

Computational Playground

Eine Rasch-Analyse zum Computational Thinking
bei Sachunterrichtsstudierenden im Lehr-Lern-Labor

Martin Brämer, Daniel Rehfeldt & Hilde Köster

Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin
braemer@zedat.fu-berlin.de, hilde.koester@fu-berlin.de

Kurzfassung

„Informatische Bildung“ soll, wie von der KMK oder auch von sachunterrichtsdidaktischer Seite gefordert, Teil des Sachunterrichts werden. Dieser Vorgang stellt einen Transfer einer Innovation in ein bestehendes Bildungssystem und somit eine besondere Herausforderung dar: Für diesen Transfer muss eine entsprechende Qualifizierung von angehenden Lehrkräften im Lehramtsstudium sichergestellt werden. Da informatische Fachkompetenzen bei Grundschullehramtsstudierenden jedoch bisher wenig erforscht wurden, wurde im QL-B-Projekt K2teach an der Freien Universität Berlin eine Studie durchgeführt, die einerseits auf die Kreuzvalidierung eines Tests zum sogenannten Computational Thinking bei Studierenden fokussiert und andererseits darauf, inwiefern die Teilnahme an einem entsprechend ausgerichteten Seminar im Lehr-Lern-Labor-Format Einfluss auf dessen Ausprägung nimmt. Die Ergebnisse zeigen sowohl eine Eignung des Tests auf Basis der Rasch-Modellierung als auch einen signifikanten Anstieg der Fähigkeiten im Bereich des Computational Thinking durch den Besuch des Lehr-Lern-Labors.

1. Problemaufriss

Aktuelle gesellschaftliche Entwicklungen in Zusammenhang mit der Digitalisierung führen uns vor Augen, dass praktisch alle Bereiche des Lebens und der Gesellschaft von digitalen Wandlungsprozessen betroffen sind und es kaum noch möglich ist, sich diesem Einfluss zu entziehen (vgl. Schaar, 2017, S. 105). In der Literatur herrscht daher Einigkeit darüber, dass in einer hoch technologisierten Welt die Fähigkeit, mit Informatiksystemen (wie bspw. dem Smartphone, Tablets oder anderen Computern) umzugehen, eine Kompetenz darstellt, die über die reine Nutzung hinaus eine kritische Haltung sowie Wissen über Funktionsweisen und Auswirkungen dieser umfassen sollte (vgl. Gallenbacher 2015; Román-González, Pérez-González, Jiménez-Fernández, 2017, S. 678; Resnick, 2017; Papert, 1982). Die KMK fordert daher schon seit dem Schuljahr 2018/19, dass auch Grundschulkindern „Kompetenzen in einer digitalen Welt“ erwerben sollen (2017, S. 19). Dies bedeutet jedoch, dass neue Inhalte wie die Informatik in den Sachunterricht der Grundschule und somit auch in die Ausbildung von angehenden Grundschullehrkräften integriert werden müssen (bspw. KMK, 2017, S. 25).

Da es sich bei der Informatik jedoch um ein neues Themengebiet für die Grundschule handelt, kann hierbei von einem „Transfer einer Innovation“ (Gräsel, 2010, S. 8) in ein bestehendes Bildungssystem gesprochen werden. Im Bereich der Transferforschung konnte anhand verschiedener Studien gezeigt werden, dass insbesondere motivationale Komponenten sowie fach- und unterrichtsbezogene Kompetenzen der

Lehrkräfte in diesem neuen Bereich entscheidend dafür sind, ob dieser nachhaltig in den Unterricht transferiert wird (Trempler, Schellenbach-Zel & Gräsel, 2013, S. 344, Jäger, 2004). Sollen also zukünftig informatische Inhalte in den Sachunterricht integriert werden, sind eine entsprechend hohe Motivation sowie domänenspezifische Kompetenzen möglichst bereits im Studium anzulegen.

An der Freien Universität (FU) Berlin wurde daher im Rahmen des durch das BMBF geförderten Projektes K2teach (Qualitätsoffensive Lehrerbildung – siehe Anhang) für den Bachelorstudiengang Grundschulpädagogik im Fach Sachunterricht eine neue Lehrveranstaltung mit informatischen Inhalten im Format eines sog. *Lehr-Lern-Labors* (LLL; Köster, Mehrstens, Brämer & Steger, 2019; Brämer, Rehfeldt, Bauer, & Köster, 2020a) entwickelt. Grundlegend für dieses LLL mit dem Titel *Computational Playground* ist das Konzept des *Computational Thinking* (CT; Wing, 2006), welches vereinfacht als (kognitive) Fähigkeit zum problemlösenden (inkl. prozeduralem, algorithmischem) Denken sowie Konzeptwissen im Bereich der Informatik beschrieben werden kann.

Nach einer Entwicklungs- und Optimierungszeit des neuartigen LLL sollte untersucht werden, ob die Teilnahme am LLL die Fähigkeit zum CT bei den Studierenden fördert. Um diese Frage beantworten zu können, sind Kenntnisse darüber erforderlich, inwiefern die Studierenden über entsprechende CT-Fähigkeiten verfügen. Dazu liegen in Deutschland bisher jedoch noch keine Studienergebnisse vor (vgl. Hsu, Chang & Hung, 2018). Wir entschlossen uns daher, auf einen

bereits verfügbaren und validierten CT-Test zum ‚Informatischen Denken‘ (vgl. Wing, 2006) zurückzugreifen. Dieser wurde jedoch bisher lediglich für Schüler:innen in der Sekundarstufe validiert (Román González et al., 2017; Guggemos, Seufert & Román-González, 2019). Unsere Untersuchung fokussiert daher zunächst auf eine Kreuzvalidierung dieses CT-Tests in Bezug auf Grundschullehrer:innen-Studierende. Daran anschließend wurde mithilfe einer Interventionsstudie im Pre-Post-Design festgestellt, ob die Teilnahme am Lehr-Lern-Labor-Format das Computational Thinking (CT) bei den Studierenden fördert.

2. Theoretischer Bezugsrahmen

Die theoretische Basis der Untersuchungen bilden die Theorie zum Professionswissen nach Shulman (1987) sowie nach Baumert und Kunter (2006) und das Konstrukt des Computational Thinking (Wing, 2006, 2017; Papert, 1982).

2.1 Fachwissen als Facette professioneller Handlungskompetenz von Lehrkräften

Inzwischen hat sich die Unterteilung des sog. *Professionswissens* (als ein Element professioneller Kompetenz von Lehrkräften) nach Shulman (1987) in die drei zentralen Facetten *Fachwissen*, *Fachdidaktisches Wissen* und *Pädagogisches Wissen* durchgesetzt (neben dem *Organisations-* und *Beratungswissen*; s. Baumert & Kunter, 2006). Im deutschsprachigen Raum wird daher insbesondere in der fachdidaktischen und bildungswissenschaftlichen Forschung oftmals auf Shulman's Modell zurückgegriffen.

Nach Shulman (1987) bezeichnet das fachspezifische Inhaltswissen bzw. das Fachwissen das Verständnis von Lehrenden hinsichtlich der Struktur eines fachwissenschaftlichen Gegenstandes, seiner „salient concepts, relations among concepts“ (ebd.) sowie deren Konventionen, welche allgemein in der jeweiligen Fachdisziplin gelten. Goodson, Hopmann und Riquarts (1999) bezeichnen das Fachwissen daher auch gemeinsam mit dem fachdidaktischen Wissen als einen „Handlungsrahmen der Lehrertätigkeit“ (ebd.).

Im Rahmen der professionellen Handlungskompetenz von Lehrerinnen und Lehrern weisen Baumert und Kunter (2006) dem Fachwissen eine grundlegende Bedeutung zu als „notwendige aber nicht hinreichende Bedingung für qualitativollen Unterricht und Lernfortschritte der Schülerinnen und Schüler [...]. Fachwissen ist die Grundlage, auf der fachdidaktische Beweglichkeit entstehen kann“ (ebd., S. 496). Dem Erwerb von fachspezifischem Inhalts-/Konzeptwissen im Bereich der Informatik wird für angehende Grundschullehrkräfte daher zukünftig eine höhere Bedeutung zukommen.

2.2 Computational Thinking

Sollen Lehrkräfte oder Studierende Kompetenzen bzgl. informatischer Inhalte erwerben, werden neben einem informatischen „Konzeptwissen“ (vgl. Döbeli Honegger, 2017, S. 60) auch grundlegende

Handlungskompetenzen relevant. Diese Aspekte lassen sich im Bereich der Informatik unter dem von Papert (1982) eingeführten Begriff *Computational Thinking* fassen, welcher international auch als Schlüsselkonzept für die sogenannten ‚21st-century skills‘ gilt (Wing, 2006, Voogt, Fisser, Good, Mishra & Yadav, 2015).

In der deutschsprachigen Informatikdidaktik wurde mit den *GI-Standards* der Gesellschaft für Informatik e.V. (GI, 2019) bereits ein sehr ausdifferenziertes Modell für informatische Bildungsstandards etabliert, welches aber lediglich einzelne Aspekte des CT enthält (Steiner & Himpsl-Gutermann, 2020, S. 7). Diese Standards bieten jedoch aufgrund ihrer stark am systematischen Aufbau als Fachdisziplin orientierten Struktur zwar eine sehr gute Orientierung, jedoch eher weniger interdisziplinäre Anknüpfungspunkte an bspw. die Medienbildung, Digital Literacy oder andere Fachdisziplinen. Auch dem Sachunterricht und seinem vielperspektivischen Ansatz ist das Konzept des CT daher u. E. näher als die GI-Standards (vgl. GDSU, 2021; Brämer, Straube, Köster, & Romeike, 2020b; Straube, Brämer, Köster & Romeike, 2018).

Noch fehlt es zwar bisher am Konsens über eine umfassende Definition des CT-Konzepts (Gouws, Bradshaw & Wentworth, 2013; Kalelioglu, Gülbahar & Kukul, 2016) und es besteht noch Uneinigkeit darüber, wie das CT in Curricula integriert werden sollte (Lye & Koh, 2014), es existiert aber eine Definition, auf die sich ein Großteil von Autor:innen einigen konnten (vgl. bspw. Weigend, 2009; Engbring, 2017; Serafini, 2011; Guggemos et al., 2019). Diese Definition nach Wing beschreibt CT als: „*Computational thinking is the thought processes involved in formulating a problem and expressing its solution(s) in such a way that a computer — human or machine — can effectively carry out.*“ (Wing, 2017, S. 8).

CT beschreibt dementsprechend die (kognitive) Fähigkeit, Denkprozesse zum Problemlösen so zu formulieren, dass deren Lösung in Form von „computational steps“ (Aho, 2012, S. 832) bzw. Algorithmen dargestellt werden kann, welche wiederum von Computern, Menschen oder einer Kombination aus beidem ausgeführt werden können. Die Maschinen bzw. Computer sind in dieser Definition entsprechend nicht als konstitutiv zu betrachten.

3. Forschungsstand zum CT bei Studierenden im Grundschullehramt / Sachunterricht

In Deutschland fehlt es derzeit noch an quantitativen Erhebungen der Fähigkeit zum CT bei Grundschullehrer:innen-Studierenden. Im Bereich des informatischen Fachwissens wurden jedoch bereits erste Studien durchgeführt. So fand Gläser (2020) bei 33 Sachunterrichtsstudierenden mittels strukturierender qualitativer Inhaltsanalysen heraus, dass das geäußerte Fachwissen zum Thema Digitalisierung „überwiegend zwischen rudimentären Anfangswissen und Reproduzieren von elementarem Faktenwissen einzustufen“

(ebd., S. 318) sei. „Erklärungen, die grundlegende Kenntnisse aufzeigen, sind dagegen selten.“ (ebd., S. 318) Insbesondere der Begriff ‚digital‘ wird hierbei häufig nicht mit digitalen Daten in Verbindung gebracht, sondern bspw. synonym mit Begriffen wie Strom/Energie, neue Medien oder Internet (ebd., S. 318).

Dengel und Heuer (2017) befragten 167 Lehramtsstudierende (von denen 126 für das Grundschullehramt studierten) mithilfe eines qualitativen Fragebogens zu ihren Vorstellungsbildern hinsichtlich des Internets. Auffällig war hierbei ein signifikanter Zusammenhang ($\chi^2 = 5,656$, $p = .017$) „zwischen den Eigenschaften ‚studiert Grundschullehramt‘ und ‚hat eine falsche Vorstellung‘. Grundschullehramtsstudierende zeigten dementsprechend häufiger fachliche Fehlannahmen als Sekundarstufenlehramtsstudierende (ebd., S. 7). Von 87 Grundschullehramtsstudierenden hatten 78 als ‚falsch‘ kategorisierte Vorstellungen und nur 18 zutreffende (ebd., S. 93).

Eine eigene in 2019 durchgeführte quantitative Fragebogenstudie zu informatischen Fähigkeiten (Brämer, Rehfeldt & Köster, 2021) mit $n = 61$ Grundschullehramtsstudierenden im Sachunterricht zeigte eine geringe selbsteingeschätzte Programmierfähigkeit. Lediglich 21% der Studierenden gaben an, entsprechende Fähigkeiten zu besitzen.¹

Auch international existieren bisher keine quantitativen Erhebungen von CT-Fähigkeiten bei Grundschullehramtsstudierenden. Zwei quantitative Studien, bei denen jeweils CT-Tests zur Anwendung kamen, wurden lediglich mit Sekundarstufenschüler:innen durchgeführt (Román-González, 2017; ICILS, 2018).

Eine Untersuchung aus der Schweiz zu den Programmierkenntnissen von 85 Grundschulpädagogikstudierenden stützt die Ergebnisse der in Deutschland durchgeführten Studien, da diese Studierenden ähnliche Angaben machten: lediglich 10% gaben an, Programmierkenntnisse zu besitzen; auch hier zeigten sich wiederholt fachliche Fehlannahmen, und der Begriff ‚Informatik‘ wurde häufig mit dem Umgang mit Medien oder einer Medienkompetenz gleichgesetzt (Döbeli Honegger & Hielscher, 2017, S. 103).

Eine der wenigen Studien, welche explizit CT in Lehramtsstudiengängen erforscht, stammt aus den USA. Yadav, Mayfield, Zhou, Hambrusch und Korb (2014) forderten 294 Lehramtsstudierende der Primar- und Sekundarstufe u. a. dazu auf, CT zu definieren und ihre Vorstellungen zum Unterrichten in Bezug auf den Erwerb von CT-Fähigkeiten zu beschreiben. Die Ergebnisse zeigten eher oberflächliche Vorstellungen von CT und damit einhergehend auch darüber, wie ein entsprechender Unterricht gestaltet

werden könnte. Auch hier wurde CT oft mit dem Umgang mit Medien gleichgesetzt. Die Lehrpersonen, die CT als den ‚Umgang mit Medien‘ verstanden, dachten außerdem, dass der Einsatz von Medien im Unterricht CT fördert. Eine an diese Erhebung anschließende kurze Intervention (Dauer unter zwei Stunden) zeigte allerdings bereits einen Wissenszuwachs in Bezug auf die Definition von CT sowie einen Zusammenhang mit der Bereitschaft, CT zu unterrichten (ebd.).

4. Forschungsfragen und Hypothesen

4.1 Kreuzvalidierung CT-Test

Zusammenfassend lassen sich auf Basis des Forschungsstandes folgende Schlüsse ziehen:

CT-Fähigkeiten wurden in Deutschland bisher noch nicht quantitativ bei Studierenden für das Grundschullehramt erhoben. Für die Sekundarstufe existiert jedoch ein bereits validierter Test.

Bei den zitierten Studien wurde insgesamt ein eher niedriges Niveau an informatischem Fach-/Konzeptwissen detektiert. Dies lässt vermuten, dass auch die Fähigkeit zum CT bei Grundschullehramtsstudierenden eher gering ausgeprägt sein dürfte. Der bereits validierte CT-Test aus der Sekundarstufe könnte daher auch für diese Stichprobe geeignet sein. Ob diese Annahme zutrifft, sollte jedoch auf Basis der neuen Stichprobe validiert werden.

Es stellt sich daher als **erste Forschungsfrage** (FF 1):

FF 1 (Kreuzvalidierung des CT-Tests): Ist der CT-Test aus der Sekundarstufe auch zur Messung von CT-Fähigkeiten angehender Sachunterrichtslehrkräften geeignet?

4.2 CT-Förderung (LLL-Intervention)

Im zweiten Schritt sollte untersucht werden, ob die Teilnahme am LLL die Fähigkeit zum CT bei den Studierenden fördert. Zur Förderung der Fähigkeit zum CT bei Studierenden im Grundschullehramt bzw. des Sachunterrichts existieren nach unseren Recherchen bislang keine Untersuchungen. Die Ergebnisse von Yadav et al. (2014) geben jedoch Hinweise darauf, dass eine Förderung sogar bereits im Rahmen kurzer Interventionen möglich ist. Eine Förderung im Rahmen eines einsemestrigen Seminars im LLL-Format sollte daher umso besser gelingen.

Diese optimistische Annahme resultiert auch aus den Ergebnissen, die im Rahmen der oben bereits zitierten (eigenen) (Interventions-) Studie (Brämer, Rehfeldt & Köster, 2021) mit $n = 61$ Grundschullehramtsstudierenden bzgl. der Förderung des Interesses gewonnen wurden: Durch die Teilnahme an einem Lehr-Lern-Labor (LLL) mit informatischen Inhalten

¹ Im Rahmen dieser Untersuchung wurde außerdem das für einen späteren Transfer relevante Interesse der Studierenden an Informatik erhoben. Es zeigte sich ein eher gering ausgeprägtes Interesse ($d = 1.53***$) in Relation zur Eichstichprobe, wobei diese Ergebnisse mit Vorsicht zu interpretieren sind; genaueres siehe Brämer, Rehfeldt, Bauer & Köster, 2020a.

konnten die zuvor geringen *informatikbezogenen Interessen* signifikant gesteigert werden ($d = 0.65^*$; ebd.).

Ob ein solches LLL neben dem Interesse jedoch auch die CT-Fähigkeiten positiv beeinflussen bzw. fördern kann, wurde bislang noch nicht untersucht. Im naturwissenschaftsbezogenen Sachunterricht und in anderen fachlichen/-didaktischen Domänen konnten Studien zur Wirksamkeit des Formats Lehr-Lern-Labor aber bereits einen Zuwachs des jeweils adressierten Fachwissens durch LLL-Interventionen nachweisen (bspw. Völker & Trefzger, 2010; 2011; Anthofer & Tepner, 2016; Kobl & Tepner, 2019; Köster et al., 2019).

Die **zweite Forschungsfrage** folgt daher der Annahme, dass ein LLL (mit informatikbezogenen Inhalten) auch den Erwerb der Fähigkeit zum CT bei Lehramtsstudierenden im Grundschullehramt fördern kann:

FF 2 (Interventionsstudie im LLL): Inwiefern lässt sich die Fähigkeit zum CT in einem Lehr-Lern-Labor fördern?

5. LLL-Intervention ‚Computational Playground‘

An der FU Berlin sind LLL im Studienfach Sachunterricht universitäre Seminare, in denen Lehramtsstudierende fachbezogene und didaktische Kompetenzen erwerben und darauf gründend theoriegeleitet Lernumgebungen für Grundschulkindern entwickeln. Diese werden in komplexitätsreduzierter, universitärer Umgebung mit Kindergruppen (i.d.R. ganze Grundschulklassen der Jahrgänge 1-4) praxiserprobt, dann gemeinsam im Seminar reflektiert, ggf. optimiert und erneut in einer Praxisphase erprobt (vgl. Brämer & Köster, 2021, S. 6; Rehfeldt, Seibert, Klempin, Lücke, Sambanis & Nordmeier, 2018). Aufgrund dieser Verzahnung der Theorie- und Praxisphasen kann davon ausgegangen werden, dass die Motivations- und Herausforderungslage im Sinne eines problemorientierten Lernens (Reusser, 2005) und konkreter eines generativen Problemlösens (Klauser, 1998, S. 278) bei den Studierenden günstiger ausfallen kann, als bei vergleichbaren ‚Theorieseminaren‘, wo der Anwendungskontext bzw. die Praxis fehlt oder eher abstrakt ausfällt. Damit beugen LLL dem Erwerb trägen Wissens (Renkl, 1996) vor, wie auch die Studien von Seibert, Rehfeldt, Klempin, Mehrrens, Nordmeier, Sambanis, Köster & Lücke (2019) nahelegen.

Das hier untersuchte LLL mit dem (bewusst mögliche Hemmschwellen bei der Wahl des Seminars berücksichtigenden) Titel ‚Computational Playground‘ besteht aus drei aufeinander folgenden Phasen: In der ersten Phase erwerben die Studierenden Fachwissen aus dem Bereich der Informatik bzw. des CT. Hierbei wurde zunächst auf Basis von Alltagsphänomenen (wie bspw. dem Prozess des Händewaschens) ein grundlegendes Verständnis über die zugrundeliegenden Konzeptionen (Schleifen, Bedingungen, und

Sequenzierungen bspw.) von CT in Anlehnung an Brennan und Resnick (2012) gelegt. Diese nun bekannten Konzepte finden im Anschluss Anwendung in verschiedenen grafischen Programmierungsumgebungen sowie bei der Programmierung von Robotern, Mikrokontrollern u.ä.. Im letzten Teil dieser Phase wird das Wissen anhand von Fachtexten aus dem Bereich des CT vertieft. Der Grad der Abstraktion nimmt in dieser Phase dementsprechend stetig zu (vom Alltagswissen zum Fachtext aus der Informatik).

In der zweiten Phase werden das erworbene Fachwissen und die grundlegenden Konzepte aus dem Bereich des CT auf fachdidaktische Inhalte bezogen. In diesem Zusammenhang beschäftigen sich die Studierenden mit Gemeinsamkeiten und Unterschieden zwischen CT und dem Forschenden Lernen (Köster & Galow, 2017) sowie mit möglichen curricularen Einbettungen informatischer Inhalte in den Perspektivrahmen (GDSU, 2013) und dem Diskurs um eine digitale Perspektive im Sachunterricht (Brämer et al., 2020b; Straube et al., 2018).

Die dritte Phase umfasst zunächst die Planung der zu gestaltenden Lernumgebungen sowie die parallel dazu stattfindenden Beobachtungsaufträge im Rahmen der Praxisphase mit Kindergruppen an der Universität. Die Lernumgebungen werden in Anlehnung an die gemeinsam erarbeiteten curricularen Zielstellungen (digitale Perspektive), dem Konzept des Forschenden Lernens (Köster & Galow, 2017) sowie einer fachdidaktischen Aufschlüsselung von Konzepten und Fähigkeiten im CT (Brennan & Resnick, 2012) angelegt. Die Beobachtungen finden wiederum auf Basis dieser fachdidaktischen Grundlegung statt. Hierfür werden zunächst möglichst ausführliche Textvignetten auf Basis eigener Beobachtungen erstellt und erst im Anschluss der Praxisphasen kriteriengeleitet auf Basis der besprochenen Konzepte ausgewertet. In diesen letzten beiden Phasen nimmt das Abstraktionsniveau dementsprechend kontinuierlich ab (vom Fachtext aus der Informatik zur Reflexion von Lernprozessen von Grundschulkindern auf Basis von Handlungen).

6. Design und Forschungsmethoden

Die LLL-Intervention Computational Playground fand im Sommersemester 2019 im 6. Semester des Studienfachs Sachunterricht statt und wurde mittels Pre-Post-Verfahren evaluiert. Dabei wurde der CT-Test von Román-González et al. (2017) mittels Online-Fragebogen und einem Zeitlimit von 20 Minuten eingesetzt.

Da dieses LLL insbesondere auf die Verzahnung der Theorie- und Praxisphasen sowie deren Einfluss auf CT fokussiert, wurde zusätzlich mit einer Parallelgruppe (PG) ein Seminar mit gleichem Inhalt, aber ohne Schüler:innenbesuche durchgeführt. Eine Baseline-Erhebung in einer Kontrollgruppe (KG) erfolgte im Rahmen eines Seminars mit nicht informatischen, sondern naturwissenschaftlichen Inhalten. Bei allen

Erhebungen wurde mit einem quasiexperimentellen Design (ohne Randomisierung) gearbeitet. Beide Teilstudien wurden im Jahr 2019 mittels Pen-and-Paper Fragebogen erhoben. Die LLL-Intervention (Teilstudie 2) fand im Sommersemester 2019 im 6. Semester des Studienfachs Sachunterricht statt und wurde mittels Pre-Post-Verfahren evaluiert. Da diese insbesondere auf die Verzahnung der Theorie- und Praxisphasen sowie deren Einfluss auf CT fokussiert, wurde mit einer Parallelgruppe (PG – Seminar mit ähnlichem Inhalt ohne Schüler*innenbesuche) sowie einer Baseline-Erhebung (KG – Seminar mit nicht informatischen, sondern naturwissenschaftlichen Inhalten) in einem quasiexperimentellen Design (ohne Randomisierung) gearbeitet.

7. Stichprobenbeschreibung

Innerhalb dieser Untersuchung werden zwei unterschiedliche Stichproben betrachtet. Bei beiden Gruppen handelt es sich um Studierende im sechsten Bachelorsemester an der Freien Universität Berlin im SoSe 2019.

Die Gesamtstichprobe mit $n = 111$ Studierenden wird lediglich für die Kreuzvalidierung des Testes sowie die Schätzung der Itemschwierigkeiten verwendet, wohingegen die Teilstichprobe mit $n = 71$ Studierenden zur Schätzung der Personenfähigkeiten im Rahmen der Intervention dient. Die Teilstichprobe ist entsprechend der Beschreibung in der Gesamtstichprobe enthalten. Beide Stichproben zeigen kaum Unterschiede in Bezug auf relevante Kovariaten (siehe Tabelle 1).

8. Computational Thinking-Test

Der CT-Test von Román-González et al. (2017) wurde lediglich für die Sekundarstufe validiert. Auf Basis des Forschungsstandes (Kap. 3) kann jedoch von einem ähnlichem CT-Fähigkeitsniveau bei unserer Stichprobe ausgegangen werden. Aufgrund der erwarteten Passung zu unserer Stichprobe sowie einer ausführlichen Validierung im Rahmen der Item-Response-Theorie (Guggemos et al., 2019), haben wir uns daher für die Verwendung dieses CT-Tests entschieden (Kap. 4). Diesem Test liegt zudem die bereits diskutierte (Kap. 2) CT-Definition von Wing (2017, S.8) in ihrer ursprünglichen Formulierung (Wing, 2006) zugrunde.

Der CT-Test wurde bereits bei Sekundarstufenschüler:innen der 5.-10. Klassenstufe in Spanien ($n = 1251$, $\alpha = .79$; ebd.) auf Basis der klassischen Testtheorie, sowie in der Schweiz in einer deutschen Übersetzung in einer 11. Klassenstufe auf Basis der Item-Response-Theorie ($n = 202$) validiert (Guggemos et al., 2019).

Die Konstruktvalidität des CT-Tests wurde bei Román-González et al. (2017) durch den Vergleich

mit den beiden standardisierten Instrumenten *Primary Mental Abilities (PMA) battery* und dem *RP30 problem-solving test* untersucht (ebd., S. 682). Es zeigte sich eine starke Korrelation mit der Problemlösefähigkeit $r = .67$, eine mittlere Korrelation mit der Argumentationsfähigkeit sowie dem räumlichen Vorstellungsvermögen mit je $r = .44$ (ebd.). Diese Ergebnisse zeigen, dass CT zwar eine hohe Ähnlichkeit zum Problemlösen aufzeigt, es sich jedoch nicht um dasselbe Konstrukt handelt.

Im Rahmen der Validierung auf Basis der Item-Response-Theorie wurde bereits ein Test auf Subgruppenunabhängigkeit anhand von externen Merkmalen mit dem Ergebnis durchgeführt, dass der Test keine Subgruppeninvarianz in Bezug auf das Alter, das Geschlecht, das Fähigkeitsniveau im Bereich der Computational Literacy (Medienkompetenz) sowie der jeweiligen Leistungsniveaus im Test besitzt (Guggemos et al., 2019). Der Test gilt somit zunächst als stichprobenunabhängig und bevorteilt bspw. keine älteren bzw. jüngeren Personen (vgl. Koller, Alexandrowicz & Hatzinger, 2012). Außerdem führten Guggemos et al. (2019) eine konfirmatorische Faktorenanalyse auf Basis eines Faktors (CT) durch und konnten somit bereits die Unidimensionalität des Tests belegen. Die Autor:innen kommen auf Basis dieser und weitere Überprüfungen des Tests zu dem Ergebnis, das der CT-Test sowohl raschkonform als auch gut geeignet für die Erhebung im Rahmen von Längsschnittstudien in deutschsprachigen Ländern ist (Guggemos et al., 2019, S. 189).

8.1 Auswertungsmethode - Kreuzvalidierung

Die Kreuzvalidierung (FF 1) findet im Rahmen der probabilistischen Testtheorie sowie einer Raschmodellierung statt und baut somit auf der Vorarbeit von Guggemos et al. (2019) auf. Guggemos und Kolleg:innen orientierten sich hierbei am Rahmenmodell von Bühner (2011, S. 547) zur Validierung der Raschkonformität. Dieser Rahmen bietet auch eine gute Orientierung für die von uns durchgeführte Kreuzvalidierung, auch wenn auf einige Tests aufgrund der ausführlichen Vorarbeit von Guggemos et al. (2019) verzichtet werden konnte. So wurde bspw. keine konfirmatorische Faktorenanalyse durchgeführt (ausführlicher zur Validierung siehe ebd.).

Die Daten wurden zunächst auf Stichprobenunabhängigkeit mittels Andersen's Likelihood Ratio Test (LRT) (Andersen, 1973)² untersucht (Koller et al. 2012, S. 19, S. 70ff). Dieser testet, ob die identifizierten Leistungsgruppen (gesplittet anhand des Medians) jeweils vom Test bevor- oder benachteiligt werden.

² Der Test wurde in der Statistiksoftware R mittels 'eRm 0.16-2' Packet (Mair & Hatzinger, 2007) gerechnet.

Darauf folgte ein Test auf Itemhomogenität bzw. lokale stochastische Unabhängigkeit mittels Yens Q3 (Yen, 1984; Christensen, Makransky & Horton, 2017)³ innerhalb der Stichprobe.

Nachdem die Raschkonformität der Daten bzw. des Testes erneut überprüft wurden, wurde mittels einer 1pl Marginal Maximum Likelihood-Schätzung (Bock & Lieberman, 1970; Bock & Aitken, 1981) die Raschanalyse mittels R package ‘TAM 3.1-45’ (Robitzsch, Kiefer, & Wu, 2019) durchgeführt. Im Rahmen dieser werden verschiedene Kennwerte untersucht, um den Modellfit sowie die Reliabilität angeben zu können. Hierfür werden die Modellfitindizes Infit und RMSD (Root-Mean-Square Deviation) herangezogen. Zusätzlich wird als Maß für die Reliabilität die EAP/PV-Reliabilität angegeben, welche sich ähnlich wie ein vergleichbares Cronbachs Alpha interpretieren lässt (Guggemos et al. 2019, S. 186) und die Messgenauigkeit der Personenkennwerte angibt.

Zum Umgang mit fehlenden Werten wurde, angelehnt an das Vorgehen bei Ludlow & O’Leary (1999), eine Strategie verwendet, welche auch in größeren Untersuchungen wie bspw. TIMMS oder ACER Verwendung findet (Adams, Wu & Macaskill, 1997). Diese Verfahren können hier Anwendung finden, da auf Basis von Bedingungen wie bspw. dem hier gesetzten Zeitlimit davon ausgegangen werden kann, dass fehlende Daten nicht zufällig (missing at random) auftreten, sondern in Form von ‚übersprungenen‘ (da zu schweren Aufgaben sowie Zeitdruck) oder ‚nicht geschafften‘ Items (aufgrund von Zeitmangel) des Tests Aussagen über die Personenfähigkeit, jedoch nicht über die Aufgabenschwierigkeit erlauben. Hierfür wurden folgende drei Schritte vollzogen:

- Zuerst wurde die Itemschwierigkeit des Datensatzes geschätzt, wobei alle *ausgelassenen Items als falsch und nicht erreichte Items als nicht bearbeitet* behandelt wurden.
- Im Anschluss wurde die *Itemkennwerte/Itemschwierigkeit normiert*, d. h. die hier berechneten Itemkennwerte wurden für die anschließende Berechnung übernommen.
- Erst im dritten Schritt wurde die Personenfähigkeit auf Basis der Itemnormierung berechnet. Hierbei wurden nun *ausgelassene und nicht erreichte Items als falsch* behandelt.

8.2 Auswertungsmethode – LLL-Intervention

Im Rahmen der LLL-Intervention (FF 2) in der Teilstichprobe wurden die vorher errechneten Itemkennwerte/Itemschwierigkeiten übernommen und innerhalb der Pre-Post-Messung verwendet. Somit konnte

auf eine genauere Schätzung der Itemschwierigkeit auf Basis der größeren Gesamtstichprobe zurückgegriffen werden. Der Umgang mit fehlenden Werten entsprach dem dritten Schritt im bereits beschriebenen Verfahren (s. o.).

Im Rahmen der Pre-Post-Analyse wurden zunächst je Zeitpunkt eine Personenfähigkeitsschätzung mittels 1pl Marginal Maximum Likelihood-Schätzung (Bock & Lieberman, 1970; Bock & Aitken, 1981) mit genormten Itemschwierigkeiten durchgeführt. Die daraus resultierenden Personenfähigkeiten (EAP-Werte) für die zwei Messzeitpunkte wurden im Anschluss je Gruppe auf Mittelwertsunterschiede mittels Wilcoxon-Rangsummen-Test (vgl. Bauer, 1972; Hollander & Wolfe, 1973) für gepaarte Stichproben untersucht. Außerdem wurde auch hier ein erneuter Reliabilitätswert auf Basis der EAP-Werte zur Messgenauigkeit der Personenwerte angegeben.

9. Ergebnisse

9.1 Kreuzvalidierung CT-Test (FF 1)

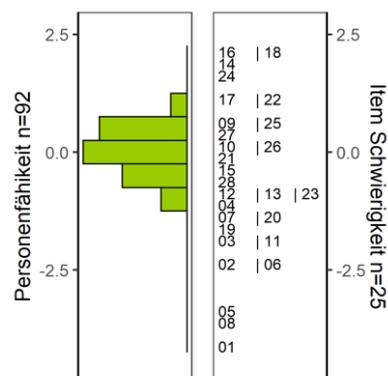


Abb. 1: Wrightmap der Gesamtstichprobe im CT-Test mit allen Aufgaben

In Bezug auf die FF 1 zeigen die Daten der Gesamtstichprobe, dass von den 28 Items im Durchschnitt 16.32 Items richtig gelöst wurden (SD = 3.51, median = 16.5, min = 8, max = 23).

Das Ergebnis des Andersen’s Likelihood Ratio Tests (LRT) (Andersen, 1973) war nichtsignifikant, weshalb hier eine Stichprobenunabhängigkeit in Bezug auf die Leistung angenommen werden kann.

Auch der Test zum Yen Q3-Kriterium war größtenteils unauffällig. Lediglich bei Item 5 und Item 6 lag der Wert mit 0.3 genau auf dem konventionellen Grenzwert, was vermutlich der geringen Stichprobengröße zugeschrieben werden kann (Christensen et al., 2017, S. 183ff).

³ Der Test wurde in der Statistiksoftware R mittels TAM Version 3.5-19 (Robitzsch, Kiefer, & Wu, 2019) durchgeführt.

| Gruppe | n | w | m | d | Altersdurchschnitt (SD) | Fachsemester (SD) |
|--|----|----|---|---|-------------------------|-------------------|
| Teilstichprobe (Intervention) | 71 | 63 | 6 | 0 | 24.92 (4.33) Jahre | 5.63(0.83) |
| Gesamtstichprobe (ohne Teilstichprobe) | 40 | 31 | 5 | 0 | 25.11 (4.45) Jahre | 5.14(0.82) |

Tab. 1: Stichprobenbeschreibung

| | wMNSQ | RMSD | θ | s.e. θ | M |
|------------|-------|------|----------|---------------|------|
| Aufgabe 1 | 0.86 | .03 | -4.00 | 0.72 | 0.98 |
| Aufgabe 2 | 0.98 | .08 | -2.38 | 0.36 | 0.90 |
| Aufgabe 3 | 0.79 | .07 | -1.94 | 0.31 | 0.86 |
| Aufgabe 4 | 0.83 | .07 | -1.19 | 0.25 | 0.75 |
| Aufgabe 5 | 0.94 | .06 | -3.27 | 0.52 | 0.96 |
| Aufgabe 6 | 0.85 | .03 | -2.38 | 0.36 | 0.90 |
| Aufgabe 7 | 0.85 | .06 | -1.32 | 0.26 | 0.77 |
| Aufgabe 8 | 1.02 | .04 | -3.55 | 0.59 | 0.97 |
| Aufgabe 9 | 1.05 | .03 | 0.63 | 0.23 | 0.36 |
| Aufgabe 10 | 1.04 | .04 | 0.12 | 0.22 | 0.48 |
| Aufgabe 11 | 0.93 | .06 | -1.99 | 0.32 | 0.87 |
| Aufgabe 12 | 1.05 | .07 | -0.79 | 0.23 | 0.68 |
| Aufgabe 13 | 0.90 | .06 | -0.79 | 0.23 | 0.68 |
| Aufgabe 14 | 1.07 | .03 | 1.93 | 0.31 | 0.14 |
| Aufgabe 15 | 0.96 | .04 | -0.42 | 0.22 | 0.60 |
| Aufgabe 16 | 1.18 | .11 | 2.02 | 0.32 | 0.13 |
| Aufgabe 17 | 1.15 | .10 | 1.10 | 0.25 | 0.27 |
| Aufgabe 18 | 1.12 | .08 | 2.12 | 0.33 | 0.12 |
| Aufgabe 19 | 0.90 | .07 | -1.69 | 0.29 | 0.83 |
| Aufgabe 20 | 0.90 | .07 | -1.31 | 0.26 | 0.78 |
| Aufgabe 21 | 1.00 | .04 | -0.10 | 0.22 | 0.53 |
| Aufgabe 22 | 0.95 | .05 | 1.15 | 0.25 | 0.26 |
| Aufgabe 23 | 0.93 | .05 | -0.85 | 0.24 | 0.69 |
| Aufgabe 24 | 1.04 | .03 | 1.64 | 0.28 | 0.18 |
| Aufgabe 25 | 0.97 | .02 | 0.61 | 0.23 | 0.37 |
| Aufgabe 26 | 0.98 | .04 | 0.02 | 0.23 | 0.51 |
| Aufgabe 27 | 1.01 | .09 | 0.39 | 0.25 | 0.42 |
| Aufgabe 28 | 1.16 | .15 | -0.63 | 0.27 | 0.65 |

Tab. 2: Ergebnisse der Raschanalyse für alle Aufgaben in der Gesamtstichprobe (N = 111)

Die Analyse der Modellfitindizes ergab bei keinem der 28 Items Auffälligkeiten. So lag der Infit-Wert (wMNSQ) zwischen den Vorgaben von $0.8 < \text{Infit} < 1.2$ (Guggemos et al., 2019; Wright & Linacre, 1994)

und die RMSD zwischen $0 < \text{RMSD} < 0.3$ (MacCallum et al., 1996) (siehe Tabelle 2).

Betrachtet man die Wrightmap aus Abbildung 1, so erkennt man die Aufgabenschwierigkeiten in

Relation zu den Personenfähigkeiten der Gesamtstichprobe. Der Test scheint insgesamt ein wenig zu leicht und die Aufgaben 1, 5 und 8 scheinen aufgrund ihres niedrigen Schwierigkeitsgrades (θ) kaum aussagekräftig in Relation zur vorliegenden Stichprobe.

Da Aufgabe 5 bereits bei der Analyse des Yens Q3-Kriterium auffällig war, gibt es in Bezug auf diese Aufgabe sogar zwei Gründe, diese zu entfernen. Die Aufgaben 1, 5 und 8 wurden dementsprechend aus den Daten entfernt. Alle nachfolgenden Analysen wurden lediglich mit 25 Items durchgeführt.

Die Raschkonformität kann somit angenommen werden, und der Test scheint nach dem Entfernen zu leichter Aufgaben eine gute Passung zur Stichprobe zu besitzen.

Mit einer Reliabilität von $EAP_{Rel} = 0.72$ weist der Test außerdem eine akzeptable mittlere Messgenauigkeit in Bezug auf die Schätzung der Personenfähigkeiten auf (vgl. Guggemos et al., 2019, S.187).

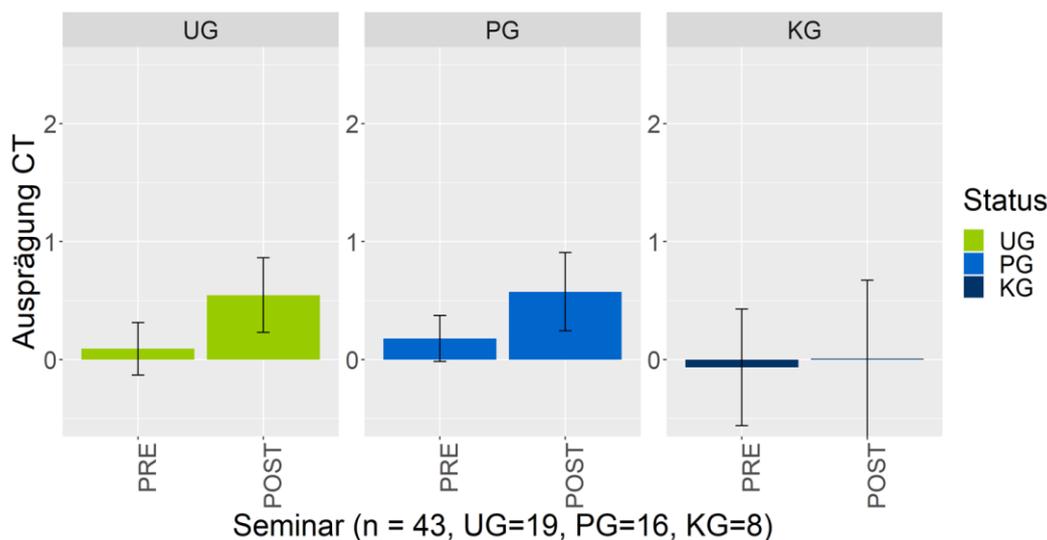


Abb. 2: Entwicklung von CT je Gruppen:

UG = Untersuchungsgruppe (LLL ‚Computational Playgrounds‘ mit Schüler:innenbesuchen)

PG = Parallelgruppe (Seminar mit informatischem Schwerpunkt, jedoch ohne Schüler:innenbesuchen)

KG = Kontrollgruppe (Seminar mit naturwissenschaftlichem Schwerpunkt)

(Fehlerbalken stellen längsschnittkorrigierten CI dar).

9.2 Intervention im LLL (FF 2)

In Abbildung 2 sind die Entwicklungen der CT-Fähigkeiten der untersuchten Studierendengruppen dargestellt.

Die Veränderungen der Personenfähigkeiten im CT zeigen vergleichbare signifikante sowie große Mittelwertunterschiede in der Untersuchungsgruppe (UG LLL) ($\Delta M = 0.46$; $p < 0.001$; $d = 1.17$) und in der Parallelgruppe (PG Theorieseminar) ($\Delta M = 0.40$; $p = 0.007$; $d = 1.03$).

In der Kontrollgruppe (KG) ließ sich kein Unterschied feststellen ($\Delta M = 0.08$; $p > 0.38$). Die Kontrollgruppe zeigt außerdem, dass die Testwerte über die Zeit stabil bleiben.

Die mittlere Messgenauigkeit der Personenkenwerte anhand der Schätzung ist auch hier mit $EAP_{rel} = 0.74$ akzeptabel.

10. Zusammenfassung und Diskussion

Die Ergebnisse zur **ersten Forschungsfrage (FF 1)** zeigen, dass von einer Eignung des CT-Tests für Grundschullehramtsstudierende ausgegangen werden kann, auch wenn der Test ein paar zu leichte

Aufgaben enthält. Werden diese entfernt, so erscheint die Passung laut Wrightmap gut (Abb. 1). Die Studierenden scheinen außerdem etwas homogener und andererseits etwas CT-fähiger zu sein als die untersuchten Sekundarschüler:innen der 11. Klasse in der Schweiz (vgl. Guggemos et al., 2019). Dieser Vergleich ist jedoch nur bedingt haltbar, da die Schüler:innen einerseits mehr Zeit zum Lösen der Aufgaben (45 Min.) hatten und andererseits drei schwerere Aufgaben im Test vorkamen. Dementsprechend scheint die Aussagekraft über die differente Homogenität valider.

Die Ergebnisse zur **zweiten Forschungsfrage (FF 2)** zeigen einen deutlichen Anstieg der CT-Fähigkeit innerhalb der LLL-Untersuchungs-, aber auch der Parallelgruppe im Theorieseminar. Dieser Befund ist erstaunlich, da im LLL aufgrund der Praxisphasen insgesamt weniger Sitzungen für die Aneignung von theoretischen Inhalten bzw. vertiefendem Fachwissen vorhanden sind. Dieses Ergebnis könnte einerseits darüber erklärt werden, dass die Praxisphasen mit Kindern die CT-Fähigkeiten genauso zu fördern scheinen, wie der theoretische Diskurs und die Arbeit mit verschiedenen Robotikmaterialien in der

Parallelgruppe. Andererseits kann auch der Umstand, die theoretisch gelernten Inhalte später innerhalb der Praxisphasen anwenden zu dürfen/müssen, zu einem höheren Lernerfolg innerhalb der ersten Theorie-Sitzungen im LLL geführt haben (vgl. Reusser, 2005; Klauser, 1998, S. 278). Dies müsste eine weiterführende Untersuchung von Praxisphasen mit Pre-Post-Design oder ein weiterer Messzeitpunkt innerhalb der LLL in der Inter-Phase klären.

Erweitert man den Blick über die hier beantworteten Forschungsfragen hinaus, so ist das LLL auf Basis bisheriger Forschungsbefunde einer rein theoretischen Vermittlung der informatischen Inhalte sogar ‚überlegen‘, da im LLL zusätzlich zu einem ähnlichen Kompetenzzuwachs im Bereich des Fachwissens auch die Interessen und somit die intrinsische Motivation der Studierenden in Bezug auf das Themenfeld Informatik im Sachunterricht gesteigert bzw. gefördert werden konnten (Brämer et al., 2020a).

11. Literatur

- Adams, R. J., Wu, M. L., & Macaskill, G. (1997). Scaling methodology and procedures for the mathematics and science scales. In M. O. Martin & D. L. Kelly (Eds.), *TIMSS technical report, Volume II: Implementation and analysis* (S. 111-145). Chestnut Hill, MA: Boston College.
- Aho, A. V. (2012). Computation and computational thinking. *The Computer Journal*, 55(7), 832-835. <http://dx.doi.org/10.1093/comjnl/bxs074> (Stand 2/2019)
- Andersen, E. B. (1973). A goodness of fit test for the Rasch model. *Psychometrika*, 38(1), S. 123-140.
- Anthofer, S. & Tepner, O. (2016). Experimentell-fachdidaktisches Wissen und Handeln von Chemie-Lehramtsstudierenden. In C. Maurer (Hrsg.), *Au-thentizität und Lernen – das Fach in der Fachdidaktik: Gesellschaft für Didaktik der Chemie und Physik. Jahrestagung in Berlin 2015* (S. 316 – 318). Universität Regensburg.
- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9 (4), S. 469-520.
- Bock, R. D., & Lieberman, M. (1970). Fitting a response model for n dichotomously scored items. *Psychometrika*, 35(2), S. 179-197.
- Bock, R.D. and Aitkin, M. (1981) Marginal Maximum Likelihood Estimation of Item Parameters Application of an EM Algorithm. *Psychometrika*, 46, S. 443-459.
- Brämer, M. & Köster, H. (2021 – im Druck). Nährböden für den digitalen Wandel? Eine Studie zur Förderung von Selbstwirksamkeitserwartungen und Interessen bei Sachunterrichtsstudierenden und -lehrkräften bezüglich informatischer Inhalte im Lehr-Lern-Labor. *GDSU-Journal* 2021. Heft 11. GDSU.
- Brämer, M., Straube, P., Köster, H. & Romeike, R. (2020b). Eine digitale Perspektive für den Sachunterricht. ein Vorschlag zur Diskussion. *GDSU-Journal*, Juli 2020(10), S. 9-19.
- Brämer, M.; Rehfeldt, D.; Bauer, C.; Köster, H. (2020a) Vorerfahrungen, Interessen und Selbstwirksamkeitserwartungen von Grundschullehramtsstudierenden und -lehrkräften bezüglich informatischer Inhalte. In Nordmeier, V.; Grötzebauch, H. (Hrsg.): *PhyDid B, Didaktik der Physik, Beiträge zur DPG-Frühjahrstagung*, Bonn 2020, S. 97-105.
- Brennan, K. & Resnick, M. (2012). New frameworks for studying and assessing the development of computational thinking. In *Proceedings of the 2012 annual meeting of the american educational research association*. <http://scratched.gse.harvard.edu/ct/files/AERA2012.pdf> (Stand 2/2019)
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson Studium.
- Christensen, K. B., Makransky, G. & Horton, M. (2017). Critical Values for Yen's Q3: Identification of Local Dependence in the Rasch Model Using Residual Correlations. *Applied Psychological Measurement*, 41(3), S. 178-194.
- Bauer, D. F. (1972). Constructing Confidence Sets Using Rank Statistics. *Journal of the American Statistical Association*, 67(339), S. 687-690. <https://doi.org/10.1080/01621459.1972.10481279> (Stand 6/2021)
- Dengel, A.; Heuer, U. (2017). Aufbau des Internets: Vorstellungsbilder angehender Lehrkräfte. In Diethelm, I. (Hrsg.) (2017). *Informatische Bildung zum Verstehen und Gestalten der digitalen Welt*. Bonn: Köllen.
- Döbeli Honegger, B. (2016): *Mehr als 0 und 1 - Schule in einer digitalisierten Welt*. Bern: hep.
- Döbeli Honegger, B., Hielscher, M. (2017). Vom Lehrplan zur LehrerInnenbildung - Erste Erfahrungen mit obligatorischer Informatikdidaktik für angehende Schweizer PrimarlehrerInnen. In Diethelm, I. (Hrsg.) (2017). *Informatische Bildung zum Verstehen und Gestalten der digitalen Welt*. Bonn: Gesellschaft für Informatik.
- Engbring, D. (2017). Aller Anfang ist schwer! Wie gelingt der Einstieg in den Informatikunterricht? In I. Diethelm (Hrsg.), *Informatische Bildung zum Verstehen und Gestalten der*

- digitalen Welt 17. (S. 227–236). Bonn: Gesellschaft für Informatik.
- Gallenbacher, J. (Hrsg.). (2015). Informatik 2015. Informatik allgemeinbildend begreifen. In I. Diethelm (Hrsg.) 16. GI-Fachtagung Informatik und Schule. Bonn: Gesellschaft für Informatik.
- GDSU (2021): Positionspapier Sachunterricht und Digitalisierung. Erarbeitet von der AG Medien & Digitalisierung der Gesellschaft für Didaktik des Sachunterrichts – GDSU (Markus Peschel, Friedrich Gervé, Inga Gryl, Thomas Irion, Daniela Schmeinck, Philipp Straube). Online-Publikation, [http://www.gdsu.de/wb/media/Medien und Digitalisierung/GDSU 2021 Positionspapier Sachunterricht und Digitalisierung deutsch de.pdf](http://www.gdsu.de/wb/media/Medien%20und%20Digitalisierung/GDSU%202021%20Positionspapier%20Sachunterricht%20und%20Digitalisierung%20deutsch%20de.pdf) (Stand 8/2021).
- GDSU (2013). Perspektivrahmen Sachunterricht. Bad Heilbrunn: Verlag Julius Klinkhardt.
- Gesellschaft für Informatik [GI] (2019). Kompetenzen für informatische Bildung im Primarbereich. [https://dl.gi.de/bitstream/handle/20.500.12116/20121/61-GI-Empfehlung Kompetenzen informatische Bildung Primarbereich.pdf?sequence=1&isAllo-wed=y](https://dl.gi.de/bitstream/handle/20.500.12116/20121/61-GI-Empfehlung_Kompetenzen_informatische_Bildung_Primarbereich.pdf?sequence=1&isAllo-wed=y) (Stand 8/2021).
- Gläser, E. (2020): Professionswissen von Sachunterrichtsstudierenden zu Digitaler und Informatischer Bildung. In Skorsetz, N./Bonanati, M./Kucharz, D. (Hrsg.): Diversität und soziale Ungleichheit: Herausforderungen an die Integrationsleistung der Grundschule. (S. 315–319) Wiesbaden: Springer Fachmedien.
- Goodson, I. (H.), Hopmann, S. (H.) & Riquarts, K. (H.). (1999). Das Schulfach als Handlungsrahmen. Vergleichende Untersuchung zur Geschichte und Funktion der Schulfächer (Bildung und Erziehung Beiheft, Bd. 7). Köln: Böhlau.
- Gouws, L. A., Bradshaw, K. & Wentworth, P. (2013). Computational thinking in educational activities. An evaluation of the educational game light-bot. In Proceedings of the 18th ACM conference on innovation and technology in computer, S. 10–15.
- Gräsel, C. (2010). Stichwort: Transfer und Transferforschung im Bildungsbereich. Zeitschrift für Erziehungswissenschaft, 13(1), S. 7-20. <https://doi.org/10.1007/s11618-010-0109-8> (Stand 8/2021)
- Guggemos, J., Seufert, S. & Román-González, M. (2019). Measuring computational thinking - adapting a performance test and a self-assessment instrument for german-speaking countries. In Proceedings of the 16th International Conference on Cognition and Exploratory Learning in Digital Age (CELDA 2019) (S. 183–191). IADIS Press.
- Hsu, T.-C., Chang, S.-C., & Hung, Y.-T. (2018). How to learn and how to teach computational thinking: Suggestions based on a review of the literature. Computers & Education, 126, S. 296–310. <https://doi.org/10.1016/j.compedu.2018.07.004> (Stand 8/2021)
- ICILS 2018 - Eickelmann, B. (2019). Measuring Secondary School Students' Competence in Computational Thinking in ICILS 2018—Challenges, Concepts, and Potential Implications for School Systems Around the World. In S.-C. Kong & H. Abelson (Eds.), Computational thinking education (S. 53–64). Singapore: Springer.
- Jäger, M. (2004). Transfer in Schulentwicklungsprojekten. Wiesbaden: VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-322-83388-4> (Stand 8/2021)
- Kalelioglu, F., Gülbahar, Y. & Kukul, V. (2016). A framework for computational thinking based on a systematic research review. Baltic Journal of Modern, S. 583–596.
- Klauser, F. (1998). Problem-based learning. Ein curricularer und didaktisch-methodischer Ansatz zur innovativen Gestaltung der kaufmännischen Ausbildung. Zeitschrift für Erziehungswissenschaft, 1(2), S. 273–293.
- KMK - Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2017). „Bildung in der digitalen Welt“. Strategie der Kultusminister-konferenz. https://www.kmk.org/fileadmin/Dat-eien/pdf/PresseUndAktuelles/2018/Digital-strategie_2017_mit_Weiterbildung.pdf (Stand 8/2021)
- Knaus, T. (2017). Verstehen – Vernetzen – Verantworten: Warum Medienbildung und informatische Bildung uns alle angehen und wir sie gemeinsam weiterentwickeln sollten. In I. Diethelm (Hrsg.), Informatische Bildung zum Verstehen und Gestalten der digitalen Welt. Lecture Notes in Informatics (LNI). (S. 31–48). Bonn: Gesellschaft für Informatik.
- Kobl, C. & Tepner, O. (2019). Förderung der Reflexionskompetenz von Chemielehramtsstudierenden. In C. Maurer (Hrsg.), Naturwissenschaftliche Bildung als Grundlage für berufliche und gesellschaftliche Teilhabe. Gesellschaft für Didaktik der Chemie und Physik Jahrestagung in Kiel 2018 (S. 325 – 328). IPN Kiel.
- Koller, I., Alexandrowicz, R. & Hatzinger, R. (2012). Das Rasch-Modell in der Praxis. Eine Einführung mit eRm. Wien: Facultas.

- Kommer, S. & Biermann, R. (2012). Der mediale Habitus von (angehenden) LehrerInnen. Medienbezogene Dispositionen und Medienhandeln von Lehramtsstudierenden. In R. Schulz-Zander, B. Eickelmann, H. Moser, H. Niesyto & P. Grell (Hrsg.), *Jahrbuch Medienpädagogik 9* (S. 81-108). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Köster, H. & Galow, P. (2014). Forschendes Lernen initiieren. Hintergründe und Modelle offenen Experimentierens. *Naturwissenschaften im Unterricht*, 25(144), S. 24-26.
- Köster, H., Mehrtens, T., Brämer, M. & Steger, J. (2019). Forschendes Lernen im zyklischen Prozess – Entwicklung eines neuen Lehr-Lern-Formats im Studienfach Sachunterricht. In B. P. J. Roth (Hrsg.), *Lehr-Lern-Labore - Innovationsmotor in der MINT- Lehrpersonenbildung*. (S. 99–111) Wiesbaden.: Springer Spektrum.
- Ludlow, L. H. & O’leary, M. (1999). Scoring Omitted and Not-Reached Items: Practical Data Analysis Implications. *Educational and Psychological Measurement*, 59(4), S. 615–630.
- Macaskill, G. (1997). Scaling methodology and procedures for the mathematics and science scales. In M. O. Martin & D. L. Kelly (Eds.), *TIMSS technical report, Volume II: Implementation and analysis* (S. 111-145). Chestnut Hill, MA: Boston College.
- MacCallum, R. C., Browne, M. W. & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, 1(2), S. 130–149.
- Mair, P., & Hatzinger, R. (2007). CML based estimation of extended Rasch models with the eRm package in R. *Psychology Science*, 49(1), S. 26–43.
- Myles Hollander and Douglas A. Wolfe (1973). *Non-parametric Statistical Methods*. New York: John Wiley & Sons., S. 27–33.
- Rehfeldt, D., Seibert, D., Klempin, C., Lücke, M., Sambanis, M. & Nordmeier, V. (2018). Mythos Praxis um jeden Preis? Die Wurzeln und Modellierung des Lehr-Lern-Labors. *die hochschullehre*, 4, S. 90-113.
- Renkl, A. (1996). Träges Wissen: Wenn Erlerntes nicht genutzt wird. *Psychologische Rundschau*, 47 (2), S. 78–92.
- Resnick, M. (2017). *Lifelong Kindergarten. Cultivating creativity through projects, passion, peers, and play*. Cambridge, Massachusetts, London: The MIT Press.
- Papert, S. (1982). *Mindstorms. Kinder, Computer und Neues Lernen*. <http://dx.doi.org/10.1007/978-3-0348-5357-6> (Stand 8/2021)
- Reusser, K. (2005). Problemorientiertes Lernen – Tiefenstruktur, Gestaltungsformen, Wirkung. Bei-träge zur Lehrerinnen- und Lehrerbildung, 23 (2), S. 159–182.
- Robitzsch, A., Kiefer, T. & Wu, M. (2020). TAM: Test Analysis Modules. Verfügbar unter <https://CRAN.R-project.org/package=TAM>
- Román-González, M., Pérez-González, J.-C., & Jiménez-Fernández, C. (2017). Which cognitive abilities underlie computational thinking? Criterion validity of the computational thinking test. *Computers in Human Behavior*, 72, S. 678–691. <https://doi.org/10.1016/j.chb.2016.08.047> (Stand 8/2021)
- Schaar, P. (2017). Wie die Digitalisierung unsere Gesellschaft verändert. In M. Schröder & A. Schwanebeck (Hrsg.), *Big Data - In den Fängen der Datenkraken. Die (un-)heimliche Macht der Algorithmen* (2. Auflage, S. 105–122). Baden-Baden: Nomos Verlagsgesellschaft mbH & Co. KG.
- Seibert, D., Rehfeldt, D., Klempin, C., Mehrtens, T., Sambanis, M., Köster, H. et al. (2019). Theoretisches Wissen gleich träges Wissen? Praxisrelevanz von fachdidaktischem Wissen in Lehr-Lern-Labor-Seminaren. *die hochschullehre*, 5, S. 355–382.
- Serafini, G. (2011). Teaching Programming at Primary Schools. Visions, Experiences, and Long-Term Research Prospects. In I. Kalas & R. Mittermeir (Hrsg.), *Informatics in Schools. Contributing to 21st Century Education*, Jg. 7013. In *Lecture Notes in Computer Science* (S. 143–154). Berlin Heidelberg: Springer.
- Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard educational review*, 57 (1), S. 1–23.
- Steiner, M. & Himpl-Gutermann, K. (2020) *Computational Thinking und Kontextorientierung*. *medienimpulse*, 58, (1), 2020.
- Straube, P., Brämer, M., Köster, H. & Romeike, R. (2018). Eine digitale Perspektive für den Sachunterricht? Fachdidaktische Überlegungen und Implikationen. Nr. 24. Oktober 2018. <https://www.widerstreit-sachunterricht.de/ebene1/superworte/zumsach/straubetal.pdf> (Stand 8/2021)
- Trempler, K., Schellenbach-Zel, J., Gräsel, C. (2013). Der Einfluss der Motivation von Lehrpersonen auf den Transfer von Innovationen. In: Rürup, M.; Bormann, I. (Hrsg.) (2013). *Innovationen im Bildungswesen*. (S. 329-347). Wiesbaden: Springer.
- Völker, M. & Trefzger, T. (2010). Lehr-Lern-Labore zur Stärkung der universitären Lehramtsausbildung. *PhyDid B, Didaktik der Physik*, Bei-träge zur DPG-Frühjahrstagung.

- Völker, M. & Trefzger, T. (2011). Ergebnisse einer explorativen empirischen Untersuchung zum Lehr-Lern-Labor im Lehramtsstudium. *Phy-Did B – Di-daktik der Physik – Beiträge zur DPG-Frühjahrstagung*.
- Voogt, J., Fisser, P., Good, J., Mishra, P., & Yadav, A. (2015). Computational thinking in compulsory education: Towards an agenda for research and practice. *Education and Information Technologies*, 20(4), S. 715–728. <https://doi.org/10.1007/s10639-015-9412-6> (Stand 8/2021)
- Weigend, M. (2009). Algorithmik in der Grundschule. In B. (H.) Koerber (Hrsg.), *Zukunft braucht Herkunft – 25 Jahre „INFOS - Informatik und Schule“*. (S. 97–108). Bonn: Gesellschaft für Informatik e.V.
- Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, 49(3), 33. https://dl.acm.org/doi/pdf/10.1145/1118178.1118215?casa_token=q0MOybfCnUE-AAAAA:fcDN7NGLCoCuMkE4yWyUM-ShqNNnrxtT5Hx2slEvjNap25wfWO8CPbIveyUOa1KwPQUdAmOQe5SB (Stand 8/2021)
- Wing, J.M. (2017). Computational thinking's influence on research and education for all. *Italian Journal of Educational Technology*, 25(2), S. 7-14. <https://ijet.itd.cnr.it/article/view/922> (Stand 8/2021)
- Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8, S. 370-371.
- Yadav, A., Mayfield, C., Zhou, N., Hambrusch, S. & Korb, J. T. (2014). Computational Thinking in Elementary and Secondary Teacher Education. *ACM Transactions on Computing Education*, 14(1), S. 1–16.
- Yen, W. M. (1984). Effects of local item dependence on the fit and equating performance of the three parameter logistic model. *Applied Psychological Measurement*, 8, S. 125-145.

Anhang

Das Projekt K2teach wird im Rahmen der gemeinsamen „Qualitätsoffensive Lehrerbildung“ von Bund und Ländern aus Mitteln des Bundesministeriums für Bildung und Forschung gefördert (FKZ: 01JA1802).