

Entwicklung und Validierung eines Situational Judgement Test zur Erfassung des Umgangs von Lehrkräften mit digitalen Problemsituationen im Unterricht

Lukas Schulze-Vorberg^{id} und Holger Horz

Institut für Psychologie, Pädagogische Psychologie, Goethe-Universität Frankfurt am Main, Deutschland

Zusammenfassung: Lehrkräfte sind immer stärker gefordert, digitale Medien kompetent in Lehr-Lern-Prozesse einzubeziehen. Es mangelt jedoch an objektiven Verfahren zur Erfassung digitaler Kompetenzen bei dieser Zielgruppe. In diesem Beitrag wird die Entwicklung und Validierung eines *Situational Judgement Test* zur Erfassung des Umgangs mit digitalen Problemsituationen im Unterricht vorgestellt. In einem mehrschrittigen Entwicklungsverfahren wurden auf Grundlage einer Vorerhebung ($n = 72$) Aufgaben entwickelt und in einer Lautes-Denken-Studie ($n = 20$) optimiert. Durch eine Expert_innenbefragung ($n = 8$) wurden Inhaltsvalidität und Scoring-Key bestimmt. In zwei Pretests mit Lehramtstudierenden ($n = 83$; $n = 58$) wurde auf Grundlage von Itemanalysen ein finales Itemset erstellt. Eine Validierungsstudie mit Lehrkräften ($n = 238$) bestätigte die postulierte Faktorenstruktur in einer konfirmatorischen Faktorenanalyse. Die Kriteriumsvalidität konnte durch erwartungskonforme Korrelationen mit externen Kriterien nachgewiesen werden. Die prognostische Validität konnte durch die signifikante Vorhersage von Problemauftrittshäufigkeiten bei der Nutzung digitaler Medien durch Lehrkräfte in einem Strukturgleichungsmodell bestätigt werden.

Schlüsselwörter: Digitale Medien, Situational Judgement Test, Lehrkräfte, Medienkompetenz

The Development and Validation of an Instrument for Assessing Teachers' Competence to Deal With Digital Problems Situations in the Classroom

Abstract: Teachers are increasingly called upon to skillfully integrate digital media into the teaching-learning processes. However, there is a lack of objective procedures for measuring digital skills among this target group. This paper presents the development and validation of a Situational Judgement Test to assess the handling of digital problem situations in the classroom. Employing a multistep development process, we developed a set of problems based on a preliminary survey ($n = 72$) and optimized them using a Loud Thinking Study ($n = 20$). We determined content validity and scoring key through an expert survey ($n = 8$). Two pretests with teachers-in-training ($n = 83$ and $n = 58$) established a final itemset based on item analyses. A validation study with teachers ($n = 238$) confirmed the postulated factor structure utilizing confirmatory factor analysis. The criterion validity was demonstrated through expected correlations with external criteria. We confirmed the prognostic validity with a structural equation model by significantly predicting the frequency of problems occurring when teachers use digital media.

Keywords: digital media, Situational Judgement Test, teacher, ICT literacy

Durch gesellschaftliche Digitalisierungs- und Medialisierungsprozesse werden vermehrt Kompetenzen im Umgang mit digitalen Medien relevant, um die angebotenen Entwicklungschancen und gesellschaftlichen Teilhabeoptionen angemessen und selbstbestimmt nutzen zu können (Kultusministerkonferenz, 2016). Diese Entwicklungen fordern, dass der Bildungsauftrag von Schulen um die Vermittlung erforderlicher Kompetenzen im Umgang

mit digitalen Medien erweitert wird, was auch durch Bund-Länder-Programme wie dem Digitalpakt gefordert und gefördert wird (Bundesrepublik Deutschland, 2019). Lehrkräfte und Schulen sind mit den Herausforderungen des Onlineunterrichts, wie dem nötigen technischen und pädagogischen Wissen und den teilweise unzureichenden digitalen Infrastrukturen an Schulen, konfrontiert (Eickelmann et al., 2019). In diesem Zuge erhalten

Kompetenzen des funktionalen Umgangs mit Problemsituationen beim Unterrichten mit digitalen Medien einen neuen Stellenwert. Der vorliegende Beitrag beschreibt daher die Definition des Konstrukts des *Umgangs mit digitalen Problemsituationen von Lehrkräften im Unterricht* und dessen Operationalisierung durch die systematische Entwicklung und Validierung eines *Situational Judgement Tests* (siehe Campion, Ployhart & MacKenzie, 2014).

Herausforderungen bei der digitalen Mediennutzung in der Schule

Die sinnvolle digitale Medienintegration in den Unterricht bietet unter Berücksichtigung des Primats des Pädagogischen lernförderliches Potential für Schüler_innen (Hillmayr, Reinhold, Ziernwald & Reiss, 2017). Das kann durch adaptive Lehr- bzw. Lernprozesse und das Schaffen von orts- und zeitunabhängigen Lerngelegenheiten, beispielsweise durch die Hinzunahme mobiler, digitaler Endgeräte, ermöglicht werden (Sung, Chang & Liu, 2016). Doch die störungsfreie und sinnvolle Hinzunahme digitaler Medien in den Lehr-Lern-Prozess kann durch Faktoren wie didaktische, organisationale, individuelle und technische Defizite erschwert werden (z. B. entlang des DO-IT-Modells; Horz & Schulze-Vorberg, 2017).

So schätzen Lehrkräfte in Deutschland ihre digitalisierungsbezogenen Kompetenzen gegenüber dem internationalen Vergleichswert überwiegend schlechter ein und auch die tägliche Mediennutzung ist signifikant geringer gegenüber den meisten Teilnehmerländern der *International Computer and Information Literacy Study* (ICILS; Eickelmann et al., 2019). In Studien zu Herausforderungen und Schwierigkeiten von digitalem Lernen an Schulen berichten Lehrkräfte (74 %) und Schulleitungen (73 %) zudem am häufigsten die mangelnde Zuverlässigkeit der technischen Geräte und Systeme (Schmid, Goertz & Behrens, 2017). Die schnellelebige Weiterentwicklung von digitaler Hard- und Software kann dazu führen, dass die eigenständige Fehler- und Problembehebung selbst bei medienkompetenten Lehrkräften heraus- oder sogar überfordernd wird. Die beschriebenen Voraussetzungen begünstigen das Auftreten von Problemsituationen bei der Nutzung digitaler Medien im Unterricht sowohl bei technisch affinen Lehrkräften als auch bei Lehrkräften mit wenig Erfahrung. Es ist daher nicht verwunderlich, dass in aktuellen Studien zwei Drittel der Lehrkräfte technische Ausfälle und Unterbrechungen im Unterricht als Hinderungsgrund für den sinnvollen Einsatz digitaler Medien nennen (Mußmann & Hardwig, 2021). Um einen möglichst störungsarmen Unterricht zu gewährleisten, sollten Lehrkräfte in auftretenden digitalen Problemsituationen im Unterricht adäquat reagieren können. Der Aspekt des

Umgangs mit Problemsituationen bei der *unterrichtlichen, digitalen Mediennutzung* (UDP) findet allerdings in der Forschung bisher kaum Beachtung.

Definition und Verortung des Konstrukts: Umgang mit digitalen Problemsituationen im Unterricht

Der Umgang mit digitalen Problemsituationen im Unterricht soll als kontextspezifische, kognitive Leistungsdispositionen in Form von Kenntnissen, Fertigkeiten oder Routinen (Hartig & Klieme, 2006) definiert werden, die zur Bewältigung von auftretenden Problemen beim Einsatz digitaler Medien im Unterricht von Lehrkräften genutzt werden. Lehrkräfte mit einer hohen Ausprägung des Konstrukts können digitale Problemsituationen im Unterricht sowie deren Ausmaße erkennen und wissen, wie man diese schnell und störungsarm selbst oder durch das Hinzuziehen geeigneter Unterstützung beseitigen oder auf alternative analoge oder digitale Unterrichtsmethoden zurückgreifen kann.

Unter digitalen Problemsituationen werden hier Probleme verstanden, die sich aufgrund der Hinzunahme von Informations- und Kommunikationstechnologien im Unterricht ergeben (angelehnt an die Definition technischer Probleme; Naumann, Goldhammer, Rölke & Stelter, 2014). Ein Problem definiert sich als ungewollte Aufgabe oder Zustand, bei der/dem nicht die Mittel zur Verfügung stehen, diese/-n in den gewünschten Zielzustand zu überführen (Dörner, 1979). Zur besseren Operationalisierbarkeit des Konstrukts wurde die Definition an den Kompetenzbegriff von Hartig und Klieme (2006) angelehnt, in dem affektiv-motivationale Aspekte nicht eingeschlossen sind. Die Gewährleistung einer höchstmöglichen aktiven Lernzeit für Schüler_innen durch einen störungsarmen Unterricht lässt eine enge Verwandtschaft von UDP mit Aspekten von Klassenführung (Kounin, 1976) vermuten. Die Reibungslosigkeit und der Schwung des Unterrichts durch das Verhindern von Unterbrechungen durch technische Defekte können als eine Dimension von Klassenführung in digitalgestützten Unterrichtsszenarien verstanden werden. Das beschriebene Konstrukt eint demnach Elemente von etabliertem Lehrkräfteprofessionswissen (Pädagogisches Wissen: Klassenführung; Voss, Kunter & Baumert, 2011) mit Wissen über digitale Medien (technologisches Wissen). Eine geeignete theoretische Möglichkeit zur Verortung des Konstrukts bietet daher das *Technological Pedagogical Content Knowledge* Modell (TPACK) von Mishra und Koehler (2006). Diese Rahmung bietet durch die Anlehnung an etablierte Konzepte des Professionswissens (Shulman, 1986) eine gute Anschlussfähigkeit an aktuelle Forschung zur professionellen Kompetenz

von Lehrkräften (z. B. Voss et al., 2011). Im TPACK-Modell werden das inhaltliche Wissen (*Content knowledge*, CK), das pädagogische Wissen (*Pedagogical knowledge*, PK) sowie deren Schnittmengen um die Facette des technologischen Wissens (*Technical knowledge*, TK) erweitert. Das technologische Wissen wird dabei vor allem als Wissen über neue technologiebasierte Medien verstanden sowie die Fähigkeit, neue Technologien zu erlernen und zu adaptieren (Mishra & Koehler, 2006). Das technologisch-pädagogische Wissen (TPK) wird als Kenntnis über die Möglichkeiten verschiedener Technologien beschrieben und wie diese in Lehr-Lernszenarien genutzt werden können. Die Facette des TPK bietet sich aufgrund der Verbindung pädagogischer und technologischer Wissensaspekte besonders für eine Verortung des definierten Konstrukts an. Bisher mangelt es jedoch an konsistenten Ausdifferenzierungen dieser und weiterer Schnittmengen, was die Vergleichbarkeit wissenschaftlicher Studienergebnisse erschwert (z. B. Chai, Koh & Tsai, 2013). In bisherigen Operationalisierungen und Definitionen wurde TPK vorwiegend als digitales „Unterrichtsmethodenwissen“ umgesetzt. TPK sollte anknüpfend an Vorarbeiten professioneller Kompetenz von Lehrkräften allerdings weitere Aspekte des pädagogischen Wissens beinhalten (Unterrichtsmethodenwissen, Leistungsbeurteilung, Klassenführung, Heterogenität; Voss et al., 2011). Die hier vorgeschlagene Konstruktdefinition des *Umgangs mit digitalen Problemsituationen im Unterricht* oder die digitale „Inkompetenzkompensationskompetenz“ (Marquard, 2005, S. 23) von Lehrkräften adressiert diese Forderung, indem ein weiterer Aspekt von PK (Klassenführung) in technologiegestützten (TK) Unterrichtssituationen definiert wird.

Zielsetzung

Bislang existieren für die Messung von technologisch-pädagogischem Wissen im Umgang mit digitalen Medien kaum praxisnahe und valide Testverfahren (Drummond & Sweeney, 2017). Diese sind allerdings nötig, um beispielsweise Lernzuwächse durch Lehrkräftefortbildungen oder Schüler_innenoutcomes in Abhängigkeit von Lehrkräftewissen in diesem Bereich valide abzubilden. Die meisten Erhebungsinstrumente im Bereich von TPACK sind Fragebögen (z. B. Schmidt et al., 2009), die den bekannten Problematiken der Selbsteinschätzung unterliegen (Zell & Krizan, 2014). Arbeiten von Baier und Kunter (2020) oder Lachner, Backfisch und Stürmer (2019) erfassen konzeptuelles Wissen (de Jong & Ferguson-Hessler, 1996) im Bereich TPK über offene Testfragen oder Mehrfachwahlaufgaben, allerdings ohne Bezug zu konkreten Unterrichtssituationen. Da Unterricht als komple-

xe soziale, unvorhersehbare und simultane Situation definiert ist (Doyle, 2006), wird ein stärkerer Kontextbezug bei TPACK-Operationalisierungen gefordert (Rosenberg & Koehler, 2015). Um den Forderungen einer verstärkten Berücksichtigung des Kontextbezugs nachzukommen und um das definierte Konstrukt praxisnah abzubilden, wurde die Operationalisierung des Konstrukts in Form eines *Situational Judgement Test* gewählt (Campion et al., 2014). Dieser methodische Ansatz, der den simulationsbasierten Verfahren zuzuordnen ist, bietet durch die Darbietung von konkreten video- oder textbasierten Unterrichtssituationen die Möglichkeit, strategisches Wissen von UDP zu erfassen (Stemler & Sternberg, 2006).

Testentwicklung

Die Aufgaben (eine Beispielaufgabe findet sich im ESM 4) für den Situational Judgement Test zur Erfassung des Wissens über UDP wurden in einem mehrschrittigen Prozess entwickelt, der in den nachfolgenden Abschnitten detailliert beschrieben wird. In einer Vorstudie (1) wurden Lehrkräfte zu möglichen Problemsituationen beim digitalen Medieneinsatz im Unterricht und dazugehörigen Lösungsstrategien befragt. Die aus den Daten entwickelten Aufgabenstämme und Handlungsalternativen wurden mit der Methode des lauten Denkens (2) auf Verständlichkeit und Revisionsbedarf überprüft. In einer Expert_innenbefragung (3) wurden die Passung der entwickelten Aufgaben zum Konstrukt sowie der Auswertungsschlüssel des Instruments entwickelt. In zwei Pretests mit Lehramtsstudierenden (4) wurden Aufgaben mit geringer Trennschärfe und Passung zum Konstrukt eliminiert und eine Auswahl getroffen. In einer weiteren Erprobung der Aufgabenauswahl wurde in einer Validierungsstudie das Instrument hinsichtlich seiner Gütekriterien, der postulierten einfaktoriellem Struktur sowie hypothesenkonformer Zusammenhänge mit verwandten Konstrukten überprüft.

Vorstudien

(1) Itementwicklung auf Grundlage von *Critical Incidents* und Lösungsstrategien bei der digitalen Mediennutzung im Unterricht

Um möglichst realitätsnah und umfänglich kritische Situationen und Handlungsalternativen zu ermitteln, wurde eine onlinebasierte Vorstudie mit 72 Gymnasiallehrkräften (44 weiblich; $M = 42.85$ Jahre; $SD = 9.68$) durchgeführt. Die teilnehmenden Lehrkräfte sollten dabei die

Nutzungshäufigkeit von unterschiedlichen digitalen Medien im Unterricht sowie die Ursachen und die Häufigkeiten von Problemauftritten angeben (Eigenverschulden vs. Fremdverschulden; Hardware vs. Software). Zudem sollten die teilnehmenden Lehrkräfte Problemsituationen (*Critical-Incidents-Technique*; Flanagan, 1954) bei der digitalen Mediennutzung im Unterricht nennen und mögliche, sowohl kompetente als auch inkompetente, Handlungsstrategien beschreiben. Aus den genannten Problemsituationen wurden 38 Aufgabenstämme mit jeweils fünf bis sieben Handlungsoptionen extrahiert. Bei der Entwicklung der Situationen wurden unterschiedliche Unterrichtsphasen, Problemquellen (Hard- und Software), Medien und Lösungsstrategien berücksichtigt. Die Entwicklung der Antwortoptionen wurde anhand eines aus den Daten abgeleiteten Kategoriensystems vorgenommen. Als Handlungsstrategien wurden beispielsweise alternative (analoge) Unterrichtsgestaltungen, das Umgehen der Probleme durch Raumtausch oder unproduktive Strategien, wie das Abwarten oder die langwierige Befassung mit dem Problem, entwickelt. Um das Konstrukt sowohl bei der Nutzung digitaler Medien bei Expert_innen als auch bei Laien erfassen zu können, wurden bei der Entwicklung der Problemsituationen unverschuldete technische Problemsituationen fokussiert (z.B. nicht bootender Computer). Eine Situationsbeschreibung mit einer intendierten vorangegangenen fehlerhaften Nutzung der Lehrkraft hätte die Realitätsnähe und Beantwortung von technisch affinen Lehrkräften beeinflusst, da diese vermutlich schon in der Entstehung der Problemsituation anders gehandelt hätten.

Die Instruktion des Instruments wurde standardisiert, um eine Vergleichbarkeit der Ergebnisse trotz heterogener schulischer Ausstattung mit digitalen Medien gewährleisten zu können. In der Instruktion wurden die Lehrkräfte gebeten, sich die Anstellung an einer fiktiven Schule mit einer vorgegebenen digitalen Infrastruktur und beschriebenen Möglichkeiten eines IT-Supports vorzustellen.

(2) Lautes-Denken-Studie: Testoptimierung

Ziel der Studie war es, Unklarheiten und Optimierungsbedarfe zu lokalisieren. Die Aufgaben wurden 20 Gymnasiallehrkräften unterschiedlichen Alters und unterschiedlicher Berufserfahrung (14 weiblich; $M = 35.65$ Jahre, $SD = 11.09$; Berufserfahrung: $M = 11.85$ Jahre, $SD = 8.41$) mit der Methode des lauten Denkens (Charters, 2003) vorgelegt. Zudem wurden Effekte des Einsatzes von Multimediaelementen in Form von Bildern der Fehlermeldungen im Aufgabenstamm, der Formulierung der Handlungsalternativen (Ich-Perspektive vs. dritte Person) so-

wie des Wordings der Skalierung (effektiv vs. kompetent) systematisch überprüft.

Hierfür wurden über 20 Stunden Interviewmaterial transkribiert, systematisch ausgewertet und revisionsbedürftige Aufgabenstämme und Handlungsalternativen überarbeitet. Aufgrund der Rückmeldungen der Proband_innen wurde aufgabenübergreifend die Ich-Perspektive (13 Nennungen, 4 Dritte-Person, 3 egal) in den Handlungsmöglichkeiten verwendet, die Unterstützung mit Bildmaterial (13 Nennungen) bei ausgewählten Aufgabenbeschreibungen festgelegt sowie der Begriff kompetent (8 Nennungen, 7 effektiv, 5 egal) für die Skalierung der Likert-Skala genutzt.

(3) Expert_innenbefragung: Realitätsnähe, Verständlichkeit, Inhaltsvalidität und Scoring-Key

Ziel der Expertenbefragung war es, die Inhaltsvalidität des Itemuniversums, den *Scoring-Key*, die Realitätsnähe und Verständlichkeit auf Aufgabenebene sowie die Passung der Aufgaben zum Konstrukt zu bestimmen. Die überarbeiteten 38 Aufgaben wurden acht Expert_innen (fünf weiblich; vier Professor_innen, eine habilitiert, zwei Postdocs, eine wissenschaftliche Mitarbeiterin) aus den Bereichen der Bildungsforschung und Digitalisierung an Universitäten in Deutschland vorgelegt. Neben den Aufgaben wurde den Expert_innen auch die Definition des Konstrukts UDP präsentiert und gebeten, jede der 38 Aufgabenstämme mit den jeweiligen Handlungsmöglichkeiten auf einer sechsstufigen Likertskala (1 = *stimme voll und ganz zu* bis 6 = *stimme überhaupt nicht zu*) hinsichtlich der Passung der Aufgaben zum Konstrukt ($M = 1.81$; $SD = 0.78$), der Realitätsnähe ($M = 1.56$; $SD = 0.67$) und der Verständlichkeit ($M = 1.5$; $SD = 0.70$) zu bewerten. Die Rückmeldungen der Expert_innen bezüglich der Konstruktpassung der entwickelten Aufgabenstämme und Handlungsalternativen sprechen für eine gute Inhaltsvalidität (für eine detaillierte tabellarische Übersicht über die Expert_innenrückmeldungen siehe ESM 1).

Für die Ermittlung des Scoring-Keys wurden alle Handlungsalternativen durch die Expert_innen hinsichtlich eines kompetenten Handelns auf einer sechsstufigen Skala (1 = *sehr kompetent* bis 6 = *sehr inkompetent*) bewertet. Der Wert des Übereinstimmungsmaßes (*Intraclass Correlation*; ICC) für das Scoring der Handlungsalternativen lag mit .54 in einem moderaten Bereich (Koo & Li, 2016; Shrout & Fleiss, 1979). Um dem moderaten ICC bei der Erstellung des Scoring-Keys zu begegnen, wurde der Modalwert (*Mode Consensus*) der Expert_innen herangezogen (Weng, Yang, Lievens & McDaniel, 2018). Wenn der Modalwert bei einem Item nicht klar gebildet werden

konnte, wurden zur Entscheidungsfindung für die Scoring-Keys die Ergebnisse der Proband_innen der Lautes-Denken-Studie hinzugezogen.

(4) Pretests: Itemanalysen und -auswahl

Ziel der Pretests war es, aus den 38 entwickelten Aufgaben eine Auswahl für die Validierungsstudie zu ermitteln. In einer ersten Testung wurden 83 Lehramtsstudierenden (38 weiblich; $M = 23.3$ Jahre, $SD = 4.66$) zwei Bündel à acht bis zehn Aufgaben in einem Testheftdesign mit randomisierter Zuteilung vorgelegt. Die durchschnittliche Bearbeitungszeit für eine Aufgabe lag bei etwa 1.5 Minuten. Die Auswertung erfolgte automatisiert mittels vorgefertigter Syntax mit der Software SPSS (Version 26.0): Bei einer Übereinstimmung mit dem Modalwert der Expert_innenmeinung wurde ein Punkt pro Item vergeben, bei Abweichungen kein Punkt. Die Punkte wurden pro Aufgabe aufsummiert, dass je nach Anzahl der möglichen Handlungsstrategien zwischen null und maximal sieben Punkte erreicht werden konnten. Die aufsummierten Punkte jeder Aufgabe wurden als Grundlage für die Itemanalysen genutzt. Insgesamt wurden nach der ersten Testung 22 Aufgaben anhand ihrer Trennschärfe ($r_{i(i-1)} > .30$) und inhaltlicher Passung zum Konstrukt ausgewählt. In einer zweiten Testung mit 58 Lehramtsstudierenden (38 weiblich; $M = 22.9$ Jahre, $SD = 3.01$) wurden jeder Testperson in randomisierter Reihenfolge 22 Aufgaben vorgelegt. Anhand von empirischer Trennschärfe, Schwierigkeit und inhaltlicher Passung (Expert_innenrating) zum Konstrukt wurden 16 Aufgaben ausgewählt (siehe auch ESM 1). Die ausgewählten Aufgaben zeigten bei vertiefenden Analysen eine gute interne Konsistenz ($\alpha = .77$). Analog zu den Ergebnissen bereits durchgeführter Studien zum Informationsgehalt von Situational-Judgement-Items waren dies ebenfalls mehrheitlich Handlungsalternativen mit Scoring-Keys in den Extrembereichen und weniger im mittleren Bereich der Likertskala (McDaniel, Psotka, Legree, Yost & Weekley, 2011).

Validierungsstudie

Nach vielversprechenden Ergebnissen der Vorstudien sollte der aus 16 Aufgaben bestehende Test weiter erprobt und erste Hinweise für dessen Validität gesammelt werden. Dabei sollten insbesondere die faktorielle Struktur des Tests sowie die Konstrukt- und Kriteriumsvalidität empirisch überprüft werden.

Faktorielle Struktur

Das Ergebnis von UDP sollte eine Vermeidung unnötiger Unterbrechungen und damit einhergehend eine Gewährleistung höchstmöglicher aktiver Lernzeit für Schüler_innen sein (Reibungslosigkeit und Schwung in digitalgestütztem Unterricht). Die Instruktionen, die entwickelten Aufgabenbeschreibungen und die dazugehörigen Handlungsalternativen zielen alle auf diesen Aspekt ab.

H1: Der entwickelte UDP-Test sollte eine eindimensionale Struktur aufweisen.

Diskriminante und inkrementelle Validität

Abgeleitet aus der postulierten Konstruktdefinition sollen in Hinblick auf die diskriminante und inkrementelle Validität (Hartig, Frey & Jude, 2008) des UDP-Tests Hypothesen in Bezug zu verwandten, jedoch inhaltlich verschiedenen, Konstrukten überprüft werden:

Technologisches Wissen. TK beinhaltet Wissen über die Nutzung von Hard-/Software und allgemein den Umgang mit technischen Situationen, die keinen konkreten Unterrichtsbezug aufweisen. Studien konnten allerdings zeigen, dass hohe Selbsteinschätzungen im Bereich des technologischen Wissens einen positiven Einfluss auf die selbst eingeschätzten Kompetenzen im Bereich von Klassenführung haben (Saritepeci, 2021).

H2: Es wird daher eine schwache bis moderate Korrelation des UDP-Tests mit dem technologischen Wissen erwartet.

Technologisches Unterrichtsmethodenwissen. Technologisches-(Unterrichtsmethodenwissen) als Subfacette von TPK bezieht sich auf den didaktisch sinnvollen Einsatz digitaler Medien und damit nicht konkret auf Aspekte von Klassenführung bei technologiegestütztem Unterricht. Bisherige Studien konnten allerdings zeigen, dass zwischen Wissen zu (analogen) Unterrichtsmethoden und Klassenführungstechniken signifikante, positive Korrelationen bestehen (Voss et al., 2011).

H3: Es wird eine schwache bis moderate Korrelation zwischen dem UDP-Test und digitalem Unterrichtsmethodenwissen erwartet.

Momentum. Das Momentum wird als Vermeidung unnötiger Unterbrechungen und Störungen des Unterrichtsflusses definiert (Gold & Holodyski, 2015). Als ein Bestandteil von Klassenführung (Kounin, 1976) ist das Momentum damit im Bereich der PK verortet. Da UDP als ein Aspekt von Klassenführung in technologiegestützten Unterrichtsettings (TPK) verstanden werden kann, ist ein Zusammenhang mit dem Momentum (PK) zu erwarten.

H4: Die Facette Momentum von Klassenführung korreliert schwach bis moderat mit dem UDP-Test.

Bewältigungsverhalten (Coping). Bewältigungsverhalten (engl. *Coping*) sind die Bemühungen einer Person, sowohl durch kognitive Prozesse als auch Verhalten mit Anforderungen umzugehen, welche die eigenen Ressourcen fordern oder übersteigen (Lazarus & Folkman, 1984). Für Lehrkräfte liegen Untersuchungen zu Bewältigungsstrategien und dem Zusammenhang mit Klassenführungs Kompetenzen vor (Romi, Lewis & Roache, 2013). Es ist nahe liegend, dass Lehrkräfte zur Gewährleistung von aktiver Lernzeit vorwiegend funktionale Bewältigungsstrategien im Bereich der Problemlösung oder des Hilfesuchens nutzen sollten und weniger die dysfunktionale Strategie der Vermeidung von Lehr-Lernsituationen unter Einbezug digitaler Medien.

H5: Es wird eine schwache bis moderate positive Korrelation zwischen dem UDP-Test und dem Bewältigungsverhalten des Problemlösens erwartet.

H6: Es wird eine schwache bis moderate positive Korrelation zwischen dem UDP-Test und dem Bewältigungsverhalten der Suche nach sozialer Unterstützung erwartet.

H7: Es wird eine schwache bis moderate negative Korrelation zwischen dem UDP-Test und dem Bewältigungsverhalten der Vermeidung erwartet.

Selbstwirksamkeitserwartung im Umgang mit digitalen Medien. Selbstwirksamkeitserwartung wird als das Zutrauen verstanden, herausfordernde Situationen durch eigene Handlungen erfolgreich bewältigen zu können (Bandura, 1977). Die Selbstwirksamkeitserwartung im Umgang mit digitalen Medien im Unterricht wird daher als das Zutrauen verstanden, auftretende digitale Problemsituationen funktional bewältigen zu können. Ein positiver Zusammenhang zwischen der Kompetenzeinschätzung von Lehrkräften und einem praxisnahen Wissenstest in diesem Bereich liegen nahe. Studien im Bereich der Lehrkräftekompetenzmessung konnten allerdings nur geringe Korrelationen von Selbsteinschätzungen und Testwerten zeigen (Kunter & Klusmann, 2010).

H8: Es wird eine schwache bis moderate positive Korrelation des Ergebnisses des UDP-Tests und der Selbstwirksamkeitserwartung erwartet.

Digitale Problemsituationen. Aufgaben oder Situationen werden als problematisch wahrgenommen, wenn diese sich nicht oder nicht ausschließlich durch den Einsatz von Routineoperationen auflösen lassen. Personen mit einem funktionalen Umgang mit digitalen Problemsituationen im Unterricht sollten weniger digitale Problemsituationen berichten, da sie über funktionale Strategien und Mittel zur Kompensation verfügen, um diese zu umgehen oder zu lösen.

H9: Der Testwert des UDP-Instruments liefert über die Selbstwirksamkeitserwartung hinaus einen inkrementellen Beitrag zur Vorhersage der wahrgenommenen, digitalen Problemsituationen im Unterricht.

Stichprobe

Die Stichprobe umfasste 238 Lehrkräfte (133 weiblich; Berufserfahrung: $M = 13.58$ Jahre; $SD = 8.84$) von sieben Schulen mit gymnasialer Oberstufe im Raum Frankfurt am Main.

Design und Vorgehen

Für die Validierungsstudie wurden alle Kooperationschulen der Goethe-Universität Frankfurt angeschrieben und zur Teilnahme eingeladen. Für die Teilnahme wurde den Schulen die Auswertung der ermittelten Daten für die Erstellung beziehungsweise Überarbeitung eines schuleigenen Medienbildungskonzeptes angeboten. Die Studienteilnahme der Schulen war freiwillig und wurde online über die Software Unipark (Questback GmbH, 2017) entweder per Fragebogenlink in Präsenz oder über einen personalisierten Link zur asynchronen Teilnahme distribuiert. Die statistischen Analysen wurden mit der Software MPlus 6.1 (Muthén & Muthén, 1998–2010) durchgeführt. Aufgrund der Robustheit gegenüber der Verletzung der Normalverteilung und der kategorialen Datengrundlage wurde der *Weighted Least Square Mean and Variance* (WLMSV) Schätzer herangezogen (Brown, 2006; Li, 2016).

Instrumente

Zur Überprüfung der Validität sowie der postulierten Konstruktdefinition wurden folgende Instrumente eingesetzt (für eine detaillierte tabellarische Übersicht der eingesetzten Instrumente und Korrelationen siehe ESM 2):

Technologisches Wissen. Das *technologische Wissen* (sieben Items, $\alpha = .91$, $M = 3.46$, $SD = 0.84$) wurde auf einer fünfstufigen Likertskala (1 = *stimme überhaupt nicht zu* bis 5 = *stimme voll und ganz zu*) mittels Selbsteinschätzung erhoben (Schmidt et al., 2009).

Technologisch-(pädagogisches) Wissen. Das *technologisch-pädagogische (Unterrichtsmethoden-) Wissen* (neun Items, $\alpha = .86$, $M = 3.48$, $SD = 0.67$) wurde auf einer fünfstufigen Likertskala (1 = *stimme überhaupt nicht zu* bis 5 = *stimme voll und ganz zu*) mittels Selbsteinschätzung erhoben (Schmidt et al., 2009). Zudem wurde das *technologisch-pädagogische (Unterrichtsmethoden-) Wissen* objektiv (sieben Items, $\alpha = .74$, $M = 6.82$, $SD = 4.63$) durch das Instrument von Baier und Kunter (2020) erfasst. Der Test besteht aus acht Fragen, die offen beantwortet werden müssen. Ein Item wurde aufgrund schlechter Kennwerte für weitere Analysen ausgeschlossen.

Bewältigungsstrategien (Coping). Das Bewältigungsverhalten wurde mit einer adaptierten Version des *Brief COPE* ohne die im Kontext nicht relevanten Subskalen *Religion* und *Alkohol-/Drogenkonsum* auf einer vierstufigen Likertskala (1 = *überhaupt nicht* bis 4 = *sehr*) erhoben (Carver, 1997; Knoll, Rieckmann & Schwarzer, 2005). Analog zum Vorgehen von Baumstarck und Kolleg_innen (2017) wurden die Items in eine vierfaktorielle Struktur zu den Facetten *Hilfesuchen* (sechs Items, $\alpha = .78$, $M = 2.31$, $SD = 0.53$), *Vermeidung* (acht Items, $\alpha = .86$, $M = 1.54$, $SD = 0.44$), *Problemlösen* (vier Items, $\alpha = .77$, $M = 2.99$, $SD = 0.60$) und *Positives Denken* (sechs Items, $\alpha = .72$, $M = 2.67$, $SD = 0.55$) aggregiert.

Klassenführung (Managing Momentum). Zur Erfassung des *Momentums* wurde die Subskala aus dem von Gold und Holodynski (2015) entwickelten Situational Judgment Test zur Erfassung von Klassenführung eingesetzt (drei Items, $\alpha = .69$, $M = 0.77$, $SD = 0.17$). Lehrkräfte bewerten auf einer sechsstufigen Skala (1 = *sehr gut* bis 6 = *ungenügend*) verschiedene Handlungsoptionen bei vorgegebenen Facetten. Über Paarvergleiche werden Punkte vergeben und damit die prozentuale Übereinstimmung mit dem Rating der Expert_innen ermittelt (.00 keine Übereinstimmung bis 1.00 volle Übereinstimmung).

Digitale Selbstwirksamkeitserwartung. Um der Kritik der zu engen Betrachtung der Computer-Selbstwirksamkeit (Scherer, Siddiq & Tondeur, 2019) zu begegnen, wurde eine sprachlich adaptierte Fassung der Selbstwirksamkeitserwartungsskala von Schwarzer und Jerusalem (1999) eingesetzt (zehn Items, $\alpha = .93$, $M = 2.67$, $SD = 0.63$). Diese reduziert den digitalen Medieneinsatz nicht nur auf den Einsatz von Computern, sondern erfasst die generelle Selbstwirksamkeit bezüglich der digitalen Mediennutzung im Unterricht auf einer vierstufigen Likertskala (1 = *stimmt nicht* bis 4 = *stimmt genau*).

Berichtete Problemhäufigkeit. Die Erfassung der Problemhäufigkeit beim Einsatz von Computer, Notebook, Tablet, Digitalkamera, interaktivem Whiteboard, Beamer, Visualizer und Smartphone im Unterricht erfolgte auf einer selbstentwickelten, fünfstufigen Likertskala (1 = *nie* bis 5 = *immer*). Lehrkräfte wurden nur zu der Problemhäufigkeit befragt, wenn sie in einem vorherigen Item bestätigt haben, dass sie das digitale Medium im Unterricht verwenden. Aus den Angaben wurde ein aggregierter Mittelwert der Problemhäufigkeit erstellt (acht Items, $\alpha = .85$, $M = 3.55$, $SD = 0.71$).

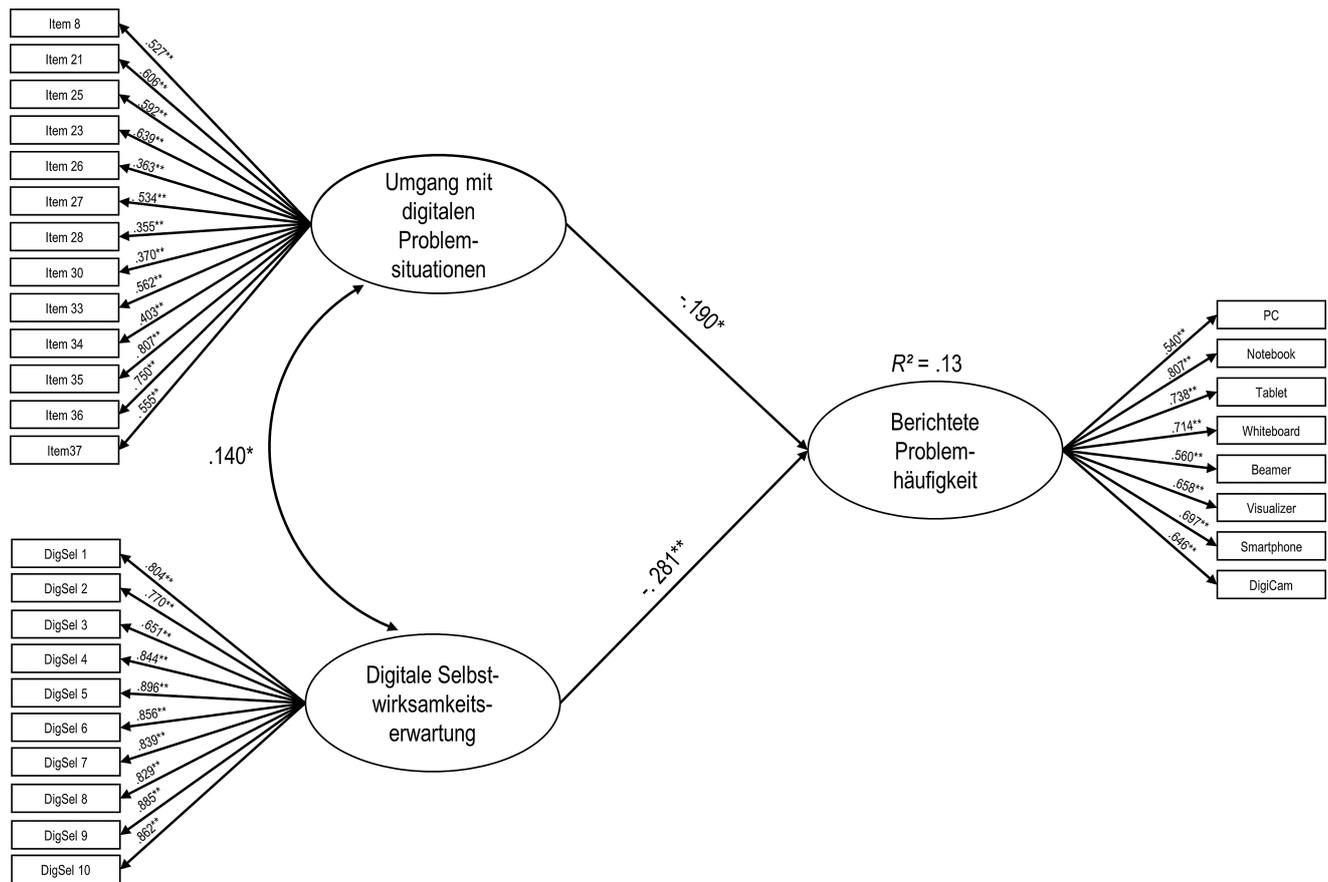
Ergebnisse

Prüfung der faktoriellen Struktur des entwickelten Tests

Um die postulierte einfaktorielle Struktur des Instruments zur Erfassung des Umgangs mit digitalen Problemsituationen zu überprüfen, wurden konfirmatorische Faktorenanalysen durchgeführt. Die Ergebnisse des einfaktoriellen Messmodells mit 16 Aufgaben zeigten keinen zufriedenstellenden Modellfit ($\chi^2_{[104]} = 176.30$, $p = .00$, CFI = .886, RMSEA = .054, TLI = .868). Die nachträgliche Modifikation von Modellen auf Grundlage von Ergebnissen konfirmatorischer Faktorenanalysen wird in Hinblick auf den Verlust des konfirmatorischen Charakters kritisch diskutiert (z.B. Bühner, 2011). Die Pretests wurden mit Lehramtsstudierenden durchgeführt, daher wurden drei Items mit nicht signifikanter bzw. schwacher Faktorladung in der Validierungsstudie mit Lehrkräften nachträglich ausgeschlossen (für eine detaillierte tabellarische Übersicht der Faktorladungen siehe ESM 3). Das optimierte Modell mit 13 Aufgaben und insgesamt 73 Handlungsalternativen zeigt eine sehr gute Modellpassung ($\chi^2_{[65]} = 79.89$, $p = .101$, CFI = .990, RMSEA = .031, TLI = .998). Um die einfaktorielle Struktur zu bestätigen, wurde ein alternatives, zweifaktorielles Modell berechnet, welches die digitalen Problemsituationen in zwei übergeordnete Faktoren teilt: Problemsituationen mit klassischen digitalen Endgeräten (Computer, Beamer) und Problemsituationen mit neueren digitalen Medien (Tablet, interaktives Whiteboard) ($\chi^2_{[64]} = 78.199$, $p = .109$, CFI = .991, RMSEA = .031, TLI = .988). Da bei der Verwendung des WLSMV-Schätzers die einfachen Werte der χ^2 -Statistik nicht zu Modellvergleichen genutzt werden können, wurde über die DIFFTEST-Funktion in MPlus 6.1 (Muthén & Muthén, 1998–2010) ein Modellvergleich mit angepasster χ^2 -Statistik herangezogen (Asparouhov & Muthén, 2006). Die Ergebnisse der durchgeführten Analysen zeigten bei der Beibehaltung der restriktiveren Struktur keine signifikante Verschlechterung der Modellpassung ($\chi^2_{\text{diff [1]}} = 2.693$, $p = .101$). Aufgrund der hohen Korrelation der Faktoren in der zweifaktoriellen Lösung ($r = .92$) sowie der Ergebnisse des Chi-Quadrat-Differenztests wurde die einfaktorielle Struktur ($\alpha = .81$) beibehalten. Hypothese H1 kann demnach bestätigt werden.

Konstruktvalidität des entwickelten Tests

Um die Validität des Konstrukts (Hartig et al., 2008) empirisch nachzuweisen, wurden auf latenter Ebene Korrelationen mit den Konstrukten des *technologischen* und



Anmerkung: * $p < .05$; ** $p < 0.01$

Abbildung 1. Grafische Darstellung des Strukturgleichungsmodells zur Vorhersage von wahrgenommenen Problemsituationen durch Lehrkräfte beim Einsatz digitaler Medien.

technologisch-pädagogischen Wissens, des Momentums, der Bewältigungsstrategien und der digitalen Selbstwirksamkeitserwartung betrachtet. Der entwickelte UDP-Test (ohne Aufgaben 9, 24 und 29) zeigt schwache bis moderate, aber signifikante Korrelationen (Cohen, 1988) mit dem technischen ($r = .23$, $p < .001$) und technologisch-pädagogischen (Methoden-)Wissen (Selbsteinschätzung: $r = .16$, $p < .05$; objektive Erfassung: $r = .37$, $p < .001$), dem Momentum ($r = .34$, $p < .001$), sowie der digitalen Selbstwirksamkeitserwartung ($r = .14$, $p < .001$) auf. Mit den Skalen der Bewältigungsstrategien (Carver, 1997) ergaben sich moderate, signifikante Korrelationen bei den Skalen des Problemlösens ($r = .41$, $p < .001$) und der Vermeidung ($r = -.37$, $p < .001$). Die Skalen Hilfesuchen ($r = -.08$, $p = .315$) und positives Denken ($r = -.07$, $p = .328$) zeigten hingegen keine signifikanten Korrelationen. Die Hypothesen H2-H8 können demnach mit Ausnahme von H6 bestätigt werden.

Kriteriumsvalidität des entwickelten Tests

Zur Überprüfung der inkrementellen Validität des entwickelten Tests wurde dessen Einfluss auf die wahrgenommenen Problemhäufigkeiten bei der Verwendung digitaler Medien im Unterricht betrachtet. In Abbildung 1 wird ersichtlich, dass das Instrument zur Erfassung des Umgangs mit digitalen Problemsituationen im Unterricht als Prädiktor für die berichtete Häufigkeit von Problemsituationen beim Einsatz digitaler Medien dient. Die Modellpassung zeigt gute Fitindizes ($\chi^2_{[431]} = 534.338$, $p = .001$, CFI = .985, RMSEA = .032, TLI = .983). Die Hinzunahme des Instruments zur Erfassung des Umgangs mit digitalen Problemsituationen im Unterricht ergibt eine signifikante Zunahme aufgeklärter Varianz von $\Delta R^2 = .033$. Hypothese H9 kann demnach bestätigt werden.

Diskussion und Ausblick

Das Ziel des vorliegenden Beitrags war die systematische Entwicklung eines praxisnahen Erhebungsinstruments zur Erfassung des Umgangs von Lehrkräften mit digitalen Problemsituationen in Form eines Situational Judgement Test.

Inhaltliche Validität und faktorielle Struktur

Die entwickelten Aufgabenstämme samt Handlungsalternativen weisen nach sorgfältiger Entwicklung eine sehr gute inhaltliche Validität auf. Lehrkräfte sowie Expert_innen aus den Bereichen der Digitalisierung und Bildungsforschung bestätigen die Realitätsnähe und die Qualität der konstruierten Aufgaben. Letztere bestätigten auch die inhaltliche Passung zum postulierten Konstrukt. Die Überprüfung der faktoriellen Struktur der in Vortests selektierten Aufgaben zeigt keine signifikante Überlegenheit einer zweidimensionalen gegenüber der postulierten eindimensionalen Struktur.

Konstruktvalidität

Die Ergebnisse zur Überprüfung der diskriminanten Validität (H2-H8) des UDP-Tests fielen größtenteils hypothesenkonform aus. Es konnten signifikante, schwache bis moderate positive Korrelationen von selbsteingeschätztem technologischen Wissen, technologisch-pädagogischen Unterrichtsmethodenwissen, digitaler Selbstwirksamkeitserwartung, Problemlösen sowie dem *Momentum* mit den UDP-Testergebnissen gezeigt werden. Entsprechend der Annahmen zeigte sich auch eine schwache, negative aber signifikante Korrelation mit der Bewältigungsstrategie der Vermeidung. Das Testergebnis korreliert erwartungskonform nicht mit der Bewältigungsstrategie des *positiven Denkens*, da diese im entwickelten Test nicht adressiert wurde. Einzig die Subfacette *Suche nach sozialer Unterstützung* des Bewältigungsverhaltens korreliert nicht erwartungskonform mit den Ergebnissen des UDP-Tests. Da die entwickelten Handlungsmöglichkeiten zum Aufsuchen von Unterstützung durch den IT-Support beinahe ausschließlich mit einer Unterbrechung des Unterrichts und der Verletzung der Aufsichtspflicht verbunden waren, ist die ausgebliebene Korrelation erklärbar. Die weitestgehend hypothesenkonformen Ergebnisse unterstreichen die Bedeutung der ausgewählten Konstrukte für den Umgang mit digitalen Problemsituationen und verdeutlichen darüber hinaus, dass durch das neu definierte Konstrukt und dessen beschriebene Operationali-

sierung weitere Aspekte im Bereich von TPK(-Klassenführung) praxisnah erfasst werden konnten.

Kriteriumsvalidität

Zur Überprüfung der Kriteriumsvalidität wurde mittels eines Strukturgleichungsmodells der Einfluss des Umgangs mit digitalen Problemsituationen auf die Häufigkeit der wahrgenommenen Probleme beim digitalen Medieneinsatz der Lehrkräfte erfasst. Die Hinzunahme des entwickelten Situational Judgement Test als zusätzlichen Prädiktor neben der Selbstwirksamkeit im Umgang mit digitalen Problemsituationen zeigt eine signifikante, geringe inkrementelle Varianzaufklärung. Da sowohl die Selbstwirksamkeitserwartung im Umgang mit digitalen Problemsituationen als auch die berichteten Problemsituationen auf Selbsteinschätzungen basieren und daher konfundiert sein werden, ist die geringe zusätzliche Varianzaufklärung durch das entwickelte Testverfahren bedeutsam. Lehrkräfte, die eine hohe Kompetenz im Umgang mit digitalen Problemsituationen haben, erleben technische Defekte anscheinend weniger als Problemsituation, da sie funktionale Strategien zur Bewältigung haben. Dieser Befund verdeutlicht die Wichtigkeit eines funktionalen Umgangs mit Problemsituationen im Umgang mit digitalen Medien: Strategien zur funktionalen Bewältigung können die Wahrnehmung einer solchen Situation beeinflussen und damit potentiell negative Einflüsse auf Volition und Motivation der digitalen Mediennutzung durch Lehrkräfte reduzieren.

Limitationen und Ausblick

Bei den beschriebenen Ergebnissen und Schlussfolgerungen ist einschränkend festzuhalten, dass die Überprüfung der praktischen Relevanz des Testverfahrens noch aussteht. Hierfür sollten weitere Außenkriterien zur Validierung herangezogen werden (Hartig et al., 2008). Dies könnte beispielsweise in experimentellen Settings realisiert werden, in denen Lehrpersonen in einer Unterrichtseinheit vor Problemsituationen gestellt werden, oder durch Schüler_inneneinschätzungen, um den Umgang mit digitalen Problemsituationen und die Korrelation zum entwickelten Instrument empirisch zu überprüfen. In weiteren Untersuchungen sollte zudem der Einfluss des Umgangs mit digitalen Problemsituationen auf den tatsächlichen Einsatz digitaler Medien und Variablen, die diesen begünstigen, überprüft werden (wie beim *Technology Acceptance Model*; Scherer et al., 2019). Das Instrument erfasst zudem nur Problemsituationen, die durch unverschul-

dete, technische Defekte entstanden sind. Um das Konstrukt umfassender mit der Facette des Umgangs mit selbst verschuldeten digitalen Fehlersituationen abzubilden, könnte der Test in computergestützten, dynamischen Testumgebungen, wie beispielsweise bei der *International Computer and Information Literacy Study* (ICILS; Eickelmann et al., 2019), umgesetzt werden.

In künftigen Untersuchungen bietet es sich zudem an, die psychometrische Prüfung des Instruments zum Umgang mit digitalen Problemsituationen von Lehrkräften im Unterricht auch bei unterschiedlichen Schulstufen differenziert zu betrachten. Es ist davon auszugehen, dass beispielsweise die Arbeit mit jüngeren Schüler_innen (Primarstufe) auch Auswirkungen auf die Wahl der Bewältigungsstrategie hat (bspw. direkt zum Support gehen, was das Alleinlassen der Schüler_innen impliziert). Um die Ergebnisse des entwickelten Testverfahrens zu bestätigen, sollte die Validität daher auch bei Lehrkräften weiterer Schulformen verifiziert werden.

Weiterhin sollte überprüft werden, inwieweit sich die im wirtschaftlichen Kontext nachgewiesenen Befunde zu Vorteilen von Fortbildungen zum Umgang gegenüber der Vermeidung von Fehler- und Problemsituationen auf den Lehrberuf übertragen lassen (Keith & Frese, 2008). Hierfür könnten in Folgestudien beispielsweise die Verbindung mit Basisdimensionen guten Unterrichts (Klieme, Funke, Leutner, Reimann & Wirth, 2001) sowie Lernerfolg der Schüler_innen und UDP durch das entwickelte Testinstrument betrachtet werden. Sollten sich positive Effekte nachweisen lassen, könnte der Umgang mit digitalen Problemsituationen Eingang in Lehrkräftefortbildungen finden. Lehrkräfte könnten beispielsweise in Schulungsmaßnahmen die von den Expert_innen in der Testentwicklung als effektiv gekennzeichneten Handlungsstrategien und Heuristiken für ihre Unterrichtspraxis vermittelt bekommen. Der entwickelte SJT könnte hier ebenfalls zur Überprüfung der eigenen Strategien in solchen Situationen und als Ausgangspunkt für Diskussionen über mögliche Handlungsalternativen eingesetzt werden. Mit der Berücksichtigung des Umgangs mit digitalen Problemsituationen in solchen Veranstaltungen würden auch Forderungen aktueller Studien zur kontinuierlichen Weiterentwicklung von Fortbildungen im Bereich digitaler Medien adressiert werden (z. B. Eickelmann et al., 2019; Schulze-Vorberg, Krille, Fabriz & Horz, 2021).

Der Einsatz von Selbsteinschätzungen wird kritisch diskutiert, da Studien zeigen konnten, dass zwischen selbst eingeschätzten und objektiven Testwerten im Bereich von TPACK nur geringe Zusammenhänge zu finden sind (z. B. Drummond & Sweeney, 2017). Im Bereich digitaler Kompetenzen sind fortlaufende Studien notwendig, um einen Überblick über Wissens- und Kompetenzstände von Lehrkräften zu geben und Ansatzpunkte für die Lehrkräf-

teaus- und forbildung abzuleiten. Um ein realistisches Abbild zu erhalten, bedarf es allerdings Messinstrumente, die über fehlerbehaftete Selbsteinschätzungsskalen hinausgehen. Das entwickelte Testverfahren adressiert diese Problematik und bietet die Möglichkeit, Studien zu Medienkompetenzen von Lehrkräften mit praxisnahen Daten anzureichern.

Die erfolgreiche Operationalisierung in Form eines Situational Judgement Test zeigt auch Möglichkeiten für die Erfassung von weiteren Facetten der Medienkompetenz mit dieser Methode auf. Damit adressiert dieser Ansatz zur Messung von digitalen Wissensaspekten im Gegensatz zu Ansätzen, die auf Selbsteinschätzungen basieren, die Forderung nach methodischer Qualität in diesem hochrelevanten und aktuellen Forschungsfeld (Scheiter & Lachner, 2019).

Elektronische Supplemente (ESM)

Die elektronischen Supplemente sind mit der Online-Version dieses Artikels verfügbar unter <https://doi.org/10.1026/0012-1924/a000304>

ESM 1. Tabelle E1: Übersicht über Expert_innenbefragungsergebnisse und empirische Trennschärfe aus Vorerhebung.

ESM 2. Tabelle E2: Deskriptivstatistik der eingesetzten Instrumente sowie bivariate Korrelation mit dem Instrument zur Erfassung des Umgangs mit digitalen Problemsituationen.

ESM 3. Tabelle E3: Standardisierte Faktorladungen und Standardfehler aller Aufgaben des Instruments zur Erfassung des Umgangs mit digitalen Problemsituationen.

ESM 4. Beispielitem aus dem Instrument zur Erfassung des Umgangs mit digitalen Problemsituationen.

Literatur

- Asparouhov, T. & Muthén, B. (2006). Robust chi square difference testing with mean and variance adjusted test statistics. *Mplus Web Notes: No. 10*. Verfügbar unter: <https://www.statmodel.com/download/webnotes/webnote10.pdf>
- Baier, F. & Kunter, M. (2020). Construction and validation of a test to assess (pre-service) teachers' technological pedagogical knowledge (TPK). *Studies in Educational Evaluation*, 67, 100936. <https://doi.org/10.1016/j.stueduc.2020.100936>
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191–215. <https://doi.org/10.1037/0033-295X.84.2.191>
- Baumstarck, K., Alessandrini, M., Hamidou, Z., Auquier, P., Leroy, T. & Boyer, L. (2017). Assessment of coping: A new french four-factor structure of the brief COPE inventory. *Health and quality*

- of life outcomes, 15(1), 8. <https://doi.org/10.1186/s12955-016-0581-9>
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. Guilford Press. Retrieved from <http://www.loc.gov/catdir/enhancements/fy0629/2006001103-b.html>
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson Studium.
- Bundesrepublik Deutschland. (2019). *Verwaltungsvereinbarung DigitalPakt Schule 2019 bis 2024*. Verfügbar unter: https://www.bmbf.de/files/VV_DigitalPaktSchule_Web.pdf
- Campion, M. C., Ployhart, R. E. & MacKenzie, W. I. (2014). The state of research on situational judgment tests: A content analysis and directions for future research. *Human Performance*, 27, 283–310. <https://doi.org/10.1080/08959285.2014.929693>
- Carver, C. S. (1997). You want to measure coping but your protocol's too long: Consider the brief COPE. *International Journal of Behavioral Medicine*, 4(1), 92–100. https://doi.org/10.1207/s15327558ijbm0401_6
- Chai, C. S., Koh, J. H.-L. & Tsai, C.-C. (2013). A review of technological pedagogical content knowledge. *Educational Technology & Society*, 16(2), 31–51.
- Charters, E. (2003). The use of think-aloud methods in qualitative research an introduction to think-aloud methods. *Brock Education Journal*, 12(2), 68–82. <https://doi.org/10.26522/bricked.v12i2.38>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale: Erlbaum.
- De Jong, T. & Ferguson-Hessler, M. G. M. (1996). Types and qualities of knowledge. *Educational Psychologist*, 31(2), 105–113. https://doi.org/10.1207/s15326985ep3102_2
- Dörner, D. (1979). *Problemlösen als Informationsverarbeitung* (2. Aufl., Kohlhammer-Standards Psychologie Studententext). Stuttgart: Kohlhammer.
- Doyle, W. (2006). Ecological approaches to classroom management. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management: Research, practice, and contemporary issues* (pp. 97–125). Routledge.
- Drummond, A. & Sweeney, T. (2017). Can an objective measure of technological pedagogical content knowledge (TPACK) supplement existing TPACK measures? *British Journal of Educational Technology*, 48, 928–939. <https://doi.org/10.1111/bjet.12473>
- Eickelmann, B., Bos, W., Gerick, J., Goldhammer, F., Schaumburg, H., Schwippert, K. et al. (Hrsg.). (2019). *ICILS 2018 #Deutschland: Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im zweiten internationalen Vergleich und Kompetenzen im Bereich Computational Thinking*. Münster: Waxmann. Verfügbar unter: https://kw.uni-paderborn.de/fileadmin/fakultaet/Institute/erziehungswissenschaft/Schulpaedagogik/ICILS_2018__Deutschland_Berichtsband.pdf
- Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin*, 51, 327–358. <https://doi.org/10.1037/h0061470>
- Gold, B. & Holodynski, M. (2015). Development and construct validation of a situational judgment test of strategic knowledge of classroom management in elementary schools. *Educational Assessment*, 20, 226–248. <https://doi.org/10.1080/10627197.2015.1062087>
- Hartig, J., Frey, A. & Jude, N. (2008). Validität. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion: Mit 43 Tabellen* (Springer-Lehrbuch, S. 135–163). Berlin: Springer. https://doi.org/10.1007/978-3-540-71635-8_7
- Hartig, J. & Klieme, E. (2006). Kompetenz und Kompetenzdiagnostik. In K. Schweizer (Hrsg.), *Leistung und Leistungsdiagnostik* (S. 127–143). Berlin: Springer. https://doi.org/10.1007/3-540-33020-8_9
- Hillmayr, D., Reinhold, F., Ziernwald, L. & Reiss, K. (2017). *Digitale Medien im mathematisch-naturwissenschaftlichen Unterricht der Sekundarstufe: Einsatzmöglichkeiten, Umsetzung und Wirksamkeit*. Münster: Waxmann. Verfügbar unter: <https://www.waxmann.com/?elD=texte&pdf=3766Volltext.pdf&typ=zusatztext>
- Horz, H. & Schulze-Vorberg, L. (2017). Digitalisierung in der Hochschullehre. *Analysen & Argumente*, 283, 57–71. Verfügbar unter: http://www.kas.de/wf/doc/kas_50782-544-1-30.pdf?171123080940
- Keith, N. & Frese, M. (2008). Effectiveness of error management training: a meta-analysis. *The Journal of Applied Psychology*, 93(1), 59–69. <https://doi.org/10.1037/0021-9010.93.1.59>
- 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. <https://doi.org/10.25656/01:5272>
- Knoll, N., Rieckmann, N. & Schwarzer, R. (2005). Coping as a mediator between personality and stress outcomes: a longitudinal study with cataract surgery patients. *European Journal of Personality*, 19, 229–247. <https://doi.org/10.1002/per.546>
- Koo, T. K. & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15, 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Kounin, J. S. (1976). *Techniken der Klassenführung* (Abhandlungen zur pädagogischen Psychologie, Neue Folge, Bd. 6.) Bern: Huber.
- Kultusministerkonferenz. (2016). *Strategie der Kultusministerkonferenz „Bildung in der digitalen Welt“*. Verfügbar unter: https://www.kmk.org/fileadmin/Dateien/pdf/PresseUndAktuelles/2017/Strategie_neu_2017_datum_1.pdf
- Kunter, M. & Klusmann, U. (2010). Kompetenzmessung bei Lehrkräften – Methodische Herausforderungen. *Unterrichtswissenschaft*, 38(1), 68–86.
- Lachner, A., Backfisch, I. & Stürmer, K. (2019). A test-based approach of Modeling and Measuring Technological Pedagogical Knowledge. *Computers & Education*, 142, 103645. <https://doi.org/10.1016/j.compedu.2019.103645>
- Lazarus, R. S. & Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer.
- Li, C.-H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, 48, 936–949. <https://doi.org/10.3758/s13428-015-0619-7>
- Marquard, O. (2005). *Abschied vom Prinzipiellen: Philosophische Studien* (Reclams Universal-Bibliothek, Bd. 7724). Stuttgart: Reclam.
- McDaniel, M. A., Psotka, J., Legree, P. J., Yost, A. P. & Weekley, J. A. (2011). Toward an understanding of situational judgment item validity and group differences. *The Journal of Applied Psychology*, 96, 327–336. <https://doi.org/10.1037/a0021983>
- Mishra, P. & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108, 1017–1054. <https://doi.org/10.1111/j.1467-9620.2006.00684.x>
- Mußmann, F. & Hardwig, T. (2021, 1. Juni). *Digitalisierung im Schulsystem Herausforderung für Arbeitszeit und Arbeitsbelastung von Lehrkräften*. Vortrag auf der Pressekonferenz zur Vorstellung der Studienergebnisse, Göttingen & Frankfurt. Verfügbar unter: https://www.gew.de/fileadmin/media/sonstige_downloads/hv/Service/Presse/2021/Digitalisierung-im-Schulsystem—Studie.pdf
- Muthén, L. K. & Muthén, B. O. (1998–2010). *Mplus User's Guide* (6th ed.). Los Angeles, CA: Muthén & Muthén.
- Naumann, J., Goldhammer, F., Rölke, H. & Stelter, A. (2014). Erfolgreiches Problemlösen in technologiebasierten Umgebungen: Wechselwirkungen zwischen Interaktionsschritten und Aufgabenanforderungen. *Zeitschrift für Pädagogische Psychologie*, 28, 193–203. <https://doi.org/10.1024/1010-0652/a000134>

- Questback GmbH. (2017). *EFS Survey* (Version Summer 2017) [Computer software]. Köln: Questback GmbH.
- Romi, S., Lewis, R. & Roache, J. (2013). Classroom management and teachers' coping strategies: Inside classrooms in Australia, China and Israel. *PROSPECTS*, 43(2), 215–231. <https://doi.org/10.1007/s11125-013-9271-0>
- Rosenberg, J. M. & Koehler, M. J. (2015). Context and Technological Pedagogical Content Knowledge (TPACK): A Systematic Review. *Journal of Research on Technology in Education*, 47, 186–210. <https://doi.org/10.1080/15391523.2015.1052663>
- Saritepeci, M. (2021). Modelling the effect of TPACK and computational thinking on classroom management in technology enriched courses. *Technology, Knowledge and Learning*, 1–15. <https://doi.org/10.1007/s10758-021-09529-y>
- Scheiter, K. & Lachner, A. (2019). DigitalPakt – was nun? Eine Positionierung aus Sicht der Lehr-Lernforschung. *Unterrichtswissenschaft*, 47, 547–564. <https://doi.org/10.1007/s42010-019-00059-2>
- Scherer, R., Siddiq, F. & Tondeur, J. (2019). The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. *Computers & Education*, 128, 13–35. <https://doi.org/10.1016/j.compedu.2018.09.009>
- Schmid, U., Goertz, L. & Behrens, J. (2017). *Monitor Digitale Bildung: Die Schulen im digitalen Zeitalter*. Gütersloh: Bertelsmann Stiftung. Verfügbar unter: https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/BSt_MD_B3_Schulen_web.pdf
- Schmidt, D. A., Baran, E., Thompson, A. D., Mishra, P., Koehler, M. J. & Shin, T. S. (2009). Technological Pedagogical Content Knowledge (TPACK). *Journal of Research on Technology in Education*, 42, 123–149. <https://doi.org/10.1080/15391523.2009.10782544>
- Schulze-Vorberg, L., Krille, C., Fabriz, S. & Horz, H. (2021). Hinweise und Empfehlungen für die Konzeption von Lehrkräftefortbildungen zu digitalen Medien. *Zeitschrift für Erziehungswissenschaft*, 24, 1113–1142. <https://doi.org/10.1007/s11618-021-01046-z>
- Schwarzer, R. & Jerusalem, M. (Hrsg.). (1999). *Skalen zur Erfassung von Lehrer- und Schülermerkmalen: Dokumentation der psychometrischen Verfahren im Rahmen der wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen*. Berlin: R. Schwarzer.
- Shrout, P. E. & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86, 420–428. <https://doi.org/10.1037/0033-2909.86.2.420>
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14.
- Stemler, S. E. & Sternberg, R. J. (2006). Using situational judgment tests to measure practical intelligence. In J. A. Weekley & R. E. Ployhart (Eds.), *Situational judgment tests: Theory, measurement, and application* (pp. 107–131). Lawrence Erlbaum Associates Publishers.
- Sung, Y.-T., Chang, K.-E. & Liu, T.-C. (2016). The effects of integrating mobile devices with teaching and learning on students' learning performance: A meta-analysis and research synthesis. *Computers & Education*, 94, 252–275. <https://doi.org/10.1016/j.compedu.2015.11.008>
- Voss, T., Kunter, M. & Baumert, J. (2011). Assessing teacher candidates' general pedagogical/psychological knowledge: Test construction and validation. *Journal of Educational Psychology*, 103, 952–969. <https://doi.org/10.1037/a0025125>
- Weng, Q., Yang, H., Lievens, F. & McDaniel, M. A. (2018). Optimizing the validity of situational judgment tests: The importance of scoring methods. *Journal of Vocational Behavior*, 104, 199–209. <https://doi.org/10.1016/j.jvb.2017.11.005>
- Zell, E. & Krizan, Z. (2014). Do people have insight into their abilities? A metasynthesis. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, 9(2), 111–125. <https://doi.org/10.1177/1745691613518075>

Historie

Onlineveröffentlichung: 01.11.2022

Förderung

Open Access-Veröffentlichung ermöglicht durch die Universitätsbibliothek Johann Christian Senckenberg.

ORCID

Lukas Schulze-Vorberg

 <https://orcid.org/0000-0003-2443-990X>

Dr. Lukas Schulze-Vorberg

Institut für Psychologie
Arbeitsbereich Pädagogische Psychologie
Goethe-Universität Frankfurt am Main
Theodor-W.-Adorno-Platz 6
60629 Frankfurt am Main
Deutschland
schulze-vorberg@psych.uni-frankfurt.de