lawrence Erlbaum.

Christensen, G. & König, H. W. (2011). *Problemlösen im Mathematikunterricht. Sachbezogene Aufgaben aus dem Alltag* (Bergedorfer Unterrichtsideen). Buxtehude: Persen.

Chu, Y., Li, Z., Su, Y. & Pizlo, Z. (2010). Heuristics in problem solving. The role of direction in controlling search space. *Journal of Problem Solving, 3,* 27–51.

Clarke, T., Ayres, P. & Sweller, J. (2005). The impact of sequencing and prior knowledge on learning mathematics through spreadsheet applications. *Educational Technology Research and Development, 53,* 15–24.

Clement, J. & Konold, C. (1989). Fostering basic problem-solving skills in mathematics. *For the learning of mathematics, 9* (3), 26–30.

Cobb, P. & Merkel, G. (1989). Thinking strategies. Teaching arithmetic through problem solving. In P. R. Trafton (Hrsg.), *New directions for elementary school mathematics* (2. Aufl., S. 70–81). Reston, VA: National Council of Teachers of Mathematics, U.S.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. Aufl.). Hillsdale, NJ: Lawrence Erlbaum.

Collet, C. (2009). *Förderung von Problemlösekompetenzen in Verbindung mit Selbstregulation. Wirkungsanalysen von Lehrerfortbildungen* (Empirische Studien zur Didaktik der Mathematik, Bd. 2). Münster: Waxmann.

Collet, C. & Bruder, R. (2008). Lernen durch Selbstbeobachtung im problemlösenden Mathematikunterricht. *Praxis Schule 5-10* (3), 34–39.

Cooper, G. & Sweller, J. (1987). Effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology, 79,* 347–362.

CRP Henri Tudor & University of Luxembourg. (2011) TAO - Open and versatile computer-based assessment platform [Computer software]. Verfügbar unter http://www.tao.lu/

Dambeck, H. (2013, 3. Dezember). Pisa-Spitzenreiter: Das Geheimnis von Asiens Mathe-Genies. *Spiegel Online.* Zugriff am 27.12.2013. Verfügbar unter http://www.spiegel.de/schulspiegel/wissen/pisa-studie-2013-mathematik-erfolgsgeheimnis-asiatischer-schueler-a-935718.html

Danner, D., Hagemann, D., Holt, D. V., Hager, M., Schankin, A., Wüstenberg, S. et al. (2011). Measuring performance in dynamic decision making. Reliability and validity of the tailorshop simulation. *Journal of Individual Differences, 32,* 225–233.

Davidson, J. E., Deuser, R. & Sternberg, R. J. (1994). The role of metacognition in problem solving. In J. Metcalfe & A. P. Shimamura (Hrsg.), *Metacognition: Knowing about knowing* (S. 207–226). Cambridge, MA: MIT Press.

Davidson, J. E. & Sternberg, R. J. (1998). Smart problem solving: How metacognition helps. In D. J. Hacker, J. Dunlosky & A. C. Graesser (Hrsg.), *Metacognition in educational theory and practice* (S. 47–68). Mahwah, NJ: Lawrence Erlbaum.

Davier, A. A. von (Hrsg.). (2011). *Statistical models for test equating, scaling, and linking. With a foreword by Paul W. Holland* (Statistics for Social and Behavioral Sciences). New York, NY: Springer.

De Ayala, R. J. (2009). *The theory and practice of item response theory*. New York, NY: Guilford Press.

De Groot, A. D. (1978). *Thought and choice in chess* (2. Aufl.). The Hague: Mouton Publishers.

De Jong, T. (2005). Problem solving methodologies. In Kempf-Leonard (Hrsg.), *Encyclopedia of social measurement* (Bd. 3, S. 171–177). San Diego, CA: Academic Press.

DeMars, C. E. (2012). Confirming testlet effects. *Applied Psychological Measurement, 36,* 104–121.

Denvir, B., Brown, M. & Stolz, C. (1982). *Low attainers in mathematics 5-16. Policies and practices in schools*. London: Methuen Educational (Schools Council Working Paper 72).

Depaepe, F., De Corte, E. & Verschaffel, L. (2010). Teachers' metacognitive and heuristic approaches to word problem solving. Analysis and impact on students' beliefs and performance. *ZDM Mathematics Education, 42,* 205–218.

OECD PISA Deutschland. (2003). *Erfassung fächerübergreifender Problemlösekompetenzen in PISA* (Deutsches PISA Konsortium, Hrsg.). Zugriff am 23.10.2013. Verfügbar unter http://www.mpib-berlin.mpg.de/Pisa/Problemloesen.pdf

Deutsches PISA-Konsortium. (2006). *PISA 2003. Dokumentation der Erhebungsinstrumente*. Münster: Waxmann.

Dewey, J. (1933). *How do we think. A restatement of the relation of reflective thinking to the educative process*. Boston, MA: Heath.

Doorman, M., Drijvers, P., Dekker, T., van den Heuvel-Panhuizen, M. I., Lange, J. & Wijers, M. (2007). Problem solving as a challenge for mathematics education in The Netherlands. *ZDM Mathematics Education, 39,* 405–418.

Dörner, D. (1976). *Problemlösen als Informationsverarbeitung*. Stuttgart: Kohlhammer.

Dörner, D. (1984). Denken, Problemlösen und Intelligenz. *Psychologische Rundschau, 35,* 10–20.

Dörner, D., Kreuzig, H. W., Reither, F. & Stäudel, T. (1983). *Lohhausen. Vom Umgang mit Unbestimmtheit und Komplexität*. Bern: Huber.

Dreher, M. & Oerter, R. (1987). Action planning competencies during adolescence and early adulthood. In S. L. Friedman, E. K. Scholnick & R. R. Cocking (Hrsg.), *Blueprints for thinking. The role of planning in cognitive development* (S. 321–355). Cambridge: Cambridge University Press.

Duchhardt, C. & Gerdes, A. (2013, April). *NEPS technical report for mathematics. Scaling results of starting cohort 4 in ninth grade* (NEPS Working Papers Nr. 22). Bamberg: National Educational Panel Study (NEPS). Zugriff am 04.12.2013. Verfügbar unter https://www.neps-data.de/Portals/0/Working%20Papers/WP_XXII.pdf

Dunbar, K. (1998). Problem solving. In W. Bechtel & G. Graham (Hrsg.), *A companion to cognitive science* (S. 289–298). London: Blackwell.

Edelmann, W. & Wittmann, S. (2012). *Lernpsychologie* (7. Aufl.). Weinheim: Beltz, PVU.

Ehmann, T. (2008). *Erfassung und Förderung metakognitiver und motivationaler Fähigkeiten. Ein halbstandardisiertes Lerntagebuch für Schülerinnen und Schüler mit Migrationshintergrund*. Dissertation, Universität Potsdam. Potsdam. Zugriff am 13.06.2012. Verfügbar unter http://opus.kobv.de/ubp/volltexte/2011/5122/

Eiter, T. & Lukasiewicz, T. (2000). Default reasoning from conditional knowledge bases. Complexity and tractable cases. *Artificial Intelligence, 124,* 169–241.

Embretson, S. E. & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah: Lawrence Erlbaum.

Enzensberger, H. M. (n.d.). *Die Mathematik im Jenseits der Kultur.* Zugriff am 24.08.2013. Verfügbar unter http://www.mathe.tu-freiberg.de/~hebisch/cafe/zugbruecke.html

FastTEST Web (c/o ASC). (2011) FastTEST Web [Computer software]. Verfügbar unter www.fasttestweb.com

Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. (2007). G*Power 3. A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175–191.

Faul, F., Erdfelder, E., Buchner, A. & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1. Tests for correlation and regression analyses. *Behavior Research Methods, 41,* 1149–1160.

Field, A., Miles, J. & Field, Z. (2012). *Discovering statistics using R*. Los Angeles, CA: SAGE.

Field, A. P. (2009). *Discovering statistics using SPSS* (3. Aufl.). Los Angeles, CA: Sage Publications.

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Hrsg.), *The nature of intelligence* (S. 231–235). Hillsdale, NJ: Lawrence Erlbaum.

Fleischer, J., Buchwald, F., Wirth, J., Rumann, S. & Leutner, D. (in preparation). Analytical problem solving. Potentials and manifestations. In B. Csapó, J. Funke & A. Schleicher (Hrsg.), *The nature of problem solving.* Paris: OECD.

Fleischer, J., Wirth, J. & Leutner, D. (2009, August). *Problem solving as both cross-curricular and subject-specific competence.* Vortrag gehalten auf der 13th Biennal Conference of the European Association for Research on Learning and Instruction (Earli), Amsterdam, Niederlande.

Fleischer, J., Wirth, J. & Leutner, D. (2014). Effekte der kontextuellen Einkleidung von Testaufgaben auf die Leistungen von Schülerinnen und Schülern im Problemlösen und in der Mathematik. Manuskript in Vorbereitung.

Fleischer, J., Wirth, J., Rumann, S. & Leutner, D. (2010). Strukturen fächerübergreifender und fachlicher Problemlösekompetenz. Analyse von Aufgabenprofilen. Projekt Problemlösen. *Zeitschrift für Pädagogik* (Beiheft 56), 239–248.

Fox, J.-P. & Verhagen, A. J. (2010). Random item effects modeling for cross-national survey data. In E. Davidov, P. Schmidt & J. Billiet (Hrsg.), *Cross-cultural analysis. Methods and applications* (S. 467–488). London: Routledge.

Fox, J.-P. (2010). *Bayesian item response modeling. Theory and applications*. New York, NY: Springer.

Frensch, P. A. & Funke, J. (1995). *Complex problem solving. The european perspective*. Hillsdale, NJ: Lawrence Erlbaum.

Freudenthal, H. (1968). Why to teach mathematics so as to be useful. *Educational Studies in Mathematics, 1,* 3–8.

Freudenthal, H. (1969). ICMI report on mathematical contests in secondary education (Olympiads) I. *Educational Studies in Mathematics, 2,* 80–114.

Freudenthal, H. (1981). Major problems of mathematics education. *Educational Studies in Mathematics, 12,* 133–150.

Freudenthal, H. (1991). *Revisiting mathematics education. China lectures* (Mathematics education library, Bd. 9). Dordrecht: Kluwer Academic Publishers.

Freund, P. A., Kuhn, J.-T. & Holling, H. (2011). Measuring current achievement motivation with the QCM. Short form development and investigation of invariance. *Personality and Individual Differences, 51,* 629–634.

Frey, A. (2012). Adaptives Testen. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (2. Aufl., S. 275–293). Berlin: Springer.

Frey, A. & Bernhardt, R. (2012). On the importance of using balanced booklet designs in PISA. *Psychological Test and Assessment Modeling, 54,* 397–417.

Frey, A., Heinze, A., Mildner, D., Hochweber, J. & Asseburg, R. (2010). Mathematische Kompetenz von PISA 2003 bis PISA 2009. In E. Klieme (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt.* Münster: Waxmann.

Frey, A. & Seitz, N.-N. (2009). Multidimensional adaptive testing in educational and psychological measurement: Current state and future challenges. *Studies in Educational Evaluation* (35), 89–94.

Fritz, A. & Funke, J. (2003). Planungsfähigkeit bei lernbehinderten Kindern. Grundsätzliche Überlegungen zum Konstrukt sowie zu dessen Diagnostik und Training. In G. Ricken, A. Fritz & C. Hofmann (Hrsg.), *Diagnose: Sonderpädagogischer Förderbedarf* (S. 416–439). Lengerich: Pabst Science.

Fritz, A. & Hussy, W. (2001). Training der Planungsfähigkeit bei Grundschulkindern. Eine Evaluationsstudie. In K. J. Klauer (Hrsg.), *Handbuch kognitives Training* (2. Aufl., S. 97–127). Göttingen: Hogrefe.

Fuchs, L. S., Fuchs, D., Finelli, R., Courey, S. J., Hamlett, C. L., Sones, E. M. et al. (2006). Teaching third graders about real-life mathematical problem-solving: A randomized controlled study. *The Elementary School Journal, 106,* 293–311.

Funke, J. (2003). *Problemlösendes Denken*. Stuttgart: Kohlhammer.

Funke, J. (2006a). Komplexes Problemlösen. In J. Funke (Hrsg.), *Denken und Problemlösen* (S. 375–445). Göttingen: Hogrefe.

Funke, J. (2006b). Lösen komplexer Probleme. In J. Funke & P. A. Frensch (Hrsg.), *Handbuch der Allgemeinen Psychologie. Kognition* (S. 439–445). Göttingen: Hogrefe.

Funke, J. & Fritz, A. (Hrsg.). (1995a). *Neue Konzepte und Instrumente zur Planungsdiagnostik*. Bonn: Dt. Psychologen-Verlag.

Funke, J. & Fritz, A. (1995b). Über Planen, Problemlösen und Handeln. In J. Funke & A. Fritz (Hrsg.), *Neue Konzepte und Instrumente zur Planungsdiagnostik* (S. 1–45). Bonn: Dt. Psychologen-Verlag.

Funke, J. & Glodowski, A.-S. (1990). Planen und Problemlösen. Überlegungen zur neuropsychologischen Diagnostik von Basiskompetenzen beim Planen. *Zeitschrift für Neuropsychologie, 1,* 139–148.

Funke, J. & Krüger, T. (1995). "Plan-A-Day": Konzeption eines modifizierbaren Instruments zur Führungskräfte-Auswahl sowie erste empirische Befunde. In J.

Funke & A. Fritz (Hrsg.), *Neue Konzepte und Instrumente zur Planungsdiagnostik* (S. 97–120). Bonn: Dt. Psychologen-Verlag.

Funke, J., Krüger, T. & Schulze, F. (1999, 30. September). *"Plan A Day" mit RehaCom. Version 1.1.* Zugriff am 28.08.2012. Verfügbar unter http://www.psychologie.uni-heidelberg.de/ae/allg/forschun/planaday/Rehacom_Manual_PLAN.pdf

Fyfe, E. R., Rittle-Johnson, B. & DeCaro, M. S. (2012). The effects of feedback during exploratory mathematics problem solving. Prior knowledge matters. *Journal of Educational Psychology, 104,* 1094–1108.

Garner, R. (1990). When children and adults do not use learning strategies: Toward a theory of settings. *Review of Educational Research, 60,* 517–529.

Gärtner, H.-J. & Scharf, V. (2001). *Chemische "Egg-Races" in Theorie und Praxis. 17 Vorschläge zur Gruppenarbeit von Mädchen und Jungen im Chemieunterricht der Sekundarstufe I.* Speyer: Staatliches Institut für Lehrerfort- und -weiterbildung des Landes Rheinland-Pfalz. Zugriff am 26.11.2013. Verfügbar unter http://www.chemie-biologie.uni-siegen.de/chemiedidaktik/dokumente/service/fundgrube/chemrace.pdf

Gick, M. L. & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12,* 306–355.

Gick, M. L. & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15,* 1–38.

Gick, M. L. (1986). Problem-solving strategies. *Educational Psychologist, 21,* 99–120.

Gilsdorf, R. & Kistner, G. (1998). *Kooperative Abenteuerspiele. Eine Praxishilfe für Schule und Jugendarbeit* (5. Aufl.). Seelze-Velber: Kallmeyer.

Globalpark AG. (2011) EFS Survey (Enterprise Feedback Suite) [Computer software]: Globalpark AG. Verfügbar unter www.unipark.info

Gobet, F. (2005). Chunking models of expertise. Implications for education. *Applied Cognitive Psychology, 19,* 183–204.

Goertzen, J. (2014). *Schülersicht des Trainings "Planen, Problemlösen, Präsentieren".* Unveröffentlichte schriftliche Hausarbeit im Rahmen der Ersten

Staatsprüfung für das Lehramt an Gymnasien und Gesamtschulen, Universität Duisburg-Essen. Essen.

Gonzales, E. & Rutkowski, L. (2010) Principles of multiple matrix booklet designs and parameter recovery in large-scale assessments. In M. von Davier & D. Hastedt (Hrsg.), *IERI monograph series. Issues and methodologies in large-scale assessments: Volume 3* (S. 125–156). Verfügbar unter http://www.ierinstitute.org/fileadmin/Documents/IERI_Monograph/IERI_Monograph_Volume_03_Chapter_6.pdf

Graham, J. W. (2012). *Missing data. Analysis and design* (Statistics for Social and Behavioral Sciences). New York, NY: Springer.

Graham, J. W., Olchowski, A. E. & Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science, 8,* 206–213.

Greefrath, G. (2010). Problemlösen und Modellieren. Zwei Seiten der gleichen Medaille. *Mathematikunterricht (MU), 56* (3), 44–56.

Greeno, J. G., Collins, A. M. & Resnick, L. B. (1996). Cognition and learning. In Berliner, David C., Calfee, Robert C. (Hrsg.), *Handbook of educational psychology* (S. 15–46). New York, NY: Simon & Schuster Macmillan.

Greiff, S. & Fischer, A. (2013). Der Nutzen einer komplexen Problemlösekompetenz. Theoretische Überlegungen und empirische Befunde. *Zeitschrift für Pädagogische Psychologie, 27,* 27–39.

Greiff, S., Holt, D. V. & Funke, J. (2013). Perspectives on problem solving in educational assessment. Analytical, interactive, and collaborative problem solving. *Journal of Problem Solving, 5,* 71–91.

Grömping, U. (2006). Relative Importance for Linear Regression in R: The Package relaimpo. *Journal of Statistical Software, 17* (1), 1–27. Zugriff am 03.12.2014. Verfügbar unter http://www.jstatsoft.org/v17/i01/paper

Grömping, U. (2007). Estimators of relative importance in linear regression based on variance decomposition. *The American Statistician, 61* (2), 139–147.

Gürtler, T., Perels, F., Schmitz, B. & Bruder, R. (2002). Training zur Förderung selbstregulativer Fähigkeiten in Kombination mit Problemlösen in Mathematik. *Zeitschrift für Pädagogik* (Beiheft 45), 222–239.

Gürtler, T., Perels, F., Schmitz, B. & Bruder, R. (2004). Training zur Förderung selbstregulativer Fähigkeiten in Kombination mit Problemlösen in Mathematik. In J. Doll & M. Prenzel (Hrsg.), *Bildungsqualität von Schule. Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsverbesserung* (S. 222–239). Münster: Waxmann.

Gut-Glanzmann, C. (2012). *Modellierung und Messung experimenteller Kompetenz. Analyse eines large-scale Experimentiertests*. Inauguraldissertation, Universität Basel. Basel. Zugriff am 06.12.2014. Verfügbar unter http://www.phzh.ch/dotnetscripts/MAPortrait_Data/158541/1/Gut_2012_Experimentelle_Kompetenz.pdf

Gut-Glanzmann, C. (2014, Dezember). *Experimentelle Kompetenzen auf der Sekundarstufe I. Problemtypenbasierte Modellierung und Messung mittels interdisziplinärer Experimentiertests*. Universität Duisburg-Essen: ZeB-Forschungskolloquium, Essen.

Halpern, D. F. (2003). *Thought & knowledge. An introduction to critical thinking* (4. Aufl.). Mahwah, NJ: Lawrence Erlbaum.

Hambrick, D. Z. (2005). The role of domain knowledge in higher-order cognition. In O. Wilhelm & R. W. Engle (Hrsg.), *Handbook of understanding and measuring intelligence* (S. 361–372). Thousand Oaks, CA: Sage Publications.

Hammann, M., Hoi Phan, T. & Bayrhuber, H. (2007). Experimentieren als Problemlösen. Lässt sich das SDDS-Modell nutzen, um unterschiedliche Dimensionen beim Experimentieren zu messen? *Zeitschrift für Erziehungswissenschaft, Sonderheft 8,* 33–49.

Hasselhorn, M. (1992). Metakognition und Lernen. In G. Nold (Hrsg.), *Lernbedingungen und Lernstrategien. Welche Rolle spielen kognitive Verstehensstrukturen?* (Tübinger Beiträge zur Linguistik, Bd. 366, S. 35–63). Tübingen: Narr.

Hasselhorn, M. (2010). Metakognition. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (4. Aufl., S. 541–547). Weinheim: Beltz, PVU.

Hasselhorn, M. & Gold, A. (2006). *Pädagogische Psychologie* (3. Aufl.). Stuttgart: Kohlhammer.

Hasson, F., Keeney, S. & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing, 32,* 1008–1015.

Heller, K. A. & Perleth, C. (2000). *Kognitiver Fähigkeitstest (KFT 4-12+R)* (3. Aufl.). Göttingen: Beltz, PVU.

Helmke, A. & Aken, M. A. G. van. (1995). The causal ordering of academic achievement and self-concept of ability during elementary school. A longitudinal study. *Journal of Educational Psychology, 87,* 624–637.

Helms, L. L. (2009). *Potential theory*. Dordrecht: Springer.

Hertwig, R. (2006). Strategien und Heuristiken. In J. Funke & P. A. Frensch (Hrsg.), *Handbuch der Allgemeinen Psychologie. Kognition* (S. 461–469). Göttingen: Hogrefe.

Hesse, C. (2014, 06. März). *Math up your life! Goethe im Duell mit Newton,* ZEIT ONLINE. Zugriff am 06.03.2014. Verfügbar unter http://blog.zeit.de/mathe/allgemein/mathe-statt-goethe/#more-150

Higgins, K. M. (1997). The Effect of year-long instruction in mathematical problem solving on middle-school students' attitudes, beliefs, and abilities. *Journal of Experimental Education, 66,* 5–28.

Ho, K. F. & Hedberg, J. G. (2005). Teachers’ pedagogies and their impact on students’ mathematical problem solving. *Journal of Mathematical Behavior, 24,* 238–252.

Hoffman, B. & Schraw, G. (2009). The influence of self-efficacy and working memory capacity on problem-solving efficiency. *Learning and Individual Differences, 19,* 91–100.

Hoffmann, A. (2012). *Planungskompetenztraining mit Bachelorstudierenden. Studentische Planungsleistung und Vorschlag einer Transfersitzung*. Unveröffentlichte Bachelorarbeit, RWTH Aachen. Aachen.

Hohenwarter, M. (2006, 12. Januar). *GeoGebra – didaktische Materialien und Anwendungen für den Mathematikunterricht*. Dissertation, Paris-Lodron-

Universität Salzburg. Salzburg. Zugriff am 13.01.2014. Verfügbar unter http://www.geogebra.org/publications/mhohen_diss.pdf

Holling, H., Preckel, F., Vock, M. & Schulze Willbrennig, B. (2004). *Schulische Begabtenförderung in den Ländern. Maßnahmen und Tendenzen* (Materialien zur Bildungsplanung und zur Forschungsförderung Nr. 121). Bonn: Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung (BLK). Zugriff am 10.12.2012. Verfügbar unter http://www.pedocs.de/volltexte/2008/328/pdf/heft121.pdf

Huff, K. & Goodmann, D. P. (2007). The demand for cognitive diagnostic assessment. In J. P. Leighton & M. J. Gierl (Hrsg.), *Cognitive diagnostic assessment for education. Theory and applications* (S. 19–60). Cambridge: University Press.

Hummel, J. E. & Holyoak, K. J. (1997). Distributes representations of structure. A theory of analogical access and mapping. *Psychological Review, 104,* 427–466.

Humphreys, L. G. (1976). A factor model for research on intelligence and problem solving. In L. B. Resnick (Hrsg.), *The nature of intelligence* (S. 329–339). Hillsdale, NJ: Lawrence Erlbaum.

Institut für Wissensmedien (IWM) (Hesse, F. W., Hrsg.). (2011). *e-teaching.org,* IWM - Institut für Wissensmedien. Zugriff am 30.06.2011. Verfügbar unter http://www.e-teaching.org

Jansen, M., Schroeders, U. & Stanat, P. (2013). Motivationale Schülermerkmale in Mathematik und den Naturwissenschaften. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle & C. Pöhlmann (Hrsg.), *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (S. 348–365). Münster: Waxmann.

Jarque, C. M. & Bera, A. K. (1987). A Test for normality of observations and regression residuals. *International Statistical Review, 55,* 163–172.

Johnson, H. J., Barnard-Brak, L., Saxon, T. F. & Johnson, M. K. (2012). An experimental study of the effects of stereotype threat and stereotype lift on men and women's performance in mathematics. *Journal of Experimental Education, 80,* 137–149.

Jonassen, D. H. (2012). Problem typology. In N. M. Seel (Hrsg.), *Encyclopedia of the sciences of learning* (S. 2683–2686). New York, NY: Springer.

Kajamies, A., Vauras, M. & Kinnunen, R. (2010). Instructing low-achievers in mathematical word problem solving. *Scandinavian Journal of Educational Research, 54,* 335–355.

Kalyuga, S. (2005). Prior knowledge principle in multimedia learning. In R. E. Mayer (Hrsg.), *The Cambridge handbook of multimedia learning* (S. 325–338). New York, NY: Cambridge University Press.

Kalyuga, S. (2009). Instructional designs for the development of transferable knowledge and skills. A cognitive load perspective. *Computers in Human Behavior, 25,* 332–338.

Katzenbach, M. (2004). Dem Fehler auf der Spur. Kinder als Fehlerdetektive. *Die neue Schulpraxis, 12,* 4–8. Zugriff am 28.03.2014. Verfügbar unter http://sinus-transfer.uni-bayreuth.de/fileadmin/MaterialienBT/Fehler_MK.pdf

Keeler, M. L. & Swanson, H. L. (2001). Does strategy knowledge influence working memory in children with mathematical disabilities? *Journal of Learning Disabilities, 34,* 418–434.

Keller, G. (2005). *Lern-Methodik-Training. Ein Übungsmanual für die Klassen 5-10* (2. Aufl.). Göttingen: Hogrefe.

Kiefer, T., Robitzsch, A. & Wu, M. L. (2014). *TAM: Test Analysis Modules* (R package version 1.0-2.1). Zugriff am 29.03.2014. Verfügbar unter http://CRAN.R-project.org/package=TAM

Kießwetter, K. (1983). Modellierung von Problemlöseprozessen. *Mathematikunterricht (MU), 29* (3), 71–101.

King, A. (1991). Effects of training in strategic questioning on children's problem-solving performance. *Journal of Educational Psychology, 83,* 307–317.

Klauer, K. J. (1987). *Kriteriumsorientierte Tests. Lehrbuch der Theorie und Praxis lehrzielorientierten Messens*. Göttingen: Hogrefe.

Klauer, K. J. (1989). Die Messung von Transferdistanzen. Ein Verfahren zur Bestimmung der Unähnlichkeit von Aufgabenanforderungen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 21,* 146–166.

Klauer, K. J. (2001). Trainingsforschung. Ansätze - Theorien - Ergebnisse. In K. J. Klauer (Hrsg.), *Handbuch kognitives Training* (2. Aufl., S. 3–66). Göttingen: Hogrefe.

Klauer, K. J. (2010). Schädliche Interferenzen beim Training hochkomplexer Lernstrategien? *Zeitschrift für Pädagogische Psychologie, 24,* 235–239.

Klauer, K. J. (2011). *Transfer des Lernens. Warum wir oft mehr lernen als gelehrt wird*. Stuttgart: Kohlhammer.

Klieme, E. (2004). Assessment of cross-curricular problem-solving competencies. In J. H. Moskowitz & M. Stephens (Hrsg.), *Comparing learning outcomes. International assessments and education policy* (S. 81–107). London: Routledge.

Klieme, E., Artelt, C. & Stanat, P. (2001). Fächerübergreifende Kompetenzen. Konzepte und Indikatoren. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (2. Aufl., S. 203–218). Weinheim: Beltz, PVU.

Klieme, E., Funke, J., Leutner, D., Reimann, P. & Wirth, J. (2001). Problemlösen als fächerübergreifende Kompetenz. Konzeption und erste Resultate aus einer Schulleistungsstudie. *Zeitschrift für Pädagogik, 47,* 179–200.

Klieme, E. & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Beschreibung eines neu eingrichteten Schwerpunktprogramms der DFG. *Zeitschrift für Pädagogik, 52,* 876–903.

Klippert, H. (2002). *Methoden-Training* (Übungsbausteine für den Unterricht, Bd. 1, 13. Aufl.). Weinheim: Beltz, PVU.

Klippert, H. (2008). *Kommunikations-Training. Bausteine für den Unterricht* (Neu ausgestattete Sonderausgabe). Weinheim: Beltz, PVU.

KMK. (2003). *Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss. Beschluss vom 4.12.2003*. Beschlüsse der Kultusministerkonferenz. Zugriff am 14.02.2012. Verfügbar unter http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2003/2003_12_04-Bildungsstandards-Mathe-Mittleren-SA.pdf

KMK. (2004a). *Bildungsstandards im Fach Mathematik für den Hauptschulabschluss. Beschluss vom 15.10.2004.* Beschlüsse der Kultusministerkonferenz.

Zugriff am 14.02.2012. Verfügbar unter http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_10_15-Bildungsstandards-Mathe-Haupt.pdf

KMK. (2004b). *Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss. Beschluss vom 4.12.2003*. Neuwied: Luchterhand.

KMK. (2005a). *Beschlüsse der Kultusministerkonferenz – Bildungsstandards im Fach Biologie für den mittleren Schulabschluss (Beschluss von 16. Dezember 2004)*. München: Wolters Kluwer Deutschland.

KMK. (2005b). *Beschlüsse der Kultusministerkonferenz – Bildungsstandards im Fach Chemie für den mittleren Schulabschluss (Beschluss von 16. Dezember 2004)*. München: Wolters Kluwer Deutschland.

KMK. (2005c). *Beschlüsse der Kultusministerkonferenz – Bildungsstandards im Fach Physik für den mittleren Schulabschluss (Beschluss von 16. Dezember 2004)*. München: Wolters Kluwer Deutschland.

KMK. (2005d). *Bildungsstandards im Fach Mathematik für den Primarbereich (Jahrgangsstufe 4). Beschluss vom 15.10.2004*. Neuwied: Luchterhand.

KMK. (2012). *Bildungsstandards im Fach Mathematik für die Allgemeine Hochschulreife. Beschluss der Kultusministerkonferenz vom 18.10.2012*. Zugriff am 16.02.2014. Verfügbar unter http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2012/2012_10_18-Bildungsstandards-Mathe-Abi.pdf

Knapp, W., Pfaff, H. & Werner, S. (2010). Verstehen durch Schreiben. Anlage einer empirischen Studie zum produktiven Umgang mit mathematischen Textaufgaben. In B. Ahrenholz (Hrsg.), *Fachunterricht und Deutsch als Zweitsprache* (2. Aufl., S. 239–256). Tübingen: Narr.

Knoblich, G. & Öllinger, M. (2006). Einsicht und Umstrukturierung beim Problemlösen. In J. Funke (Hrsg.), *Denken und Problemlösen* (S. 1–85). Göttingen: Hogrefe.

Koeppen, K., Hartig, J., Klieme, E. & Leutner, D. (2008). Current issues in competence modeling and assessment. *Zeitschrift für Psychologie / Journal of Psychology, 216,* 61–73.

Köhler, W. (1963). *Intelligenzprüfungen am Menschenaffen* (2. Aufl.). Berlin: Springer.

Kolen, M. J. & Brennan, R. L. (2014). *Test equating. Methods and practices* (3. Aufl.). New York, NY: Springer.

Köller, O. (2006, 14. November). *Stellungnahme zum Text von Joachim Wuttke: "Fehler, Verzerrungen, Unsicherheiten in der PISA-Auswertung".* Zugriff am 29.05.2012. Verfügbar unter http://www.phil-fak.uni-duesseldorf.de/fileadmin/Redaktion/Institute/Sozialwissenschaften/BF/Lehre/WiSe11_12/Pisa/PISA_Koeller.pdf

Köller, O. & Saß, S. (2013). Wie viel g ist in Mathematiktests? Konfirmatoriscbe Faktorenanalysen zur Konstruktvalidität von Mathematiktests am Ende der Primarstufe. In N. McElvany & H. G. Holtappels (Hrsg.), *Empirische Bildungsforschung. Theorien, Methoden, Befunde und Perspektiven (Festschrift für Wilfried Bos)* (S. 173–187). Münster: Waxmann.

Komorek, E., Bruder, R., Collet, C. & Schmitz, B. (2006). Inhalte und Ergebnisse einer Intervention um Mathematikunterricht der Sekundarstufe I mit einem Unterrichtskonzept zur Förderung mathematischen Problemlösens und von Selbstregulationskonzeptem. In M. Prenzel & L. Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms* (BiQua, S. 240–267). Münster: Waxmann.

Kraha, A., Turner, H., Nimon, K., Zientek, L. R. & Henson, R. K. (2012). Tools to support interpreting multiple regression in the face of multicollinearity. *Frontiers in psychology, 3* (March), 1–16.

Krulik, S. & Rudnick, J. A. (1987). *Problem solving. A handbook for teachers* (2. Aufl.). Boston: Allyn & Bacon.

Krulik, S. & Rudnick, J. A. (1996). *The new sourcebook for teaching reasoning and problem solving in Junior and Senior High School*. Boston, MA: Allyn & Bacon.

Krulik, S. & Rudnick, J. A. (1998). *Assessing reasoning and problem solving. A sourcebook for elementary school teachers*. Boston: Allyn & Bacon.

Kubinger, K. D., Rasch, D. & Yanagida, T. (2011). *Statistik in der Psychologie. Vom Einführungskurs bis zur Dissertation*. Göttingen: Hogrefe.

Kunina-Habenicht, O., Wilhelm, O., Matthes, F. & Rupp, A. A. (2010). Kognitive Diagnosemodelle. Theoretisches Potential und methodische Probleme. Projekt Kognitive Diagnosemodelle. *Zeitschrift für Pädagogik* (Beiheft 56), 75–85.

Kuntze, S. & Prediger, S. (2005). Ich schreibe, also denk' ich - Über Mathematik schreiben. *Praxis der Mathematik in der Schule, 47* (5), 1–6.

Laker, D. R. (1990). Dual dimensionality of training transfer. *Human Resource Development Quarterly, 1,* 209–223.

Lehmann, D. & Magidor, M. (1992). What does a conditional knowledge base entail? *Journal of Artificial Intelligence, 55,* 1–60. Zugriff am 29.03.2014. Verfügbar unter http://arxiv.org/abs/cs/0202022v1

Lesch, H. (2011, 22. Juli). *Wir irren uns empor … oder warum die Physik so erfolgreich ist?* Bayreuth: Physikalisches Kolloquium der Studierenden, Universität Bayreuth. Zugriff am 22.01.2014. Verfügbar unter http://www.youtube.com/watch?v=u29--YNGMyg

Lester, F. K. (1994). Musings about mathematical problem-solving research: 1970-1994. *Journal for research in mathematics education, 25,* 660–675.

Leuders, T. (2010). Problemlösen. In T. Leuders (Hrsg.), *Mathematik-Didaktik. Praxishandbuch für die Sekundarstufe I und II* (5. Aufl., S. 119–135). Berlin: Cornelsen Scriptor.

Leuders, T. & Sodian, B. (2013). Inwiefern sind Kompetenzmodelle dazu geeignet kognitive Prozesse von Lernenden zu beschreiben? *Zeitschrift für Erziehungswissenschaft, 16* (Sonderheft 18), 27–33.

Leutner, D. (1992a). *Adaptive Lehrsysteme. Instruktionspsychologische Grundlagen und experimentelle Analysen* (Fortschritte der psychologischen Forschung, Bd. 13). Weinheim: Psychologie-Verl.-Union.

Leutner, D. (1992b). Das Testlängendilemma in der lernprozess-begleitenden Wissensdiagnostik. *Zeitschrift für Pädagogische Psychologie, 6,* 233–238.

Leutner, D. (1993). Das gleitende Testfenster als Lösung des Testlängendilemmas: Eine Robustheitsstudie. *Zeitschrift für Pädagogische Psychologie, 7,* 33–45.

Leutner, D. (2004). Instructional-design principles for adaptivity in open learning environments. In N. M. Seel & S. Dijkstra (Hrsg.), *Curriculum, plans, and processes in instructional design. International perspectives* (S. 289–307). Mahwah, NJ: Lawrence Erlbaum.

Leutner, D. (2013a, Oktober). *Projekt „Analytisches Problemlösen" – Phase 3.* Jahreskolloquium des DFG-Schwerpunktprogramms „Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen", Frankfurt am Main.

Leutner, D. (2013b, November). *Problemlösen als fachliche und fächerübergreifende Kompetenz.* Forschungstag PH FHNW, Basel.

Leutner, D., Fleischer, J., Spoden, C. & Wirth, J. (2007). Landesweite Lernstandserhebungen zwischen Bildungsmonitoring und Individualdiagnostik. *Zeitschrift für Erziehungswissenschaft, Sonderheft 8,* 149–167.

Leutner, D., Fleischer, J. & Wirth, J. (2006). Problemlösekompetenz als Prädiktor für zukünftige Kompetenz in Mathematik und in den Naturwissenschaften. In PISA-Konsortium Deutschland (Hrsg.), *PISA 2003. Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres* (S. 119–137). Münster: Waxmann.

Leutner, D., Fleischer, J., Wirth, J., Greiff, S. & Funke, J. (2012). Analytische und dynamische Problemlösekompetenz im Lichte internationaler Schulleistungsvergleichsstudien. *Psychologische Rundschau, 63,* 34–42.

Leutner, D., Klieme, E., Meyer, K. & Wirth, J. (2004). Problemlösen. In PISA-Konsortium Deutschland (Hrsg.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland - Ergebnisse des zweiten internationalen Vergleichs* (S. 147–175). Münster: Waxmann.

Leutner, D., Klieme, E., Meyer, K. & Wirth, J. (2005). Die Problemlösekompetenz in den Ländern der Bundesrepublik Deutschland. In M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand et al. (Hrsg.), *PISA*

2003. Der zweite Vergleich der Länder in Deutschland - Was wissen und können Jugendliche? (S. 125–146). Münster: Waxmann.

Leutner, D., Rumann, S. & Wirth, J. (2009). *Antrag auf Gewährung einer Sachbeihilfe im Rahmen des Schwerpunktprogramms 1293 "Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen". Fortsetzungsantrag für die zweite Förderphase*, Essen.

Leutner, D., Rumann, S. & Wirth, J. (2011, Januar). *Antrag auf Gewährung einer Sachbeihilfe im Rahmen des Schwerpunktprogramms 1293 „Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen". Fortsetzungsantrag für die dritte Förderphase*, Essen.

Levy, F. & Murnane, R. J. (2005). *The new division of labor. How computers are creating the next job market*. Princeton, NJ: Princeton University Press.

Lezak, M. D. (1995). *Neuropsychological assessment* (3. Aufl.). New York, NY: Oxford University Press.

Lind, G. (1996). *Konstanz method of dilemma-discussion.* Zugriff am 23.08.2012. Verfügbar unter http://www.uni-konstanz.de/ag-moral/moral/dildisk-e.htm

Lind, G. (2002). *Ist Moral lehrbar? Ergebnisse der modernen moralpsychologischen Forschung* (2. Aufl.). Berlin: Logos.

Lind, G. (2009). *Moral ist lehrbar. Handbuch zur Theorie und Praxis moralischer und demokratischer Bildung* (2. Aufl.). München: Oldenbourg.

Lindeman, R. H., Merenda, P. F. & Gold, R. Z. (1980). *Introduction to bivariate and multivariate analysis*. Scott, Foresman, Glenview, IL.

Lindquist, M. M. (1989). It's time to change. In P. R. Trafton (Hrsg.), *New directions for elementary school mathematics* (2. Aufl., S. 1–13). Reston, VA: National Council of Teachers of Mathematics, U.S.

Lingel, K., Götz, L., Artelt, C. & Schneider, W. (2014). *MAESTRA 5-6+. Mathematisches Strategiewissen für 5. und 6. Klassen*. Göttingen: Hogrefe.

Luchins, A. S. (1942). Mechanization in problem solving. The effect of Einstellung. *Psychological Monographs: General and Applied, 54,* 1–95.

Lüdtke, O., Robitzsch, A., Trautwein, U. & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung. *Psychologische Rundschau, 58,* 103–117.

Maaß, J. & Siller, H.-S. (Hrsg.). (2014). *Neue Materialien für einen realitätsbezogenen Mathematikunterricht 2* (ISTRON-Schriftenreihe). Wiesbaden: Springer Spektrum.

Mähler, C. & Stern, E. (2010). Transfer. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (4. Aufl., S. 859–869). Weinheim: Beltz, PVU.

Mandl, H., Prenzel, M. & Gräsel, C. (1992). Das Problem des Lerntransfers in der betrieblichen Weiterbildung. *Unterrichtswissenschaft, 20,* 126–143.

Marewski, J. N. & Schooler, L. J. (2011). Cognitive niches. An ecological model of strategy selection. *Psychological Review, 118,* 393–437.

Marschner, J. (2011). *Adaptives Feedback zur Unterstützung des selbstregulierten Lernens durch Experimentieren*. Dissertation, Universität Duisburg-Essen. Essen. Zugriff am 24.02.2014. Verfügbar unter http://duepublico.uni-duisburg-essen.de/servlets/DerivateServlet/Derivate-27679/Diss_Marschner.pdf

Marschner, J., Thillmann, H., Wirth, J. & Leutner, D. (2012). Wie lässt sich die Experimentierstrategie-Nutzung fördern? *Zeitschrift für Erziehungswissenschaft, 15,* 77–93.

Marsh, H. W. (1990a). Causal ordering of academic self-concept and academic achievement. A multiwave, longitudinal panel analysis. *Journal of Educational Psychology, 82,* 646–656.

Marsh, H. W. (1990b). The structure of academic self-concept. The Marsh/Shavelson Model. *Journal of Educational Psychology, 82,* 623–636.

Mattes, W. (2011). *Methoden für den Unterricht. Kompakte Übersichten für Lehrende und Lernende*. Paderborn: Schöningh.

Max-Planck-Institut für Bildungsforschung, Berlin. (n.d.). *Die Effekte von Anreizen auf Motivation und Leistungen in Mathematiktests.* Zugriff am 24.02.2014. Verfügbar unter https://www.mpib-berlin.mpg.de/Pisa/KurzberichtMotivation.pdf

Mayer, R. E. (1992). *Thinking, problem solving, cognition* (2. Aufl.). New York, NY: Freeman.

Mayer, R. E. (1998). Cognitive, metacognitive, and motivational aspects of problem solving. *Instructional Science, 26,* 49–63.

Mayer, R. E. (1999). Fifty years of creativity research. In R. J. Sternberg (Hrsg.), *Handbook of creativity* (S. 449–460). New York, NY: Cambridge University Press.

Mayer, R. E. (Hrsg.). (2005). *The Cambridge handbook of multimedia learning*. New York, NY: Cambridge University Press.

Mayer, R. E. (2008). *Learning and instruction* (2. Aufl.). Upper Saddle River, NJ: Merrill Prentice Hall.

Mayer, R. E. & Wittrock, M. C. (1996). Problem-solving transfer. In Berliner, David C., Calfee, Robert C. (Hrsg.), *Handbook of educational psychology* (S. 47–62). New York, NY: Simon & Schuster Macmillan.

McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology, 60,* 577–605.

McDaniel, M. A. & Nguyen, N. T. (2001). Situational judgment tests. A review of practice and constructs assessed. *International Journal of Selection and Assessment, 9,* 103–113.

McLeod, D. B. (1988). Affective issues in mathematical problem solving: Some theoretical considerations. *Journal for research in mathematics education, 19,* 134–141.

Meyerhöfer, W. (2006, 29. November). *Zur Stellungnahme von Prof. Dr. Olaf Köller vom 14.11.2006 zum Text von Joachim Wuttke: Fehler, Verzerrungen, Unsicherheiten in der PISA-Auswertung*. Zugriff am 29.05.2012. Verfügbar unter http://www.phil-fak.uni-duesseldorf.de/fileadmin/Redaktion/Institute/Sozialwissenschaften/BF/Lehre/Materialien/Pisa/Wolfram%20Meyerhoefer%20zu%20Olaf%20Koeller.pdf

Middleton, H. (2002). Complex problem solving in a workplace setting. *International Journal of Educational Research, 37,* 67-84.

Miller, G. A., Galanter, E. & Pribram, K. H. (1960). *Plans and the structure of behavior*. London: Holt, Rinehart and Winston.

Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen. (2005a). *Lernstand 9. Mathematik Aufgabenheft A für Schülerinnen und Schüler.* Zugriff am 25.10.2011. Verfügbar unter http://www.standardsicherung.schulministerium.nrw.de/lernstand8/upload/download/Testaufgaben/mathematik_schuelerheft_A_05.pdf

Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen. (2005b). *Lernstand 9. Mathematik Aufgabenheft B für Schülerinnen und Schüler.* Zugriff am 25.10.2011. Verfügbar unter http://www.standardsicherung.schulministerium.nrw.de/lernstand8/upload/download/Testaufgaben/mathematik_schuelerheft_B_05.pdf

Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen. (2005c). *Lernstand 9. Mathematik Auswertungsanleitung A1/A2 für Lehrerinnen und Lehrer.* Zugriff am 25.10.2011. Verfügbar unter http://www.standardsicherung.schulministerium.nrw.de/lernstand8/upload/download/Testaufgaben/mathematik_lehrerheft_A_05.pdf

Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen. (2005d). *Lernstand 9. Mathematik Auswertungsanleitung B1/B2 für Lehrerinnen und Lehrer.* Zugriff am 25.10.2011. Verfügbar unter http://www.standardsicherung.schulministerium.nrw.de/lernstand8/upload/download/Testaufgaben/mathematik_lehrerheft_B_05.pdf

Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (Hrsg.). (2007). *Kernlehrplan für das Gymnasium - Sekundarstufe I (G8) in Nordrhein-Westfalen. Mathematik* (Schule in NRW, Heft 3401 (G8)). Frechen: Ritterbach.

Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen. (2010a, 07. September). *Systematisches Probieren.* Zugriff am 26.11.2013. Verfügbar unter http://www2.standardsicherung.de/lernstand8/upload/download/mat_mathematik/Systematisches_Probieren.pdf

Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen. (2010b, 07. September). *Vorwärtsarbeiten (Schlussfolgern).* Zugriff am 04.02.2013. Verfügbar unter http://www2.standardsicherung.de/lernstand8/upload/download/mat_mathematik/Vorwrtsarbeiten.pdf

Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (Hrsg.). (2011). *Kernlehrplan und Richtlinien für die Hauptschule in Nordrhein-Westfalen. Mathematik* (Schule in NRW, Heft 3203). Frechen: Ritterbach.

Ministerium für Schule, Jugend und Kinder des Landes Nordrhein-Westfalen (Hrsg.). (2004a). *Kernlehrplan für die Gesamtschule – Sekundarstufe I in Nordrhein-Westfalen. Mathematik* (Schule in NRW, Heft 3106). Frechen: Ritterbach.

Ministerium für Schule, Jugend und Kinder des Landes Nordrhein-Westfalen (Hrsg.). (2004b). *Kernlehrplan für die Hauptschule in Nordrhein-Westfalen. Mathematik* (Schule in NRW, Heft 3203). Frechen: Ritterbach.

Ministerium für Schule, Jugend und Kinder des Landes Nordrhein-Westfalen (Hrsg.). (2004c). *Kernlehrplan für die Realschule in Nordrhein-Westfalen. Mathematik* (Schule in NRW, Heft 3302). Frechen: Ritterbach.

Möller, J., Bonerad, E.-M. & Pohlmann, B. (2006). Antwortmuster und Itemkennwerte. Ein unlösbares Item im KFT 4-12+R. *Diagnostica, 52,* 73–75.

Murayama, K., Pekrun, R., Lichtenfeld, S. & Vom Hofe, R. (2013). Predicting long-term growth in students' mathematics achievement. The unique contributions of motivation and cognitive strategies. *Child Development, 84,* 1475–1490.

NCTM (National Council of Teachers of Mathematics). (2000). *Principles and standards for school mathematics*. Reston, VA: NCTM.

Neubrand, M. (2003). "Mathematical literacy"/„Mathematische Grundbildung". Der Weg in die Leistungstests, die mathematikdidaktische Bedeutung, die Rolle als Interpretationshintergrund für den PISA-Test. *Zeitschrift für Erziehungswissenschaft, 6,* 338–356.

Neubrand, M., Biehler, R., Blum, W., Cohors-Fresenborg, E., Flade, L., Knoche, N. et al. (2001). Grundlagen der Ergänzung des internationalen PISA-Mathematik-Tests in der deutschen Zusatzerhebung. *ZDM Mathematics Education, 33,* 45–59. Deutsche PISA-Expertengruppe Mathematik.

Neubrand, M., Blum, W., Ehmke, T., Jordan, A., Senkbeil, M., Ulfig, F. et al. (2005). Mathematische Kompetenz im Ländervergleich. In M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand et al. (Hrsg.), *PISA 2003. Der zweite Vergleich der Länder in Deutschland - Was wissen und können Jugendliche?* (S. 51–84). Münster: Waxmann.

Neuenhaus, N., Artelt, C., Lingel, K. & Schneider, W. (2010). Fifth graders metacognitive knowledge: general or domain-specific? *European Journal of Psychology of Education, 26,* 163–178.

Newell, A. & Simon, H. A. (1972). *Human problem solving* (2. Aufl.). Englewood Cliffs, NJ: Prentice-Hall.

Nimon, K., Lewis, M., Kane, R. & Haynes, R. M. (2008). An R package to compute commonality coefficients in the multiple regression case. An introduction to the package and a practical example. *Behavior Research Methods, 40,* 457–466.

Nöllke, C. & Schmettkamp, M. (2011). *Präsentieren*. Freiburg (Breisgau): Haufe.

Norman, G. R. & Schmidt, H. G. (1992). The psychological basis of problem-based learning. A review of the evidence. *Academic Medicine, 67,* 557–565.

OECD. (2003). *The PISA 2003 assessment framework. Mathematics, reading, science and problem solving knowledge and skills*. Paris: OECD. Verfügbar unter http://www.oecd.org/dataoecd/46/14/33694881.pdf

OECD. (2004a). *Learning for tomorrow's world. First results from PISA 2003*. Paris: OECD.

OECD. (2004b). *Problem solving for tomorrow's world. First measures of cross-curricular competencies from PISA 2003*. Paris: OECD.

OECD. (2005). *PISA 2003 technical report*. Paris: OECD. Zugriff am 05.11.2013. Verfügbar unter http://www.oecd.org/edu/school/programmeforinternationalstudentassessmentpisa/35188570.pdf

OECD. (2013). *OECD skills outlook 2013. First results from the survey of adult skills*. Paris: OECD.

OECD. (2014). *PISA 2012 results: Creative problem solving. Students' skills in tackling real-life problems (Volume V)*. Paris: PISA, OECD Publishing.

Ohlsson, S. (2012). The problems with problem solving. Reflections on the rise, current status, and possible future of a cognitive research paradigm. *Journal of Problem Solving, 5,* 101–128.

Öllinger, M. & Knoblich, G. (2006). Lösen einfacher Probleme. In J. Funke & P. A. Frensch (Hrsg.), *Handbuch der Allgemeinen Psychologie. Kognition* (S. 431–438). Göttingen: Hogrefe.

O'Neill, T., Peabody, M., Tan, R. J. B. & Du, Y. (2013). How much item drift is too much? *Rasch Measurement Transactions, 27,* 1423–1424. Zugriff am 02.05.2014. Verfügbar unter http://www.rasch.org/rmt/rmt273a.htm

Opfermann, M., Azevedo, R. & Leutner, D. (2012). Metacognition and hypermedia learning. How do they relate? In N. M. Seel (Hrsg.), *Encyclopedia of the sciences of learning* (S. 2224–2228). New York, NY: Springer.

Otto, B., Perels, F. & Schmitz, B. (2008). Förderung mathematischen Problemlösens anhand eines Selbstregulationstrainings. *Zeitschrift für Pädagogische Psychologie, 22,* 221–232.

Otto, B., Perels, F., Schmitz, B. & Bruder, R. (2006). Längsschnittliche und prozessuale Evaluation eines Trainingsprogramms zur Förderung sachspezifischer und fächerübergreifender (selbstregulativer) Kompetenzen. In M. Prenzel & L. Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms* (BiQua, S. 211–239). Münster: Waxmann.

Paas, F. G. W. C. & van Merriënboer, J. J. G. (1993). The efficiency of instructional conditions. An approach to combine mental-effort and performance measures. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 35,* 737–743.

Paas, F. G. W. C. & van Merriënboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem solving skills. A cognitive load approach. *Journal of Educational Psychology, 86,* 122–133.

Paris, S. G., Lipson, M. Y. & Wixson, K. K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology, 8,* 293–316.

Pascha, A., Schöppe, B. & Hacker, W. (2001). Was macht Planen kompliziert? Zum Einfluss von Aufgabenmerkmalen auf die Schwierigkeit von Abfolgeplanung. *Zeitschrift für Psychologie / Journal of Psychology, 209,* 245–276.

Pehkonen, E. (1991). Developments in the understanding of problem solving. *ZDM Mathematics Education, 23,* 46–50.

Pelkner, A.-K. & Boehnke, K. (2003). Streber als Leistungsverweigerer? *Zeitschrift für Erziehungswissenschaft, 6,* 106–125.

Pelkner, A.-K., Günther, R. & Boehnke, K. (2004). Die Angst vor sozialer Ausgrenzung als leistungshemmender Faktor. In J. Doll & M. Prenzel (Hrsg.), *Bildungsqualität von Schule. Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsverbesserung* (S. 326–340). Münster: Waxmann.

Perels, F., Schmitz, B. & Bruder, R. (2005). Lernstrategien zur Förderung von mathematischer Problemlösekompetenz. In C. Artelt & B. Moschner (Hrsg.), *Lernstrategien und Metakognition. Implikationen für Forschung und Praxis* (S. 155–175). Münster: Waxmann.

Perels, F., Gürtler, T. & Schmitz, B. (2005). Training of self-regulatory and problem-solving competence. *Learning and Instruction, 15,* 123–139.

Perels, F., Schmitz, B. & Bruder, R. (2003). Trainingsprogramm zur Förderung der Selbstregulationskompetenz von Schülern der achten Gymnasialklasse. *Unterrichtswissenschaft, 31,* 23–37.

Perkins, D. N. & Salomon, G. (1989). Are cognitive skills context-bound? *Educational Researcher, 18,* 16–25.

Perler, D. & Wild, M. (2005). *Der Geist der Tiere. Philosophische Texte zu einer aktuellen Diskussion* (Suhrkamp Taschenbuch Wissenschaft, Bd. 1741). Frankfurt am Main: Suhrkamp.

Peterßen, W. H. (2000). *Handbuch Unterrichtsplanung. Grundfragen, Modelle, Stufen, Dimensionen* (9. Aufl.). München: Oldenbourg.

PISA-Konsortium Deutschland (Hrsg.). (2006). *PISA 2003. Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres*. Münster: Waxmann.

Pólya, G. (1945). *How to solve it. A new aspect of mathematical method*. Princeton: University Press.

Pólya, G. (1973). *How to solve it. A new aspect of mathematical method* (2. Aufl.). Princeton: University Press.

Popper, K. R. (2005). *Logik der Forschung* (Gesammelte Werke in deutscher Sprache, Bd. 3, 11. Aufl.). Tübingen: Mohr Siebeck.

Prenzel, M. & Allolio-Näcke, L. (2006). Das DFG-Schwerpunktprogramm Bildungsqualität von Schule – ein Überblick. In M. Prenzel & L. Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms* (BiQua, S. 7–20). Münster: Waxmann.

Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., Neubrand, M. et al. (2005). *PISA 2003: Ergebnisse des zweiten Ländervergleichs. Zusammenfassung,* IPN Kiel. Zugriff am 21.05.2012. Verfügbar unter http://pisa.ipn.uni-kiel.de/PISA2003_E_Zusammenfassung.pdf

Prenzel, M. & Drechsel, B. (2004). "Wie PISA zeigt…" - Konsequenzen aus internationalen Leistungsvergleichen und ihre Begründung. In K. Reiss (Hrsg.), *Beiträge zum Mathematikunterricht 2004* (S. 31–39). Hildesheim: Franzbecker.

Prenzel, M., Sälzer, C., Klieme, E. & Köller, O. (Hrsg.). (2013a). *PISA 2012. Fortschritte und Herausforderungen in Deutschland*. Münster: Waxmann.

Prenzel, M., Sälzer, C., Klieme, E. & Köller, O. (Hrsg.). (2013b, 03. Dezember). *PISA 2012: Fortschritte und Herausforderungen in Deutschland. Zusammenfassung*. Zugriff am 03.12.2013. Verfügbar unter http://www.pisa.tum.de/fileadmin/w00bgi/www/Berichtband_und_Zusammenfassung_2012/PISA_Zusammenfassung_online.pdf

Prenzel, M., Walter, O. & Frey, A. (2007). PISA misst Kompetenzen. *Psychologische Rundschau, 58,* 128–136.

Pugalee, D. K. (2001). Writing, mathematics, and metacognition. Looking for connections through students' work in mathematical problem solving. *School Science and Mathematics, 101,* 236–245.

Pugalee, D. K. (2004). A comparison of verbal and written descriptions of students' problem solving processes. *Educational Studies in Mathematics, 55,* 27–47.

QuestBack AG. (2012) Enterprise Feedback Suite: EFS Survey 8.2 [Computer software]. Köln: QuestBack AG. Verfügbar unter http://www.unipark.info

QuestBack GmbH. (2013) Enterprise Feedback Suite: EFS Survey 10 [Computer software]. Köln: QuestBack GmbH. Verfügbar unter http://unipark.info/

R Core Team. (2014). *R: A language and environment for statistical computing,* R Foundation for Statistical Computing. Verfügbar unter http://www.R-project.org

Rathmell, E. C. & Huinker, D. M. (1989). Using "part-whole" language to help children represent and solve word problems. In P. R. Trafton (Hrsg.), *New directions for elementary school mathematics* (2. Aufl., S. 99–109). Reston, VA: National Council of Teachers of Mathematics, U.S.

Reiss, K. & Törner, G. (2007). Problem solving in the mathematics classroom. The German perspective. *ZDM Mathematics Education, 39,* 431–441.

Renkl, A. (2012). Modellierung von Kompetenzen oder von interindividuellen Kompetenzunterschieden. *Psychologische Rundschau, 63,* 50–53.

Renkl, A. & Stern, E. (1994). Die Bedeutung von kognitiven Eingangsvoraussetzungen und schulischen Lerngelegenheiten für das Lösen von einfachen und komplexen Textaufgaben. *Zeitschrift für Pädagogische Psychologie, 8,* 27–39.

Rey, G. D. (2009). *E-Learning. Theorien, Gestaltungsempfehlungen, Forschung und Anwendungen*. Bern: Huber.

Rey, G. D. & Buchwald, F. (2011). The expertise reversal effect. Cognitive load and motivational explanations. *Journal of Experimental Psychology: Applied, 17,* 33–48.

Rheinberg, F., Vollmeyer, R. & Burns, B. D. (2001). FAM. Ein Fragebogen zur Erfassung aktueller Motivation in Lern- und Leistungssituationen. *Diagnostica, 47,* 57–66.

Rieß, W., Wirtz, M., Barzel, B. & Schulz, A. (Hrsg.). (2012). *Experimentieren im mathematisch-naturwissenschaftlichen Unterricht. Schüler lernen wissenschaftlich denken und arbeiten*. Münster: Waxmann.

Rindermann, H. (2006). Was messen internationale Schulleistungsstudien? *Psychologische Rundschau, 57,* 69–86.

Rittle-Johnson, B., Star, J. R. & Durkin, K. (2009). The importance of prior knowledge when comparing examples. Influences on conceptual and procedural knowledge of equation solving. *Journal of Educational Psychology, 101,* 836–852.

Robertson, S. I. (2001). *Problem solving*. Philadelphia, PA: Psychology Press.

Robitzsch, A. (2014a). *miceadds. Some additional multiple imputation functions, especially for mice* (R package version 0.12-9). Zugriff am 10.05.2014. Verfügbar unter http://CRAN.R-project.org/package=miceadds

Robitzsch, A. (März 2014b). *Missing Data in R* (Frühjahrsakademie 2014 "Methoden der empirischen Bildungsforschung"). Berlin: IQB - Institut zur Qualitätsentwicklung im Bildungswesen, Forschungsdatenzentrum.

Rocklin, T. R. (1994). Self-Adapted Testing. *Applied Measurement in Education, 7,* 3–14.

Rodewald, K., Rentrop, M., Holt, D. V., Roesch-Ely, D., Backenstrass, M., Funke, J. et al. (2011). Planning and problem-solving training for patients with schizophrenia. A randomized controlled trial. *BMC Psychiatry, 11,* 1–11.

Rost, D. H. & Sparfeldt, J. R. (2002). Facetten des schulischen Selbstkonzeptes. *Diagnostica, 48,* 130–140.

Rourke, A. & Sweller, J. (2009). The worked-example effect using ill-defined problems. Learning to recognise designers' styles. *Learning and Instruction, 19,* 185–199.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: Wiley.

Rumann, S., Fleischer, J., Stawitz, H., Wirth, J. & Leutner, D. (2010). Vergleich von Profilen der Naturwissenschafts- und Problemlöse-Aufgaben der PISA 2003-Studie. *Zeitschrift für Didaktik der Naturwissenschaften, 16,* 315–327.

Schafer, J. L. & Olsen, M. K. (1998). Multiple imputation for multivariate missing-data problems. A data analyst's perspective. *Multivariate Behavioral Research, 33* (4), 545–571.

Schank, R. & Abelson, R. (1977). *Scripts, plans, goals, and understanding*. Hillsdale, NJ: Lawrence Erlbaum. Zugriff am 17.02.2014. Verfügbar unter http://colinallen.dnsalias.org/Secure/Q240/1977-SchankAbelson.pdf

Schaper, N. (2004). Förderung des Wissens- und Fähigkeitstransfers beim Einsatz von Lernsoftware. In D. M. Meister (Hrsg.), *Online-Lernen und Weiterbildung* (Bildung und neue Medien, Bd. 5, S. 105–136). Wiesbaden: VS, Verlag für Sozialwissenschaften.

Scherer, R. (2012). *Analyse der Struktur, Messinvarianz und Ausprägung komplexer Problemlösekompetenz im Fach Chemie. Eine Querschnittstudie in der Sekundarstufe I und am Übergang zur Sekundarstufe II* (Studien zum Physik- und Chemielernen, Bd. 141). Berlin: Logos.

Scherer, R. & Tiemann, R. (2012). Factors of problem-solving competency in a virtual chemistry environment. The role of metacognitive knowledge about strategies. *Computers & Education, 59,* 1199–1214.

Schiefele, U. & Csikszentmihalyi, M. (1995). Motivation and ability as factors in mathematics experience and achievement. *Journal for research in mathematics education, 26,* 163–181.

Schlagmüller, M. & Schneider, W. (2007). *WLST 7-12. Würzburger Lesestrategie-Wissenstest für die Klassen 7-12. Ein Verfahren zur Erfassung metakognitiver Kompetenzen bei der Verarbeitung von Texten* (Deutsche Schultests). Göttingen: Hogrefe.

Schlagmüller, M., Visé, M. & Schneider, W. (2001). Zur Erfassung des Gedächtniswissens bei Grundschulkindern. Konstruktionsprinzipien und empirische Bewährung der Würzburger Testbatterie zum deklarativen Metagedächtnis. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 33,* 91–102.

Schmid, U. (2008). Computermodelle des Problemlösens. In J. Müsseler (Hrsg.), *Allgemeine Psychologie* (2. Aufl., S. 600–628). Berlin: Spektrum.

Schmid, U., Wirth, J. & Polkehn, K. (2003). A closer look at structural similarity in analogical transfer. *Cognitive Science Quarterly, 3,* 57–89.

Schmidt-Atzert, L. & Amelang, M. (2012). *Psychologische Diagnostik* (5. Aufl.). Berlin: Springer.

Schmitz, C., Le Meur, T., Olivier, D., Cleeland, J., Wills, E., Duy Hai, M. et al. (2012) LimeSurvey [Computer software]. Verfügbar unter www.limesurvey.org

Schneider, W. & Artelt, C. (2010). Metacognition and mathematics education. *ZDM Mathematics Education, 42,* 149–161.

Schneider, W., Baumert, J., Becker-Mrotzek, M., Hasselhorn, M., Kammermeyer, G., Rauschenbach, T. et al. (2012). *Expertise „Bildung durch Sprache und Schrift (BISS)“,* Bund-Länder-Initiative zur Sprachförderung, Sprachdiagnostik und Leseförderung. Zugriff am 12.02.2014. Verfügbar unter http://www.bmbf.de/pubRD/BISS_Expertise.pdf

Schoenfeld, A. H. (1982). Some thoughts on problem-solving research and mathematics education. In F. K. Lester, Jr. & J. Garofalo (Hrsg.), *Mathematical problem solving. issues and research.* Philadelphia, PA: The Franklin Institute Press.

Schoenfeld, A. H. (1983a). Episodes and executive decisions in mathematical problem solving. In R. Lesh & M. Landau (Hrsg.), *Acquisition of mathematics concepts and processes* (S. 345–395). New York, NY: Academic Press.

Schoenfeld, A. H. (1983b). The wild, wild, wild, wild, wild world of problem solving. A review of sorts. *For the learning of mathematics, 3* (3), 40–47.

Schoenfeld, A. H. (1985). *Mathematical problem solving*. Orlando: Academic Press.

Schoenfeld, A. H. (1987). Pólya, problem solving, and education. *Mathematics Magazine, 60,* 283–291.

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

Schraw, G. J. (1998). Promoting general metacognitive awareness. *Instructional Science, 26,* 113–125.

Schraw, G. & Moshman, D. (1995). Metacognitive Theories. *Educational Psychology Review, 7,* 351–371.

Schroeder, T. L. & Lester, F. K. (1989). Developing understanding in mathematics via problem solving. In P. R. Trafton (Hrsg.), *New directions for elementary school mathematics* (2. Aufl., S. 31–42). Reston, VA: National Council of Teachers of Mathematics, U.S.

Schupp, H. (1988). Anwendungsorientierter Mathematikunterricht in der Sekundarstufe I zwischen Tradition und neuen Impulsen. *Mathematikunterricht (MU), 34* (6), 5–16.

Schütte, M. (2012). *Selbstreguliertes Lernen aus Sachtexten. Modellierung und Erfassung der erforderlichen Teilkompetenzen*. Dissertation, Universität Duisburg-Essen. Essen. Zugriff am 13.06.2013. Verfügbar unter http://duepublico.uni-duisburg-essen.de/servlets/DerivateServlet/Derivate-30745/Diss_schuette.pdf

Schütte, M., Wirth, J. & Leutner, D. (2010). Selbstregulationskompetenz beim Lernen aus Sachtexten. Entwicklung eines Kompetenzstrukturmodells. *Zeitschrift für Pädagogik* (Beiheft 56), 249–257.

Schütte, M., Wirth, J. & Leutner, D. (2012). Lernstrategische Teilkompetenzen für das selbstregulierte Lernen aus Sachtexten. *Psychologische Rundschau, 63,* 26–33.

Schwartz, J. L. & Yerushalmy, M. (1995). On the need for a bridging language for mathematical modelling. *For the learning of mathematics, 15* (2), 29–35.

Segerer, R., Marx, A. & Marx, P. (2012). Unlösbare Items im KFT 4-12+R. *Diagnostica, 58,* 45–50.

Selter, C. (1994). Jede Aufgabe hat eine Lösung. Vom rationalen Kern irrationalen Vorgehens. *Grundschule* (3), 20–22.

Selter, C. (2009). Kinder rechnen anders. Ein Projekt der TU Dortmund und der Deutschen Telekom Stiftung. *Grundschulzeitschrift, 25* (222/223), 20.

Shahat, M. A., Ohle, A., Treagust, D. F. & Fischer, H. E. (2013). Design, development and validation of a model of problem solving for egyptian science classes. *International Journal of Science and Mathematics Education, 11,* 1157–1181.

Short, E. J. & Weissberg-Benchell, J. A. (1989). The triple alliance for learning. Cognition, metacognition, and motivation. In C. B. McCormick, G. Miller & M. Pressley (Hrsg.), *Cognitive strategy research. From basic research to educational application* (S. 33–63). New York, NY: Springer.

Sieprinska, A. (2004). Research in mathematics education through a keyhole. Task problematization. *For the learning of mathematics, 24* (2), 7–15.

Silver, E. A. (1994). On mathematical problem posing. *For the learning of mathematics, 14* (1), 19–28.

Silver, E. A. & Cai, J. (1996). An analysis of arithmetic problem posing by middle school students. *Journal for research in mathematics education, 27,* 521–539.

Simmons, J. P., Nelson, L. D. & Simonsohn, U. (2011). False-positive psychology. Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22,* 1359–1366.

Singley, M. K. & Anderson, J. R. (1989). *Transfer of cognitive skill*. Cambridge, MA: Harvard University Press.

Sonnleitner, P., Brunner, M., Greiff, S., Funke, J., Keller, U., Martin, R. et al. (2012). The Genetics Lab. Acceptance and psychometric characteristics of a computer-based microworld assessing complex problem solving. *Psychological Test and Assessment Modeling, 54,* 54–72.

Sonntag, W. (2004). Der Einfluss des Klauerschen Denktrainings auf mathematisches Denken und Lernen von lernbehinderten Sonderschülern. *Zeitschrift für Pädagogische Psychologie, 18,* 101–111.

Souvignier, E., Förster, N. & Salaschek, M. (2014). quop: Ein Ansatz internetbasierter Lernverlaufsdiagnostik mit Testkonzepten für Lesen und Mathematik.

In M. Hasselhorn, W. Schneider & U. Trautwein (Hrsg.), *Lernverlaufsdiagnostik* (Tests und Trends, Bd. 12, S. 239–256). Göttingen: Hogrefe.

Spencer, S. J., Steele, C. M. & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35,* 4–28.

Spinath, B., Pelster-Stiensmeier, J., Schöne, C. & Dickhäuser, O. (2002). *SELLMO. Skalen zur Erfassung der Lern- und Leistungsmotivation*. Göttingen: Hogrefe.

Stacey, K. (1991). Linking application and acquisition of mathematical ideas through problem solving. *ZDM Mathematics Education, 23,* 8–14.

Stacey, K. (2005). The place of problem solving in contemporary mathematics curriculum documents. *Journal of Mathematical Behavior, 24,* 341–350.

Stawitz, H., Rumann, S., Fleischer, J. & Wirth, J. (2008). Analyse von Naturwissenschaften- und Problemlöse-Aufgaben bei PISA 2003. In D. Höttecke (Hrsg.), *Kompetenzen, Kompetenzmodelle, Kompetenzentwicklung. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Essen 2007* (Gesellschaft für Didaktik der Chemie und Physik, Bd. 28, S. 407–409). Berlin: LIT-Verlag.

Stein, B. S. (1986). Constraints on spontaneous transfer in problem-solving tasks. *Memory & Cognition, 14,* 432–441.

Stern, E. (1992). Warum werden Kapitänsaufgaben "gelöst"? Das Verstehen von Textaufgaben aus psychologischer Sicht. *Mathematikunterricht (MU), 38* (5), 7–29.

Stevens, R. H. & Thadani, V. (2007). Quantifying student's scientific problem solving efficiency and effectiveness. *Technology Instruction Cognition and Learning, 5,* 325–337.

Stillman, G. & Mevarech, Z. (2010). Metacognition research in mathematics education. From hot topic to mature field. *ZDM Mathematics Education, 42,* 145–148.

Süß, H. M. (1996). *Intelligenz, Wissen und Problemlösen*. Göttingen: Hogrefe.

Swanson, H. L. (1990). Influence of metacognitive knowledge and aptitude on problem solving. *Journal of Educational Psychology, 82,* 306–314.

Swanson, H. L. (2011). Working memory, attention, and mathematical problem solving. A longitudinal study of elementary school children. *Journal of Educational Psychology, 103,* 821–837.

Swanson, H. L. & Cooney, J. B. (1993). The influence of working memory and classification ability on childrens word problem solution. *Journal of Experimental Child Psychology, 55,* 374–395.

Taylor, K. L. & Dionne, J. P. (2000). Accessing problem-solving strategy knowledge. The complementary use of concurrent verbal protocols and retrospective debriefing. *Journal of Educational Psychology, 92,* 413–425.

Templ, M., Alfons, A., Kowarik, A. & Prantner, B. (2013). *VIM: Visualization and imputation of missing values* (R package version 4.0.0). Zugriff am 30.05.2014. Verfügbar unter http://CRAN.R-project.org/package=VIM

Tenorth, H.-E. (2011, Oktober). *Einheit in der Vielfalt - oder: Kann man die Kompetenzforschung noch an schulische Arbeit und pädagogische Reflexion rückbinden?* Jahreskolloquium des DFG-Schwerpunktprogramms „Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen“, Berlin.

Teong, S. K. (2003). The effect of metacognitive training on mathematical word-problem solving. *Journal of Computer Assisted Learning, 19,* 46–55.

Thillmann, H. (2007). *Selbstreguliertes Lernen durch Experimentieren. Von der Erfassung zur Förderung*. Dissertation, Universität Duisburg-Essen. Essen.

Thomas, J. C. (1974). An analysis of behavior in the Hobbits-Orcs problem. *Cognitive Psychology, 6,* 257–269.

Todd, P. M. & Gigerenzer, G. (2000). Précis of "Simple heuristics that make us smart". *Behavioral and Brain Sciences, 23,* 727–780.

Tomic, W. (1995). Training in inductive reasoning and problem solving. *Contemporary Educational Psychology, 20,* 483–490.

Törner, G. & Zielinski, U. (1992). Problemlösen als integraler Bestandteil des Mathematikunterrichts. Einblicke und Konsequenzen. *Journal für Mathematik-Didaktik, 13,* 253–270.

Tricot, A. & Sweller, J. (2014). Domain-specific knowledge and why teaching generic skills does not work. *Educational Psychology Review, 26,* 265–283.

Van Buuren, S. & Groothuis-Oudshoorn, K. (2011). mice. Multivariate imputation by chained equations in R. *Journal of Statistical Software, 45* (3), 1–67. Verfügbar unter http://www.jstatsoft.org/v45/i03/

Van den Heuvel-Panhuizen, M. I. (2000). *Mathematics education in the Netherlands. A guided tour*. Freudenthal Institute Cd-rom for ICME9. Utrecht: Utrecht University. Zugriff am 20.12.2013. Verfügbar unter http://dme.colorado.edu/fius/rme_tour.pdf

Van den Heuvel-Panhuizen, M. I. (2003). The didactical use of models in realistic mathematics education. An example from a longitudinal trajectory on percentage. *Educational Studies in Mathematics, 54,* 9–35.

Van den Heuvel-Panhuizen, M. I. & Bodin-Baarends, C. (2004). All or nothing. Problem solving by high achievers in mathematics. *Journal of the Korea Society of Mathematical Education, 8,* 115–121.

Van der Linden, W. J. & Glas, C. A. (Hrsg.). (2000). *Computerized adaptive testing. Theory and practice*. New York, NY: Springer.

Van der Linden, W. J. & Glas, C. A. (2010). *Elements of adaptive testing* (Statistics for Social and Behavioral Sciences). New York, NY: Springer.

Van Someren, M. W., Barnard, Y. F. & Sandberg, J. A. C. (1994). *The think aloud method. A practical guide to modelling cognitive processes*. London: Academic Press.

Vlcek, R. (2000). *Workshop Improvisationstheater. Übungs- und Spielesammlung für Theaterarbeit, Ausdrucksfindung und Gruppendynamik* (2. Aufl.). München: Auer.

Vollmer, G. (1992). *Gelöste, ungelöste und unlösbare Probleme. Zu den Bedingungen wissenschaftlichen Fortschritts* (Veröffentlichung der Joachim Jungius-Gesellschaft der Wissenschaften, Bd. 68). Göttingen: Vandenhoeck & Ruprecht.

Vollmer, G. (1995). Wir irren uns empor. Zum Tode des Philosophen Karl Raimund Popper. *Skeptiker, 8,* 4–6.

Vollmeyer, R. & Funke, J. (1999). Personen- und Aufgabenmerkmale beim komplexen Problemlösen. *Psychologische Rundschau, 50,* 213–219.

Von der Groeben, A. & Kaiser, I. (2011). Rampe, Fächer, Blüte, Gerüst. Aufgabendifferenzierung (1). *Pädagogik, 63,* 40–45.

Wagner, J. (2010). The fraction of missing information as a tool for monitoring the quality of survey data. *Public Opinion Quarterly, 74* (2), 223–243.

Walter, O., Ramm, G., Zimmer, K., Heidemeier, H. & Prenzel, M. (2006). PISA 2003 - Kompetenzen von Jungen und Mädchen mit Migrationshintergrund in Deutschland. Ein Problem ungenutzter Potentiale? *Unterrichtswissenschaft, 34,* 146–169.

Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika, 54,* 427–450.

Weidenmann, B. (2008). *Handbuch Active Training. Die besten Methoden für lebendige Seminare* (2. Aufl.). Weinheim: Beltz, PVU.

Weinert, F. E. (1999). *Concepts of competence. Definition and selection of competencies*. Paris: OECD. Zugriff am 18.03.2014. Verfügbar unter http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.111.1152&rep=rep1&type=pdf

Welsh, M. C. & Huizinga, M. (2005). Tower of Hanoi disk-transfer task. Influences of strategy knowledge and learning on performance. *Learning and Individual Differences, 15,* 283–298.

Wernicke, J. (2013, 29. November). "Schluss mit dem Geteste". Interview mit Wolfram Meyerhöfer. *TAZ.* Zugriff am 30.11.2013. Verfügbar unter http://www.taz.de/Mathedidaktiker-ueber-Pisa/!128233/

Wilhelm, O. & Nickolaus, R. (2013). Was grenzt das Kompetenzkonzept von etablierten Kategorien wie Fähigkeit, Fertigkeit oder Intelligenz ab? *Zeitschrift für Erziehungswissenschaft, 16* (Sonderheft 18), 23–26.

Winkelmann, H., Robitzsch, A., Stanat, P. & Köller, O. (2012). Mathematische Kompetenzen in der Grundschule. *Diagnostica, 58,* 15–30.

Wirth, J. & Klieme, E. (2003). Computer-based assessment of problem solving competence. *Assessment in Education, 10,* 329–345.

Wu, M. L., Adams, R. J., Wilson, M. R. & Haldane, S. A. (2007). *ACER ConQuest Version 2. Generalised item response modelling software*. Camberwell: The Australian Council for Educational Research.

Wüstenberg, S., Greiff, S. & Funke, J. (2012). Complex problem solving — More than reasoning? *Intelligence, 40,* 1–14.

Wuttke, J. (2006). Fehler, Verzerrungen, Unsicherheiten in der PISA-Auswertung. In T. Jahnke & W. Meyerhöfer (Hrsg.), *PISA & Co - Kritik eines Programmes* (S. 101–154). Hildesheim: Franzbecker.

Yoshida, H., Verschaffelt, L. & De Corte, E. (1997). Realistic considerations in solving problematic word problems. Do Japanese and Belgian children have the same difficulties. *Learning and Instruction, 7,* 329–338.

Young, J. W. A. (1924). *The Teaching of Mathematics in the Elementary and the Secondary School*. New York, NY: Longmans & Co.

Zimmer, K., Burba, D. & Rost, J. (2004). Kompetenzen von Jungen und Mädchen. In PISA-Konsortium Deutschland (Hrsg.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland - Ergebnisse des zweiten internationalen Vergleichs* (S. 211–224). Münster: Waxmann.

Zimmerman, B. J. (1995). Self-regulation involves more than metacognition. A social cognitive perspective. *Educational Psychologist, 30,* 217–221.

Zimmerman, B. J. (2008). Investigating self-regulation and motivation. Historical background, methodological developments, and future prospects. *American Educational Research Journal, 45,* 166-183.

Zöttl, L., Heinze, A. & Reiss, K. (2007). Problemlösen im Kontext. Unterschiede in der Bearbeitung von Alltagsproblemen und mathematischen Problemen. In A. Peter-Koop & A. Bikner-Ahsbahs (Hrsg.), *Mathematische Bildung - mathematische Leistung. Festschrift für Michael Neubrand zum 60. Geburtstag* (S. 217–232). Hildesheim: Franzbecker.

Zucker, S., Sassman, C. & Case, B. J. (2004, 20. Februar). *Cognitive Labs. Technical Report*. San Antonio, TX: Pearson Inc. Zugriff am 16.12.2011. Verfügbar unter http://www.pearsonassessments.com/NR/rdonlyres/E5CD33E6-D234-46F3-885A-9358575372FB/0/CognitiveLabs.pdf

Anhang

Anhang A. Motivationale Unterschiede in der Einschätzung von Items der Domänen Problemlösen und Mathematik?

In der PISA-Studie 2003 zeigten sich für einige Länder und Subgruppen Kompetenzunterschiede zwischen den Domänen Problemlösen und Mathematik (OECD, 2004b). War bei PISA 2003 womöglich die Testmotivation der SuS bei der Bearbeitung der Mathematik-Items geringer als bei den Problemlöse-Items? Finden SuS Problemlösen „toll“ und Mathematik „öde“ (Leutner, 2013b)?

Testmotivation als Teil der Leistungsmotivation ist ein Augenschein-valides, theoretisch plausibles Konstrukt, das bei *low stakes testing* oft statistisch ignoriert wird, d. h. bei Testungen, die für die Getesteten relativ frei von Risiken und negativen Folgen im Falle des Scheiterns sind. Insbesondere bei leistungsschwächeren SuS ist fraglich, ob sie ihre theoretisch erwartete maximale Testleistung zeigen, wenn eine Testung für sie nicht persönlich bedeutsam ist (Asseburg & Frey, 2013).

Experimentelle Befunde (Max-Planck-Institut für Bildungsforschung, Berlin, n.d.) in Kontext von PISA 2000 zeigten für einen Mathematiktest, dass sich SuS, die glaubten, an einer internationalen Bildungsstudie teilzunehmen, in ihrer selbst berichteten Anstrengungsbereitschaft und in ihrer Testleistung weder von SuS unterscheiden,

- die glaubten, an einem Mathematiktest teilzunehmen, zu dem sie eine Rückmeldung von ihrem Lehrer über die Anzahl gelöster Aufgaben bekommen werden,
- noch von SuS, die glaubten, der Mathematiktest würde von ihrem Mathematiklehrer benotet,
- noch von SuS, die glaubten, sie würden für 10 Deutsche Mark an einem Mathematiktest teilnehmen.

Diese Befunde können jedoch nicht die Fragen beantworten, ob sich die SuS im Test maximal angestrengt haben und ob es bei Problemlösetests andere Ergebnisse bezüglich motivationaler Faktoren gibt als bei Mathematiktests (vgl. Kontexteffekte).

Um den teilweise gefundenen Unterschied von Problemlösekompetenz und Mathematikkompetenz durch Testmotivation (partiell) zu erklären, müsste es zwischen den teilnehmenden Ländern differentielle motivationale Effekte der teilnehmenden SuS hinsichtlich der beiden Domänen geben. Empirische Belege für differentielle motivationale Effekte für die Testdomänen Mathematik bzw. Problemlösen lassen sich auch aus den bei PISA 2003 erhobenen Daten nicht gewinnen, da affektiv-motivationale Variablen nicht pro Testdomäne erfasst worden sind (Deutsches PISA-Konsortium, 2006). Die These domänen-spezifischer motivationaler Effekte ist mit den PISA-2003-Daten nicht zu beantworten. Sie wird daher anhand von Umfrage-Daten einer deutschen Realschüler-Stichprobe exploriert. Wenn domänenspezifische Unterschiede in der Testmotivation gefunden werden, spricht dies eher für eine mangelnde Potenzial*nutzung* (Fleischer et al., 2014) als für eine mangelnde Potenzialausschöpfung (Leutner et al., 2004), wobei sich beide Hypothesen nicht gegenseitig ausschließen.

Fragestellungen

Um Aspekte einer motivationalen Erklärung der bei PISA 2003 beobachteten Diskrepanz fächerübergreifender Problemlösekompetenz und mathematischer Kompetenz in Deutschland zu erforschen, wird mittels eines Fragebogens geprüft, ob Aufgaben aus den Domänen Problemlösen und Mathematik aus PISA 2003 (OECD, 2004a) von SuS unterschiedlich wahrgenommen werden. Dazu wird untersucht, ob sich die spontane Einschätzung von ähnlich schwierigen Mathematik- und Problemlöseaufgaben aus PISA 2003 hinsichtlich der eingeschätzten Schwierigkeit und der vier leistungsmotivationalen Faktoren Interesse, Hoffnung auf Erfolg, Misserfolgsbefürchtung und Herausforderung bzw. Anstrengungsbereitschaft (Rheinberg et al., 2001) unterscheiden. Konkret werden folgende Hypothesen untersucht: Mathematikaufgaben werden als schwieriger eingeschätzt als Problemlöseaufgaben (Hypothese 1). Bei Mathematikaufgaben ist die Hoffnung auf Erfolg geringer als bei Problemlöseaufgaben (Hypothese 2). Mathematikaufgaben werden als weniger interessant eingeschätzt als Problemlöseaufgaben (Hypothese 3). Mathematikaufgaben rufen mehr Misserfolgsbefürchtung hervor als Problemlöseaufgaben (Hypothese 4). Mathematikaufgaben werden als herausfordernder eingeschätzt als Problemlöseaufgaben (Hypothese 5).

Methode

Stichprobe

Es nahmen 117 SuS an der Online-Umfrage teil (45 Mädchen, 57 Jungen, 15 Personen machten keine Angabe zu ihrem Geschlecht). Das durchschnittliche Alter der SuS beträgt 14.92 Jahre (*SD* = 1.00). Die SuS besuchten überwiegend Realschulen (111x Realschule, 1x Gymnasium, 1x Hauptschüler, 4x keine Angabe).

Instrumente

Die Umfrage besteht aus fünf Problemlöse- und fünf Mathematikaufgaben aus PISA 2003 (OECD, 2004b, 2005), die jeweils einzeln auf einer HTML-Seite präsentiert wurden. Die Reihenfolge der Aufgabenblöcke (Problemlösen vs. Mathematik) und der Aufgaben innerhalb jedes Blocks wurden randomisiert. Die SuS sollten die Aufgaben nicht lösen, sondern nach dem Lesen einer Aufgabe jeweils fünf Items beantworten: Ein Schwierigkeitsrating (*Wie schwierig ist diese Aufgabe für dich?* Skala: 1 = *sehr leicht*, bis 6 = *sehr schwer*) und je ein Item der vier Skalen des Fragebogens zur Erfassung aktueller Motivation in Lern- und Leistungssituationen (FAM, Rheinberg et al., 2001; Skala: 1 = *trifft nicht zu* bis 7 = *trifft zu*):

1. *Nach dem Lesen erscheint mir die Aufgabe sehr interessant* (Interesse),
2. *Ich bin fest entschlossen, mich bei dieser Aufgabe voll anzustrengen* (Herausforderung),
3. *Ich fürchte mich ein wenig davor, dass ich mich hier blamieren könnte* (Misserfolgsbefürchtung),
4. *Ich glaube, der Schwierigkeit dieser Aufgabe gewachsen zu sein* (Hoffnung auf Erfolg).

Nach der Einschätzung der 10 Aufgaben wurden demographische Angaben (Geschlecht, Alter, Schulform) erfasst.

Aufgabenauswahl: Um eine Vergleichbarkeit der Aufgaben hinsichtlich der Schwierigkeit und Oberflächen-Merkmalen zu gewährleisten, wurde ein *matching* vorgenommen, dass jeder Problemlöseaufgabe je eine Mathematikaufgabe zuordnet (Tabelle A-1), die dieser hinsichtlich Aufgabenschwierigkeit, Vorhandensein von Abbildungen bzw. Tabellen, und Textmenge möglichst ähnlich ist. Die Problemlöseaufgaben enthalten allerdings etwas mehr Text (Fleischer et al., 2010).

Tabelle A-1. Verwendete Items aus PISA 2003 (OECD, 2005) mit Angabe von Lösungsquote und Standardabweichung

Problemlösen			*Mathematik*		
Name	*p [%]*	*SD*	*Name*	*p [%]*	*SD*
Bibliothekensystem 1	74.8	0.25	Das beste Auto 1	72.91	0.24
Bibliothekensystem 2	14.32	0.16	Das beste Auto 2	25.42	0.21
Studienplan	31.06	0.23	Testergebnisse	32.21	0.25
Anschlusszüge	24.15	0.17	Raubüberfälle	29.5	0.22
Bewässerung 1	62.88	0.25	Skateboard 1	72.01	0.25
Bewässerung 2	51.34	0.28	Skateboard 2	45.53	0.26
Bewässerung 3	54.44	0.28	Skateboard 3	49.78	0.29
Ferienlager	40.09	0.23	Erdbeben	46.48	0.22

Durchführung

Die Umfrage wurde im Frühjahr 2014 telefonisch und per E-Mail an Schulen in Nordrhein-Westfalen beworben. Es wurde den Schulen überlassen, ob die SuS in der Schule oder Zuhause an der Umfrage teilnehmen. Die Logfiles legen nahe, dass die Umfrage hauptsächlich während der Schulzeiten im Klassenverband absolviert wurde. Es wurde keine Belohnung gezahlt. Schulen und Schülern wurde die Möglichkeit angeboten, nach Abschluss des Projekts über die Ergebnisse informiert zu werden. Die Erhebung fand in Form einer Online-Umfrage mit EFS Survey 10.1 (QuestBack GmbH, 2013) statt. Jede Aufgabe wurde auf einer HTML-Seite präsentiert. Unter der jeweiligen Aufgabe waren jeweils die fünf Fragen zu beantworten (siehe Instrumente). Die geschätzte Umfragezeit lag bei 20 Minuten, eine gute Schätzung der realen Bearbeitungszeit ($M = 11.80$, $SD = 5.64$). Es wurde darauf hingewiesen, spontan und ehrlich zu antworten.

Ergebnisse

Die Histogramme der – jeweils per Mittelwertsbildung über die Aufgaben – aggregierten Einschätzungen für Mathematik und Problemlösen sind in Abbildung A-1 zu finden. Die eingeschätzte Schwierigkeit ist deskriptiv für Mathematikaufgaben ($M = 3.53$, $SD = 1.12$) niedriger als für Problemlöseaufgaben ($M = 3.72$, $SD = 1.23$). In dieses Bild passt auch, dass die Hoffnung auf Erfolg bei den Mathematikaufgaben etwas größer eingeschätzt wird ($M = 4.23$, $SD = 1.25$) als bei den Problemlöseaufgaben ($M = 4.20$, $SD = 1.43$). Bei der Interessensfrage fällt auf, dass die Verteilungen eher rechtsschief sind, mit Häufungen im mittleren und unteren Skalenbereich. Mathematikaufgaben ($M = 3.35$, *SD* = 1.37) werden deskriptiv als etwas interessanter als Problemlöseaufgaben (*M* = 3.28, *SD* = 1.44) eingeschätzt. Bei der Einschätzung der Misserfolgsbefürchtung streuen die Häufigkeiten der Problemlöseaufgaben breit im Skalenbereich von 1 bis 5, wobei der Skalenbereich 6 bis 7 kaum besetzt ist. Für die Mathematikaufgaben zeigt sich bzgl. der Misserfolgsbefürchtung dasselbe Muster mit einer Ausnahme: Bei der Einschätzung der Misserfolgsbefürchtung bezüglich der Mathematikaufgaben gibt es eine Häufung von Personen am Minimum (1) der Skala, d. h. von Personen mit sehr niedriger eingeschätzter Misserfolgsbefürchtung. Diese Häufung ist bei der Einschätzung der Problemlöseaufgaben nicht zu beobachten. Im Mittel ist kein Unterschied zwischen den Domänen Problemlösen (*M* = 2.97, *SD* = 1.44) und Mathematik (*M* = 2.98, *SD* = 1.37) zu beobachten. Die eingeschätzte Herausforderung oder Anstrengungsbereitschaft ist für Mathematikaufgaben (*M* = 3.91, *SD* = 1.48) ähnlich hoch wie für Problemlöseaufgaben (*M* = 3.89, *SD* = 1.53).

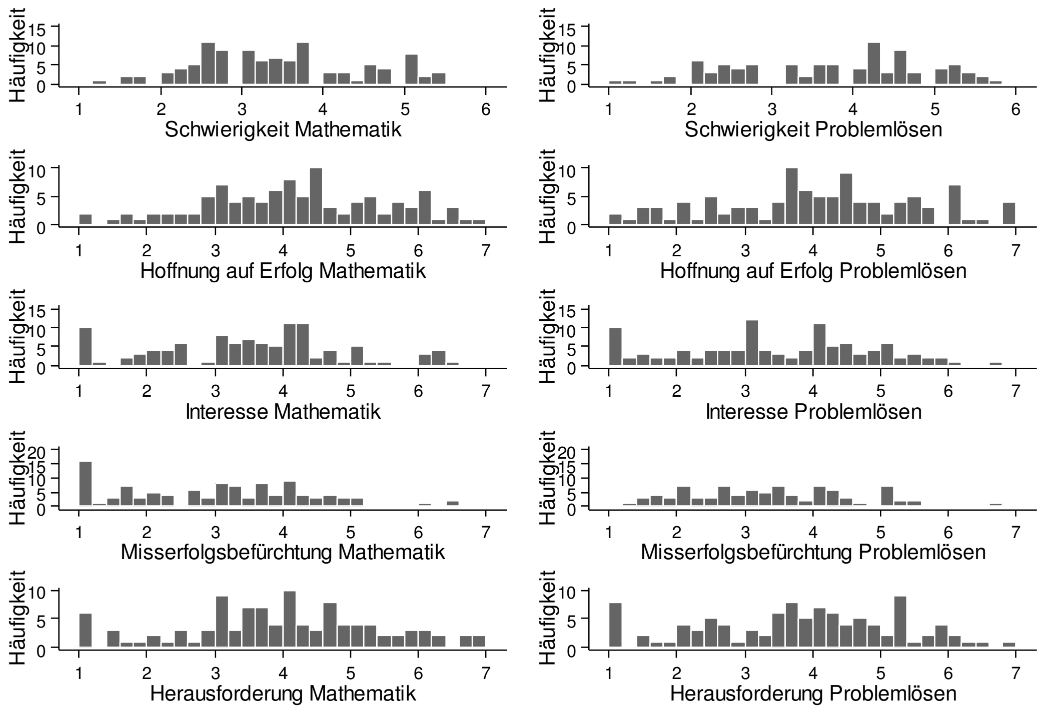


Abbildung A-1. Histogramme der mittleren Itemeinschätzungen getrennt nach Domäne

Anmerkungen. Schwierigkeitsskala: 1 = *sehr leicht*, bis 6 = *sehr schwer*; sonstige Skalen: 1 = *trifft nicht zu*, bis 7 = *trifft zu.*

Das Spinnennetzdiagramm (Abbildung A-2) fasst die Ähnlichkeit der Beurteilungen der Mathematik- und Problemlöseaufgaben auf Basis der mittleren Beurteilung gut zusammen. Die Domänen unterscheiden sich nicht in der mittleren Einschätzung von Schwierigkeit, Hoffnung auf Erfolg, Interesse, Misserfolgsbefürchtung und Herausforderung. Angesichts der deskriptiven Ergebnisse, die allen fünf Hypothesen widersprechen, erübrigt sich eine inferenzstatistische Prüfung der Ergebnisse.

Weitere Befunde: Bivariate Zusammenhänge

Da die univariate Analysen noch nicht berücksichtigen, dass jede Person sowohl Problemlöseaufgaben als auch Mathematikaufgaben beurteilt, betrachten wir zusätzlich bivariate Zusammenhänge. Während die univariaten Analysen eher für eine homogene Einschätzung der Aufgaben der beiden Domänen sprechen, zeigten die Streudiagramme (Abbildung A-3), dass die Zusammenhänge der mittleren Einschätzungen von Problemlöse- und Mathematikaufgaben im mittleren Größenbereich liegen – und zwar für alle fünf Items (Interesse, Herausforderung, Hoffnung auf Erfolg, Misserfolgsbefürchtung, Schwierigkeit). Es ist keine Häufung in den Quadranten erkennbar, die eine positivere Einschätzung der Problemlöseitems und negativere Einschätzung der Mathematikitems bedeuten.

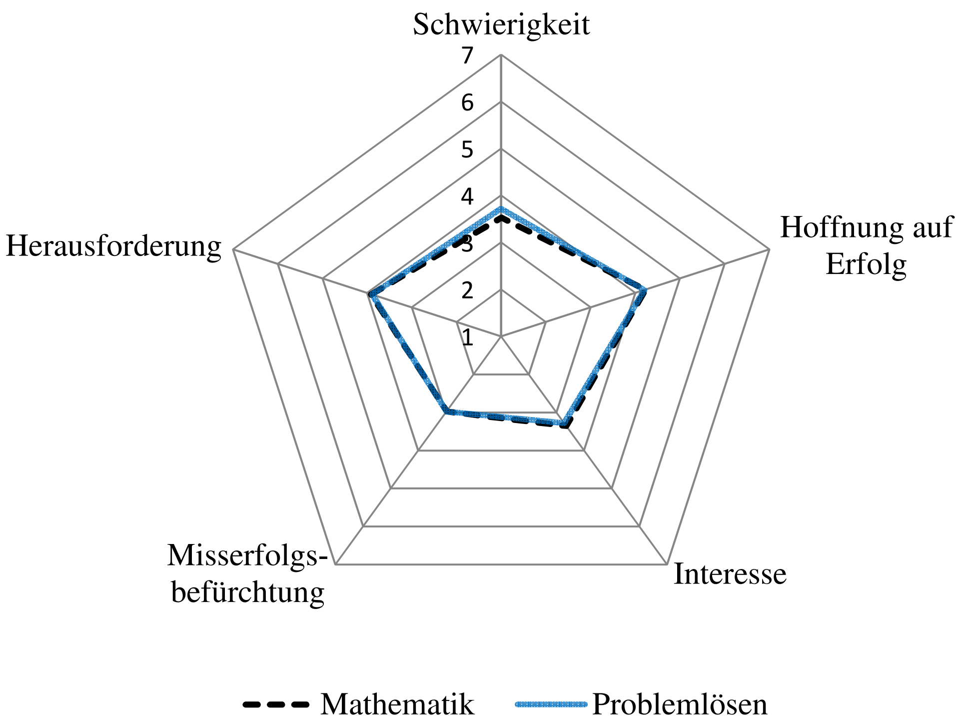


Abbildung A-2. Radar-Graph der mittleren Einschätzung der Mathematik- und Problemlöseaufgaben

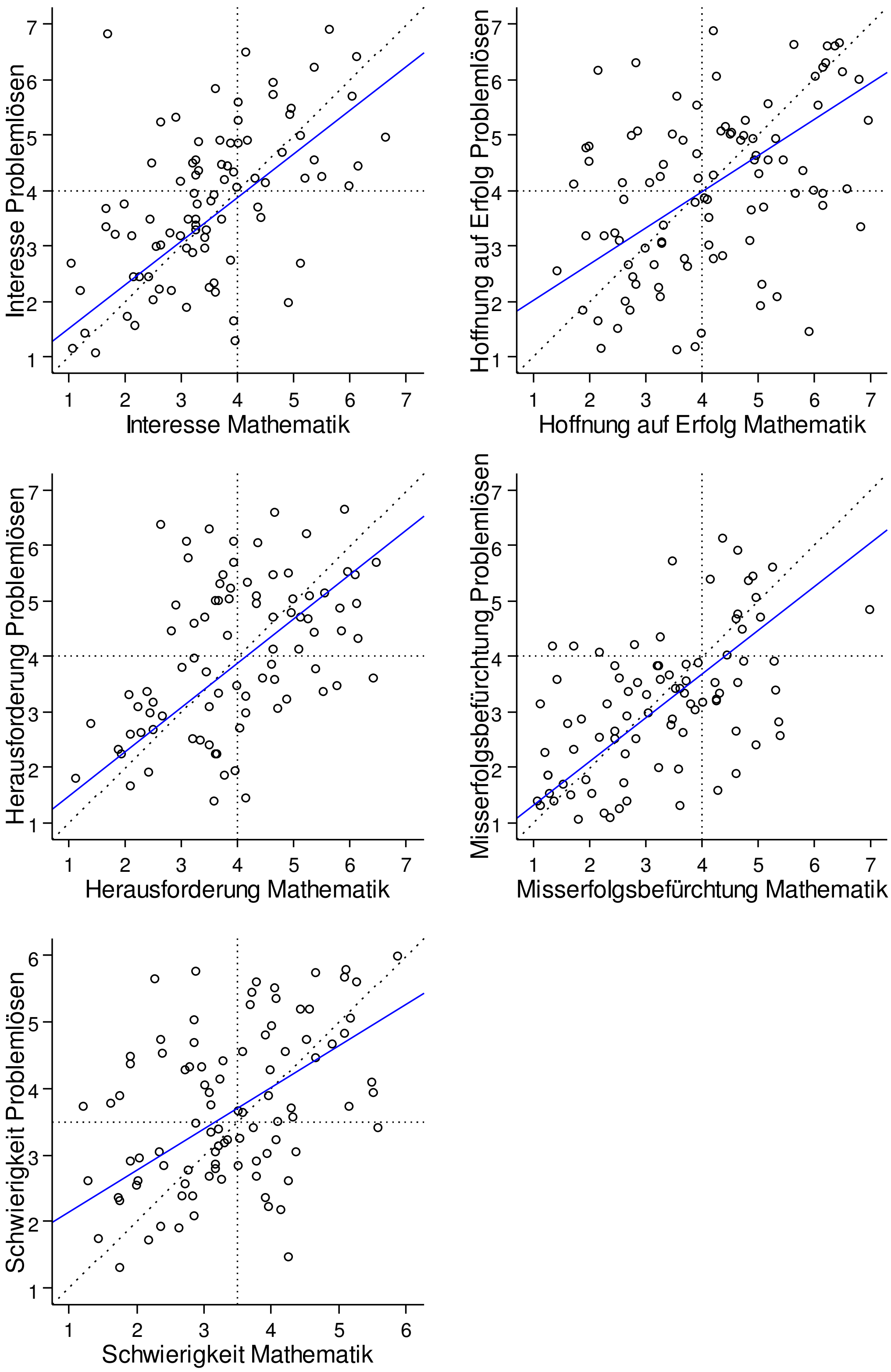


Abbildung A-3. Zusammenhänge der mittleren Einschätzung der Problemlöse- und Mathematikaufgaben bzgl. Interesse, Herausforderung, Hoffnung auf Erfolg, Misserfolgsbefürchtung und Schwierigkeit

Diskussion

Die verwendete Teilstichprobe von PISA-2003-Items wird über beide Domänen hinweg vergleichbar bzgl. der eingeschätzten Schwierigkeit, des Interesses, der Herausforderung bzw. Anstrengungsbereitschaft, und der Misserfolgsbefürchtung eingeschätzt. Diese Studie kann die Idee unterschiedlicher Motivationseffekte bei Problemlöse- und Mathematikaufgaben daher nicht belegen.

Entgegen der Hypothese, dass Mathematikaufgaben negativer beurteilt werden als Problemlöseaufgaben, zeigten die deskriptiven Ergebnisse dieser Stichprobe sogar, dass die Problemlöseaufgaben als etwas schwieriger beurteilt werden. Möglicherweise ist dies durch die höhere Textmenge der Problemlöseaufgaben (Fleischer et al., 2010) sowie dadurch zu erklären, dass die SuS die Aufgaben nur einschätzen und nicht lösen mussten. Wie präzise die Schülerurteile sind und wie sie mit Testleistungen in den Domänen korrelieren, kann in dieser Studie nicht geprüft werden, da die SuS die Aufgaben nur eingeschätzt haben.

Einschränkend muss man weiterhin beachten, dass die Daten weder der PISA-2003-Erhebung entstammen noch eine internationale Stichprobe zugrunde liegt, sondern hauptsächlich eine anfallende Stichprobe von SuS verschiedener Realschulen aus Nordrhein-Westfalen. Die Schülerstichprobe ist damit nicht repräsentativ. Ferner basiert die Itemauswahl bzw. das Matching der Itemschwierigkeiten der Problemlöse- und Mathematikitems auf der internationalen PISA-Metrik von 2003, deren Aktualität und Gültigkeit für die befragte Realschüler-Stichprobe zehn Jahre später zumindest fraglich ist. Zudem wurde nur eine Teilstichprobe der PISA-Items verwendet, die mit Ein-Item-Skalen eingeschätzt wurden, für die keine Reliabilität berechenbar ist.

Die univariaten Ergebnisse lassen vermuten, dass die SuS nicht zwischen den Problemlöse- und Mathematikaufgaben differenzieren. Im Gegensatz zu Fleischer et al. (2014), die die Aufgaben in unterschiedliche Testkontexte (Mathematik vs. Problemlösen) eingebettet haben, bleibt hier jedoch offen, ob die SuS die Aufgaben eher als Mathematik- bzw. Problemlöseaufgaben wahrgenommen haben, da die Instruktion bewusst allgemein gehalten wurde („Wir wollen wissen, wie Schülerinnen und Schüler bestimmte Aufgaben einschätzen.“). Die bivariaten Streudiagramme (Abbildung A-3) zeigten keine Häufung in den Quadranten, die eine positivere Einschätzung der Problemlöseitems und negativere Einschätzung der Mathematikitems bedeuten.

Insgesamt liefert diese Umfrage somit keine Hinweise dafür, dass bei PISA 2003 in Deutschland die Diskrepanz zwischen fächerübergreifender und mathematischer Problemlösekompetenz durch unterschiedliche motivationale Effekte in den Testdomänen entstanden sein könnte.

Anhang B. Verzweigte Lernumgebungen und Tests mit EFS Survey

Dieses Kapitel skizziert die technische Eignung der kommerziellen Online-Umfrage-Software EFS Survey (Globalpark AG, 2011; QuestBack AG, 2012) für die Anforderungen der computerbasierten Experimente 1 und 2. Es werden Vorteile und Einschränkungen von EFS Survey vorgestellt. Das Kapitel basiert auf einer Manuskriptversion[23] des folgenden Artikels:

> Buchwald, F., Spoden, C., Fleischer, J. & Leutner, D. (2013). Verzweigte Lernumgebungen und Tests mit EFS Survey 8. *Diagnostica, 59, 113-117.* DOI: 10.1026/0012-1924/a000080.

EFS Survey

EFS Survey (Globalpark AG, 2011) ist als kommerzielle Umfrage-Software prinzipiell in vielen Anwendungsbereichen nutzbar. An dieser Stelle wird aus pragmatischen Gründen ausschließlich auf Anwendungsszenarien aus dem psychologischen Bereich eingegangen, die ohne Programmierkenntnisse realisierbar sind. Zunächst wird EFS Survey 8.2, orientiert an den Kriterien der Produktsteckbriefe von *e-teaching.org* (Institut für Wissensmedien (IWM), 2011), vorgestellt. Anschließend wird die Erstellung verzweigter Lerneinheiten, des gleitenden Testfensters (Leutner, 1992b) sowie nicht-adaptiver, selbstadaptierter und verzweigter Tests skizziert.

Einsatzgebiet

EFS Survey erlaubt die Entwicklung webbasierter Umfragen, Experimente, Lernumgebungen und Tests. Der Einsatz von EFS Survey erfordert einen aktuellen Webbrowser und eine Internetverbindung.

Vorteile von EFS Survey

- *Großer Funktionsumfang*. EFS Survey gehört zu den mächtigeren Umfrage-Programmen, deren Funktionsumfang auch Ansprüchen wissenschaftlicher Forschung (u. a. Randomisierung, *data tracking, routing*, Filter und Trigger, anpassbares Layout) gerecht wird.

[23] Diese Artikelfassung entspricht nicht vollständig dem in der Zeitschrift veröffentlichten Artikel. Dies ist nicht die Originalversion des Artikels und kann daher nicht zur Zitierung herangezogen werden.

- *Geringe Einstiegshürde für Anwender*. Die Einstiegshürde zur Benutzung der graphischen Oberfläche ist gering, auch wenn manche Optionen etwas versteckt sind. Zur Einführung gibt es kurze Video-Tutorials. Während des Arbeitens mit EFS Survey sind über ein Fragezeichensymbol einblendbare Hinweise nützlich. Zum Lernen und Nachschlagen eignet sich die mit zahlreichen Screenshots versehene Dokumentation.
- *Programmieren: Nicht nötig, aber möglich.* Seiten, Fragen, Verzweigungen, Sprungstellen und Abbruchkriterien können im Webbrowser über die graphische Benutzeroberfläche erstellt werden. Das Verwenden von HTML-Code, eigenen Flash-Aufgaben, JavaScript und LUA-Code ist möglich, für viele Zwecke aber nicht nötig.
- *Items und Itemformate*. EFS Survey verfügt neben gebundenen Aufgabenformaten – wie *single choice* und *multiple choice*, Matrix-Fragen, semantischem Differential, Ranking-Aufgaben – über offene Antwortformate in Form von Freitextfeldern mit festlegbarer Größe. Dabei sind sowohl beim Aufgabenstamm als auch bei den Antwortoptionen Formatierungen sowie Abbildungen in den gängigen Grafik-Formaten (u. a. GIF, PNG, JPEG, TIFF) möglich. Darüber hinaus ist der Im- und Export von Seiten und Items zwischen verschiedenen mit EFS Survey erstellten Projekten möglich.
- *Einfacher Datenexport*. Die User-Daten, d. h. Nutzereingaben und Logfiles, können in Standardformate (z. B. XLS, SAV) exportiert werden, die mit verbreiteter Statistik-Software (z. B. R, MS Excel, IBM SPSS Statistics) weiterverarbeitet werden können. Die EFS-Survey-Projektdateien können zur Archivierung und Wiederverwendung im XML-Format exportiert werden.

Grenzen von EFS Survey

EFS Survey ist ursprünglich eine Software für Online-Umfragen. Daraus ergeben sich für Zwecke des Lernens und Testens eine Reihe von „natürlichen“ Grenzen. Auf einige relevante Beschränkungen soll kurz hingewiesen werden:

- *Keine Offline-Nutzung*. Da es sich bei EFS Survey um eine Online-Umfrage-Software handelt, bedarf es bei den Standard-Lizenzmodellen einer bestehenden Internetverbindung.

- *Eingeschränkte Multimedia-Plattform.* Auch wenn multimediale Lehr-Lernumgebungen realisiert werden können, ist EFS Survey insbesondere für Szenarien mit vielen multimedialen Inhalten nur eingeschränkt geeignet, da die Speicherkapazität für multimediale Inhalte stark beschränkt ist. Durch das Einbetten von Videos und anderer Medien über einfachen HTML-Code kann man diese Einschränkung jedoch leicht umgehen. Für das Lehren und Lernen bzw. Testen formellastiger Domänen ist EFS Survey nur bedingt geeignet, da kein eigener Formeleditor existiert.
- *Keine IRT-Skalierung.* Da EFS Survey keine Schätzalgorithmen unterstützt, ermöglicht EFS Survey kein *tailored testing*.
- *Keine Reaktionszeiten.* EFS Survey ist nicht dazu konzipiert, präzise Bearbeitungs- bzw. Reaktionszeiten zu erfassen.

Beispiele psychologischer Anwendungsszenarien

Im Folgenden wird die Erstellung verzweigter Lerneinheiten und Tests mit EFS Survey skizziert.

Lerneinheiten mit EFS Survey

Mit EFS Survey sind klassische, nicht-verzweigte, multimediale Lerneinheiten realisierbar (vgl. z. B. Rey & Buchwald, 2011). Im Folgenden fokussieren wir auf die Möglichkeiten, die sich durch die Verwendungen von verzweigten Programmabläufen ergeben. In EFS Survey können sich die Bedingungen für Verzweigungen auf beliebige (im Projekt angelegte) Variablen beziehen.

Verzweigte Lerneinheiten. EFS Survey ermöglicht es, ohne Programmierkenntnisse bedingte Sprünge zu verwenden (Abbildung B-1), wodurch Verzweigungen möglich sind. Bedingte Sprünge sind hier Regeln, die angeben, unter welcher Bedingung welcher Inhalt dem Lernenden präsentiert wird.

a)

b)

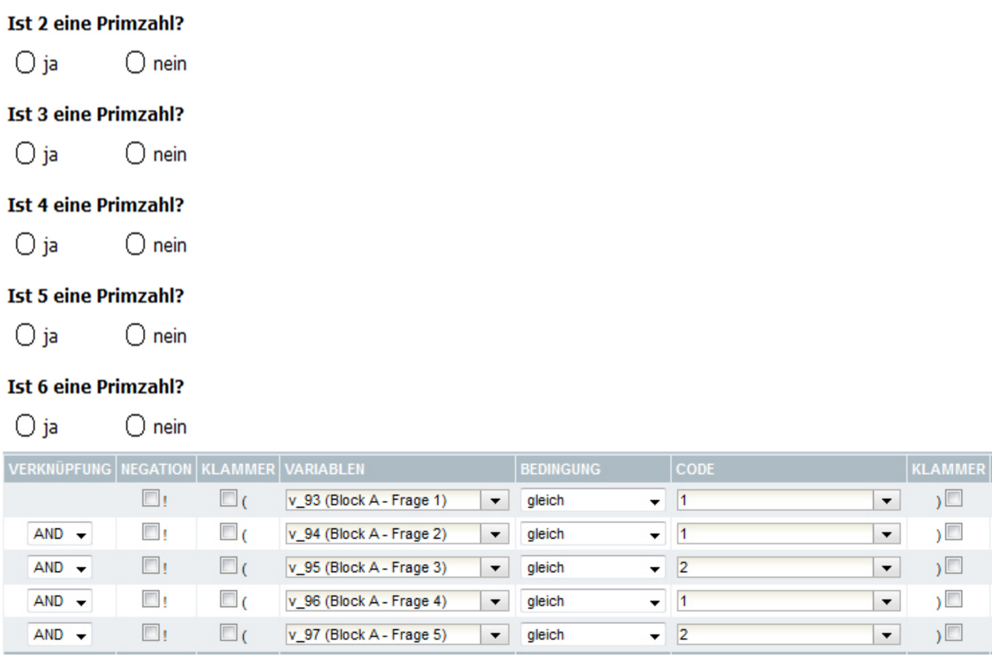


Abbildung B-1. Fiktives Beispiel für die Erstellung einer Sprungbedingung

Anmerkungen. Um zu prüfen, ob die fünf Fragen (a) alle korrekt beantwortet wurden, muss eine Bedingung (b) erstellt werden (Code: 1 = *ja*, 2 = *nein*). Außerdem muss später das Sprungziel, d. h. die gewünschte Zielseite, angegeben werden. Dass die Variablennummerierung hier mit v_93 beginnt, hat keine Bedeutung. Eine Eingabe der Bedingung als Code-Zeile ist auch möglich:
if(v_93 = 1 AND v_94 = 1 AND v_95 = 2 AND v_96 = 1 AND v_97 = 2).

Die folgende Aufzählung gibt Beispiele für kombinierbare Anwendungsmöglichkeiten, die mittels bedingter Sprünge möglich sind:

- *Gezieltes Üben*. Als diagnostische Basis für die Vorgabe neuen Materials in Lerneinheiten wird meist der Erfolg bzw. Misserfolg beim Lösen eines oder mehrerer zuvor bearbeiteter Items genutzt (Leutner, 2004; Rey, 2009). In Abhängigkeit von gegebenen Antworten oder Lösungen werden bestimmte Inhalte (Lerneinheiten oder Aufgaben) vorgegeben oder übersprungen.

- *Bekämpfung von Fehlvorstellungen*. Verwendet man als Bedingung die Auswahl bestimmter falscher Antwortoptionen, die spezifische Fehlvorstellungen indizieren, so kann man diese Fehlvorstellungen gezielt thematisieren. Dabei können wieder bedingte Sprünge zum Einsatz kommen. Alternativ können Filter verwendet werden, die fehlerspezifisches Material, falls nötig, anzeigen oder andernfalls überspringen.

- *Differenziertes Feedback*. Der Lernende erhält differenziertes Feedback zu seinen Aufgabenlösungen. Dies kann pro Item oder pro Aufgabenblock geschehen. Je nach Antwortverhalten wird individuelles Feedback präsentiert.
- *Selbstgewählte Bearbeitungsreihenfolge*. Der Lernende kann selbst entscheiden, welche Teile des Programms er in welcher Reihenfolge bearbeitet. Um dieses Szenario zu realisieren, muss der Lernende gefragt werden, mit welchem Thema er fortfahren möchte. In Abhängigkeit von der gegebenen Antwort wird dann auf das gewünschte Thema weitergeleitet.
- *Selbstgewählte Vertiefungen*. Der Lernende kann nicht nur die Reihenfolge der Materialbearbeitung auswählen, sondern stattdessen oder zusätzlich Vertiefungsthemen wählen.

Auf diese Weise kann ein Lernprogramm erstellt werden, das dem Konzept der Binnendifferenzierung genügt, in dem Lernende Inhalte, Aufgaben und Feedback erhalten, die dem individuellen Leistungsniveau angepasst sind.

Lernerfolgskontrolle mit dem gleitenden Testfenster.[24] Das gleitende Testfenster (Leutner, 1992b, 1993) ist ein einfaches Entscheidungsverfahren, mit dem bestimmt werden kann, ob ein bestimmter Lehrstoff mit einer bestimmten Sicherheit beherrscht wird. Es basiert auf der Analyse der jeweils letzten *x* bearbeiteten Items eines Blocks A, bestehend aus y Übungsaufgaben (Leutner, 1993). Sobald eine bestimmte Anzahl *a* der *x* zuletzt bearbeiteten Items richtig gelöst worden ist, wird der Block A beendet und der Lernende gelangt zum nächsten Aufgabenblock B oder zum Programmende. Um die Prozedur mit Hilfe eines Computerprogramms zu realisieren, bedarf es, konzeptuell betrachtet, nur einer Reihe von verzweigten Seiten und des Zugriffs auf die zwischengespeicherten Antworten der Lernenden. Liegt beispielsweise das Kriterium „a = 5, x = 5“ vor, dann lautet die konkrete Regel: Falls der Lernende fünf Aufgaben nacheinander richtig gelöst hat, dann gehe weiter zu Block B (vgl. Leutner, 1993). Um diese Regel zu implementieren, muss nach dem fünften Item zum ersten Mal geprüft werden, ob alle fünf vorhergehenden Items korrekt beantwortet worden sind. Abbildung B-2 zeigt exemplarisch die graphische Oberfläche, mit deren Hilfe der benötigte Code erzeugt wird, um ein gleitendes Testfenster mit EFS Survey zu erstellen.

[24] Diese Prozedur war im DFG-Antrag für Experiment 1 vorgesehen. Eine Begründung, warum sie nicht zum Einsatz kommt, ist in Anhang C zu finden.

Testen mit EFS Survey

Im Folgenden werden Beispiele verschiedener Testarten skizziert, die mit EFS Survey umsetzbar sind.

Nicht-adaptive Tests. Klassische nicht-adaptive psychologische Tests sind mit EFS Survey im Prinzip realisierbar, sofern sie sich in eine multimediale Abfolge von HTML-Seiten bringen lassen und das Aufgabenformat von EFS Survey unterstützt wird. Dies ist etwa bei Persönlichkeitstests der Fall, da diese strukturell meist die Form von Fragebögen haben. Aber auch Leistungstests wie Matrizentests sind im Prinzip implementierbar, indem zusätzlich Abbildungen verwendet werden.

Selbstadaptierte Tests (Rocklin, 1994). Diese Tests, bei denen der Proband die Schwierigkeit der nächsten Aufgabe selbst auswählt, sind über bedingte Sprünge ebenfalls in EFS Survey realisierbar. Dazu muss zunächst eine Frage nach der gewünschten Schwierigkeit der nächsten Aufgabe angegeben werden. Jede Antwortmöglichkeit, beispielsweise „einfach“, „mittel“ und „schwierig“ wird dann als Bedingung für den weiteren Programmverlauf verwendet.

Fest-verzweigte Tests. Aus Überlegungen zur Messeffizienz, Präzision und diskriminanten Validität kann es wichtig sein, den weiteren Verlauf eines individualdiagnostischen Leistungstests in Abhängigkeit von bestimmten Kriterien adaptiv vorzugeben (Bühner, 2010; Frey, 2012). Im Rahmen des adaptiven Testens unterscheidet man *branched testing* und *tailored testing*. Beim *branched testing* wird vor Testbeginn festgelegt, welche Items bei welchem Antwortverhalten vorgelegt werden, während beim *tailored testing* diese Entscheidung erst während der Testung erfolgt (van der Linden & Glas, 2000, 2010). EFS Survey ermöglicht kein tailored testing, da keine Schätzalgorithmen zur Echtzeitschätzung von Fähigkeitsparametern zur Verfügung stehen.[25] Sofern ein geeignet skalierter Itempool für *branched testing* zur Verfügung steht, sind fest verzweigte Tests mit zwei- oder mehrstufiger Strategie hingegen in EFS Survey über bedingte Sprünge bzw. Filter implementierbar. Dabei muss eine Zählvariable angelegt werden, die die erreichten Rohpunkte zählt (siehe Abbildung B-2). In Abhängigkeit von der erreichten Rohpunktzahl, die vom Testautor einer bestimmten Leistungsstufe zuge-

[25] Für webbasiertes *tailored testing* sei auf die kommerzielle CAT-Software *FastTEST Web* (FastTEST Web (c/o ASC), 2011) verwiesen. Diese verfügt beispielsweise auch über einen integrierten Formeleditor.

ordnet wird, werden dann auf der nächsten Stufe unterschiedlich schwierige Items vorgegeben.

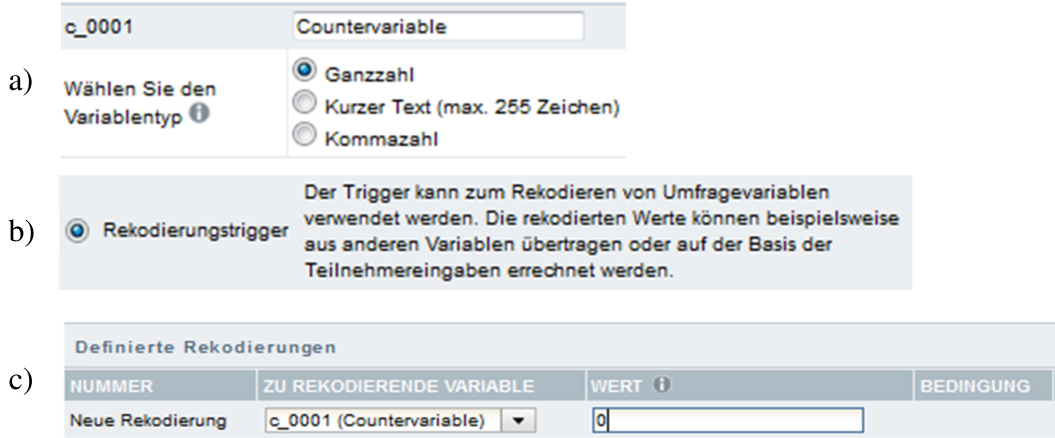


Abbildung B-2. Anlegen einer Zählvariablen in EFS Survey

Anmerkungen. Zunächst muss unter „Projekteigenschaften“ eine „benutzer-definierte Variable“ mit dem Typ „Ganzzahl“ angelegt werden (a). Anschließend wird auf einer Seite ein „Rekodierungstrigger“ angelegt (b). Dabei definiert man die Zählvariable (Countervariable), hier c_0001, zunächst als Nullvektor (c). Die Aufdatierung der Countervariable erfolgt über bedingte Rekodierungstrigger auf weiteren Seiten. Mit der Eingabe #c_001# +1 in das Feld „Wert“ erhöht man den Wert der Variable c_001 um 1. Damit die Anzahl der richtig gelösten Items gezählt wird, darf der Rekodierungstrigger nur dann ausgeführt werden, wenn ein Item richtig gelöst wurde. Um dies zu erreichen, wird eine Bedingung für die Ausführung des Rekodierungstriggers angelegt (vgl. Abbildung B-1). Die so angelegte Countervariable kann das Ziel von Filtern und Sprungbefehlen sein.

Open-Source-Alternativen zu EFS Survey

Mit *LimeSurvey* (Schmitz et al., 2012), dem früheren PHP Surveyor, steht ein Open-Source-Umfragetool zur Verfügung. Dieses Tool erreicht allerdings in der aktuellen stabilen Version 1.92+ noch nicht die Benutzerfreundlichkeit und die Qualität der Dokumentation von EFS Survey. LimeSurvey kann mit Hilfe lokaler Server im Gegensatz zu EFS Survey auch offline genutzt werden.

Zur Erstellung komplexer und dynamischer Testaufgaben eignet sich die on- und offline verwendbare Open-Source-Plattform *TAO* (CRP Henri Tudor & University of Luxembourg, 2011).

Anhang C. Experiment 1: Begründung der Nicht-Verwendung des gleitenden Testfensters

Im DFG-Antrag (Leutner et al., 2009) war die Verwendung des gleitenden Testfensters (Leutner, 1993) für das Experiment 1 vorgesehen. Das gleitende Testfenster (Leutner, 1993) erscheint jedoch aus folgenden Gründen für dieses Experiment ungeeignet:

1. Berechnungen der Größe des benötigten Itemuniversums basierend auf einer Umformung der Formel zur Berechnung der Testlänge eines kriteriumsorientierten Tests (Klauer, 1987, siehe Tabelle C-1) mit dem in Tabelle C-2 dokumentierten R-Code zeigten, dass wesentlich mehr Items benötigt würden, als zur Verfügung stehen und mit vertretbarem Aufwand generierbar und zu pilotieren wären.
2. Das zugrunde liegende allgemeine Binomialmodell geht von gleichen Lösungswahrscheinlichkeiten der Items aus (Klauer, 1987), was z. B. für die Problemlöseaufgaben nicht gegeben ist. Man beachte jedoch, dass das verallgemeinerte Binomialmodell unterschiedliche Lösungswahrscheinlichkeiten berücksichtigen kann (Klauer, 1987).
3. Es ist unklar, wie sich die zweite Chance der Aufgabenbearbeitung, die die SuS erhalten sollen, auf die Lösungswahrscheinlichkeiten genau auswirkt.
4. Das experimentelle Design erfordert einen Posttest, der für Experimental- und Kontrollgruppe gleich ist. Dadurch wird die Trennung von Übungs- und Testaufgaben, der Vorteil des gleitenden Testfensters, teilweise aufgehoben.

Das gleitende Testfenster wird daher für das Experiment 1 nicht verwendet.

Tabelle C-1. Herleitung der Formel zur Berechnung der Größe des benötigten Itemuniversums eines kriteriumsorientierten Tests

Nach Klauer (1987), der auf Berk (1980) zurückgreift, berechnet sich die Länge N^* eines kriteriumsorientierten Test nach der Formel

$$N^* = \frac{\frac{z^2\pi_z(1-\pi_z)}{e^2}}{1+\frac{1}{N}\left(\frac{z^2\pi_z(1-\pi_z)}{e^2}-1\right)}$$

wobei

z = z-Wert des Konfidenzintervalls (z. B. z = 1.96 für 95 %, 2-seitig)
π = geforderter Kompetenzgrad (z. B. 0.9)
e = Fehlerschranke um π_z
N = Größe der Grundgesamtheit der Aufgaben

Setze $y \coloneqq \frac{z^2\pi_z(1-\pi_z)}{e^2}$. Mit dieser Schreibweise gilt für die Länge N^* eines kriteriumsorientierten Test: $N^* = \frac{y}{1+\frac{1}{N}(y-1)}$.

Wir wollen die Formel im Folgenden nach *N*, d. h. nach der Größe der benötigten Grundgesamtheit der Aufgaben auflösen:

$$N^*\left(1+\frac{1}{N}(y-1)\right) = y \Leftrightarrow 1+\frac{1}{N}(y-1) = \frac{y}{N^*} \Leftrightarrow \frac{1}{N}(y-1) = \frac{y}{N^*}-1$$

$$\Leftrightarrow \frac{1}{N} = \frac{\frac{y}{N^*}-1}{y-1} \Leftrightarrow N = \frac{y-1}{\frac{y}{N^*}-1}$$

Wir ersetzen jetzt wieder *y*:

$$N = \frac{\frac{z^2\pi_z(1-\pi_z)}{e^2}-1}{\frac{\frac{z^2\pi_z(1-\pi_z)}{e^2}}{N^*}-1}$$

Diese Formel gibt *N*, d. h. die Größe der benötigten Grundgesamtheit der Aufgaben, in Abhängigkeit von z, π, e, N^* aus.

Tabelle C-2. R-Code zur Berechnung der Größe des benötigten Itemuniversums eines kriteriumsorientierten Tests

```
# Die nachfolgende R-Funktion universe_size liefert bei Eingabe
      # der Länge n des benötigten Tests (in Items)
  # und der Größe des Konfidenzintervalls (z. B. k = 95 %)
  # und des geforderten Kompetenzgrades (z. B. piz = 0.9)
  # und der akzeptablen Fehlerschranke um den geforderten
       Kompetenzgrad (z. B. e = 0.1)
      # die Größe der benötigten Grundgesamtheit N
       (Itemuniversum)eines kriteriumsorientierten Tests.
# Ausgangspunkt: Klauer (1987)
# R-Programmidee: Florian Buchwald
# R-Implementierung: Florian Buchwald, Volker Zischka
# Version 1.0, Mai 2011

# Input: n = Länge des benötigten Tests
# Input: piz = geforderter Kompetenzgrad (z. B. 0.9)
# Input: k = Prozentwert zum z-Wert
# Input: e = Fehlerschranke um den geforderten Kompetenzgrad piz
# Output: N = Größe der benötigten Grundgesamtheit des
       kriteriumsorientierten Tests (Itemuniversum)


universe_size <- function (n, piz, e, k)
{
# Berechnung des z-Wertes zum gewünschten Konfidenzintervall
# zweiseitige Tests
 z = qnorm(0.5*(1+k))
# Einführung einer Hilfsvariable h
 h = z*z*piz*(1-piz)/(e*e)
# Berechne die benötigte Grundgesamtheit N
 N = (h-1)/((h/n)-1)
#Rückgabe des Ergebnisfeldes
 return(N)
}
```

Anhang D. Experiment 1: Paarvergleiche der Skalen zum konditionalen Wissen und zur Planungsfähigkeit im Posttest

Beim konditionalen Wissen wird von der Notenskala (1 = *sehr gut*; bis 6 = *ungenügend*) ausgegangen, d. h. z. B. c > a bedeutet, dass die Experten Option a für besser als Option c einschätzen. Beim konditionalen Wissen bedeutet C > A, dass gemäß dem Expertenurteil Teilschritt C nach Teilschritt A auszuführen ist.

Tabelle D-1. Paarvergleiche Konditionales Wissen Mathematik

Item	*Verwendete Paarvergleiche*	*Anzahl Paarvergleiche*
KW_MA_x1 (Szenario 1 von Artelt)	c>a, c>e, c>d, b>e, b>d	5
KW_MA_x4 (Szenario 5 von Artelt)	a>e, a>c, b>e, b>c, d>e, d>c	6
KW_MA_x5 (Szenario 6 von Artelt)	c>a, c>b, d>a, d>b, e>a, e>b	6

Tabelle D-2. Paarvergleiche Fächerübergreifendes Konditionales Wissen

Item	*Verwendete Paarvergleiche*	*Anzahl Paarvergleiche*
KW_PL_x1 (Stillarbeit)	a<b, a<c, b>d, b>e, c>d, c>e	6
KW_PL_x3 (Mannschaftsführer)	a>b, a>c, b<d, c<d	4
KW_PL_x4 (Musik)	a>b, a<c, a>d, a>e, b<c, b>e, c>d, c>e, d>e	9
KW_PL_x6 (Buchausleihe)	a<b, a<c, b>c, b>d, c>d	5
KW_PL_x7 (Werbung)	a>b, a>c, a>d, b<c, b<d	5

Tabelle D-3. Paarvergleiche Planungsfähigkeit Mathematik

Item	*Verwendete Paarvergleiche*	*Anzahl Paarvergleiche*
SE_*M*_X3 (Dreieckskonstruktion)	F<A, F<B, F<C,F<D, F<E, D<B, D<A, D<C,D<E, B<C, B<A, B<E, A<E, C<E	14
SE_*M*_X5 (Oberfläche)	F<A, F<B, F<C, F<D, F<E, F<G, G<E, G<D,G<A, G<B, C<E, C<D, C<A, C<B, E<D, E<A, E<B, D<A, D<B, A<B	20
SE_*M*_X6 (Zahl)	B<A, B<C, B<D, B<E, B<F, B<G, C<A, C<E, C<D, C<F, G<A, G<E, G<D, G<F,A<E, A<D, A<F, E<D, E<F, D<F	20

Tabelle D-4. Paarvergleiche Fächerübergreifende Planungsfähigkeit

Item	*Verwendete Paarvergleiche*	*Anzahl Paarvergleiche*
SE_PL_X4 (Kühlschrank)	A>B, A<C, A<D, A<E, A>F, B<C, B<D, B<E, B>F, C>D, C>E, C>F, D>F, E>F	14
SE_PL_X5 (Kuchen backen)	A>B, A>C, A>D, A>E, A<F, A<G, A>H, B<C, B>E, B<F, B<G, C>D, C>E, C<F, C<G, C>H, D>E, D<F, D<G, D<H, E<F, E<G, E<H, F>H, G>H	25
SE_PL_X7 (Unterschriftenaktion)	A>B, A>C, A>D, A>E, A>F, B>D, B>E, B<F, C<B, C<D, C<E, D<E, D<F, E<F	14
SE_PL_X8 (Schul-AG)	A>B, A>C, A>D, A>E, A>F, B<D, B>E, B>F, C<D, C>E, C>F, D>E, D>F	13

Anhang E. Experiment 1: Experimentalgruppentreatment

Die Screenshots wurden aus Druckgründen auf eine Seitenbreite von 14 cm skaliert, was ungefähr 2/3 der Originalgröße entspricht. Aus Platzgründen kennzeichnet ein schwarzer Rahmen jeweils eine Bildschirmseite, damit mehrere Screenshots auf einer Seite dokumentiert werden können.

Übersicht

Herzlich willkommen zu den *Essener Selbstlernmodulen zum problemlösenden Denken* (E-SePro).

E-SePro gibt dir anhand zahlreicher Beispiele und Aufgaben einen Einblick in das spannende Feld des Problemlösens.

E-SePro besteht derzeit aus folgenden Teilen, die du der Reihe nach durchlaufen wirst.

Teil 1

- Einführung in das Thema Problemlösen
- Begriffsklärungen
- Gut geplant ist halb gelöst
- Wann mache ich was?
- Umgang mit Text, Abbildungen und Tabellen

Teil 2

- Übungsaufgaben

Was ist ein Problem?

Ein Problem besteht aus
- einer *Ausgangssituation*,
- einem *Ziel*,
- und einem *nicht direkt offensichtlichen Lösungsweg*

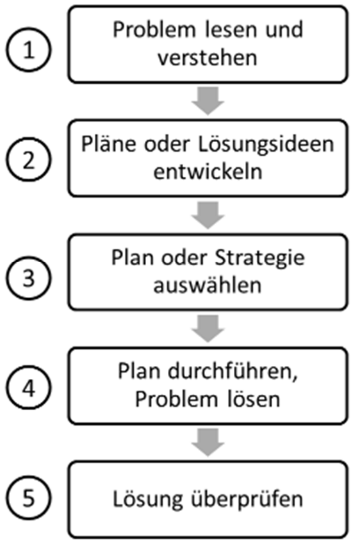
Beim Problemlösen geht es also darum, ein Ziel zu erreichen. Man weiß jedoch nicht sofort, wie man dorthin kommt.

An dieser Stelle war ein Comic zusehen (http://static.nichtlustig.de/toondb/031016.html).

Weiter

Problemlösen, ein Prozess

- Problemlösen lässt sich als Prozess darstellen.
- Bitte sieh dir dazu die nebenstehende Abbildung gut an
- Jede der fünf Phasen des Problemlösens ist wichtig, wenn man ein Problem lösen möchte.
- Im Verlaufe von E-SePro wirst du dich mit den einzelnen Phasen des Problemlösens beschäftigen.
  - Dazu gibt es eine ganze Reihe verschiedener Aufgaben.



Wann ist ein Problem gelöst?

Ein Problem ist *gelöst*, wenn das vorgegebene Ziel erreicht ist.

<u>Beispiel:</u>

- Wenn du zum ersten Mal dein kaputtes Fahrrad ohne Hilfe reparieren möchtest, ist das für dich wahrscheinlich ein Problem.
- Dieses Problem ist gelöst, wenn du dein Fahrrad alleine wieder in einen fahrtüchtigen Zustand versetzt hast.

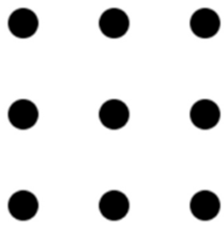
Im Rahmen dieses Lernprogramms üben wir den Umgang mit bestimmten Problemen. Diese Probleme haben meist einen Bezug zu Alltagssituationen oder sind allgemeine Knobelaufgaben.

Auf der nächsten Seite beginnen wir mit einer kleinen Knobelaufgabe.

Das Neun-Punkte-Problem

Aus diesem Beispiel kannst du Einiges über Probleme und Problemlösungen lernen. Man kann es so formulieren:

Versuche, die neun Punkte in nebenstehender Abbildung durch maximal vier gerade Linien zu verbinden, ohne dabei den Stift abzusetzen und eine Linie mehrfach zu ziehen.



Kennst du das Neun-Punkte-Problem bereits?

○ ja ○ nein

Glaubst du, dass man die neun Punkte auf die gewünschte Art verbinden kann?

○ ja ○ nein

Begründe bitte kurz, warum du glaubst bzw. nicht glaubst, dass man die neun Punkte wie gefordert verbinden kann.

Ausgangssituation, Mittel, Ziele

Neben dem sprachlichen Verständnis der Aufgabenstellung muss man versuchen zu verstehen,

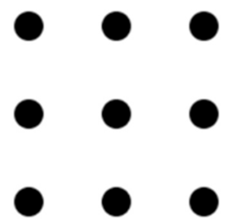
- worin, das Problem besteht,
- wie das Ziel bzw. die Lösung aussieht,
- welche Schritte zur Lösung erlaubt bzw. verboten sind.

Ohne dieses Verständnis ist es schwierig, ein Problem zu lösen.

Man muss sich also bewusst machen, wo man steht, wo man hin möchte und wie man dorthin kommt. Dazu ist es wichtig die Situation, die Regeln und Bedingungen sowie das Ziel zu kennen!

Daher üben wir jetzt die Bestimmung dieser Bestandteile eines Problems. Betrachte deshalb erneut das Neun-Punkte-Problem.

Versuche, die neun Punkte in nebenstehender Abbildung durch maximal vier gerade Linien zu verbinden, ohne dabei den Stift abzusetzen und eine Linie mehrfach zu ziehen.



Denke kurz über folgende Fragen nach:

- Worin besteht die Ausgangssituation?
- Was sind die Regeln und Einschränkungen?
- Was ist das Ziel?

Entscheide, um welches Element eines Problems es sich bei den folgenden Bestandteilen des Neun-Punkte-Problems jeweils handelt.

	Ausgangssituation	Regel	Ziel
Die Anordnung der neun Punkte.	○	○	○
Das durchgehende Zeichnen ohne den Stift abzusetzen.	○	○	○
Die Maximalzahl von vier geraden Linien.	○	○	○
Das Verbot des mehrfachen Abfahrens einer Linie.	○	○	○

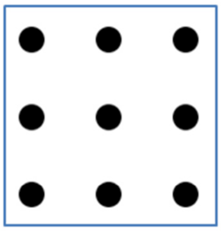
Regeln

Viele Menschen, die versuchen, das Neun-Punkte-Problem zu lösen, legen sich zusätzliche Beschränkungen auf, die sie am Finden einer Lösung hindern.

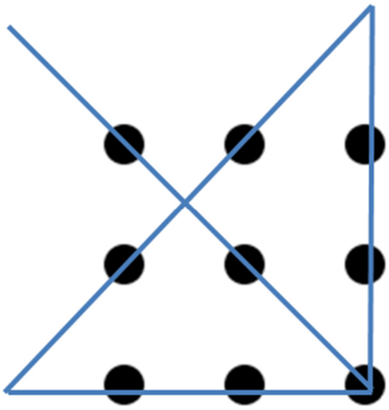
An dieser Stelle war ein Comic zusehen (http://static.nichtlustig.de/toondb/110914.html).

Selbst auferlegte Regeln

Viele Menschen, die versuchen das Neun-Punkte-Problem zu lösen, denken, sie dürften eine selbst auferlegte Box nicht überschreiten. Das wird in der Aufgabenstellung jedoch nicht gefordert.



Erst wenn man diese selbst auferlegte Regel verwirft, kann man eine Lösung des Neun-Punkte-Problems finden:



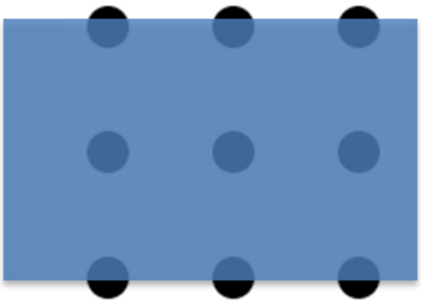
, bist du auf diese Lösung gekommen?

○ ja ○ nein

Selbst auferlegte Regeln

Jetzt kennen wir schon eine Lösung für das Neun-Punkte-Problem. Es gibt aber noch andere Lösungen.

Beispielsweise steht in der Aufgabenstellung nicht, wie breit der Stift sein darf. Daher kann man das Neun-Punkte-Problem sogar mit einer ganz, ganz dicken Linie lösen!



, bist du auf diese Lösung gekommen?

○ ja ○ nein

Zwischenfazit

Um ein Problem erfolgreich zu lösen, muss man sich zumindest klar machen,

- welche Ausgangssituation vorliegt und
- wie man diese Situation verändern kann und darf,
- um ein zuvor bestimmtes Ziel zu erreichen.

Block: Sachwissen

Problemlösen, ein Prozess

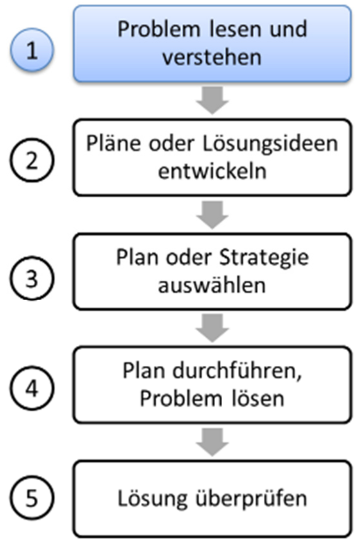
Erinnerst du dich an den Problemlöseprozess?

Das Verständnis der Begriffe steht am Anfang. Es ist eine Voraussetzung um ein Problem zu verstehen.

Dies geschieht in Phase 2 des Problemlöseprozesses (siehe Abbildung).

Um ein Problem zu lösen, muss man es verstehen! Das beginnt damit, dass man die Sprache verstehen muss, in der es formuliert ist.

Wer eine Aufgabe lösen möchte, muss zudem meist die Begriffe verstehen, die verwendet werden.



Damit du gute Chancen hast, die hier vorgestellten Probleme zu verstehen, findest du auf den nächsten Seiten zunächst ein kleines Quiz. Es enthält Begriffe aus den anschließend folgenden Aufgaben und Problemen.

Was versteht man bei Kinofilmen unter "Altersbegrenzung" oder "Altersfreigabe"?
Kreuze die richtige Antwort an.

- ○ Für Personen, die älter als die angegebene Altersgrenze sind, ist der Kinoeintritt kostenlos.
- ○ Für Personen, die jünger als die angegebene Altersgrenze sind, ist der Film nicht geeignet.
- ○ Personen, die älter als die angegebene Altersgrenze sind, finden den Film wahrscheinlich langweilig.
- ○ Für Personen, die jünger als die angegebene Altersgrenze sind, steht nur eine begrenzte Anzahl Kinokarten zur Verfügung.

Was bedeutet bei einem Kinoprogramm die Angabe "nur Mo-Fr"?
Kreuze die richtige Antwort an.

- ○ Der Film wird nur Montag und Freitag gezeigt.
- ○ Der Film wird nur Montag oder Freitag gezeigt.
- ○ Der Film wird an allen Tagen gezeigt nur nicht Montag bis Freitag.
- ○ Der Film wird nur Montag bis Freitag gezeigt.

Wie viele Minuten sind 1,5 Stunden?
Kreuze die richtige Antwort an.

- ○ 75 Minuten.
- ○ 90 Minuten.
- ○ 120 Minuten.
- ○ 150 Minuten.

Wie viele Minuten sind 2 Stunden?
Kreuze die richtige Antwort an.

- ○ 120 Minuten.
- ○ 100 Minuten.
- ○ 160 Minuten.
- ○ 200 Minuten.

Die folgende Abbildung stellt den Ausschnitt eines U-Bahn-Planes dar.

Wofür stehen die weißen Kreise auf den farbigen Linien in der Abbildung?
Kreuze die richtige Antwort an.

- ◯ Sie markieren U-Bahnstationen.
- ◯ Sie markieren U-Bahntunnel.
- ◯ Sie markieren die Länge der U-Bahnzüge.

Wofür stehen die weißen Doppelkreise auf den farbigen Linien in der Abbildung?
Kreuze die richtige Antwort an.

- ◯ Sie markieren U-Bahntunnel, ab denen die U-Bahnzüge oberirdisch fahren.
- ◯ Sie markieren U-Bahnstationen, an denen man von einer Linie in eine andere Linie umsteigen kann.
- ◯ Sie markieren U-Bahnstationen, an denen nachts keine U-Bahnzüge halten.

Wofür stehen die farbigen Linien in der Abbildung?
Kreuze die richtige Antwort an.

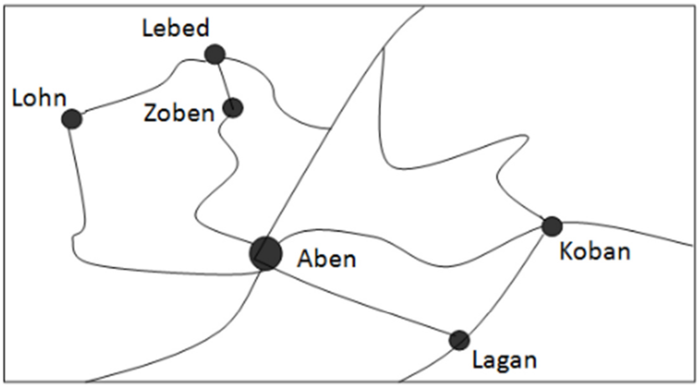
- ◯ Sie geben die Länge der U-Bahnzüge an.
- ◯ Sie geben die Schnelligkeit der U-Bahnzüge an.
- ◯ Sie geben den Fahrpreis für ein U-Bahnticket an.
- ◯ Sie geben den Streckenverlauf einer U-Bahnlinie an.

Gibt es eine Möglichkeit direkt von der roten auf die lila Linie umzusteigen?

Wovon hängt der Fahrpreis für ein Ticket beim U-Bahn-Fahren üblicherweise ab?
Kreuze die richtige Antwort an.

- ◯ von der Anzahl der angefahrenen U-Bahnstationen
- ◯ von der Geschwindigkeit des U-Bahnzuges
- ◯ von der Größe des U-Bahnzuges
- ◯ von der Anzahl der Sitze im U-Bahnzug

Die folgende Abbildung stellt eine Landkarte der Straßen zwischen verschiedenen Städten dar.



Welche Stadt ist die südlichste?
Kreuze die richtige Antwort an.

- ◯ Kobahn
- ◯ Lohn
- ◯ Lagan
- ◯ Aben

Welche Stadt ist die westlichste?
Kreuze die richtige Antwort an.

- ◯ Koban
- ◯ Aben
- ◯ Lagan
- ◯ Lohn

Gibt es eine Möglichkeit von Lohn nach Koban zu fahren ohne dabei durch Aben fahren zu müssen? Falls ja, gib die Städte an, die dabei durchfahren werden müssen.

Wie wird die Entfernung zwischen zwei Städten üblicherweise gemessen?
Kreuze die richtige Antwort an.

- ◯ in Metern
- ◯ in Grad
- ◯ in Kilometern
- ◯ in Höhenmetern

Die folgende Tabelle stellt Angaben zum Temperaturregler eines Gefrierschranks dar.

Regler-Position	Temperatur
1	-16 °C
2	-18 °C
3	-20 °C
4	-22 °C

Gib für jede Aussage an, ob sie wahr oder falsch ist.

	wahr	falsch
Je höher die Regler-Position eingestellt ist, umso niedriger wird die Temperatur.	○	○
Die Temperatur ist unabhängig von der Regler-Position.	○	○
Je höher die Regler-Position eingestellt ist, umso höher wird die Temperatur.	○	○

Was passiert, wenn man einen Gefrierschrank bei einer Raumtemperatur von 18°C häufig und lange öffnet?
Gib für jede Aussage an, ob sie wahr oder falsch ist.
Die Temperatur im Gefrierschrank...

	wahr	falsch
...sinkt, da frische Luft schneller gekühlt werden kann	○	○
...steigt, da warme Luft hinein kommt.	○	○
...bleibt gleich, da sich die Luft nicht bewegt.	○	○

Welche der folgenden Temperaturangaben ist die niedrigste (kälteste)?
Kreuze die richtige Antwort an.

- ○ 3°C
- ○ 16°C
- ○ -13°C
- ○ -16°C

Was versteht man in einer Bibliothek unter "Ausleihfrist"?
Kreuze die richtige Antwort an.

- ○ Die Zeit, die man ein Buch oder eine Zeitschrift ausleihen darf.
- ○ Die Zeit, die man warten muss, bis man ein Buch oder eine Zeitschrift ausleihen darf.
- ○ Die Zeit, die ein Buch oder eine Zeitschrift mindestens in der Bibliothek bleiben muss, bevor das Buch oder die Zeitschrift erneut ausgeliehen werden kann.
- ○ Die maximale Anzahl an Büchern oder Zeitschriften, die man gleichzeitig ausleihen darf.

Was versteht man in einer Bibliothek unter "überfälligen Ausleihen"?
Man versteht darunter Bücher oder Zeitschriften, ...

- ○ ...die lange nicht mehr ausgeliehen wurden.
- ○ ...die nur in der Bibliothek gelesen werden dürfen, aber nicht ausgeliehen werden können.
- ○ ...deren Ausleihfrist bereits abgelaufen ist.
- ○ ...von denen es mehrere Exemplare in der Bibliothek gibt.

Was versteht man in einer Bibliothek unter "vorbestellten Ausleihen"?
Man versteht darunter Bücher oder Zeitschriften, ...

- ○ ...die die Bibliothek bestellt hat, die jedoch noch nicht geliefert wurden.
- ○ ...die ausgeliehen sind, aber bereits von einer anderen Person zur Ausleihe vorgemerkt wurden.
- ○ ...für die eine längere Ausleihfrist gilt als für andere.
- ○ ...die nur in der Bibliothek gelesen werden dürfen, aber nicht ausgeliehen werden können.

Musterlösung

Bitte sieh dir die richtigen Antworten auf die Fragen der letzten Seiten an.

Was versteht man bei Kinofilmen unter "Altersbegrenzung" oder "Altersfreigabe"?
- Für Personen, die jünger als die angegebene Altersgrenze sind, ist der Film nicht geeignet.

Was bedeutet bei einem Kinoprogramm die Angabe "nur Mo-Fr"?
- Der Film wird nur Montag bis Freitag gezeigt.

Wie viele Minuten sind 1,5 Stunden?
- 90 Minuten.

Wie viele Minuten sind 2 Stunden?
- 120 Minuten.

Wofür stehen die weißen Kreise auf den farbigen Linien des U-Bahn-Planes?
- Sie markieren U-Bahnstationen.

Wofür stehen die weißen Doppelkreise auf den farbigen Linien in der Abbildung?
- Sie markieren U-Bahnstationen, an denen man von einer Linie in eine andere Linie umsteigen kann.

Wofür stehen die farbigen Linien in der Abbildung?
- Sie geben den Streckenverlauf einer U-Bahnlinie an.

Wovon hängt der Fahrpreis für ein Ticket beim U-Bahn-Fahren üblicherweise ab?
- von der Anzahl der angefahrenen U-Bahnstationen

Wie wird die Entfernung zwischen zwei Städten üblicherweise gemessen?
- in Kilometern

Was versteht man in einer Bibliothek unter "Ausleihfrist"?
- Die Zeit, die man ein Buch oder eine Zeitschrift ausleihen darf.

Was versteht man in einer Bibliothek unter "überfälligen Ausleihen"?
- Man versteht darunter Bücher oder Zeitschriften, deren Ausleihfrist bereits abgelaufen ist.

Was versteht man in einer Bibliothek unter "vorbestellten Ausleihen"?
- Man versteht darunter Bücher oder Zeitschriften, die ausgeliehen sind, aber bereits von einer anderen Person zur Ausleihe vorgemerkt wurden.

Sehr schön, , jetzt kennst du wichtige Begriffe für die folgenden Aufgaben.

Anmerkung. Zwischen den beiden Kommata erschien das Pseudonym, das die SuS angeben konnten.

Block: Planungsfähigkeit

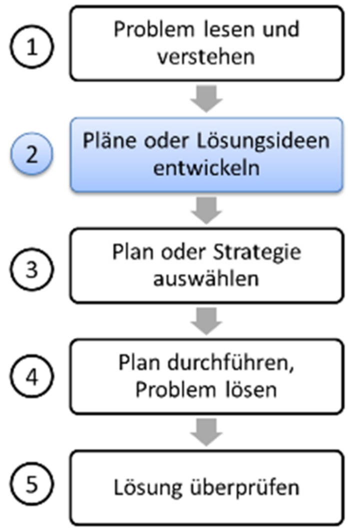
Problemlösen, ein Prozess

Erinnerst du dich an den Problemlöseprozess?

Wenn du Ideen hast, wie man ein Problem löst, kommt es oft darauf an, dass du bestimmte Schritte des Lösungsweges in einer sinnvollen Reihenfolge erledigst.

Dies geschieht in Phase 2 des Problemlöseprozesses (siehe Abbildung).

Mit diesem Thema wollen wir uns jetzt beschäftigen.



Planen

Wenn du ein Problem lösen möchtest, brauchst du einen **Plan**.

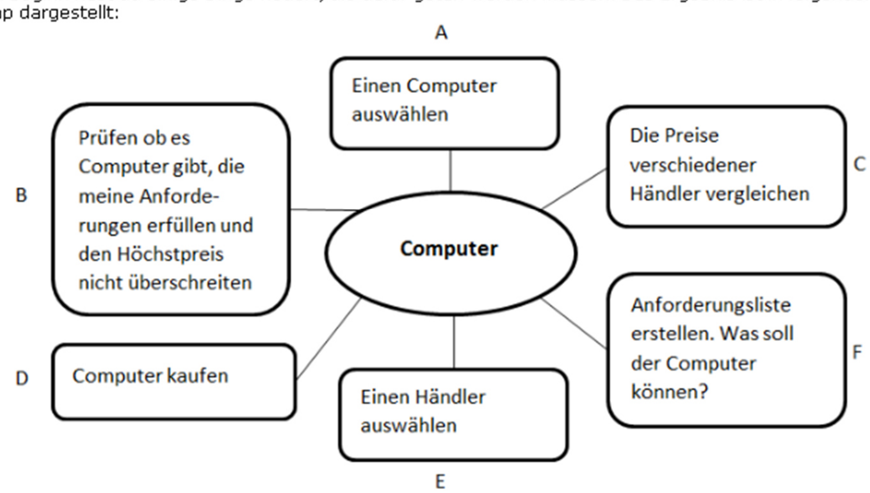
Manchmal reicht es, sich den Plan vorzustellen. Oft ist es aber hilfreich, sich den Plan zu notieren.

Einen Plan zu entwickeln bedeutet, sich Gedanken über mögliche Lösungen des Problems zu machen.

- Wie könnte man das Problem lösen?
- Welche alternativen Lösungswege gibt es?
- Welchen Lösungsweg soll ich wählen, wenn es mehrere Lösungswege gibt?
- Was mache ich, wenn ich bei der Problemlösung auf neue Schwierigkeiten stoße?

Auf den folgenden Seiten findest du je eine Aufgabe mit zugehörigen Lösungsschritten in Form einer Mindmap. Deine Aufgabe ist es die jeweiligen Teilschritte zur Lösung der Aufgabe, in eine sinnvolle Reihenfolge zu bringen. Wenn du glaubst, B ist der erste Teilschritt, dann wähle B im Feld hinter „1. Schritt" aus, usw.

Du möchtest unbedingt einen neuen Computer haben. Deine Eltern haben zugestimmt den Computer zu kaufen, allerdings nur unter der Bedingung, dass du dich gut informierst und eine Entscheidung triffst, welches Gerät gekauft werden soll. Außerdem haben sie dir einen Höchstpreis genannt, den der Computer nicht übersteigen darf. Zu Beginn hast du einige Dinge notiert, die dafür getan werden müssen. Das Ergebnis ist in folgender Mindmap dargestellt:

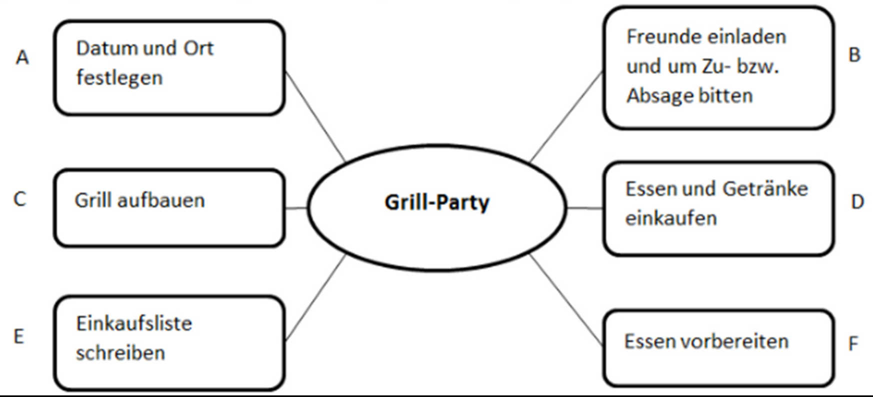


Deine Aufgabe ist es, die in der Mindmap angegebenen Teilschritte in eine sinnvolle Reihenfolge zu bringen.
Ordne dazu den einzelnen Teilschritten einen Buchstaben zu.

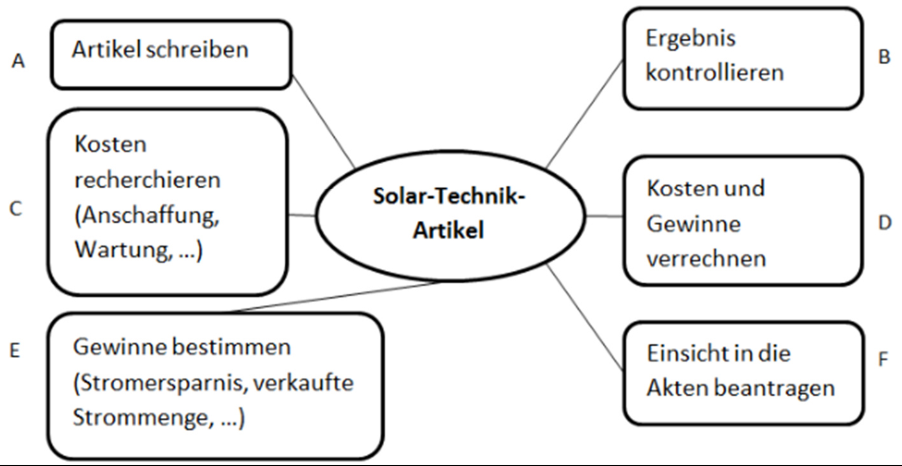
	zugeordneter Buchstabe
1. Schritt:	-
2. Schritt:	-
3. Schritt:	-
4. Schritt:	-
5. Schritt:	-
6. Schritt:	-

Aus Platzgründen und um Redundanz zu vermeiden, wird bei den nachfolgenden analogen Aufgaben nur noch der Aufgabenstamm samt Mindmap abgedruckt.

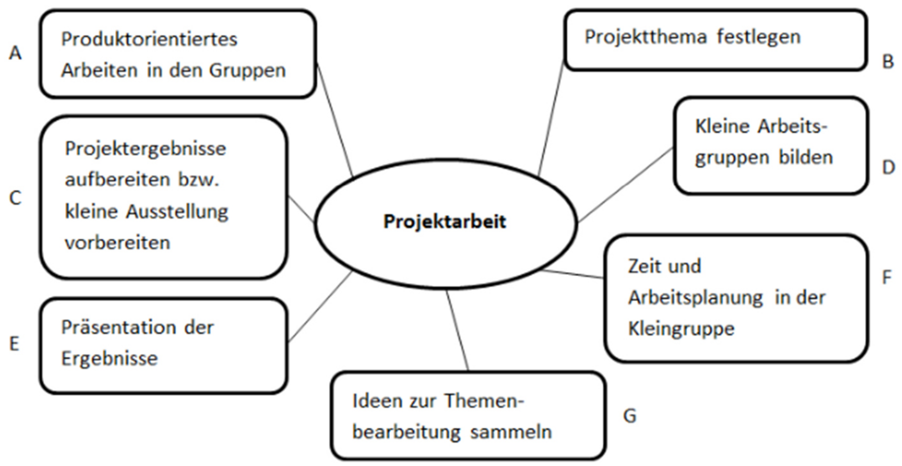
Du willst gemeinsam mit deinen Freunden eine Grill-Party veranstalten. Zu Beginn hast du mögliche Teilschritte des Vorgehens gesammelt. Das Ergebnis ist in folgender Mindmap dargestellt:



Lea will einen Artikel für die Schülerzeitung schreiben, der der Frage nachgeht, ob sich die vor fünf Jahren angeschaffte Solar-Technik an ihrer Schule finanziell gelohnt hat. Hilf ihr die Teilschritte ihres Vorgehens in eine sinnvolle Reihenfolge zu bringen:

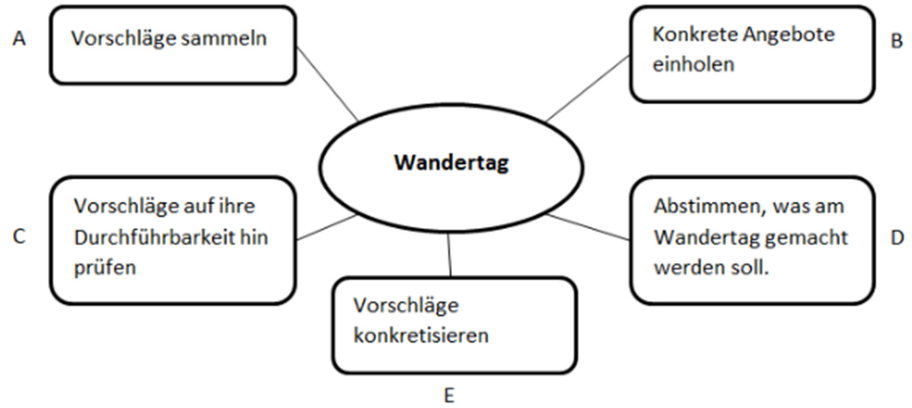


„Ein Projekt zeichnet sich dadurch aus, daß eine Arbeitsgruppe ein vorgegebenes oder selbst gewähltes Arbeitsvorhaben schrittweise plant, organisiert, durchführt und auswertet. Die wichtigsten Arbeitsschritte sind im folgenden ungeordnet aufgeführt“ (Klippert, 2002).



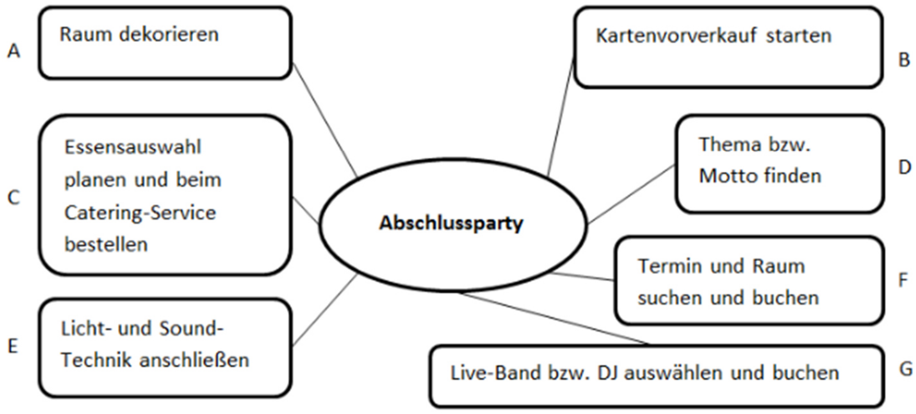
Anmerkung. Der Inhalt der Aufgabe *Projektarbeit* basiert auf einer Aufgabe von Klippert (2002).

In eurer Klasse steht in vier Wochen der Wandertag bevor. Dieses Jahr sollt ihr Eigeninitiative zeigen und den begleitenden Lehrern eine Planung vorlegen. Zu Beginn habt ihr einige Dinge notiert, die dafür getan werden müssen. Das Ergebnis ist in folgender Mindmap dargestellt:



Du bist im Organisationsteam der Abschlussparty deiner Jahrgangsstufe.

Zu Beginn habt ihr einige Dinge notiert, die zur Planung notwendig sind. Das Ergebnis ist in folgender Mindmap dargestellt:



Zwischenfazit und Überleitung

Sehr schön, du hast einen weiteren Teil von E-SePro beendet.

Pläne sind hilfreich, um Probleme zu lösen. Oft gibt es aber mehr als einen Plan. Welchen Plan man umsetzen soll, ist jedoch nicht immer leicht zu entscheiden.

Beispiel: Stuttgart 21

- In Stuttgart will man seit längerem den Hauptbahnhof verbessern. Dazu gab es einen Plan der deutschen Bahn AG namens „Stuttgart 21", der einen Umbau des Kopfbahnhofs in einen Durchgangsbahnhof vorsieht. Die Gegner von „Stuttgart 21" lehnen dieses Projekt jedoch ab und schlugen vor, den bestehenden Kopfbahnhof zu erweitern.

Block: Konditionales Wissen

Problemlösen, ein Prozess

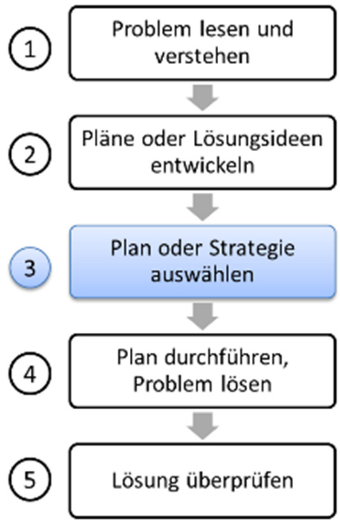
Betrachte bitte erneut den Problemlöseprozess

Wir beschäftigen uns jetzt mit Beispielen, in denen es mehrere Handlungsmöglichkeiten gibt.

Das bedeutet, dass du dich für einen Plan oder eine Strategie entscheiden musst.

Dies entspricht Phase 3 des Problemlöseprozesses (siehe Abbildung).

Mit diesem Thema wollen wir uns jetzt beschäftigen.



Wann mache ich was?

Wenn man vor einem Problem steht, weiß man manchmal nicht, welches Vorgehen zum Ziel führt. Dafür kann es verschiedene Gründe geben:

- Entweder man hat keine Idee
- oder man hat verschiedene Pläne, weiß aber nicht, welcher der Beste ist.

Jetzt wollen wir uns mit der Wahl des Vorgehens beschäftigen, wenn man mehrere Optionen oder Pläne hat.

Aufgabenstellung

Im Folgenden werden verschiedene Situationen beschrieben, zu denen du verschiedene Handlungsmöglichkeiten bewerten sollst. Sieh dir die Situation und die Handlungsmöglichkeiten bitte genau an. Bewerte dann jede einzelne Handlungsmöglichkeit mit einer Schulnote (von 1 = „sehr gut" bis 6 = „ungenügend"). Je besser eine Handlungsmöglichkeit deiner Meinung nach ist, umso besser sollte deine Benotung sein.

- "Gute" Handlungsmöglichkeiten, sollst du mit 1 oder 2 bewerten.
- "Mittelmäßige" Handlungsmöglichkeiten, sollst du mit 3 oder 4 bewerten.
- "Schlechte" Handlungsmöglichkeiten, sollst du mit 5 oder 6 bewerten.

Selbstverständlich kannst du bei deiner Bewertung die gleiche Schulnote mehrmals vergeben, wenn du Handlungsmöglichkeiten gleich gut findest.

Rückmeldungen

Wenn deine Antworten unvollständig sind oder deine Einschätzung der Handlungsmöglichkeiten deutlich von der Expertenmeinung abweicht, zeigt dir E-SePro nach deinem zweiten Lösungsversuch eine Musterlösung an. Dabei kommt folgendes Smiley-System zum Einsatz:

	=	Diese Handlungsmöglichkeit ist gut, sinnvoll, angemessen.
	=	Diese Handlungsmöglichkeit ist unvollständig oder verbesserungsfähig.
	=	Diese Handlungsmöglichkeit ist ungeeignet oder unangebracht.

Jonas überlegt, sich ein Smartphone zuzulegen. Dazu möchte er die aktuelle Version der Betriebssysteme Android und iOS vergleichen. Was könnte er tun?
Bewerte die folgenden Möglichkeiten, indem du pro Antwort eine Schulnote ankreuzt.

	1	2	3	4	5	6
Jonas geht in ein Geschäft und probiert Smartphones mit beiden Betriebssystemen aus.	○	○	○	○	○	○
Jonas informiert sich im Internet.	○	○	○	○	○	○
Jonas fragt im Bekanntenkreis nach, ob jemand Erfahrung mit einem der beiden Betriebssysteme hat.	○	○	○	○	○	○
Jonas lässt sich im Fachhandel beraten.	○	○	○	○	○	○
Jonas sucht in Fachzeitschriften nach einem Vergleichsartikel	○	○	○	○	○	○

Die Lösung einer komplizierten Hausaufgabe erfordert mehrere Schritte. Bei einem dieser Schritte kommst du nicht weiter. Wie gehst du vor?
Bewerte die folgenden Möglichkeiten, indem du pro Antwort eine Schulnote ankreuzt.

	1	2	3	4	5	6
Ich suche nach einer Beispielaufgabe, bei der ich diesen Schritt nachvollziehen kann.						
Ich denke darüber nach, ob andere Möglichkeiten bestehen die Aufgabe zu lösen.						
Um keine Zeit zu verlieren, fange ich sofort mit der nächsten Aufgabe an.						
Ich frage jemanden, ob er mir bei der Lösung helfen kann.						

Johanna hat den Verdacht, dass ihr W-LAN-Router nicht richtig funktioniert. Denn seit einer Stunde kommt sie mit ihrem Laptop nicht ins Internet. Was könnte sie sinnvollerweise als nächstes tun?
Bewerte die folgenden Möglichkeiten, indem du pro Antwort eine Schulnote ankreuzt.

	1	2	3	4	5	6
Einen neuen Router kaufen.						
Versuchen mit einem anderen Gerät per W-LAN ins Internet zu gelangen.						
Neustarten des W-LAN-Routers.						
Abwarten. Das Problem löst sich möglicherweise von selbst.						
Den Wartungsdienst anrufen, um einen Techniker zu bestellen.						
Dem Laptop gut zureden, damit er sich mit dem W-LAN verbindet.						

Du musst eine schriftliche Bewerbung für ein Praktikum schreiben. Wie gehst du vor?
Bewerte die folgenden Möglichkeiten, indem du pro Antwort eine Schulnote ankreuzt.

	1	2	3	4	5	6
Ich suche im Internet nach Bewerbungsunterlagen und orientiere mich daran.						
Ich orientiere mich an in der Schule besprochenen Bewerbungsunterlagen.						
Ich schreibe ohne Vorlage eine Bewerbung, so wie ich es für richtig halte.						
Ich bitte meine Eltern eine Bewerbung für mich zu schreiben.						

Pia darf in der kommenden Woche zum ersten Mal wählen. Sie ist politisch bislang wenig interessiert. Weil sie jedoch wählen möchte, überlegt sie, was sie tun sollte.
Bewerte die folgenden Möglichkeiten, indem du pro Antwort eine Schulnote ankreuzt.

	1	2	3	4	5	6
Da eine Woche nicht ausreicht, um sich politisch zu informieren, sollte sie überhaupt nicht wählen.						
Sie fragt ihre Eltern, was diese wählen, und wählt dasselbe.						
Sie informiert sich im Internet auf den Webseiten der Parteien über Kandidaten und Inhalte, bevor sie wählt.						
Sie sieht sich die Wahlplakate an und wählt den sympathischeren Kandidaten und seine Partei.						
Da eine Woche nicht ausreicht, um sich politisch zu informieren, macht sie einfach zufällig ihre Kreuze in der Wahlkabine.						
Sie nutzt den Wahl-O-Mat, ein Computerprogramm das zeigt, „welche zu einer Wahl zugelassene Partei der eigenen politischen Position am nächsten steht“ und wählt dem Ergebnis entsprechend.						
Sie besucht die Informationsstände der Parteien in der Fussgängerzone.						

Der 20-jährige Kölner Max W. will seine Kondition verbessern. Nun ist er auf der Suche nach einem guten Fitness-Studio, um dort regelmäßig zu trainieren.
Bewerte die folgenden Möglichkeiten, indem du pro Antwort eine Schulnote ankreuzt.

	1	2	3	4	5	6
Max fragt in seinem Freundeskreis nach Erfahrungen.	○	○	○	○	○	○
Max achtet auf die Anzeigen der Fitness-Studios in der Lokalzeitung.	○	○	○	○	○	○
Max recherchiert im Internet.	○	○	○	○	○	○
Max geht zum Probetraining in jedem Fitness-Studio in seiner Stadt.	○	○	○	○	○	○
Max geht zum Probetraining im nächstgelegenen Fitness-Studio.	○	○	○	○	○	○

Zwischenfazit

Wir haben uns gerade mit Beispielen beschäftigt, bei denen es mehrere Handlungsmöglichkeiten gibt.

Dabei hast du gesehen,

- dass es sich lohnt, über verschiedene Lösungsmöglichkeiten nachzudenken.
- dass verschiedene Wege zum Ziel führen können.
- dass sich verschiedene Optionen in ihrer Qualität unterscheiden.
- dass es manchmal gar nicht so einfach ist, sich für eine Option zu entscheiden.

Block: Handlungswissen

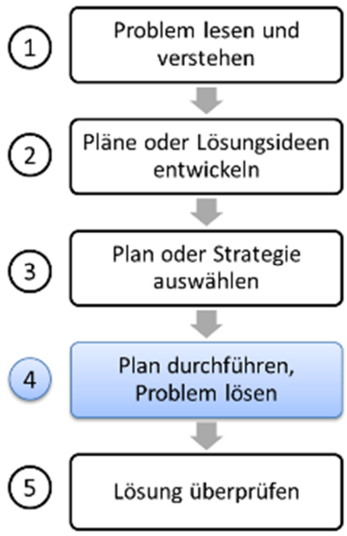
Problemlösen, ein Prozess

Die Abbildung zeigt wieder den Problemlöseprozess.

Der Umgang mit Text, Tabellen und Abbildungen kann in allen Phasen wichtig sein.

Oftmals beginnt dies schon beim Lesen und Verstehen der Ausgangssituation.

Auch beim konkreten Lösen eines Problems spielt der Umgang mit Text, Abbildungen und Tabellen eine wichtige Rolle (Phase 4) . Damit beschäftigen wir uns jetzt.



Text, Abbildungen und Tabellen

Grundlegend für das Lösen von Aufgaben ist neben dem Verständnis der Aufgabenstellung oftmals die Berücksichtigung von Abbildungen oder Tabellen.

Medien wie deine Schulbücher, Zeitungen oder Webseiten im Internet enthalten neben geschriebenen Text oftmals Abbildungen und Tabellen. Das hast du auch beim Neun-Punkte-Problem bereits gesehen.

Auf den nächsten Seiten folgen verschiedene Aufgaben, in denen neben Text Abbildungen oder Tabellen vorkommen. Bitte versuche die Aufgaben zu lösen. Wenn du mit den Aufgaben auf einer Seite fertig bist, klicke wie bisher auf „weiter". Wenn deine Lösung richtig ist, siehst du auf der nächsten Seite folgenden Smiley:

Wenn deine Lösung nicht vollständig richtig ist, erhältst du Tipps und einen zweiten Versuch, die Aufgabe zu lösen. Solltest du dann keinen besseren Lösungsvorschlag machen, siehst du die Musterlösung.

Mandy möchte einen neuen Toaster kaufen. Auf einem Vergleichsportal im Internet findest sie folgende Tabelle zum Vergleich dreier Toaster.

	Toasty Mach 2	Toastfix	Toasterwave
Optik	Sehr gut	befriedigend	gut
Handhabung	ausreichend	gut	befriedigend
Toastergebnis	befriedigend	gut	sehr gut
Stromverbrauch	mittel	mittel	gering
Gesamtnote	befriedigend (2,7)	gut (2,3)	gut (2,0)
Kindersicherung	ja	nein	nein
Preis	59,95 €	29,99 €	39,95 €

Frage 1: Welches ist der günstigste Toaster?

- ◯ Toasty Mach 2
- ◯ Toastfix
- ◯ Toasterwave
- ◯ Alle Geräte sind gleich teuer.

Frage 2: Ihrer Freundin Elena ist vor allem das Design wichtig. Welcher Toaster ist wohl ihr Favorit?

- ◯ Toasty Mach 2
- ◯ Toastfix
- ◯ Toasterwave
- ◯ Das kann man mit den Angaben in der Tabelle nicht beantworten.

Hier siehst du eine Übersicht über die Fehlstunden einiger Schüler der Jahrgangsstufe 13 während eines gesamten Schuljahres.

Name	Fehlstunden in den Fächern:			
	Mathe	Sport	Deutsch	Physik
Franka	3	4	2	8
Stefan	6	2	8	4
Pauline	2	10	4	2
Ahmet	6	6	2	-
Pia	-	-	-	-
Eddy	5	-	2	8
Franziska	-	2	-	2
Sengül	4	4	-	2
Patricia	6	-	4	-
Sascha	2	4	-	2

I1: Gibt es eine Schülerin oder einen Schüler ohne Fehlstunden? Wenn ja, wie lautet der Name?

I2: Wer hat im Sportunterricht die meisten Fehlstunden?

- Franka
- Stefan
- Pauline
- Ahmet
- Pia
- Eddy
- Franziska
- Sengül
- Patricia
- Sascha

I3: Wie viele Fehlstunden hat Sascha in Physik?

I4: Welche Schüler haben jeweils 6 Fehlstunden in Mathe?
Schreibe ihre Namen in das Textfeld.

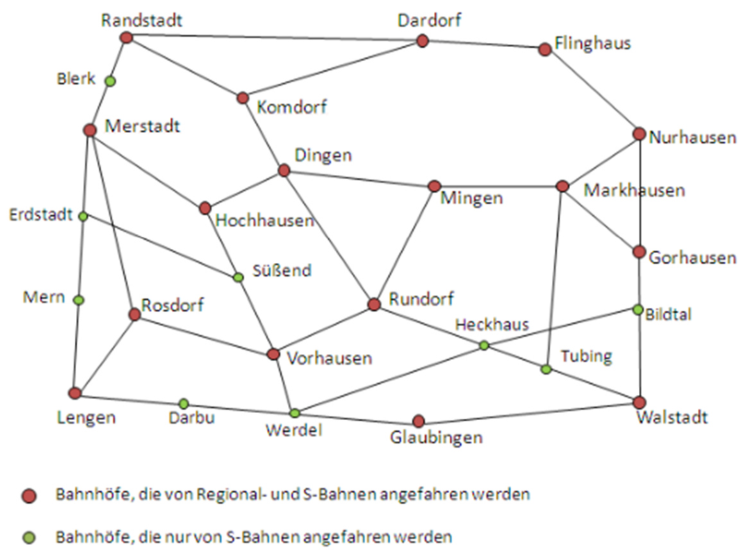
I5: Wer hat mehr Fehlstunden in Sport: Franka oder Patricia?

- Franka
- Patricia

I6: In welchem Fach hat Eddy die meisten Fehlstunden?

- Mathe
- Sport
- Deutsch
- Physik

Hier siehst du ein Streckennetz der Bahn:



Wie viele Bahnhöfe, die nur von S-Bahnen angefahren werden, gibt es auf der Karte?

Wie viele Bahnhöfe, die nur von S-Bahnen angefahren werden, liegen auf der Strecke Dingen-Rundorf-Heckhaus-Tubing-Walstadt?

Wie viele Bahnhöfe, die nur von S-Bahnen angefahren werden, liegen auf der Strecke Randstadt-Blerk-Merstadt-Erdstadt-Mern?

Unten siehst du den Jubiläumsflyer der Pizzeria Italia.

Pizzeria Italia	Tel. 05555 – 42 42 42
Altstraße 42	pizzeria-italia@www.de
55555 Kleinstadt	

Pizzeria Italia:
Seit 10 Jahren ihre Nummer 1!

Neue Jubiläumsangebote!

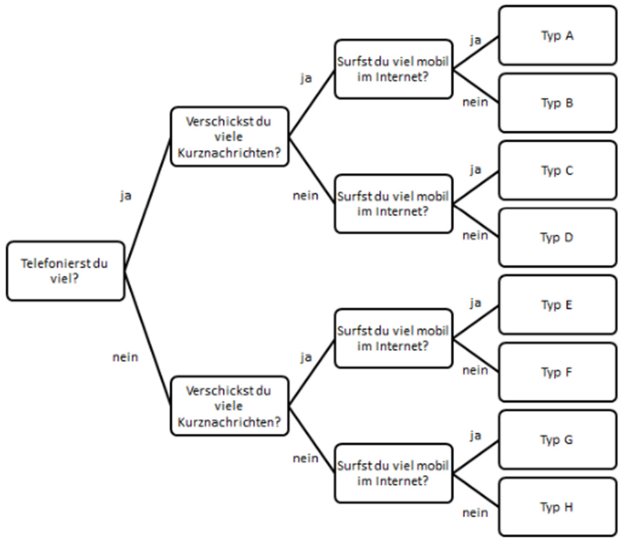
	Pizzen	clein (Ø 26 cm	mittel (Ø 32 cm)	groß (Ø 36 cm)
101	**Pizza Capricciosa** mit Schinken, Champignons und Artischocken	4,50 €	6,50 €	7,50 €
102	**Pizza Hawaii** mit Schinken und Ananas	3,00 €	5,00 €	6,00 €
103	**Pizza Vier Jahreszeiten** mit Champignons, Paprika, Schinken und Salami	3,50 €	5,50 €	6,50 €
104	Pizza Broccoli mit Broccoli, Tomatenscheiben, Knoblauch und Gorgonzola	4,50 €	6,50 €	7,50 €
105	**Pizza Meeresfrüchte** mit Krabben, Tintenfisch, Miesmuscheln, Zwiebeln und Knoblauch	5,50 €	4,50 €	8,50 €

Ab einem Bestellwert von 15 Euro liefern wir frei Haus!

Bevor der Jubiläumsflyer in den Druck gehen kann, muss er überprüft werden. Das ist deine Aufgabe.
Gib an, ob die folgenden Aspekte des Jubiläumsflyers mindestens einen Fehler enthalten.

Layout (u. a. Anordnung der grafischen Bestandtteile)	○ fehlerhaft	○ fehlerfrei
Preise für kleine Pizzen	○ fehlerhaft	○ fehlerfrei
Preise für mittlere Pizzen	○ fehlerhaft	○ fehlerfrei
Preise für große Pizzen	○ fehlerhaft	○ fehlerfrei
Rechtschreibung	○ fehlerhaft	○ fehlerfrei
Formatierung (z. B. einheitliche Schriftart)	○ fehlerhaft	○ fehlerfrei

Als Informatikprojekt wollen Udo und Mareike eine Internetseite erstellen, die zunächst erfragt, welchem Handynutzertyp der Besucher entspricht, und ihm dann Vorschläge für Handys bzw. Smartphones unterbreitet. Udo und Mareike befragen ihre Freunde und Bekannten. Im Wesentlichen unterscheiden die Befragten in ihren Antworten auf die folgenden drei Fragen: Telefonierst du viel? Verschickst du viele Kurznachrichten? Surfst du viel mobil im Internet?
Um sich das Programm besser vorstellen zu können, visualisieren Udo und Mareike ihre Ergebnisse.



Gib an, zu welchem Handynutzertyp die Personen jeweils gehören.

	Typ A	Typ B	Typ C	Typ D	Typ E	Typ F	Typ G	Typ H
Melinda telefoniert kaum, verschickt aber viele Kurznachrichten und surft viel im Internet.	○	○	○	○	○	○	○	○
Edgar findet Internet auf seinem Handy überflüssig, telefoniert täglich und schickt viele SMS.	○	○	○	○	○	○	○	○
Jason nutzt sein Handy nur zum Verschicken von Kurznachrichten.	○	○	○	○	○	○	○	○

Laura telefoniert bis zu zwei Stunden täglich und hört viel Internetradio. Zu welchen Handynutzertypen könnte sie gehören?
Kreuze alle Handynutzertypen an, zu denen Laura gehören könnte.

☐ Typ A ☐ Typ B ☐ Typ C ☐ Typ D ☐ Typ E ☐ Typ F ☐ Typ G ☐ Typ H

Maxi möchte für ihre Geburtstagsfeier multikulturell kochen. Die folgenden Gerichte (mit Zutatenliste) kommen in ihre engere Auswahl:

- **Spaghetti Carbonara** (Nudeln, Zwiebeln, Eigelb, Schmand, roher Schinken, Pfeffer, Knoblauch, Olivenöl)
- **Bifteki** (Hackfleisch, Schafskäse, Zwiebeln, Paniermehl, Eier, Pfeffer, Salz, Knoblauch, Olivenöl)
- **Lahmacun** (Mehl, Wasser, Zucker, Salz, Trockenhefe, Hackfleisch, Zwiebeln, Paprika, Salz, Chilipulver, Tomatenmark, Tomaten)
- **Fish ´n Chips** (Kartoffeln, Meersalz, Fischfilet, Bier, Mehl, Speisestärke, Backpulver, Salz, Pfeffer)
- **Vanille-Pudding** (Vanillepuddingpulver, Milch, Zucker)
- **Obstsalat** (Bananen, Äpfel, Orangen, Weintrauben, Zitronensaft, Zucker)

Bei der endgültigen Entscheidung will sie sich nach den Besonderheiten ihrer Gäste richten: Sie weiß, dass Simon eine Knoblauch-Allergie hat und Tina keinerlei Milchprodukte verträgt. Auf die übrigen Gäste muss Maxi bei der Essensauswahl nicht achten.

Frage 1: Welche Gerichte könnte Maxi für ihre Gäste 1:1 nach Rezept zubereiten, wenn sie darauf achtet, was ihre Gäste essen können und was nicht?

☐ Spaghetti Carbonara
☐ Bifteki
☐ Lahmacun
☐ Fish ´n Chips
☐ Vanille-Pudding
☐ Obstsalat

Frage 2: Überraschend sagt Tina zwei Tage vorher wegen Krankheit ab. Welche Gerichte könnte Maxi jetzt zusätzlich vorbereiten.

☐ Spaghetti Carbonara
☐ Bifteki
☐ Lahmacun
☐ Fish ´n Chips
☐ Vanille-Pudding
☐ Obstsalat

Anhang F. Experiment 1: Kontrollgruppentreatment

Die Screenshots wurden aus Druckgründen auf eine Seitenbreite von 14 cm skaliert, was ungefähr 2/3 der Originalgröße entspricht. Aus Platzgründen kennzeichnet ein schwarzer Rahmen jeweils eine Bildschirmseite, damit mehrere Screenshots auf einer Seite dokumentiert werden können.

Einleitung

Heute wirst du zwei kleine Lernprogramme ausprobieren.

Im Laufe des ersten Lernprogramms lernst du die kostenlose Mathematik-Software **GeoGebra** kennen.

GeoGebra 4

GeoGebra bietet viele Funktionen. Daher ist GeoGebra von der Grundschule bis zur Univerisität einsetzbar. In unserer kleinen Einführung lernst du die Benutzung im Bereich Geometrie kennen. Das bedeutet, dass du lernst, wie man mit GeoGebra Punkte, Strecken, Dreiecke und andere geometrische Figuren zeichnet.

Beim zweiten Lernprogramm darfst du verschiedenene Aufgaben zum problemlösenden Denken bearbeiten. Dazu erfährst du später mehr.

Kennst du GeoGebra bereits?

○ ja ○ nein

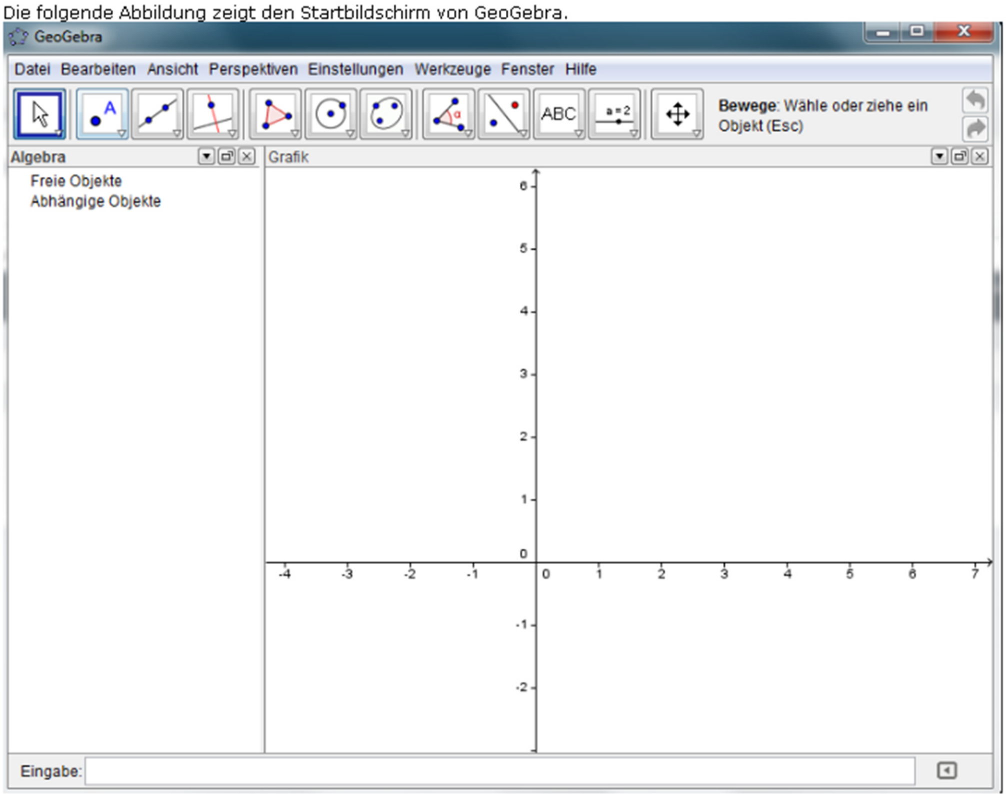
Habt ihr GeoGebra schon einmal im Unterricht verwendet?

○ ja ○ nein

Welche anderen Computerprogramme habt ihr im Mathematikunterricht schon benutzt?

Startbildschirm

Die folgende Abbildung zeigt den Startbildschirm von GeoGebra.

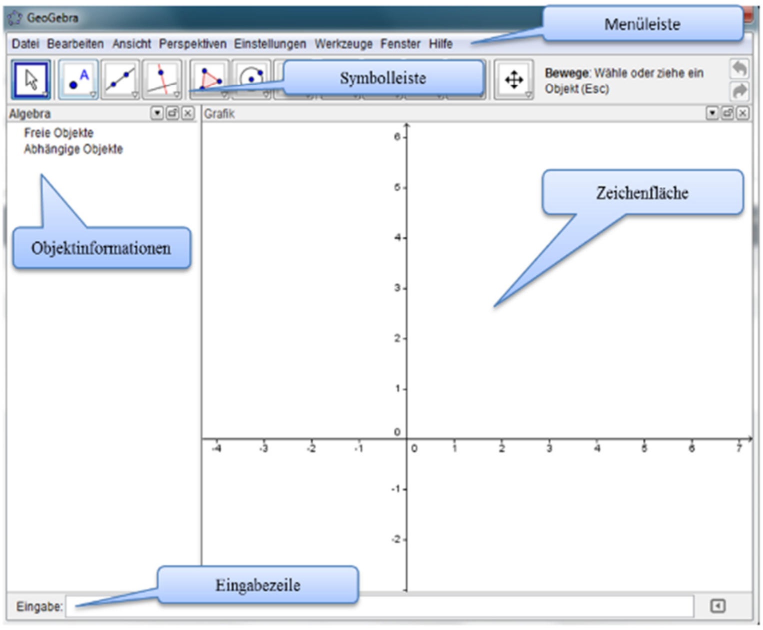


Auf der nächsten Seite erfährst du, aus welchen Bestandteilen die Benutzeroberfläche von GeoGebra besteht.

Bestandteile der Benutzeroberfläche

Die Benutzeroberfläche von GeoGebra besteht im Wesentlichen aus folgenden fünf Bereichen (siehe Abbildung):

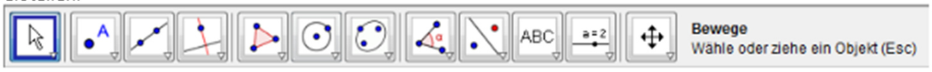
- Menüleiste
- Symbolleiste
- Objektinformationen
- Zeichenfläche
- Eingabezeile



In dieser Einführung brauchen wir hauptsächlich die Symbolleiste und die Zeichenfläche. Auf der nächsten Seite wird dir zunächst die Symbolleiste vorgestellt.

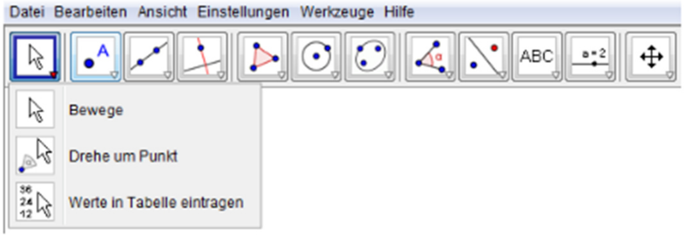
Symbolleiste

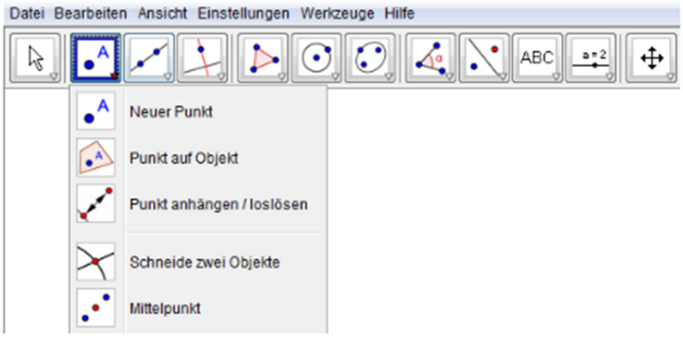
Die Symbolleiste enthält alle wichtigen Möglichkeiten, um geometrische Objekte wie Punkte, Linien und Kreise zu erstellen.

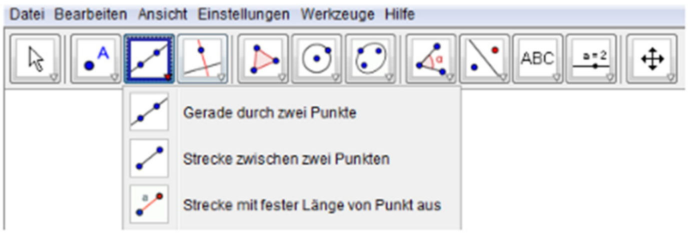


Bestandteile der Symbolleiste (Teil 1)

Sieh dir bitte die folgenden Abbildungen gut an, um zu verstehen, wo du welche Programmbefehle findest:

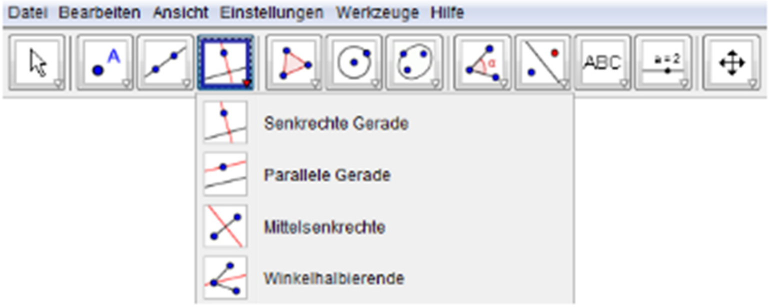


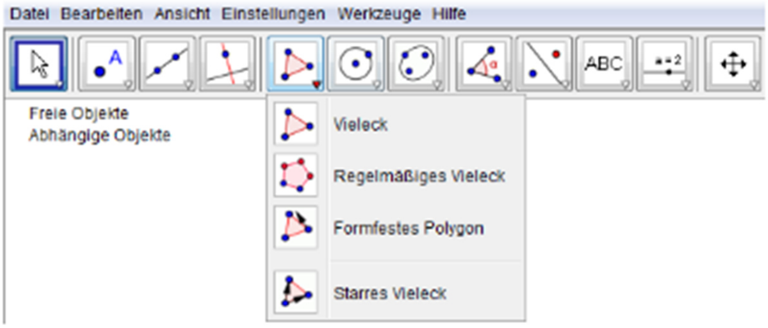


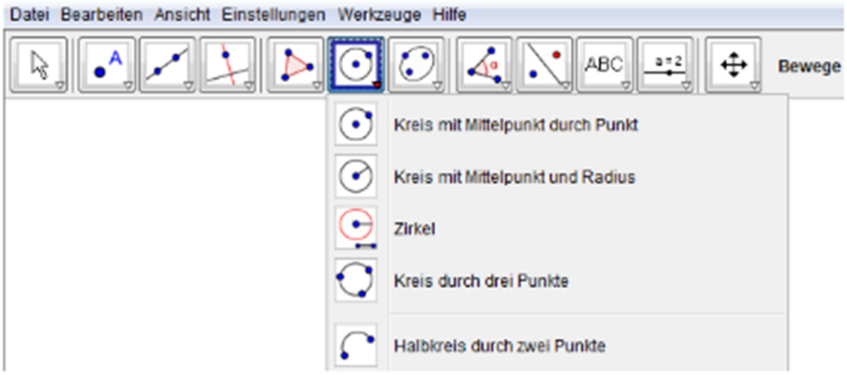


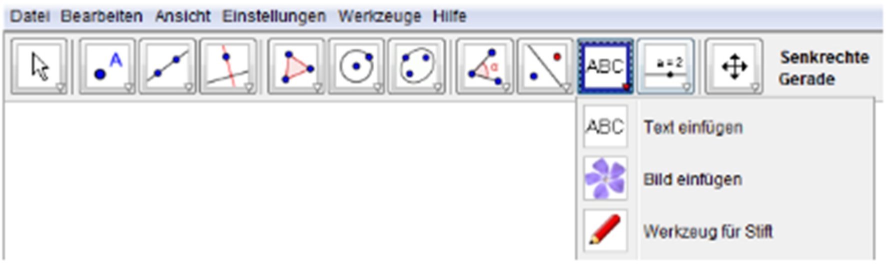
Bestandteile der Symbolleiste (Teil 2)

Sieh dir bitte die folgenden Abbildungen gut an, um zu verstehen, wo du welche Programmbefehle findest:









Die übrigen Symbole, die hier nicht vorgestellt worden sind, brauchen wir bei unseren Beispielen nicht.

Ausprobieren

Nachdem du die Symbole theoretisch kennen gelernt hast, ist es Zeit, GeoGebra praktisch auszuprobieren.

GeoGebra kann für viele mathematische Bereiche verwendet werden. In dieser Einführung lernst du nur die Verwendung in der Geometrie kennen.

GeoGebra siehst du weiter unten auf dieser Seite zwischen den zwei schwarzen Linien. Eventuell musst du dazu etwas nach unten scrollen.

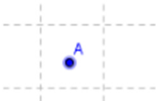
Um die Oberfläche für unsere Beispiele etwas übersichtlicher zu gestalten, klicke im Menü „Ansicht"

- einmal auf „Achsen",
- einmal auf „Koordinatengitter",
- und einmal auf „Algebra".

Damit du GeoGebra kennen lernen kannst, beginnen wir mit einigen Übungsaufgaben:

Übung 1: Punkte zeichnen

a) Wir fangen ganz einfach und wollen einen Punkt zeichnen. Klicke dazu mit der linken Maustaste auf das Symbol  und dann auf die Zeichenfläche. Das Ergebnis sieht dann etwa so aus:



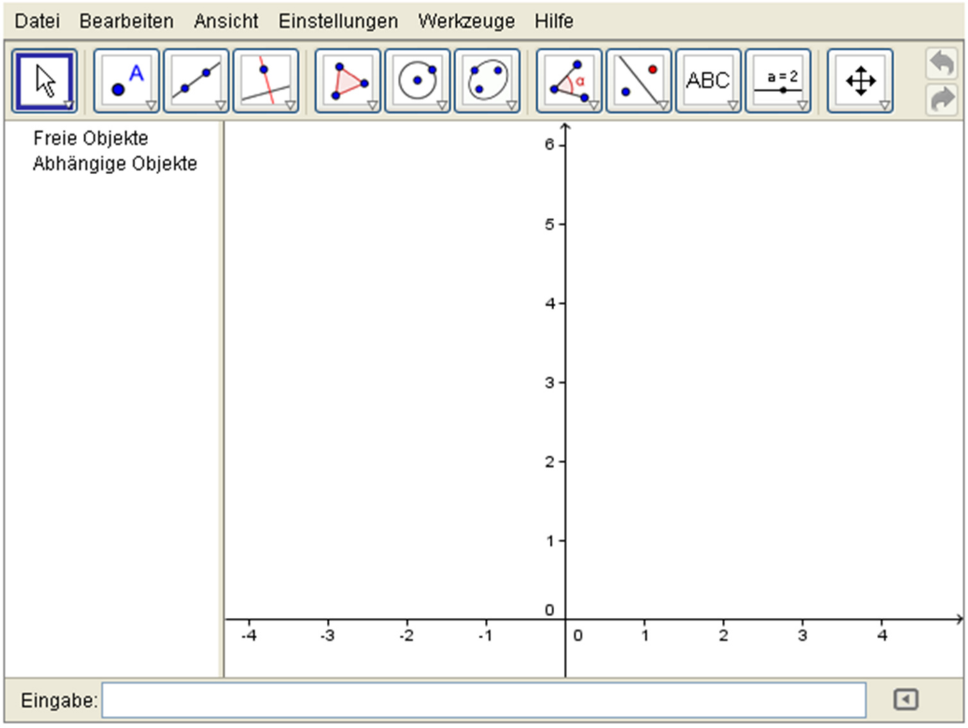
b) Wenn du den gezeichneten Punkt verschieben möchtest, klicke in der Symbolleiste auf das Mauszeigersymbol und dann auf den Punkt. Jetzt kannst du ihn auf der Zeichenfläche verschieben. Probiere das bitte jetzt aus.

GeoGebra

Übung 2: Eine Strecke zwischen zwei Punkten zeichnen

Als zweites zeichne bitte eine Strecke. Um eine Strecke zwischen zwei Punkten zu zeichnen, musst du zunächst zwei Punkte zeichnen. Dann wähle im Menü das Symbol „Strecke zwischen zwei Punkten"  aus.

GeoGebra



Wie leicht oder schwer war die Übung 2 zu verstehen?

sehr leicht ○ ○ ○ ○ ○ ○ ○ **sehr schwer**

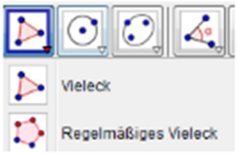
Beim Bearbeiten und Verstehen der Übung 2 war meine Denk-Anstrengung ...

sehr gering ○ ○ ○ ○ ○ ○ ○ **sehr hoch**

Übung 3: Ein Dreieck zeichnen

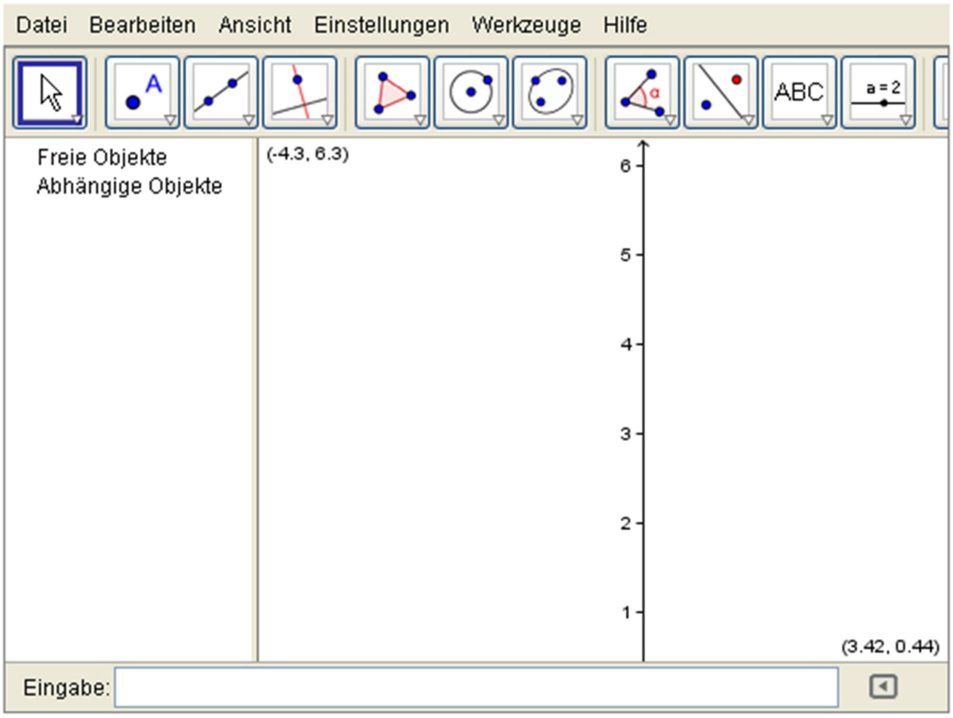
a) Da du bereits Strecken zeichnen kannst, kannst du auch schon Dreiecke zeichnen, indem du drei Punkte zeichnest und diese mit Strecken verbindest. Versuche das bitte jetzt.

b) Es gibt jedoch noch andere, einfachere Möglichkeiten, Dreiecke (und andere Vielecke) zu zeichnen. Dazu klickt man auf das Symbol „Vieleck". Dann klickt man an die Stellen, an denen die Eckpunkte entstehen sollen. Achte darauf, dass du am Ende nochmals auf den Startpunkt, d.h. die erste Ecke, klicken musst. Erstelle auf diese Weise ein weiteres Dreieck.



c) Wenn du ein Dreieck oder ein anderes Vieleck mit gleich langen Seiten erstellen möchtest, wählst du das „regelmäßige Vieleck" aus und folgst den Anweisungen in der rechten oberen Ecke. Versuche jetzt ein Quadrat zu zeichnen.

GeoGebra



Wie leicht oder schwer war die Übung 3 zu verstehen?

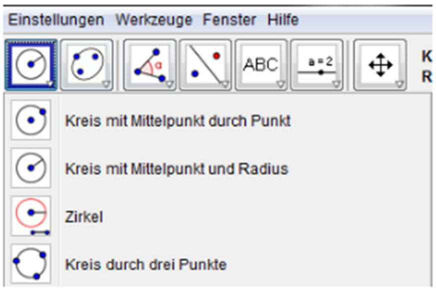
sehr leicht ○ ○ ○ ○ ○ ○ ○ **sehr schwer**

Beim Bearbeiten und Verstehen der Übung 3 war meine Denk-Anstrengung ...

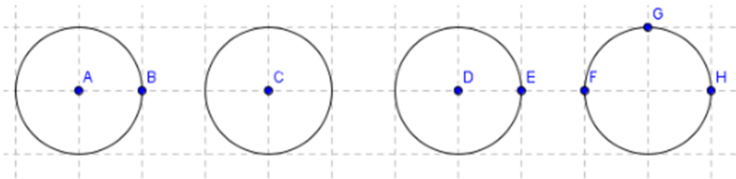
sehr gering ○ ○ ○ ○ ○ ○ ○ **sehr hoch**

Übung 4: Einen Kreis zeichnen

Um mit GeoGebra einen Kreis zu zeichnen, gibt es verschiedene Möglichkeiten, die du über das Kreissymbol aufrufen kannst:

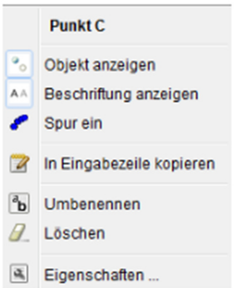


Die folgende Abbildung zeigt dir das Ergebnis der verschiedenen Möglichkeiten. Je nach Vorgehen werden dabei unterschiedlich viele Punkte angelegt.



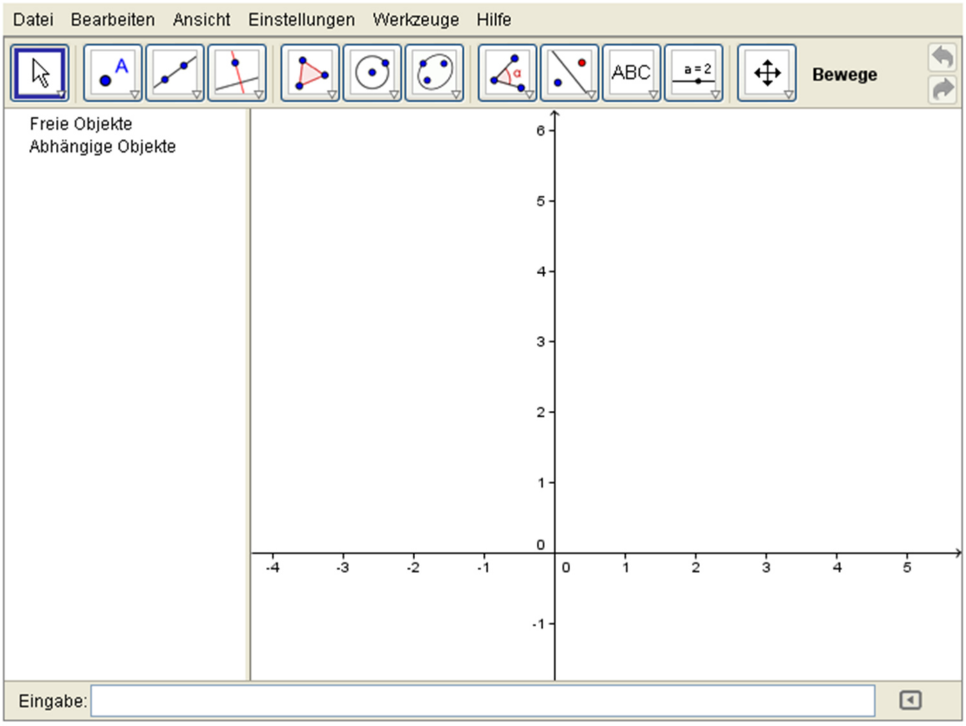
Versuche nun die vier Kreise aus der Abbildung mit GeoGebra zu erzeugen.

Hinweis: Wenn man die Punkte nicht sehen möchte, kann man sie ausblenden. Dazu klickt man mit der rechten Maustaste auf den gewünschten Punkt. Dadurch öffnet sich folgendes Fenster:



Wenn du einmal auf "Objekt anzeigen" klickst, wird das Häckchen davor entfernt und das Objekt ausgeblendet. Dieses Ausblenden funktioniert nicht nur für Punkte, sondern für jedes gezeichnete Objekt.

GeoGebra



Wie leicht oder schwer war die Übung 4 zu verstehen?

sehr leicht ○ ○ ○ ○ ○ ○ ○ **sehr schwer**

Beim Bearbeiten und Verstehen der Übung 4 war meine Denk-Anstrengung ...

sehr gering ○ ○ ○ ○ ○ ○ ○ **sehr hoch**

In den Übungen hast du gelernt,

- wie du Punkte zeichnest,
- wie du Strecken und Linien zeichnest,
- wie du Dreiecke, Quadrate und andere Flächen zeichnest,
- wie du Objekte ausblendest und wieder einblendest.

Mit diesem Wissen bist du gut vorbereitet, um verschiedene geometrische Figuren zu zeichnen und geometrische Fragestellungen zu visualisieren.

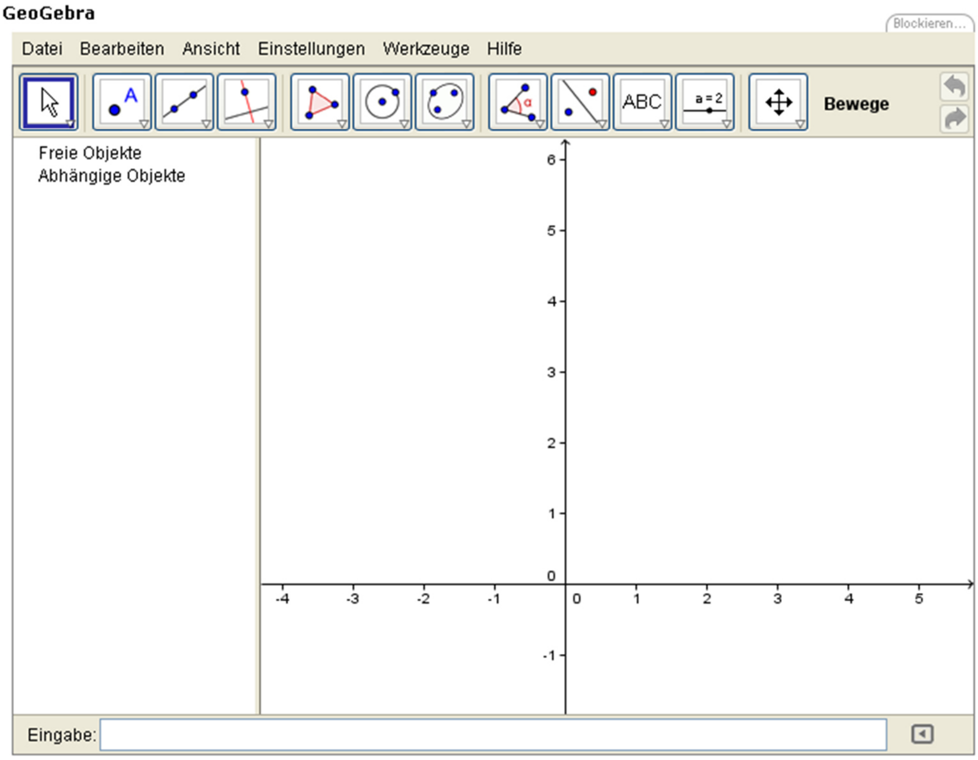
Selbstverständlich kannst du auch Text zu deinen Zeichnungen hinzufügen. Dabei gibt es verschiedene Möglichkeiten.

a) Du kannst Text einfügen, indem du auf das Symbol "ABC" klickst und den Bildschirmanweisungen folgst.

b) Du kannst das Stiftsymbol verwenden, um mit gedrückter linker Maustaste zu zeichnen und zu schreiben.

Probiere jetzt bitte beide Möglichkeiten aus.

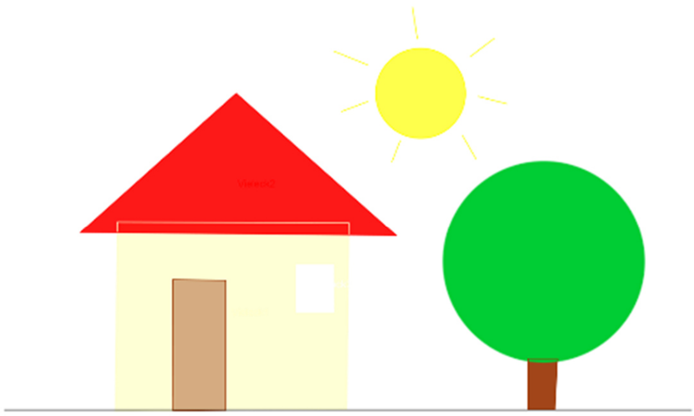
GeoGebra



Übung 5: Skizzen anfertigen und Malen mit GeoGebra

Mit GeoGebra kann man auch Zeichnungen anfertigen und Bilder malen, die aus geometrischen Formen bestehen. Die folgenden zwei Bilder zeigen dir Beispiele:

Beispiel 1

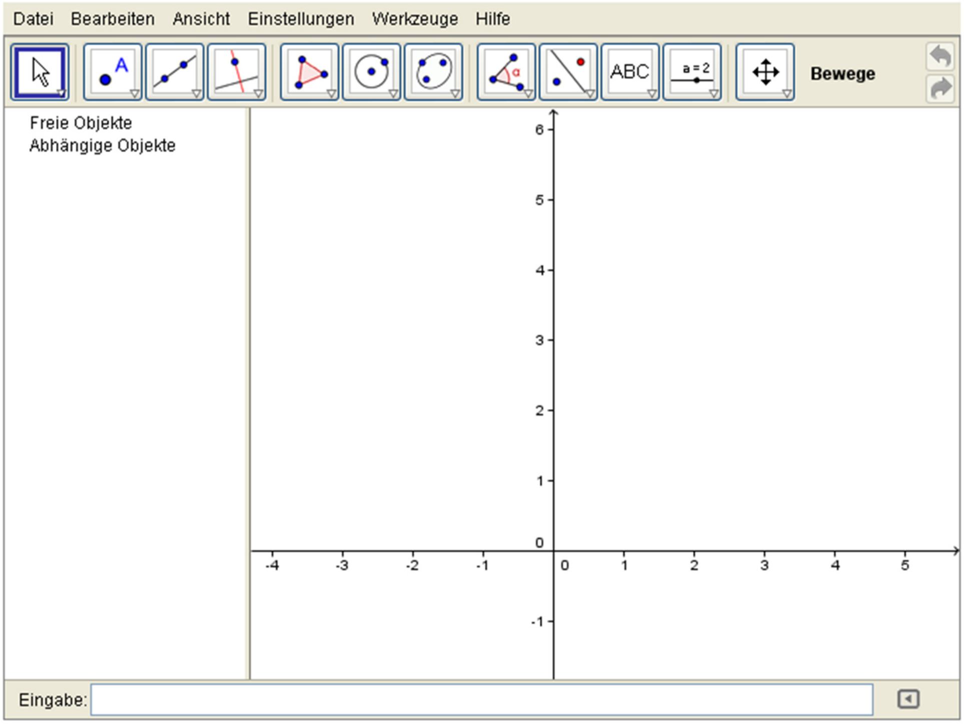


Beispiel 2

Wie man geometrische Figuren mit GeoGebra erzeugt, weißt du bereits. Wenn du die Farbe eines Objekts ändern möchtest, musst du es mit der rechten Maustaste anklicken und dann "Eigenschaften" anklicken. Dort kannst du die Farbe ändern.

Versuche nun selbst eine Zeichnung anzufertigen, damit du GeoGebra besser kennen lernst. Nimm dir dafür ca. 5 Minuten Zeit.

GeoGebra



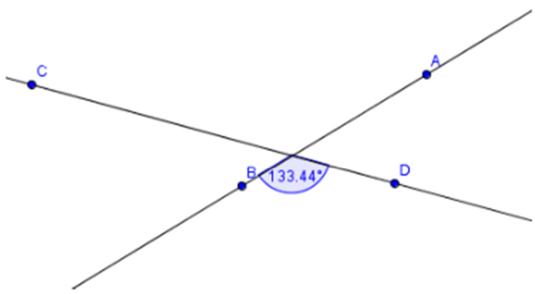
Was hast du gezeichnet?

Hinweis: Der Weiter-Button wird erst nach 5 Minuten auf dieser Seite eingeblendet.

Nachdem du GeoGebra besser kennengelernt hast, versuche bitte folgende Frage zu beantworten:

Wie hätte GeoGebra im aktuellen Schuljahr sinnvoll in deinem Mathematikunterricht eingesetzt werden können?

Unsere kleine Einführung hat sich hauptsächlich mit der Benutzeroberfläche von GeoGebra beschäftigt, die du im Bereich Geometrie brauchen kannst.



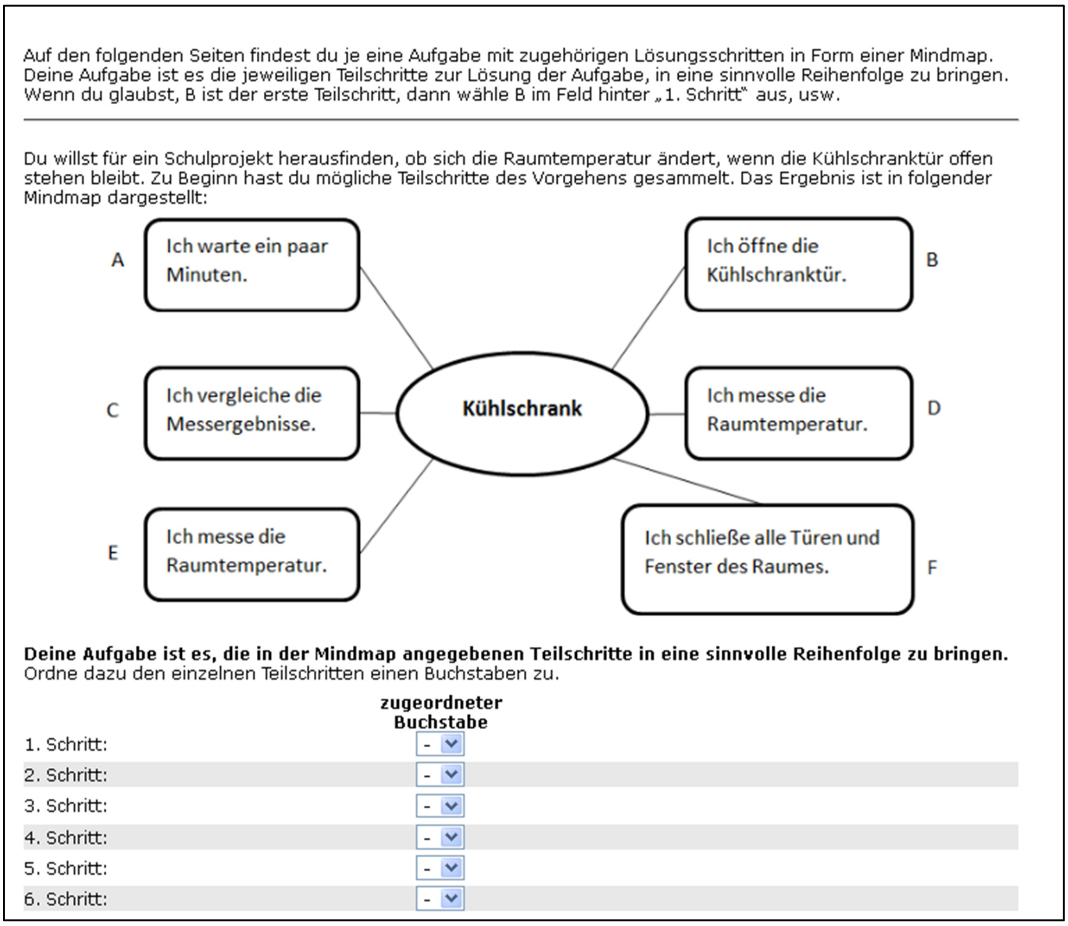
GeoGebra kann man noch für viele andere Bereiche der Mathematik nutzen, die wir aus Zeitgründen hier jedoch nicht thematisieren. Mit Hilfe der Formelzeile, kannst du z. B. Funktionen zeichnen lassen.

Wenn du GeoGebra Zuhause nutzen möchtest, gib Zuhause "GeoGebra" in eine Suchmaschine ein oder Frage deinen Mathematiklehrer.

Anhang G. Experiment 1: Posttest Problemlösen

Die Screenshots wurden aus Druckgründen auf eine Seitenbreite von 14 cm skaliert, was ungefähr 2/3 der Originalgröße entspricht. Aus Platzgründen kennzeichnet ein schwarzer Rahmen jeweils eine Bildschirmseite, damit mehrere Screenshots auf einer Seite dokumentiert werden können.

Block: Planungsfähigkeit


Auf den folgenden Seiten findest du je eine Aufgabe mit zugehörigen Lösungsschritten in Form einer Mindmap. Deine Aufgabe ist es die jeweiligen Teilschritte zur Lösung der Aufgabe, in eine sinnvolle Reihenfolge zu bringen. Wenn du glaubst, B ist der erste Teilschritt, dann wähle B im Feld hinter „1. Schritt" aus, usw.

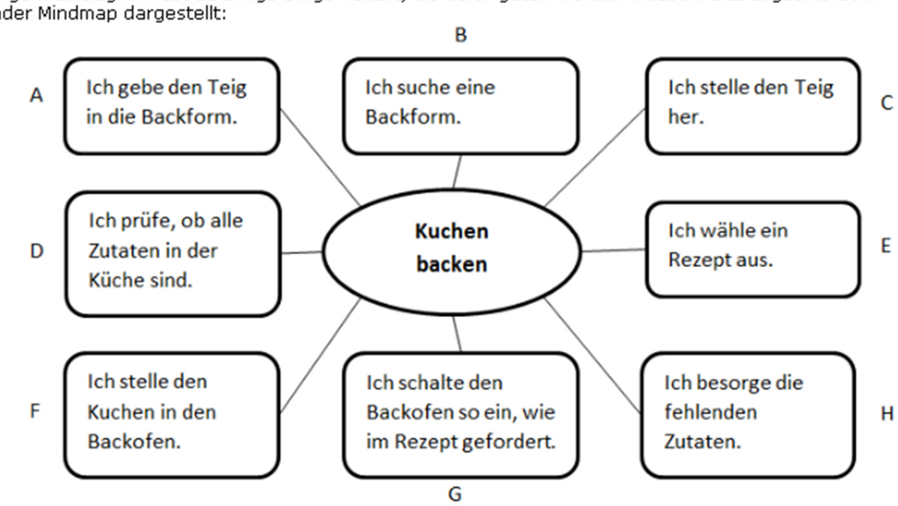
Du willst für ein Schulprojekt herausfinden, ob sich die Raumtemperatur ändert, wenn die Kühlschranktür offen stehen bleibt. Zu Beginn hast du mögliche Teilschritte des Vorgehens gesammelt. Das Ergebnis ist in folgender Mindmap dargestellt:

Deine Aufgabe ist es, die in der Mindmap angegebenen Teilschritte in eine sinnvolle Reihenfolge zu bringen.
Ordne dazu den einzelnen Teilschritten einen Buchstaben zu.

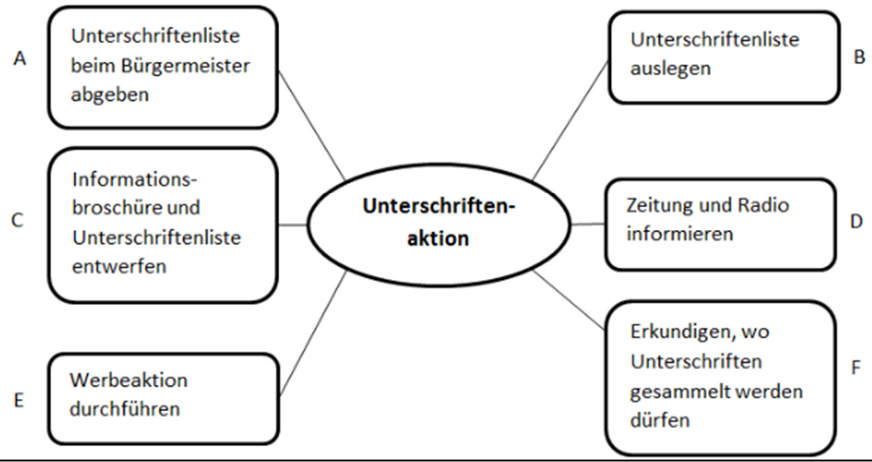
	zugeordneter Buchstabe
1. Schritt:	-
2. Schritt:	-
3. Schritt:	-
4. Schritt:	-
5. Schritt:	-
6. Schritt:	-

Aus Platzgründen und um Redundanz zu vermeiden, wird bei den nachfolgenden analogen Aufgaben nur noch der Aufgabenstamm samt Mindmap abgedruckt.

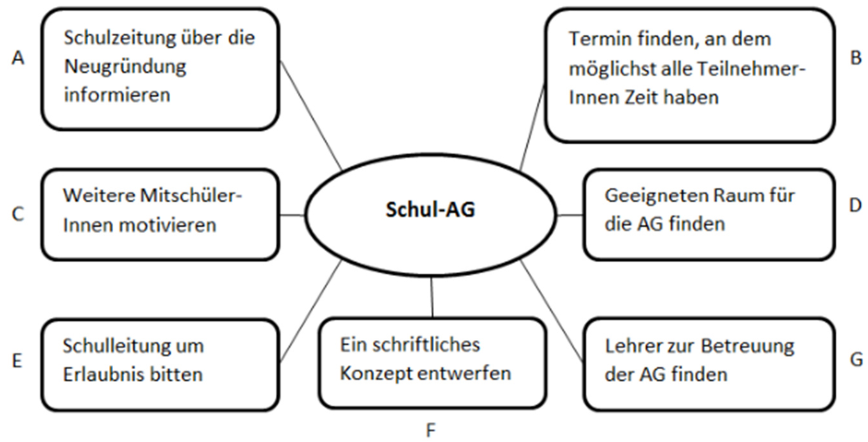
Du bist auf eine Geburtstagsparty eingeladen und willst als Geschenk einen selbstgebackenen Kuchen mitbringen. Zu Beginn hast du einige Dinge notiert, die dafür getan werden müssen. Das Ergebnis ist in folgender Mindmap dargestellt:



In deiner Stadt soll das einzige Schwimmbad geschlossen werden. Du hast mit deinen Freunden beschlossen, eine Unterschriftenaktion gegen die Schließung des Schwimmbads zu organisieren. Zu Beginn habt ihr mögliche Teilschritte des Vorgehens gesammelt. Das Ergebnis ist in folgender Mindmap dargestellt:



Du möchtest mit deinen Freunden unbedingt eine neue Schul-AG an eurer Schule gründen. Zu Beginn habt ihr einige Dinge notiert, die dafür getan werden müssen. Das Ergebnis ist in folgender Mindmap dargestellt:



Block: Konditionales Wissen

Im Folgenden werden verschiedene Situationen beschrieben, zu denen du verschiedene Handlungsmöglichkeiten bewerten sollst. Sieh dir die Situation und die Handlungsmöglichkeiten bitte genau an. Bewerte dann jede einzelne Handlungsmöglichkeit mit einer Schulnote (von 1 = „sehr gut" bis 6 = „ungenügend"). Je besser eine Handlungsmöglichkeit deiner Meinung nach ist, umso besser sollte deine Benotung sein.

- "Gute" Handlungsmöglichkeiten, sollst du mit 1 oder 2 bewerten.
- "Mittelmäßige" Handlungsmöglichkeiten, sollst du mit 3 oder 4 bewerten.
- "Schlechte" Handlungsmöglichkeiten, sollst du mit 5 oder 6 bewerten.

Selbstverständlich kannst du bei deiner Bewertung die gleiche Schulnote mehrmals vergeben, wenn du Handlungsmöglichkeiten gleich gut findest.

Dein Lehrer bzw. deine Lehrerin stellt euch eine neue Aufgabe zur Bearbeitung in Stillarbeit. Wie gehst du vor?
Bewerte die folgenden Möglichkeiten, indem du pro Antwort eine Schulnote ankreuzt.

	1	2	3	4	5	6
Ich verschaffe mir einen Überblick, indem ich die Aufgabe mehrfach lese, und bewerte dann die verschiedenen Wege, die zur Lösung führen könnten.	○	○	○	○	○	○
Ich achte auf die in der Aufgabe gegebenen Informationen und arbeite mit ihnen so, wie im Unterricht gehabt.	○	○	○	○	○	○
Ich suche nach einem Schlüsselwort, das den Lösungsweg anzeigt. Meist steht es in der letzten Zeile bzw. im letzten Satz.	○	○	○	○	○	○
Auch wenn ich mir sicher bin, auf dem richtigen Weg zu sein, überdenke ich beim Lösen der Aufgaben den gemachten Lösungsweg.	○	○	○	○	○	○
Ich versuche, mich an den Lösungsweg einer ähnlichen Aufgabe zu erinnern und diesen auf die neue Aufgabe zu übertragen.	○	○	○	○	○	○

Du bist in der Bibliothek und möchtest dir einen neuen Roman ausleihen. Da die Bibliothek in 20 Minuten schließt, hast du nur wenig Zeit, dich zu entscheiden, welches Buch du ausleihen willst. Wie gehst du vor?
Bewerte die folgenden Möglichkeiten, indem du pro Antwort eine Schulnote ankreuzt.

	1	2	3	4	5	6
Ich lese mir bei einigen Büchern die Zusammenfassung auf der Rückseite durch.	○	○	○	○	○	○
Ich lese bei einigen Büchern die ersten Kapitel.	○	○	○	○	○	○
Ich lese bei einigen Büchern zufällig eine Seite.	○	○	○	○	○	○
Ich schaue nach Büchern eines Autors, den ich gerne mag.	○	○	○	○	○	○

In deiner Sportmannschaft muss ein neuer Mannschaftsführer bzw. eine neue Mannschaftsführerin gewählt werden. Du willst unbedingt gewählt werden. Wie gehst du vor?
Bewerte die folgenden Möglichkeiten, indem du pro Antwort eine Schulnote ankreuzt.

	1	2	3	4	5	6
Ich bleibe still und hoffe, dass jemand mich vorschlägt.	○	○	○	○	○	○
Ich schlage vor, dass sich alle Interessenten melden und melde mich dann selbst auch.	○	○	○	○	○	○
Ich mache schon vorher Werbung für mich und erkläre, wofür ich eintrete.	○	○	○	○	○	○
Ich verspreche alles Mögliche, um gut da zu stehen.	○	○	○	○	○	○

Du willst ein aktuelles Lied aus den Musikcharts auswendig lernen. Wie gehst du vor?
Bewerte die folgenden Möglichkeiten, indem du pro Antwort eine Schulnote ankreuzt

	1	2	3	4	5	6
Immer wenn das Lied im Radio läuft, summe ich mit.	○	○	○	○	○	○
Ich besorge mir den Liedtext und lerne ihn durch wiederholtes leises Lesen auswendig.	○	○	○	○	○	○
Ich schreibe eine Zusammenfassung des Liedtextes und lese diese häufig durch.	○	○	○	○	○	○
Ich besorge mir den Liedtext und singe mir das Lied häufig laut vor.	○	○	○	○	○	○
Ich besorge mir die Musik sowie den Liedtext und singe häufig laut mit.	○	○	○	○	○	○

Beim Surfen im Internet stört dich die Werbung. Was tust du?
Bewerte die folgenden Möglichkeiten, indem du pro Antwort eine Schulnote ankreuzt.

	1	2	3	4	5	6
Ich versuche, nicht auf die Werbung zu achten.	○	○	○	○	○	○
Ich installiere ein Werbe-Blocker-Programm.	○	○	○	○	○	○
Ich steige auf einen anderen Browser um.	○	○	○	○	○	○
Ich frage jemanden, wie ich die Werbung loswerde.	○	○	○	○	○	○

Block: Handlungswissen

Bei den folgenden Aufgaben sollst du Aufgaben und Fragen zu gegebenen Tabellen und Abbildungen bearbeiten.

Hier siehst du einen Ausschnitt aus einer Abfahrtstafel am Weltener Hauptbahnhof.

Zeit	Linie	*Über*	Ziel	**Gleis**
07:32	S 17	Komdorf Ratsdorf Windhaus Dardorf	Flinghaus	**2**
07:38	IC 23	Blerk Merstadt Erdstadt Rosdorf	Mern	**6**
07:44	S 17	Komdorf Ratsdorf Windhaus Dardorf	Flinghaus	**2**
07:45	RB 3	Dardorf Flinghaus Nurhausen Markhausen	Mingen	**10**
07:48	IC 67	Kugeldorf Dingen Ratsdorf Begel	Windhaus	**5**
07:54	S 16	Gorhausen Bildtal Heckhaus Werdel	Glaubingen	**1**
08:08	IC 23	Blerk Merstadt Erdstadt Rosdorf	Mern	**6**
08:15	RB 6	Merstadt Lengen Darbu	Hochhausen	**10**
08:18	IC 67	Kugeldorf Dingen Ratsdorf Begel	Windhaus	**5**
08:24	S 16	Gorhausen Bildtal Heckhaus Werdel	Glaubingen	**1**
08:32	S 17	Komdorf Ratsdorf Windhaus Dardorf	Flinghaus	**2**
08:34	IC 40	Vorhausen Bildtal Süßend Rundorf	Heckhaus	**5**
08:38	IC 23	Blerk Merstadt Erdstadt Rosdorf	Mern	**6**
08:45	RB 6	Merstadt Lengen Darbu	Hochhausen	**10**

Zu welchen Zeiten fährt der IC 23 ab?

Auf welchem Gleis fährt die Linie RB 6 ab?

Was ist das Ziel der Linie S 16?

Mit welchen Linien kann man nach Merstadt fahren?

Welche Linie fährt über Begel?

Wie viele Linien fahren über Ratsdorf?

Hier siehst du einen Ausschnitt aus dem Einzelverbindungsnachweis der Handyrechnung von Ilka.

Datum/ Uhrzeit	Dauer (Std:Min:Sek)	Ziel
03.04.09, 13:47	00:10:14	0178 4322xxx
03.04.09, 14:40	00:44:32	02131 530xxx
03.04.09, 19:15	00:24:55	0174 6709xxx
03.04.09, 20:00	01:03:20	0211 7809xxx
03.04.09, 21:23	02:15:01	0174 6709xxx
04.04.09, 18:02	00:18:40	0178 4322xxx
04.04.09, 19:30	00:03:01	01577 632xxx
05.04.09, 10:33	00:12:36	02131 530xxx
06.04.09, 11:13	02:01:24	02137 6875xxx
06.04.09, 16:52	00:12:11	0178 4322xxx
06.04.09, 20:06	00:04:43	0800 334xxx
07.04.09, 08:01	00:18:12	0178 4322xxx
07.04.09, 15:11	00:20:03	030 786xxx
08.04.09, 12:05	01:03:29	0174 6709xxx

Wie häufig telefonierte Ilka mit der Nummer 01577 632xxx?

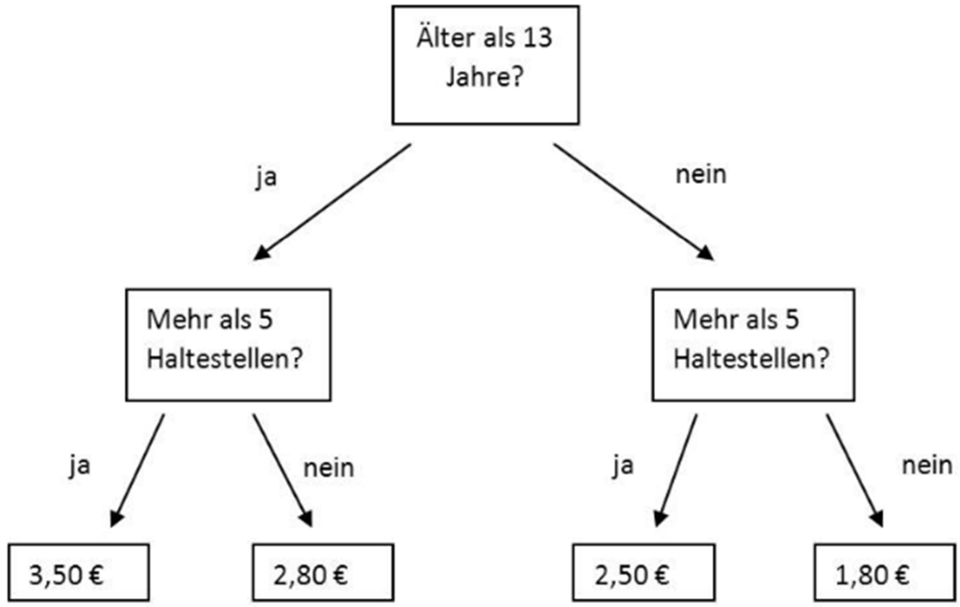
Wie lang dauerte Ilkas kürzestes Gespräch?

Wie lange dauerte das Gespräch mit der Nummer 030 786xxx am 07.04.09?

Welche Nummer rief Ilka am 04.04.09 um 18:02 an?

Wann fand das längste Gespräch mit der Nummer 0178 4322xxx statt?

Die untenstehende Übersicht verdeutlicht, wie sich die Busfahrpreise in Gibstadt zusammensetzen:



Tina ist 13 Jahre alt und fährt 4 Haltstellen. Wie hoch sind ihre Fahrtkosten?

Sven ist 16 Jahre alt und fährt 8 Haltestellen. Wie hoch sind seine Fahrtkosten?

Lukas ist 11 Jahre alt und fährt 5 Haltestellen. Wie hoch sind seine Fahrtkosten?

Der Preis für einen Kinofilm setzt sich folgendermaßen zusammen:

Man unterscheidet ob der Besucher älter oder jünger als 12 Jahre alt ist.

→ *12 oder jünger*: Grundpreis 2,50 €; *älter als 12*: Grundpreis 3 €

Man unterscheidet, ob der Film länger als 90 Minuten dauert.

→ 1 € extra, falls der *Film länger als 90 Minuten* dauert.

Man unterscheidet, ob der Film im 3-D-Kino läuft oder nicht.

→ 2 € extra, falls der *Film im 3-D-Kino* läuft.

Wie oft muss bei der Berechnung des Preises zwischen 2 Möglichkeiten entschieden werden?

Welches der beiden Merkmale „Überlange" oder „3D-Kino" erhöht den Kinopreis um einen höheren Betrag?

Was ist nach der oberen Übersicht der günstigste Eintrittspreis für einen Kinofilm?

Block: Problemlösen

Es wurden folgende Aufgaben zum Problemlösen aus PISA 2003 (OECD, 2004b) eingesetzt: Kinobesuch, Anschlusszüge, Gefrierschrank, Bewässerung, Bibliothekensystem (nur Teil 1), Studienplan, Urlaub. Zusätzlich wurde folgende Aufgabe verwendet:

Bei Familie Neustein steht eine Familienfeier bevor. Mutter Andrea hat Verwandte eingeladen, Vater Ingo das Essen bestellt. Nur die Sitzordnung bereitet den Neusteins noch Kopfzerbrechen. Denn Teile der Verwandtschaft sind zerstritten:

- Bei den Cousinen Susanne und Michelle herrscht gerade Zickenkrieg.
- Tante Frieda und Onkel Peter sind seit Kindertagen zerstritten.
- Mandy, Michelle und Ingo können sich nicht leiden.

Um die Feier nicht zu stören, sollten an einem Tisch die Streithähne nicht direkt nebeneinander oder direkt gegenüber sitzen. Auch Rücken an Rücken sollten diese Personen nicht sitzen.

Anderseits möchten bestimmte Personen gerne direkt nebeneinander, gegenüber oder schräg gegenüber sitzen, um sich gut unterhalten zu können:

- Paul, Mandy und Frieda wollen über einen möglichen gemeinsamen Urlaub sprechen.
- Andrea und ihr Bruder Heinz haben sich lange nicht gesehen und möchten nebeneinander sitzen.
- Die Kinder Jason, Kevin und Eva möchten gerne zusammen sitzen.

Es steht fest, dass die Tische folgendermaßen stehen müssen:

Tisch 1

1	2	3
4	5	6

Tisch 2

7	8	9
10	11	12

Mache einen Vorschlag für eine Verteilung der Neusteins auf die beiden Tische, d. h. gib eine mögliche Sitzordnung an.

Tisch 1	-	-	-
	-	-	-
Tisch 2	-	-	-
	-	-	-

Anhang H. Experiment 1: Posttest Mathematik

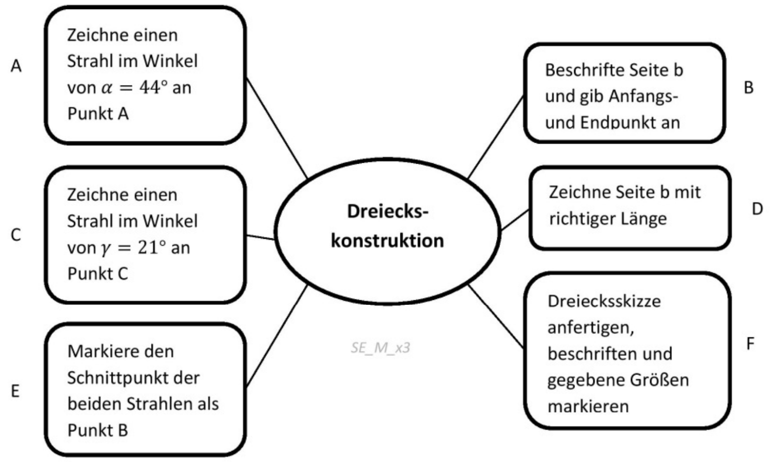
Die Screenshots wurden aus Druckgründen auf eine Seitenbreite von 14 cm skaliert, was ungefähr 2/3 der Originalgröße entspricht. Aus Platzgründen kennzeichnet ein schwarzer Rahmen jeweils eine Bildschirmseite, damit mehrere Screenshots auf einer Seite dokumentiert werden können.

Block: Planungsfähigkeit

Auf den folgenden Seiten findest du je eine mathematische Aufgabe. Du sollst das Ergebnis **nicht** ausrechnen. Stattdessen findest du die zugehörigen Lösungsschritte in Form einer Mindmap. *Deine Aufgabe ist es die jeweiligen Teilschritte zur Lösung der mathematischen Aufgabe, in eine sinnvolle Reihenfolge zu bringen. Wenn du glaubst, B ist der erste Teilschritt, dann trage B in das Feld hinter „1. Schritt" ein, usw.*

A6: Konstruiere ein Dreieck ABC aus folgenden Angaben: b = 4,8 cm, α = 44°, γ = 21°.

Als Hilfestellung für die Lösung der Aufgabe erhaltet ihr folgende Mindmap mit Teilschritten:



Deine Aufgabe ist es, die in der Mindmap angegebenen Teilschritte in eine sinnvolle Reihenfolge zu bringen. Trage dazu den Buchstaben des entsprechenden Teilschritts ein:

1. Schritt: ______

2. Schritt: ______

3. Schritt: ______

4. Schritt: ______

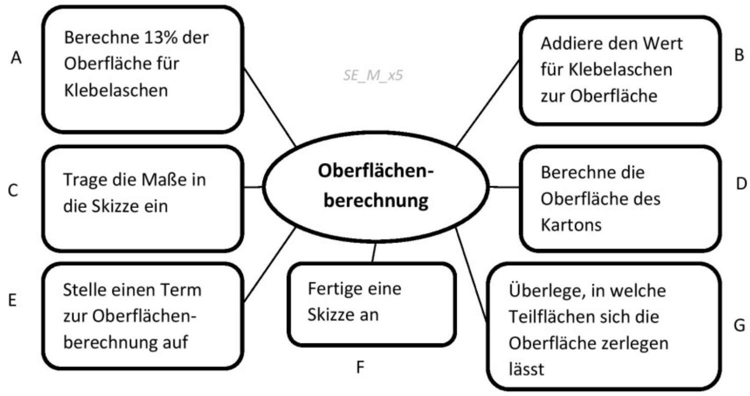
5. Schritt: ______

6. Schritt: ______

Aus Platzgründen und um Redundanz zu vermeiden, wird bei den nachfolgenden analogen Aufgaben nur noch der Aufgabenstamm samt Mindmap abgedruckt.

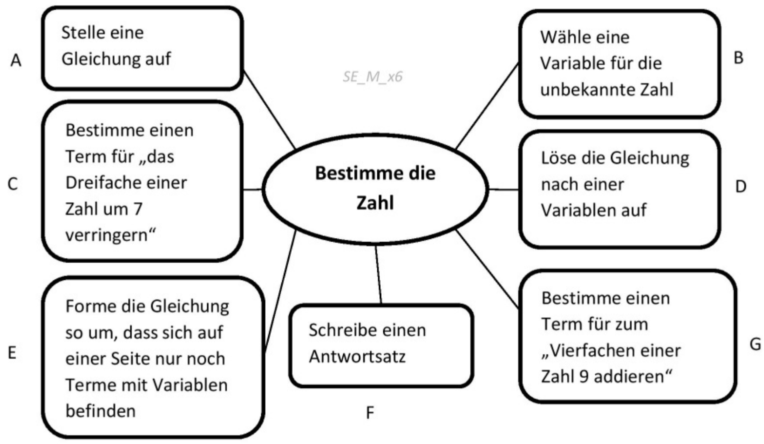
A7: Um eine rohrförmige Verpackung aus Pappe mit 18 cm Länge und 6 cm Durchmesser herzustellen, braucht man zusätzlich 13 Prozent Material für Klebelaschen. Wie hoch ist der Bedarf an Pappe insgesamt?

Als Hilfestellung für die Lösung der Aufgabe erhaltet ihr folgende Mindmap mit Teilschritten:



A8: Addiert man zum Vierfachen einer Zahl 9, so erhält, man dasselbe, als wenn man das Dreifache einer Zahl um 7 verringert. Welche Zahl ist das?

Als Hilfestellung für die Lösung der Aufgabe erhaltet ihr folgende Mindmap mit Teilschritten:



Block: Konditionales Wissen

Beim *mathematischen konditionalen Wissen* wurden drei Szenarien verwendet (Deutsches PISA-Konsortium, 2006), die inzwischen in einer weiter entwickelten Version publiziert sind (Lingel et al., 2014).

Block: Handlungswissen

A1: Längeneinheiten kann man wie folgt umrechnen:

1 km	=	1000 m
1 m	=	10 dm
	=	100 cm
	=	1000 mm
1 dm	=	10 cm

HW_M_x16

Rechne in die angegebene Einheit um!

a) 52 km sind m.

b) 3 m sind mm.

c) 32 dm sind cm.

A2: Gegeben sei die Funktion $f(x) = 0{,}5x - 3$. Welche Werte nimmt die Funktion bei folgenden Werten an?

HW_M_x8

a) $x = 24$

Antwort:

b) $x = 7$

Antwort:

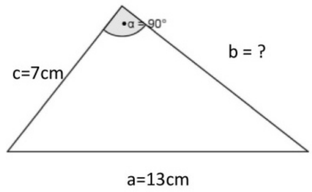
A3: Eine viereckige Design-Kaffeetasse hat eine rechteckige Grundfläche mit den Kantenlängen $a = 7\ cm$ und $b = 3\ cm$. Die maximale Füllhöhe beträgt $h = 12\ cm$. Berechne das Volumen der Tasse in cl (Zentilitern) mit den angegebenen Formeln.

Antwort:

$V = a \cdot b \cdot h$, $1\ \mathrm{cl} = 10\ \mathrm{cm}^3$

HW_M_x3

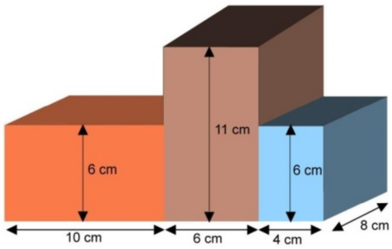
A4: Berechne die Länge der fehlenden Seite! Runde auf eine Dezimalstelle. Zur Erinnerung findest du unten die dafür notwendige Formel. HW_M_x13_2



Es gilt: $a^2=b^2+c^2$

Antwort:___________________________

A5: In der folgenden Aufgabe geht es um Berechnung von Flächeninhalt, Oberflächeninhalt und Rauminhalt (Volumen). Die dafür benötigten Formeln sind unter den Aufgaben angegeben.



a) Wie groß ist die Grundfläche („Bodenfläche“) dieses Körpers?

Antwort:___________________________ cm^2 HW_M_x17_1

b) Wie groß ist die Oberfläche des mittleren Quaders?

Antwort:___________________________cm^2 HW_M_x17_2

c) Welches Volumen hat der linke Quader?

Antwort:___________________________cm^3 HW_M_x17_3

Quader: Volumen $V = a \cdot b \cdot c$, Oberfläche $O = 2 \cdot (a \cdot b + a \cdot c + b \cdot c)$

Rechteckfläche: $F = a \cdot b$

Block: Problemlösen

Es wurden folgende Aufgaben aus PISA 2003 OECD (2004a) eingesetzt: Auswahl, Testergebnisse, Skateboard, Die dritte Seite, Das beste Auto, Autofahrt, Zimmermann, Lotterien, Sportplatz.

Anhang I. Experiment 1: Deskriptivstatistik

Tabelle I-1. Deskriptivstatistik: Punktzahlen im Posttest zum fächerübergreifenden Problemlösen in Experiment 1

Skala	*Bedingung*	*Schulform*	*M*	*SD*	*N*
Planungsfähigkeit	Experimentalgruppe	Hauptschule	18.25	10.28	28
		Gymnasium	24.51	9.14	43
(maximal 42 Punkte)		Gesamt	22.04	10.02	71
	Kontrollgruppe	Hauptschule	23.65	7.27	31
		Gymnasium	28.15	4.54	41
		Gesamt	26.21	6.24	72
	Gesamt	Hauptschule	21.08	9.16	59
		Gymnasium	26.29	7.45	84
		Gesamt	24.14	8.56	143
Konditionales Wissen	Experimentalgruppe	Hauptschule	11.11	6.77	28
		Gymnasium	15.72	4.83	43
(maximal 29 Punkte)		Gesamt	13.90	6.07	71
	Kontrollgruppe	Hauptschule	12.58	6.94	31
		Gymnasium	16.15	5,25	41
		Gesamt	14.61	6.25	72
	Gesamt	Hauptschule	11.88	6.84	59
		Gymnasium	15.93	5.01	84
		Gesamt	14.26	6.15	143
Handlungswissen	Experimentalgruppe	Hauptschule	9.50	5.58	28
		Gymnasium	15.95	5.41	43
(maximal 27 Punkte)		Gesamt	13.41	6.30	71
	Kontrollgruppe	Hauptschule	14.10	5.83	31
		Gymnasium	18.88	4.59	41
		Gesamt	16.82	5.65	72
	Gesamt	Hauptschule	11.92	6.12	59
		Gymnasium	17.38	5.21	84
		Gesamt	15.13	6.20	143
PISA 2003	Experimentalgruppe	Hauptschule	1.07	1.15	28
		Gymnasium	3.05	2.25	43
(maximal 9 Punkte)		Gesamt	2.27	2.12	71
	Kontrollgruppe	Hauptschule	2.94	2.45	31
		Gymnasium	3.95	1.99	41
		Gesamt	3.51	2.24	72
	Gesamt	Hauptschule	2.05	2.14	59
		Gymnasium	3.49	2.16	84
		Gesamt	2.90	2.26	143

Tabelle I-2. Bearbeitungszeit in Minuten nach Bedingung (EG = Experimentalgruppe, KG = Kontrollgruppe) und Schulform bei den fächerübergreifenden Tests

Test	*Bedingung*	*Schulform*	*M*	*SD*	*N*
Planungsfähigkeit	EG	Hauptschule	3.71	2.42	28
		Gymnasium	4.87	2.00	43
		Gesamt	4.41	2.23	71
	KG	Hauptschule	7.57	3.30	31
		Gymnasium	6.52	1.66	41
		Gesamt	6.97	2.54	72
	Gesamt	Hauptschule	5.74	3.49	59
		Gymnasium	5.67	2.01	84
		Gesamt	5.70	2.71	143
Handlungswissen	EG	Hauptschule	6.07	3.17	28
		Gymnasium	7.24	2.35	43
		Gesamt	6.78	2.74	71
	KG	Hauptschule	10.60	3.56	31
		Gymnasium	9.72	2.97	41
		Gesamt	10.10	3.24	72
	Gesamt	Hauptschule	8.45	4.05	59
		Gymnasium	8.45	2.93	84
		Gesamt	8.45	3.42	143
Konditionales Wissen	EG	Hauptschule	1.85	1.02	28
		Gymnasium	2.46	0.82	43
		Gesamt	2.22	0.94	71
	KG	Hauptschule	3.57	1.13	31
		Gymnasium	3.87	1.48	41
		Gesamt	3.74	1.34	72
	Gesamt	Hauptschule	2.75	1.38	59
		Gymnasium	3.15	1.38	84
		Gesamt	2.99	1.39	143
Problemlösen	EG	Hauptschule	4.12	3.83	28
		Gymnasium	8.33	4.70	43
		Gesamt	6.67	4.82	71
	KG	Hauptschule	12.40	5.97	31
		Gymnasium	16.09	6.90	41
		Gesamt	14.50	6.73	72
	Gesamt	Hauptschule	8.47	6.53	59
		Gymnasium	12.12	7.03	84
		Gesamt	10.61	7.04	143

Tabelle I-3. Deskriptivstatistik: Effizienz beim Posttest zum fächerübergreifenden Problemlösen in Experiment 1

Skala	*Bedingung*	*Schulform*	*M*	*SD*	*N*
Effizienz der Skala Handlungswissen	Experimentalgruppe	Hauptschule	1.47	0.72	28
		Gymnasium	2.37	1.04	42
		Gesamt	2.01	1.02	70
	Kontrollgruppe	Hauptschule	1.29	0.57	31
		Gymnasium	2.06	0.74	41
		Gesamt	1.73	0.77	72
	Gesamt	Hauptschule	1.37	0.65	59
		Gymnasium	2.22	0.91	83
		Gesamt	1.87	0.91	142
Effizienz der Skala Planungsfähigkeit	Experimentalgruppe	Hauptschule	5.68	4.43	28
		Gymnasium	5.40	2.90	42
		Gesamt	5.51	3.57	70
	Kontrollgruppe	Hauptschule	3.49	2.03	31
		Gymnasium	4.55	1.11	41
		Gesamt	4.10	1.65	72
	Gesamt	Hauptschule	4.53	3.53	59
		Gymnasium	4.98	2.24	83
		Gesamt	4.79	2.84	142
Effizienz der Skala PISA 2003	Experimentalgruppe	Hauptschule	0.35	0.59	28
		Gymnasium	0.33	0.24	42
		Gesamt	0.34	0.42	70
	Kontrollgruppe	Hauptschule	0.21	0.15	31
		Gymnasium	0.26	0.17	41
		Gesamt	0.24	0.16	72
	Gesamt	Hauptschule	0.28	0.43	59
		Gymnasium	0.30	0.21	83
		Gesamt	0.29	0.32	142
Effizienz der Skala Konditionales Wissen	Experimentalgruppe	Hauptschule	6.05	3.64	28
		Gymnasium	6.73	2.29	42
		Gesamt	6.46	2.90	70
	Kontrollgruppe	Hauptschule	3.51	1.72	31
		Gymnasium	4.54	1.65	41
		Gesamt	4.10	1.75	72
	Gesamt	Hauptschule	4.72	3.06	59
		Gymnasium	5.65	2.28	83
		Gesamt	5.26	2.66	142

Anmerkung. Für die Skalen zum mathematischen Problemlösen liegen keine Bearbeitungszeiten pro Schüler pro Skala vor (Gruppenerhebung mit Papier und Stift).

Tabelle I-4. Deskriptivstatistik: Punktzahlen im Posttest zum mathematischen Problemlösen in Experiment 1

Skala	*Bedingung*	*Schulform*	*M*	*SD*	*N*
Planungsfähigkeit	Experimentalgruppe	Hauptschule	27.43	14.84	28
		Gymnasium	37.33	16.55	43
(maximal 54 Punkte)		Gesamt	33.42	16.52	71
	Kontrollgruppe	Hauptschule	31.48	13.53	31
		Gymnasium	35.83	19.12	41
		Gesamt	33.96	16.97	72
	Gesamt	Hauptschule	29.56	14.19	59
		Gymnasium	36.60	17.75	84
		Gesamt	33.69	16.69	143
Handlungswissen	Experimentalgruppe	Hauptschule	3.00	2.29	28
		Gymnasium	6.81	2.15	43
(maximal 12 Punkte)		Gesamt	5.31	2.89	71
	Kontrollgruppe	Hauptschule	3.19	2.30	31
		Gymnasium	6.78	2.77	41
		Gesamt	5.24	3.12	72
	Gesamt	Hauptschule	3.10	2.28	59
		Gymnasium	6.80	2.46	84
		Gesamt	5.27	3.00	143
Konditionales Wissen	Experimentalgruppe	Hauptschule	4.86	2.65	28
		Gymnasium	7.23	2.92	43
(maximal 17 Punkte)		Gesamt	6.30	3.03	71
	Kontrollgruppe	Hauptschule	5.19	2.82	31
		Gymnasium	7.05	3.54	41
		Gesamt	6.25	3.36	72
	Gesamt	Hauptschule	5.03	2.72	59
		Gymnasium	7.14	3.22	84
		Gesamt	6.27	3.19	143
PISA 2003	Experimentalgruppe	Hauptschule	2.46	1.55	28
		Gymnasium	6.37	1.92	43
(maximal 14 Punkte)		Gesamt	4.83	2.61	71
	Kontrollgruppe	Hauptschule	2.74	1.86	31
		Gymnasium	6.29	2.30	41
		Gesamt	4.76	2.76	72
	Gesamt	Hauptschule	2.61	1.71	59
		Gymnasium	6.33	2.10	84
		Gesamt	4.80	2.68	143

Anhang J. Experiment 1: Korrelation der Leistungs- und Effizienzmaße

				Fächerübergreifende Skalen								*Mathematische Skalen*			
			Zeit	E-PLA	L-PLA	E-KW	L-KW	E-HW	L-HW	E-PISA	L-PISA	L-PLA	L-KW	L-HW	L-PISA
Fächerübergreifende Skalen	Zeit Posttest fächer-	*r*		-.274	.518	-.292	.301	-.144	.534	-.160	.593	-.062	.028	.010	.194
	übergreifende Skalen (Zeit)	*p*		.001	.000	.000	.000	.044	.000	.029	.000	.233	.373	.454	.011
	Effizienz	*r*	-.268		.267	.380	.080	.163	-.023	.547	.135	.054	.028	.110	.061
	Planungsfähigkeit (E-PLA)	*p*	.001		.001	.000	.174	.027	.393	.000	.055	.263	.373	.098	.238
	Leistung	*r*	.532	.266		.139	.470	.080	.478	.106	.453	.069	.160	.158	.263
	Planungsfähigkeit (L-PLA)	*p*	.000	.001		.050	.000	.172	.000	.105	.000	.209	.029	.031	.001
	Effizienz Konditionales	*r*	-.266	.381	.160		.474	.217	-.005	.141	-.004	-.028	.050	.199	.182
	Wissen (E-KW)	*p*	.001	.000	.028		.000	.005	.475	.047	.483	.371	.280	.009	.015
	Leistung Konditionales	*r*	.333	.082	.498	.485		.138	.292	.042	.293	.097	.156	.118	.226
	Wissen (L-KW)	*p*	.000	.165	.000	.000		.052	.000	.311	.000	.127	.032	.082	.003
	Effizienz	*r*	-.071	.162	.144	.252	.253		.379	.118	.054	.130	.004	.203	.290
	Handlungswissen (E-HW)	*p*	.201	.027	.043	.001	.001		.000	.082	.262	.063	.479	.008	.000
	Leistung	*r*	.539	-.005	.501	.055	.387	.480		.017	.489	-.005	.020	.162	.371
	Handlungswissen L-HW)	*p*	.000	.474	.000	.259	.000	.000		.420	.000	.475	.407	.027	.000
	Effizienz Problemlösen	*r*	-.151	.548	.114	.148	.057	.130	.040		.382	.055	-.026	.000	.155
	(E-PISA)	*p*	.036	.000	.088	.039	.250	.061	.318		.000	.260	.378	.499	.033
	Leistung Problemlösen	*r*	.596	.134	.481	.054	.394	.211	.577	.369		.042	.121	.025	.352
	(L-PISA)	*p*	.000	.055	.000	.260	.000	.006	.000	.000		.311	.077	.386	.000
Mathematische Skalen	Leistung Planungsfähigkeit	*r*	-.008	.046	.117	.000	.184	.210	.087	.060	.145		.128	.334	-.011
	(L-PLA)	*p*	.464	.293	.082	.498	.014	.006	.152	.239	.042		.066	.000	.448
	Leistung Konditionales	*r*	.078	.037	.211	.094	.255	.140	.161	-.005	.248	.198		.156	.171
	Wissen (L-KW)	*p*	.178	.332	.006	.132	.001	.047	.027	.475	.001	.009		.033	.021
	Leistung	*r*	.105	.101	.237	.237	.301	.390	.369	.033	.279	.416	.328		.300
	Handlungswissen (L-HW)	*p*	.106	.115	.002	.002	.000	.000	.000	.349	.000	.000	.000		.000
	Leistung Problemlösen	*r*	.246	.062	.316	.224	.389	.462	.525	.151	.519	.173	.351	.593	
	(L-PISA)	*p*	.002	.230	.000	.004	.000	.000	.000	.036	.000	.019	.000	.000	

Anmerkungen. Untere Diagonalmatrix: Interkorrelation der Skalen. Obere Diagonalmatrix: Partialkorrelation der Skalen bei Berücksichtigung der Skala *Figurenanalogien* (N2) des KFT-4-12+R (Heller & Perleth, 2000). Signifikante Korrelationen sind grau hinterlegt ($p < .05$, einseitige Signifikanztestung). *P*-Werte sind auf drei Nachkommastellen gerundet.

Anhang K. Tests of strategy knowledge: Critical aspects and need for further research

This appendix is based on an improved version of my internal report (Buchwald, 2014).[26]

Introduction

Tests of strategy knowledge are used with increasing frequency in educational psychology and science education (Artelt et al., 2009; Artelt et al., 2012; Lingel et al., 2014; Scherer & Tiemann, 2012; Schlagmüller & Schneider, 2007; Schlagmüller, Visé & Schneider, 2001; Shahat et al., 2013; Thillmann, 2007). However, problems concerning the validity, the construction and the scoring are seldom reported. In the following comment we will sketch the typical way of test construction, illustrate some problematic aspects with fictive examples and give a short outlook on possible ways to study these problems empirically.

Tests of strategy knowledge

The tests of strategy knowledge to be discussed consist of a short scenario description followed by a limited number of action possibilities (strategies) that have to be rated on a Likert Scale (Neuenhaus et al., 2010). The scoring is typically done by pairwise comparisons of the strategies. But not all possible $\frac{n(n-1)}{2}$ pairs of n strategies are used for the scoring. The scoring rules are based on those pairs of strategies resulting from a prior expert rating. If most of the experts agreed that strategy S_i is better than strategy S_j ($S_i < S_j$) or worse ($S_i > S_j$) the pair (S_i, S_j) is used as an item in the tests of strategy knowledge. The following scoring rules for each item are:

- Score 1 (agreement with the expert rating): If test taker's greater-lesser-relation (e. g. $S_i < S_j$) of his rating of two strategies i and j is equal to the expert rating (e. g. $S_i < S_j$).
- Score 0.5 (equal rating of strategies): If test taker's ratings of two strategies are equal ($S_i = S_j$). This rule is independent of the expert rating.

[26] Thanks to Prof. Dr. Joachim Wirth, Prof. Dr. Günter Daniel Rey, Markus Thielgen and Andreas Fischer who encouraged me to write some arguments down.

- Score 0 (disagreement with the expert rating): If test taker's greater-lesser-relation (e. g. $S_i < S_j$) of his rating of two strategies i and j is contrary to the expert rating (e. g. $S_i > S_j$).

The range of the resulting sum, the knowledge score (Artelt et al., 2009), is between zero (no agreement or concordance) and the number of used pairwise comparisons (perfect agreement or concordance). Therefore, the scale is sometimes constructed by calculating the mean performance, i. e. dividing the knowledge score by the number of used pairwise comparisons (Shahat et al., 2013).

Problematic aspects

The following discussion of problematic scoring aspects concerns especially those tests that used the scoring methods above. The order of the presented problems is arbitrary.

Problem 1: Agreement of experts and consequences for test development

Who counts as an expert in the domain that is tested – researchers, teachers? What qualifies them as experts in a domain (domain-specific knowledge, number of papers, years of experience, academic fame)? Are they really experts or just well-known collaborators or even some guys from the own lab? As in most surveys opinions differ or (technically spoken) there is variance. How much expert agreement is needed in order to use a specific pairwise comparison for the test (more than 50 percent, more than 90 percent)? How many pairwise comparisons are discarded? Does the specific paper address anything about these questions?

Problem 2: Deviation of the rated level between experts and test takers

Assume that the experts agreed that an answer option A is a very good strategy (mark = 1) and answer option B is a good strategy (mark = 2). Therefore the relationship A < B is used for pairwise comparison ("IF A < B THEN Score = Score + 1"; Schlagmüller & Schneider, 2007). Assume further that a test taker named Peter rated answer option A as a bad strategy (mark = 5) and answer option B as a very bad strategy (mark = 6). According to the scoring rule "IF A < B THEN Score = Score + 1" Peter receives one point regardless of his deviation concerning the *quality* of the strategies. To sum up, possible discrepancies concerning the level or quality of two strategies A and B are irrelevant for scoring as long as the relationship A < B (or A > B) is the base of the scoring algorithm. This can result in a possible loss of information. For example, Artelt et al. (2009) distinguishes theoretically between good, medium and bad strategy options, but their relative scoring – as described above – does not take these kind of absolute quality of a strategy into account (cf. problem 3).

Problem 3: Little differentiation of the answer quality between test takers

Assume, as in example 1, that the relationship “A better B” (A < B) is used for pairwise comparison (“IF A < B THEN Score = Score + 1) and that Peter rated answer option A as a bad strategy (mark = 5) and answer option B as a very bad strategy (mark = 6). Assume, Anne, another test taker, has rated answer A as a very good strategy (mark = 1) and answer option B as a good strategy (mark = 2). Both, Peter and Anne, will receive one point for the pair (A, B). The similarity to problem 2 and 3 is evident.

Problem 4: Consideration of potential equivalence of strategies

The scoring procedure, which only considers strategic pairs (A, B) of which the majority of experts opted for their superiority, considers the aspect of the potential equality of strategies. This point concerns the construct validity of the strategy knowledge test. They do not cover items of which an equality of strategies A and B is assumed, or they do not take them into account for the scoring. This has to be seen critically particularly when it comes to strategy knowledge in mathematical fields, since exceedingly there, different adequate solutions for sufficient problematic tasks exist (Bruder, 2000). Even in procedures which award 0.5 points per equality rating there is a potential validity problem: If the test taker rates two strategies A and B equally, this is done a) independently from the specific strategies or the expert rating and b) without regarding the level of quality (cf. problem 2).

Problem 5: Disability to decide or refusal of serious answering or regular answer

We assume that Peter rated all answer options of a scenario *S* as medium strategies (mark 3). Because of the equality of all ratings Tim receives half of the scores of the scenario *S* by marking one column. If a test is normed – like the WLST 7-12 (Schlagmüller & Schneider, 2007) – this is only a small problem. With regard to the content the question remains if Peter rated all strategies equally, (1) because he refused a serious answering or (2) because he answered seriously and was not able to take a decision about the different quality of the strategies (2a) or because he rated consciously for the equality of all available strategies (2b). In group settings, where the tests are mainly used, the question can hardly be answered. In individual test settings, where tests of strategies are – with the exceptional case of the WLST 7-12 – hardly administered, the diagnostician can use further indicators (observations, further inquiry).

Problem 6: Probability of guessing

If one has to judge if a score of x is a good or medium or weak performance in a specific test, regarding guessing probability or tabled norms (as in the case of the WLST 7-12; Schlagmüller & Schneider, 2007) or other criteria is useful.

If Peter reached a raw score of x in a specific strategy knowledge test, is the result coincidentally? Is he just guessing? When closed formats are used for testing purpose it is useful to regard scoring by chance or blind guessing. How high is the probability of reaching a score x by chance for tests of strategy knowledge? Which percentile rank corresponds a score x under the assumption of complete guessing? Only the WLST 7-12 mentioned the information about the probability of guessing (p = .50 per pairwise comparison; Schlagmüller & Schneider, 2007). One reason might be that in contrast to multiple-choice formats the use of pairwise comparisons, for which experts agree, do not make the probability of guessing very obvious. But it is computable. Assume, we want to compute the probability of guessing for the comparison of option A and B and use a x-point Likert scale. Then, there are x^2 possible outcomes, each with a probability of $\frac{1}{x^2}$. The probabilities of the three possible results $A = B$, $A > B$, $A < B$ are $P(A = B) = \frac{x}{x^2}$ and $P(A > B) = P(A < B) = \frac{(x^2-x)}{2} \cdot \frac{1}{x^2}$. Therefore the expected value E for one pairwise comparison based on the typical scoring is

$$E = \frac{(x^2 - x)}{2} \cdot \frac{1}{x^2} \cdot 1 + \frac{x}{x^2} \cdot 0.5 = 0.5$$

Therefore, under the assumption of complete guessing 50 percent of the maximum score is expected for this kind of tests with this scoring rule and independent from the experts rating. In case of the WLST 7-12, for example, with a score of 40 (50 percent of the maximum) the percentile rank is – dependent of gender and age – between 23 for male seven-graders to 1 for students in grade 11 and 12

(Schlagmüller & Schneider, 2007). This shows that the absolute reached score has to be handled with care, if there are no tabled norms available.

Problem 7: Testlet Effects?

> „A testlet is a cluster of items that share a common passage, scenario, or other context. These items might measure something in common beyond the trait measured by the test as a whole; if so, the model for the item responses should allow for this testlet trait. But modeling testlet effects that are negligible makes the model unnecessarily complicated and risks capitalization on chance, increasing the error in parameter estimates. Checking each testlet to see if the items within the testlet share something beyond the primary trait could therefore be useful“ (DeMars, 2012).

However, the scoring of the strategy knowledge tests ignores this potential source of method effects that could be interpreted as context effects. For an exceptional case with a similar test format that should assess cognitive styles see Bertling (2012), who used structural equation modeling to test testlet effects.

Problem 8: What is measured? Are the tests uni- or multidimensional?

Empirical results for situational judgement tests concerning the question of dimensionality are mixed and depend on the concrete test used (cf. McDaniel & Nguyen, 2001). Strategy knowledge is theoretical related to the domain-specific knowledge. Furthermore, reading competence and verbal intelligence seems to be measured to a certain degree.

Future research

To investigate, if the theoretical problematic aspects influence psychometric characteristics (e.g. reliability, concurrent or prognostic validity), empirical studies with real or simulated data are necessary.

Problem 1 could be investigated by comparing different samples of experts and the consequences for test takers test scores. Aspects of problems 2 and 3 could be studied by comparing the described scoring method with a scoring method that take the mean squared deviation into account. Comparing a scoring with 0.5 points per equal judgment of two options with one without the 0.5-rule can test aspects of problem 4. Furthermore, although several scoring methods for strategy knowledge tests and situational judgement tests are known (see McDaniel & Nguyen, 2001), empirical comparisons of the different approaches are needed. The identification of persons who marked only one column of one or more scenarios (problem 5) is possible by just counting the marked answers. However, as mentioned above, it remains unclear, why a person has marked this pattern. Testlet effects (problem 7) and potential multidimensionality (problem 8) can be investigated within the structural equation modelling framework.

Anhang L. Experiment 2: Experimentalgruppentreatment

Die Screenshots wurden aus Druckgründen auf eine Seitenbreite von 14 cm skaliert, was ungefähr 2/3 der Originalgröße entspricht. Aus Platzgründen kennzeichnet ein schwarzer Rahmen jeweils eine Bildschirmseite, damit mehrere Screenshots auf einer Seite dokumentiert werden können.

Übersicht

Herzlich willkommen zu den *Essener Selbstlernmodulen zum problemlösenden Denken* (E-SePro).

E-SePro besteht derzeit aus folgenden Teilen, die du der Reihe nach durchlaufen wirst.

Teil 1 (1. Stunde)

In der ersten Unterrichtsstunde erhältst du anhand von Beispielen und Aufgaben einen **Einblick in das spannende Feld des Problemlösens**

- Einführung in das Thema Problemlösen
- Gut geplant ist halb gelöst
- Wann mache ich was?
- Umgang mit Text, Abbildungen und Tabellen

Teil 2 (2. Stunde)

Im Anschluss an den ersten Teil möchten wir sehen, wie du dich bei ähnlichen Aufgaben schlägst.

- Test!

Die Aufgaben sind von unterschiedlichem Schwierigkeitsgrad. Versuche dennoch bitte alle Aufgaben zu bearbeiten.

Was ist ein Problem?

Ein Problem besteht aus

- einer *Ausgangssituation*,
- einem *Ziel*,
- und einem *nicht direkt offensichtlichen Lösungsweg*

Beim Problemlösen geht es also darum, ein Ziel zu erreichen. Man weiß jedoch nicht sofort, wie man dorthin kommt.

An dieser Stelle war ein Comic zu sehen (http://static.nichtlustig.de/toondb/031016.html).

Mit freundlicher Genehmigung von Bulls Press/©Joscha Sauer/Distr. Bulls.

Problemlösen, ein Prozess

- Problemlösen lässt sich als Prozess darstellen.
- Bitte sieh dir dazu die nebenstehende Abbildung gut an
- Jede der fünf Phasen des Problemlösens ist wichtig, wenn man ein Problem lösen möchte.
- Im Verlaufe von E-SePro wirst du dich mit den einzelnen Phasen des Problemlösens beschäftigen.
  - Dazu gibt es eine ganze Reihe verschiedener Aufgaben.

Wann ist ein Problem gelöst?

Ein Problem ist *gelöst*, wenn das vorgegebene Ziel erreicht ist.

Beispiel:

- Wenn du zum ersten Mal dein kaputtes Fahrrad ohne Hilfe reparieren möchtest, ist das für dich wahrscheinlich ein Problem.
- Dieses Problem ist gelöst, wenn du dein Fahrrad alleine wieder in einen fahrtüchtigen Zustand versetzt hast.

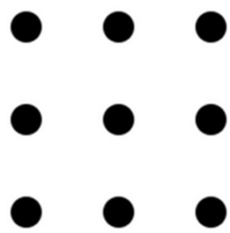
Im Rahmen dieses Lernprogramms üben wir den Umgang mit bestimmten Problemen. Diese Probleme haben meist einen Bezug zu Alltagssituationen oder sind allgemeine Knobelaufgaben.

Auf der nächsten Seite beginnen wir mit einer kleinen Knobelaufgabe.

Das Neun-Punkte-Problem

Aus diesem Beispiel kannst du Einiges über Probleme und Problemlösungen lernen. Man kann es so formulieren:

Versuche, die neun Punkte in nebenstehender Abbildung durch maximal vier gerade Linien zu verbinden, ohne dabei den Stift abzusetzen und eine Linie mehrfach zu ziehen.



Kennst du das Neun-Punkte-Problem bereits?

○ ja ○ nein

Glaubst du, dass man die neun Punkte auf die gewünschte Art verbinden kann?

○ ja ○ nein

Begründe bitte kurz, warum du glaubst bzw. nicht glaubst, dass man die neun Punkte wie gefordert verbinden kann.

Ausgangssituation, Mittel, Ziele

Neben dem sprachlichen Verständnis der Aufgabenstellung muss man versuchen zu verstehen,

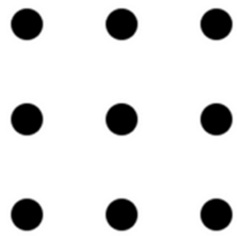
- worin, das Problem besteht,
- wie das Ziel bzw. die Lösung aussieht,
- welche Schritte zur Lösung erlaubt bzw. verboten sind.

Ohne dieses Verständnis ist es schwierig, ein Problem zu lösen.

Man muss sich also bewusst machen, wo man steht, wo man hin möchte und wie man dorthin kommt. Dazu ist es wichtig die Situation, die Regeln und Bedingungen sowie das Ziel zu kennen!

Daher üben wir jetzt die Bestimmung dieser Bestandteile eines Problems. Betrachte deshalb erneut das Neun-Punkte-Problem.

Versuche, die neun Punkte in nebenstehender Abbildung durch maximal vier gerade Linien zu verbinden, ohne dabei den Stift abzusetzen und eine Linie mehrfach zu ziehen.



Denke kurz über folgende Fragen nach:

- Worin besteht die Ausgangssituation?
- Was sind die Regeln und Einschränkungen?
- Was ist das Ziel?

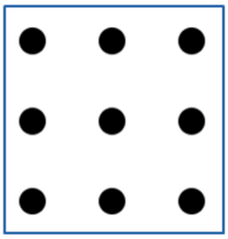
Regeln

Viele Menschen, die versuchen, das Neun-Punkte-Problem zu lösen, legen sich zusätzliche Beschränkungen auf, die sie am Finden einer Lösung hindern.

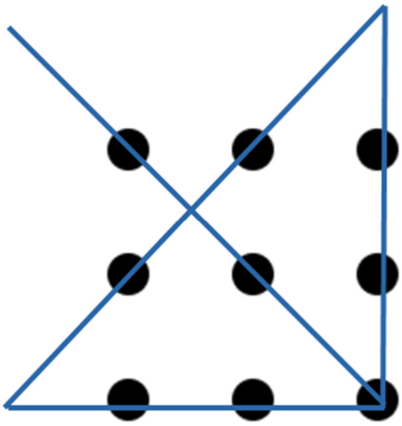
An dieser Stelle war ein Comic zusehen (http://static.nichtlustig.de/toondb/110914.html).

Selbst auferlegte Regeln

Viele Menschen, die versuchen das Neun-Punkte-Problem zu lösen, denken, sie dürften eine selbst auferlegte Box nicht überschreiten. Das wird in der Aufgabenstellung jedoch nicht gefordert.



Erst wenn man diese selbst auferlegte Regel verwirft, kann man eine Lösung des Neun-Punkte-Problems finden:



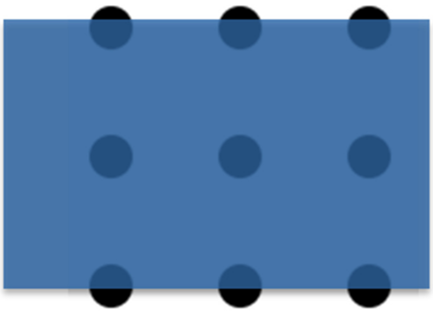
, bist du auf diese Lösung gekommen?

○ ja ○ nein

Selbst auferlegte Regeln

Jetzt kennen wir schon eine Lösung für das Neun-Punkte-Problem. Es gibt aber noch andere Lösungen.

Beispielsweise steht in der Aufgabenstellung nicht, wie breit der Stift sein darf. Daher kann man das Neun-Punkte-Problem sogar mit einer ganz, ganz dicken Linie lösen!



, bist du auf diese Lösung gekommen?

○ ja ○ nein

Zwischenfazit

Um ein Problem erfolgreich zu lösen, muss man sich zumindest klar machen,

- welche Ausgangssituation vorliegt und
- wie man diese Situation verändern kann und darf,
- um ein zuvor bestimmtes Ziel zu erreichen.

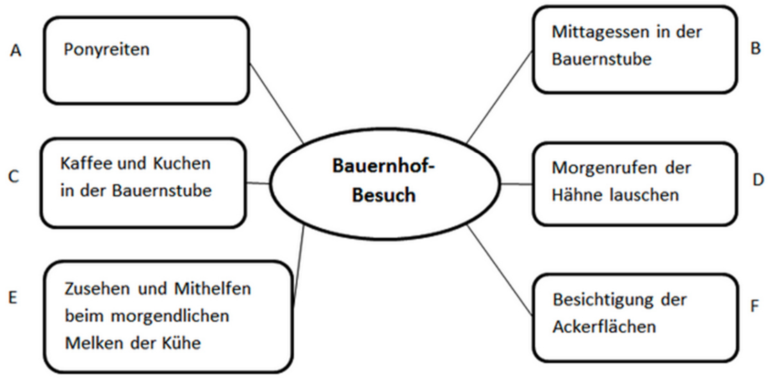
Block: Planungsfähigkeit

Beispielaufgabe

Bevor du die Übungsaufgaben bearbeitest, gehen wir ein Beispiel durch:

Beispiel

Eine Grundschulklasse möchte ihren Wandertag auf einem Bauernhof verbringen. Ein Brainstorming der Wünsche für den Besuch auf dem Bauernhof liefert folgende Vorschläge:



Deine Aufgabe ist es, die in der Mindmap angegebenen Teilschritte in eine sinnvolle Reihenfolge zu bringen. Ordne dazu den einzelnen Teilschritten einen Buchstaben zu.

	zugeordneter Buchstabe
1. Schritt:	-
2. Schritt:	-
3. Schritt:	-
4. Schritt:	-
5. Schritt:	-
6. Schritt:	-

Bei diesen Aufgaben kann es mehrere richtige Reihenfolgen geben. Wenn du dich an die Lösung solcher Aufgaben machst, denke bei der Bearbeitung über folgende Frage nach:

- Habe ich mir die Teilschritte gut durchgelesen?
- Was sollte man in der gegebenen Situation als Erstes tun?
- Was sollte man in der gegebenen Situation als Letztes tun?
- Habe ich wirklich pro Schritt einen Buchstaben zugeordnet oder eine Eingabe vergessen?
- Habe ich mehreren Schritten denselben Buchstaben zugeordnet?
- Habe ich am Ende noch einmal meine Schrittreihenfolge kontrolliert?

Wenn die Schülerinnen und Schüler beim ersten Versuch, die Planungsaufgaben zu lösen, keine optimale Lösung angegeben haben, erhielten Sie folgenden Hinweis:

Hinweise zur Lösung

Du hast noch keine optimale Reihenfolge gefunden.

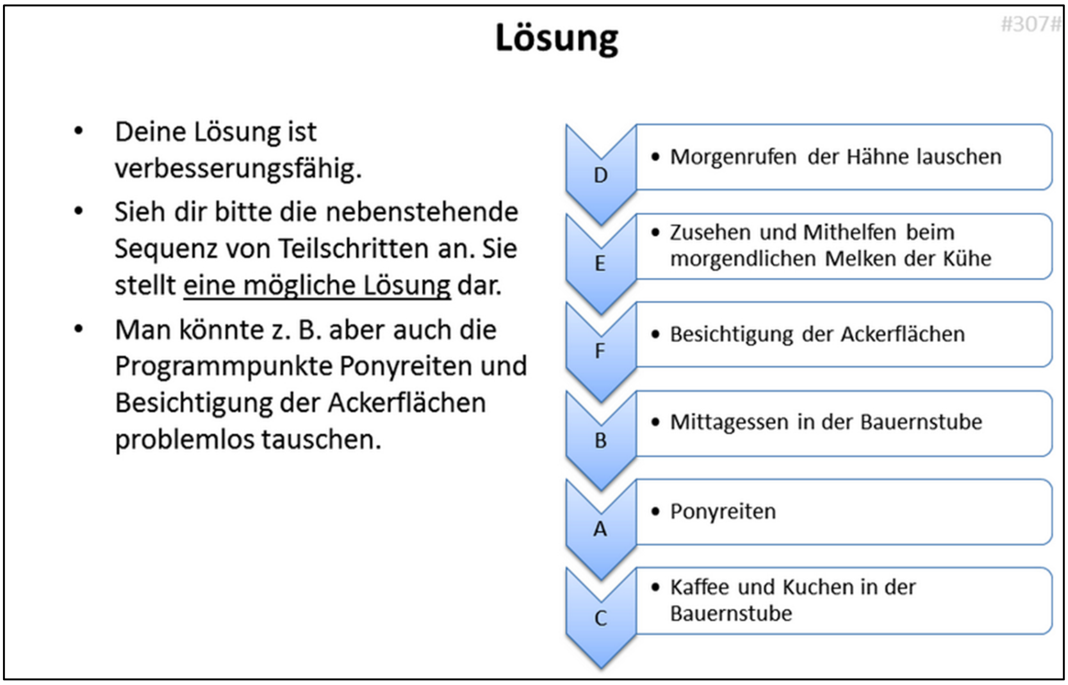
Bevor du einen neuen Lösungsvorschlag machst, denke bitte über folgende Fragen nach:

- Hast du wirklich pro Schritt einen Buchstaben zugeordnet oder eine Eingabe vergessen?
- Hast du mehreren Schritten denselben Buchstaben zugeordnet?
- Hast du am Ende noch einmal deine Schrittreihenfolge kontrolliert?
- Hast du dir die Teilschritte gut durchgelesen?
- Was sollte man in der gegebenen Situation als Erstes tun?
- Was sollte man in der gegebenen Situation als Letztes tun?

Es gibt bei allen Teilschritten bestimmte Beschränkungen in der Reihenfolge, auf die man leicht kommen kann.

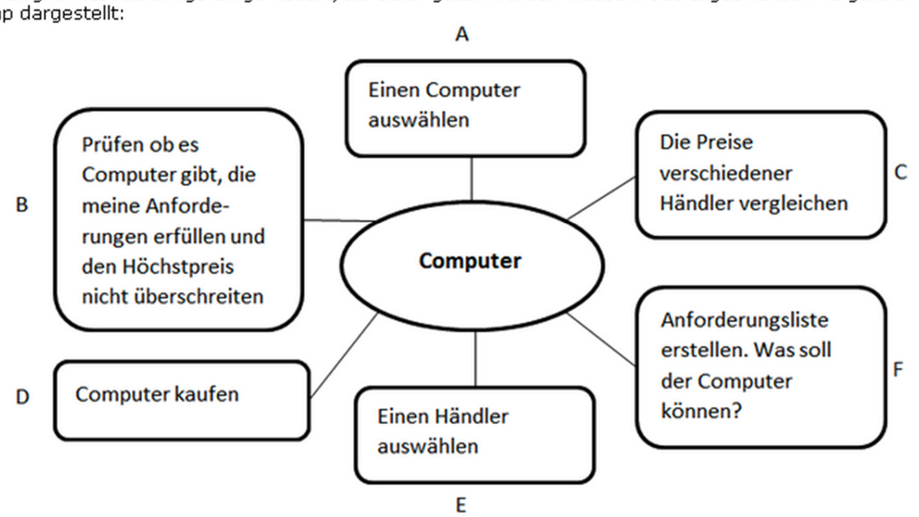
- Beispielsweise solltest du erst einen Computer auswählen, nachdem du eine Anforderungsliste erstellt hast.

Wenn Sie auch beim zweiten Versuch, keine optimale Lösung angegeben haben, wurde eine Lösung eingeblendet, z. B.:



Aus Platzgründen und um Redundanz zu vermeiden, wird bei den nachfolgenden analogen Aufgaben nur noch der Aufgabenstamm samt Mindmap sowie die Lösung abgedruckt.

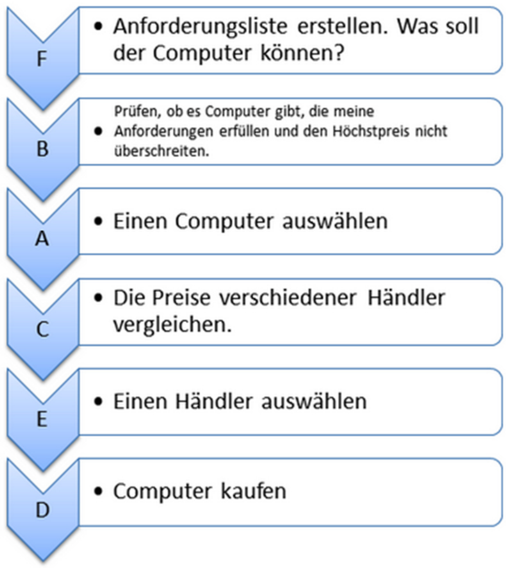
Du möchtest unbedingt einen neuen Computer haben. Deine Eltern haben zugestimmt den Computer zu kaufen, allerdings nur unter der Bedingung, dass du dich gut informierst und eine Entscheidung triffst, welches Gerät gekauft werden soll. Außerdem haben sie dir einen Höchstpreis genannt, den der Computer nicht übersteigen darf. Zu Beginn hast du einige Dinge notiert, die dafür getan werden müssen. Das Ergebnis ist in folgender Mindmap dargestellt:



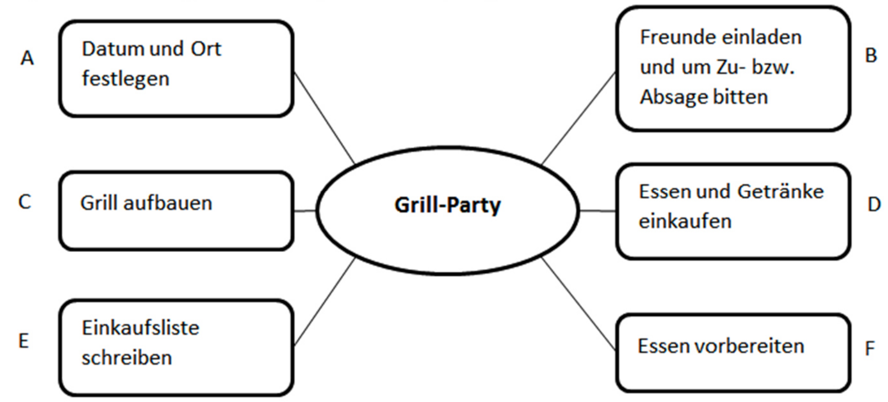
Lösung

#309#

- Deine Lösung ist verbesserungsfähig.
- Sieh dir bitte die nebenstehende Sequenz von Teilschritten an. Sie stellt eine mögliche Lösung dar.



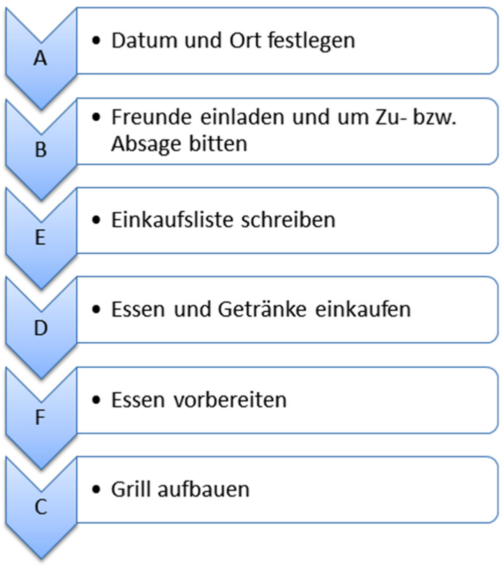
Du willst gemeinsam mit deinen Freunden eine Grill-Party veranstalten. Zu Beginn hast du mögliche Teilschritte des Vorgehens gesammelt. Das Ergebnis ist in folgender Mindmap dargestellt:



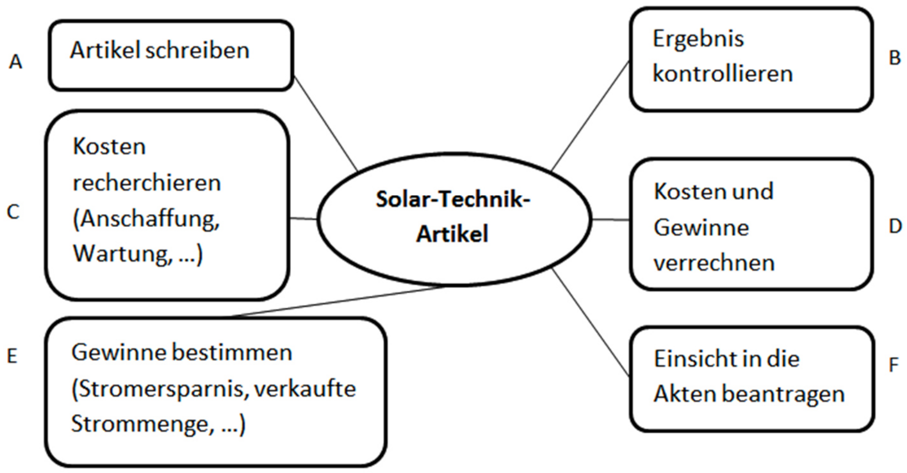
Lösung

#305#

- Deine Lösung ist verbesserungsfähig.
- Sieh dir bitte die nebenstehende Sequenz von Teilschritten an. Sie stellt <u>eine mögliche Lösung</u> dar.



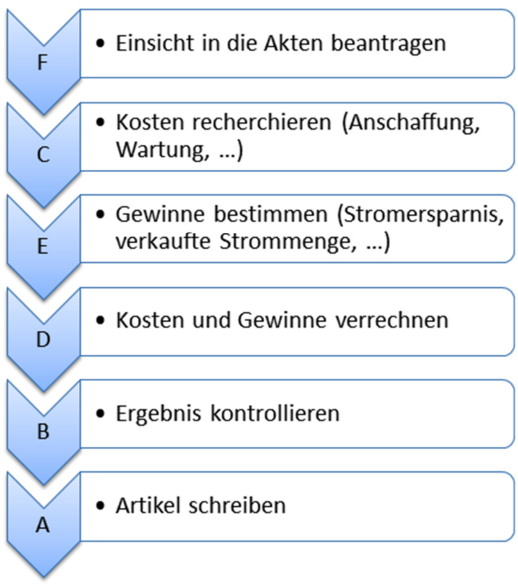
Lea will einen Artikel für die Schülerzeitung schreiben, der der Frage nachgeht, ob sich die vor fünf Jahren angeschaffte Solar-Technik an ihrer Schule finanziell gelohnt hat. Hilf ihr die Teilschritte ihres Vorgehens in eine sinnvolle Reihenfolge zu bringen:



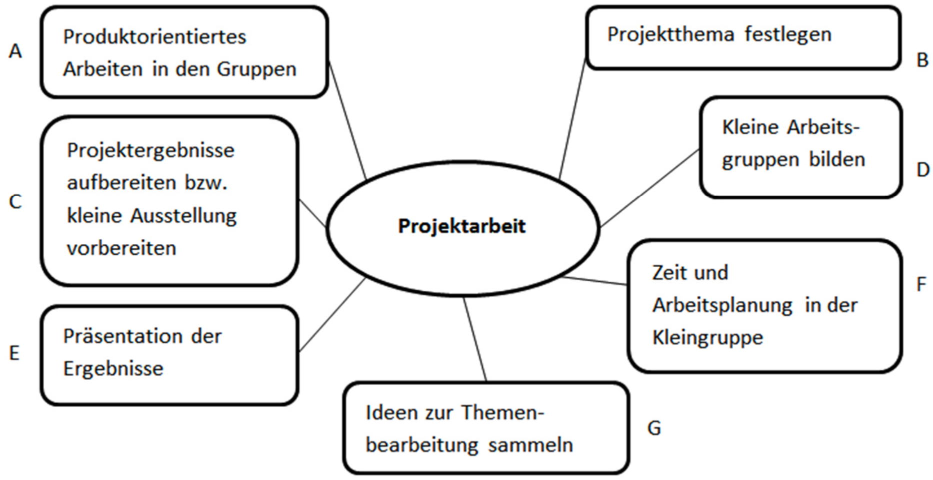
Lösung

#317#

- Deine Lösung ist verbesserungsfähig.
- Sieh dir bitte die nebenstehende Sequenz von Teilschritten an. Sie stellt eine mögliche Lösung dar.



„Ein Projekt zeichnet sich dadurch aus, daß eine Arbeitsgruppe ein vorgegebenes oder selbst gewähltes Arbeitsvorhaben schrittweise plant, organisiert, durchführt und auswertet. Die wichtigsten Arbeitsschritte sind im folgenden ungeordnet aufgeführt“ (Klippert, 2002).

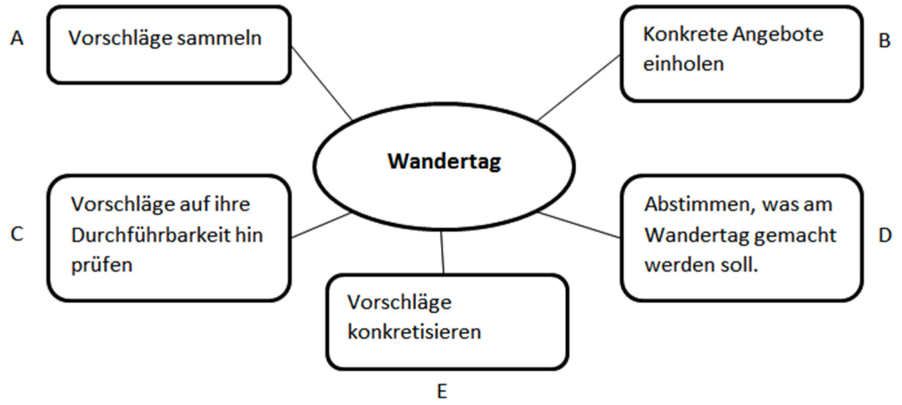


Lösung

#321#

- Deine Lösung ist verbesserungsfähig.
- Sieh dir bitte die nebenstehende Sequenz von Teilschritten an. Sie stellt eine mögliche Lösung dar.

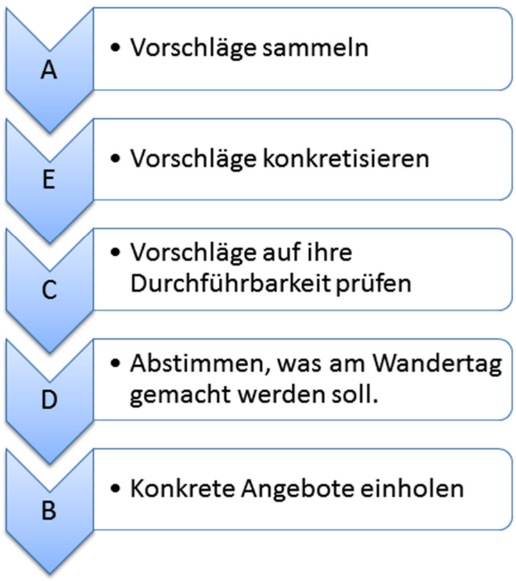
In eurer Klasse steht in vier Wochen der Wandertag bevor. Dieses Jahr sollt ihr Eigeninitiative zeigen und den begleitenden Lehrern eine Planung vorlegen. Zu Beginn habt ihr einige Dinge notiert, die dafür getan werden müssen. Das Ergebnis ist in folgender Mindmap dargestellt:



#313#

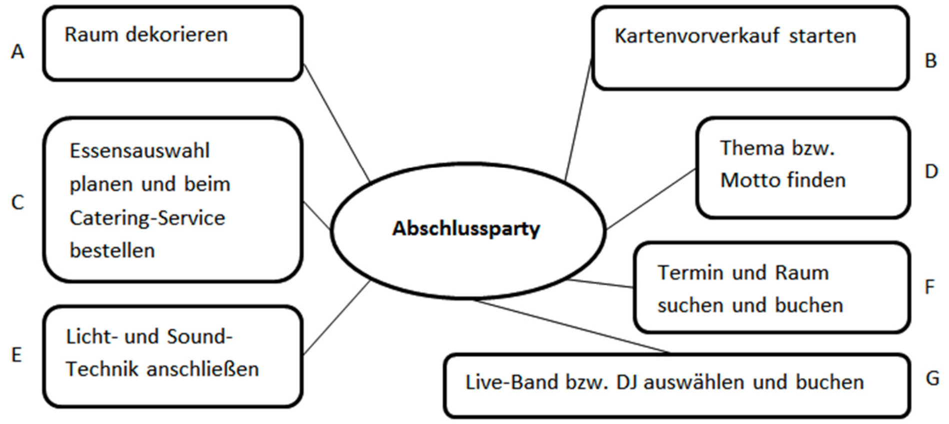
Lösung

- Deine Lösung ist verbesserungsfähig.
- Sieh dir bitte die nebenstehende Sequenz von Teilschritten an. Sie stellt eine mögliche Lösung dar.



Du bist im Organisationsteam der Abschlussparty deiner Jahrgangsstufe.

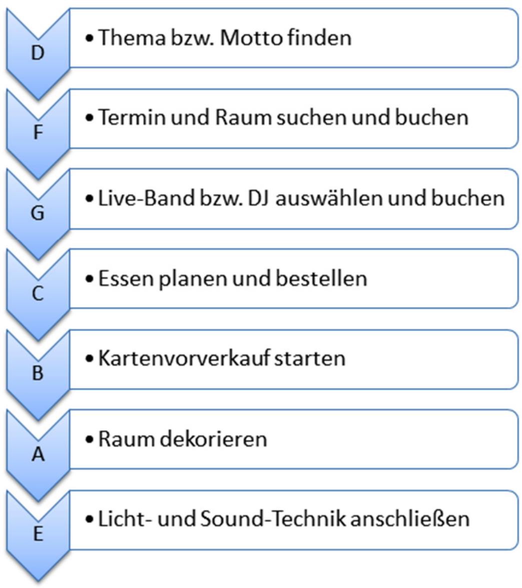
Zu Beginn habt ihr einige Dinge notiert, die zur Planung notwendig sind. Das Ergebnis ist in folgender Mindmap dargestellt:



#319#

Lösung

- Deine Lösung ist verbesserungsfähig.
- Sieh dir bitte die nebenstehende Sequenz von Teilschritten an. Sie stellt eine mögliche Lösung dar.



Zwischenfazit und Überleitung

Sehr schön, du hast einen weiteren Teil von E-SePro beendet.

Pläne sind hilfreich, um Probleme zu lösen. Oft gibt es aber mehr als einen Plan. Welchen Plan man umsetzen soll, ist jedoch nicht immer leicht zu entscheiden.

Beispiel: Stuttgart 21

- In Stuttgart will man seit längerem den Hauptbahnhof verbessern. Dazu gab es einen Plan der deutschen Bahn AG namens „Stuttgart 21", der einen Umbau des Kopfbahnhofs in einen Durchgangsbahnhof vorsieht. Die Gegner von „Stuttgart 21" lehnen dieses Projekt jedoch ab und schlugen vor, den bestehenden Kopfbahnhof zu erweitern.

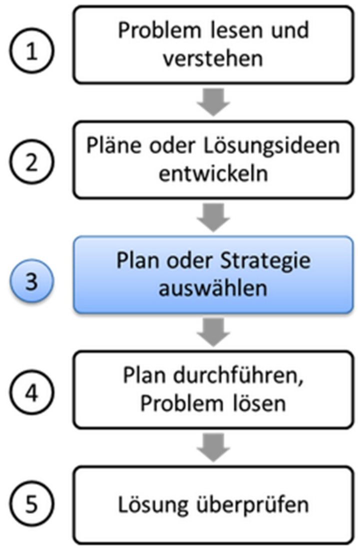
Block: Konditionales Wissen

Problemlösen, ein Prozess

Betrachte bitte erneut den Problemlöseprozess

In Phase 3 des Problemlöseprozesses musst du dich zunächst für eine Handlungsmöglichkeit entscheiden (siehe Abbildung).

Das bedeutet, dass du **Alternativen abwägen** und dich **begründet für einen Plan oder eine Strategie entscheiden** musst.



Wann mache ich was?

Wenn man vor einem Problem steht, weiß man manchmal nicht, welches Vorgehen zum Ziel führt. Dafür kann es verschiedene Gründe geben:

- Entweder man hat keine Idee
- oder man hat verschiedene Pläne, weiß aber nicht, welcher Plan der Beste ist.

Hier wollen wir uns mit der Wahl des Vorgehens beschäftigen, wenn man mehrere Optionen oder Pläne hat.

Aufgabenstellung

Im Folgenden werden verschiedene Situationen beschrieben, zu denen du verschiedene Handlungsmöglichkeiten bewerten sollst. Sieh dir die Situation und die Handlungsmöglichkeiten bitte genau an. Bewerte dann jede einzelne Handlungsmöglichkeit mit einer Schulnote (von 1 = „sehr gut“ bis 6 = „ungenügend“). Je besser eine Handlungsmöglichkeit deiner Meinung nach ist, umso besser sollte deine Benotung sein.

- "Gute" Handlungsmöglichkeiten, sollst du mit 1 oder 2 bewerten.
- "Mittelmäßige" Handlungsmöglichkeiten, sollst du mit 3 oder 4 bewerten.
- "Schlechte" Handlungsmöglichkeiten, sollst du mit 5 oder 6 bewerten.

Selbstverständlich kannst du bei deiner Bewertung die gleiche Schulnote mehrmals vergeben, wenn du Handlungsmöglichkeiten gleich gut findest.

Rückmeldungen

Wenn deine Antworten unvollständig sind oder deine Einschätzung der Handlungsmöglichkeiten deutlich von der Musterlösung abweichen, erhältst du einen 2. Lösungsversuch! Wenn du dann noch nicht die beste Lösung gefunden hast, zeigt dir E-SePro eine Musterlösung an. Dabei kommt folgendes Smiley-System zur Bewertung der jeweiligen Handlungsmöglichkeiten zum Einsatz:

	=	Diese Handlungsmöglichkeit ist gut, sinnvoll, angemessen.
	=	Diese Handlungsmöglichkeit ist unvollständig oder verbesserungsfähig.
	=	Diese Handlungsmöglichkeit ist ungeeignet oder unangebracht.

Beispielaufgabe

Bevor du die Übungsaufgaben bearbeitest, gehen wir ein Beispiel durch:

Beispiel

Du hast dir von deinem Weihnachtsgeld ein neues Regal in einem großen Möbelhaus gekauft. Nun möchtest du es alleine in deinem Zimmer zusammenbauen. Nachdem du schon fast alles zusammengebaut hast, stellst du fest, dass dir ein paar Kleinteile fehlen. Was tust du?

Bewerte die folgenden Handlungsmöglichkeiten, indem du pro Antwort eine Schulnote ankreuzt.

a) Ich fahre zurück zum Möbelhaus und reklamiere die Ware.	1 2 3 4 5 6
b) Ich kaufe einfach ein neues Regal.	1 2 3 4 5 6
c) Ich kaufe die Kleinteile nach.	1 2 3 4 5 6
d) Ich sehe nach, ob wir die benötigten Kleinteile irgendwo haben.	1 2 3 4 5 6

Lösung (vergleiche die Abbildung)

- a) und d) sind gute Handlungsmöglichkeiten. Daher solltest du sie mit 1 oder 2 bewerten.
- b) ist keine gute Idee. Daher solltest du sie mit 5 oder 6 bewerten.
- c) ist eine mittelmäßige Lösung. Daher solltest c) mit 3 oder 4 bewerten.

Einschätzung der Alternativen

a) Ich fahre zurück zum Möbelhaus und reklamiere die Ware.
Das ist die direkteste Lösungsmöglichkeit.

b) Ich kaufe einfach ein neues Regal.
Erneut Geld auszugeben, ohne das alte Regal zurückzugeben, ist eine schlechte Idee.

c) Ich kaufe die Kleinteile nach.
Wenn das Möbelhaus etwas weiter weg ist, ist das eine zeitsparende Möglichkeit. Aber es kostet zusätzlich Geld.

d) Ich sehe nach, ob wir die benötigten Kleinteile irgendwo haben.
Das ist in der Regel eine zeit- und kostengünstige Möglichkeit.

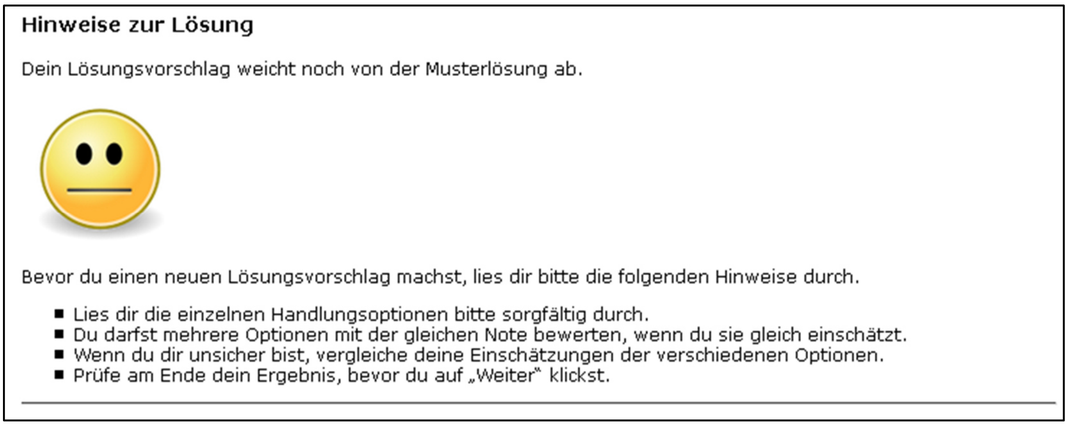
Hinweis: Eine oder mehrere Fragen sind für den weiteren Verlauf des Fragebogens wichtig.

Du musst eine schriftliche Bewerbung für ein Praktikum schreiben. Wie gehst du vor?
Bewerte die folgenden Möglichkeiten, indem du pro Antwort eine Schulnote ankreuzt.

	1	2	3	4	5	6
Ich suche im Internet nach Bewerbungsunterlagen und orientiere mich daran.	○	○	○	○	○	○
Ich orientiere mich an in der Schule besprochenen Bewerbungsunterlagen.	○	○	○	○	○	○
Ich schreibe ohne Vorlage eine Bewerbung, so wie ich es für richtig halte.	○	○	○	○	○	○
Ich bitte meine Eltern eine Bewerbung für mich zu schreiben.	○	○	○	○	○	○

Weiter

Anmerkung. Der Hinweis „Eine oder mehrere Fragen sind für den Verlauf des Fragebogens wichtig“ wurde eingeblendet, wenn auf „Weiter“ geklickt worden war, bevor alle Fragen einer Seite beantwortet worden waren.



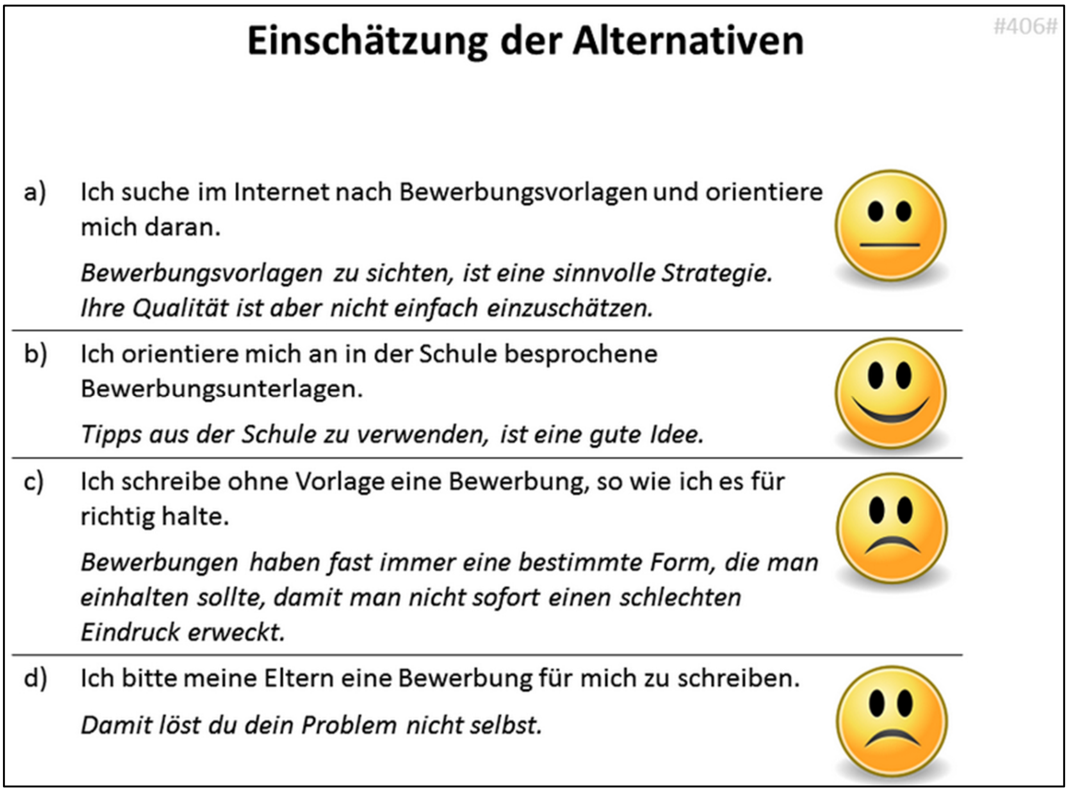
Hinweise zur Lösung

Dein Lösungsvorschlag weicht noch von der Musterlösung ab.

Bevor du einen neuen Lösungsvorschlag machst, lies dir bitte die folgenden Hinweise durch.

- Lies dir die einzelnen Handlungsoptionen bitte sorgfältig durch.
- Du darfst mehrere Optionen mit der gleichen Note bewerten, wenn du sie gleich einschätzt.
- Wenn du dir unsicher bist, vergleiche deine Einschätzungen der verschiedenen Optionen.
- Prüfe am Ende dein Ergebnis, bevor du auf „Weiter" klickst.

Anmerkung. Diese Seite erschien, wenn eine optimale Lösung einer Aufgabe zum konditionalen Wissen beim ersten Lösungsversuch nicht erreicht wurde. Sie wird im Folgenden nicht erneut abgedruckt.



#406#

Einschätzung der Alternativen

a) Ich suche im Internet nach Bewerbungsvorlagen und orientiere mich daran.
Bewerbungsvorlagen zu sichten, ist eine sinnvolle Strategie. Ihre Qualität ist aber nicht einfach einzuschätzen.

b) Ich orientiere mich an in der Schule besprochene Bewerbungsunterlagen.
Tipps aus der Schule zu verwenden, ist eine gute Idee.

c) Ich schreibe ohne Vorlage eine Bewerbung, so wie ich es für richtig halte.
Bewerbungen haben fast immer eine bestimmte Form, die man einhalten sollte, damit man nicht sofort einen schlechten Eindruck erweckt.

d) Ich bitte meine Eltern eine Bewerbung für mich zu schreiben.
Damit löst du dein Problem nicht selbst.

Anmerkung. Eine Seite mit einer Einschätzung der Handlungsalternativen wurde eingeblendet, wenn eine optimale Lösung einer Aufgabe zum konditionalen Wissen auch beim zweiten Lösungsversuch nicht erreicht wurde.

Jonas überlegt, sich ein Smartphone zuzulegen. Dazu möchte er die aktuelle Version der Betriebssysteme Android und iOS vergleichen. Was könnte er tun?
Bewerte die folgenden Möglichkeiten, indem du pro Antwort eine Schulnote ankreuzt.

	1	2	3	4	5	6
Jonas geht in ein Geschäft und probiert Smartphones mit beiden Betriebssystemen aus.	○	○	○	○	○	○
Jonas informiert sich im Internet.	○	○	○	○	○	○
Jonas fragt im Bekanntenkreis nach, ob jemand Erfahrung mit einem der beiden Betriebssysteme hat.	○	○	○	○	○	○
Jonas lässt sich im Fachhandel beraten.	○	○	○	○	○	○
Jonas sucht in Fachzeitschriften nach einem Vergleichsartikel	○	○	○	○	○	○

#410#

Einschätzung der Alternativen

a) Jonas geht in ein Geschäft und probiert Smartphones mit beiden Betriebssystemen aus.
Der aktive Vergleich der Alternativen ist eine gute Idee.

b) Jonas informiert sich im Internet.
Selbstständige Lösungsversuche sind immer eine gute Idee.

c) Jonas fragt im Bekanntenkreis nach, ob jemand Erfahrung mit einem der beiden Betriebssysteme hat.
Das kann ein guter Startpunkt für die Entscheidungsfindung sein.

d) Jonas lässt sich im Fachhandel beraten.
Fachhändler wollen Produkte verkaufen. Daher ist ihre Objektivität manchmal fragwürdig.

e) Jonas sucht in Fachzeitschriften nach einem Vergleichsartikel.
Fachzeitschriften sind eine gute Informationsgrundlage.

Die Lösung einer komplizierten Hausaufgabe erfordert mehrere Schritte. Bei einem dieser Schritte kommst du nicht weiter. Wie gehst du vor?
Bewerte die folgenden Möglichkeiten, indem du pro Antwort eine Schulnote ankreuzt.

	1	2	3	4	5	6
Ich suche nach einer Beispielaufgabe, bei der ich diesen Schritt nachvollziehen kann.	○	○	○	○	○	○
Ich denke darüber nach, ob andere Möglichkeiten bestehen die Aufgabe zu lösen.	○	○	○	○	○	○
Um keine Zeit zu verlieren, fange ich sofort mit der nächsten Aufgabe an.	○	○	○	○	○	○
Ich frage jemanden, ob er mir bei der Lösung helfen kann.	○	○	○	○	○	○

Weiter

#405#

Einschätzung der Alternativen

a) Ich suche nach einer Beispielaufgabe, bei der ich diesen Schritt nachvollziehen kann.

Bekannte Lösungen zu verwenden, ist eine sinnvolle Strategie.

b) Ich denke darüber nach, ob andere Möglichkeiten bestehen die Aufgabe zu lösen.

Selbstständige Lösungsversuche sind eine gute Idee.

c) Um keine Zeit zu verlieren, fange ich sofort mit der nächsten Aufgabe an.

Sofort aufzugeben, hilft bei der Lösung von Problemen nicht weiter.

d) Ich frage jemanden, ob er mir bei der Lösung helfen kann.

Wenn du in einer angemessenen Zeit nicht selbst auf eine Lösungsidee kommst, ist das eine Möglichkeit.

Johanna hat den Verdacht, dass ihr W-LAN-Router nicht richtig funktioniert. Denn seit einer Stunde kommt sie mit ihrem Laptop nicht ins Internet. Was könnte sie sinnvollerweise als nächstes tun?
Bewerte die folgenden Möglichkeiten, indem du pro Antwort eine Schulnote ankreuzt.

	1	2	3	4	5	6
Einen neuen Router kaufen.	○	○	○	○	○	○
Versuchen mit einem anderen Gerät per W-LAN ins Internet zu gelangen.	○	○	○	○	○	○
Neustarten des W-LAN-Routers.	○	○	○	○	○	○
Abwarten. Das Problem löst sich möglicherweise von selbst.	○	○	○	○	○	○
Den Wartungsdienst anrufen, um einen Techniker zu bestellen.	○	○	○	○	○	○
Dem Laptop gut zureden, damit er sich mit dem W-LAN verbindet.	○	○	○	○	○	○

#411#

Einschätzung der Alternativen

a) Einen neuen Router kaufen.

Das ist als Sofortmaßnahme übertrieben.

b) Versuchen mit einem anderen Gerät per W-LAN ins Internet zu gelangen.

Damit kann man prüfen, ob das Problem möglicherweise am Laptop und nicht am W-LAN-Router liegt.

c) Neustarten des W-LAN-Routers.

Das ist ein sinnvoller und naheliegender Lösungsversuch.

d) Abwarten. Das Problem löst sich möglicherweise von selbst.

Das klingt seltsam, hilft aber in diesem Fall manchmal tatsächlich.

e) Den Wartungsdienst anrufen, um einen Techniker zu bestellen.

Das ist als Sofortmaßnahme übertrieben.

f) Dem Laptop gut zureden, damit er sich mit dem W-LAN verbindet.

Das hilft nun wirklich nicht.

Pia darf in der kommenden Woche zum ersten Mal wählen. Sie ist politisch bislang wenig interessiert. Weil sie jedoch wählen möchte, überlegt sie, was sie tun sollte.
Bewerte die folgenden Möglichkeiten, indem du pro Antwort eine Schulnote ankreuzt.

	1	2	3	4	5	6
Da eine Woche nicht ausreicht, um sich politisch zu informieren, sollte sie überhaupt nicht wählen.	○	○	○	○	○	○
Sie fragt ihre Eltern, was diese wählen, und wählt dasselbe.	○	○	○	○	○	○
Sie informiert sich im Internet auf den Webseiten der Parteien über Kandidaten und Inhalte, bevor sie wählt.	○	○	○	○	○	○
Sie sieht sich die Wahlplakate an und wählt den sympathischeren Kandidaten und seine Partei.	○	○	○	○	○	○
Da eine Woche nicht ausreicht, um sich politisch zu informieren, macht sie einfach zufällig ihre Kreuze in der Wahlkabine.	○	○	○	○	○	○
Sie nutzt den Wahl-O-Mat, ein Computerprogramm das zeigt, „welche zu einer Wahl zugelassene Partei der eigenen politischen Position am nächsten steht" und wählt dem Ergebnis entsprechend.	○	○	○	○	○	○
Sie besucht die Informationsstände der Parteien in der Fussgängerzone.	○	○	○	○	○	○

Einschätzung der Alternativen

a) Da eine Woche nicht ausreicht, um sich politisch zu informieren, sollte sie überhaupt nicht wählen.

Nicht zu wählen, ist keine gute Option.

b) Sie fragt ihre Eltern, was diese wählen, und wählt dasselbe.

Das ist keine gute Option, da sie nicht über ihre Wahl nachdenkt.

c) Sie informiert sich im Internet auf den Webseiten der Parteien über Kandidaten und Inhalte, bevor sie wählt.

Selbstständige Lösungsversuche sind eine gute Idee.

d) Sie sieht sich die Wahlplakate an und wählt den sympathischeren Kandidaten und seine Partei.

Sympathie ist zwar häufig ein Kriterium, aber weder notwendig noch hinreichend für politischen Erfolg.

e) Da eine Woche nicht ausreicht, um sich politisch zu informieren, macht sie einfach zufällig ihre Kreuze in der Wahlkabine.

Das ist keine rationale Wahl.

f) Sie nutzt den Wahl-O-Mat, ein Computerprogramm das zeigt, „welche zu einer Wahl zugelassene Partei der eigenen politischen Position am nächsten steht" und wählt dem Ergebnis entsprechend.

Sich zu informieren, ist eine gute Idee, danach ohne nachzudenken zu wählen nicht.

g) Sie besucht die Informationsstände der Parteien in der Fussgängerzone.

Sich zu informieren, ist eine gute Idee.

Der 20-jährige Kölner Max W. will seine Kondition verbessern. Nun ist er auf der Suche nach einem guten Fitness-Studio, um dort regelmäßig zu trainieren.
Bewerte die folgenden Möglichkeiten, indem du pro Antwort eine Schulnote ankreuzt.

	1	2	3	4	5	6
Max fragt in seinem Freundeskreis nach Erfahrungen.	○	○	○	○	○	○
Max achtet auf die Anzeigen der Fitness-Studios in der Lokalzeitung.	○	○	○	○	○	○
Max recherchiert im Internet.	○	○	○	○	○	○
Max geht zum Probetraining in jedem Fitness-Studio in seiner Stadt.	○	○	○	○	○	○
Max geht zum Probetraining im nächstgelegenen Fitness-Studio.	○	○	○	○	○	○

#409#

Einschätzung der Alternativen

a) Max fragt in seinem Freundeskreis nach Erfahrungen.

Persönliche Erfahrungen von Bekannten können ein guter Startpunkt sein.

b) Max achtet auf die Anzeigen der Fitness-Studios in der Lokalzeitung.

Werbung in Zeitungen ist kein Zeichen für Qualität.

c) Max recherchiert im Internet.

Selbstständiges Informieren bei mehreren Quellen ist eine gute Idee.

d) Max geht zum Probetraining jedes Fitness-Studios seiner Stadt.

Das funktioniert nur in kleinen Orten.
In Großstädten wie Köln ist das viel zu aufwändig.

e) Max geht zum Probetraining im nächstgelegenen Fitness-Studio.

Das ist eine naheliegende Option. Ob das Fitness-Studio allerdings gut ist, bleibt unklar.

Zwischenfazit

Wir haben uns gerade mit Beispielen beschäftigt, bei denen es mehrere Handlungsmöglichkeiten gibt.

Dabei hast du gesehen,

- dass es sich lohnt, über verschiedene Lösungsmöglichkeiten nachzudenken.
- dass verschiedene Wege zum Ziel führen können.
- dass sich verschiedene Optionen in ihrer Qualität unterscheiden.
- dass es manchmal gar nicht so einfach ist, sich für eine Option zu entscheiden.

Block: Handlungswissen

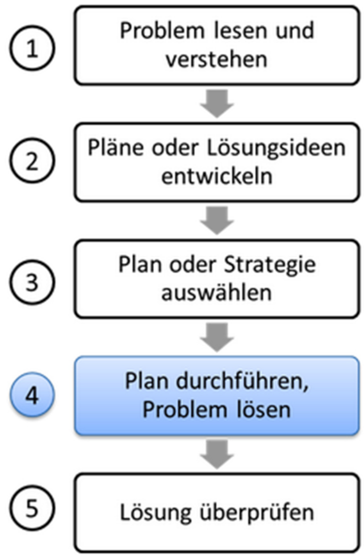
Problemlösen, ein Prozess

Die Abbildung zeigt wieder den Problemlöseprozess.

Der Umgang mit **Text, Tabellen und Abbildungen** kann in allen Phasen wichtig sein.

Oftmals beginnt dies schon beim **Lesen und Verstehen der Ausgangssituation**. Das hast du auch beim Neun-Punkte-Problem bereits gesehen.

Auch beim **konkreten Lösen eines Problems** spielt der Umgang mit Text, Abbildungen und Tabellen eine wichtige Rolle (Phase 4). Beispielsweise enthalten Aufgaben in Schulbüchern neben geschriebenen Text oftmals Abbildungen und Tabellen.



Vorbemerkung zu den Aufgaben

Auf den nächsten Seiten folgen verschiedene Aufgaben, in denen neben Text Abbildungen oder Tabellen vorkommen. Bitte versuche die Aufgaben zu lösen. Wenn du mit den Aufgaben auf einer Seite fertig bist, klicke wie bisher auf „weiter".

- Wenn deine Lösung richtig ist, siehst du auf der nächsten Seite folgenden Smiley:

- Wenn deine Lösung nicht vollständig richtig ist, erhältst du Tipps und einen zweiten Versuch, die Aufgabe (besser) zu lösen. Wenn du die Lösung beim zweiten Versuch nicht findest, wir dir die Musterlösung angezeigt.

Mandy möchte einen neuen Toaster kaufen. Auf einem Vergleichsportal im Internet findest sie folgende Tabelle zum Vergleich dreier Toaster.

	Toasty Mach 2	Toastfix	Toasterwave
Optik	Sehr gut	befriedigend	gut
Handhabung	ausreichend	gut	befriedigend
Toastergebnis	befriedigend	gut	sehr gut
Stromverbrauch	mittel	mittel	gering
Gesamtnote	befriedigend (2,7)	gut (2,3)	gut (2,0)
Kindersicherung	ja	nein	nein
Preis	59,95 €	29,99 €	39,95 €

Frage 1: Welches ist der günstigste Toaster?

- Toasty Mach 2
- Toastfix
- Toasterwave
- Alle Geräte sind gleich teuer.

Frage 2: Ihrer Freundin Elena ist vor allem das Design wichtig. Welcher Toaster ist wohl ihr Favorit?

- Toasty Mach 2
- Toastfix
- Toasterwave
- Das kann man mit den Angaben in der Tabelle nicht beantworten.

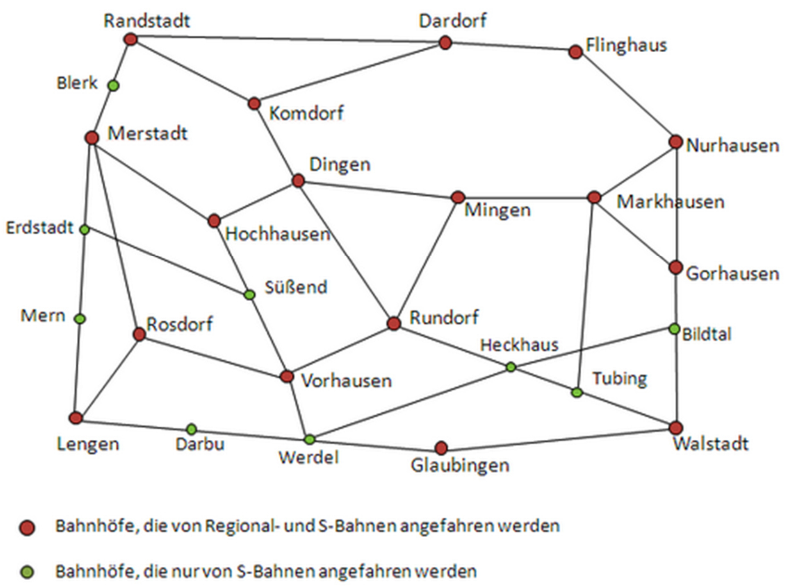
Hinweise zur Lösung

Du hast noch keine optimale Lösung gefunden.

Bevor du einen neuen Lösungsvorschlag machst, lies dir bitte die folgenden Hinweise durch.

- Frage 1: Um den günstigsten Toaster zu finden, sieh in die Zeile „Preis". Welches Gerät kostet am wenigsten?
- Frage 2: Elena wählt wohl den Toaster mit der besten Optik. Suche also in der Zeile „Optik" nach der besten Bewertung.

Hier siehst du ein Streckennetz der Bahn:



Wie viele Bahnhöfe, die nur von S-Bahnen angefahren werden, gibt es auf der Karte?

Wie viele Bahnhöfe, die nur von S-Bahnen angefahren werden, liegen auf der Strecke Dingen-Rundorf-Heckhaus-Tubing-Walstadt?

Wie viele Bahnhöfe, die nur von S-Bahnen angefahren werden, liegen auf der Strecke Randstadt-Blerk-Merstadt-Erdstadt-Mern?

Hinweise zur Lösung

Du hast noch keine optimale Lösung gefunden.

Beachte bei deinem neuen Versuch folgende Hinweise:

- Frage 1: Sieh in der Legende nach, wie S-Bahnhöfe gekennzeichnet sind. Zähle dann in der Karte diese Symbole ab.
- Frage 2 und 3: Suche zunächst die angegebene Strecke auf der Karte. Zähle dann die S-Bahnhöfe auf dieser Strecke.

Unten siehst du den Jubiläumsflyer der Pizzeria Italia.

Pizzeria Italia	Tel. 05555 – 42 42 42
Altstraße 42	pizzeria-italia@www.de
55555 Kleinstadt	

Pizzeria Italia:
Seit 10 Jahren ihre Nummer 1!

Neue Jubiläumsangebote!

	Pizzen	dein (Ø 26 cm	mittel (Ø 32 cm)	groß (Ø 36 cm)
101	**Pizza Capricciosa** mit Schinken, Champignons und Artischocken	4,50 €	6,50 €	7,50 €
102	**Pizza Hawaii** mit Schinken und Ananas	3,00 €	5,00 €	6,00 €
103	**Pizza Vier Jahreszeiten** mit Champignons, Paprika, Schinken und Salami	3,50 €	5,50 €	6,50 €
104	Pizza Broccoli mit Broccoli, Tomatenscheiben, Knoblauch und Gorgonzola	4,50 €	6,50 €	7,50 €
105	**Pizza Meeresfrüchte** mit Krabben, Tintenfisch, Miesmuscheln, Zwiebeln und Knoblauch	5,50 €	4,50 €	8,50 €

Ab einem Bestellwert von 15 Euro liefern wir frei Haus!

Bevor der Jubiläumsflyer in den Druck gehen kann, muss er überprüft werden. Das ist deine Aufgabe.
Gib an, ob die folgenden Aspekte des Jubiläumsflyers mindestens einen Fehler enthalten.

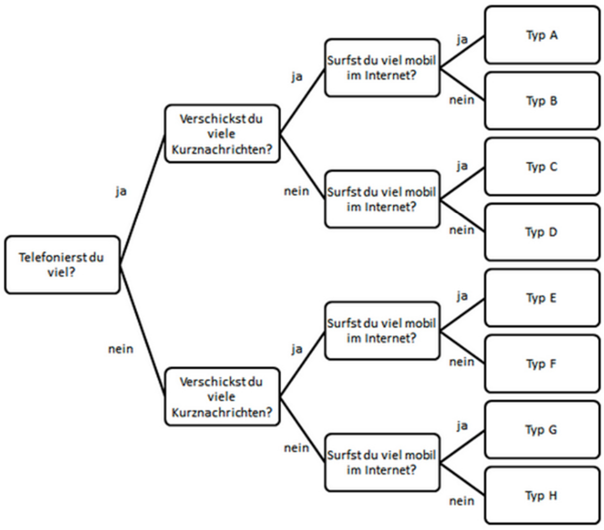
Layout (u. a. Anordnung der grafischen Bestandteile)	○ fehlerhaft	○ fehlerfrei
Preise für kleine Pizzen	○ fehlerhaft	○ fehlerfrei
Preise für mittlere Pizzen	○ fehlerhaft	○ fehlerfrei
Preise für große Pizzen	○ fehlerhaft	○ fehlerfrei
Rechtschreibung	○ fehlerhaft	○ fehlerfrei
Formatierung (z. B. einheitliche Schriftart)	○ fehlerhaft	○ fehlerfrei

Du hast mindestens eine Teilaufgabe nicht beantwortet!

Deine Lösung ist noch nicht vollständig richtig.

Beachte bei deinem zweiten Lösungsversuch bitte folgende Hinweise:
- Hast du keine Antwort ausgelassen?
- Prüfe die verschiedenen Aspekte einzeln.
- Sobald du einen Fehler bei einem Aspekt findest, kreuze "fehlerhaft" an. Du brauchst dann diesen Aspekt nicht auf weitere Fehler kontrollieren.

Als Informatikprojekt wollen Udo und Mareike eine Internetseite erstellen, die zunächst erfragt, welchem Handynutzertyp der Besucher entspricht, und ihm dann Vorschläge für Handys bzw. Smartphones unterbreitet. Udo und Mareike befragen ihre Freunde und Bekannten. Im Wesentlichen unterscheiden die Befragten in ihren Antworten auf die folgenden drei Fragen: Telefonierst du viel? Verschickst du viele Kurznachrichten? Surfst du viel mobil im Internet?
Um sich das Programm besser vorstellen zu können, visualisieren Udo und Mareike ihre Ergebnisse.



Gib an, zu welchem Handynutzertyp die Personen jeweils gehören.

	Typ A	Typ B	Typ C	Typ D	Typ E	Typ F	Typ G	Typ H
Melinda telefoniert kaum, verschickt aber viele Kurznachrichten und surft viel im Internet.	○	○	○	○	○	○	○	○
Edgar findet Internet auf seinem Handy überflüssig, telefoniert täglich und schickt viele SMS.	○	○	○	○	○	○	○	○
Jason nutzt sein Handy nur zum Verschicken von Kurznachrichten.	○	○	○	○	○	○	○	○

Laura telefoniert bis zu zwei Stunden täglich und hört viel Internetradio. Zu welchen Handynutzertypen könnte sie gehören?
Kreuze alle Handynutzertypen an, zu denen Laura gehören könnte.

☐ Typ A ☐ Typ B ☐ Typ C ☐ Typ D ☐ Typ E ☐ Typ F ☐ Typ G ☐ Typ H

Hinweise zur Lösung

Du hast noch keine optimale Lösung gefunden.

Beachte bei deinem neuen Versuch folgende Hinweise:

Edgar findet Internet auf seinem Handy überflüssig, telefoniert viel und verschickt Hunderte von SMS. Welcher Nutzertyp ist Edgar?
- Gehe das Diagramm der Reihe nach ab: Telefoniert Edgar viel? Verschickte Edgar viele Kurznachrichten? Surft Edgar viel mobil im Internet?

Jason nutzt sein Handy nur zum Verschicken von Kurznachrichten. Welcher Nutzertyp ist Jason?
- Wenn Jason sein Handy nur zum Verschicken von Kurznachrichten nutzt, dann telefoniert er damit nicht und surft auch nicht mobil im Internet. Versuche mit diesen Hinweisen den passenden Nutzertyp zu finden.

Laura telefoniert bis zu zwei Stunden täglich und hört viel Internetradio. Welcher Nutzertyp könnte sie sein?
- Laura telefoniert viel und nutzt das mobile Internet. Ob sie viele Kurznachrichten schreibt oder nicht, ist nicht bekannt. Betrachte beide Möglichkeiten. Zu welchem Nutzertypen könnte sie somit gehören?

Maxi möchte für ihre Geburtstagsfeier multikulturell kochen. Die folgenden Gerichte (mit Zutatenliste) kommen in ihre engere Auswahl:

- **Spaghetti Carbonara** (Nudeln, Zwiebeln, Eigelb, Schmand, roher Schinken, Pfeffer, Knoblauch, Olivenöl)
- **Bifteki** (Hackfleisch, Schafskäse, Zwiebeln, Paniermehl, Eier, Pfeffer, Salz, Knoblauch, Olivenöl)
- **Lahmacun** (Mehl, Wasser, Zucker, Salz, Trockenhefe, Hackfleisch, Zwiebeln, Paprika, Salz, Chilipulver, Tomatenmark, Tomaten)
- **Fish ´n Chips** (Kartoffeln, Meersalz, Fischfilet, Bier, Mehl, Speisestärke, Backpulver, Salz, Pfeffer)
- **Vanille-Pudding** (Vanillepuddingpulver, Milch, Zucker)
- **Obstsalat** (Bananen, Äpfel, Orangen, Weintrauben, Zitronensaft, Zucker)

Bei der endgültigen Entscheidung will sie sich nach den Besonderheiten ihrer Gäste richten: Sie weiß, dass Simon eine Knoblauch-Allergie hat und Tina keinerlei Milchprodukte verträgt. Auf die übrigen Gäste muss Maxi bei der Essensauswahl nicht achten.

Frage 1: Welche Gerichte könnte Maxi für ihre Gäste 1:1 nach Rezept zubereiten, wenn sie darauf achtet, was ihre Gäste essen können und was nicht?

- [ ] Spaghetti Carbonara
- [ ] Bifteki
- [ ] Lahmacun
- [ ] Fish ´n Chips
- [ ] Vanille-Pudding
- [ ] Obstsalat

Frage 2: Überraschend sagt Tina zwei Tage vorher wegen Krankheit ab. Welche Gerichte könnte Maxi jetzt zusätzlich vorbereiten.

- [ ] Spaghetti Carbonara
- [ ] Bifteki
- [ ] Lahmacun
- [ ] Fish ´n Chips
- [ ] Vanille-Pudding
- [ ] Obstsalat

Anmerkung. Diese Aufgabe wurde auch in Experiment 3 als Teil des kleinen Transfertests nach dem Planungskompetenztraining verwendet.

Hinweise zur Lösung

Du hast noch keine optimale Lösung gefunden.

Beachte bei deinem neuen Versuch folgende Hinweise:

Frage 1:
- Simon verträgt keinen Knoblauch. Daher scheiden alle Gerichte mit Knoblauch aus.
- Tina verträgt keine Milchprodukte. Daher scheiden alle Gerichte aus, die Zutaten enthalten, die aus Milch hergestellt werden. In welchen der angegebenen Zutaten ist Milch enthalten?

Frage 2:
- Welche Gerichte enthalten Milchprodukte? Welche Beschränkung der Speisenauswahl gilt weiterhin?

Bei Familie Neustein steht eine Familienfeier bevor. Mutter Andrea hat Verwandte eingeladen, Vater Ingo das Essen bestellt. Nur die Sitzordnung bereitet den Neusteins noch Kopfzerbrechen. Denn Teile der Verwandtschaft sind zerstritten:

- Bei den Cousinen Susanne und Michelle herrscht gerade Zickenkrieg.
- Tante Frieda und Onkel Peter sind seit Kindertagen zerstritten.
- Mandy, Michelle und Ingo können sich nicht leiden.

Um die Feier nicht zu stören, sollten an einem Tisch die Streithähne nicht direkt nebeneinander oder direkt gegenüber sitzen. Auch Rücken an Rücken sollten diese Personen nicht sitzen.

Anderseits möchten bestimmte Personen gerne direkt nebeneinander, gegenüber oder schräg gegenüber sitzen, um sich gut unterhalten zu können:

- Paul, Mandy und Frieda wollen über einen möglichen gemeinsamen Urlaub sprechen.
- Andrea und ihr Bruder Heinz haben sich lange nicht gesehen und möchten nebeneinander sitzen.
- Die Kinder Jason, Kevin und Eva möchten gerne zusammen sitzen.

Es steht fest, dass die Tische folgendermaßen stehen müssen:

Tisch 1	1	2	3
	4	5	6

Tisch 2	7	8	9
	10	11	12

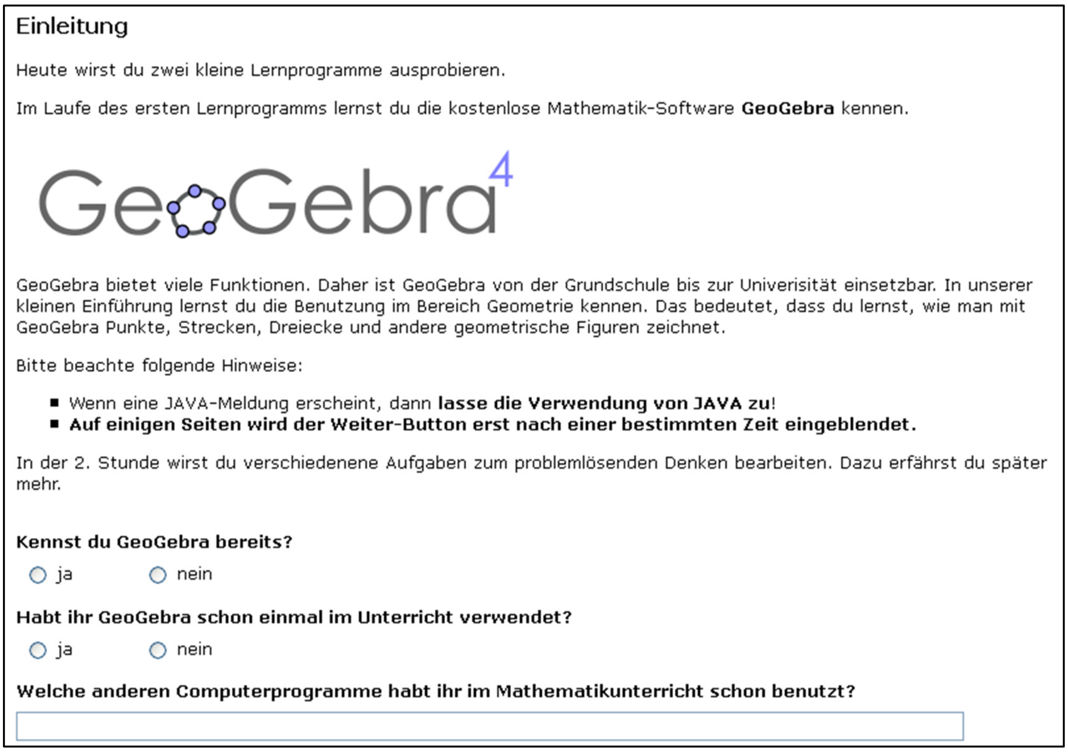
Mache einen Vorschlag für eine Verteilung der Neusteins auf die beiden Tische, d. h. gib eine mögliche Sitzordnung an.

Tisch 1	-	-	-
	-	-	-
Tisch 2	-	-	-
	-	-	-

Anhang M. Experiment 2: Änderungen im Kontrollgruppentreatment

Die Screenshots wurden aus Druckgründen auf eine Seitenbreite von 14 cm skaliert, was ungefähr 2/3 der Originalgröße entspricht. Aus Platzgründen kennzeichnet ein schwarzer Rahmen jeweils eine Bildschirmseite, damit mehrere Screenshots auf einer Seite dokumentiert werden können.

Um Redundanz zu vermeiden, enthält dieser Anhang nur Seiten, die im Vergleich zum Experiment 2 in Anhang F neu dazu gekommen sind oder geändert worden sind. Die Einleitungsseite wurde um zwei Hinweise ergänzt:



Einleitung

Heute wirst du zwei kleine Lernprogramme ausprobieren.

Im Laufe des ersten Lernprogramms lernst du die kostenlose Mathematik-Software **GeoGebra** kennen.

GeoGebra bietet viele Funktionen. Daher ist GeoGebra von der Grundschule bis zur Univerisität einsetzbar. In unserer kleinen Einführung lernst du die Benutzung im Bereich Geometrie kennen. Das bedeutet, dass du lernst, wie man mit GeoGebra Punkte, Strecken, Dreiecke und andere geometrische Figuren zeichnet.

Bitte beachte folgende Hinweise:

- Wenn eine JAVA-Meldung erscheint, dann **lasse die Verwendung von JAVA zu**!
- **Auf einigen Seiten wird der Weiter-Button erst nach einer bestimmten Zeit eingeblendet.**

In der 2. Stunde wirst du verschiedenene Aufgaben zum problemlösenden Denken bearbeiten. Dazu erfährst du später mehr.

Kennst du GeoGebra bereits?

○ ja ○ nein

Habt ihr GeoGebra schon einmal im Unterricht verwendet?

○ ja ○ nein

Welche anderen Computerprogramme habt ihr im Mathematikunterricht schon benutzt?

Nach Übung 2 wurde eine Übung zum Speichern von GeoGebra-Dateien hinzugefügt:

Übung: Datei speichern

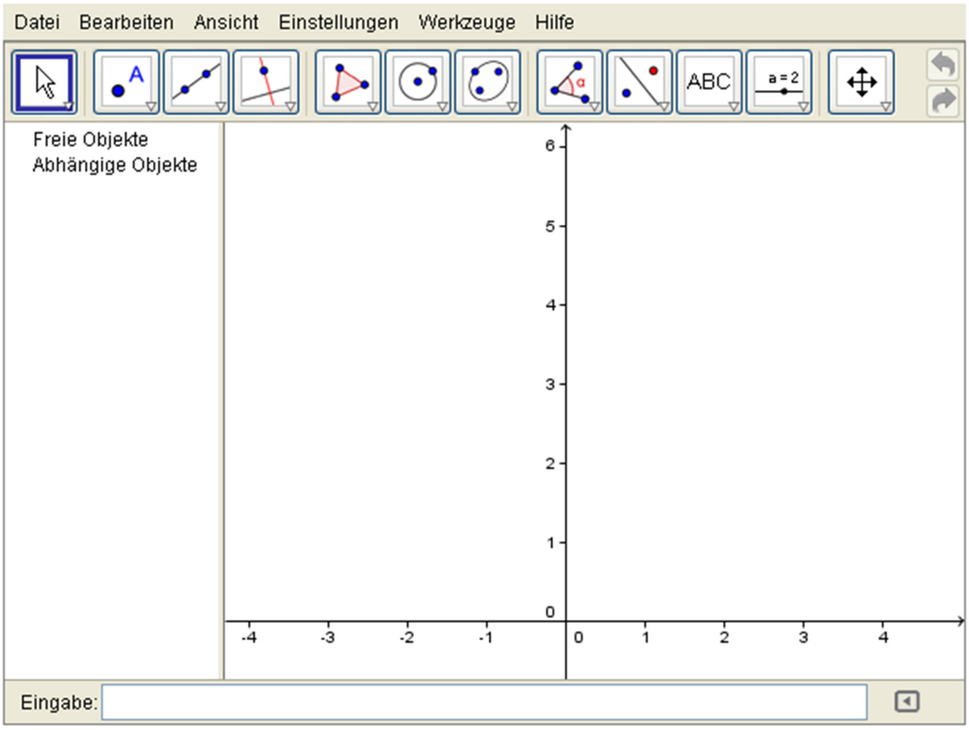
Du kannst deine GeoGebra-Datei natürlich auch speichern.

Um eine GeoGebra-Datei zu speichern, macht man einfach Folgendes:

1. Klicke in der Menüleiste auf „Datei".
2. Klicke auf „Speichern unter...".
3. Es öffnet sich ein Fenster.
4. Gib einen Dateinamen ein.
5. Wähle den Speicherort aus.
6. Klicke auf „Speichern".

Aufgabe: Erzeuge zwei beliebige Objekte in der Zeichenfläche und speichere die GeoGebra-Datei dann wie oben beschrieben.

GeoGebra



Wie kann man eine GeoGebra-Datei speichern?
Gib bitte die Reihenfolge der Schritte an.

	Schritt
1.	Bitte wähle eine Option aus.
2.	Bitte wähle eine Option aus.
3.	Bitte wähle eine Option aus.
4.	Bitte wähle eine Option aus.
5.	Bitte wähle eine Option aus.
6.	Bitte wähle eine Option aus.

Wie leicht oder schwer war die Übung zu verstehen?

sehr leicht ○ ○ ○ ○ ○ ○ ○ **sehr schwer**

Beim Bearbeiten und Verstehen der Übung war meine Denk-Anstrengung ...

sehr gering ○ ○ ○ ○ ○ ○ ○ **sehr hoch**

Nach dem Zwischenfazit wurden zwei neue Übungen ergänzt:

Beschriftungen

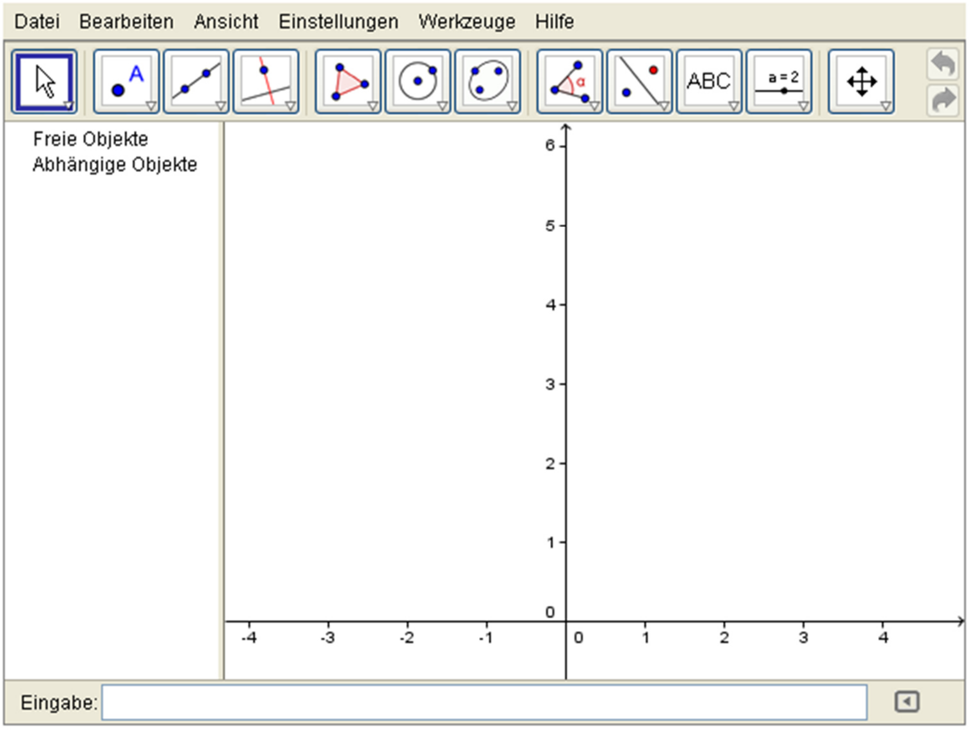
Selbstverständlich kannst du auch Text zu deinen Zeichnungen hinzufügen. Dabei gibt es verschiedene Möglichkeiten.

a) Du kannst Text einfügen, indem du auf das Symbol "ABC" klickst und den Bildschirmanweisungen folgst.

b) Du kannst das Stiftsymbol verwenden, um mit gedrückter linker Maustaste zu zeichnen und zu schreiben.

Aufgabe: Probiere jetzt bitte beide Möglichkeiten aus.

GeoGebra



Wie leicht oder schwer war die Übung zu verstehen?

sehr leicht ○ ○ ○ ○ ○ ○ ○ **sehr schwer**

Beim Bearbeiten und Verstehen der Übung war meine Denk-Anstrengung ...

sehr gering ○ ○ ○ ○ ○ ○ ○ **sehr hoch**

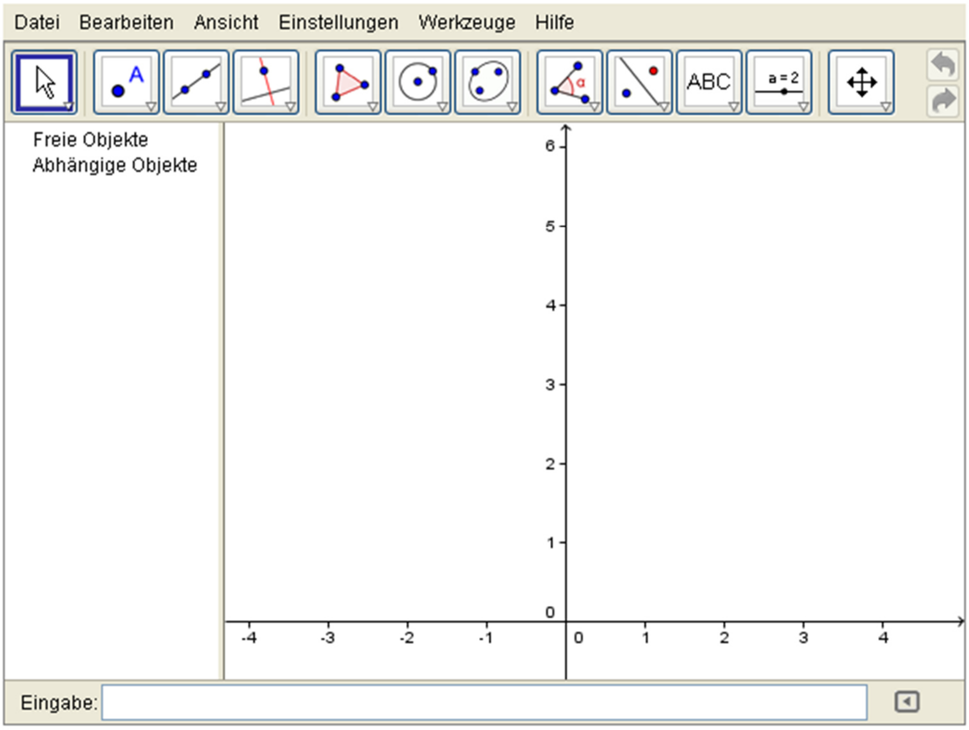
Übung: Farben

Wie man geometrische Figuren mit GeoGebra erzeugt, weißt du bereits. Jetzt sehen wir uns an, wie man die Farbe von Objekten ändern kann.

Wenn du die Farbe eines Objekts ändern möchtest, gehe wie folgt vor:

1. Erstelle ein Objekt, dessen Farbe du ändern möchtest, z. B. ein Dreieck.
2. Klicke mit der rechten Maustaste auf das Objekt
3. Klicke dann auf "Eigenschaften..." . Es öffnet sich ein Fenster. Achtung: Das kann im Browser einen Moment dauern.
4. Klicke im neuen Fenster auf "Farben".
5. Wähle per Linksklick die gewünschte Farbe aus
6. Klicke dann unten rechts auf "Schließen".

GeoGebra



Wie leicht oder schwer war die Übung zu verstehen?

sehr leicht ○ ○ ○ ○ ○ ○ ○ **sehr schwer**

Beim Bearbeiten und Verstehen der Übung war meine Denk-Anstrengung ...

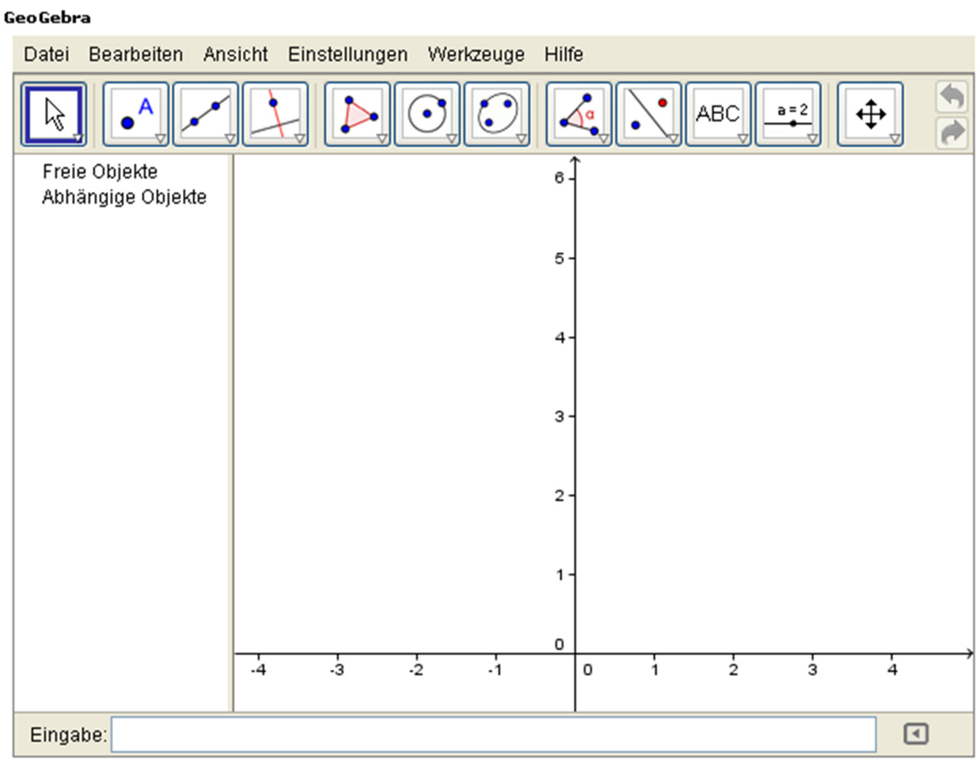
sehr gering ○ ○ ○ ○ ○ ○ ○ **sehr hoch**

Nach Übung 5 wurde eine wiederholende Übung hinzugefügt:

letzte Übung

Zum Abschluss wiederholen wir noch einmal das Zeichnen der wichtigsten Objekte.

1. Zeichne zwei Punkte.
2. Zeichne eine Gerade durch diese zwei Punkte
3. Zeichne ein grünes Dreieck
4. Gib deiner Zeichnung eine Überschrift
5. Speichere deine Datei



Wie leicht oder schwer war die Übung zu verstehen?

sehr leicht ○ ○ ○ ○ ○ ○ ○ **sehr schwer**

Beim Bearbeiten und Verstehen der Übung war meine Denk-Anstrengung ...

sehr gering ○ ○ ○ ○ ○ ○ ○ **sehr hoch**

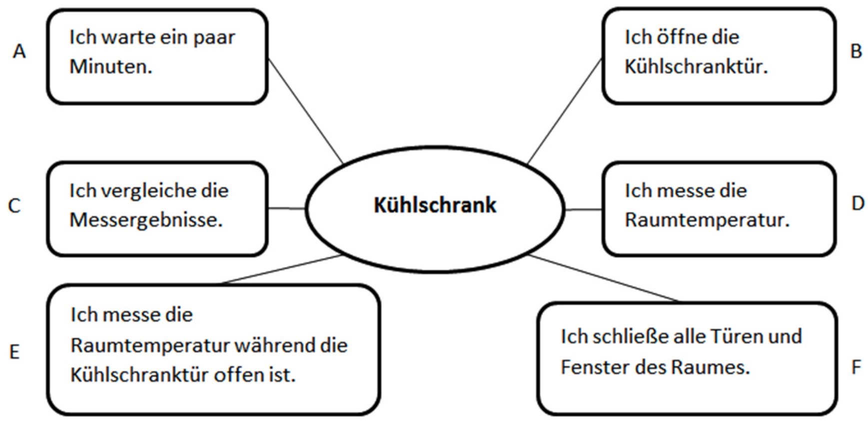
Anhang N. Experiment 2: Posttest Problemlösen

Die einzelnen Testblöcke und die Aufgaben innerhalb eines Blocks wurden randomisiert vorgegeben, um Reihenfolgeeffekte zu vermeiden.

Block: Planungsfähigkeit

Die Aufgabe Kühlschrank wurde im Vergleich zum Experiment geringfügig geändert.

Du willst für ein Schulprojekt herausfinden, ob sich die Raumtemperatur ändert, wenn die Kühlschranktür offen stehen bleibt. Zu Beginn hast du mögliche Teilschritte des Vorgehens gesammelt. Das Ergebnis ist in folgender Mindmap dargestellt:



Deine Aufgabe ist es, die in der Mindmap angegebenen Teilschritte in eine sinnvolle Reihenfolge zu bringen.

Die übrigen Aufgaben wurden wie in Experiment 1 eingesetzt (siehe Anhang G).

Block: Konditionales Wissen

Es wurden dieselben Aufgaben wie in Experiment 1 verwendet (siehe Anhang G).

Block: Handlungswissen

Es wurden dieselben Aufgaben wie in Experiment 1 verwendet (siehe Anhang G) mit Ausnahme der Aufgaben *Fehlstunden und Einzelverbindungsnachweis*, die gestrichen wurden, da sie zu leicht waren.

Block Problemlösen

Es wurden wie in Experiment 1 folgende Aufgaben zum Problemlösen aus PISA 2003 (OECD, 2004b) eingesetzt: Kinobesuch, Anschlusszüge, Gefrierschrank, Bewässerung, Bibliothekensystem (nur Teil 1), Studienplan, Urlaub. Die selbst erstellte Aufgabe *Sitzordnung* wurde vom Posttest in das Treatment verschoben.

Anhang O. Experiment 2: Bearbeitungszeit pro Gruppe

Tabelle O-1. Bearbeitungszeit des Treatments und Posttests in Minuten getrennt nach experimenteller Bedingung in Experiment 2

Zeit	*experimentelle Bedingung*	*M*	*SD*	*n*
Treatment	Konditionales Wissen + Planungsfähigkeit	35.85	8.15	53
	Handlungswissen + Planungsfähigkeit	38.11	7.89	54
	Handlungswissen + konditionales Wissen	34.99	8.28	53
	Kontrolltreatment	36.25	8.45	43
	Gesamtsumme	36.31	8.21	203
Posttest	Konditionales Wissen + Planungsfähigkeit	27.34	10.17	53
	Handlungswissen + Planungsfähigkeit	26.38	10.00	54
	Handlungswissen + konditionales Wissen	27.31	9.42	53
	Kontrolltreatment	29.09	10.89	43
	Gesamtsumme	27.45	10.06	203

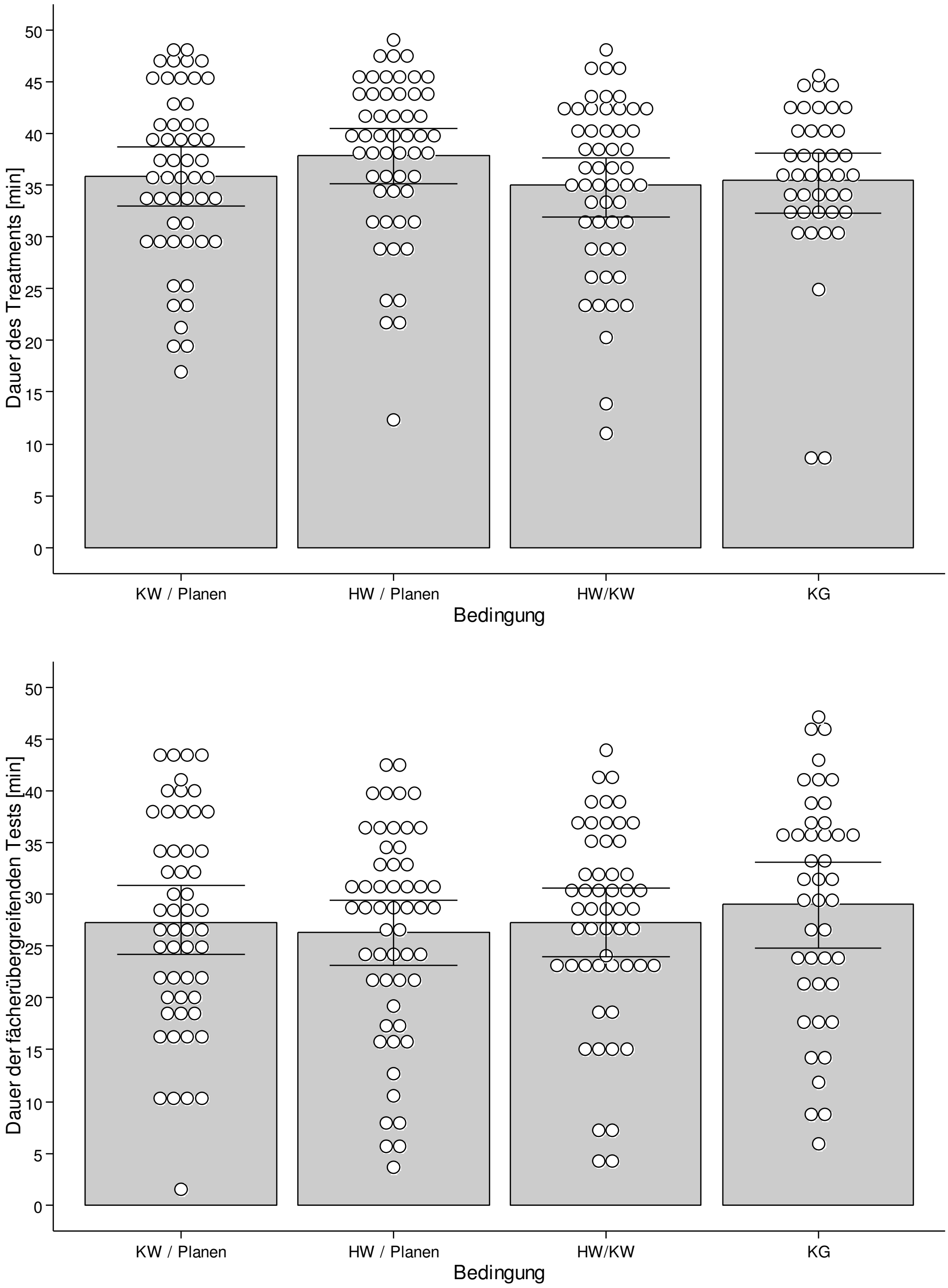


Abbildung O-1. Dauer des Treatments und des Posttests in Experiment 2 gruppiert nach experimenteller Bedingung als Dotplot (weiß) und Balkendiagramme (grau)

Anmerkungen. KW = Konditionales Wissen, Planen = Planungsfähigkeit, HW = Handlungswissen. Das schwarze Dreieck kennzeichnet den Mittelwert. Fehlerbalken repräsentieren das 99 %-Bootstrap-Konfidenzintervall des Mittelwerts (mit 1000 Replikationen). Die Abbildung wurde mit dem R-Paket *ggplot 2* (Wickham, 2009) erzeugt.

Anhang P. Experiment 2: Deskriptivstatistik und Analysen

Tabelle P-1. Deskriptivstatistik Handlungswissen

experimentelle Bedingung	*Schulform*	*M*	*SD*	*N*
Gruppe 1 (KW und PLA)	Hauptschule	-0.18	0.73	15
	Realschule	-0.06	0.96	28
	Gymnasium	0.08	1.22	5
	Gesamt	-0.08	0.91	48
Gruppe 2 (HW und PLA)	Hauptschule	-0.06	1.00	15
	Realschule	0.00	1.06	29
	Gymnasium	0.75	0.77	4
	Gesamt	0.04	1.02	48
Gruppe 3 (HW und KW)	Hauptschule	0.08	0.89	15
	Realschule	-0.33	0.84	28
	Gymnasium	0.80	0.80	4
	Gesamt	-0.10	0.90	47
Gruppe 4 (Kontrollgruppe)	Hauptschule	0.38	0.70	6
	Realschule	-0.12	1.16	29
	Gymnasium	-0.22	1.54	4
	Gesamt	-0.05	1.13	39
Gesamt	Hauptschule	0.00	0.85	51
	Realschule	-0.13	1.01	114
	Gymnasium	0.34	1.11	17
	Gesamt	-0.05	0.98	182

Anmerkungen. Alle Variablen sind z-standardisiert. KW = Konditionales Wissen, PLA = Planungsfähigkeit, HW = Handlungswissen.

Tabelle P-2. Deskriptivstatistik konditionalen Wissens

experimentelle Bedingung	*Schulform*	*M*	*SD*	*N*
Gruppe 1 (KW und PLA)	Hauptschule	0.20	0.93	15
	Realschule	0.16	1.13	28
	Gymnasium	-0.83	1.20	5
	Gesamt	0.07	1.10	48
Gruppe 2 (HW und PLA)	Hauptschule	-0.32	1.03	15
	Realschule	-0.15	0.95	29
	Gymnasium	0.39	0.92	4
	Gesamt	-0.16	0.97	48
Gruppe 3 (HW und KW)	Hauptschule	0.36	1.01	15
	Realschule	0.19	0.93	28
	Gymnasium	1.34	0.52	4
	Gesamt	0.34	0.97	47
Gruppe 4 (Kontrollgruppe)	Hauptschule	0.00	0.89	6
	Realschule	-0.33	0.96	29
	Gymnasium	-0.61	1.12	4
	Gesamt	-0.31	0.95	39
Gesamt	Hauptschule	0.07	0.99	51
	Realschule	-0.04	1.01	114
	Gymnasium	0.02	1.27	17
	Gesamt	0.00	1.02	182

Anmerkungen. Alle Variablen sind z-standardisiert. KW = Konditionales Wissen, PLA = Planungsfähigkeit, HW = Handlungswissen.

Tabelle P-3. Deskriptivstatistik Planungsfähigkeit

experimentelle Bedingung	*Schulform*	*M*	*SD*	*N*
Gruppe 1 (KW und PLA)	Hauptschule	-0.13	0.89	15
	Realschule	0.13	1.08	28
	Gymnasium	-0.38	1.22	5
	Gesamt	0.00	1.03	48
Gruppe 2 (HW und PLA)	Hauptschule	-0.06	0.85	15
	Realschule	-0.15	1.14	29
	Gymnasium	0.61	0.26	4
	Gesamt	-0.06	1.02	48
Gruppe 3 (HW und KW)	Hauptschule	0.52	0.73	15
	Realschule	-0.42	1.07	28
	Gymnasium	0.38	0.84	4
	Gesamt	-0.05	1.04	47
Gruppe 4 (Kontrollgruppe)	Hauptschule	0.37	0.50	6
	Realschule	0.20	0.86	29
	Gymnasium	-0.09	1.57	4
	Gesamt	0.19	0.88	39
Gesamt	Hauptschule	0.14	0.83	51
	Realschule	-0.06	1.06	114
	Gymnasium	0.10	1.07	17
	Gesamt	0.01	1.00	182

Anmerkungen. Alle Variablen sind z-standardisiert. KW = Konditionales Wissen, PLA = Planungsfähigkeit, HW = Handlungswissen.

Tabelle P-4. Follow-up-ANCOVAs zum Treatment-Check der Komponenten fächerübergreifenden Problemlösens

Quelle der Variation	*Abhängige Variable*	*SS*	*Df*	*MS*	*F*	*Sig.*	η_p^2
KFT-Skala N2	KW [a]	15.139	1	15.139	16.867	.000	.091
	PLA [b]	24.756	1	24.756	29.770	.000	.150
	HW [c]	24.182	1	24.182	29.067	.000	.147
Bedingung	KW	13.719	3	4.573	5.095	.002	.083
	PLA	1.660	3	0.553	0.665	.575	.012
	HW	1.466	3	0.489	0.588	.624	.010
Schulform	KW	0.160	2	0.080	0.089	.915	.001
	PLA	1.373	2	0.687	0.826	.440	.010
	HW	2.547	2	1.274	1.531	.219	.018
Bedingung * Schulform	KW	8.567	6	1.428	1.591	.153	.053
	PLA	12.155	6	2.026	2.436	.028	.080
	HW	4.938	6	0.823	0.989	.434	.034
Fehler	KW	151.693	169	0.898			
	PLA	140.539	169	0.832			
	HW	140.601	169	0.832			
Gesamt	KW	189.269	182				
	PLA	180.600	182				
	HW	174.856	182				

Anmerkungen. KW = Konditionales Wissen, HW = Handlungswissen, PLA = Planungsfähigkeit.
[a] R^2 = .199 (Angepasstes R^2 = .142). [b] R^2 = .222 (Angepasstes R^2 = .166), [c] R^2 = .194 (Angepasstes R^2 = .137).

Tabelle P-5. Paarweise Vergleiche der Skala fächerübergreifendes konditionales Wissen (mit Bonferoni-Korrektur)

Gruppe i	*Gruppe j*	$M_i - M_j$	*SE*	*Sig.*	*95 %-KI Untere Grenze*	*95 %-KI Obere Grenze*
Gruppe 1 (KW und PLA)	Gruppe 2 (HW und PLA)	-0.110	0.255	1.000	-0.792	0.572
	Gruppe 3 (HW und KW)	-0.814	0.256	.010	-1.496	-0.132
	Gruppe 4 (KG)	0.169	0.274	1.000	-0.563	0.901
Gruppe 2 (HW und PLA)	Gruppe 1 (KW und PLA)	0.110	0.255	1.000	-0.572	0.792
	Gruppe 3 (HW und KW)	-0.704	0.265	.052	-1.412	0.004
	Gruppe 4 (KG)	0.279	0.283	1.000	-0.476	1.034

Anmerkung. KW = Konditionales Wissen, HW = Handlungswissen, PLA = Planungsfähigkeit, KG = Kontrollgruppe mit GeoGebra-Treatment.

Tabelle P-6. Deskriptivstatistik: fächerübergreifendes Problemlösen

experimentelle Bedingung	*Schulform*	*M*	*SD*	*N*
Gruppe 1 (KW und PLA)	Hauptschule	-0.04	0.73	15
	Realschule	-0.26	0.89	28
	Gymnasium	-0.22	1.41	5
	Gesamt	-0.19	0.89	48
Gruppe 2 (HW und PLA)	Hauptschule	-0.01	0.96	15
	Realschule	-0.36	0.77	29
	Gymnasium	1.50	0.44	4
	Gesamt	-0.10	0.95	48
Gruppe 3 (HW und KW)	Hauptschule	0.14	0.74	15
	Realschule	-0.21	0.78	28
	Gymnasium	0.84	1.69	4
	Gesamt	-0.01	0.90	47
Gruppe 4 (Kontrollgruppe)	Hauptschule	-0.11	0.94	6
	Realschule	0.08	1.17	29
	Gymnasium	0.51	1.10	4
	Gesamt	0.09	1.11	39
Gesamt	Hauptschule	0.01	0.81	51
	Realschule	-0.19	0.92	114
	Gymnasium	0.61	1.31	17
	Gesamt	-0.06	0.96	182

Anmerkungen. Alle Variablen sind z-standardisiert. KW = Konditionales Wissen, PLA = Planungsfähigkeit, HW = Handlungswissen.

Tabelle P-7. Paarweise Vergleiche der Skala fächerübergreifendes Problemlösen (mit Bonferoni-Korrektur)

					95 %-KI	
Gruppe i	*Gruppe j*	$M_i - M_j$	*SE*	*Sig.*	*Untere Grenze*	*Obere Grenze*
Hauptschule	Realschule	0.157	0.154	.928	-0.215	0.528
	Gymnasium	-.0601*	0.248	.049	-1.200	-0.002
Realschule	Hauptschule	-0.157	0.154	.928	-0.528	0.215
	Gymnasium	-0.758*	0.226	.003	-1.304	-0.212
Gymnasium	Hauptschule	0.601*	0.248	.049	0.002	1.200
	Realschule	0.758*	0.226	.003	0.212	1.304

Anmerkung. Basierend auf geschätzten Randmitteln.

Tabelle P-8. Deskriptivstatistik: mathematisches konditionales Wissen

experimentelle Bedingung	*Schulform*	*M*	*SD*	*N*
Gruppe 1 (KW und PLA)	Hauptschule	0.17	0.95	15
	Realschule	-0.17	1.08	28
	Gymnasium	0.11	0.82	5
	Gesamt	-0.04	1.01	48
Gruppe 2 (HW und PLA)	Hauptschule	-0.45	1.13	15
	Realschule	0.00	0.84	29
	Gymnasium	0.64	0.51	4
	Gesamt	-0.09	0.96	48
Gruppe 3 (HW und KW)	Hauptschule	-0.30	0.93	15
	Realschule	-0.01	0.88	28
	Gymnasium	1.05	1.09	4
	Gesamt	-0.01	0.96	47
Gruppe 4 (Kontrollgruppe)	Hauptschule	-0.42	0.40	6
	Realschule	-0.12	1.02	29
	Gymnasium	0.57	0.47	4
	Gesamt	-0.10	0.93	39
Gesamt	Hauptschule	-0.22	0.97	51
	Realschule	-0.08	0.95	114
	Gymnasium	0.57	0.78	17
	Gesamt	-0.06	0.96	182

Anmerkungen. Alle Variablen sind z-standardisiert. KW = Konditionales Wissen, PLA = Planungsfähigkeit, HW = Handlungswissen.

Tabelle P-9. Deskriptivstatistik: mathematische Planungsfähigkeit

experimentelle Bedingung	*Schulform*	*M*	*SD*	*N*
Gruppe 1 (KW und PLA)	Hauptschule	-0.20	1.03	15
	Realschule	0.22	0.82	28
	Gymnasium	0.13	1.16	5
	Gesamt	0.08	0.92	48
Gruppe 2 (HW und PLA)	Hauptschule	0.36	0.39	15
	Realschule	-0.20	1.04	29
	Gymnasium	-0.27	1.09	4
	Gesamt	-0.03	0.91	48
Gruppe 3 (HW und KW)	Hauptschule	0.40	0.85	15
	Realschule	-0.21	1.02	28
	Gymnasium	1.25	0.17	4
	Gesamt	0.11	1.02	47
Gruppe 4 (Kontrollgruppe)	Hauptschule	-0.20	1.02	6
	Realschule	-0.16	1.19	29
	Gymnasium	0.34	0.71	4
	Gesamt	-0.11	1.11	39
Gesamt	Hauptschule	0.14	0.85	51
	Realschule	-0.09	1.03	114
	Gymnasium	0.35	0.99	17
	Gesamt	0.02	0.98	182

Anmerkungen. Alle Variablen sind z-standardisiert. KW = Konditionales Wissen, PLA = Planungsfähigkeit, HW = Handlungswissen.

Tabelle P-10. Deskriptivstatistik: mathematisches Handlungswissen

experimentelle Bedingung	*Schulform*	*M*	*SD*	*N*
Gruppe 1 (KW und PLA)	Hauptschule	-0.40	0.74	15
	Realschule	-0.22	0.81	28
	Gymnasium	0.91	0.78	5
	Gesamt	-0.16	0.86	48
Gruppe 2 (HW und PLA)	Hauptschule	-0.45	0.71	15
	Realschule	0.01	0.89	29
	Gymnasium	0.91	0.34	4
	Gesamt	-0.06	0.88	48
Gruppe 3 (HW und KW)	Hauptschule	-0.62	0.70	15
	Realschule	-0.09	0.78	28
	Gymnasium	1.85	0.53	4
	Gesamt	-0.09	0.97	47
Gruppe 4 (Kontrollgruppe)	Hauptschule	-0.48	0.63	6
	Realschule	-0.21	0.95	29
	Gymnasium	2.06	0.21	4
	Gesamt	-0.02	1.11	39
Gesamt	Hauptschule	-0.49	0.69	51
	Realschule	-0.13	0.86	114
	Gymnasium	1.40	0.73	17
	Gesamt	-0.08	0.94	182

Anmerkungen. Alle Variablen sind z-standardisiert. KW = Konditionales Wissen, PLA = Planungsfähigkeit, HW = Handlungswissen.

Tabelle P-11. Follow-up-ANCOVAs (Komponenten mathematischer Problemlöse-kompetenz)

Quelle der Variation	*Abhängige Variable*	*SS*	*Df*	*MS*	*F*	*Sig.*	η_p^2
KFT-Skala N2	KW [a]	1.711	1	1.711	1.924	.167	.011
	PLA [b]	3.938	1	3.938	4.309	.039	.025
	HW [c]	15.232	1	15.232	27.440	.000	.140
Bedingung	KW	0.940	3	0.313	0.353	.787	.006
	PLA	4.733	3	1.578	1.726	.164	.030
	HW	2.430	3	0.810	1.459	.228	.025
Schulform	KW	8.367	2	4.183	4.705	.010	.053
	PLA	2.949	2	1.475	1.613	.202	.019
	HW	42.839	2	21.419	38.588	.000	.313
Bedingung * Schulform	KW	5.226	6	0.871	0.980	.441	.034
	PLA	11.805	6	1.968	2.153	.050	.071
	HW	5.164	6	0.861	1.551	.165	.052
Fehler	KW	15.256	169	0.889			
	PLA	154.464	169	0.914			
	HW	93.809	169	0.555			
Gesamt	KW	166.536	182				
	PLA	175.177	182				
	HW	162.591	182				

Anmerkungen. KW = Konditionales Wissen, HW = Handlungswissen, PLA = Planungsfähigkeit.
[a] R^2 = .095 (Angepasstes R^2 = .030). [b] R^2 = .118 (Angepasstes R^2 = .055).
[c] R^2 = .418 (Angepasstes R^2 = .377).

Tabelle P-12. Paarweise Vergleiche der mathematischen Skalen (mit Bonferoni-Korrektur)

Skala	*Gruppe i*	*Gruppe j*	$M_i - M_j$	*SE*	*Sig.*	*95 %-KI* Untere Grenze	*95 %-KI* Obere Grenze
KW	Hauptschule	Realschule	-0.182	0.168	.838	-0.588	0.224
		Gymnasium	-0.827	0.271	.008	-1.481	-0.172
	Realschule	Hauptschule	0.182	0.168	.838	-0.224	0.588
		Gymnasium	-0.644	0.247	.030	-1.241	-0.047
	Gymnasium	Hauptschule	0.827	0.271	.008	0.172	1.481
		Realschule	0.644	0.247	.030	0.047	1.241
PLA	Hauptschule	Realschule	0.170	0.170	.957	-0.242	0.582
		Gymnasium	-0.244	0.275	1.000	-0.908	0.420
	Realschule	Hauptschule	-0.170	0.170	.957	-0.582	0.242
		Gymnasium	-0.414	0.250	.299	-1.020	0.191
	Gymnasium	Hauptschule	0.244	0.275	1.000	-0.420	0.908
		Realschule	0.414	0.250	.299	-0.191	1.020
HW	Hauptschule	Realschule	-0.379	0.133	.014	-0.700	-0.058
		Gymnasium	-1.864	0.214	.000	-2.381	-1.346
	Realschule	Hauptschule	0.379	0.133	.014	0.058	0.700
		Gymnasium	-1.485	0.195	.000	-1.956	-1.013
	Gymnasium	Hauptschule	1.864	0.214	.000	1.346	2.381
		Realschule	1.485	0.195	.000	1.013	1.956

Anmerkungen. Basierend auf geschätzten Randmitteln. KW = Konditionales Wissen, HW = Handlungswissen, PLA = Planungsfähigkeit.

Tabelle P-13. Deskriptivstatistik: mathematisches Problemlösen

experimentelle Bedingung	*Schulform*	*M*	*SD*	*N*
Gruppe 1 (KW und PLA)	Hauptschule	-0.25	0.78	15
	Realschule	-0.22	1.11	28
	Gymnasium	0.30	1.40	5
	Gesamt	-0.17	1.04	48
Gruppe 2 (HW und PLA)	Hauptschule	-0.41	0.51	15
	Realschule	-0.04	0.54	29
	Gymnasium	1.20	1.22	4
	Gesamt	-0.05	0.72	48
Gruppe 3 (HW und KW)	Hauptschule	-0.16	0.82	15
	Realschule	-0.25	0.87	28
	Gymnasium	0.74	1.27	4
	Gesamt	-0.14	0.91	47
Gruppe 4 (Kontrollgruppe)	Hauptschule	-0.22	0.76	6
	Realschule	-0.15	1.02	29
	Gymnasium	-0.07	1.19	4
	Gesamt	-0.15	0.98	39
Gesamt	Hauptschule	-0.27	0.71	51
	Realschule	-0.16	0.90	114
	Gymnasium	0.53	1.25	17
	Gesamt	-0.13	0.91	182

Anmerkungen. Alle Variablen sind z-standardisiert. KW = Konditionales Wissen, PLA = Planungsfähigkeit, HW = Handlungswissen.

Tabelle P-14. Paarweise Vergleiche der Skala mathematisches Problemlösen (mit Bonferoni-Korrektur)

					95 %-KI	
Gruppe i	*Gruppe j*	$M_i - M_j$	*SE*	*Sig.*	*Untere Grenze*	*Obere Grenze*
Hauptschule	Realschule	-0.122	0.148	1.000	-0.480	0.236
	Gymnasium	-0.741^*	0.239	.007	-1.319	-0.163
Realschule	Hauptschule	0.122	0.148	1.000	-0.236	0.480
	Gymnasium	-0.619^*	0.218	.015	-1.146	-0.092
Gymnasium	Hauptschule	0.741^*	0.239	.007	0.163	1.319
	Realschule	0.619^*	0.218	.015	0.092	1.146

Anmerkung. Basierend auf geschätzten Randmitteln

Anhang Q. Experiment 3: Überblick der Treatments

Experimentalgruppe:

	Einheit 1	Einheit 2	Einheit 3	Einheit 4	Einheit 5	Einheit 6	Einheit 7	Einheit 8	Einheit 9	Einheit 10	Einheit 11	Einheit 12
Zeit	KW 6	KW 8	KW 9	KW 10	KW 11	KW 12	KW15	KW16	KW 17	KW 18	KW 19	KW 20
5	Einstieg	Einstieg	Einstieg	Einstieg	Einstieg	Einstieg	Einstieg	Einstieg	Einstieg	Einstieg	Einstieg	Einstieg
10	Vorstellungs-runde	Test T1: Problem-lösen und Mathe-matik	Ziele	Rück-meldung Tour-Planer	Rück-meldung EP	Test Routen-Planer (RP)	Rück-meldung RP	Besprechung HA	Besprechung HA und Wdh.	Besprechung HA	Feedback-Bogen	Einsammeln HA
15			Test Tour-Planer (TP)	Übung Einkaufs-Planer (EP) in Partner-arbeit	Übung Urlaubs-Planer (UP) in Partner-arbeit		Eierflug-modell	Instruktion Genetics Lab			Übung Prompting-Card	Test T2 Problem-lösen und Mathematik
20	Vorstellung der Inhalte (Mindmap)											
25									Wasserkrug-Aufgabe	Besprechung Übungsheft Problemlösen	Besprechung Arbeitsblatt Handy-nutzertypen	
30								Vorstellung Strategien				
35	Fragebogen (Demo-graphisches, Selbstein-schätzung)									Wdh. Heuristiken		
40								Übung Genetics Lab		Übung 1 Vor-wärtsarbeiten Sudoku 1	Pause	
45									Pause		Übung „Nachdenken über Alternativen	
50				Pause	Pause				Überblick Heuristiken			
55				Ergebnis-blätter EP	Ergebnis-blätter UP	Pause				Pause		
60						Start Problem-lösen			Vorwärts-arbeiten	Übung 2 Vor-wärtsarbeiten Sudoku 2		
65	Pause		Pause			Übungsheft Problemlösen			Übungen zum Vorwärts-arbeiten & Besprechung		Aufgabe Hobbits & Orks	
70	KFT-N2	Pause	Planspiel-Reflexion & Ergebnis-sicherung	Reflexion	Reflexion					Arbeitsblatt Handy-nutzertypen		
75		Rhetorik-Übung		Arbeits-blatt Planspiel-Transfer				Pause				Zusatzzeit
80					Transfer-übung Kuchen-stücke			Nach-besprechung			Fragebogen „Freude am Denken/ Selbst-einschätzung	
85	Ergebnis-sicherung					Meinungsbild						Pause
90						Merkblatt	HA					Quartals-noten
95	Ausblick		Ausblick	Ausblick			Pause	HA & Ausblick	Feedback & HA	Feedback & HA	HA & Ausblick	Abschluss

Kontrollgruppe:

Zeit	Einheit 1	Einheit 2	Einheit 3	Einheit 4	Einheit 5	Einheit 6	Einheit 7	Einheit 8	Einheit 9	Einheit 10	Einheit 11	Einheit 12
	KW 6	**KW 8**	**KW 9**	**KW 10**	**KW 11**	**KW 12**	**KW15**	**KW16**	**KW 17**	**KW 18**	**KW 19**	**KW 20**
5	Einstieg	Einstieg	Einstieg	Einstieg	Einstieg	Einstieg	Einstieg	Einstieg	Einstieg	Einstieg	Einstieg	Einstieg
10	Vorstellungsrunde	Test T1 Problemlösen und Mathematik	Test Tour-Planer (TP)	Rückmeldung TP	5-Stufen-Modell	Test Routen-Planer (RP)	Rückmeldung RP	Wdh. Merkblatt Bertholt/ Diehl/Kühne	Power-Point: Vertraut machen & Brainstorming	Feedback zu Präsentationen	Feedback-Bogen	Einsammeln HA
15				5-Stufen-Modell	Einstieg Körpersprache, Übung Spiegelbild		Arbeitsblatt Rhetorik-Stationen				Schülervorträge Präsentieren mit PowerPoint	Test T2: Problemlösen und Mathematik
20	Vorstellung der Inhalte (Mindmap)			„Sprechhemmungen und Redeängste“ Selbsteinschätzung & Gruppenarbeit zum Thema Lampenfieber				Vortrag der Kurzreferate in Gruppen		Wdh. mit Merkblatt „PPP gestalten“		
25					Übung Körpersprache deuten				Einführung Power-Point-Präsentation (PPP)			
30										Überarbeitung der PPP		
35	Fragebogen (Demographisches, Selbsteinschätzung)											
40					Übung Haltung & Stand					Einführung Masterfolie, Bilder etc.		
45							Pause					
50					Pause		Arbeitsblatt Kurzreferat	Pause	Pause	Pause	Pause	
55					Arbeitsblatt „Körpersprache bewusst einsetzen“	Pause		Vortrag der Kurzreferate in Gruppen	Erstellung eigener PPP	Übung Präsentationen zum Thema Europa erstellen	Schülervorträge Präsentieren mit PowerPoint	
60			Pause	Pause		Test Übungsheft März	10 Regeln für den guten Vortrag					
65	Pause		Planspiel-Reflexion & Ergebnissicherung	Übungen gegen Lampenfieber								
70	KFT	Pause					Vorbereitung der Kurzreferate					
75		Rhetorik-Übung	Einstieg Präsentieren		Übung Spiegelbild							Zusatzzeit
80			Übung Sprechcollagen			Wdh. Übungen der Vorwoche		Besprechung				
85	Ergebnissicherung				Übung „sicherer Stand“					Ergebnispräsentation	Fragebogen „Freude am Denken/ Selbsteinschätzung	Pause
90												Quartalsnoten
95	Ausblick		Ergebnissicherung & Ausblick	Ergebnissicherung & Ausblick	Ausblick		Ausblick & HA	Ausblick & HA	Feedback & Ausblick	Ausblick & HA	Ausblick	Abschluss

Anhang R. Experiment 3: Arbeitsblatt „Kuchenstücke“

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

Aufgabe: Kuchenstücke

Zum Spielenachmittag hast du Freunde eingeladen und einen runden Sandkuchen gekauft. Ursprünglich hattest du sechs Freunde erwartet, aber jetzt kommen sieben. Kannst du den Kuchen mit drei Schnitten so schneiden, dass jeder von euch ein Stück bekommt? Begründe deine Antwort.

Buchwald, F. (2012). *Arbeitsblatt „Kuchenstücke“*. Essen Seite 1

Anmerkung. Die Grundidee der Aufgabe stammt von Keller (2005).

UNIVERSITÄT DUISBURG ESSEN
Offen im Denken

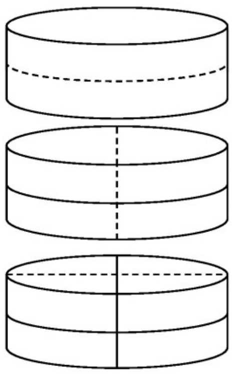
Tipps

- In wie viele Stücke sollst du den Kuchen aufteilen?
- Markiere die wichtigen Angaben in der Aufgabenstellung.
- Versuche die Aufgabe zeichnerisch zu lösen.

Lösung zur Aufgabe Kuchenstücke

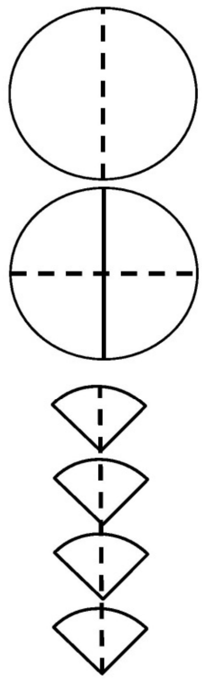
Lösung 1

Ein horizontaler Schnitt und zwei vertikale Schnitte liefern acht Stücke.



Lösung 2

Kuchen mit zwei Schnitten vierteln. Dann die Viertel in einer Reihe anordnen und mit einem Schnitt halbieren.



Arbeitsauftrag (Partnerarbeit, 5 min)

Seht euch die beiden Lösungen gut an und vergleiche sie anschließend. Dabei können folgende Fragen helfen:

a) Ist die Form des Kuchens egal? Funktioniert die Aufteilung auch mit einem Gugelhupf?
b) Funktioniert das Aufteilen auch mit einer Sahnetorte?

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

Didaktischer Kommentar

Die Aufgabenstellung enthält eigentlich keine komplizierten Begriffe. Gegebenenfalls ist durch Erklärung (Sandkuchen = Marmorkuchen) oder Zeigen der Bilder (Seite 6) für Klarheit zu sorgen.

Wichtig ist zu erkennen, dass eine Aufteilung in 8 Stücke gesucht ist (7 Gäste + Gastgeber).

Einige SuS werden die vermeintliche Lösung für 7 Stücken finden.

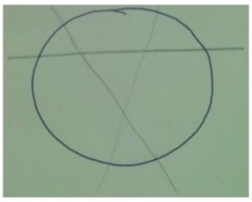


Abbildung 1. Aufteilung in 7 (ungleiche) Stücke

Ziel ist jedoch eine Aufteilung in (mindestens) 8 Stücke. Wenn damit das kleine Problem verstanden ist, kann man sich an den Plan machen. Hier reicht ein kurzer gedanklicher Plan, z. B. „Zeichne einen Kreis und teile ihn mit drei Strichen in 8 Teile auf".

Typischerweise beginnen Personen bei dieser Aufgabe mit dem Zeichnen von Kreisen und drei Strichen (vgl. Abbildung 2). Das führt jedoch hier nicht direkt zum Ziel.

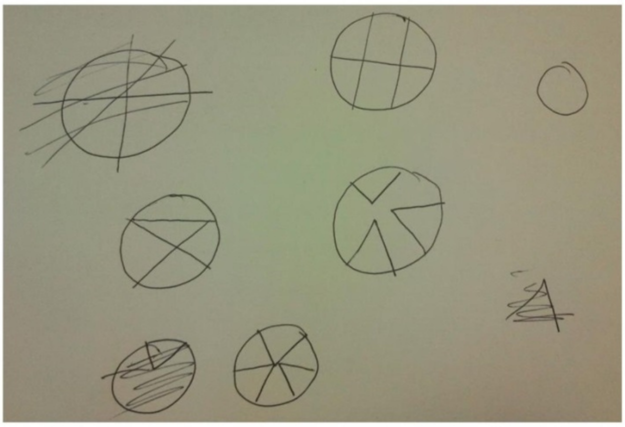


Abbildung 2. Lösungsversuche

Auch das Erkennen, dass der momentane Weg nicht zum Ziel führt ist wichtig. → Plan ändern

Hier: Stelle dir einen Kuchen vor. Er ist kreisförmig, aber dreidimensional und nicht flach wie ein Kreis → 3-dimensionale Lösung suchen!

Buchwald, F. (2012). *Arbeitsblatt „Kuchenstücke"*. Essen Seite 4

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

Anmerkungen (insbesondere wenn viele glauben, eine Aufteilung in 8 Stücke sei nicht möglich):

Man kann die Aufgabe auch ein Stück weit tabellarisch lösen:

Anzahl Schnitte	Anzahl Kuchenstücke
0	1
1	2
2	4
3	8

ODER algebraisch:

Der Kuchen soll auf 8 Personen aufgeteilt werden, d. h. Ziel ist 1 in Achtel zu zerlegen. Annahme: Ein Schnitt verdoppelt die Anzahl der Kuchenstücke. → $\frac{1}{1} = \frac{2}{2} = \frac{4}{4} = \frac{8}{8}$ (An das Gleichheitszeichen „1. Schnitt“, „2. Schnitt“, „3. Schnitt“ schreiben).

In beiden Fällen resultiert die Vermutung: Es geht mit drei Schnitten (da $2^3 = 8$).

Bei beiden Wegen bleibt die Frage: WIE sollen die Schnitte gesetzt werden? → Zeichnen.

Anmerkung zu den Lösungen

Zu Lösung 1: Das Stapeln im 3. Schritt ist nicht nötig . Man kann den Kuchen auch erst mit 2 Schnitten vierteln und dann den horizontalen Schnitt ansetzen.

Es gibt noch eine weitere Lösungsidee: Man viertelt zunächst den Kuchen mit zwei Schnitten. Anschließend schneidet man mit einem „runden“ Schnitt einen Innenkreis heraus. Aber: Unterschiedlich große Teile, z. B. Gugelhupf → Diskussionsfrage: Sind „runde“ Schnitte erlaubt?

Anmerkung zu den Aufgaben

Zu a)

Bei Lösung 1: Symmetrie als Voraussetzung gleich großer Stücke; Kuchen muss stapelbar sein (nicht bei der Variante ohne Stapeln)

Bei Lösung 2: Ja, solange der Kuchen in Länge und Breite achsensymmetrisch ist.

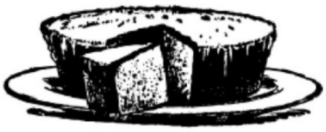
Gegebenenfalls erklären, wie ein Gugelhupf aussieht oder Abbildung zeigen.

Weiterführende Aufgaben / Fragen

Welche impliziten, d. h. stillschweigenden, Annahmen hast du getroffen? Wie sinnvoll waren diese?

- gleich große Stücke
- eine gerade Linie pro Schnitt
- nur vertikale (natürliche) Kuchenschnitte und keine horizontalen
- nicht-bröselnder Sandkuchen
- Was ist „ein Schnitt“?
- Sind auch „runde“ Schnitte zulässig?

Sandkuchen:



Gugelhupf:

Sahnetorte:

Bildquellen:

- http://openclipart.org/detail/167853/cake-by-johnny_automatic
- http://openclipart.org/detail/172461/cake-by-frankes-172461
- http://openclipart.org/detail/173485/cake-by-bleonafoniqi-173485

Anhang S. Experiment 3: Arbeitsblatt „Eierflugmodell“

Institut für Psychologie (RWTH-Aachen) & Lehrstuhl für Lehr-Lernpsychologie (Universität Duisburg-Essen)

Übung: „Eierflugmodell“

Es geht darum als Gruppe gemeinsam eine Aufgabe zu erfüllen. Ziel ist es, ein „Eierflugmodell“ zu bauen. Bewertungskriterien sind erfolgreiche Planung, Konstruktion, und Präsentation unter Einhaltung bestimmter Vorgaben (Zeit, Qualität, begrenzte Arbeitsmaterialien).

Material pro Kleingruppe:

- 1 (Überraschungs-)Ei
- 1 m Schnur
- 1x (Krepp-)klebeband
- 1x Doppelbögen Zeitungen
- 2x Büroklammern
- 2x Luftballons
- 1x Schere
- 1x Planungsbogen für die Gruppenidee

Pro Person:

1 Schmierblatt für die Planungsphase, das nicht verbastelt werden darf!

Ablauf der Übung

Phase	Inhalt	Zeit
Einführung		5
Gruppenbildung	Zufällige Aufteilung der Gruppe in 3 Kleingruppen	3
Materialausgabe	Siehe oben	2
Planungsphase	Diskussion in den Gruppen: Wie gehen wir vor? Wie soll unser Eierflugmodell aussehen? → Plan notieren	10
Bauphase	Herstellen des Eierflugmodells	15
Praxistest	„Funktioniert“ unser Eierflugmodell? Üben der Präsentation	10
Pause		
Präsentationsphase	Jede Gruppe stellt der Reihe nach ihr Eierflugmodell vor.	5 min pro Gruppe
Nachbesprechung	Diskussion im Plenum: Wie hat Eure Gruppe zusammen an der gestellten Aufgabe gearbeitet? Wie ist die Planungsphase verlaufen? Wie hat die Umsetzung des Plans funktioniert? Was ist Euch aufgefallen während Eurer Zusammenarbeit? Bei Bedarf detailliertere Nachfragen der Trainer, ohne die Gruppe auf den Trainer alleine antworten zu lassen – lediglich um die Kommunikation in Gang zu halten: Wer hat welche Aufgabe übernommen? Wurden Aufgaben verteilt (Regelwächter, Zeitwächter Etc.)? Hat jemand die Koordination/Steuerung übernommen? Wie kam es dazu? Wessen Ideen wurde umgesetzt? Welche Absprachen wurden getroffen? Wie kam es dazu? In welcher Lautstärke?	15-20

Anmerkung. Die Übung basiert auf einer Übung von Gilsdorf und Kistner (1998), die stark modifiziert worden ist.

Bewertungsraster

	Gruppe 1	Gruppe 2	Gruppe 3
Vornamen der Schüler			
Planung			
Konstruktion			
Präsentation			
Gesamtbewertung			

Übung „Eierflugmodell"

Aufgabenstellung
Eure Aufgabe ist es, ein „Eierflugmodell" zu planen, zu basteln und zu präsentieren. Die Bewertungskriterien sind Planung, Bau sowie Präsentation.

Ablauf:

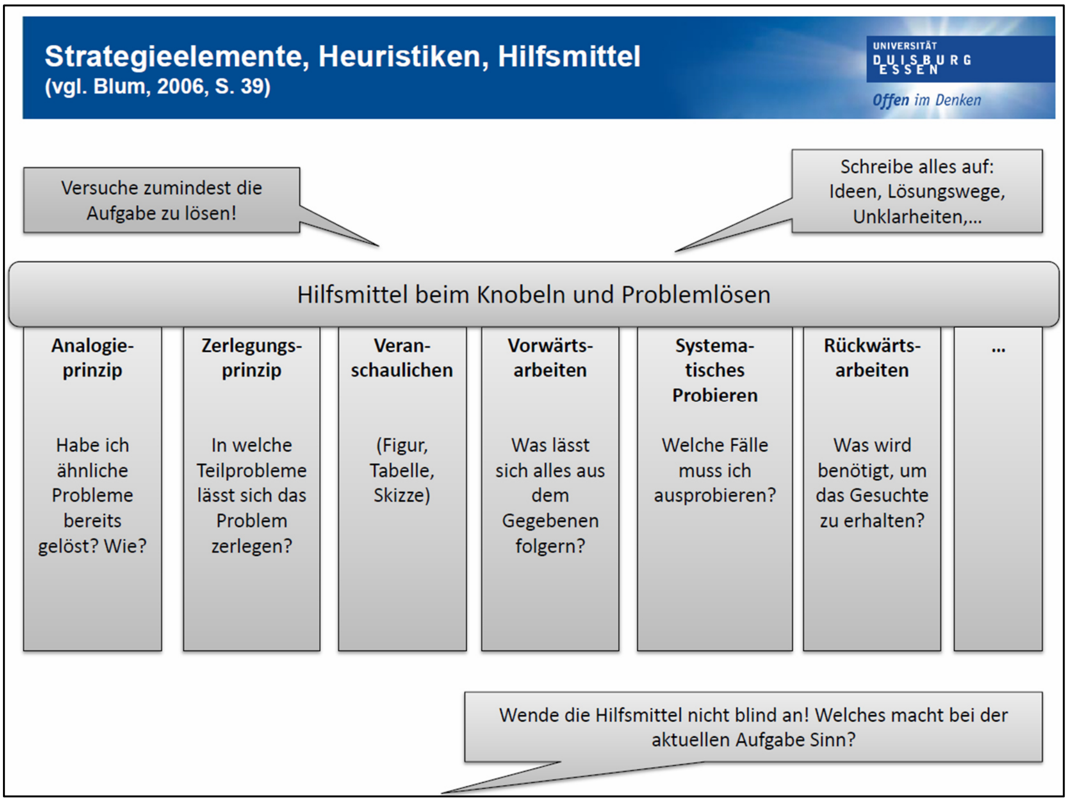
- Das Material, das ihr verwenden dürft, erhaltet ihr gleich. Die Schere könnt ihr benutzen, aber nicht verbasteln.
- Planungsphase (10 min) → Gruppendiskussion: Wie gehen wir vor? Wie soll unser Eierflugmodell aussehen?
  - Haltet euren Plan schriftlich fest, bevor ihr in die Bauphase übergeht.
- Bauphase (15 min) → Bastelt das Eierflugmodell!
- Gruppeninterner Praxistest des Eierflugmodells (10 min): „Funktioniert" unser Eierflugmodell? + Üben der Präsentation
- Pause
- Präsentation: Jede Gruppe hat 5 min Zeit, ihr „Eierflugmodell" zu präsentieren
- Gemeinsame Nachbesprechung mit allen Gruppen

Platz für Notizen:

Gruppenmitglieder:

Unser Plan für das Eierflugmodell

Anhang T. Experiment 3: Folie „Heuristiken“



Anmerkung. Die Referenz auf der Folie bezieht sich auf Blum et al. (2006).

Anhang U. Experiment 3: Arbeitsblatt „Systematisches Probieren“

Offen im Denken

Aufgabe 1a

Schreibe aus deinem Mathe-Buch eine Aufgabe heraus, die du durch systematisches Ausprobieren lösen kannst. Gib auch die Quelle an (Buchtitel, Seitenzahl, ISBN-Nr.).

Aufgabe:

Quelle:

Aufgabe 1b

Begründe für die Aufgabe aus 1a, warum man die Aufgabe durch systematisches Ausprobieren lösen kann.

Aufgabe 2

Denke dir eine Mathe-Aufgabe aus, die man durch systematisches Ausprobieren lösen kann.

Buchwald, F. (2012). Arbeitsblatt „Systematisches Probieren“. Essen.

Anhang V. Experiment 3: Arbeitsblatt „Sudoku“

Sudoku

“Eines der beliebtesten Rätsel der Welt: Sudoku! Ein Sudoku besteht aus drei mal drei Quadraten, die jeweils wieder dreimal drei Felder haben. In jedem dieser Neuner-Quadrate [dicker umrandet], aber auch in jeder Zeile und jeder Spalte müssen alle Zahlen von 1 bis 9 [genau einmal] vorkommen” (http://www.sueddeutsche.de/app/spiele/sudoku/).

5	9					3	7	1
			9				6	
	3	8	1	4				
	8	1	3	6		4		
9	6	5	2				8	
	4			8		6		7
		7	4		2		3	6
	2	9			8		1	5
3	1	6		9		2	4	

http://www.surfpoeten.de/sudoku/generator

Aufgaben

1. Erinnere dich an die letzte Woche: Welche Heuristik/Strategie kannst du bei Sudokus anwenden?
2. Versuche das Sudoku zu lösen.
3. Erkläre deinem Sitznachbarn, wann und wie du die Heuristik/Strategie bei deiner Lösung benutzt hast.

Übungsaufgaben

Leicht

		4			6		2	3
			9	2	5		4	
2	8	5	7			9		
9		7		4				
		3			8	2		9
			3	9				4
5		6	4		9			
	4	9		5		8		2
							9	

mittel

4		6				1		
		3	5	6	2			
5			4		9			
						5		
	8							1
			8		3	7		
8			9					
				8		9		3
1	7		2		4			8

schwer

		7			2			
					7			
	2		9	4				1
1	8	4			9			
5	7						8	
2	9		4					
9		8	5	2		4		
3					8		6	
		1		3			2	

Anmerkungen. Mögliche Strategien beim Sudoku sind z. B. Versuch und Irrtum, Ausschlussprinzip, systematisches Vorgehen (Reihen, Zeilen), Teilziele (sichere Kästchen zuerst, vervollständige einen Block), Notizen machen (z. B. kleine Zahlen an den Rand) und Kombinationen davon. Für spezifische Strategien beim Sudoku siehe z. B. http://www.sudoku-space.de/sudoku-loesungstechniken/.

Anhang W. Experiment 3: Arbeitsblatt „Handynutzertypen“

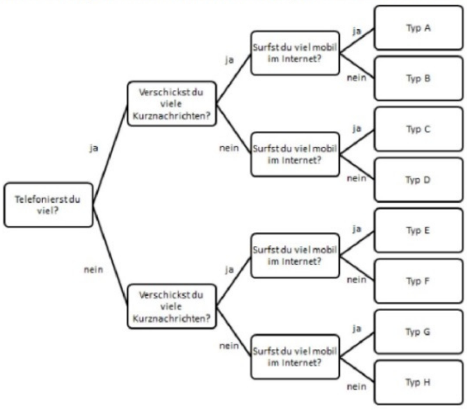
UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

Handynutzertypen

A1

Als Informatikprojekt wollen Udo und Mareike eine Internetseite erstellen, die zunächst erfragt, welchem Handynutzertyp der Besucher entspricht, und ihm dann Vorschläge für Handys bzw. Smartphones unterbreitet. Udo und Mareike befragen ihre Freunde und Bekannten. Im Wesentlichen unterscheiden die Befragten in ihren Antworten auf die folgenden drei Fragen: Telefonierst du viel? Verschickst du viele Kurznachrichten? Surfst du viel mobil im Internet?
Um sich das Programm besser vorstellen zu können, visualisieren Udo und Mareike ihre Ergebnisse.



Gib an, zu welchem Handynutzertyp die Personen jeweils gehören.

	Typ A	Typ B	Typ C	Typ D	Typ E	Typ F	Typ G	Typ H
Edgar findet Internet auf seinem Handy überflüssig, telefoniert regelmäßig und schickt viele SMS.	○	○	○	○	○	○	○	○
Jason nutzt sein Handy nur zum Verschicken von Kurznachrichten.	○	○	○	○	○	○	○	○
Melinda telefoniert kaum, verschickt aber viele Kurznachrichten und surft viel im Internet.	○	○	○	○	○	○	○	○

Laura telefoniert bis zu zwei Stunden täglich und hört gerne Internetradio. Zu welchen Handynutzertypen könnte sie gehören?

☐ Typ A ☐ Typ B ☐ Typ C ☐ Typ D ☐ Typ E ☐ Typ F ☐ Typ G ☐ Typ H

A2

Denke dir selbst eine Aufgabe aus, die der aus Aufgabe A1 ähnlich ist. Schreibe sie auf die Rückseite und stelle sie deinem Nachbarn.

A3

Gibt es noch weitere Kriterien, die Udo und Mareike für ihr Programm verwenden können?

Buchwald, F. (2012). Arbeitsblatt „Handynutzertypen“. Essen.

Anhang X. Experiment 3: Arbeitsblatt „Nachdenken über Alternativen“

Nachdenken über Alternativen

UNIVERSITÄT DUISBURG ESSEN
Offen im Denken

Heute beschäftigen wir uns mit Beispielen, bei denen es mehrere Handlungsmöglichkeiten gibt. Das kennst du auch aus dem Alltag. Auch in der Schule musst du überlegen, wie du eine konkrete Aufgabe bearbeitest. Dabei musst du dich zunächst für ein Vorgehen (einen Plan, eine Option, eine Strategie) entscheiden.

- Oft können verschiedene Wege zum Ziel führen.
- Diese Lösungswege unterscheiden sich in ihrer Qualität.
- Daher lohnt es sich, über verschiedene Lösungsmöglichkeiten nachzudenken.

Abbildung 1. Der Problemlöseprozess

Übungsaufgaben

Partnerarbeit: Lest euch die folgenden Beispiele durch. Diskutiert dann leise darüber, wie gut die jeweiligen Handlungsmöglichkeiten in der beschriebenen Situation geeignet sind. Benotet die Möglichkeiten anschließend.

Johanna hat den Verdacht, dass ihr W-LAN-Router nicht richtig funktioniert. Denn seit einer Stunde kommt sie mit ihrem Laptop nicht ins Internet. Was könnte sie sinnvollerweise als nächstes tun?
Bewerte die folgenden Möglichkeiten, indem du pro Antwort eine Schulnote ankreuzt.

	1	2	3	4	5	6
Einen neuen Router kaufen.	○	○	○	○	○	○
Versuchen mit einem anderen Gerät per W-LAN ins Internet zu gelangen.	○	○	○	○	○	○
Neustarten des W-LAN-Routers.	○	○	○	○	○	○
Abwarten. Das Problem löst sich möglicherweise von selbst.	○	○	○	○	○	○
Den Wartungsdienst anrufen, um einen Techniker zu bestellen.	○	○	○	○	○	○
Dem Laptop gut zureden, damit er sich mit dem W-LAN verbindet.	○	○	○	○	○	○

Platz für Anmerkungen:

Buchwald, F. (2012). Arbeitsblatt „Nachdenken über Alternativen“ (Version 1.1). Essen.

Nachdenken über Alternativen

UNIVERSITÄT DUISBURG ESSEN
Offen im Denken

Jonas überlegt, sich ein Smartphone zuzulegen. Dazu möchte er die aktuelle Version der Betriebssysteme Android und iOS vergleichen. Was könnte er tun?
Bewerte die folgenden Möglichkeiten, indem du pro Antwort eine Schulnote ankreuzt.

	1	2	3	4	5	6
Jonas geht in ein Geschäft und probiert Smartphones mit beiden Betriebssystemen aus.	○	○	○	○	○	○
Jonas informiert sich im Internet.	○	○	○	○	○	○
Jonas fragt im Bekanntenkreis nach, ob jemand Erfahrung mit einem der beiden Betriebssysteme hat.	○	○	○	○	○	○
Jonas lässt sich im Fachhandel beraten.	○	○	○	○	○	○
Jonas sucht in Fachzeitschriften nach einem Vergleichsartikel	○	○	○	○	○	○

Platz für Anmerkungen:

Der 20-jährige Kölner Max W. will seine Kondition verbessern. Nun ist er auf der Suche nach einem guten Fitness-Studio, um dort regelmäßig zu trainieren.
Bewerte die folgenden Möglichkeiten, indem du pro Antwort eine Schulnote ankreuzt.

	1	2	3	4	5	6
Max fragt in seinem Freundeskreis nach Erfahrungen.	○	○	○	○	○	○
Max achtet auf die Anzeigen der Fitness-Studios in der Lokalzeitung.	○	○	○	○	○	○
Max recherchiert im Internet.	○	○	○	○	○	○
Max geht zum Probetraining in jedem Fitness-Studio in seiner Stadt.	○	○	○	○	○	○
Max geht zum Probetraining im nächstgelegenen Fitness-Studio.	○	○	○	○	○	○

Platz für Anmerkungen:

Buchwald, F. (2012). Arbeitsblatt „Nachdenken über Alternativen“ (Version 1.1). Essen.

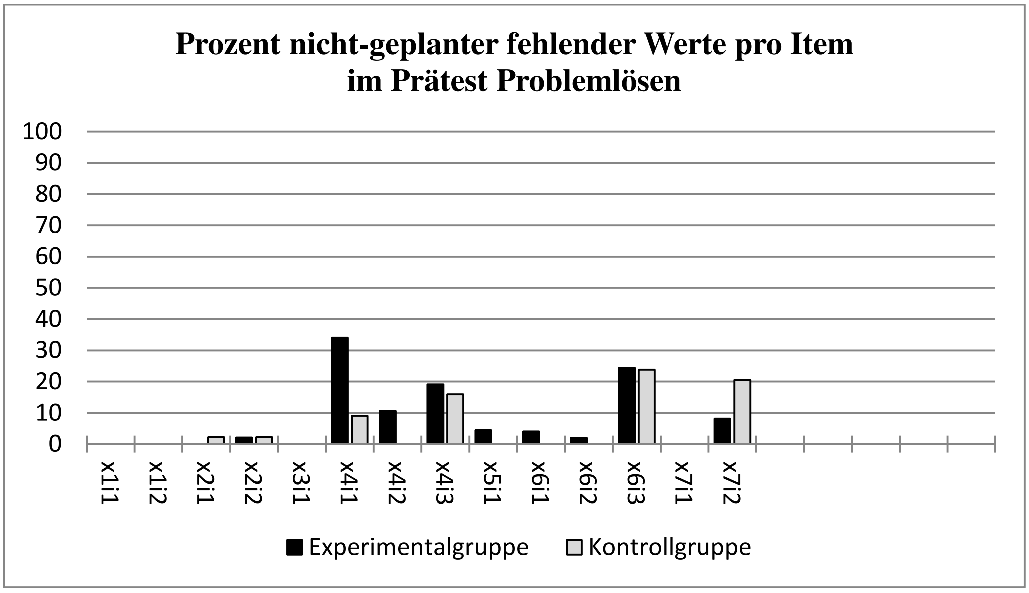
Anhang Y. Experiment 3: Eingesetzte Testaufgaben

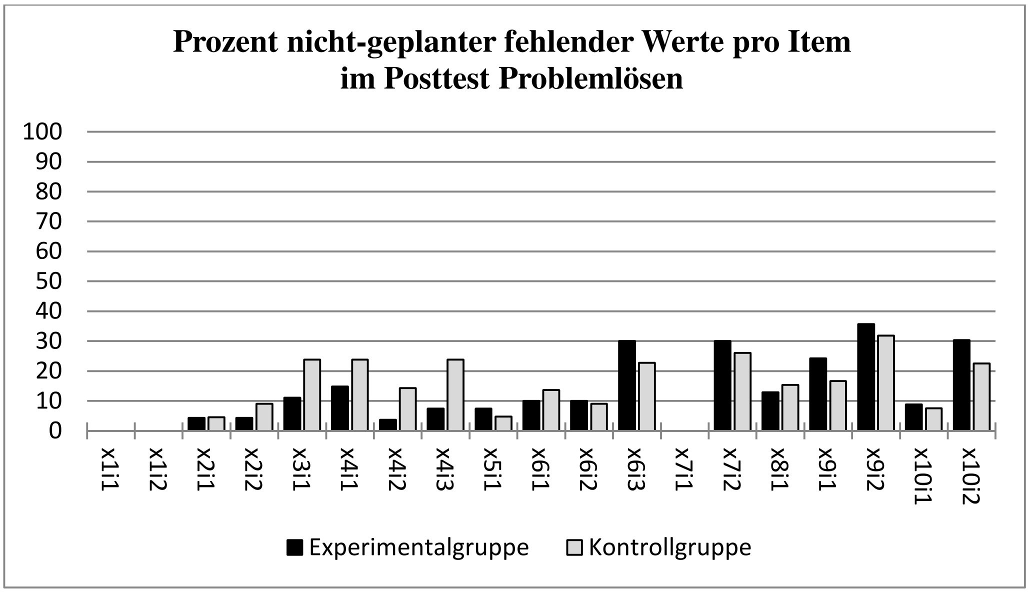
interner Code	*Cluster*	*Aufgabe*	*Item*	*Quelle*
x11i1	M1	Auswahl	M510Q01T	OECD (2004a)
x12i1	M1	Hölzchen und Boxen	B016	Lernstand 9 [a]
x12i2	M1	Hölzchen und Boxen	B015r	Lernstand 9 [a]
x13i3	M1	Blindenschrift	M442Q02	OECD (2004a)
x14i1	M2	Der Geburtstagskuchen	A024r	Lernstand 9 [a]
x15i1	M2	Physiktests	M468Q01T	OECD (2004a)
x16i1	M2	Würfelspiel	X196r	Lernstand 9 [a]
x16i2	M2	Würfelspiel	X195r	Lernstand 9 [a]
x17i1	M2	Ansicht eines Zimmers	M033Q01	OECD (2004a)
x18i1	M3	Quadrat im Kreis	B126r	Lernstand 9 [a]
x19i1	M3	Skateboard Frage 1	M520Q01T	OECD (2004a)
x19i2	M3	Skateboard Frage 2	M520Q02	OECD (2004a)
x19i3	M3	Skateboard Frage 3	M520Q03T	OECD (2004a)
x20i1	M4	Schnurr	B136	Lernstand 9 [a]
x21i1	M4	Dreieckszahlen	pug3	Pugalee (2001)
x22i1	M4	Zimmernummern	pug1	Pugalee (2001)
x23i1	M4	Die dritte Seite	M462Q01T	OECD (2004a)
x1i1	PL1	Gefrierschrank	X423Q01T	OECD (2004b)
x1i2	PL1	Gefrierschrank	X423Q02T	OECD (2004b)
x2i1	PL1	Kinobesuch	X601Q01	OECD (2004b)
x2i2	PL1	Kinobesuch	X601Q02	OECD (2004b)
x3i1	PL1	Studienplan	X414Q01	OECD (2004b)
x4i1	PL2	Bewässerung	X603Q02T	OECD (2004b)
x4i2	PL2	Bewässerung	X603Q03	OECD (2004b)
x4i3	PL2	Bewässerung	X603Q01	OECD (2004b)
x5i1	PL2	Anschlusszüge	X415Q01T	OECD (2004b)
x6i1	PL3	Design by Numbers	X412Q02	OECD (2004b)
x6i2	PL3	Design by Numbers	X412Q01	OECD (2004b)
x6i3	PL3	Design by Numbers	X412Q03	OECD (2004b)
x7i1	PL3	Energiebedarf	X430Q01	OECD (2004b)
x8i1	PL4	Ferienlager	X417Q01	OECD (2004b)
x9i1	PL4	Urlaub	X602Q01	OECD (2004b)
x9i2	PL4	Urlaub	X602Q02	OECD (2004b)
x10i1	PL4	Bibliothekensystem	X402Q01T	OECD (2004b)
x10i2	PL4	Bibliothekensystem	X402Q02T	OECD (2004b)

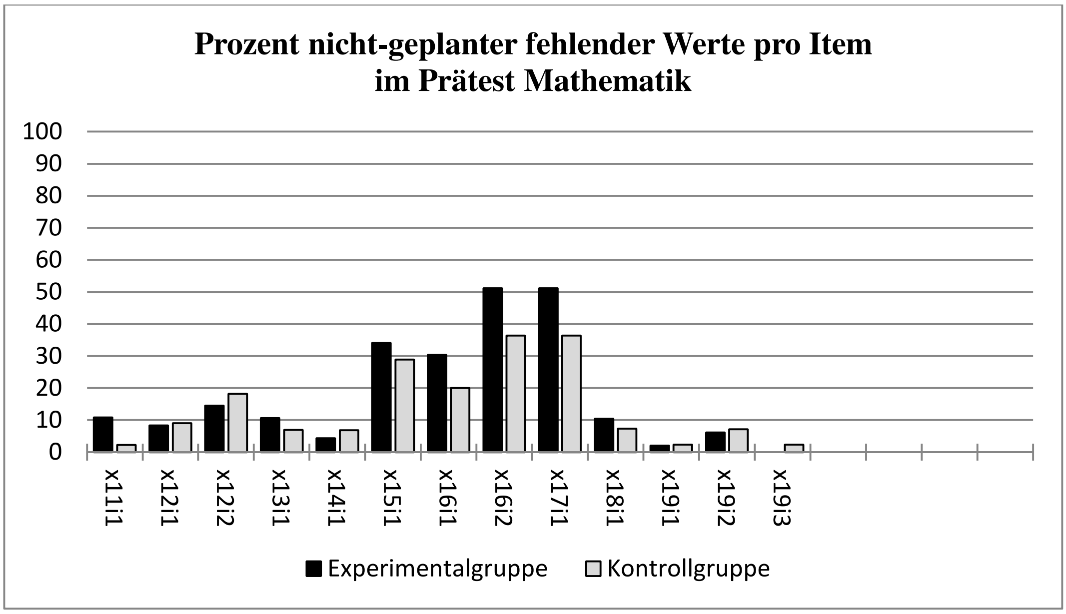
Anmerkung.
[a] Quelle: Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2005a, 2005b.

Anhang Z. Experiment 3: Nicht-geplante fehlende Werte pro Item

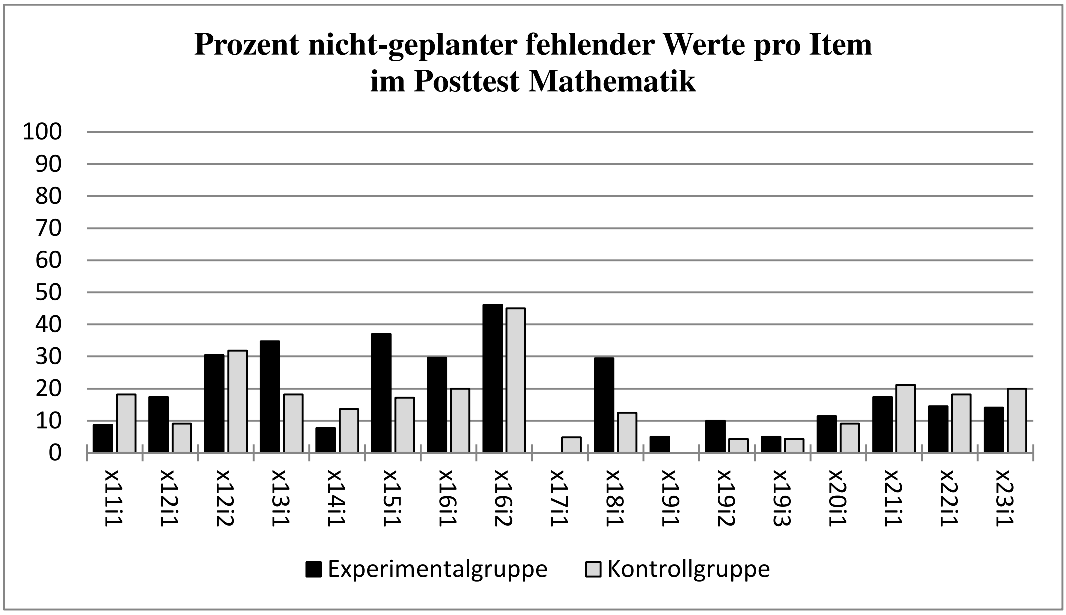
Anmerkung: Man beachte, dass die Prozentangaben von Prä- und Posttest nur bedingt vergleichbar sind, da sie auf einem multi-matrix Testheftdesign mit unterschiedlich vielen Aufgaben pro Itemcluster beruhen, wobei niemand dasselbe Item in Prä- und Posttest bearbeitet hat.







Prozent nicht-geplanter fehlender Werte pro Item
im Prätest Mathematik
100
90
80
70
60
50
40
30
20
10
0
x11i1
x12i1
x12i2
x13i1
x14i1
x15i1
x16i1
x16i2
x17i1
x18i1
x19i1
x19i2
x19i3
Experimentalgruppe
Kontrollgruppe



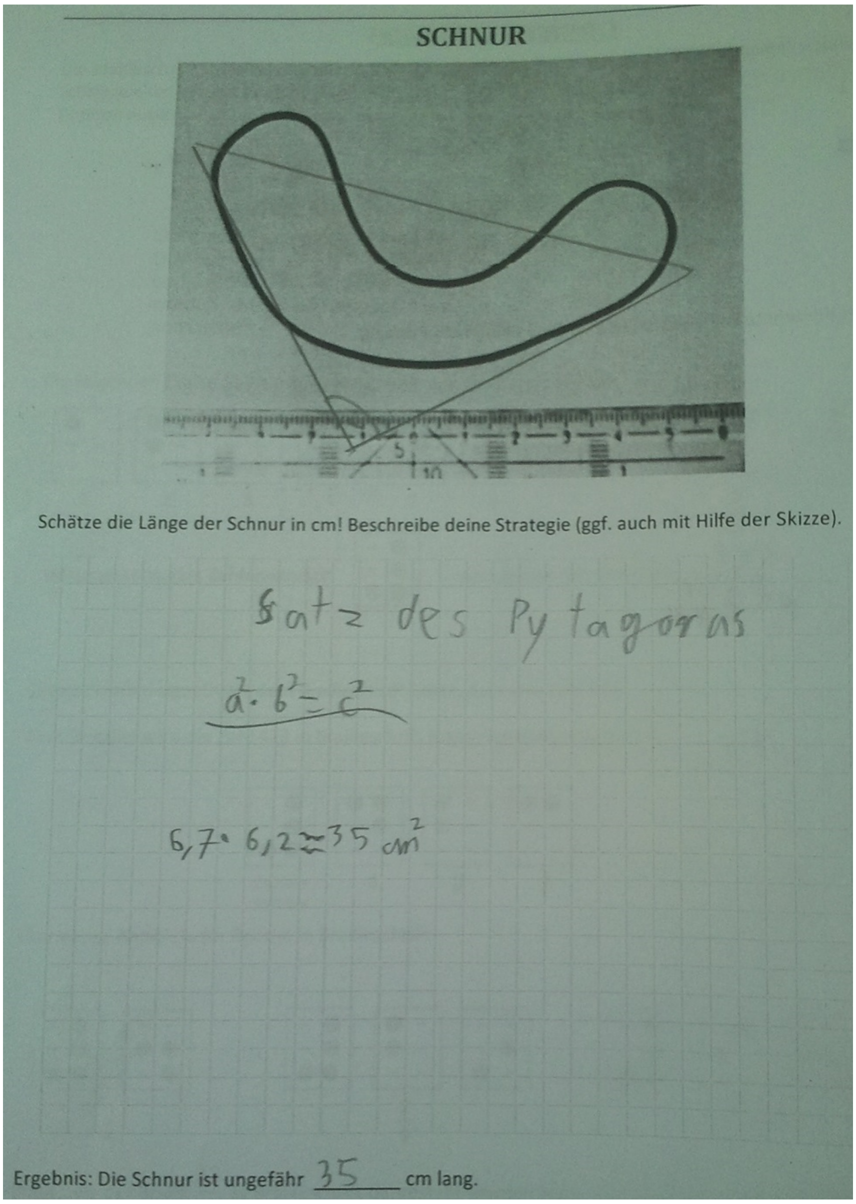
Prozent nicht-geplanter fehlender Werte pro Item
im Posttest Mathematik
100
90
80
70
60
50
40
30
20
10
0
x11i1
x12i1
x12i2
x13i1
x14i1
x15i1
x16i1
x16i2
x17i1
x18i1
x19i1
x19i2
x19i3
x20i1
x21i1
x22i1
x23i1
Experimentalgruppe
Kontrollgruppe

Anhang AA. Experiment 3: Typische Schülerfehler bei den Testaufgaben

Code	*Aufgabe*	*Typische Fehler*
x23i1	Die dritte Seite	- Angabe einer konkreten Länge statt eines Wertebereichs - Mangelndes Sachwissen („Wertebereich“) - Implizite Annahme, dass das Dreieck rechtwinklig sei.
x9i1	Urlaub	- Mangelndes Verständnis der Entfernungstabelle - Nur eine Strecke berechnet, nicht die kürzeste Strecke
x22i1	Zimmernummern	- Mangelndes Sachwissen (Ziffern und Zahlen verwechselt) - Nullen bei den Ziffern vergessen
x8i1	Ferienlager	- Aufsichtspersonen (Betreuer) nicht berücksichtigt

Anmerkung. Die Tabelle beruht auf den Rückmeldungen der Trainer in einer Nachbesprechung.

Anhang BB. Experiment 3: Beispiel für negativen Transfer



Anmerkung. Der Schüler approximiert zunächst das Objekt durch ein rechtwinkliges Dreieck. Der Umfang dieses Dreiecks würde eine gute Schätzung der Länge der Schnurr liefern. Statt die Seitenlängen des Dreiecks zu schätzen bzw. zu messen und zu addieren, versucht der Schüler den „Satz des Pytagoras [sic!]" anzuwenden. Dabei wird eine falsche Formel verwendet ($a^2 \cdot b^2 = c^2$), in die Längen ohne Quadrate (und Einheiten) eingesetzt werden. Das ungenaue Schätzergebnis (35 cm²) von c² wird dann als Länge der Schnurr angegeben.

Anhang CC. Experiment 3: Skalierung des Prä- und Posttests

Die Items des Prä- und Posttests wurden dichotom gescort und anschließend pro Domäne mit dem R-Paket *TAM* (Kiefer et al., 2014) skaliert.[27] Da die Stichprobe für IRT-Zwecke eher klein ist und die Qualität von *seperate calibration* von der Präzision der Schätzungen in der Referenzstichprobe abhängt, wird die effiziente Methode des *concurrent calibration* verwendet, bei der alle Testzeitpunkte gemeinsam skaliert werden, um die maximal verfügbare Information zu nutzen (Kolen & Brennan, 2014). Dabei wurden die Itemschwierigkeit von Items, die zu beiden Testzeitpunkten vorkamen, zunächst auf Gleichheit fixiert. Da sich einige Items in ihrer Schwierigkeit beim Prä- und Posttest deutlich unterscheiden (displacement, O'Neill, Peabody, Tan & Du, 2013), wurde die entsprechenden Fixierungen aufgehoben, um bessere Fit-Indices (infit und vor allem outfit) zu erhalten.[28] Tabelle BB-1 enthält die jeweiligen Itemschwierigkeiten der finalen Skalierung. Abbildung BB-1 zeigt die Histogramme der Itemschwierigkeiten und WLEs für Prä- und Posttest der beiden Domänen Problemlösen und Mathematik. Auffällig ist, dass für beide Domänen die Items eher zu schwer sind. Tabelle BB-2 und BB-3 beinhalten die Zusammenfassung der Skalierungsergebnisse.

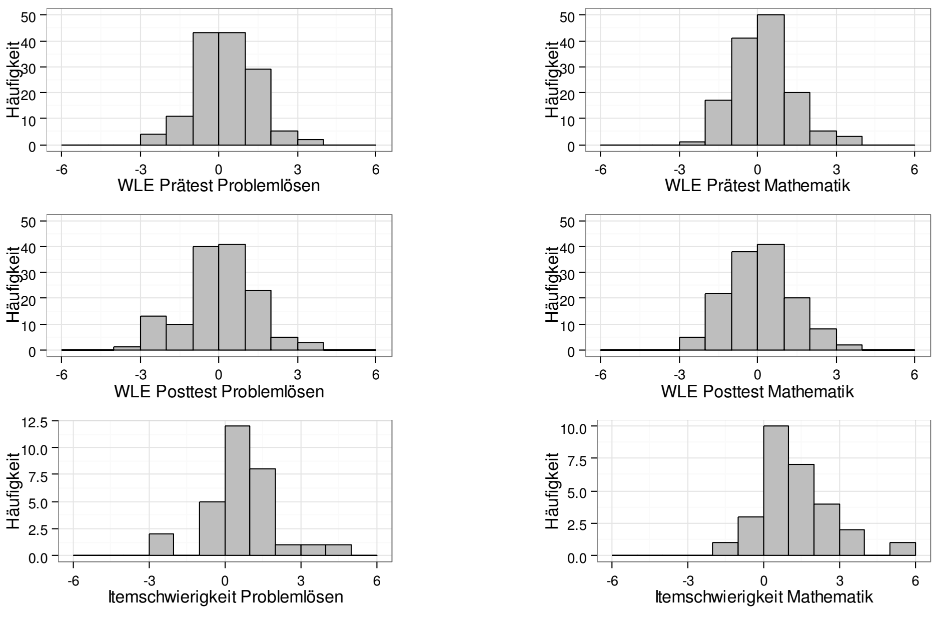


Abbildung BB-1. Histogramme der Itemschwierigkeiten und WLEs für Prä- und Posttest der Domäne Problemlösen und Mathematik

[27] Ich danke Thomas Kiefer für die Beantwortung technischer Rückfragen.
[28] Mögliche Gründe für unterschiedliche Itemschwierigkeiten in Prä- und Posttest können technischer Art (eher kleine Stichprobe für IRT-Zwecke, Stichprobenfehler) oder inhaltlicher Art sein (Lerngewinn der SuS, andere Motivation in Prä- oder Posttest).

Tabelle BB-1. Tabelle der Itemschwierigkeiten (Xsi) und Standardfehler (SE) der Domäne Problemlösen und Mathematik

Domäne	*Prätest*			*Posttest*		
Problemlösen	*Item*	*Xsi*	*SE*	*Item*	*Xsi*	*SE*
	i1t1	.67	.23	i1t2	.51	.33
	i2t1	.93	.23	i2t2	.85	.35
	i3t1	-.22	.23	i3t2	-.46	.33
	i4t1	-.01	.22	i4t2	.18	.33
	i5t1*	1.61	.22			
	i6t1	-.28	.23	i6t2	.08	.31
	i7t1	.92	.24	i7t2	.48	.32
	i8t1	.34	.23	i8t2	.18	.31
	i9t1*	2.84	.31			
	i10t1	1.34	.25	i10t2	.52	.34
	i11t1	.56	.23	i11t2	1.02	.36
	i12t1	1.34	.25	i12t2	1.15	.37
	i13t1	-2.02	.34	i13t2	-2.96	.62
	i14t1*	3.12	.36			
				i15t2	1.98	.25
				i16t2	1.17	.21
				i17t2	1.86	.24
				i18t2	-.99	.20
				i19t2	4.16	.59
Mathematik	*Item*	*Xsi*	*SE*	*Item*	*Xsi*	*SE*
	i1t1	.01	.24	i1t2	-.12	.33
	i2t1	1.62	.28	i2t2	2.20	.44
	i3t1	.00	.23	i3t2	.65	.34
	i4t1	1.27	.27	i4t2	1.68	.39
	i5t1	.27	.24	i5t2	.31	.32
	i6t1*	2.07	.25			
	i7t1*	.99	.20			
	i8t1	2.08	.32	i8t2	1.58	.39
	i9t1	2.08	.32	i9t2	-1.51	.37
	i10t1	1.87	.30	i10t2	3.29	.77
	i11t1	.02	.23	i11t2	-.70	.35
	i12t1	1.19	.26	i12t2	.36	.35
	i13t1	.87	.25	i13t2	.24	.34
				i14t2	1.54	.22
				i15t2	.99	.20
				i16t2	3.40	.40
				i17t2	5.38	1.01

Anmerkungen. Bei mit * gekennzeichnete Items wurde die Schwierigkeit in Prä- und Posttest auf Gleichheit fixiert.

Tabelle BB-2. Zusammenfassung der Skalierung Problemlösen (R-Output)

```
## TAM 1.0-2 (2014-03-26)
##
## Date of Analysis: 2014-04-24 13:54:32
## Time difference of 26.09 secs
## Computation time: 26.09
## R version 3.0.2 (2013-09-25) x86_64, mingw32 | nodename = DUE-
MOBIL | login = buchwald
##
## Multidimensional Item Response Model in TAM
##
## IRT Model 1PL
## ------------------------------------------------------------
## Number of iterations = 522
## Numeric integration with 441 integration points
##
## Deviance =  2666  | Log Likelihood =  -1333
## Number of persons =  173
## Number of items =  33
## Number of estimated parameters =  34
##     Item threshold parameters  =  30
##     Item slope parameters      =  0
##     Regression parameters      =  1
##     (Co)Variance parameters    =  3
##
## AIC  =  2734  | penalty = 68    | AIC = -2*LL + 2*p
## AICc =  2752  | penalty = 85.25    | AICc = -2*LL + 2*p +
   2*p*(p+1)/(n-p-1)  (bias corrected AIC)
## BIC  =  2842  | penalty = 175.2    | BIC = -2*LL + log(n)*p
## aBIC  =  2733  | penalty = 66.76    | aBIC = -2*LL + log((n-
   2)/24)*p  (adjusted BIC)
## CAIC =  2876  | penalty = 209.2   | CAIC = -2*LL +
   [log(n)+1]*p  (consistent AIC)
##
## ------------------------------------------------------------
## EAP Reliability
##  Dim1  Dim2
## 0.480 0.491
## ------------------------------------------------------------
## Covariances and Variances
##       V1    V2
## V1 0.709 0.499
## V2 0.499 0.924
## ------------------------------------------------------------
## Correlations and Standard Deviations (in the diagonal)
##       V1    V2
## V1 0.842 0.617
## V2 0.617 0.961
## ------------------------------------------------------------
## Regression Coefficients
##          V1 V2
## [1,] 0.2282  0
## ------------------------------------------------------------
```

Tabelle BB-3. Zusammenfassung der Skalierung Mathematik (R-Output)

```
## TAM 1.0-2 (2014-03-26)
##
## Date of Analysis: 2014-04-24 13:47:23
## Time difference of 16.18 secs
## Computation time: 16.18
## R version 3.0.2 (2013-09-25) x86_64, mingw32 | nodename = DUE-
   MOBIL | login = buchwald
##
## Multidimensional Item Response Model in TAM
##
## IRT Model 1PL
## ------------------------------------------------------------
## Number of iterations = 301
## Numeric integration with 441 integration points
##
## Deviance =  2242  | Log Likelihood =  -1121
## Number of persons =  173
## Number of items =  30
## Number of estimated parameters =  32
##     Item threshold parameters  =  28
##     Item slope parameters      =  0
##     Regression parameters      =  1
##     (Co)Variance parameters    =  3
##
## AIC  =  2306  | penalty = 64    | AIC = -2*LL + 2*p
## AICc =  2321  | penalty = 79.09    | AICc = -2*LL + 2*p +
   2*p*(p+1)/(n-p-1)  (bias corrected AIC)
## BIC  =  2407  | penalty = 164.9    | BIC = -2*LL + log(n)*p
## aBIC  =  2305  | penalty = 62.84    | aBIC = -2*LL + log((n-
   2)/24)*p  (adjusted BIC)
## CAIC =  2439  | penalty = 196.9   | CAIC = -2*LL +
   [log(n)+1]*p  (consistent AIC)
##
## ------------------------------------------------------------
## EAP Reliability
##  Dim1  Dim2
## 0.576 0.541
## ------------------------------------------------------------
## Covariances and Variances
##       V1    V2
## V1 1.303 0.959
## V2 0.959 1.115
## ------------------------------------------------------------
## Correlations and Standard Deviations (in the diagonal)
##       V1    V2
## V1 1.141 0.796
## V2 0.796 1.056
## ------------------------------------------------------------
## Regression Coefficients
##           V1 V2
## [1,] 0.01491  0
```

Anhang DD. Experiment 3: Dokumentation der multiplen Imputation

Die multiple Imputation in Experiment 3 wurde mit R (R Core Team, 2014) und den R-Paketen mice (van Buuren & Groothuis-Oudshoorn, 2011) durchgeführt. Folgende Schritte sind dabei zu beachten (van Buuren & Groothuis-Oudshoorn, 2011):

1. Prüfe, ob die MAR-Annahme (missing at random) plausibel ist: Durch multi matrix sampling entstandene fehlende Werte, d. h. pro Testheftversion nicht-administrierte Testaufgaben (*missing by design*), sind per constructionem als *missing completely at random* (MCAR) annehmbar (Lüdtke et al., 2007), d. h. die MCAR-Voraussetzung zur Ersetzung dieser missings auf Grundlage von *multipler Imputation* (MI) ist gegeben (Graham, 2012). Um explorativ zu prüfen, ob es zu selektiven Ausfällen im Trainingsverlauf gekommen ist, ob beispielsweise leistungsschwächere SuS dem nicht-angekündigten Posttest fernblieben, werden *margin plots* mit dem R-Paket *VIM* (Templ et al., 2013) betrachtet. Die grauen Punkte an den linken bzw. unteren Rändern in Abbildung DD-1 streuen über den gesamten Leistungsbereich, was gegen selektiven Stichprobenausfall im Prä- und Posttest spricht.
2. Wähle die Form des Imputationsmodells: Als Imputationstechnik für kategoriale Variablen wurde das *multinomial logit model* und für metrische Variablen *predictive mean matching* verwendet (van Buuren & Groothuis-Oudshoorn, 2011).
3. Spezifiziere die Prädiktormatrix im Imputationsmodell (Tabelle DD-1 und DD-2): Welche Variablen sollen zur Ersetzung fehlender Werte pro Variable verwendet werden? Basierend auf einer Betrachtung der Korrelationsmatrix des unvollständigen Datensatzes wurden diejenigen Variablen v_i zur Schätzung der imputierten Werte der Variable x ausgewählt, die mindestens .30 mit x korrelieren. Zusätzlich wurden die Hypothesen-relevanten Variablen Bedingung, Prä- und Posttestleistungen der Domänen Problemlösen und Mathematik sowie die passiv imputierten Interaktionsterme potentieller ATI-Effekte im Imputationsmodell verwendet (van Buuren & Groothuis-Oudshoorn, 2011).

4. Wähle den Umgang mit transformatierten Variablen wie Interaktionsvariablen: Gemäß den Empfehlungen von van Buuren und Groothuis-Oudshoorn (2011) wurden Interaktionsterme passiv imputiert.
5. Lege die Reihenfolge fest, in der die Variablen imputiert werden: Tabelle DD-3 dokumentiert die verwendete Reihenfolge der Variablen.
6. Wähle die Eigenschaften der Start-Imputation und die Anzahl der Iterationen, bevor ein imputierter Datensatz herausgeschrieben wird: Es wurden jeweils 15 Iterationen zwischen den Imputationen durchgeführt.
7. Wähle die gewünschte Anzahl imputierter Datensätze: Es wurden 20 Datensätze imputiert, da die Voreinstellung von 5 Datensätzen, die auf eine Faustregel Rubins (1987) zurückgeht, inzwischen als suboptimal betrachtet wird (Graham et al., 2007; Robitzsch, 2014b).

Zur graphischen Überprüfung der Qualität der Konvergenz des MI-Algorithmus wurden *trace line plots* mit dem R-Paket *mice* (van Buuren & Groothuis-Oudshoorn, 2011) erzeugt (Abbildung DD-2). Trace line plots sind Liniendiagramme, die die geschätzten Werte für Mittelwert und Standardabweichung gegen die Nummer der Iteration des MI-Algorithmus‘ darstellen. Eine Linie repräsentiert dabei einen imputierten Datensatz. Der Linienverlauf in Abbildung DD-2 spricht für gutes Konvergenzverhalten (vgl. van Buuren & Groothuis-Oudshoorn, 2011).

Mit dem R-Paket *mice* (van Buuren & Groothuis-Oudshoorn, 2011) wurden zusätzlich Boxplots erstellt, um die Qualität der Imputation der Leistung im Prä- und Posttest für die Domänen Problemlösen und Mathematik graphisch zu beurteilen (Abbildung DD-3 und DD-4). Die Qualität der Boxplots ist unauffällig. Alle Boxplots überschneiden sich deutlich, auch wenn Median und Boxgröße durch den Algorithmus per constructionem schwanken. Das zeigt auch wieder, dass fünf Imputationen zu wenig sein können.

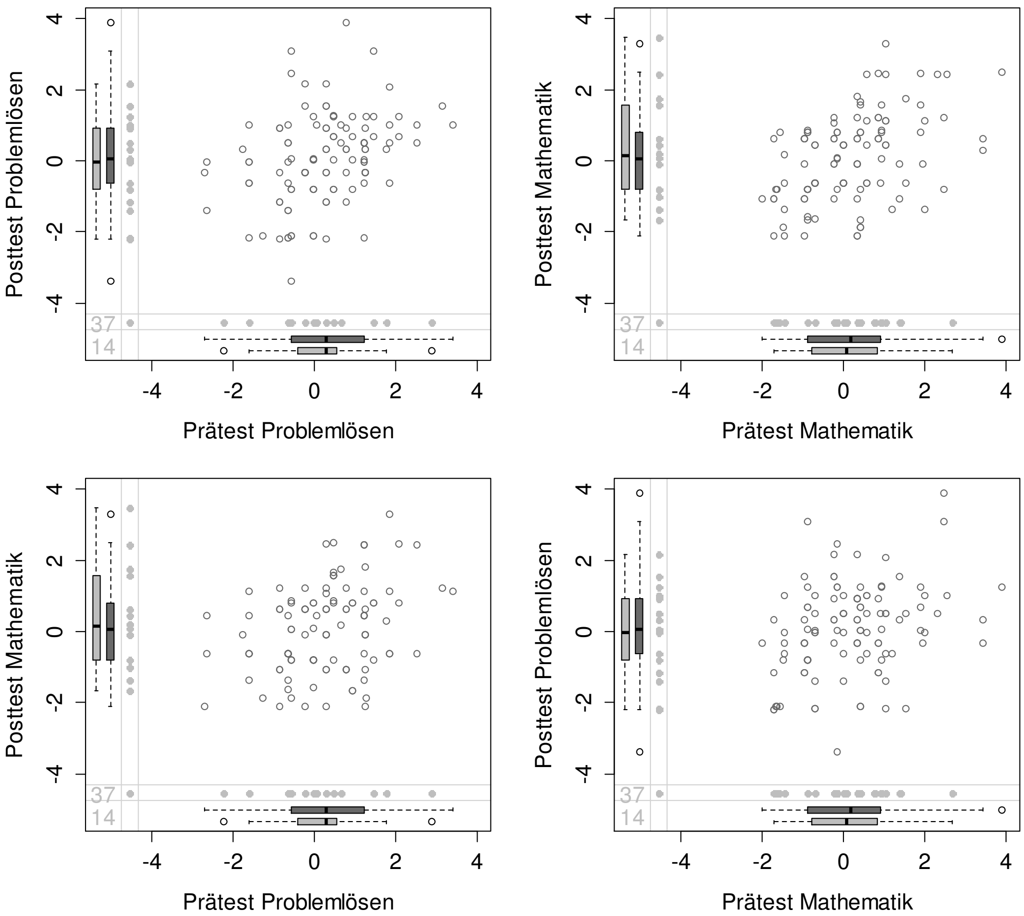


Abbildung DD-1. Margin plots für Prä- und Posttest

Tabelle DD-1. Prädiktormatrix, Teil 1

Variable	*Nr.*	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
Geschlecht	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Bedingung	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Trainingsgruppe	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Alter	6	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0
Sprache	7	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Note.Deutsch	8	0	0	1	0	0	0	0	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Note.Englisch	9	0	0	1	0	0	0	1	0	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Note.Mathematik	10	0	0	1	0	0	0	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Note.Biologie	11	0	0	1	0	0	0	1	1	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Note.Chemie	12	0	0	1	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Note.Geschichte	14	0	0	1	0	0	0	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Note.Sport	15	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
KFT.N2	16	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
KFT.V3	17	0	0	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FAM.PKT.T2	18	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FAM.PKT.T1	19	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
T2.Sorgfalt	20	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	1	1	1	0	1	0	0	0	0
T2.Konzentration	21	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
T2.Muehe	22	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
SJT.sysPL	23	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
SJT.unsysPL	24	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SJT.hilfe	25	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Testheft	26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
missingsPL.T2	27	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0
missingsM.T2	28	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
missingsPL.T1	29	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0
missingsM.T1	30	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
Selbsteinschätzung1	31	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0
Selbsteinschätzung2	32	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	1	0	1	1	1	1	0	1	0	0
Selbsteinschätzung3	33	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	1	1	0	1	1	1	0	0	0	0
Selbsteinschätzung4	34	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	0	0	1	0
Selbsteinschätzung5	35	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	1	0	0	0	0
Selbsteinschätzung6	36	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0
PKT.T1	37	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	0	0
PKT.T2	38	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	1	0	0	0
FB.Klippert	39	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
DISK.Mathematik	40	0	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DISK.Deutsch	41	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
DISK.Englisch	42	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DISK.Biologie	43	0	0	1	0	0	0	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
DISK.Chemie	44	0	0	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
Angst.Mathematik	45	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
SW.NW	46	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
SW.Mathematik	47	0	0	1	0	0	0	0	0	1	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1
NFC	48	0	0	1	0	0	0	0	0	1	1	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	1
FAM.T2.AN	49	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FAM.T2.IN	50	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FAM.T2.CH	51	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
FAM.T2.PS	52	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SELLMO.LZ	53	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
SELLMO.VL	54	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
SELLMO.AZ	55	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
SELLMO.AV	56	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0
Motivationsindikatior	59	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Problemlösen.T1	65	0	0	1	0	0	0	0	0	1	1	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
Problemlösen.T2	66	0	0	1	0	0	0	0	0	1	1	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
Mathematik.T1	74	0	0	1	0	0	0	0	0	1	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
Mathematik.T2	75	0	0	1	0	0	0	0	0	1	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
ATI.PLxBedingung	78																																							
ATI.MxBedingung	79																																							
PLxM	80																																							
ATI.PLxMxBedingung	81																																							

Fortsetzung

Tabelle DD-2. Prädiktormatrix, Teil2

Variable	*Nr.*	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81
Geschlecht	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Bedingung	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Trainingsgruppe	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Alter	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Sprache	7	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
Note.Deutsch	8	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
Note.Englisch	9	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
Note.Mathematik	10	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
Note.Biologie	11	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
Note.Chemie	12	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
Note.Geschichte	14	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
Note.Sport	15	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
KFT.N2	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
KFT.V3	17	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
FAM.PKT.T2	18	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
FAM.PKT.T1	19	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
T2.Sorgfalt	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
T2.Konzentration	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
T2.Muehe	22	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
SJT.sysPL	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
SJT.unsysPL	24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
SJT.hilfe	25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
Testheft	26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
missingsPL.T2	27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
missingsM.T2	28	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
missingsPL.T1	29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
missingsM.T1	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
Selbsteinschätzung1	31	0	0	0	0	0	0	0	1	0	0	1	0	1	0	0	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
Selbsteinschätzung2	32	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
Selbsteinschätzung3	33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
Selbsteinschätzung4	34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
Selbsteinschätzung5	35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
Selbsteinschätzung6	36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
PKT.T1	37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
PKT.T2	38	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
FB.Klippert	39	1	0	1	1	0	1	1	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
DISK.Mathematik	40	0	0	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
DISK.Deutsch	41	0	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
DISK.Englisch	42	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
DISK.Biologie	43	1	0	0	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
DISK.Chemie	44	1	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
Angst.Mathematik	45	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
SW.NW	46	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
SW.Mathematik	47	0	0	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
NFC	48	0	0	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
FAM.T2.AN	49	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
FAM.T2.IN	50	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
FAM.T2.CH	51	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
FAM.T2.PS	52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
SELLMO.LZ	53	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
SELLMO.VL	54	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
SELLMO.AZ	55	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
SELLMO.AV	56	1	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
Motivationsindikatior	59	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Problemlösen.T1	65	1	0	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
Problemlösen.T2	66	1	0	1	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
Mathematik.T1	74	1	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
Mathematik.T2	75	1	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0	1	1	1	1
ATI.PLxBedingung	78																																									
ATI.MxBedingung	79																																									
PLxM	80																																									
ATI.PLxMxBedingung	81																																									

Anmerkungen. Ausgeblendete Variablen wurden weder zur Imputation verwendet noch selbst imputiert. AxB kennzeichnet passiv imputierten Interaktionsterm (Produktvariable) von A und B.

Tabelle DD-3. Reihenfolge der Imputation der Variablen

```
visitseq.vars <- c( "Person-
Scores.Dim01.pl","PersonMax.Dim01.pl","N.items.pl","missingsPLT1",
"er-
ror.Dim01.pl","PL.T1","ati.pl.treat","PersonScores.Dim01.ma","Pers
on-
Max.Dim01.ma","N.items.ma","missingsMT1","error.Dim01.ma","M.T1","
ati.m.treat","pot",
"ati.pot.treat","anw.ps.treatment","T2.Sorgfalt","T2.Konzentration
",
"T2.Muehe","T2.fam.AN","T2.fam.IN","T2.fam.CH","T2.fam.PS","Person
Scores.Dim02.pl","PersonMax.Dim02.pl","missingsPLT2","error.Dim02.
pl","PL.T2","PersonScores.Dim02.ma","PersonMax.Dim02.ma","missings
MT2","error.Dim02.ma","M.T2","id","geschlecht","klasse","inv.Motiv
ationsindika-
tor","gruppe","Trainingsgruppe","alter","sprache.r2","note.d","not
e.e","note.m","note.bio","note.ch","note.ph","note.ge","note.sp","
ZSka-
la.DISK.M","ZSkala.DISK.D","ZSkala.DISK.E","ZSkala.DISK.Bio","ZSka
la.DISK.Ch","ZSkala.Matheangst","ZSkala.Selbstwirksamkeit.NW","ZSk
a-
la.Selbstwirksamkeit.Mathe","sellmo.skala.lz","sellmo.skala.vl","s
ellmo.skala.az","sellmo.skala.av",
"KFT.N2","KFT.V3","tp.fam.mean","tp.score.final","anw.pkt.treatmen
t","rp.fam.mean","rp.score.final","SJT.sysPL","SJT.unsysPL","SJT.h
il-
fe","Testheft","N.Motivation","N.Anstrengung","N.Konzentration","N
.MMA","N.SMA","N.Lernerfolg","ZSkala.Klippert","Zskala.nfc"
)
```

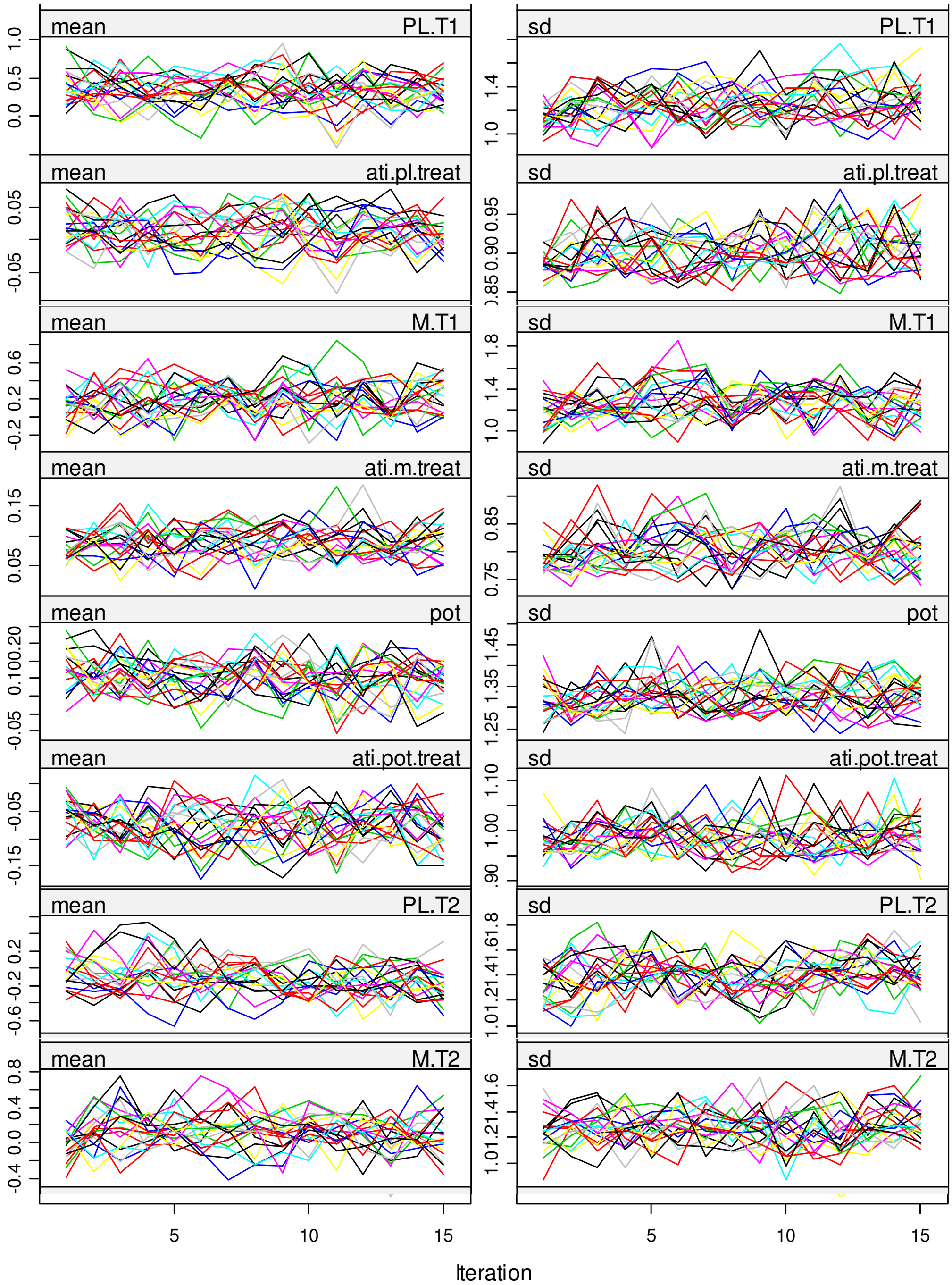
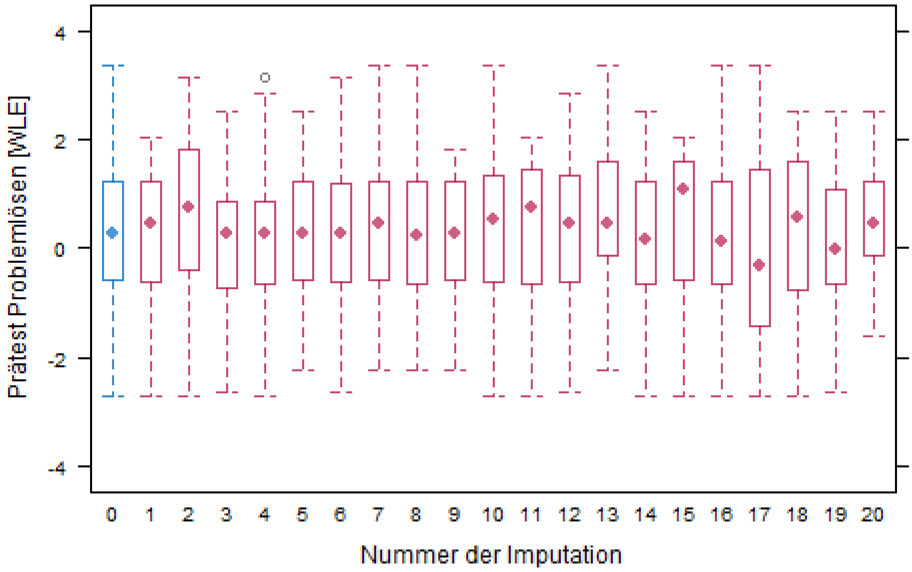


Abbildung DD-2. Trace line plots für Problemlösen und Mathematik im Prä- und Posttest und die Interaktionsterme

Anmerkungen. PL.T1 = Prätest Problemlösen. ati.pl.treat = Bedingung x Prätest Problemlösen. M.T1 = Prätest Mathematik. ati.m.treat = Bedingung x Prätest Mathematik. pot = Prätest Problemlösen x Prätest Mathematik. ati.pot.treatment = Bedingung x Prätest Problemlösen x Prätest Mathematik. PL.T2 = Posttest Problemlösen. M.T2 = Posttest Mathematik.



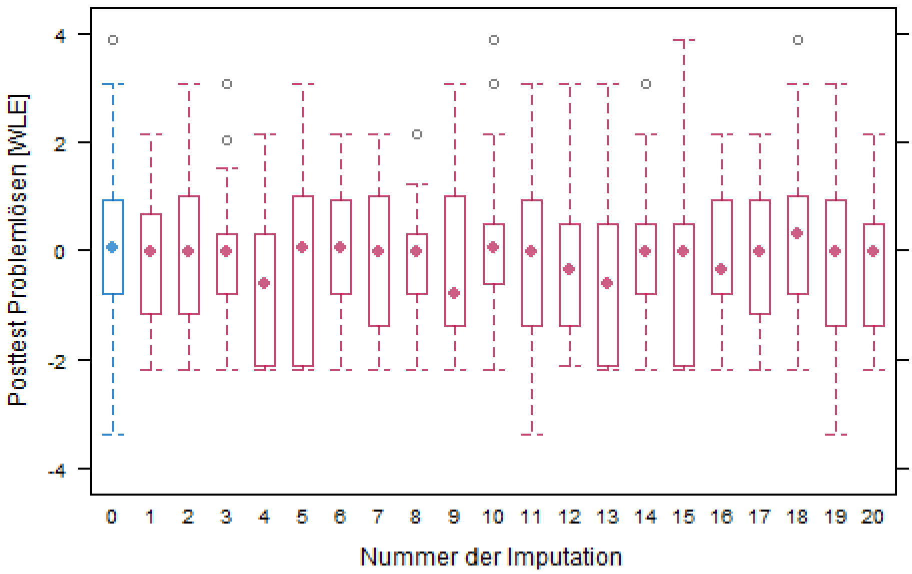


Abbildung DD-3. Boxplots für den Prä- und Posttest Problemlösen für den Originaldatensatz (Imputationsnummer 0) und die 20 imputierten Datensätze

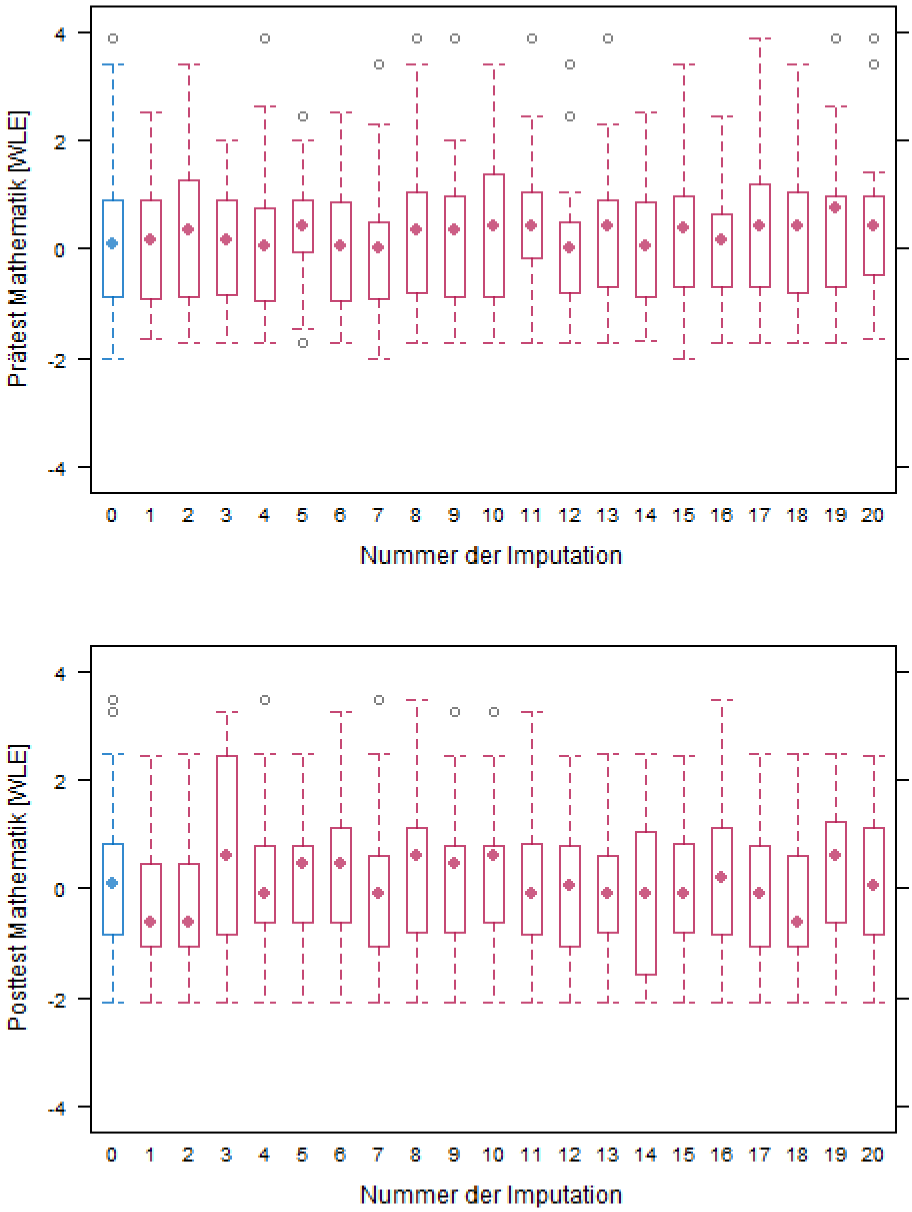


Abbildung DD-4. Boxplots für den Prä- und Posttest Mathematik für den Originaldatensatz (Imputationsnummer 0) und die 20 imputierten Datensätze