



Transferwissen von Lehrenden in der berufsbezogenen Weiterbildung: Entwicklung und Pilotierung eines Testinstruments

Alisha Koch  · Susanne Wißhak  · Claudio Spener  ·
Alexander Naumann  · Sabine Hochholdinginger 

Eingegangen: 15. Dezember 2021 / Überarbeitet: 18. März 2022 / Angenommen: 25. April 2022 / Online
publiziert: 15. Mai 2022
© Der/die Autor(en) 2022

Zusammenfassung Transfer gilt als zentrales Ziel berufsbezogener Weiterbildungsmaßnahmen. Obwohl Lehrende in der berufsbezogenen Weiterbildung Transfer anbahnen und unterstützen können, ist über ihr dafür benötigtes professionelles Wissen bislang wenig bekannt, und es fehlt an einem Messinstrument zu dessen Erfassung. Zu diesem Zweck entwickelten wir einen Test, welcher wissenschaftlich fundiertes Wissen über die Determinanten und die Förderung von Transfer erfasst. Der vorliegende Beitrag thematisiert die Entwicklung und Pilotierung des Instruments. Auf Basis der Item-Response-Theory (IRT) wurden die Testantworten von $N=105$ Lehrenden skaliert und interpretiert. Die 32 Items zeigen eine gute Passung hinsichtlich des eindimensionalen Partial-Credit-Modells mit Infit-Werten zwischen 0,93 und 1,16. Die EAP/PV-Reliabilität liegt bei 0,81. Die Ergebnisse weisen darauf hin, dass sich Transferwissen mithilfe des Instruments operationalisieren und reliabel messen lässt.

Schlüsselwörter Lehrende in der Weiterbildung · Trainingstransfer · Testentwicklung · IRT-Skalierung · Partial-Credit-Modell · Validierung

Alisha Koch (✉) · Susanne Wißhak · Claudio Spener · Sabine Hochholdinginger
Universität Konstanz, Konstanz, Deutschland
E-Mail: alisha.koch@uni-konstanz.de

Alexander Naumann
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Frankfurt a.M., Deutschland

Transfer knowledge of trainers in continuing vocational education and training: Construction and piloting of a test instrument

Abstract The goal of work-related training is transfer. Although trainers in work-related continuing education can initiate and support transfer, little is known about the professional knowledge they need to do so, and there is a lack of a measurement tool to capture this knowledge. For this reason, we constructed an instrument as a means to assessing scientifically based knowledge about the determinants of transfer and about its facilitation. This paper reports the test construction and the piloting of the instrument. Based on item response theory (IRT), the answers of $N=105$ trainers were scaled and interpreted. The 32 items showed a good fit with respect to the one-dimensional partial credit model with infit values ranging from 0.93 to 1.16. The EAP/PV reliability was 0.81. The results indicate that transfer knowledge can be operationalized and reliably measured using the instrument.

Keywords Teachers in continuing education · Transfer of training · Test development · IRT-scaling · Partial credit model · Validation

1 Einleitung

Lebenslanges Lernen soll Erwachsene nicht nur dazu befähigen, auf aktuelle Transformationen wie etwa die Digitalisierung zu reagieren, sondern diese aktiv und selbstbestimmt mitzugestalten (Autorengruppe Bildungsberichterstattung 2020). In diesem Kontext ist insbesondere auch die berufsbezogene Weiterbildung von Bedeutung (Kuhnhenne 2020).

Ein wichtiges Ziel berufsbezogener Weiterbildungsmaßnahmen ist der Transfer, der als langfristige Anwendung des Gelernten am Arbeitsplatz definiert wird (Hense und Mandl 2011). Verschiedene Befunde weisen darauf hin, dass das erfolgreiche Aneignen von Lerninhalten noch nicht dazu führt, dass das Gelernte später auch verwendet wird (Alliger et al. 1997). Ein fortwährendes Problem besteht darin, dass das Gelernte keinen Eingang in den Arbeitsalltag der Teilnehmenden findet und der Transfer missglückt (Kanning 2014). Die internationale Trainingstransferforschung fokussiert sich daher seit Jahren auf die Erforschung der Determinanten von Transfer (Salas et al. 2012). Dabei liegt der Fokus jedoch bisher kaum auf den Unterstützungsmöglichkeiten durch die Lehrenden, weswegen vorgeschlagen wurde, den Blick zukünftig stärker auf diese zu richten (Baldwin et al. 2017; Wißhak 2022).

Die professionelle Handlungskompetenz von Lehrenden gilt als wesentlicher Einflussfaktor für gelingende Lehr-/Lernprozesse und den Lernerfolg der Teilnehmenden (Strauch et al. 2020). Studien weisen zudem darauf hin, dass Lehrende auch den Transfererfolg ihrer Teilnehmenden beeinflussen können (z. B. Russ-Eft et al. 2010; Donovan und Darcy 2011). Dennoch ist über ihr professionelles Wissen über Weiterbildungstransfer bislang wenig bekannt. Fraglich ist, ob die Lehrenden, die sich ihr pädagogisches Wissen häufig informell oder non-formal aneignen (Strauch et al. 2020), mit den Befunden der Trainingstransferforschung vertraut sind, oder

ob hierbei von einer Kluft zwischen Forschung und Praxis (Hutchins et al. 2010) auszugehen ist.

Vor dem Hintergrund dieser Forschungslücke beschreibt der vorliegende Beitrag die Entwicklung und Pilotierung eines Testinstruments, welches das Wissen von Lehrpersonen aus der berufsbezogenen Weiterbildung über die Determinanten und Fördermöglichkeiten von Transfer erfassen soll. Im Beitrag stellen wir die Testentwicklung und erste Validierungsstudien, die Aufgabenkonstruktion und die Skalierung vor. Langfristig sollen aus den Befunden Handlungsempfehlungen für die Professionalisierung der Lehrenden in der berufsbezogenen Weiterbildung abgeleitet werden.

2 Theoretischer Rahmen und Forschungslage

2.1 Transfer als zentrales Ziel berufsbezogener Weiterbildung

Berufsbezogene Weiterbildungsmaßnahmen finden meist in Form sogenannter Trainings statt, und die Lehrenden werden auch als Trainerinnen und Trainer bezeichnet (Fuchs 2011). Im Rahmen von Trainings sollen sich die Teilnehmenden Wissen, Fähigkeiten oder Einstellungen aneignen, die sie für die Bewältigung beruflicher Herausforderungen und die Mitgestaltung o. g. Transformationen benötigen (Salas et al. 2012). Damit der Transfer gelingt, muss das Gelernte generalisiert, also auf neue Situationen übertragen und über einen längeren Zeitraum am Arbeitsplatz angewendet werden (Baldwin und Ford 1988).

Es existieren bereits einige Modelle, die versuchen, Determinanten von Transfer abzubilden und damit Ansätze zur Erklärung und Förderung von Transfer zu liefern (Baldwin und Ford 1988; Holton et al. 2000; Burke und Hutchins 2008; Blume et al. 2019; Sandmeier et al. 2021). Das wohl bekannteste Rahmenmodell liefern Baldwin und Ford (1988), welche zwischen den drei Einflussfaktoren *Teilnehmendenmerkmale*, *Weiterbildungsdesign* und *Arbeitsumfeld* differenzieren, die jeweils direkt oder indirekt auf den Transfer wirken. Zu den Teilnehmendenmerkmalen zählen sie die Fähigkeiten der Teilnehmenden sowie deren Persönlichkeitsmerkmale und Motivation. Unter das Weiterbildungsdesign fallen die Lernprinzipien und -methoden. Zum Arbeitsumfeld zählen die Unterstützung am Arbeitsplatz, die die Teilnehmenden im Hinblick auf die Transferförderung erfahren, sowie die Anwendungsmöglichkeiten der Weiterbildungsinhalte nach dem Training. Das Modell diene als Grundlage für eine Vielzahl empirischer Untersuchungen zu den Determinanten von Weiterbildungstransfer (Wißhak 2022). Obwohl das Modell von Baldwin und Ford (1988) wichtige und empirisch bestätigte Transferdeterminanten enthält, wurden Vorschläge zu seiner Weiterentwicklung gemacht. So wird in dem erweiterten Modell von Burke und Hutchins (2008) zusätzlich die Relevanz der Lehrperson betont. Darüber hinaus berücksichtigt es die übrigen beteiligten Akteure wie Peers und Führungskräfte, sowie die zeitlichen Phasen vor, während und nach einem Training.

Auch Baldwin et al. (2017) weisen darauf hin, dass die Lehrenden wahrscheinlich eine größere Rolle spielen als ihre mangelnde Berücksichtigung in der Transferforschung vermuten lässt. Wie genau die Lehrpersonen Transfer fördern sollen, bleibt

jedoch weitestgehend unklar. Studien, welche sich direkt auf den Zusammenhang zwischen den Kompetenzen der Lehrenden und Transfer beziehen, gibt es bisher nur vereinzelt. Dennoch weisen diese darauf hin, dass Trainerinnen und Trainer durch ihr professionelles Handeln den Erfolg des Trainings grundsätzlich mitbestimmen (Russ-Eft et al. 2010; Donovan und Darcy 2011; Rasli et al. 2012). Zudem wird in diversen Studien die Trainerkompetenz indirekt über das transferförderliche Weiterbildungsdesign erhoben (Grohmann et al. 2014).

2.2 Transferwissen als Teil der professionellen Handlungskompetenz von Lehrenden in der berufsbezogenen Weiterbildung

Transferwissen wird in diesem Beitrag definiert als das evidenzbasierte, also empirisch fundierte Wissen über die Determinanten und Fördermöglichkeiten von Transfer. Trotz seiner Relevanz ist das Transferwissen der Lehrenden bisher weitgehend unerforscht, was damit zusammenhängen könnte, dass es sich bei der Lehre in der Weiterbildung um keinen formalen Beruf handelt und zumindest in Deutschland die Qualifikationen und Anforderungen für die Tätigkeit nicht reglementiert und sehr heterogen sind (Bonnes und Hochholding 2016; Wißhak et al. 2020). Entsprechend wenig Forschung existiert zu ihrem Professionswissen. Dieses gilt jedoch als wesentlicher Faktor der *professionellen Handlungskompetenz* von Lehrenden, zu welcher auch die *Überzeugungen, Werthaltungen, motivationalen Orientierungen* und *selbstregulierenden Fähigkeiten* zählen (Baumert und Kunter 2006, S. 482). Obwohl solche Kompetenzdimensionen bei der Transferförderung u.U. auch eine Rolle spielen, ist im Hinblick auf das in diesem Beitrag fokussierte Transferwissen vor allem das Professionswissen relevant. In Anlehnung an Shulman (1986) wird das Professionswissen von Baumert und Kunter (2006, S. 482) in die Bereiche *Fachwissen, fachdidaktisches Wissen, Organisationswissen, Beratungswissen* und *generisches pädagogisch-psychologisches Wissen* unterteilt. Letzteres wird definiert „als Wissen zur erfolgreichen Gestaltung und Optimierung von Lehr-/Lernsituationen“ (Marx et al. 2017, S. 167) und scheint im Hinblick auf den Trainingserfolg von besonders hoher Relevanz zu sein.

Mit dem *GRETA-Kompetenzmodell* wurde darüber hinaus vom Deutschen Institut für Erwachsenenbildung (DIE) ein Modell professioneller Handlungskompetenz für Lehrende in der Erwachsenen- und Weiterbildung entwickelt, welches in Anlehnung an Baumert und Kunter (2006, S. 7) die vier Kompetenzaspekte *fach- und feldspezifisches Wissen* (z. B. *Wissen über Adressaten*), *berufspraktisches Wissen und Können*, *professionelle Werthaltungen und Überzeugungen* und *professionelle Selbststeuerung* (Strauch et al. 2019) umfasst. Besonders im Rahmen des Kompetenzaspekts *berufspraktisches Wissen und Können* werden einige Facetten aufgeführt, die den Transfer fördern können. Beispielhaft zu nennen wären die genaue Abstimmung der Lehr-/Lernmethoden auf die Weiterbildungsziele oder die Teilnehmendenorientierung (Wißhak 2022). Dennoch wird Transfer als Ziel von Weiterbildung nicht explizit berücksichtigt. So bezieht sich die Facette *Outcomeorientierung* lediglich auf die Lernergebnisse.

2.3 Die Lücke zwischen Forschung und Praxis und ein erstes Instrument

Trotz der vielfältigen Einflussmöglichkeiten durch die Lehrenden gibt es bisher kaum Untersuchungen darüber, ob sie die Forschungsergebnisse zu den Determinanten von Weiterbildungstransfer kennen und ob sie wissen, wie sie Transfer fördern können. Studienergebnisse von Wißhak und Hochholding (2018) weisen darauf hin, dass Lehrende Transferförderung zwar als wichtig für ihren Beruf einordnen. Gleichzeitig empfanden sie es jedoch als weniger wichtig, empirische Weiterbildungsforschung zu kennen. Im Rahmen der wbmonitor Umfrage zum Thema Wissenstransfer gaben 64 % der befragten Weiterbildungsanbieter an, einen generellen Bedarf an Forschungsergebnissen zu haben. Gleichzeitig bejahten 47 % der Befragten aber, ihre praktischen Arbeitsanforderungen auch ohne wissenschaftliche Erkenntnisse ausführen zu können (Christ et al. 2019). Diese Befunde decken sich mit den Ergebnissen von Hutchins et al. (2010). Hier gaben die Lehrenden an, sich ihr Wissen über Weiterbildungstransfer weniger über wissenschaftliche Fachzeitschriften anzueignen, sondern vielmehr über die Berufserfahrung, den persönlichen Austausch mit Kolleginnen und Kollegen sowie durch Bücher und das Internet. Die Autorinnen sprechen dabei von einem „*research-practice gap*“ womit gemeint ist, dass die Befunde aus der Wissenschaft nur unzureichend in der Praxis wahrgenommen werden (Hutchins et al. 2010, S. 606). Mögliche Gründe dafür könnten Zeitmangel sowie der fehlende Zugang zu den Fachartikeln sein. Der *research-practice gap* ist besonders dahingehend problematisch, als die Forderung nach langfristig wirksamer Weiterbildung immer lauter wird (Holton 2004; Hutchins et al. 2010; Hetfleisch et al. 2017). Um dieses Problem angehen zu können, wird zunächst ein geeignetes Messinstrument für das Transferwissen Lehrender benötigt.

Für die Erfassung des pädagogisch-psychologischen Wissens von Lehrenden in Schule und Weiterbildung wurden bereits mehrere Instrumente entwickelt und erprobt (König und Blömeke 2009; Voss et al. 2011; Marx et al. 2017). Das Transferwissen wird jedoch in keinem dieser Instrumente fokussiert. Der bislang einzige Versuch, das Transferwissen von Trainerinnen und Trainern zu erfassen, stammt von Hutchins und Burke (2007). Dafür konstruierten die Autorinnen auf Basis eines Literaturreviews 32 Items, die jeweils eine Aussage zu Transferdeterminanten enthalten. Die Befragten, 139 Mitglieder eines US-amerikanischen Weiterbildungsverbands, bewerteten jedes Item als *richtig*, *falsch* oder *unsicher*. Insgesamt wurden 78 % der Items *richtig* beantwortet. Zudem korrelierte das erfasste Transferwissen signifikant positiv mit dem Bildungsniveau der Befragten, der hierarchischen Position im Unternehmen und dem Besitz eines Trainerzertifikates. Die Befunde legen weiterhin nahe, dass das Wissen der Befragten über das Weiterbildungsdesign und das Arbeitsumfeld der Lernenden ausgeprägter war als das Wissen über Teilnehmendenmerkmale und Trainingsevaluation. Aufgrund des Aufgabenformats kann jedoch von einer hohen Ratewahrscheinlichkeit ausgegangen werden.

Der hier vorgestellte Forschungsstand bildet den Ausgangspunkt für die Entwicklung eines Messinstruments für Transferwissen. Der vorliegende Beitrag fokussiert die Testentwicklung und -pilotierung und liefert erste Hinweise darauf, inwiefern eine valide Interpretation der Testergebnisse möglich ist.

3 Methode

Für die Testentwicklung wurden zunächst verschiedene qualitative Validierungsschritte vorgenommen. Im Rahmen der Pilotstudie wurden darüber hinaus die Testantworten von 105 Lehrenden auf Basis der Item-Response-Theory (IRT) skaliert und im Hinblick auf die Optimierung des Instruments interpretiert.

3.1 Testentwicklung und erste Validierungsschritte

Bei dem entwickelten Instrument handelt es sich um einen psychometrischen Test (Döring und Bortz 2016), welcher Transferwissen erfassen soll. Die konzeptuelle Grundlage für das latente Konstrukt Transferwissen lieferten zum einen die Transfermodelle von Baldwin und Ford (1988) sowie Burke und Hutchins (2008). Zum anderen basiert es auf einem systematischen Literaturreview, welches belastbare empirische Befunde aus 19 Metaanalysen zu Transferdeterminanten zusammenfasste und im Hinblick auf die Transferförderung durch Lehrende auswertete (Wißhak 2022). Das Literaturreview orientierte sich an den methodischen Richtlinien des PRISMA-Statement (Liberati et al. 2009). Die statistisch bedeutsamen Befunde der untersuchten Metaanalysen, die als für Lehrpersonen relevant eingeschätzt wurden, wurden zunächst extrahiert und in die drei Kategorien nach Baldwin und Ford (1988), *Teilnehmendenmerkmale*, *Weiterbildungsdesign* und *Arbeitsumfeld* eingeordnet. In einer anschließenden Synthese wurden Handlungsmöglichkeiten der Lehrenden abgeleitet. Aus den Befunden des Literaturreviews ergeben sich beobachtbare Ausprägungen des latenten Konstrukts Transferwissen. So sind sich Lehrende mit einer hohen Ausprägung zunächst darüber im Klaren, dass Transfer vor, während und nach einer Maßnahme gefördert werden muss. Transfer gelingt beispielsweise eher, wenn im Vorfeld eine Bedarfsanalyse durchgeführt wird, um die Maßnahme auf die Lernenden zuzuschneiden. Darüber hinaus sollten Lehrende wissen, welche *Teilnehmendenmerkmale* sich positiv oder negativ auf Transfer auswirken können. So korrelieren die Motivation und Selbstwirksamkeit deutlich mit Transfer und sollten daher nach Möglichkeit gefördert werden (Gegenfurtner 2011; Gegenfurtner et al. 2013). Im Hinblick auf das *Weiterbildungsdesign* sollten Lehrende verschiedene transferförderliche Trainingsformen und -methoden kennen und zielgerichtet einsetzen können. Auch generelle transferförderliche Prinzipien der Trainingsgestaltung sollten ihnen bekannt sein, wie Feedback, das sogenannte Overlearning, also das wiederholte Üben oder die transferförderliche Sequenzierung von Lehr-/Lernangeboten. Zudem sollten Lehrende sich der wichtigen Rolle bewusst sein, die auch das *Arbeitsumfeld* der Lernenden spielt. Sie sollten wissen, wie sie die wichtigsten Akteure wie Kolleginnen und Kollegen sowie Führungskräfte einbinden und bereits im Training Unterstützungsmechanismen anbahnen können. Schließlich sollten die Lehrenden wissen, dass Transfer dadurch wahrscheinlicher wird, dass er evaluiert und gewürdigt wird (Wißhak 2022).

Die Basis für die Aufgabenkonstruktion lieferten die Befunde des o. g. Literaturreviews (Wißhak 2022). Folglich beziehen sich die Aufgaben auf die drei Inhaltsbereiche *Teilnehmendenmerkmale*, *Weiterbildungsdesign* und *Arbeitsumfeld*. Für die Aufgabenkonstruktion wurden Determinanten verwendet, die einen eindeutigen Zu-

Kreuzen Sie bitte nur eine Antwortoption an, und zwar diejenige, welche die Frage am besten beantwortet. Die übrigen Antworten müssen nicht falsch sein, die Zusammenhänge mit Transfer sind aber geringer. Falls Sie die richtige Antwort nicht kennen, bitten wir Sie, das Kästchen "weiß nicht" anzuklicken und nicht zu raten.

Die erlernten Trainingsinhalte sollen im Anschluss auf verschiedene neue Anwendungssituationen übertragen werden, die sich stark von der Trainingssituation unterscheiden können. Wie würden Sie vorgehen?

Die Lernenden sehen sich Videos zum Thema an.

Ich gebe den Lernenden eine Schritt für Schritt Anleitung.

Ich lasse die Lernenden die neuen Inhalte selbst erkunden.

Ich erarbeite die Inhalte mit den Lernenden durch gezieltes Fragen.

weiß nicht

Abb. 1 Beispielitem Forced-Choice (FC). (Korrekte Antwort kursiv gekennzeichnet)

Hier kreuzen Sie bitte für jede der vier Auswahloptionen an, ob sie richtig oder falsch ist. Sie haben außerdem wieder die Möglichkeit, die Kategorie "weiß nicht" anzuklicken.

Welche der folgenden Aussagen über Trainingsergebnisse ist/sind richtig?

Dass den Teilnehmenden ein Training gefällt, ist wichtiger für den Transfer, als dass sie das Training nützlich finden.

richtig falsch weiß nicht

Der Erwerb von Wissen ist eine ausreichende Bedingung für Transfer in den Arbeitsalltag.

richtig falsch weiß nicht

Der Erwerb von Fähigkeiten ist eine ausreichende Bedingung für Transfer in den Arbeitsalltag.

richtig falsch weiß nicht

Wenn die Teilnehmenden den Trainingsergebnissen einen hohen Wert beimessen, steigert dies die Wahrscheinlichkeit, dass Transfer stattfindet.

richtig falsch weiß nicht

Abb. 2 Beispielitem Multiple-Choice (MC). (Korrekte Antwort kursiv gekennzeichnet)

sammenhang mit der Variable Transfer aufwiesen. Zum anderen wurden Distraktoren (falsche Antwortmöglichkeiten) entwickelt, um die Ratewahrscheinlichkeit möglichst gering zu halten, die Itemschwierigkeit zu erhöhen und die Trennschärfe zu verbessern. Als Distraktoren wurden Konzepte verwendet, die gemäß den Befunden des Literaturreviews (Wißhak 2022) keine oder negative Zusammenhänge mit Transfer aufwiesen oder typische Fehlvorstellungen der Lehrenden widerspiegeln (Lindner et al. 2015).

Daraus resultierten zunächst 80 Aufgaben, welche für jeden Inhaltsbereich Wissen über die Determinanten sowie Handlungswissen zur Förderung von Transfer erfassen sollten. Aus Gründen der Auswertungsobjektivität wählten wir zunächst ein geschlossenes Aufgabenformat (Rost 2004). Jede der Forced-Choice-(FC) und Multiple-Choice-(MC) Aufgaben setzt sich aus einem Itemstamm, welcher die Frage oder die Problemstellung enthält, und dazugehörigen Antwortoptionen zusammen. Die FC-Aufgaben werden dichotom (richtig/falsch) bewertet. Dafür wurden die Lehrenden in den Instruktionen gebeten, nur eine Auswahloption zu wählen und zwar diejenige, welche die Frage am besten beantwortet (s. Abb. 1). Zusätzlich erhielten sie den Hinweis, dass die anderen Auswahloptionen nicht falsch sein müssen, die nachgewiesene Bedeutung für Transfer aber geringer ist. Als richtig wird die Antwort codiert, welche auf Basis des Literaturreviews nachweislich die größte Bedeutung für Transfer aufwies. Für die MC-Aufgaben wählten wir eine polytome Kodierung. Das heißt, die Lehrenden konnten die Aufgabe auch teilweise korrekt beantworten. Hier wurden sie gebeten, jede der vier Auswahloptionen innerhalb einer Aufgabe als *richtig* oder *falsch* (s. Abb. 2) zu beurteilen. Die Auswahloptionen

enthielten sowohl Determinanten, die auf Basis des Literaturreviews nachweislich eine Bedeutung für Transfer haben, aber auch Distraktoren, die von den Lehrenden als falsch erkannt werden mussten. Als Score ergab sich dann ein Summenwert basierend auf der Anzahl der korrekt gelösten Items innerhalb einer Aufgabe. Bei beiden Aufgabenformaten gab es zudem die Option *weiß nicht*.

Um erste Hinweise auf die *Inhaltsvalidität* zu erhalten, wurde die erste Testversion zunächst mittels einer kleinen Stichprobe ($N=10$) erprobt. Die Lehrenden bearbeiteten in Einzelsitzungen jeweils die Hälfte der Items (40 Aufgaben) mittels der Methode des Lauten Denkens (Downing und Haladyna 1997). Hierbei wurden sie gebeten, ihre Gedankenprozesse beim Lösen der Aufgaben simultan zu verbalisieren. Die gesammelten Daten wurden im Anschluss mithilfe der qualitativen Inhaltsanalyse nach Mayring (2015) nach Kriterien wie Verständlichkeit, Vollständigkeit, wahrgenommener Schwierigkeit, Eindeutigkeit, Domänenabhängigkeit und allgemeiner Testgestaltung ausgewertet. Dadurch wurden Aufgaben identifiziert, bei denen die Befragten noch Bearbeitungsschwierigkeiten aufgrund von Störvariablen (z. B. Verständnisprobleme aufgrund der Formulierung) hatten. Gleichzeitig konnten Aufgaben und Auswahloptionen identifiziert werden, die noch zu einfach oder zu schwierig waren. Des Weiteren ließen wir die Aufgaben von 25 Lehrenden in Form eines Online-Fragebogens validieren. Mithilfe von Beurteilungsfragen (z. B. „Inwieweit ist dieses Thema wichtig für alle Trainer*innen – unabhängig vom Trainingsthema und der Branche, in der das Training stattfindet?“) und einer 6-stufigen Likert-Skala (1 = *stimme gar nicht zu* bis 6 = *stimme voll und ganz zu*) bewerteten sie die Items u. a. nach Kriterien der Domänenunabhängigkeit, Verständlichkeit und Relevanz (Voss et al. 2011). Items, die nur von Hardskill- oder von Softskill-Trainerinnen und Trainern bearbeitet werden konnten, wurden eliminiert, um ein domänenunabhängiges Instrument zu erhalten.

Um weitere Hinweise auf die *Inhaltsvalidität* des Testinstruments zu erhalten, bewerteten fünf Expertinnen und Experten aus der Wissenschaft und Praxis jedes Item hinsichtlich deren Relevanz und Repräsentation des Konstrukts Transferwissen auf der Basis einer 5-stufigen Likert-Skala (1 = *stimme zu* bis 5 = *stimme nicht zu*). Außerdem prüften sie unsere Bewertung der Antwortoptionen als richtig oder falsch. Sie hatten zudem die Möglichkeit, die Aufgaben sowie die einzelnen Auswahloptionen zu kommentieren und Transferdeterminanten zu nennen, die aus ihrer Sicht im Testinstrument fehlten.

Die Auswertungsergebnisse der Methode des Lauten Denkens, der fragebogenbasierten Beurteilung der Aufgaben sowie der Einschätzung der Expertinnen und Experten wurden zur Überarbeitung und Selektion der Items genutzt. Um die Ratewahrscheinlichkeit bei den Aufgaben zu verringern, wurden drei zusätzliche Items in einem offenen Aufgabenformat (OF) entwickelt (Rost 2004). Hier wurden die Lehrenden gebeten, transferförderliche Maßnahmen hinsichtlich der Teilnehmendenmerkmale, des Weiterbildungsdesigns und des Arbeitsumfelds in Stichworten zu nennen. Als Erwartungshorizont gaben wir fünf leere Felder vor.

In die quantitative *Pilotstudie* nahmen wir schließlich 39 Items auf. Diese setzten sich aus drei OF-, 14 FC- sowie 22 MC-Aufgaben zusammen.

3.2 Datenerhebung

Die Pilotierung eines Testinstruments anhand einer kleineren Stichprobe gilt als obligatorische Grundlage für die Testentwicklung (Brandt und Moosbrugger 2020). Dafür recherchierten wir in einem ersten Schritt über Online-Datenbanken, Karrierenetzwerke sowie Weiterbildungsträger (z. B. Industrie- und Handelskammern) nach geeigneten Trainerinnen und Trainern. Im nächsten Schritt kontaktierten wir sie per E-Mail, klärten sie über das Forschungsvorhaben auf und baten sie um die freiwillige und anonyme Teilnahme. Die Datenerhebung fand vollständig anonym und mithilfe der Online-Plattform SoSci Survey (Leiner 2020) statt. Der Vorteil der Plattform ist die Gewährleistung der Anonymität und des Datenschutzes, da jeder Person bzw. jedem Fall eine ID, bestehend aus Zahlen zugeordnet wird und die IP-Adressen der Teilnehmenden nicht aufgezeichnet werden (Leiner 2020). Zu Beginn des Tests wurden die Lehrenden nochmals über das Forschungsvorhaben, den Umgang mit den erhobenen Daten sowie die Form der Veröffentlichung informiert. Die Informationen zum Datenschutz nach Art. 13 DS-GVO und die Einwilligungserklärung konnten sie sich als PDF-Dokument herunterladen. Mittels Mausclick mussten die Lehrenden daraufhin ihre Einwilligung zur Teilnahme an der Pilotstudie und die damit verbundene Datenverarbeitung geben. Für die Bearbeitung konnten sich die Lehrenden die Zeit nehmen, die sie benötigten und den Test auch zwischenzeitlich unterbrechen. Um ein realistisches Bild über das Transferwissen der Lehrenden zu erhalten, wurden sie gebeten, die Aufgaben ohne Hilfsmittel zu bearbeiten. Die mittlere Bearbeitungszeit lag bei $M = 32,47$ min ($SD = 12,67$).

3.3 Stichprobenmerkmale

Nach Abschluss der Erhebung lagen insgesamt 181 Fälle vor. Davon wurden diejenigen Fälle ausgeschlossen, die abgebrochen wurden, ein auffälliges Antwortverhalten aufwiesen ($n = 68$) oder bei denen es sich nicht um Trainerinnen oder Trainer ($n = 8$) handelte. Vereinzelt übersprungene Aufgaben in abgeschlossenen Fällen wurden als *falsch* codiert.

Zur Datenanalyse lagen schlussendlich die Testergebnisse von $N = 105$ Lehrenden vor, wovon 47,6 % Personen weiblich und 52,4 % männlich waren. Das Durchschnittsalter der Lehrenden lag bei 52,8 Jahren ($SD = 9,8$). Knapp die Hälfte (41,9 %) übte die Trainertätigkeit hauptberuflich aus, 39 % nebenberuflich und 19 % gaben die Trainertätigkeit als Teil ihres Hauptberufs an.

Ein hoher Anteil (81,9 %) der Lehrenden arbeitete extern, z. B. auf Honorarbasis, für Unternehmen. Nur 10,5 % waren als interne Lehrende tätig.

Über die Hälfte (61,9 %) gab als höchsten berufsqualifizierenden Abschluss einen Hochschulabschluss an (z. B. Diplom in Betriebswirtschaft). Lediglich 12,4 % der Lehrenden gaben einen pädagogischen Studienabschluss (z. B. Erziehungswissenschaften) an. Es verfügten 14,3 % über eine Fachausbildung (z. B. Kommunikationsfachwirtin oder -fachwirt). Weitere 11,4 % hatten eine Berufsausbildung (z. B. Industriekauffrau oder -mann) absolviert. Nur 5,7 % hatten promoviert und 1,9 % der Lehrenden gaben an, keinen Abschluss zu besitzen. Zudem hatte der Großteil (70,5 %) eine Trainerausbildung absolviert.

3.4 Datenanalyse

Die Skalierung erfolgte mittels IRT (Embretson und Reise 2000). Die IRT ermöglicht die Verortung der Personenfähigkeiten und der Itemschwierigkeiten auf einer gemeinsamen Skala (Masters 1982). Aufgrund des mehrstufigen Antwortformats nutzten wir das Partial-Credit-Modell (Masters 1982). Das Ziel dieser Skalierung war es, zu prüfen, ob sich das Konstrukt Transferwissen auf einer eindimensionalen Skala abbilden lässt.

Um das Modell berechnen zu können, erfolgte in einem ersten Schritt die Codierung der offenen Antworten mittels der qualitativen Inhaltsanalyse nach Mayring (2015). Entscheidend für die Auswertung der Pilotierungsdaten und die Berechnung des Modells waren die als *richtig* codierten Antworten der Lehrenden, da diese letztlich bepunktet wurden oder eine Null vergeben wurde, so dass es keine Minuspunkte gab. Zwei Raterinnen codierten die Antworten von 32 Lehrenden (30% des Gesamtmaterials). Zur Ermittlung der Beurteilerübereinstimmung wurde Cohens Kappa berechnet (Cohen 1960; Wirtz und Caspar 2002). Alle drei Aufgaben wiesen mindestens einen Übereinstimmungswert von 0,7 auf (OF1: 0,78; OF2: 0,81; OF3: 0,81). Gemäß Landis und Koch (1977) weisen die OF2 und OF3 damit eine sehr gute und die OF1 eine substantielle Übereinstimmung auf, sodass die übrigen Aufgaben von einer Person codiert werden konnten.

In einem nächsten Schritt wurden die Testantworten aller Aufgabenformate der 105 Lehrenden unter Anwendung des Partial-Credit-Modells skaliert und interpretiert. Das Modell eignet sich besonders gut für Aufgaben mit mehrstufigen Antwortkategorien (Bond und Fox 2007). Besonders die mehrstufigen Items ermöglichen eine genauere Einschätzung der Fähigkeit von Personen als es bei dichotomen Items möglich ist (Masters 1982). Die Modellierung erfolgte mithilfe der Software ConQuest (Wu et al. 2007). Die Messgenauigkeit des Tests wurde über die EAP/PV-Reliabilität geschätzt. Die Qualität der Items wurde mittels der klassischen Trennschärfe, der punktbiserialen Korrelationen und der IRT-Schwierigkeitsschätzer überprüft (Bond und Fox 2007). Dabei wurden zu schwere und zu leichte Items ausgeschlossen. Darüber hinaus wurden die Aussagen von $N=10$ Lehrenden, die den Test zusätzlich mittels der Methode des Lauten Denkens bearbeiteten, für die Entfernung einzelner Items herangezogen. Schlussendlich wurden daraufhin fünf FC- und zwei MC-Aufgaben ausgeschlossen. In das hier berichtete Modell flossen 32 Items ein, welche sich aus drei OF-, 9 FC- und 20 MC-Aufgaben zusammensetzen (s. Tab. 1). Der Inhaltsbereich Weiterbildungsdesign enthält die größte Anzahl an Items, da im Zuge des Literaturreviews festgestellt wurde, dass die Lehrenden hierauf am meisten Einfluss nehmen können (Wißhak 2022).

Tab. 1 Übersicht über die Anzahl der Items und Aufgabenformate pro Inhaltsbereich

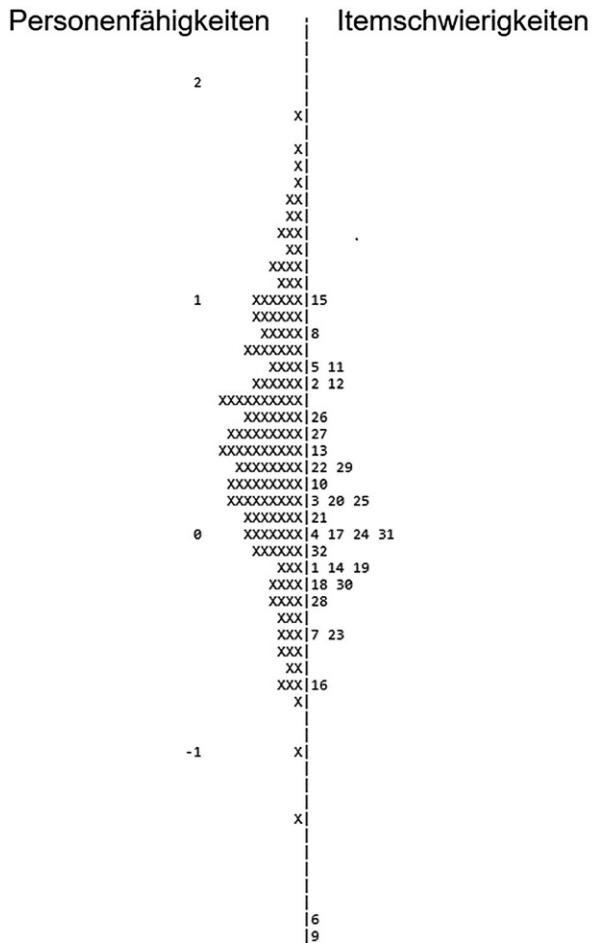
Einflussfaktor	OF-Aufgaben	FC-Aufgaben	MC-Aufgaben
Teilnehmendenmerkmale (6 Items)	1 Item	1 Item	4 Items
Weiterbildungsdesign (17 Items)	1 Item	7 Items	9 Items
Arbeitsumfeld (9 Items)	1 Item	1 Item	7 Items
<i>Summe (32 Items)</i>	<i>3 Items</i>	<i>9 Items</i>	<i>20 Items</i>

4 Ergebnisse

Die Ergebnisse weisen darauf hin, dass sich Transferwissen mithilfe des Instruments operationalisieren und reliabel messen lässt. Die EAP/PV-Reliabilität ist ein Reliabilitätsschätzer, der ähnlich wie Cronbachs Alpha zu interpretieren ist (Rost 2004). Die EAP/PV-Reliabilität betrug 0,81 und kann damit als gut angesehen werden. Die interne Konsistenz der Personen-Testwerte (coefficient alpha) lag bei 0,79 und ist damit zufriedenstellend. Die Gesamtvarianz der latenten Variable ist mit 0,315 vergleichsweise niedrig, aber immer noch ausreichend.

Gemäß den Empfehlungen von Bond und Fox (2007) zeigten die Items eine gute Passung zum eindimensionalen Partial-Credit-Modell. Das Toleranzintervall der Item-Infit Werte war 0,80 bis 1,20 (Wu et al. 2007). Empirisch lagen die Infit-Werte der Testitems zwischen 0,93 und 1,16 und zeigten damit eine gute Passung

Abb. 3 Darstellung der Personenfähigkeiten und Itemschwierigkeiten in Form einer Wright Map



jedes X= 0,6 Fälle

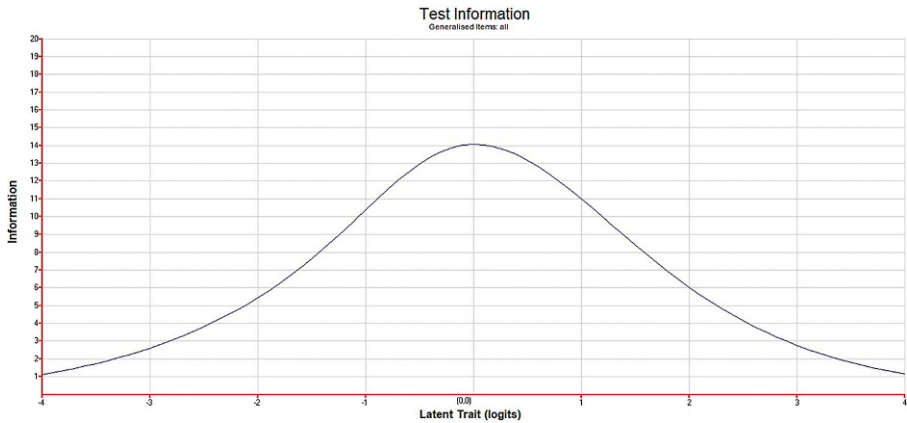


Abb. 4 Testinformationsfunktion

zum Modell. Die Trennschärfe drückt aus, wie gut ein Item zwischen verschiedenen Personenfähigkeiten unterscheidet (Rost 2004). Die Items wiesen angemessene Trennschärfen (Diskrimination) $>0,20$ auf (Haladyna 2004). Lediglich fünf Items wiesen Trennschärfen zwischen 0,12 und 0,18 auf, wurden jedoch beibehalten, da sie insgesamt zu einer besseren Messgenauigkeit des Tests (EAP/PV-Reliabilität) beitrugen. Bei den dichotomen Items lagen die Werte zwischen 0,12 und 0,44. Um sicherzugehen, dass zwischen den Abstufungen der polytomen Items tatsächlich Fähigkeitsunterschiede vorlagen, erfolgte die Punkteabstufung empirisch, und die Trennschärfen lagen zwischen 0,26 und 0,55. Die punkt-biserialen Korrelationen der Antwortkategorien lagen geordnet vor und stiegen mit jeder Stufe an (Bond und Fox 2007).

Die sogenannte Wright Map (Abb. 3) stellt die Personenfähigkeiten und Itemschwierigkeiten anschaulich in einer gemeinsamen Darstellung gegenüber. Im oberen Bereich werden Personen mit hohen Fähigkeiten und Items mit hoher Schwierigkeit dargestellt. Der untere Bereich bildet Personen mit geringeren Fähigkeiten und leichtere Items ab. Alle 32 Items lagen modellkonform zwischen $-2,0$ und $2,0$ (Bond und Fox 2007).

Die Testinformationsfunktion sagt aus, wie gut ein Test die Personenfähigkeit über den gesamten Fähigkeitsbereich misst (Baker und Kim 2017). Wie in Abb. 4 ersichtlich, hat die Funktion einen normalverteilten und gleichmäßigen Verlauf. Zwischen -1 und $+1$ weist sie ihren maximalen Wert auf. Das Transferwissen von Lehrenden, deren Fähigkeiten um diesen Bereich liegen, kann demzufolge mithilfe des Testinstruments sehr reliabel gemessen werden (Hambleton und Jones 1993). Der Informationsgehalt für Personen mit geringen und hohen Fähigkeiten ist ebenfalls zufriedenstellend, da die Items ein gutes Gleichgewicht zwischen der Anzahl von leichten und schweren Items aufweisen (s. Abb. 3).

5 Zusammenfassung und Diskussion

In diesem Beitrag wurden die Entwicklung und Erprobung eines psychometrischen Instruments dargestellt, welches bei Lehrenden in der berufsbezogenen Weiterbildung das Transferwissen als Teil des Professionswissens erfassen soll. Konzeptuelle Grundlage dafür waren etablierte Transfermodelle sowie vornehmlich ein systematisches Literaturreview, welches die empirische, quantitativ orientierte internationale Forschung zu Transferdeterminanten in der berufsbezogenen Weiterbildung anhand von Metaanalysen auswertete (Wißhak 2022). Demzufolge gibt es zahlreiche Transferdeterminanten, die die Lehrenden kennen sollten und einige Möglichkeiten der Transferförderung, für die sie das nötige Handlungswissen besitzen sollten. Dieses Wissen lässt sich, gemäß dem Transfermodell von Baldwin und Ford (1988), wiederum den Teilnehmendenmerkmalen, dem Weiterbildungsdesign oder den Merkmalen des Arbeitsumfelds zuordnen.

Für die Testentwicklung wurden zunächst insgesamt 83 Items im Forced-Choice- und Multiple-Choice-Format sowie im offenen Aufgabenformat generiert. Diese Aufgaben wurden in mehreren qualitativen Erprobungsschritten mit unterschiedlichen Stichproben von Expertinnen und Experten aus Forschung und Praxis hinsichtlich ihrer Praktikabilität und Inhaltsvalidität optimiert und selektiert.

Im Mittelpunkt des Beitrags stand eine darauf basierende Pilotierungserhebung der verbliebenen 39 Items, für die anhand einer Stichprobe aus 105 Lehrenden die Reliabilität auf der Grundlage eines eindimensionalen Partial-Credit-Modells, aus der Familie der Rasch-Modelle, ermittelt wurde. Insgesamt erzielten die Items eine gute Reliabilität, angemessene Trennschärfen und eine modellkonforme Verteilung von Personenfähigkeiten und Itemschwierigkeiten.

Damit sind die ersten qualitativen und quantitativen Entwicklungsschritte des Instruments insgesamt sehr zufriedenstellend verlaufen. Lediglich die Varianz der latenten Variable weist bislang noch einen verbesserungswürdigen Wert auf. Es kann jedoch davon ausgegangen werden, dass sich diese mittels einer größeren Stichprobe verbessert. Die IRT-Analysen zeigen, dass sich das Konstrukt Transferwissen eindimensional abbilden lässt. Darüber hinaus könnte in zukünftigen Erhebungen mit größeren Stichproben ein mehrdimensionales Modell geprüft werden, das sich beispielsweise entlang der drei Einflussbereiche nach Baldwin und Ford (1988) aufgliedert oder das deklaratives und prozedurales Wissen unterscheidet. Zusätzlich sollen mittels größerer Stichproben und DIF-Analysen Gruppenunterschiede hinsichtlich des Transferwissens (z. B. zwischen Hard- und Softskill Trainerinnen und Trainern) untersucht werden.

Eine weitere, nur mit zusätzlichen Stichproben zu beantwortende Frage ist, wie weit das auf diese Weise erfasste Transferwissen auf andere Bereiche der Erwachsenenbildung übertragbar ist. Zunächst wird jedoch keine über die berufsbezogene Weiterbildung hinaus gehende Geltung des Instruments beansprucht.

Ähnlich stellt sich auch die Frage nach der Konstruktvalidität. Konkret wäre zu prüfen, inwieweit das hier erhobene Transferwissen differenzielle Validität über das generische pädagogisch-psychologische Wissen hinaus aufweist. Hierzu könnte man in weiteren Erhebungen ergänzend das Testinstrument zur Erfassung des pädagogisch-psychologischen Wissens von Lehrkräften aus Schule und Erwachse-

nenbildung (Marx et al. 2017) einsetzen und sowohl die direkten Zusammenhänge beider Instrumente als auch die inkrementelle Kriteriumsvalidität etwa in Bezug auf den beruflichen Erfolg der Lehrenden ermitteln. Nicht zuletzt wäre es interessant, dies für Lehrkräfte aus unterschiedlichen Bereichen der Erwachsenenbildung zu vergleichen.

Wir möchten außerdem darauf hinweisen, dass für die Transferförderung durch Lehrende auch andere Kompetenzaspekte, wie ihre wahrgenommene Verantwortung für Transfer, von Bedeutung sind (Wißhak und Barth 2021). Außerdem ist denkbar, dass auch das in dem Instrument nicht erfasste meta-kognitive Wissen der Lehrenden eine Rolle spielen könnte (Hasselhorn 1992).

Sobald der Test abschließend validiert ist, kann er in der Praxis genutzt werden, um das Transferwissen von Lehrkräften in der berufsbezogenen Weiterbildung zu ermitteln. Da eine gelingende Transferförderung eine entscheidende Rolle für den Erfolg von Weiterbildungen spielt, kann ein entsprechender Test sowohl als Grundlage für die Auswahl von Lehrkräften dienen als auch als Basis für eine entsprechende, fremd- oder selbstinitiierte Weiterentwicklung der Lehrenden. Damit ist der Test nicht nur als Baustein für eine aussagekräftige Bestimmung der Professionalität von Lehrenden anzusehen, sondern kann darüber hinaus dabei unterstützen, die individuelle Professionalisierung der Lehrenden im Sinne einer Kompetenzentwicklung voranzutreiben, wie dies etwa Nittel und Seltrecht (2008) allgemeiner postulieren. Die gewonnenen Daten könnten genutzt werden, um die Ausbildung von Trainerinnen und Trainern entsprechend zu optimieren und damit der Forderung nach wissenschaftlich fundierten Fortbildungsmaßnahmen in der Erwachsenenbildung nachzugehen (Schrader 2010).

Abschließend hoffen wir, mit unserer Forschung langfristig ein Bewusstsein für die Bedeutung der Transferförderung zu schaffen. Dafür gilt es neben dem Transferwissen auch die empfundene Verantwortung, Motivation und Selbstwirksamkeit der Lehrenden dahingehend zu stärken, Transfer aktiv im Rahmen ihrer Möglichkeiten zu fördern.

Funding This study was financially supported by the University of Konstanz's Young Scholar Fund (YSF) for excellent postdocs and by an Independent Research Grant (IRG) of the Zukunftskolleg.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

Literatur

- Alliger, G. M., Tannenbaum, S. I., Bennett, W., Traver, H., & Shotland, A. (1997). A meta-analysis of the relations among training criteria. *Personnel Psychology*, *50*(2), 341–358. <https://doi.org/10.1111/j.1744-6570.1997.tb00911.x>.
- Autorengruppe Bildungsberichterstattung (2020). *Bildung in Deutschland 2020. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung in einer digitalisierten Welt*. Bielefeld: wbv Publikation. <https://doi.org/10.3278/6001820gw>.
- Baker, F. B., & Kim, S.-H. (2017). The information function. In F. B. Baker & S.-H. Kim (Hrsg.), *Statistics for social and behavioral sciences. The basics of item response theory using R* (S. 89–104). Cham: Springer. https://doi.org/10.1007/978-3-319-54205-8_6.
- Baldwin, T. T., & Ford, J. K. (1988). Transfer of training: a review and directions for future research. *Personnel Psychology*, *41*(1), 63–105. <https://doi.org/10.1111/j.1744-6570.1988.tb00632.x>.
- Baldwin, T. T., Ford, J. K., & Blume, B. D. (2017). The state of transfer of training research: moving toward more consumer-centric inquiry. *Human Resource Development Quarterly*, *28*(1), 17–28. <https://doi.org/10.1002/hrdq.21278>.
- Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, *9*(4), 469–520. <https://doi.org/10.1007/s11618-006-0165-2>.
- Blume, B. D., Ford, K. J., Surface, E. A., & Olenick, J. (2019). A dynamic model of training transfer. *Human Resource Management Review*, *29*(2), 270–283. <https://doi.org/10.1016/j.hrmr.2017.11.004>.
- Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: fundamental measurement in the human sciences*. New Jersey: Lawrence Erlbaum.
- Bonnes, C., & Hochholdinger, S. (2016). Die berufliche Entwicklung und die Wege des Qualifikations- und Kompetenzerwerbs von Lehrenden in der berufsbezogenen Weiterbildung. *Berufs- und Wirtschaftspädagogik – online*, *29*, 1–20. http://www.bwpat.de/ausgabe29/bonnes_hochholdinger_bwpat29.pdf. Zugegriffen: 18. März 2022.
- Brandt, H., & Moosbrugger, H. (2020). Planungsaspekte und Konstruktionsphasen von Tests und Fragebogen. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 39–66). Berlin, Heidelberg: Springer.
- Burke, L. A., & Hutchins, H. M. (2008). A study of best practices in training transfer and proposed model of transfer. *Human Resource Development Quarterly*, *19*(2), 107–128. <https://doi.org/10.1002/hrdq.1230>.
- Christ, J., Koscheck, S., Martin, A., & Widany, S. (2019). *Wissenstransfer – Wie kommt die Wissenschaft in die Praxis? Ergebnisse der wbmonitor Umfrage 2018*. Bonn: BIBB.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46. <https://doi.org/10.1177/0013164460020001104>.
- Donovan, P., & Darcy, D. P. (2011). Learning transfer: the views of practitioners in Ireland. *International Journal of Training and Development*, *15*(2), 121–139. <https://doi.org/10.1111/j.1468-2419.2011.00374.x>.
- Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Berlin: Springer.
- Downing, S. M., & Haladyna, T. M. (1997). Test item development: validity evidence from quality assurance procedures. *Applied Measurement in Education*, *10*(1), 61–82. https://doi.org/10.1207/s15324818ame1001_4.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. New Jersey: Lawrence Erlbaum.
- Fuchs, S. (2011). *Professionalitätswissenschaft des Weiterbildungspersonals: Tätigkeiten, Kompetenzen und Fortbildung von Trainern in der beruflichen/betrieblichen Weiterbildung*. Hamburg: Kovac.
- Gegenfurtner, A. (2011). Motivation and transfer in professional training: a metaanalysis of the moderating effects of knowledge type, instruction, and assessment conditions. *Educational Research Review*, *3*(6), 153–168. <https://doi.org/10.1016/j.edurev.2011.04.001>.
- Gegenfurtner, A., Veermans, K., & Vauras, M. (2013). Effects of computer support, collaboration, and time lag on performance self-efficacy and transfer of training: a longitudinal meta-analysis. *Educational Research Review*. <https://doi.org/10.1016/j.edurev.2012.04.001>.
- Grohmann, A., Beller, J., & Kauffeld, S. (2014). Exploring the critical role of motivation to transfer in the training transfer process. *International Journal of Training and Development*, *18*(2), 84–103. <https://doi.org/10.1111/ijtd.12030>.

- Haladyna, T. M. (2004). *Developing and validating multiple-choice test items* (3. Aufl.). New Jersey: Lawrence Erlbaum.
- Hambleton, R. K., & Jones, R. W. (1993). An NCME instructional module on. *Educational Measurement: Issues and Practice*, 12(3), 38–47. <https://doi.org/10.1111/j.1745-3992.1993.tb00543.x>.
- Hasselhorn, M. (1992). Metakognition und Lernen. In G. Narr (Hrsg.), *Lernbedingungen und Lernstrategien. Welche Rolle spielen kognitive Verstehtensstrukturen?* (S. 35–63). Tübingen: Narr.
- Hense, J., & Mandl, H. (2011). Transfer in der beruflichen Weiterbildung. In O. Zlatkin-Troitschanskaia (Hrsg.), *Stationen Empirischer Bildungsforschung* (S. 249–263). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-531-94025-0_18.
- Hetfleisch, P., Goeze, A., & Schrader, J. (2017). Wie PraktikerInnen wissenschaftliche Befunde verwenden. Selektions- und Rezeptionsprozesse bei der Implementation eines evidenzbasierten Trainingskonzepts. *Zeitschrift für Pädagogik*, 63(2), 182–205. <https://doi.org/10.25656/01:18490>.
- Holton III, E. F. (2004). Implementing evidence-based practices: time for a national movement? *Human Resource Development Review*, 3, 187–188.
- Holton III, E. F., Bates, R. A., & Ruona, W. E. A. (2000). Development of a generalized learning transfer system inventory. *Human Resource Development Quarterly*, 11(4), 333–360.
- Hutchins, H. M., & Burke, L. A. (2007). Identifying trainers' knowledge of training transfer research findings—closing the gap between research and practice. *International Journal of Training and Development*, 11(4), 236–264. <https://doi.org/10.1111/j.1468-2419.2007.00288.x>.
- Hutchins, H. M., Burke, L. A., & Berthelsen, A. M. (2010). A missing link in the transfer problem? Examining how trainers learn about training transfer. *Human Resource Management*, 49(4), 599–618. <https://doi.org/10.1002/hrm.20371>.
- Kanning, U. P. (2014). Prozess und Methoden der Personalentwicklung. In H. Schuler & U. P. Kanning (Hrsg.), *Lehrbuch der Personalpsychologie* (S. 501–562). Göttingen, Bern, Wien: Hogrefe.
- König, J., & Blömeke, S. (2009). Pädagogisches Wissen von angehenden Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 12(3), 499–527. <https://doi.org/10.1007/s11618-009-0085-z>.
- Kuhnhenne, M. (2020). Lebensbegleitendes oder lebenslanges Lernen? Herausforderungen der Weiterbildung. *Forschungsförderung Report No. 8*. <https://www.econstor.eu/bitstream/10419/233581/1/fofoe-report-008-2020.pdf>. Zugegriffen: 18. März 2022.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159. <https://doi.org/10.2307/2529310>.
- Leiner, D. J. (2020). SoSci Survey (Version 3.2.14i) [Computer software]. <https://www.sosicisurvey.de>. Zugegriffen: 18. März 2022.
- Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS medicine*, 6(7), 1–28. <https://doi.org/10.1371/journal.pmed.1000100>.
- Lindner, M. A., Strobel, B., & Köller, O. (2015). Multiple-Choice-Prüfungen an Hochschulen? *Zeitschrift für Pädagogische Psychologie*, 29(3–4), 133–149. <https://doi.org/10.1024/1010-0652/a000156>.
- Marx, C., Goeze, A., Voss, T., Hoehne, V., Klotz, V. K., & Schrader, J. (2017). Pädagogisch-psychologisches Wissen von Lehrkräften aus Schule und Erwachsenenbildung: Entwicklung und Erprobung eines Testinstruments. *Zeitschrift für Erziehungswissenschaft*, 20(S1), 165–200. <https://doi.org/10.1007/s11618-017-0733-7>.
- Masters, G. N. (1982). A rasch model for partial credit scoring. *Psychometrika*, 47(2), 149–174. <https://doi.org/10.1007/BF02296272>.
- Mayring, P. (2015). *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. Weinheim, Basel: Beltz.
- Nittel, D., & Seltrecht, A. (2008). Der Pfad der „individuellen Professionalisierung“. Ein Beitrag zur kritisch-konstruktiven erziehungswissenschaftlichen Berufsgruppenforschung. Fritz Schütze zum 65. Geburtstag. *BIOS*, 21(1), 124–145.
- Rasli, A., Tat, H. H., Chin, T. A., & Khalaf, B. (2012). Identification of factors and attributes for effective transfer of IT training to the workplace. *Procedia – Social and Behavioral Sciences*, 40, 174–181. <https://doi.org/10.1016/j.sbspro.2012.03.178>.
- Rost, J. (2004). *Lehrbuch Testtheorie – Testkonstruktion*. Bern: Huber.
- Russ-Eft, D. F., Dickison, P., & Levine, R. (2010). Taking the pulse of training transfer: Instructor quality and EMT certification examination results. *Human Resource Development Quarterly*, 21(3), 291–306. <https://doi.org/10.1002/hrdq.20052>.
- Salas, E., Tannenbaum, S. I., Kraiger, K., & Smith-Jentsch, K. A. (2012). The science of training and development in organizations: what matters in practice. *Psychological Science in the Public Interest*, 13(2), 74–101. <https://doi.org/10.1177/1529100612436661>.

- Sandmeier, A., Hanke, U., & Gubler, M. (2021). Entwicklung und Validierung eines praxistauglichen Evaluationsinstruments zur Messung und Optimierung von Lerntransfer. *Zeitschrift für Evaluation*, 20(1), 11–36.
- Schrader, J. (2010). Fortbildung von Lehrenden in der Erwachsenenbildung: Notwendig? Sinnvoll? Möglich? Bedarf und Angebote im Überblick. In J. Schrader, R. Hohmann & S. Hartz (Hrsg.), *Mediengestützte Fallarbeit. Konzepte, Erfahrungen und Befunde zur Kompetenzentwicklung von Erwachsenenbildnern* (S. 25–68). Bielefeld: wbv.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14. <https://doi.org/10.3102/0013189X015002004>.
- Strauch, A., Lencer, S., Bosche, B., Gladkova, V., Schneider, M., & Trevino-Eberhard, D. (2019). GRETA – kompetent handeln in Training, Kurs & Seminar Das GRETA Kompetenzmodell. <https://www.die-bonn.de/id/37005/about/html/>. Zugegriffen: 15. Dez. 2021.
- Strauch, A., Alberti, V., & Schneider, M. (2020). Professionelle Handlungskompetenzen Lehrender in der Erwachsenen- und Weiterbildung anerkennen. *Hessische Blätter für Volksbildung*. <https://doi.org/10.3278/HBV2001W010>.
- Voss, T., Kunter, M., & Baumert, J. (2011). Assessing teacher candidates' general pedagogical/psychological knowledge: test construction and validation. *Journal of Educational Psychology*, 103(4), 952–969.
- Wirtz, M. A., & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität: Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen*. Göttingen, Bern: Hogrefe.
- Wißhak, S. (2022). Transfer in der berufsbezogenen Weiterbildung: Systematisches Literaturreview und Synthese mit Blick auf die Handlungsmöglichkeiten der Lehrenden. *Zeitschrift für Weiterbildungsforschung*. <https://doi.org/10.1007/s40955-022-00204-y>.
- Wißhak, S., & Barth, D. (2021). Perceptions of accountability for the transfer of training by leadership trainers. *International Journal of Training and Development*. <https://doi.org/10.1111/ijtd.12255>.
- Wißhak, S., & Hochholdinger, S. (2018). Trainers' knowledge and skills from the perspective of trainers, trainees and human resource development practitioners. *International Journal of Training Research*, 16(3), 218–231. <https://doi.org/10.1080/14480220.2018.1576327>.
- Wißhak, S., Bonnes, C., Keller, I., Barth, D., & Hochholdinger, S. (2020). Qualifikationen von Lehrenden in der beruflich-betrieblichen Weiterbildung. *Zeitschrift für Bildungsforschung*, 10(1), 103–123. <https://doi.org/10.1007/s35834-020-00262-7>.
- Wu, M. L., Adams, R. J., Wilson, M., & Haldane, S. A. (2007). *ACER ConQuest. Version 2.0. Generalised item response software*. Camberwell: ACER Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.