KoKoHs Assessment-Portfolio

Testverfahren zur Modellierung und Messung generischer und domänenspezifischer Kompetenzen bei Studierenden und Hochschulabsolventen

KoKoHs

GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung

Editorial

KoKoHs Assessment-Portfolio: Testverfahren zur Modellierung und Messung generischer und domänenspezifischer Kompetenzen bei Studierenden und Hochschulabsolventen

Die heutige, sich dynamisch entwickelnde Wissensgesellschaft erfordert von HochschulabsolventInnen, fachliches Wissen und fachübergreifende Kompetenzen zu erwerben sowie transferieren und auf gesellschaftliche und professionelle Herausforderungen anwenden zu können. Um den weitreichenden Anforderungen des 21. Jahrhunderts gerecht zu werden, ist eine Kompetenzorientierung in der Hochschulpraxis unerlässlich. Erwerb von domänenspezifischen und generischen Kompetenzen ist ein zentrales Lehr-Lern-Ziel der Hochschulbildung. Um Studiengänge, Lehrveranstaltungen und Prüfungen kompetenzorientiert zu gestalten und den Lernfortschritt der Studierenden zuverlässig zu erfassen, sind objektive, reliable und valide Messungen der in der Hochschule erworbenen Kompetenzen erforderlich.

Mit diesem Ziel wurde die deutschlandweite Förderinitiative „KoKoHs: Kompetenzmodellierung und Kompetenzerfassung im Hochschulsektor" vom Bundesministerium für Bildung und Forschung (BMBF) etabliert. In zwei Förderphasen wurden von 2011 bis 2020 theoretisch-konzeptionelle Kompetenzmodelle und korrespondierende Instrumente und Verfahren zur Kompetenzerfassung von Studierenden entwickelt und umfassend validiert. In rund 40 groß angelegten Verbundprojekten wurden Testinstrumente entwickelt oder adaptiert, um akademische Kompetenzen im Hochschulbereich valide und reliabel zu erfassen. Neben fachspezifischen Kompetenzen in verschiedenen Studiendomänen (Wirtschafts- und Sozialwissenschaften, MINT, Lehrerbildung, Psychologie, Medizin) wurden auch generische Kompetenzen (z.B. Selbstregulation, akademisches Textverständnis) fokussiert. Die entwickelten Modelle und Instrumente ermöglichen die valide Erfassung der Ausprägung von Kompetenzen beim Studieneintritt und ihrer weiteren Entwicklung im Studienverlauf sowie in anschließenden berufspraktischen Phasen (wie in Lehrerbildung und Medizin). KoKoHs-Instrumente eigenen sind zur Messung der folgenden Arten von Konstrukten: *(i)* wissensbasierte und affektiv-volitionale Dispositionen als Voraussetzungen für die Leistung der Studierenden in zukünftigen beruflichen Situationen; *(ii)* im Studium erworbene situations-bezogene Fähigkeiten und Fertigkeiten sowie *(iii)* tatsächliche Performanz i.S. der Leistung in realistischen Situationen. Zur validen Erfassung authentischer professioneller Aufgaben wurden mehrere computergestützte Assessments (z.B. videobasierte Unterrichtssimulationen) entwickelt, die sich besonders für die Prozessdiagnostik eignen.

In den KoKoHs-Projekten wurden insgesamt über 50 Testinstrumente verschiedener Formate entwickelt bzw. adaptiert (inkl. innovative video-, computer- und simulationsbasierte Instrumente) und deutschlandweit erprobt. Die vorliegende Publikation bietet eine detaillierte Übersicht in Form von Steckbriefen der meisten in KoKoHs entwickelten Testinstrumente. Jeder Steckbrief enthält die zentralen Angaben auf einen Blick (Kurzfassung) sowie eine anschließende differenzierte Darstellung (Langfassung), in der alle relevanten Informationen zu den Testinstrumenten angegeben sind, wie zum konzeptionellen Hintergrund, Testaufbau, Validierung sowie Hinweise zum Einsatz in der Praxis. Mit dieser Übersicht und weiterführenden Informationen soll der Einsatz der Instrumente in weiteren Forschungs- und Transferprojekten in der Hochschulpraxis gefördert werden. In den Steckbriefen werden vielfältige bereits erprobte sowie auch perspektivische Einsatzmöglichkeiten der Kompetenztests für die Hochschulpraxis exemplarisch aufgezeigt. Neben den Instrumenten für die Eingangsdiagnostik können formative Assessments die Lehr-Lern-Prozesse und den Fortschritt im Kompetenzerwerb effektiv unterstützen und summative Assessments die Lernergebnisse der Studierenden valide erfassen. Ergebnisse der Kompetenztests können zur Gestaltung von Lehr-Lern-Settings genutzt werden, i.S. einer Passung von Curriculum, Instruktion und Assessment, und so entscheidend dazu beitragen, Lehr-Lern-Inhalte, die Gestaltung der Lehre und die Prüfungspraxis optimal aufeinander abzustimmen. Die Kompetenzassessments können auch für eine objektive Evaluierung von Studienprogrammen wie auch ganzen Fachbereichen gemeinsam mit anderen Indikatoren herangezogen werden und damit zur Qualitätsentwicklung und evidenzbasierten Steuerung im Hochschulbereich beitragen. Die Erfassung von akademischen Kompetenzen mittels KoKoHs-Kompetenztests kann ein Baustein des beständigen Prozesses der Weiterentwicklung von Lehren und Lernen in der Hochschule sein. Diese Publikation soll die Bekanntheit der Testinstrumente mitsamt Qualitätsstandards in den verschiedenen Ebenen der Hochschulforschung, -praxis und -politik erhöhen und Wege zu einem gelungenen Transfer eröffnen.

Wir bedanken uns sehr herzlich bei allen TestentwicklerInnen aus dem KoKoHs-Programm für die Bereitstellung der Instrumente und weiteren Projektinformationen zur Erstellung und Publikation der Steckbriefe. Ein besonderer Dank gilt auch dem BMBF und insbesondere *Martina Diegelmann* für die konzeptionelle und finanzielle Unterstützung dieser Publikation.

Olga Zlatkin-Troitschanskaia, Hans Anand Pant, Marie-Theres Nagel, Dimitri Molerov, Corinna Lautenbach, Miriam Toepper
Mainz, Berlin, Januar 2020

Inhalt

Editorial . . . 3

AkaTex Akademische Textkompetenzen: Wissenschaftliches Formulieren und Argumentieren . . . 6

ÄKHOM: Facets-of-Competence (FOC): Bewertungsinstrument für ärztliche und pflegerische Bewerter*innen 8

AstraLite-Koswo AST: Argumentstrukturtest . . . 10

Berliner Test zur Erfassung Fluider und Kristalliner Intelligenz (BEFKI) . . . 12

BilWiss-2.0-Test: Bildungswissenschaftliches Wissen von Lehrkräften . . . 15

DaZKom: Deutsch-als-Zweitsprache-Kompetenz von Lehrkräften . . . 18

DaZKom-Video: Deutsch-als-Zweitsprache-Kompetenz von Lehrkräften . . . 20

ELMaWi Mathe-AK: Aktionsbezogene Kompetenz von Mathematiklehrkräften . . . 22

ELMaWi Mathe-RK: Reflexive Kompetenz von Mathematiklehrkräften . . . 25

ELMaWi ARKOM-Wirtschaft: Aktionsbezogene und Reflexive Kompetenz von Wirtschaftslehrkräften . . . 28

WiWi-PCK: Fachdidaktisches Wissen von Wirtschaftslehrkräften . . . 31

EXMO: Kompetenzen von Biologielehrkräften im Vermitteln und Beurteilen von Experimenten . . . 34

HEED: Unternehmensgründungstest . . . 37

KOM@ING KAT: Kategorisierungssysteme . . . 39

KOM@ING NIA: Niedrig Inferente Analysen . . . 41

KOM@ING SEL: Studierende-Expert-Lösung . . . 43

KOM@ING: Experteninterview PARI . . . 45

KoMeWP: Fachwissen und Fachdidaktisches Wissen in Rechnungswesen . . . 47

KomPaed PAnTHoa: Praxisanforderungen von Hochschulabsolvent*innen . . . 50

KomPrü-KomFäh: Performanzbasierte Messung kommunikativer Fähigkeiten von Studierenden . . . 53

Ko-NaMa Assessment: Kompetenzen für das Nachhaltigkeitsmanagement . . . 55

KUI: Kompetenzfacetten für das Unterrichten in Informatik . . . 58

LeScEd: Forschungskompetenz in Bildungswissenschaften . . . 60

MultiTex MDC: Verständnis mehrerer Dokumente . . . 62

M³K: Medienpädagogische Kompetenz . . . 64

MokoMasch ADP: Advanced Design Project Skills im Maschinenbau . . . 66

PlanvoLL-D: Didaktische Adaptivität von Lehrkräften . . . 68

PlanvoLL-D GPK: Pädagogisches Wissen . . . 70

Profile-P PCK: Mechanik von Physiklehrkräften . . . 73

Profile-P CK: Mechanik von Physiklehrkräften . . . 76

Profile-P DET: Dialogischer Erklärtest für Physiklehrkräfte . . . 78

Pro-KomMa: KomMa-Tests zu fachlichem und pädagogischem Wissen (MCK, MPCK und GPK) von frühpädagogischen Fachkräften in Mathematik . . . 80

PRO-SRL-EVA – PRO-SRL: Mein Lerntagebuch . . . 83

PRO-SRL-EVA – PRO-SRL: Situational Judgement Test zur Motivationsregulation . . . 85

PRO-SRL-EVA – PRO-SRL Fragebogen: Selbstreguliertes Lernen – Wissen und Anwendung . . . 88

SOSCIE: Unterrichtsplanung von Geschichtslehrkräften . . . 91

TEDS-Validierung: P-PID und M-PID als situatives allgemeines pädagogisches Wissen und situatives fachdidaktisches Wissen von Mathematiklehrkräften . . . 94

TEDS-Validierung: CME Classroom Management Expertise von Lehrkräften . . . 97

ValiDiS: Ko-WADiS-Test Wissenschaftliches Denken in Naturwissenschaften . . . 100

WiKom-SoWi: Wissenschaftliche Kompetenz in Politikwissenschaften . . . 103

WiKom-SoWi: Forschungskompetenz in Psychologie . . . 105

WiKom-SoWi: Wissenschaftliche Kompetenz in Soziologie . . . 107

WiWiKom: Wirtschaftswissenschaftliche Fachkompetenz . . . 109

WiWiSET: Test zur ökonomischen Studieneingangskompetenz . . . 112

Impressum . . . 115

AkaTex Akademische Textkompetenzen: Wissenschaftliches Formulieren und Argumentieren

KURZFASSUNG

Name des Instruments: Ratingverfahren AkaTex

Fachdomäne: Produktion wissenschaftlicher Texte, Sprachdidaktik

Erfasste Kompetenzen: Akademische Textkompetenzen in wissenschaftlichem Formulieren und Argumentieren

Zielgruppe: Bachelorstudierende zu Beginn des Studiums

Testart: Bewertung eines Diskursreferats (Miniaturformat einer Seminararbeit)

Modalität: Computergestützt

Dauer: Schreibphase: 2 Wochen, durchschnittliche Bewertungsdauer (pro Diskursreferat): 20 Minuten

Testaufbau: 13 Items: 7 Rating-Items zur Kategorie „Wissenschaftliches Formulieren", 6 Rating-Items zur Kategorie „Fachlicher Gehalt und Argumentation"

Allgemeiner Testzweck: Erfassung von Kompetenzstand und -entwicklung in der Produktion wissenschaftlicher Texte (wissenschaftliches Formulieren und Argumentieren)

Konkrete Einsatzszenarien: Überprüfung von Leistungsstand und -entwicklung von Studienanfänger*innen

Nicht geeignet für: Schüler*innen vor der Sekundarstufe II

Hinweise für praktische Anwendung: —

Bisheriger Einsatz in Projekten:
AkaTex: http://www.uni-siegen.de/phil/akatex/

Kontakt / Speicherort:
Dr. Lena Decker
Universität Siegen, Fakultät I,
Germanistisches Seminar
E-Mail: decker@germanistik.uni-siegen.de
Postanschrift: Hölderlinstr. 3, 57068 Siegen
Webseite Dr. Lena Decker: https://www.uni-siegen.de/phil/germanistik/mitarbeiter/decker_lena/

ALLGEMEINE INFORMATIONEN

Name des Instruments: Ratingverfahren AkaTex
Bisheriger Einsatz in Projekten:
AkaTex: http://www.uni-siegen.de/phil/akatex/
Kontakt / Speicherort:
Dr. Lena Decker
Universität Siegen, Fakultät I, Germanistisches Seminar
E-Mail: decker@germanistik.uni-siegen.de
Postanschrift: Hölderlinstr. 3, 57068 Siegen
Webseite Dr. Lena Decker: https://www.uni-siegen.de/phil/germanistik/mitarbeiter/decker_lena/

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Akademische Textkompetenzen im wissenschaftlichen Formulieren und Argumentieren
Theoriemodell: „Akademische Textkompetenzen im engeren Sinne" auf Basis der Modelle von Ossner (2008), Becker-Mrotzek & Schindler (2007) und Schindler/ Siebert-Ott (2013)
Testart: Bewertung eines Diskursreferats (Miniaturformat einer Seminararbeit)
Modalität: Computergestützt
Testaufbau:
Itempool: 13 Items; 7 Rating-Items zur Kategorie „Wissenschaftliches Formulieren", 6 Rating-Items zur Kategorie „Fachlicher Gehalt und Argumentation". Teilnehmer*innen verfassen zu einem sprachdidaktischen Thema Diskursreferate, indem sie mehrere Texte zu einer gegebenen Aufgabenstellung auswerten, zueinander in Beziehung setzen und eine eigene Position entwickeln. Diese werden mithilfe des Ratingverfahrens AkaTex von geschulten Rater*innen beurteilt.

TESTGÜTEKRITERIEN

Reliabilität: Interrater-Reliabilität bei 2 unabhängigen Rater*innen „Wissenschaftliches Formulieren" (Cronbachs α = 0.703 – 0.902); „Fachlicher Gehalt und Argumentation" (Cronbachs α = 0.829 – 0.937)
Validität Testinhalt: In Expertenvalidierung (Expert*innen der Schreibforschung) wurde die Operationalisierung der Kriterien der interessierenden Kompetenzen als angemessen und vollständig bewertet.
Validität Antwortprozesse: keine Angaben
Validität interner Testaufbau: Die Kompetenzfacetten „Fachlicher Gehalt" und „Wissenschaftliches Formulieren" hängen statistisch signifikant zusammen.
Validität Zusammenhänge mit anderen Kompetenzindikatoren: keine Angaben
Validität Konsequenzen der Testergebnisse: keine Angaben
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: Schreibphase: 2 Wochen, durchschnittliche Bewertungsdauer (pro Diskursreferat): 20 Minuten
Benötigtes Material: Computer/Laptop
Besonderheiten: Rater*innen-Schulung erforderlich
Praxisbeispiel: Studienanfänger*innen verfassten in Einführungsseminaren zu einem sprachdidaktischen Thema ein Diskursreferat, indem sie mehrere Texte zu einer gegebenen Aufgabenstellung auswerten, zueinander in Beziehung setzen und eine eigene Position entwickeln. Diese wurden mithilfe des Ratingverfahrens AkaTex von geschulten Rater*innen beurteilt. Das Ratingschema wurde in weiterentwickelter Form für die Bewertung von Seminar- und akademischen Abschlussarbeiten in den Lehramtsstudiengängen an der Universität Siegen eingesetzt.

FACHDOMÄNE

Erprobt für: Produktion wissenschaftlicher Texte, Sprachdidaktik
Prinzipiell auch geeignet für: Alle Studienfachdomänen
Nicht geeignet für: —

ZIELGRUPPE

Erprobt für: Bachelorstudierende zu Beginn des Studiums
Prinzipiell auch geeignet für: Schüler*innen der gymnasialen Oberstufe; fortgeschrittene Studierende im Bachelorstudium
Nicht geeignet für: Schüler*innen vor der Sekundarstufe II

ALLGEMEINER TESTZWECK

Erprobt für: Erfassung von Kompetenzstand
Prinzipiell auch geeignet für: Erfassung von Kompetenzentwicklung
Nicht geeignet für: —

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: —
Prinzipiell auch geeignet für: Kognitive Kompetenzen, Noten, Bildungsabschluss
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für Kurse im zeitlichen Verlauf. Die Analysen wurden auf Einzel- und Gruppenebene durchgeführt. Die Stichprobe wurde ausschließlich an einer Universität erhoben.
Prinzipiell auch geeignet für: Einzelne Studierende im zeitlichen Verlauf; Kursebene
Nicht geeignet für: Hochschulebene, Landes- und Bundesebene, internationale Ebene

EINSATZSZENARIEN

Erprobt für: Erfassung von Leistungsstand und -entwicklung der Schreibkompetenzen von Lehramtsstudierenden
Prinzipiell auch geeignet für: Studienerfolgsvorhersage, Kursvergleiche, Hochschulentwicklung
Nicht geeignet für: —

ERLAUBTE KONSEQUENZEN

Erprobt für: Beurteilung der Studienleistung
Prinzipiell auch geeignet für: Prüfungsnoten, Abschlussnoten
Nicht geeignet für: —

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Fachstudium (Master) in Verbindung mit dem Praxissemester („writing in the disciplines" in Verbindung mit „teaching writing and assessing writing")

WEITERE WICHTIGE ANGABEN

Der Texttyp Diskursreferat wird bereits in der gymnasialen Oberstufe als Vorform der Facharbeit gelehrt.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Decker, Lena; Siebert-Ott, Gesa (2019). Schreibend an fachlichen Diskursen partizipieren: Ergebnisse einer Interventionsstudie zur Förderung der Textkompetenzen von Lehramtsstudierenden. In Helmuth Feilke; Katrin Lehnen & Martin Steinseifer (Hrsg.), *Eristische Literalität. Wissenschaftlich streiten – Wissenschaftlich schreiben.* Münster: Waxmann.
Decker, Lena; Siebert-Ott, Gesa (2018). Wissenschaft als diskursive Praxis. Schreibend an fachlichen Diskursen partizipieren. In Sabine Schmölzer-Eibinger; Bora Bushati; Christopher Ebner & Lisa Niederdorfer (Hrsg.), *Wissenschaftliches Schreiben lehren und lernen. Diagnose und Förderung wissenschaftlicher Textkompetenz in Schule und Universität* (S. 193–218). Münster: Waxmann.
Decker, Lena (2016). *Wissenschaft als diskursive Praxis. Schreibend an fachlichen Diskursen partizipieren.* (Kölner Beiträge zur Sprachdidaktik, Bd. 10). Duisburg: Gilles & Francke.

WEITERFÜHRENDE LITERATUR

Decker, Lena & Siebert-Ott, Gesa (2019). Sprachsensible Bildungsräume gestalten. Eine Professionalisierungsaufgabe in der Lehrer/-innenbildung. In Una Dirks & Kathrin Siebold (Hrsg.), *Sprachsensible DaF-/DaZ-/DaM-Bildungsräume: Sprech- & Textformen im Fokus.* Münster: Waxmann.
Decker, Lena & Oehme, Viola (2018). Inhalte angeben – sachtext- und literaturbezogen. *Der Deutschunterricht*, Heft 3/2018, S. 44–55.

ÄKHOM: Facets-of-Competence (FOC): Bewertungsinstrument für ärztliche und pflegerische Bewerter*innen

KURZFASSUNG

Name des Instruments: FOC-Bewertungsinstrument (Facets of Competence) für ärztliche und pflegerische Bewerter*innen

Fachdomäne: Medizin

Erfasste Kompetenzen: Verantwortung; Teamarbeit und Kollegialität; Kennen von persönlichen Grenzen / Möglichkeiten; Struktur und Arbeitsplanung; Umgang mit Fehlern; wissenschaftlich begründete Arbeitsmethoden; verbale Kommunikation mit Kolleg*innen, Supervisor*innen

Zielgruppe: Medizinstudierende ab dem 8. Semester

Testart: Bewertungsinstrument (Fremdevaluation von Medizinstudierenden durch Ärzt*innen und Pflegekräfte; Selbstevaluation)

Modalität: Paper-Pencil-Test (auch Online-Einsatz möglich)

Dauer: 5–10 Minuten

Testaufbau: 4 Items: Je ein Item zu jeder der 7 Kompetenzfacetten pro Fall (5-stufige Skala: 1: ungenügend, 3: akzeptabel, 5: sehr gut; „kein Urteil möglich") plus je ein Item zur Abfrage der Urteilssicherheit (5-stufige Skala: 1: unsicher, 5: sicher). Pro Item gibt es ein Freitextfeld zum Eintragen von Stichworten als Grundlage der Bewertung. Außerdem wird jeder Patient*innenfall ebenfalls mit 7 Items (eins pro Facette) bewertet.

Allgemeiner Testzweck: Erfassen des Kompetenzstands oder der Kompetenzentwicklung

Konkrete Einsatzszenarien: Summativ: Zwischenprüfung, Abschlussprüfung; formativ: vor Beginn oder am Ende des Praktischen Jahres

Nicht geeignet für: —

Hinweise für praktische Anwendung: Ankerbeispiele zur Vereinfachung des Ratings und für verbesserte Standardisierung; eine Raterschulung mit dem Bewertungsinstrument ist erforderlich.

Bisheriger Einsatz in Projekten: Erste Version: UHTRUST (Vorgängerprojekt von ÄKHOM), Weiterentwicklung: ÄKHOM t_1, Anpassung für ÄKHOM t_2

Kontakt / Speicherort:
Prof. Dr. med. Sigrid Harendza, MME
Universitätsklinikum Hamburg-Eppendorf, III. Medizinische Klinik
Postanschrift: Martinistr. 52, 20246 Hamburg
E-Mail: harendza@uke.de
Webseite: https://www.uke.de/kliniken-institute/kliniken/iii.-medizinische-klinik-und-poliklinik/forschung/ag-harendza.html

ALLGEMEINE INFORMATIONEN

Name des Instruments: FOC-Bewertungsinstrument (Facets of Competence) für ärztliche und pflegerische Bewerter*innen
Bisheriger Einsatz in Projekten: Erste Version: UHTRUST (Vorgängerprojekt von ÄKHOM), Weiterentwicklung: ÄKHOM t_1, Anpassung für ÄKHOM t_2
Kontakt / Speicherort:
Prof. Dr. med. Sigrid Harendza, MME
Universitätsklinikum Hamburg-Eppendorf, III. Medizinische Klinik
Postanschrift: Martinistr. 52, 20246 Hamburg
E-Mail: harendza@uke.de
Webseite: https://www.uke.de/kliniken-institute/kliniken/iii.-medizinische-klinik-und-poliklinik/forschung/ag-harendza.html

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Verantwortung; Teamarbeit und Kollegialität; Kennen von persönlichen Grenzen / Möglichkeiten; Struktur und Arbeitsplanung; Umgang mit Fehlern; wissenschaftlich begründete Arbeitsmethoden; verbale Kommunikation mit Kolleg*innen, Supervision
Theoriemodell: Kompetenzbegriff nach Epstein et al. (2002) und Van der Blijet et al. (2002); Kompetenzmodell abgeleitet vom CanMEDs-Modell, Kompetenzentwicklung nach Chen et al. (2015)
Testart: Bewertungsinstrument (Fremdevaluation von Medizinstudierenden durch Ärzt*innen und Pflegekräfte; Selbstevaluation)
Modalität: Paper-Pencil-Test (auch Online-Einsatz möglich)
Testaufbau:
Itempool: 4 Items: Je ein Item zu jeder der 7 Kompetenzfacetten pro Fall (5-stufige Skala: 1: ungenügend, 3: akzeptabel, 5: sehr gut; „kein Urteil möglich") plus je ein Item zur Abfrage der Urteilssicherheit (5-stufige Skala: 1: unsicher, 5: sicher). Pro Item gibt es ein Freitextfeld zum Eintragen von Stichworten als Grundlage der Bewertung. Außerdem wird jeder Patient*innenfall ebenfalls mit 7 Items (eins pro Facette) bewertet.
Testversionen: Ankerbeispiele zur Vereinfachung des Ratings und für verbesserte Standardisierung

TESTGÜTEKRITERIEN

Reliabilität: Interne Konsistenz für Oberärzt*innen =.90; Assisstenzärzt*innen = .80;Pflegekräft e= .78
Validität Testinhalt: Abgleich der (mit FOC) zu bewertenden Kompetenzen mit der Berufsrealität von Ärzt*innen an den teilnehmenden Kliniken. Alle beobachtbaren Verhaltensweisen sind aus der Praxis abgeleitet und werden realitätsnah geprüft.
Validität Antwortprozesse: Der FOC-Bogen erfasst die Stufen des Kompetenzmodells (s.o.).
Validität interner Testaufbau: keine Angaben
Validität Zusammenhänge mit anderen Kompetenzindikatoren: keine Angaben
Validität Konsequenzen der Testergebnisse: Zunächst nur formativer Einsatz mit Feedback an die Studierenden
Testfairness: Ein positives Ethikvotum liegt vor; Genderfairness ist gegeben.

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 5–10 Minuten
Benötigtes Material: Beobachtbare, simulierte Gesprächs- oder Interaktionssituation und ggf. weitere Beobachtungen eines simulierten Arbeitstags
Besonderheiten: Eine Raterschulung mit dem Bewertungsinstrument ist erforderlich; bei der Auswertung gibt es quantitative und qualitative Teile (Ratings mit zugehörigen Begründungsfeldern und Abfrage der Urteilssicherheit).
Praxisbeispiel: Das Instrument wurde zur Bewertung von Kompetenzfacetten bei der Durchführung einer kompetenzbasierten Prüfung am Ende des Medizinstudiums eingesetzt (Projekt ÄKHOM). Die Bewertung fand unter realitätsnahen Bedingungen statt (Simulationssituation eines ersten Arbeitstags im Krankenhaus) und wurde zusammen mit den Ergebnissen anderer Kompetenzerfassungsinstrumente (s.u.) an die 70 bzw. 90 Teilnehmer*innen der Studie zurückgemeldet.

FACHDOMÄNE

Erprobt für: Medizin
Prinzipiell auch geeignet für: Rechtsmedizin, Lehramt
Nicht geeignet für: —

ZIELGRUPPE

Erprobt für: Medizinstudierende ab dem 10. Semester
Prinzipiell auch geeignet für: —
Nicht geeignet für: Medizinstudierende vor dem 8. Semester

ALLGEMEINER TESTZWECK

Erprobt für: Erfassen des Kompetenzzustands zu beliebigem Zeitpunkt ab 8. Semester
Prinzipiell auch geeignet für: Wiederholten Einsatz zum Erfassen der Kompetenzentwicklung
Nicht geeignet für: —

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Wissenstest, Empathie-Bewertung, EPA-Bewertungsbogen (Entrustable professional activities)
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Die Ergebnisse sind statistisch aussagekräftig für Einzelpersonen zu einem Zeitpunkt, Kohorten von Studierenden, verschiedene Hochschulstandorte, international: Der Pilot-Fragebogen wurde in Deutschland und den Niederlanden erprobt (Projekt UHTRUST).
Prinzipiell auch geeignet für: Einsatz im zeitlichen Verlauf
Nicht geeignet für: —

EINSATZSZENARIEN

Erprobt für: Summativ: Zwischenprüfung, Abschlussprüfung; formativ: vor Beginn oder am Ende des Praktischen Jahres
Prinzipiell auch geeignet für: Ärztliche Weiterbildung
Nicht geeignet für: —

ERLAUBTE KONSEQUENZEN

Erprobt für: Kompetenz für den Berufseinstieg
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Das FOC-Bewertungsinstrument ist auf Anfrage beim Projektteam erhältlich. Das Verfahren kann auch formativ und längsschnittlich eingesetzt werden. (Das Instrument wird aktuell zur Prüfung von Verlaufsveränderungen [mehrere Zeitpunkte] eingesetzt.)

WEITERE WICHTIGE ANGABEN

Das FOC-Bewertungsinstrument muss in einer beobachtbaren Situation eingesetzt werden. Eine Raterschulung ist erforderlich.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Prediger S, Fürstenberg S, Berberat PO, Kadmon M, Harendza S. (2019). Interprofessional assessment of medical students' competences with an instrument suitable for physicians and nurses. *BMC Med Educ, 19*(1):46.

WEITERFÜHRENDE LITERATUR

Epstein, R. M. & Hundert, E. M. (2002). Defining and assessing professional competence. *JAMA, 287*, S. 226-235.
Van der Blij, M.; Boon, J.; Van Lieshout, H.; Schafer, H. & Schrijen, H. (2002). *Competentieprofielen: over schillen en knoppen.* [e-Competence profiles]. Utrecht: Digitale Universiteit.
Chen, H.C.; van den Broek, W.E. & Ten Cate, O. (2015). The case for use of entrustable professional activities in undergraduate medical education. *Acad Med., 90*(4), S. 431-436.

AstraLite-Koswo AST: Argumentstrukturtest

KURZFASSUNG

Name des Instruments: Argumentstrukturtest (AST)

Fachdomäne: Human- und Sozialwissenschaften, empirisch bzw. evidenzbasiert argumentierende Wissenschaften

Erfasste Kompetenz: Epistemisch-systematische Lesekompetenzen: Erkennen und Zuordnen von funktionalen Argumentbestandteilen

Zielgruppe: Bachelor- und Masterstudierende zu Beginn und im Verlauf des Studiums

Testart: Leistungstest; Zuordnungsaufgaben

Modalität: Computergestützt

Dauer: 30 Minuten

Testaufbau: 40 Items; Korrekte Zuordnung von 5 Argumentbestandteilen in 8 Argumenten nach Toulmin (1985). Schwierigkeitsvariation durch unterschiedliche Anzahl und Position der Argumentbestandteile

Allgemeiner Testzweck: Erfassung von Kompetenzstand und -entwicklung von Gruppen

Konkrete Einsatzszenarien: Gruppendiagnostik, Evaluation von Lehr-Lern- und Trainingsmaßnahmen, Vorhersage des Studienerfolgs

Nicht geeignet für: Individualdiagnostik, internationale Vergleiche, Auswahlentscheidungen, Nichtmuttersprachler*innen

Hinweise für praktische Anwendung: Kontrolle der Muttersprache bzw. der Sprachkompetenz

Bisheriger Einsatz in Projekten: KOSWO, AstraLite, https://www.uni-regensburg.de/psychologie-paedagogik-sport/paedagogik-1/projekte/index.html

Kontakt / Speicherort:
Prof. Dr. Tobias Richter
Institut für Psychologie, Lehrstuhl Psychologie IV,
Universität Würzburg, Deutschland
Webseite: https://go.uniwue.de/edpsych

Dr. Sebastian Schmid
Institut für Pädagogik, Universität Regensburg, Deutschland
Webseite: https://www.uni-regensburg.de/psychologie-paedagogik-sport/paedagogik-1/team/dr-sebastian-schmid/index.html

ALLGEMEINE INFORMATIONEN

Name des Instruments: Argumentstrukturtest (AST)
Bisheriger Einsatz in Projekten: Kompetenzen Studierender im Umgang mit wissenschaftlicher Originalliteratur (KOSWO), Assessment und Training von Scientific Literacy (AstraLite), https://www.uni-regensburg.de/psychologie-paedagogik-sport/paedagogik-1/projekte/index.html
Kontakt / Speicherort:
Prof. Dr. Tobias Richter, Institut für Psychologie, Lehrstuhl Psychologie IV, Universität Würzburg, https://go.uniwue.de/edpsych
Dr. Sebastian Schmid, Institut für Pädagogik, Universität Regensburg, https://www.uni-regensburg.de/psychologie-paedagogik-sport/paedagogik-1/team/dr-sebastian-schmid/index.html

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Epistemisch-systematische Lesekompetenzen: Erkennen und Zuordnen von funktionalen Argumentbestandteilen
Theoriemodell: Argumentationsmodell von Toulmin (1958)
Testart: Leistungstest; Zuordnungsaufgaben
Modalität: Computergestützt
Testaufbau:
Itempool: 8 Argumente, korrekte Zuordnung der 5 Argumentbestandteile nach Toulmin; 40 Items verfügbar, die sich unterscheiden nach: Anzahl der Bestandteile (3 = einfaches Argument vs. alle 5 = komplexes Argument), Position der Bestandteile (Argument beginnt mit Behauptung vs. Begründung).

TESTGÜTEKRITERIEN

Reliabilität: Cronbachs $\alpha = .76$ (40 Items)
Validität Testinhalt: Der AST misst curricular relevante Kompetenzen beim Lesen wissenschaftlicher Originalliteratur. Diese Kompetenzen werden in nahezu allen wissenschaftlich arbeitenden Studiengängen gefordert.
Validität Antwortprozesse: Analysen einer Blickbewegungsstudie ($N = 60$) zeigen, dass komplexere Argumente länger und häufiger gelesen werden.
Validität interner Testaufbau: Theoriekonform wurden einfache Argumente häufiger richtig gelöst als komplexe Argumente; Argumente, die mit der Behauptung beginnen, häufiger als Argumente, die mit Begründung beginnen.
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Moderate positive Zusammenhänge mit verbaler Intelligenz ($r = .40$) und der Schulabschlussnote ($r = .17$).
Validität Konsequenzen der Testergebnisse: Das Instrument ermöglicht eine Einschätzung des Lehr-Lernbedarfs zur Förderung der epistemisch-systematischen Lesekompetenzen.

Testfairness: Keine Geschlechterunterschiede. Sprachunterschiede: Die Aufgaben enthalten viel akademische Sprache. Bei Nichtmuttersprachler*innen kann sich das Sprachniveau auf das Testergebnis auswirken.

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 30 Minuten
Benötigtes Material: Computer/Laptop, Teststeuerungsprogramm
Besonderheiten: Auswertung: Testergebnis ist der Summenwert korrekt gelöster Items.
Praxisbeispiel: Der AST wurde an zwei deutschen Universitäten in drei verschiedenen Parallelformen in fünf Untersuchungen eingesetzt, die eine zufriedenstellende Reliabilität und Validität des Testinstruments zeigen. Zudem wurde der AST in zwei Studien zusammen mit einem Test zur Beurteilung der Plausibilität von wissenschaftlichen Argumenten vor und nach einer Lehrmaßnahme zur Verbesserung der epistemisch-systematischen Lesekompetenzen eingesetzt, um die Wirksamkeit der Intervention zu überprüfen (von der Mühlen, Richter, Schmid & Berthold, 2018).

FACHDOMÄNE

Erprobt für: Human- und Sozialwissenschaften
Prinzipiell auch geeignet für: Empirisch bzw. evidenzbasiert argumentierende Wissenschaften
Nicht geeignet für: Rein hermeneutische Geisteswissenschaften

ZIELGRUPPE

Erprobt für: Studierende zu Beginn des Studiums
Prinzipiell auch geeignet für: Studierende im Verlauf des Studiums
Nicht geeignet für: Wissenschaftler*innen

ALLGEMEINER TESTZWECK

Erprobt für: Forschungs- und Evaluationszwecke
Prinzipiell auch geeignet für: Erfassung von Kompetenzstand und -entwicklung im Studium
Nicht geeignet für: Erfassung allgemeiner Kompetenz im Argumentieren

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Tests zur Bewertung der Plausibilität von Argumenten, verbaler Intelligenz, soziodemografische Angaben (wie Muttersprache, Geschlecht), Noten
Prinzipiell auch geeignet für: Andere Maße der Kompetenzmessung
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für den Vergleich von Studierenden, auch über den zeitlichen Verlauf.
Prinzipiell auch geeignet für: Vergleich von Kursen, Hochschulen, Länder- und Bundesebene
Nicht geeignet für: Individualdiagnostik, internationale Vergleiche

EINSATZSZENARIEN

Erprobt für: Erfassung des aktuellen Kompetenzstands; Evaluation von Lehr-Lernmaßnahmen (Prä-Post-Messung)
Prinzipiell auch geeignet für: Vorhersage des Studienerfolgs
Nicht geeignet für: Auswahlentscheidungen

ERLAUBTE KONSEQUENZEN

Erprobt für: Einschätzung eines Lehr-Lernbedarfs
Prinzipiell auch geeignet für: Förderung des Umgangs mit wissenschaftlichen Originalarbeiten im Studium; Verbesserung der Scientific Literacy
Nicht geeignet für: Studienzulassung, Studienwahlempfehlung

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Zur Verbesserung der Lesekompetenzen von Studierenden soll der AST zu Beginn des Studiums als Teil einer Trainingsintervention eingesetzt und später im universitären Curriculum etabliert werden.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Münchow, H.; Richter, T.; von der Mühlen, S.; Schmid, S.; Bruns, K. & Berthold, K. (2020). Verstehen von Argumenten in wissenschaftlichen Texten: Reliabilität und Validität des Argumentstrukturtests (AST). *Diagnostica*.

WEITERFÜHRENDE LITERATUR

von der Mühlen, S.; Richter, T.; Schmid, S. & Berthold, K. (2018). How to improve argumentation comprehension in university students: Experimental test of a training approach. *Instructional Science, 47(2)*, 215–237. https://doi.org/10.1007/s11251-018-9471-3
Münchow, H., Richter, T. & Schmid, S. (2020). What does it take to deal with academic literature? Epistemic components of scientific literacy. In O. Zlatkin-Troitschanskaia, H. A. Pant, M. Toepper & C. Lautenbach (Eds.). *Student learning in German higher education: Innovative measurement approaches and research results.* Wiesbaden, Germany: Springer.

Berliner Test zur Erfassung Fluider und Kristalliner Intelligenz (BEFKI)

KURZFASSUNG

Name des Instruments: Berliner Test zur Erfassung Fluider und Kristalliner Intelligenz (BEFKI)

Fachdomäne: Fachübergreifend einsetzbar

Erfasste Kompetenzen: Fluide Intelligenz und Kristalline Intelligenz

Zielgruppe: Schüler*innen der Klassenstufen 3–13, Erwachsene unterschiedlicher Altersgruppen einschließlich Studierender

Testart: Leistungstest

Modalität: Paper-Pencil-Test; auch computergestützt

Dauer: 70 Minuten (bei Einsatz aller Subtests)

Testaufbau: 112 Items zu 2 übergeordneten Dimensionen mit je 3 Teildimensionen: 1. Fluide Intelligenz (48 Items zu verbalem, numerischem und figuralem schlussfolgernden Denken) und 2. Kristalline Intelligenz (64 Items zu natur-, geistes- und sozialwissenschaftlichem Wissen)

Allgemeiner Testzweck: Erfassen kognitiver Fähigkeiten (fluide und kristalline Intelligenz) einmalig oder im zeitlichen Verlauf

Konkrete Einsatzszenarien: Grundlagenforschung; Kontrollvariable bei Untersuchungen im Schulkontext; Diagnostik in der Schullaufbahn- und Berufsberatung; Vorhersage des Studienerfolgs

Nicht geeignet für: —

Hinweise für praktische Anwendung: Entwickelt für den Einsatz bei Schüler*innen; erfolgreich bei Studierenden bzw. Erwachsenen eingesetzt. Einzel- oder Gruppentestungen sind möglich.

Bisheriger Einsatz in Projekten: Eingesetzt zur Messung kognitiver Grundfähigkeiten in Studien zu sprachlichen, mathematischen und naturwissenschaftlichen Kompetenzen auf Basis der Bildungsstandards der KMK (Pilotierungs-, Normierungs- und Bildungstrendstudien des IQB); PISA-Studien seit 2012; wissenschaftliche Begleitung der Hector Kinderakademien (Hector-Institut für Empirische Bildungsforschung/ DIPF Frankfurt); Studien mit Studierenden (WiWiSET, https://www.wiwi-kompetenz.de/wiwiset-2016-2019/), (WiWiKom, https://www.wiwi-kompetenz.de/wiwikom-ii-2015-2019/)

Kontakt / Speicherort:
Dr. Stefan Schipolowski
Humboldt-Universität zu Berlin,
Institut zur Qualitätsentwicklung im Bildungswesen
E-Mail: stefan.schipolowski@iqb.hu-berlin.de
Postanschrift: Unter den Linden 6, 10099 Berlin

ALLGEMEINE INFORMATIONEN:

Name des Instruments: Berliner Test zur Erfassung Fluider und Kristalliner Intelligenz (BEFKI)
Bisheriger Einsatz in Projekten: Eingesetzt zur Messung kognitiver Grundfähigkeiten in Studien zu sprachlichen, mathematischen und naturwissenschaftlichen Kompetenzen auf Basis der Bildungsstandards der KMK (Pilotierungs-, Normierungs- und Bildungstrendstudien des IQB); PISA-Studien seit 2012; wissenschaftliche Begleitung der Hector Kinderakademien (Hector-Institut für Empirische Bildungsforschung/ DIPF Frankfurt); Studien mit Studierenden (WiWiSET (https://www.wiwi-kompetenz.de/wiwiset-2016-2019/), WiWiKom (https://www.wiwi-kompetenz.de/wiwikom-ii-2015-2019/))

Kontakt / Speicherort:
Dr. Stefan Schipolowski
Humboldt-Universität zu Berlin,
Institut zur Qualitätsentwicklung im Bildungswesen
E-Mail: stefan.schipolowski@iqb.hu-berlin.de
Postanschrift: Unter den Linden 6, 10099 Berlin

DAS INSTRUMENT

INHALT UND STRUKTUR:

Erfasste Kompetenzen: Fluide Intelligenz (gf) mit verbalem, numerischem und figuralem schlussfolgernden Denken; kristalline Intelligenz (gc) mit den Domänen Natur-, Geistes- und Sozialwissenschaften
Theoriemodell: Intelligenzstrukturtheorien von Carroll (1993; Drei-Stratum-Theorie), Horn und Cattell (Horn & Noll, 1997; gf-gc-Theorie) und McGrew (2009; CHC-Theorie)
Testart: Leistungstest
Modalität: Paper-Pencil-Test; auch computergestützt
Testaufbau:
112 Items zu 2 übergeordneten Dimensionen mit je 3 Teildimensionen: Figurale Intelligenz/gf (48 Items) mit verbalem, numerischem und figuralem schlussfolgernden Denken (jeweils 16 Items); kristalline Intelligenz/gc (64 Items) mit naturwissenschaftlichem (6 Inhaltsbereiche, 24 Items), geisteswissenschaftlichem (5 Inhaltsbereiche, 20 Items) und sozialwissenschaftlichem Wissen (5 Inhaltsbereiche, 20 Items)

TESTGÜTEKRITERIEN

Reliabilität: Psychometrische Anforderungen an Schwierigkeit, Trennschärfen und Reliabilität werden erfüllt. Interne Konsistenz bei Jugendlichen in der Oberstufe (Testform BEFKI 11+): Für die Gesamtskala gf: α = .79; für die Gesamtskala gc: α = .81. Interne Konsistenz bei Studierenden im Fach Psychologie (Testform BEFKI 11+): Für die Gesamtskala gf: α = .79; für die Gesamtskala gc: α = .76
Validität Testinhalt: Die Itemtypen und –inhalte wurden anhand einschlägiger Literatur und Vorerprobungen ausgewählt.
Validität Antwortprozesse: keine Angaben
Validität interner Testaufbau: Die Prüfung von Messmodellen zur internen Struktur anhand der Normierungsdaten von Kindern und Jugendlichen zeigt theoriekonforme Ergebnisse. Für Studierende belegen empirische Studien die Eindimensionalität der Skala zum figuralen schlussfolgernden Denken.
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Diskriminante und konvergente Validität wurden mit anderen Maßen kognitiver Fähigkeiten geprüft: Kognitiver Fähigkeits-Test KFT 4-12+R; Wortschatztest WST; Kompetenztests auf Basis der Bildungsstandards der KMK; Schulnoten; Studiennoten; Mittelwertunterschiede zwischen geprüften Gruppen entsprechen den Erwartungen (wie verschiedene Jahrgangsstufen, besuchte Schulart, Geschlecht).
Validität Konsequenzen der Testergebnisse: Vorhersage individueller Fähigkeitsausprägungen in der Berufsberatung oder im Studium
Testfairness: Messinvarianz-Analysen zu Geschlecht und besuchter Schulart mittels konfirmatorischer Faktorenanalysen bestätigen weitgehend die Annahme strikter Messinvarianz und somit eines fairen Instruments bezüglich der untersuchten Intelligenzkonstrukte.

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 70 Minuten (bei Einsatz aller Untertests)
Benötigtes Material: Druckvorlagen der Testitems, Auswertungsschlüssel
Besonderheiten: 4 verschiedene Testversionen für unterschiedliche Jahrgangsstufen bzw. Altersgruppen und je 2 Versionen des Aufgabenhefts mit umgedrehter Reihenfolge der Items bzw. Antworten sind verfügbar. Einzel- oder Gruppentestungen sind möglich. Die Testverfahren wurden für den Einsatz bei Schüler*innen entwickelt und normiert; die Testversion für die Oberstufe wurde auch erfolgreich bei Erwachsenen (einschl. Studierenden) eingesetzt.
Praxisbeispiel: Normierungsstudien mit insgesamt 18.000 Schüler*innen der Jahrgangsstufen 3 bis 12 in allen deutschen Bundesländern wurden durchgeführt; zudem Erprobungsstudien mit Studierenden verschiedener Studiengänge (u. a. 9.000 Studierenden der Sozial- und Wirtschaftswissenschaften, 400 Studierenden im Fach Psychologie sowie 240 Studierenden im Fach Chemie/Lehramt).

FACHDOMÄNE

Erprobt für: Schulkontext; Psychologie, Wirtschaftswissenschaften, Sozialwissenschaften, Chemie/Lehramtsstudium
Prinzipiell auch geeignet für: Fachübergreifend einsetzbar
Nicht geeignet für: —

ZIELGRUPPE

Erprobt für: Schüler*innen der Klassenstufen 3 bis 12/13. Die Testversion für die Mittelstufe und Oberstufe wurde in mehreren Studien auch erfolgreich bei Erwachsenen (einschl. Studierenden) eingesetzt.
Prinzipiell auch geeignet für: Erwachsene unterschiedlicher Altersgruppen einschließlich Studierende (fachübergreifend)
Nicht geeignet für: —

ALLGEMEINER TESTZWECK

Erprobt für: Erfassung kognitiver Fähigkeiten (fluide und kristalline Intelligenz) in quer- und längsschnittlichen Designs
Prinzipiell auch geeignet für: Kognitive Fähigkeiten sind für zahlreiche Outcomes wie Studien- und Berufserfolg prädiktiv valide und werden in der Studien- und Berufsberatung genutzt.
Nicht geeignet für: Zum Erfassen weiterer kognitiver Fähigkeitsfaktoren (z. B. mental speed) sind zusätzliche Testverfahren erforderlich; spezifisches Fachwissen bestimmter Studienfächer wird mit dem Allgemeinwissenstest nicht erfasst.

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Eingesetzt zur Messung kognitiver Grundfähigkeiten in Studien zu sprachlichen, mathematischen und naturwissenschaftlichen Kompetenzen auf Basis der Bildungsstandards der KMK (Pilotierungs-, Normierungs- und Bildungstrendstudien des IQB); PISA-Studien seit 2012; wissenschaftliche Begleitung der Hector Kinderakademien (Hector-Institut für Empirische Bildungsforschung/DIPF Frankfurt); Studien mit Studierenden (WiWiSET, https://www.wiwi-kompetenz.de/wiwiset-2016-2019/), (WiWiKom, https://www.wiwi-kompetenz.de/wiwikom-ii-2015-2019/)
Prinzipiell auch geeignet für: Keine Einschränkung
Nicht geeignet für: Keine Einschränkung

EIGNUNG

VERGLEICHSEBENEN:

Erprobt für: Individual- und Gruppenebene; auch im zeitlichen Verlauf
Prinzipiell auch geeignet für: Versionen in anderen Sprachen (Englisch, Arabisch, Polnisch, Russisch) werden erprobt
Nicht geeignet für: —

EINSATZSZENARIEN

Erprobt für: Grundlagenforschung; schlussfolgerndes Denken als Kontrollvariable bei Untersuchungen im Schulkontext (PISA-Studien und IQB-Bildungstrends); Vorhersage des Studienerfolgs
Prinzipiell auch geeignet für: Als Kontrollvariablen für zahlreiche Fragestellungen sowie prädiktiv für verschiedene Outcomes wie Studien- und Berufserfolg; Nutzung in der Studien- und Berufsberatung
Nicht geeignet für: Tests zu allgemeinen kognitiven Fähigkeiten sollten je nach Fragestellung um Tests zu spezifischeren Konstrukten (z. B. Fachwissenstests) ergänzt werden.

ERLAUBTE KONSEQUENZEN

Erprobt für: —
Prinzipiell auch geeignet für: Empfehlungen zur Schul-, Studien- und Berufswahl; Bewerberauswahl (jeweils in Kombination mit weiteren Informationen und Ergebnissen anderer Testverfahren)
Nicht geeignet für: Differenzierte Beurteilung fachspezifischer Kompetenzen (z. B. Studienleistung)

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Der Test wird in weiteren Sprachen (Englisch, Arabisch, Polnisch, Russisch) erprobt. Die Testverfahren wurden z. T. im Hogrefe-Verlag veröffentlicht. Noch nicht veröffentlichte Testversionen können ggf. in Absprache mit den Testentwickler*innen eingesetzt werden.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Wilhelm, O.; Schroeders, U. & Schipolowski, S. (2014). *Berliner Test zur Erfassung fluider und kristalliner Intelligenz für die 8. bis 10. Jahrgangsstufe (BEFKI 8-10).* Göttingen: Hogrefe.
Schipolowski, S.; Wilhelm, O. & Schroeders, U. (2020). *Berliner Test zur Erfassung fluider und kristalliner Intelligenz für die 11. und 12. Jahrgangsstufe (BEFKI 11+).* Göttingen: Hogrefe.
Schroeders, U.; Schipolowski, S. & Wilhelm, O. (2015). Age-related changes in the mean and covariance structure of fluid and crystalized intelligence in childhood and adolescence. *Intelligence, 48*, S. 15–29. https://doi.org/10.1016/j.intell.2014.10.006

WEITERFÜHRENDE LITERATUR

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* Cambridge: Cambridge University Press.
Horn, J. L. & Noll, J. (1997). Human cognitive capabilities: Gf-Gc theory. In D. P. Flanagan; J. L. Genshaft & P. L. Harrison (Hrsg.), *Contemporary intellectual assessment: Theories, tests and issues* (S. 53–91). New York: Guilford.
McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, 37*, S. 1–10. https://doi.org/10.1016/j.intell.2008.08.004

BilWiss-2.0-Test: Bildungswissenschaftliches Wissen von Lehrkräften

KURZFASSUNG

Name des Instruments: BilWiss-2.0-Test

Fachdomäne: Lehrerbildung

Erfasste Kompetenzen: Bildungswissenschaftliches Wissen von (angehenden) Lehrkräften

Zielgruppe: Lehrkräfte in Ausbildung (Lehramtsstudium, Vorbereitungsdienst), Lehrkräfte im Beruf

Testart: Leistungstest (Multiple-Choice-Test mit Einfach- und Mehrfachwahlaufgaben)

Modalität: Paper-Pencil-Test; Online-Version möglich

Dauer: 45 Minuten in der Kurzversion; ca. 15–20 Minuten pro Subskala in der Langversion

Testaufbau: 119 geschlossene Multiple-Choice-Items (Kurzversion: 65 Items) in den 6 Inhaltsbereichen (Skalen): Unterrichtsgestaltung; Lernen & Entwicklung; Diagnostik & Evaluation; Bildungstheorie; Schule als Bildungsinstitution; Lehrerberuf als Profession. In der Langversion jeweils 23, 24, 21, 14, 19, 18 Items pro Skala; in der Kurzversion jeweils 15, 10, 12, 9, 9, 10 Items pro Skala.

Allgemeiner Testzweck: Untersuchung des Wissensstands und der Wissensentwicklung von Gruppen von Lehrkräften

Konkrete Einsatzszenarien: Vorhersage der Leistungsentwicklung, (Kurs-)Vergleiche, Studienerfolgsvorhersage, Studienabbruchvorhersage, (Hochschul-)Evaluationen, Formatives Assessment

Nicht geeignet für: Individualdiagnostik, Zulassungsentscheidungen, Prüfungen, Akkreditierungen

Hinweise für praktische Anwendung: Der Einsatz der Langversion innerhalb einer Testsitzung sollte aufgrund der Testlänge vermieden werden; empfohlen wird der Einsatz einzelner Skalen der Langversion oder der Einsatz der Kurzversion. Der Test, Informationen zum Ablauf der Testung sowie zur Auswertung können beim BilWiss-Projektteam am Standort Frankfurt angefragt werden.

Bisheriger Einsatz in Projekten: BilWiss-UV;
Webseite: http://bilwiss.paedpsych.de

Kontakt / Speicherort:
Simone Emmenlauer
Goethe-Universität Frankfurt
Theodor-W.-Adorno-Platz 6
60629 Frankfurt am Main
E-Mail: emmenlauer@paed.psych.uni-frankfurt.de

Forschungsdaten am FDZ des IQB:
Webseite: https://www.iqb.hu-berlin.de/fdz/studies/BilWiss

ALLGEMEINE INFORMATIONEN

Name des Instruments: BilWiss-2.0-Test
Bisheriger Einsatz in Projekten: BilWiss-UV
Kontakt / Speicherort:
Webseite: http://bilwiss.paedpsych.de
Forschungsdaten am FDZ des IQB:
Webseite: https://www.iqb.hu-berlin.de/fdz/studies/BilWiss

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Bildungswissenschaftliches Wissen von (angehenden) Lehrkräften
Theoriemodell: Modell der professionellen Kompetenz von Lehrkräften (Baumert & Kunter, 2006); Klassifikation des Professionswissens von Lehrkräften nach Schulman (1986); Klassifikation bildungswissenschaftlicher Themen auf der Grundlage der BilWiss-Delphi-Studie (Kunina-Habenicht et al., 2012)
Testart: Leistungstest (Multiple-Choice-Test mit Einfach- und Mehrfachwahlaufgaben)
Modalität: Paper-Pencil-Test; Online-Version über die Plattform Unipark
Testaufbau:
Itempool: 119 geschlossene Multiple-Choice-Items in den 6 Inhaltsbereichen (Skalen): Unterrichtsgestaltung (23 Items), Lernen & Entwicklung (24 Items), Diagnostik & Evaluation (21 Items), Bildungstheorie (14 Items), Schule als Bildungsinstitution (19 Items), Lehrerberuf als Profession (18 Items)
Testversionen: Langversion (119 Items; Einsatz einzelner Skalen möglich); Kurzversion (65 Items, jeweils 15, 10, 12, 9, 9, 10 Items pro o.g. Skala)

TESTGÜTEKRITERIEN

Reliabilität: EAP/PV-Reliabilitäten der Skalen aus der IRT-Skalierung:
Langversion: Unterrichtsgestaltung (.72), Lernen & Entwicklung (.52), Diagnostik & Evaluation (.59), Bildungstheorie (.65), Schule als Bildungsinstitution (.62), Lehrerberuf als Profession (.61)
Kurzversion: Unterrichtsgestaltung (.71), Lernen & Entwicklung (.49), Diagnostik & Evaluation (.57), Bildungstheorie (.64), Schule als Bildungsinstitution (.60), Lehrerberuf als Profession (.59).
Die EAP/PV Reliabilität des gesamten Kurztests beträgt .88.
Interne Konsistenz für den Kurztest: Cronbachs $\alpha = .86$
Validität Testinhalt: Delphi-Studie mit Expert*innen aus der ersten und zweiten Phase der Lehrerbildung in NRW lieferte bildungswissenschaftliche Themen als Grundlage der Testentwicklung; große Abdeckung bildungswissenschaftlicher Themen aus der Delphi-Studie in der Langversion; Ländervertretungen der Kultusministerien bestätigen Relevanz der Testinhalte für universitäre Lehrerbildung in den jeweiligen Bundesländern; hohe inhaltliche Passung der Testitems mit Themen der KMK-Standards zur Lehrerbildung in den Bildungswissenschaften
Validität Antwortprozesse: In kognitiven Interviews konnten die Testaufgaben mit Wissen aus Veranstaltungen der universitären Lehrerbildung beantwortet werden.
Validität interner Testaufbau: Die sechs o.g. Inhaltsbereiche wurden in konfirmatorischen Faktorenanalysen bestätigt.
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Prüfung der Zusammenhänge zwischen bildungswissenschaftlichem Wissen und relevanten Outcomevariablen: Puffernde Wirkung für emotionale Erschöpfung im Vorbereitungsdienst; Wissen im Bereich Schulorganisation sagt proaktives Engagement im Lehrerberuf vorher; Bildungswissenschaftliches Wissen sagt die Qualität der Analyse von Unterricht vorher.
Validität Konsequenzen der Testergebnisse: Das Instrument ist nicht zur Individualdiagnostik geeignet, sondern wird ausschließlich für den Einsatz in der Forschung (z. B. für Gruppenvergleiche) und zum Bildungsmonitoring empfohlen.
Testfairness: Der Wissenstest ist sprachlastig. Personen mit nicht deutscher Muttersprache schneiden signifikant schlechter ab.

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 45 Minuten in der Kurzversion
Benötigtes Material: Papier und Stift bzw. Computer/Laptop/Tablet/Smartphone mit Internetzugang
Besonderheiten: Der Einsatz der Langversion innerhalb einer Testsitzung sollte aufgrund der Testlänge vermieden werden; empfohlen wird der Einsatz einzelner Skalen der Langversion oder der Einsatz der Kurzversion. Der Test, Informationen zum Ablauf der Testung sowie zur Auswertung können beim BilWiss-Projektteam am Standort Frankfurt angefragt werden. Eine Hinterlegung des Tests im Forschungsdatenzentrum am DIPF wird vorbereitet. Für die Auswertung sind eine 2-PL-IRT-Skalierung sowie eine Skalierung nach der klassischen Testtheorie möglich.
Praxisbeispiel: Anhand psychometrischer Kennwerte wurden die Aufgaben aus Vorgängerprojekten modifiziert oder neu generiert und 2016 in einer Pilotstudie mit 555 Lehramtsstudierenden erprobt. Die Passung der Testitems zum bildungswissenschaftlichen Studium wurde 2016 bei 60 Lehramtsstudierenden in kognitiven Interviews beurteilt. Nach Itemoptimierung und -auswahl erfolgte ein Feldtest (N = 600 Lehramtsanwärter*innen) und die Entwicklung der Finalversion. Seit 2017 wird eine Längsschnittstudie mit Lehramtsstudierenden an vier Hochschulen aus vier Bundesländern (zu Beginn N = 928) durchgeführt, um die Veränderung im bildungswissenschaftlichen Wissen bei Besuch inhaltlich passender Lerngelegenheiten zu erfassen.

FACHDOMÄNE

Erprobt für: Lehrerbildung – Bildungswissenschaftlicher (= fachunabhängiger) Teil der Lehramtsausbildung / Lehrerberuf
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

ZIELGRUPPE

Erprobt für: Lehrkräfte in Ausbildung (Lehramtsstudium, Vorbereitungsdienst) in vier Bundesländern (NRW, Hessen, Bayern, BaWü)
Prinzipiell auch geeignet für: Lehrkräfte im Beruf (deutschlandweit), Lehrkräfte in Ausbildung (deutschlandweit)
Nicht geeignet für: Personen ohne ausreichende Deutschkenntnisse

ALLGEMEINER TESTZWECK

Erprobt für: Untersuchung des Wissensstands und der Wissensentwicklung von Gruppen von Lehrkräften
Prinzipiell auch geeignet für: Interventionsevaluationen
Nicht geeignet für: Individualdiagnostik

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Andere Kompetenzaspekte (Überzeugungen, Motivation, Selbstregulation), Berufserfolgskriterien, kristalline und fluide Intelligenz
Prinzipiell auch geeignet für: Weitere Wissenstests (z. B. Fachwissen, fachdidaktisches Wissen)
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Die Ergebnisse sind statistisch aussagekräftig für die Kursebene / Hochschulebene.
Prinzipiell auch geeignet für: Landes- und Bundesebene
Nicht geeignet für: Einzelpersonen, internationale Vergleiche

EINSATZSZENARIEN

Erprobt für: Vorhersage der Leistungsentwicklung, (Kurs-)Vergleiche
Prinzipiell auch geeignet für: Studienerfolgsvorhersage, Studienabbruchvorhersage, (Hochschul-)Evaluationen, Formatives Assessment, Bildungsmonitoring
Nicht geeignet für: Zulassungsentscheidungen, Prüfungen, Akkreditierungen

ERLAUBTE KONSEQUENZEN

Erprobt für: Maßnahmen zur Verbesserung der Lehre
Prinzipiell auch geeignet für: Formatives Feedback
Nicht geeignet für: Zulassungsentscheidungen, Akkreditierungen

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Der Test wird regelmäßig eingesetzt, um die Entwicklung des bildungswissenschaftlichen Wissens bei Besuch von Lerngelegenheiten und für verschiedene Gruppen/Standorte zu untersuchen. Zudem wird der Einfluss des bildungswissenschaftlichen Wissens für die schulische Berufspraxis untersucht. Der Test kann für eine Nutzung bei den Projektmitarbeiterinnen am Standort Frankfurt angefragt und die Daten des BilWiss-Forschungsprogramms können beim Forschungsdatenzentrum des IQB zur Nachnutzung angefragt werden.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Kunina-Habenicht, O.; Lohse-Bossenz, H.; Kunter, M.; Dicke, T.; Förster, D.; Gößling, J. et al. (2012). Welche bildungswissenschaftlichen Inhalte sind wichtig in der Lehrerbildung? *Zeitschrift für Erziehungswissenschaft, 15*, S. 649–682. DOI: 10.1007/s11618-012-0324-6

Kunina-Habenicht, O., Maurer, C., Wolf, K., Holzberger, D., Schmidt, M., Dicke, T. et al. (2020). Der BilWiss-2.0-Test: Ein revidierter Test zur Erfassung des bildungswissenschaftlichen Wissens von (angehenden) Lehrkräften. *Diagnostica*

Kunter, M.; Kunina-Habenicht, O.; Baumert, J.; Dicke, T.; Holzberger, D.; Lohse-Bossenz, H. et al. (2016). Bildungswissenschaftliches Wissen und professionelle Kompetenz in der Lehramtsausbildung – Ergebnisse des Projekts BilWiss. In C. Gräsel & K. Trempler (Hrsg.), *Entwicklung von Professionalität pädagogischen Personals. Interdisziplinäre Betrachtungen, Befunde und Perspektiven* (S. 37–54). Wiesbaden: Springer-Online.

Linninger, C.; Kunina-Habenicht, O.; Emmenlauer, S.; Dicke, T.; Schulze-Stocker, F.; Leutner, D. et al. (2015). Assessing teachers' educational knowledge: Construct specification and validation using mixed methods. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 47*, S. 72–83. DOI: 10.1026/0049-8637/a000126

WEITERFÜHRENDE LITERATUR

Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft, 9*, S. 469–520. DOI: 10.1007/s11618-006-0165-2

DaZKom: Deutsch-als-Zweitsprache-Kompetenz von Lehrkräften

KURZFASSUNG

Name des Instruments: DaZKom-Test

Fachdomäne: Deutsch-als-Zweitsprache in der Lehrerbildung aller Schulfächer und Schulformen

Erfasste Kompetenzen: Deutsch als Zweitsprache (DaZ)-Kompetenz bei angehenden und praktizierenden Lehrkräften

Zielgruppe: Angehende und praktizierende Lehrkräfte

Testart: Kompetenzmessung

Modalität: Paper-Pencil-Test sowie computergestützt

Dauer: 40 Minuten (Kurzversion); 60 Minuten (Langversion)

Testaufbau: Stimuli mit zwei bis sechs Items; zusätzlich Abfrage von soziodemographischen Daten/DaZ-relevanten Lerngelegenheiten

Allgemeiner Testzweck: Assessment-Tool für Seminarveranstaltungen der Lehramtsausbildung, Fort- und Weiterbildungen für Lehrkräfte im Bereich DaZ

Konkrete Einsatzszenarien: Prä-post-Messungen zur Evaluation von DaZ-relevanten Lerngelegenheiten

Nicht geeignet für: Auswahlentscheidungen (Zulassung zu Studiengängen)

Hinweise für praktische Anwendung: Testleitermanual; zur Auswertung wird ein Codiermanual mit Punktesystem (0–2 Punkte) bereitgestellt.

Bisheriger Einsatz in Projekten: National: Sprachen-Bilden-Chancen, ProfaLe, BiProfessional, Weiterbildungsinitiative NRW

International: eCALLMS/ICMEE (University of Nebraska-Lincoln, USA), DivEd (University of Turku, Finnland)

Kontakt / Speicherort:
Prof. Dr. Barbara Koch-Priewe
E-Mail: bkoch-priewe@uni-bielefeld.de
Webseite: http://www.uni-bielefeld.de/lili/projekte/dazkom/

ALLGEMEINE INFORMATIONEN

Name des Instruments: DaZKom-Test
Bisheriger Einsatz in Projekten: National: Sprachen-Bilden-Chancen, ProfaLe, BiProfessional, Weiterbildungsinitiative NRW International: eCALLMS/ICMEE (University of Nebraska-Lincoln, USA), DivEd (University of Turku, Finnland)
Kontakt / Speicherort:
Prof. Dr. Barbara Koch-Priewe
E-Mail: bkoch-priewe@uni-bielefeld.de
Webseite: https://www.uni-bielefeld.de/erziehungswissenschaft/ag4/projekte/dazkom.html

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Deutsch-als-Zweitsprache (DaZ)-Kompetenz bei angehenden und praktizierenden Lehrkräften
Theoriemodell: DaZKom-Modell (Köker et al., 2015); Kompetenz als Kontinuum (Blömeke et al., 2015)
Testart: Kompetenztest
Modalität: Paper-Pencil-Test sowie computergestützt e Testvariante
Testaufbau:
40 bzw. 60 Minuten zur Durchführung des DaZKom-Tests, 10–15 Minuten für Soziodemografie, ggf. DaZ Lerngelegenheiten und Überzeugungen zu Mehrsprachigkeit

TESTGÜTEKRITERIEN

Reliabilität: WLE-Reliabilität = 0.78 (Ehmke & Hammer, 2018)
Validität Testinhalt: Im ersten Expertenreview (N = 7) beurteilten die Expert*innen auf einer vierstufigen Likert-Skala (1 = sehr wichtig bis 4 = unwichtig) die drei theoretischen Dimensionen des DaZ-Kompetenzmodells Fachregister (M = 1.71, SD = 0.49), Mehrsprachigkeit (M = 1.71, SD = 0.76) und Didaktik (M = 1.00, SD = 0.00) nach ihrer Wichtigkeit für die DaZ-Kompetenz angehender Lehrkräfte als sehr wichtig bis wichtig (Gültekin-Karakoç, 2018). Im zweiten Expertenrating (N = 6) wurden die Testitems den drei Dimensionen zugeordnet (Raterübereinstimmung: P = 63 %, Cohens Kappa = 0.43). Zudem wurden die Relevanz (M = 3.52, SD = 0.46; P = 47 %) und Repräsentativität (M = 3.31, SD = 0.67; P = 39 %) der jeweiligen Testitems für die Dimension beurteilt (vierstufige Likert-Skala: 1 = überhaupt nicht relevant / repräsentativ bis 4 = sehr relevant / repräsentativ).
Validität Antwortprozesse: keine Angaben
Validität interner Testaufbau: keine Angaben
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Statistisch signifikante Zusammenhänge der DaZ-Kompetenz mit linguistischem Wissen (r = 0.25) und pädagogischem Unterrichtswissen in der Subdimension Umgang mit Heterogenität (r = 0.21), aber keine mit mathematikdidaktischem Wissen (Hammer et al., 2015; 2016). Weitere Zusammenhänge zwischen DaZ-Kompetenz mit

Lesegeschwindigkeit (r = 0.34) und Lesefähigkeit (r = 0.26); keine statistisch signifikanten Zusammenhänge mit kognitiven Fähigkeiten (schlussfolgerndes Denken) (Hammer & Ehmke, 2018)
Validität Konsequenzen der Testergebnisse: In einem Standard-setting-Verfahren wurden basierend auf N = 1.383 Studierenden-Antworten drei Standardstufen („unter Mindeststandard", „Mindest-standard", „Regelstandard") bestimmt und auf der Grundlage der Itemanforderungen inhaltlich bzw. kriterial charakterisiert (Gültekin-Karakoç et al., 2016; Gültekin-Karakoç, 2018).
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 40 Minuten (Kurzversion); 60 Minuten (Langversion)
Benötigtes Material: Testhefte; Einverständniserklärungen zur Datenverarbeitung
Besonderheiten: Testleitermanual; die Auswertung erfolgt anhand eines Kodiermanuals mit Punktesystem (0–2 Punkte).
Praxisbeispiel: Der Einsatz erfolgte bislang national in den Projekten Sprachen-Bilden-Chancen, ProfaLe, BiProfessional und Weiterbildungs-initiative NRW sowie international in eCALLMS/ICMEE (University of Nebraska-Lincoln, USA), DivEd (University of Turku, Finnland).
Erprobt für: Alle Schulfächer, alle Schulformen (n = ca. 1.500 Lehramtsstudierende)
Nicht geeignet für: Anwendung bei Einzelpersonen, Aussagen auf individueller Ebene (Zulassungs-/ studienrelevante Erfolgsaussagen)

ZIELGRUPPE

Erprobt für: Angehende und praktizierende Lehrkräfte
Prinzipiell auch geeignet für: —
Nicht geeignet für: Auswahlentscheidungen (Zulassung zu Studiengängen)

ALLGEMEINER TESTZWECK

Erprobt für: Seminarveranstaltungen der Lehramtsausbildung sowie Fort- und Weiterbildungen für Lehrkräfte im Bereich DaZ
Prinzipiell auch geeignet für: Beschreibung der Kompetenzentwick-lung bei Gruppen von Studierenden im Verlauf einer Lehrveranstaltung, z. B. zum Zwecke der Evaluation von Lehrveranstaltungen
Nicht geeignet für: Auswahlentscheidungen (Zulassung zu Studiengängen)

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Fragebögen zu DaZ-Lerngelegenheiten, Beliefs/Über-zeugungen zu Mehrsprachigkeit in der Schule
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig auf Gruppen-ebene
Prinzipiell auch geeignet für: Kursebene, Hochschulebene, Landes- und Bundesebene
Nicht geeignet für: —

EINSATZSZENARIEN

Erprobt für: Messung der Deutsch-als-Zweitsprache-Kompetenz angehender und praktizierender Lehrkräfte
Prinzipiell auch geeignet für: Hochschulevaluation, Evaluation der Lehrkräfteaus- und weiterbildung, Bedarfsplanung in der Hochschul-lehre
Nicht geeignet für: Auswahlentscheidungen (Zulassung zu Studiengängen)

ERLAUBTE KONSEQUENZEN

Erprobt für: Einschätzung der DaZ-Kompetenz bei (angehenden) Lehrkräften
Prinzipiell auch geeignet für: Anpassung der Lehrkräfteaus-/-weiter-bildung an Bedarfe
Nicht geeignet für: Keine Konsequenzen auf individueller Ebene wie Zulassungsentscheidung, Exmatrikulation, Aussagen zum Studienerfolg, Arbeitsleistung

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Projekt DaZKom-Video

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Ehmke, T. & Hammer, S. (2018). Skalierung und dimensionale Struktur des DaZKom-Testinstruments. In T. Ehmke; S. Hammer; B. Koch-Priewe; A. Köker & U. Ohm (Hrsg.), *Professionelle Kompetenzen angehender Lehrkräfte im Bereich Deutsch als Zweitsprache* (S. 129–148). Münster: Waxmann.
Hammer, S. & Ehmke, T. (2018). Ergebnisse einer Validierungsstudie zum DaZKom-Testinstrument. In T. Ehmke; S. Hammer; B. Koch-Priewe; A. Köker & U. Ohm (Hrsg.), *Professionelle Kompetenzen angehender Lehrkräfte im Bereich Deutsch als Zweitsprache* (S. 185–200). Münster: Waxmann.

WEITERFÜHRENDE LITERATUR

Ehmke, T.; Hammer, S.; Koch-Priewe, B.; Köker, A. & Ohm, U. (2018). *Professionelle Kompetenzen angehender Lehrkräfte im Bereich Deutsch als Zweitsprache.* Münster: Waxmann.
Hammer, S.; Carlson, S. A.; Ehmke, T.; Koch-Priewe, B.; Köker, A.; Ohm, U.; Rosenbrock, S. & Schulze, N. (2015). Kompetenz von Lehramtsstudierenden in Deutsch als Zweitsprache: Validierung des GSL-Testinstruments. In S. Blömeke & O. Zlatkin-Troitschanskaia (Hrsg.), Kompetenzen von Studierenden. *Zeitschrift für Pädagogik,* Beiheft 61, 32–54. Weinheim: Beltz.
Köker, A.; Rosenbrock, S.; Ohm, U.; Carlson, S. A.; Ehmke, T.; Hammer, S.; Koch-Priewe, B. & Schulze, N. (2015). DaZKom – Ein Modell von Lehrerkompetenz im Bereich Deutsch als Zweitsprache. In B. Koch-Priewe; A. Köker; J. Seifried & E. Wuttke (Hrsg.), *Welche Kompetenzen brauchen Lehramtsstudierende und angehende ErzieherInnen? Theoretische und empirische Zugänge* (S. 177–206). Bad Heilbrunn: Klinkhardt.

DaZKom-Video: Deutsch-als-Zweitsprache-Kompetenz von Lehrkräften

KURZFASSUNG

Name des Instruments: DaZKom-Video Testinstrument

Fachdomäne: Lehrerbildung

Erfasste Kompetenzen: Deutsch-als-Zweitsprache (DaZ)-Kompetenz bei angehenden und praktizierenden Lehrkräften

Zielgruppe: angehende und praktizierende Lehrkräfte

Testart: Video-Vignetten mit mündlichen, aufgezeichneten Antworten (performanznahe Kompetenzmessung), offenes Antwortformat

Modalität: computergestützt

Dauer: 45–60 Minuten

Testaufbau: 12 Video-Vignetten mit jeweils zwei Items; Videos zeigen DaZ-relevante, situationsspezifische Unterrichtssituationen (1-2 Minuten) mit einer kurzen Kontextbeschreibung, auf die die Teilnehmer*innen direkt reagieren müssen;zusätzlich Abfrage von DaZ-relevanten Lerngelegenheiten und soziodemographischen Daten

Allgemeiner Testzweck: Assessment-Tool für Seminarveranstaltungen der Lehramtsausbildung, Fort- und Weiterbildungen für Lehrkräfte im Bereich DaZ

Konkrete Einsatzszenarien: Bisheriger Einsatz in drei Erprobungs-studien; potentiell für Hochschulevaluation, Evaluation der Lehrkräfteaus- und -weiterbildung, Bedarfsplanung in der Hochschul-lehre

Nicht geeignet für: Anwendung bei Einzelpersonen, Aussagen auf individueller Ebene (Zulassungs-/ studienrelevante Erfolgsaussagen)

Hinweise für praktische Anwendung: Für die Durchführung gibt es organisatorische und technische Hinweise, sodass eine standardisierte Testdurchführung gewährleistet werden kann. Zur Auswertung wird ein Kodiermanual mit Punktesystem (0–2 Punkte) bereitgestellt, sodass die Antworten von vorher geschulten Kodierer*innen kodiert werden können.

Bisheriger Einsatz in Projekten: Verbundprojekt „DaZKom-Video: Performanznahe Messung von Deutsch-als-Zweitsprache-Kompetenz bei (angehenden) Lehrkräften)"

Kontakt / Speicherort:
Prof. Dr. Barbara Koch-Priewe
E-Mail: bkoch-priewe@uni-bielefeld.de
Webseite: https://www.uni-bielefeld.de/erziehungswissenschaft/ag4/projekte/dazkom-video.html

ALLGEMEINE INFORMATIONEN

Name des Instruments: DaZKom-Video Testinstrument
Bisheriger Einsatz in Projekten: DaZKom-Video
Kontakt / Speicherort:
Prof. Dr. Barbara Koch-Priewe
E-Mail: bkoch-priewe@uni-bielefeld.de
Webseite: https://www.uni-bielefeld.de/erziehungswissenschaft/ag4/projekte/dazkom-video.html

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Deutsch-als-Zweitsprache (DaZ)-Kompetenz bei angehenden und praktizierenden Lehrkräften
Theoriemodell: DaZKom-Modell (Köker et al., 2015), Kompetenz als Kontinuum (Blömeke et al., 2015), Expertenmodell nach Dreyfus & Dreyfus (1989)
Testart: Performanznahe Kompetenzmessung; Video-Vignetten mit mündlichen, aufgezeichneten Antworten
Modalität: computer- und videobasiert, offline-Testung
Testaufbau:
Itempool: 12 Video-Vignetten mit jeweils zwei Items (24 Items); zusätzlich Abfrage von soziodemographischen Daten/DaZ-relevante Lerngelegenheiten

TESTGÜTEKRITERIEN

Reliabilität: IRT-Reliabilität $\alpha = 0.6$
Validität Testinhalt: keine Angaben
Validität Antwortprozesse: keine Angaben
Validität interner Testaufbau: Überprüfung dimensionaler Modelle (Lemmrich et al., 2019)
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Zusammenhänge mit anderen Kompetenzmaßen (DaZ-Zertifikat, Lehrerfahrung, Forschung im Bereich DaZ)
Validität Konsequenzen der Testergebnisse: keine Angaben
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 45–60 Minuten
Benötigtes Material: Tablets, Headsets; Einverständniserklärungen zur Studienteilnahme und Datenverarbeitung
Besonderheiten: Testdurchführung erfolgt anhand einer Testleiter-instruktion, Testauswertung erfolgt mit einem Kodiermanual (0-2 Punkte) und nach einer Kodier*innen-Schulung.
Fachdomäne: Lehrerbildung

Praxisbeispiel: In drei Erprobungsstudien wurde der Test auf Tablets bei 500 Proband*innen (Lehramtsstudierende und praktizierende Lehrkräfte) an Universitäten sowie in Fort- und Weiterbildungsveranstaltungen deutschlandweit eingesetzt. Die technische Umsetzung funktioniert einwandfrei und die Proband*innen geben positive Rückmeldungen zur Authentizität der Videoszenen. Teilnehmer*innen sehen 12 Videosequenzen, die authentische Unterrichtssituationen zeigen und werden dann aufgefordert nacheinander zwei Items zu beantworten (1. Was nehmen Sie wahr? 2. Wie würden Sie anstelle der Lehrkraft in dieser Situation handeln?)

FACHDOMÄNE

Erprobt für: Lehrkräfte aller Schulfächer und Schulformen
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

ZIELGRUPPE

Erprobt für: angehende und praktizierende Lehrkräfte
Prinzipiell auch geeignet für: Fortbildner*innen, angehende Lehrkräfte
Nicht geeignet für: Studieanfänger*innen

ALLGEMEINER TESTZWECK

Erprobt für: Seminarveranstaltungen der Lehramtsausbildung sowie Fort- und Weiterbildungen für Lehrkräfte im Bereich DaZ
Prinzipiell auch geeignet für: Beschreibung der Kompetenzentwicklung bei Gruppen von Studierenden im Verlauf einer Lehrveranstaltung
Nicht geeignet für: —

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Fragebögen zu soziodemographischen Daten, Lerngelegenheiten, Beliefs, Persönlichkeitsmerkmalen (Big Five)
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig auf Gruppenebene
Prinzipiell auch geeignet für: Kursebene, Hochschulebene, Landes- und Bundesebene
Nicht geeignet für: Eingangsdiagnostik

EINSATZSZENARIEN

Erprobt für: angehende und praktizierende Lehrkräfte aller Schulformen und Fächer
Prinzipiell auch geeignet für: Evaluation von Lehrveranstaltungen an Hochschulen, Evaluation der Lehrkräfteaus- und -weiterbildung, Bedarfsplanung in der Hochschullehre
Nicht geeignet für: Zulassung, Vorhersage des Studienerfolgs

ERLAUBTE KONSEQUENZEN

Erprobt für: Einschätzung der DaZ-Kompetenz bei (angehenden) Lehrkräften
Prinzipiell auch geeignet für: Anpassung der Lehrkräfteaus-/Weiterbildung an Bedarfe
Nicht geeignet für: Keine Konsequenzen auf individueller Ebene wie Zulassungsentscheidung, Exmatrikulation, Aussagen zum Studienerfolg, Arbeitsleistung

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Aktuell wird das Instrument in Pilotierungs- und Normierungsstudien eingesetzt. In Zukunft soll das Instrument zur Evaluation der Lerngelegenheiten sowie in Lehrkräftefort- und -weiterbildungen genutzt werden.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Hecker, S.-L. & Nimz, K. (2020). Expertinnen-Ratings zur Deutsch-als-Zweitsprache-Kompetenz von Lehrkräften: Eine Vorstudie zur Konzipierung eines videobasierten Testinstruments. *Berlin-Brandenburger Beiträge zur Bildungsforschung.*

Lemmrich, S., Hecker, S.-L., Klein, S., Ehmke, T. Köker, A., Koch-Priewe, B., & Ohm, U. (2019). Performanznahe und videobasierte Messung von DaZ-Kompetenz bei Lehrkräften. Skalierung und dimensionale Struktur des Testinstruments. In T. Ehmke, P. Kuhl & M. Pietsch (Hrsg.), *Lehrer. Bildung. Gestalten. Beiträge zur empirischen Forschung in der Lehrerbildung* (S. 188-202). Weinheim und Basel: Beltz Juventa.

Nimz, K., Hecker, S.-L. & Köker, A.(2018). Videobasierte Messung von DaZ-Kompetenz bei Lehrkräften. In: C. Caruso, J. Hofmann, A. Rohde & K. Schick (Hrsg.) *Sprache im Unterricht. Ansätze, Konzepte, Methoden* (S. 439-452). Trier: WVT.

ELMaWi Mathe-AK: Aktionsbezogene Kompetenz von Mathematiklehrkräften

KURZFASSUNG

Name des Instruments: Mathe-AK (Erfassung von Aktionsbezogener Kompetenz in Mathematik)

Fachdomäne: Lehrerbildung Mathematik

Erfasste Kompetenzen: Aktionsbezogene Kompetenz (AK) von Lehrkräften mit Fach Mathematik

Zielgruppe: (Angehende) Lehrkräfte mit Fach Mathematik der Sekundarstufe I/II (Lehramtsstudierende der Mathematik [Bachelor, Master], Referendar*innen und Lehrer*innen mit Fach Mathematik); analoge Testinstrumente wurden von den Autor*innen auch für die Verwendung im Primar- und Elementarbereich entwickelt.

Testart: Leistungstest; videobasierte Aufgaben, offenes Antwortformat

Modalität: Computer- und videogestützt; Audioaufnahme zum Erfassen aktionsbezogener Kompetenz mit Speed-Komponente

Dauer: 30 Minuten

Testaufbau: 9 Aufgaben; Themenbereiche: Algebra (6 Items) und Analysis (3 Items) für Sekundarstufe. Pro Item je ein Video mit einem prototypischen Unterrichtsszenario (30-60 Sekunden) mit kurzer Kontextbeschreibung, auf das Testteilnehmer*innen direkt reagieren müssen. Die Antworten werden als Audioaufnahme aufgezeichnet und anhand eines Manuals von geschultem Personal bewertet (zwischen 0 und 2).

Allgemeiner Testzweck: Erfassen des Kompetenzstands und der Kompetenzentwicklung (zusammen mit weiteren Maßen professioneller Lehrerkompetenz: Tests zum Fachwissen, Fachdidaktikwissen; s. auch Steckbrief Mathe-RK; Wirtschaft-RK, Wirtschaft-AK)

Konkrete Einsatzszenarien: Vergleiche des Kompetenzstands zwischen Ausbildungsphasen im Lehramt (Bachelor, Master, Referendariat, Schule); Vergleiche zwischen angehenden und praktizierenden Lehrkräften in den Fächern Mathematik und/oder Wirtschaft

Nicht geeignet für: Nicht deutschsprachige internationale Ebene; Eingangsdiagnostik; Studienanfänger*innen

Hinweise für praktische Anwendung: Organisatorische und technische Hinweise zur Sicherstellung einer standardisierten Testdurchführung im Testleitermanual; für eine objektive Auswertung wird ein Kodiermanual bereitgestellt; Raterschulung ist erforderlich.

Bisheriger Einsatz in Projekten: Verbundprojekt „Erfassung von fachspezifischen Kompetenzen bei Lehramtsstudierenden der Fächer Mathematik und Wirtschaftswissenschaften – eine quasiexperimentelle Validierungsstudie unter besonderer Berücksichtigung der Domänenspezifität (ELMaWi)"

Kontakt / Speicherort:
Colin Jeschke
IPN – Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik Kiel
E-Mail: jeschke@leibniz-ipn.de
Webseite: https://www.elmawi.de

ALLGEMEINE INFORMATIONEN

Name des Instruments: Mathe-AK (Erfassung von Aktionsbezogener Kompetenz in Mathematik)
Fachdomäne: Lehrerbildung Mathematik
Bisheriger Einsatz in Projekten: Verbundprojekt „Erfassung von fachspezifischen Kompetenzen bei Lehramtsstudierenden der Fächer Mathematik und Wirtschaftswissenschaften – eine quasiexperimentelle Validierungsstudie unter besonderer Berücksichtigung der Domänenspezifität (ELMaWi)"
Kontakt / Speicherort:
Colin Jeschke
IPN – Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik Kiel
E-Mail: jeschke@leibniz-ipn.de
Webseite: https://www.elmawi.de

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Aktionsbezogene Kompetenz von Lehrkräften mit Fach Mathematik
Zielgruppe: (Angehende) Lehrkräfte mit Fach Mathematik der Sekundarstufe I/II (Lehramtsstudierende Studierende der Mathematik [Bachelor, Master], Referendar*innen und Lehrer*innen mit Fach Mathematik)
Theoriemodell: Nach dem Kompetenzstrukturmodell mit den beiden Facetten AK und RK (s. Steckbrief Mathe-RK) auf der Basis von Lindmeier (2011) und Kuhn (2014)
Testart: Leistungstest; videobasierte Aufgaben mit offenem Antwortformat
Modalität: Computer- und videogestützt; Audioaufnahme zum Erfassen von AK

Testaufbau:
Aufgabenpool: 9 Aufgaben; Themenbereiche: Algebra (6 Items) und Analysis (3 Items) für Sekundarstufe. Pro Item je ein Video mit einem prototypischen Unterrichtsszenario (30-60 Sekunden) mit kurzer Kontextbeschreibung, auf das Testteilnehmer*innen direkt reagieren. Die Antworten werden als Audioaufnahme aufgezeichnet und anhand eines erprobten Manuals nach qualitativ-inhaltlichen und formalen Kriterien bewertet (zwischen 0 und 2).

TESTGÜTEKRITERIEN

Reliabilität: Cronbachs α für AK = .6 (9 Aufgaben)
Validität Testinhalt: keine Angaben
Validität Antwortprozesse: keine Angaben
Validität interner Testaufbau: In Abgrenzung zu CK, PCK und RK theoriekonforme eindimensionale Struktur der AK in CFA bestätigt (Jeschke et al., 2019)
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Theoriekonforme signifikante schwache Zusammenhänge mit CK und PCK in Mathematik sowie moderate Zusammenhänge mit RK Mathematik (Jeschke et al., 2019)
Validität Konsequenzen der Testergebnisse: Theoriekonforme Differenzen in der Testleistung zwischen verschiedenen Statusgruppen (Bachelor, Master, Referendariat, Lehrdienst) (Jeschke et al.,2019)
Testfairness: Wegen der Sprachlastigkeit des Tests ist ein Sprachbias anzunehmen

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 30 Minuten
Benötigtes Material: Computer/Laptop mit dem Test-Programm, Headset mit Mikrofon
Besonderheiten: Organisatorische und technische Hinweise zur Sicherstellung einer standardisierten Testdurchführung im Testleitermanual. Für eine objektive Auswertung wird ein Kodiermanual bereitgestellt; Raterschulung ist erforderlich.
Praxisbeispiel: Das Instrument wurde mit allen Statusgruppen an meheren Institutionen bundesweit eingesetzt. Es beinhaltet eine Einführung in die Benutzung der Software und eine Kontextbeschreibung zu jedem Videoitem. Zur Bearbeitung wird zuerst der Videoclip (30–60 Sekunden) einmal angesehen. Er bildet eine prototypische Unterrichtssituation aus dem Mathematikunterricht der Sekundarstufe I bzw. II ab, in der eine mathematikspezifische Anforderung auftritt (z. B. eine Schülerin stellt eine fachliche Frage, ein Schüler zeigt einen typischen Fehler). Die Teilnehmenden werden gebeten, auf die Situation unter Zeitdruck direkt in natürlicher Sprache zu reagieren (z. B.: „Geben Sie eine Erklärung!" „Geben Sie einen hilfreichen Hinweis, sodass der Lernende die Aufgabe selbst lösen kann!") und die Schüler*innen dabei anzusprechen. Die Aussagen werden per Audioaufnahme aufgezeichnet. Die Datenauswertung erfolgt anhand eines Kodiermanuals.

FACHDOMÄNE

Erprobt für: Mathematikdidaktik und Wirtschaftsdidaktik
Prinzipiell auch geeignet für: andere Lehramtsdomänen
Nicht geeignet für: —

ZIELGRUPPE

Erprobt für: Angehende und praktizierende Lehrkräfte mit Fach Mathematik der Sekundarstufe II (Lehramtsstudierende der Mathematik [Bachelor, Master], Referendar*innen und Lehrer*innen mit Fach Mathematik)
Prinzipiell auch geeignet für: Angehende und praktizierende Lehrkräfte in Mathematik der Sekundarstufe I, ggf. als Kurzskala (entsprechende Testinstrumente wurden von den Autor*innen auch für den Primar- und Elementarbereich entwickelt)
Nicht geeignet für: —

ALLGEMEINER TESTZWECK

Erprobt für: Erfassung des Kompetenzstands in den beiden Facetten AK und RK (s. Steckbrief Mathe-RK) und Vergleiche des Kompetenzstands zwischen Ausbildungsphasen im Lehramt (Bachelor, Master, Referendariat, Schule); Vergleiche zwischen angehenden und praktizierenden Lehrkräften in den Fächern Mathematik und/oder Wirtschaft
Prinzipiell auch geeignet für: Erfassung der Kompetenzentwicklung in allen Ausbildungsphasen; Vergleiche mit weiteren Schulfächern
Nicht geeignet für: —

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Videogestützte Instrumenten zum Erfassen von AK und RK im Fach Wirtschaft für Domänenvergleiche; Instrumente zur Erfassung von PCK und CK in Wirtschaft und Mathematik; Instrumente zur Erfassung von generischen Kompetenzen (wie Intelligenz, Situationsbewusstsein, Ambiguitätstoleranz) für Analysen zur Domänenspezifität (Jeschke et al., 2019)
Prinzipiell auch geeignet für: Weitere Instrumente zur Erfassung von fachbezogenen oder generischen Kompetenzen
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für die individuelle Ebene, Gruppen- bzw. Kursebene, Hochschulebene, internationale Ebene (deutschsprachig)
Prinzipiell auch geeignet für: Länder- und Bundesebene
Nicht geeignet für: Internationale Ebene (nicht deutschsprachig)

EINSATZSZENARIEN

Erprobt für: Vergleiche des Kompetenzstands zwischen Ausbildungsphasen im Lehramt (Bachelor, Master, Referendariat, Schule); Vergleiche zwischen Wirtschaft und Mathematik (bei Studierenden, Referendar*innen, Lehrer*innen)
Prinzipiell auch geeignet für: Lehr-Lern-Tool zur Abschätzung von individuellem Lernbedarf; Erfassung der Kompetenzentwicklung in allen Aus- und Weiterbildungsphasen; Vergleiche mit weiteren Schulfächern
Nicht geeignet für: Eingangsdiagnostik

ERLAUBTE KONSEQUENZEN

Erprobt für: Einschätzung des Kompetenzstands im Lehramtsstudium/Referendariat/Schule; Informationen zur Weiterentwicklung von Lehrangeboten zu den geprüften Kompetenzen
Prinzipiell auch geeignet für: Einschätzung der Kompetenzentwicklung über die verschiedenen Phasen hinweg; statistische Vorhersage des Kompetenzstandes zu einem späteren Zeitpunkt
Nicht geeignet für: Eingangsdiagnostik

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Das Instrument kann um weitere inhaltliche Dimensionen erweitert werden (z. B. in Form zusätzlicher Aufgaben). Es wird eine Weiterentwicklung der Erhebungssoftware angestrebt, um langfristig auf dem aktuellen technischen Stand zu bleiben. Das Instrument zur Erfassung von Mathe-AK kann jederzeit bei den Autor*innen angefragt werden. Die allgemeine Vorgehensweise ist in Lindmeier (2011) ausführlich dokumentiert. Das Instrument wird künftig auch im Transferprojekt „Förderung fachspezifischer Kompetenzen bei Lehramtsstudierenden der Fächer Mathematik und Wirtschaftswissenschaften mittels videobasierter ELMaWi-Tools (ELMaWi-Transfer)" eingesetzt.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z. B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Jeschke, C., Kuhn, C., Lindmeier, A., Zlatkin-Troitschanskaia, O., Saas, H. & Heinze, A. (2019). Performance assessment to investigate the domain-specificity of instructional skills among pre-service and in-service teachers of mathematics and economics. *British Journal of Educational Psychology*, 1–13. https://doi.org/10.1111/bjep.12277
Jeschke, C.; Lindmeier, A. & Heinze, A. (2018). Wie fachspezifisch sind „fachspezifische Kompetenzen" von Lehrkräften? In U. Kortenkamp & A. Kuzle (Hrsg.), *Beiträge zum Mathematikunterricht 2017* (S. 1119–1122). Münster: WTM.
Jeschke, C.; Lindmeier, A. & Heinze, A. (2017). What do mathematics pre-service teachers lack for mastering instructional demands? In B. Kaur; W. K. Ho; T. L. Toh & B. H. Choy (Hrsg.), *Proceedings of the 41st Conference of the International Group for the Psychology of Mathematics Education* (3, S. 33–40). Singapore: PME.

WEITERFÜHRENDE LITERATUR

Kuhn, C. (2014). *Fachdidaktisches Wissen von Lehrkräften im kaufmännisch-verwaltenden Bereich. Modellbasierte Testentwicklung und Validierung* (Empirische Berufsbildungs- und Hochschulforschung, Bd. 2). Landau: Verlag Empirische Pädagogik.
Lindmeier, A. (2011). *Modelling and measuring knowledge and competencies of teachers: A threefold domain-specific structure model for mathematics.* Münster: Waxmann.

ELMaWi Mathe-RK: Reflexive Kompetenz von Mathematiklehrkräften

KURZFASSUNG

Name des Instruments: Mathe-RK (Erfassung von Reflexiver Kompetenz in Mathematik)

Fachdomäne: Lehrerbildung Mathematik

Erfasste Kompetenzen: Reflexive Kompetenz (RK) von Lehrkräften mit Fach Mathematik

Zielgruppe: (Angehende) Lehrkräfte mit Fach Mathematik der Sekundarstufe I/II (Lehramtsstudierende der Mathematik [Bachelor, Master], Referendar*innen und Lehrer*innen mit Fach Mathematik); analoge Testinstrumente wurden von den Autor*innen auch für die Verwendung in Primar- und Elementarbereich entwickelt

Testart: Leistungstest; videobasierte Aufgaben, offenes Antwortformat

Modalität: Computer- und videogestützt; Audioaufnahmen und schriftliche Eingaben zum Erfassen reflexiver Kompetenz (RK)

Dauer: 30 Minuten

Testaufbau: 9 Aufgaben; eine Skala zur Planung und Reflexion von Mathematikunterricht in den Themengebieten der Sekundarstufen-Algebra (6 Items) und -Analysis (3 Items). Der Test geht auf Probleme zu Unterrichtsmaterialien (4 Items), Planung von Unterrichtseinstiegen (2 Items) und Reflexion von Schülerbearbeitungen (3 Items) ein.

Allgemeiner Testzweck: Erfassen des Kompetenzstands und der Kompetenzentwicklung (zusammen mit weiteren Maßen professioneller Lehrkompetenz: Test zum Fachwissen, Fachdidaktikwissen; s. auch Steckbrief Mathe-AK)

Konkrete Einsatzszenarien: Vergleiche des Kompetenzstands zwischen Ausbildungsphasen im Lehramt (Bachelor, Master, Referendariat, Schule); Vergleiche zwischen angehenden und praktizierenden Lehrkräften in den Fächern Mathematik und/oder Wirtschaft

Nicht geeignet für: Nicht deutschsprachige internationale Ebene; Eingangsdiagnostik; Studienanfänger*innen

Hinweise für praktische Anwendung: Organisatorische und technische Hinweise zur Sicherstellung einer standardisierten Testdurchführung im Testleitermanual; für eine objektive Auswertung wird ein Kodiermanual bereitgestellt; Raterschulung ist erforderlich.

Bisheriger Einsatz in Projekten: Verbundprojekt „Erfassung von fachspezifischen Kompetenzen bei Lehramtsstudierenden der Fächer Mathematik und Wirtschaftswissenschaften – eine quasiexperimentelle Validierungsstudie unter besonderer Berücksichtigung der Domänenspezifität (ELMaWi)"

Kontakt / Speicherort:
Colin Jeschke
IPN – Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik Kiel
E-Mail: jeschke@leibniz-ipn.de
Webseite: https://www.elmawi.de

ALLGEMEINE INFORMATIONEN

Name des Instruments: Mathe-RK (Erfassung von Reflexiver Kompetenz in Mathematik)
Fachdomäne: Lehrerbildung Mathematik
Bisheriger Einsatz in Projekten: Verbundprojekt „Erfassung von fachspezifischen Kompetenzen bei Lehramtsstudierenden der Fächer Mathematik und Wirtschaftswissenschaften – eine quasiexperimentelle Validierungsstudie unter besonderer Berücksichtigung der Domänenspezifität (ELMaWi)"
Kontakt / Speicherort:
Colin Jeschke
IPN – Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik Kiel
E-Mail: jeschke@leibniz-ipn.de
Webseite: https://www.elmawi.de

DAS INSTRUMENT

INHALT UND STRUKTUR

Theoriemodell: Nach dem Kompetenzstrukturmodell mit den beiden Facetten AK und RK (s. Steckbrief Mathe-AK) auf der Basis von Lindmeier (2011) und Kuhn (2014)
Testart: Leistungstest; videobasierte Aufgaben mit offenem Antwortformat
Modalität: Computer- und videogestützt; Audioaufnahmen und schriftliche Eingaben zum Erfassen reflexiver Kompetenz (RK)

Testaufbau:
Aufgabenpool: 9 Aufgaben; eine Skala zur Planung und Reflexion von Mathematikunterricht in den Themengebieten der Sekundarstufen-Algebra (6 Items) und -Analysis (3 Items). Der Test geht auf Probleme zu Unterrichtsmaterialien (4 Items), Planung von Unterrichtseinstiegen (2 Items) und Reflexion von Schülerbearbeitungen (3 Items) ein. Jedes Item enthält eine kurze Situationsbeschreibung und eine Abbildung bzw. einen kurzen Videoclip (30-60 Sekunden), auf welche die Teilnehmenden eingehen können (offenes Antwortformat).
Die Ergebnisse werden anhand eines erprobten Manuals nach qualitativ-inhaltlichen und formalen Kriterien bewertet (zwischen 0 und 2).

TESTGÜTEKRITERIEN

Reliabilität: Cronbachs α für RK = .68 (9 Aufgaben)
Validität Testinhalt: keine Angaben
Validität Antwortprozesse: keine Angaben
Validität interner Testaufbau: In Abgrenzung zu AK, CK, und PCK theoriekonforme eindimensionale Struktur der RK in CFA bestätigt.
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Theoriekonforme signifikante schwache Zusammenhänge mit anderen Kompetenzmaßen wie Intelligenz, Motivation, Überzeugungen, Noten
Validität Konsequenzen der Testergebnisse: Theoriekonforme Differenzen in der Testleistung zwischen verschiedenen Statusgruppen (Bachelor, Master, Referendariat, Lehrdienst)
Testfairness: Aufgrund der Sprachlastigkeit des Tests ist ein Sprachbias anzunehmen.

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 30 Minuten
Benötigtes Material: Computer/Laptop mit dem Test-Programm, Headset mit Mikrofon
Besonderheiten: Aufsicht durch eine geschulte Testleitung wird empfohlen. Die Antworten sollten von (möglichst zwei) geschulten Personen anhand des bereitgestellten Kodiermanuals ausgewertet werden.
Praxisbeispiel: Das Instrument kann an jedem Ort eingesetzt werden, der eine Verwendung passender Computer/Laptops inkl. Headset und der eigens entwickelten Software ermöglicht. Es beinhaltet eine Einführung in die Benutzung der Software und eine ausführliche Situationsbeschreibung zu jedem Item. Für jedes Item wird eine Abbildung (z. B. von Unterrichtsmaterialien) oder ein Video (z. B. von einem problematischen Stundenende) gezeigt, worauf die Teilnehmenden mit eigenen Worten eingehen. Die Datenauswertung erfolgt anhand eines Kodiermanuals.

FACHDOMÄNE

Erprobt für: Mathematikdidaktik
Prinzipiell auch geeignet für: andere Fachdomänen wie Wirtschaftsdidaktik
Nicht geeignet für: —

ZIELGRUPPE

Erprobt für: (Angehende) Lehrkräfte mit Fach Mathematik der Sekundarstufe II (Lehramtsstudierende der Mathematik [Bachelor, Master], Referendar*innen und Lehrer*innen mit Fach Mathematik)
Prinzipiell auch geeignet für: Angehende und praktizierende Lehrkräfte in Mathematik der Sekundarstufe I, ggf. als Kurzskala (entsprechende Testinstrumente wurden von den Autor*innen auch für den Primar- und Elementarbereich entwickelt).
Nicht geeignet für: —

ALLGEMEINER TESTZWECK

Erprobt für: Erfassung des Kompetenzstands in den beiden Facetten AK und RK (s. Steckbrief Mathe-AK) und Vergleiche des Kompetenzstands zwischen Ausbildungsphasen im Lehramt (Bachelor, Master, Referendariat, Schule); Vergleiche zwischen angehenden und praktizierenden Lehrkräften in den Fächern Mathematik und/oder Wirtschaft
Prinzipiell auch geeignet für: Erfassung der Kompetenzentwicklung in allen Ausbildungsphasen; Vergleiche mit weiteren Schulfächern
Nicht geeignet für: Eignungsdiagnostik

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Videogestützte Instrumente zum Erfassen von AK und RK im Fach Wirtschaft für Domänenvergleiche; mit Instrumenten zur Erfassung von PCK und CK in Wirtschaft und Mathematik; mit Instrumenten zur Erfassung von generischen Kompetenzen (wie Intelligenz, Situationsbewusstsein, Ambiguitätstoleranz) für Analysen zur Domänenspezifität (Jeschke et al., 2019)
Prinzipiell auch geeignet für: Weitere Instrumente zur Erfassung von fachbezogenen oder generischen Kompetenzen
Nicht geeignet für: Weitere aufmerksamkeits-, zeit- oder anstrengungsintensive Instrumente

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für die individuelle Ebene, Gruppen- bzw. Kursebene, Hochschulebene, internationale Ebene (deutschsprachig).
Prinzipiell auch geeignet für: Länder- und Bundesebene
Nicht geeignet für: Internationale Ebene: (nicht deutschsprachig)

EINSATZSZENARIEN

Erprobt für: Vergleiche des Kompetenzstands zwischen Ausbildungsphasen im Lehramt (Bachelor, Master, Referendariat, Schule); Vergleiche zwischen Wirtschaft und Mathematik (bei Lehramtsstudierende, Referendar*innen, Lehrer*innen)
Prinzipiell auch geeignet für: Lehr-Lern-Tool zur Abschätzung von individuellem Lernbedarf; Erfassung der Kompetenzentwicklung in allen Aus- und Weiterbildungsbildungsphasen; Vergleiche mit weiteren Schulfächern
Nicht geeignet für: Eingangsdiagnostik

ERLAUBTE KONSEQUENZEN

Erprobt für: Einschätzung des Kompetenzstands im Lehramtsstudium / Referendariat / Schule; Informationen zur Weiterentwicklung von Lehrangeboten zu den geprüften Kompetenzen
Prinzipiell auch geeignet für: Einschätzung der Kompetenzentwicklung über die verschiedenen Phasen hinweg; statistische Vorhersage des Kompetenzstandes zu einem späteren Zeitpunkt
Nicht geeignet für: Eingangsdiagnostik

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Das Instrument kann um weitere inhaltliche Dimensionen erweitert werden (z. B. in Form zusätzlicher Aufgaben). Es wird eine Weiterentwicklung der Erhebungssoftware angestrebt, um langfristig auf dem aktuellen technischen Stand zu bleiben. Das Instrument zur Erfassung von Mathe-RK kann jederzeit bei den Autor*innen angefragt werden. Die allgemeine Vorgehensweise ist in Lindmeier (2011) ausführlich dokumentiert. Das Instrument wird künftig auch im Transferprojekt „Förderung fachspezifischer Kompetenzen bei Lehramtsstudierenden der Fächer Mathematik und Wirtschaftswissenschaften mittels videobasierter ELMaWi-Tools (ELMaWi-Transfer)" eingesetzt.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Jeschke, C.; Lindmeier, A. & Heinze, A. (2018). Wie fachspezifisch sind „fachspezifische Kompetenzen" von Lehrkräften? In U. Kortenkamp & A. Kuzle (Hrsg.), *Beiträge zum Mathematikunterricht 2017* (S. 1119–1122). Münster: WTM.
Jeschke, C., Lindmeier, A., Heinze, A. (im review). Vom Wissen zum Handeln unter Druck: Wie wird fachspezifisches Wissen von Mathematiklehrkräften für Anforderungssituationen verfügbar? *Journal für Mathematik-Didaktik*.
Jeschke, C., Kuhn, C., Lindmeier, A., Zlatkin-Troitschanskaia, O., Saas, H. & Heinze, A. (2019). Performance assessment to investigate the domain-specificity of instructional skills among pre-service and in-service teachers of mathematics and economics. *British Journal of Educational Psychology*, 1–13. https://doi.org/10.1111/bjep.12277

WEITERFÜHRENDE LITERATUR

Kuhn, C. (2014). *Fachdidaktisches Wissen von Lehrkräften im kaufmännisch-verwaltenden Bereich. Modellbasierte Testentwicklung und Validierung* (Empirische Berufsbildungs- und Hochschulforschung, Bd. 2). Landau: Verlag Empirische Pädagogik.
Lindmeier, A. (2011). *Modelling and measuring knowledge and competencies of teachers: A threefold domain-specific structure model for mathematics.* Münster: Waxmann.

ELMaWi ARKOM-Wirtschaft: Aktionsbezogene und Reflexive Kompetenz von Wirtschaftslehrkräften

KURZFASSUNG

Name des Instruments: ARKOM-Wirtschaft (Erfassung von Aktionsbezogener und Reflexiver Kompetenz in Wirtschaft)

Fachdomäne: Wirtschaftspädagogik

Erfasste Kompetenzen: Aktionsbezogene und reflexive Kompetenzen von Lehrkräften mit Fach Wirtschaft

Zielgruppe: (Angehende) Lehrkräfte mit Fach Wirtschaft (Studierende der Wirtschaftspädagogik [Bachelor, Master], Referendar*innen und Lehrer*innen mit Fach Wirtschaft an berufsbildenden Studienseminaren und berufsbildenden Schulen); (angehende) Lehrkräfte von Economics im US-Raum

Testart: Leistungstest: videobasierte Aufgaben, offenes Antwortformat

Modalität: Computer- und videogestützt: Audioaufnahme zum Erfassen aktionsbezogener Kompetenz (AK) mit Speed-Komponente; schriftliche Eingabe zum Erfassen reflexiver Kompetenz (RK)

Dauer: 60 Minuten

Testaufbau: 14 videobasierte Aufgaben: AK: 7 Aufgaben; RK: 7 Aufgaben

Allgemeiner Testzweck: Erfassen des Kompetenzstands und der Kompetenzentwicklung in den beiden Facetten

Konkrete Einsatzszenarien: Vergleiche des Kompetenzstands zwischen Ausbildungsphasen im Lehramt (Bachelor, Master, Referendariat, Schule); Vergleiche zwischen angehenden und praktizierenden Lehrkräften in den Fächern Wirtschaft und/oder Mathematik; Interventionsstudie mit Prä-Post-Messdesign zum Erfassen der Kompetenzentwicklung bei Lehramtsstudierenden mit Fach Wirtschaft

Nicht geeignet für: Eingangsdiagnostik; Studienanfänger*innen

Hinweise für praktische Anwendung: Organisatorische und technische Hinweise zur Sicherstellung einer standardisierten Testdurchführung im Testleitermanual; für eine objektive Auswertung wird ein Kodiermanual bereitgestellt; Raterschulung ist erforderlich.

Bisheriger Einsatz in Projekten: Verbundprojekt „Erfassung von fachspezifischen Kompetenzen bei Lehramtsstudierenden der Fächer Mathematik und Wirtschaftswissenschaften – eine quasiexperimentelle Validierungsstudie unter besonderer Berücksichtigung der Domänenspezifität (ELMaWi) (BMBF-FKZ: 01PK15012A/B)", Lehrprojekt „Förderung handlungsnaher Kompetenzen mittels eines videobasierten Lerntools zur Simulation authentischer Unterrichtssituationen bei Studierenden der Wirtschaftspädagogik", internationales Kooperationsprojekt „Machbarkeitsstudie zu Adaption und Einsatz eines deutschsprachigen Instruments zur Erfassung der fachdidaktischen Kompetenzen von Lehrkräften mit Fach Wirtschaft in den USA".

Kontakt / Speicherort:
Dr. Christiane Kuhn
Johannes Gutenberg-Universität Mainz
E-Mail: christiane.kuhn@uni-mainz.de
Webseite: https://www.elmawi.de/

ALLGEMEINE INFORMATIONEN

Name des Instruments: ARKOM-Wirtschaft (Erfassung von Aktionsbezogener und Reflexiver Kompetenz in Wirtschaft)
Bisheriger Einsatz in Projekten: Verbundprojekt „Erfassung von fachspezifischen Kompetenzen bei Lehramtsstudierenden der Fächer Mathematik und Wirtschaftswissenschaften – eine quasiexperimentelle Validierungsstudie unter besonderer Berücksichtigung der Domänenspezifität (ELMaWi)", Lehrprojekt „Förderung handlungsnaher Kompetenzen mittels eines videobasierten Lerntools zur Simulation authentischer Unterrichtssituationen bei Studierenden der Wirtschaftspädagogik", internationales Kooperationsprojekt „Machbarkeitsstudie zu Adaption und Einsatz eines deutschsprachigen Instruments zur Erfassung der fachdidaktischen Kompetenzen von Lehrkräften mit Fach Wirtschaft in den USA"
Kontakt / Speicherort:
Dr. Christiane Kuhn
Johannes Gutenberg-Universität Mainz
E-Mail: christiane.kuhn@uni-mainz.de
Webseite: https://www.elmawi.de/

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Aktionsbezogene und reflexive Kompetenzen von Lehrkräften mit Fach Wirtschaft
Theoriemodell: Kompetenzstrukturmodell mit den beiden Facetten AK und RK auf der Basis von Kuhn (2014), Lindmeier (2011); siehe auch Kuhn et al. (2018; 2020), Zlatkin-Troitschanskaia et al. (2019)
Testart: Leistungstest: videobasierte Aufgaben mit offenem Antwortformat
Modalität: Computer- und videogestützt: Audioaufnahme zum Erfassen von AK mit Speedkomponente; schriftliche Eingabe zum Erfassen von RK
Testaufbau:
Itempool: 14 videobasierte Aufgaben: AK: 7 Aufgaben; RK: 7 Aufgaben

TESTGÜTEKRITERIEN

Reliabilität: Cronbachs α: für AK .64 (N = 390, 7 Aufgaben), für RK .61 (N = 384, 7 Aufgaben)
Validität Testinhalt: Inhaltliche Validierung der Items mittels einer standardisierten Expertenbefragung (N = 41) (Kuhn et al., 2018)
Validität Antwortprozesse: Kognitive Validierung der Items in kognitiven Interviews mittels der Methode des lauten Denkens (N = 44) (Zlatkin-Troitschanskaia et al., 2019)
Validität interner Testaufbau: Theoriekonforme eindimensionale Modelle der AK und der RK in CFA bestätigt
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Theoriekonforme Zusammenhänge mit CK und PCK in Wirtschaft und Mathematik sowie mit AK und RK in Mathematik sowie mit generischen Konstrukten (wie Intelligenz, Ambiguitätstoleranz) (Jeschke et al., 2019; Kuhn et al. 2020)
Validität Konsequenzen der Testergebnisse: Theoriekonforme Unterschiede in den Kompetenzausprägungen zwischen den verschiedenen Expertisegruppen (Bachelor, Master, Referendar*innen, Lehrkräfte) (Kuhn et al., 2020)
Testfairness: Aufgrund der Sprachlastigkeit des Tests ist ein Sprachbias anzunehmen.

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 60 Minuten
Benötigtes Material: Computer/Laptop mit dem Test-Programm, Headset mit Mikrofon
Besonderheiten: Organisatorische und technische Hinweise zur Sicherstellung einer standardisierten Testdurchführung im Testleitermanual. Für eine objektive Auswertung wird ein Kodiermanual bereitgestellt; Raterschulung ist erforderlich.
Praxisbeispiel: In einer Interventionsstudie mit Prä-Post-Messdesign wurde die Kompetenzentwicklung von AK und RK in Wirtschaft in einem universitären Kurs im Lehramtsstudium analysiert (Saas et al., 2020).

FACHDOMÄNE

Erprobt für: Wirtschaftspädagogik; kaufmännisch-verwaltender Bereich
Prinzipiell auch geeignet für: Ökonomische Bildung im allgemeinbildenden Bereich; weitere Domänen mit wirtschaftswissenschaftichen Inhalten
Nicht geeignet für: —

ZIELGRUPPE

Erprobt für: (Angehende) Lehrkräfte mit Fach Wirtschaft (Bachelor und Masterstudierende der Wirtschaftspädagogik, Referendar*innen und Lehrer*innen mit Wirtschaft an berufsbildenden Studienseminaren und Schulen deutschlandweit)
Prinzipiell auch geeignet für: (Angehende) Lehrkräfte von Economics im US-Raum: Eine adaptierte und synchronisierte Version der Economics Aufgaben und Videos liegt vor und wurde bereits im US-Raum für Validierungszwecke eingesetzt; grundsätzlich sind Adaptionen für den Einsatz in weiteren Ländern denkbar.
Nicht geeignet für: Studienanfänger*innen

ALLGEMEINER TESTZWECK

Erprobt für: Erfassung des Kompetenzstands in den beiden Facetten und Vergleiche des Kompetenzstands zwischen Ausbildungsphasen im Lehramt (Bachelor, Master, Referendariat, Schule); Vergleiche zwischen angehenden und praktizierenden Lehrkräften in den Fächern Wirtschaft und/oder Mathematik; Interventionsstudie mit Prä-Post-Messdesign zum Erfassen der Kompetenzentwicklung bei Lehramtsstudierenden mit Fach Wirtschaft
Prinzipiell auch geeignet für: Erfassung der Kompetenzentwicklung in allen Ausbildungsphasen; Vergleiche mit weiteren Schulfächern
Nicht geeignet für: Eingangsdiagnostik

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Gemeinsamen Einsatz mit videogestützten Instrumenten zum Erfassen von AK und RK im Fach Mathematik für Domänenvergleiche; mit Instrumenten zur Erfassung von PCK und CK in Wirtschaft und Mathematik; mit Instrumenten zur Erfassung von generischen Kompetenzen (wie Intelligenz, Ambiguitätstoleranz) für Analysen zur Domänenspezifität (Jeschke et al., 2019; Kuhn et al. 2020)
Prinzipiell auch geeignet für: Weitere Instrumente zur Erfassung von fachbezogenen oder generischen Kompetenzen
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für die Gruppenebene / Kursebene / Hochschulebene.
Prinzipiell auch geeignet für: Internationale Ebene
Nicht geeignet für: —

EINSATZSZENARIEN

Erprobt für: Vergleiche des Kompetenzstands zwischen Ausbildungsphasen im Lehramt (Bachelor, Master, Referendariat, Schule); Vergleiche zwischen Wirtschaft und Mathematik (bei Studierenden, Referendar*innen, Lehrer*innen); Interventionsstudie mit Prä-Post-Messdesign zum Erfassen der Kompetenzentwicklung von Studierenden in Wirtschaft in einem universitären Kurs der Wirtschaftspädagogik
Prinzipiell auch geeignet für: Erfassung der Kompetenzentwicklung in allen Ausbildungsphasen; Vergleiche mit weiteren Schulfächern
Nicht geeignet für: Eingangsdiagnostik

ERLAUBTE KONSEQUENZEN

Erprobt für: Einschätzung des Kompetenzstands im Lehramtsstudium/Referendariat/Schule
Prinzipiell auch geeignet für: Einschätzung der Kompetenzentwicklung über die verschiedenen Phasen hinweg; statistische Vorhersage des Kompetenzstandes zu einem späteren Zeitpunkt
Nicht geeignet für: Eingangsdiagnostik

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Das Instrument kann um weitere inhaltliche Dimensionen erweitert werden (z. B. in Form zusätzlicher Aufgaben). Neben einer Adaption und Synchronisation für (angehende) Lehrkräfte von Economics im U.S.-Raum, die bereits in einer Validierungsstudie in den USA eingesetzt wurde, steht das Instrument auch für Adaptionen in weiteren Ländern bereit. Das Instrument zur Erfassung der AK und RK in Wirtschaft kann beim Mainzer Projektteam angefragt werden. Das Instrument wird künftig auch im Transferprojekt „Förderung fachspezifischer Kompetenzen bei Lehramtsstudierenden der Fächer Mathematik und Wirtschaftswissenschaften mittels videobasierter ELMaWi-Tools (ELMaWi-Transfer)" sowie im Verbundprojekt „Technik- und Wirtschaft: Integrierte Didaktik (TWIND)" eingesetzt.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Kuhn, C., Zlatkin-Troitschanskaia, O., Lindmeier, A., Jeschke, C., Saas, H. & Heinze, A. (2020). Relationships between domain-specific knowledge, generic attributes, and instructional skills – Results from a comparative study with pre- and in-service teachers of mathematics and economics. In O. Zlatkin-Troitschanskaia, H. A. Pant, M. Toepper & C. Lautenbach (Eds.), *Student Learning in German Higher Education: Innovative Measurement Approaches and Research Results*. Wiesbaden: Springer.

Kuhn, C.; Zlatkin-Troitschanskaia, O.; Brückner, S. & Saas, H. (2018). A new video-based tool to enhance teaching economics. *International Review of Economics Education, 27*, 24-33. https://doi.org/10.1016/j.iree.2018.01.007

Saas, H.; Kuhn, C. & Zlatkin-Troitschanskaia, O. (2020). Ein videobasiertes Lehr-Lernformat als innovativer hochschuldidaktischer Ansatz in der wirtschaftspädagogischen Lehrerausbildung. *Zeitschrift für Erziehungswissenschaft*.

Zlatkin-Troitschanskaia, O., Kuhn, C., Brückner, S. & Leighton, J. P. (2019). Evaluating a technology-based assessment (TBA) to measure teachers' action-related and reflective skills. *International Journal of Testing (IJT)*, 19(2), 148–171. https://doi.org/10.1080/15305058.2019.1586377

WEITERFÜHRENDE LITERATUR

Kuhn, C. (2014). *Fachdidaktisches Wissen von Lehrkräften im kaufmännisch-verwaltenden Bereich. Modellbasierte Testentwicklung und Validierung* (Empirische Berufsbildungs- und Hochschulforschung, Bd. 2). Landau: Verlag Empirische Pädagogik.

Lindmeier, A. (2011). *Modelling and measuring knowledge and competencies of teachers: A threefold domain-specific structure model for mathematics.* Münster: Waxmann.

WiWi-PCK: Fachdidaktisches Wissen von Wirtschaftslehrkräften

KURZFASSUNG

Name des Instruments: WiWi-PCK: Fachdidaktisches Wissen Wirtschaft

Fachdomäne: Wirtschaftspädagogik; kaufmännisch-verwaltender Bereich

Erfasste Kompetenzen: Fachdidaktisches Wissen von (angehenden) Lehrkräften (*pedagogical content knowledge*, PCK) mit Fach Wirtschaft (kaufmännisch-verwaltender Bereich)

Zielgruppe: (Angehende) Lehrkräfte mit Fach Wirtschaft (Bachelor- und Masterstudierende der Wirtschaftspädagogik, Referendar*innen und Lehrer*innen mit Fach Wirtschaft an berufsbildenden Studienseminaren und Schulen); (angehende) Lehrkräfte von Economics im US-Raum und ggf. weiteren Ländern

Testart: Leistungstest: Situationsbezogene Aufgaben im Multiple Choice- und offenen Antwortformat

Modalität: Paper-Pencil-Test

Dauer: 45 Minuten (Langversion); 30 Minuten (Kurzversion)

Testaufbau: Langversion: 17 Items verteilt über die Kategorien „Praktische Anforderungen" (Unterrichtsplanung, Reaktion auf Schüleräußerungen) und „Kognitive Prozesse" (Anwenden & Analysieren, Kreieren). Kurzversion: 11 Items

Allgemeiner Testzweck: Erfassen des Kompetenzstands

Konkrete Einsatzszenarien: Vergleiche des Kompetenzstands zwischen Gruppen in den verschiedenen Ausbildungsphasen im Lehramt (Bachelor, Master, Referendariat, Schule); Vergleiche zwischen den Fächern Wirtschaft und Mathematik; Interventionsstudie mit Prä-Post-Messdesign zum Erfassen der Kompetenzentwicklung in Wirtschaft

Nicht geeignet für: Eingangstestung; Studienanfänger*innen

Hinweise für praktische Anwendung: Organisatorische Hinweise zur Sicherstellung einer standardisierten Testdurchführung in einem Testleitermanual; zur Sicherstellung einer objektiven Auswertung wird ein Kodiermanual bereitgestellt; darin ist auch eine Dokumentation der fachdidaktischen Aufgaben (inklusive theoretischer Verortung) enthalten.

Bisheriger Einsatz in Projekten: Konstruktion eines Messinstruments zur Erfassung der fachdidaktischen Kompetenz in der Domäne Wirtschaftswissenschaften, Verbundprojekt „Erfassung von fachspezifischen Kompetenzen bei Lehramtsstudierenden der Fächer Mathematik und Wirtschaftswissenschaften – eine quasiexperimentelle Validierungsstudie unter besonderer Berücksichtigung der Domänenspezifität (ELMaWi)", internationales Kooperationsprojekt „Machbarkeitsstudie zu Adaption und Einsatz eines deutschsprachigen Instruments zur Erfassung der fachdidaktischen Kompetenzen von Lehrkräften mit Fach Wirtschaft in den USA", Lehrprojekt „Förderung handlungsnaher Kompetenzen mittels eines videobasierten Lerntools zur Simulation authentischer Unterrichtssituationen bei Studierenden der Wirtschaftspädagogik"

Kontakt / Speicherort:
Dr. Christiane Kuhn
E-Mail: christiane.kuhn@uni-mainz.de
Webseiten:
https://www.wipaed.uni-mainz.de/lehrstuhlteam/wissenschaftliche-mitarbeiter-innen-der-lehrstuehle/dr-christiane-kuhn/
https://www.wipaed.uni-mainz.de/fachdidaktisches-wissen-wirtschaft/

ALLGEMEINE INFORMATIONEN

Name des Instruments: WiWi-PCK: Fachdidaktisches Wissen Wirtschaft
Bisheriger Einsatz in Projekten: Konstruktion eines Messinstruments zur Erfassung der fachdidaktischen Kompetenz in der Domäne Wirtschaftswissenschaften, Verbundprojekt „Erfassung von fachspezifischen Kompetenzen bei Lehramtsstudierenden der Fächer Mathematik und Wirtschaftswissenschaften – eine quasiexperimentelle Validierungsstudie unter besonderer Berücksichtigung der Domänenspezifität (ELMaWi)", internationales Kooperationsprojekt „Machbarkeitsstudie zu Adaption und Einsatz eines deutschsprachigen Instruments zur Erfassung der fachdidaktischen Kompetenzen von Lehrkräften mit Fach Wirtschaft in den USA", Lehrprojekt „Förderung handlungsnaher Kompetenzen mittels eines videobasierten Lerntools zur Simulation authentischer Unterrichtssituationen bei Studierenden der Wirtschaftspädagogik"
Kontakt / Speicherort:
Dr. Christiane Kuhn
E-Mail: christiane.kuhn@uni-mainz.de
Webseiten:
https://www.wipaed.uni-mainz.de/lehrstuhlteam/wissenschaftliche-mitarbeiter-innen-der-lehrstuehle/dr-christiane-kuhn/
https://www.wipaed.uni-mainz.de/fachdidaktisches-wissen-wirtschaft/

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Fachdidaktisches Wissen von (angehenden) Lehrkräften (PCK) in der Domäne Wirtschaft (kaufmännisch-verwaltender Bereich)
Theoriemodell: Strukturmodell zum wirtschaftsdidaktischen Wissen von Lehrkräften in der Domäne Wirtschaft (kaufmännisch-verwaltender Bereich) (Kuhn, 2014)
Testart: Leistungstest: Situationsbezogene Aufgaben im Multiple-Choice- und offenen Antwortformat
Modalität: Paper-Pencil-Test
Testaufbau:
Itempool: 17 Items verteilt über die Kategorien „Praktische Anforderungen" (Unterrichtsplanung, Reaktion auf Schüleräußerungen) und „Kognitive Prozesse" (Anwenden & Analysieren, Kreieren) (siehe auch „Weitere wichtige Angaben")
Testversionen: 1 Langversion; 1 Kurzversion: 11 Items

TESTGÜTEKRITERIEN

Reliabilität: Cronbachs α: 0.68 (N = 338, 17 Items); EAP-Reliabilität: 0.71
Validität Testinhalt: Inhaltliche Validierung der Items mittels einer standardisierten Expertenbefragung (N = 13) (Kuhn, 2014)
Validität Antwortprozesse: Kognitive Validierung der Items in kognitiven Interviews mittels der Methode des lauten Denkens (N = 6 Studierende der Wirtschaftspädagogik) (Kuhn & Brückner, 2013)
Validität interner Testaufbau: Theoriekonforme eindimensionale Struktur (konfirmatorische Faktorenanalyse [CFA], N = 338, 17 Items: $\chi 2/df = 1.34$ mit $\chi 2 = 159.60$ und df = 119; RMSEA = 0.03; CFI = 0.94; TLI = 0.93; WRMR = 0.90); Konformität mit dem Rasch-Modell für ordinale Daten (Partial Credit-Modell, N = 338): WMNSQ mit Werten (0.91; 1.14); Validierung der Antwortkategorien der MC-Items (Trennschärfen der Antwortkategorien gemäß der Kategorienfolge) (Kuhn et al., 2016)
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Theoriekonform: PCK Wirtschaft korreliert signifikant, aber mäßig mit dem betriebswirtschaftlichen (BAKT-K) und mit dem volkswirtschaftlichen Fachwissen (WBT-K) (Kuhn et al., 2014). Der Vergleich mit einer Kontrastgruppe (Referendar*innen und Lehrkräfte mit anderen Fächern als Wirtschaft) bestätigt, dass der Test nicht ein allgemein pädagogisches bzw. allgemeindidaktisches Wissen erfasst (Kuhn, 2014). Weitere theoriekonforme Zusammenhänge mit externen Variablen wie Abiturnote, absolvierte Schulpraktika, Anzahl Berufsjahre im Schuldienst (Kuhn, 2014).
Validität Konsequenzen der Testergebnisse: Theoretische Annahmen zum Expertise-Grad bestätigt: Die Testleistung steigt signifikant von Wirtschaftspädagogik-Studierenden im Bachelor zu Wirtschaftspädagogik-Studierenden im Master zu Referendar*innen mit Fach Wirtschaft; zwischen Referendar*innen und Lehrkräften mit dem Fach Wirtschaft bestehen keine signifikanten Unterschiede (Kuhn, 2014).
Testfairness: Der Test erfasst dasselbe Merkmal bei den verschiedenen Gruppen: (a) weiblich vs. männlich und (b) Studierende vs. Referendar*innen/Lehrkräfte (Kuhn, 2014).

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 45 Minuten (Langversion); 30 Minuten (Kurzversion)
Benötigtes Material: Testhefte, Stifte
Besonderheiten: Organisatorische Hinweise zur Sicherstellung einer standardisierten Testdurchführung in einem Testleitermanual; zur Sicherstellung einer objektiven Auswertung wird ein Kodiermanual bereitgestellt; darin ist auch eine Dokumentation der fachdidaktischen Aufgaben enthalten (inklusive theoretische Verortung).
Praxisbeispiel: Siehe bisheriger Einsatz in o.g. Forschungs- und Lehrprojekten

FACHDOMÄNE

Erprobt für: Wirtschaftspädagogik: Kaufmännisch-verwaltender Bereich (Domäne Wirtschaft)
Prinzipiell auch geeignet für: Ökonomische Bildung (allgemeinbildender Bereich)
Nicht geeignet für: —

ZIELGRUPPE

Erprobt für: (Angehende) Lehrkräfte der Fächer Wirtschaft (Bachelor- und Masterstudierende der Wirtschaftspädagogik, Referendar*innen und Lehrkräfte mit dem Fach Wirtschaft an berufsbildenden Studienseminaren und Schulen)
Prinzipiell auch geeignet für: (Angehende) Lehrkräfte von Economics im US-Raum (eine adaptierte Version der VWL-Aufgaben liegt vor und wurde bereits im US-Raum für Validierungszwecke eingesetzt); grundsätzlich sind Adaptionen für den Einsatz in weiteren Ländern denkbar.
Nicht geeignet für: Studienanfänger*innen

ALLGEMEINER TESTZWECK

Erprobt für: Erfassen des Kompetenzstands in den verschiedenen Gruppen
Prinzipiell auch geeignet für: Erfassung der Kompetenzentwicklung
Nicht geeignet für: —

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Instrumente zur Erfassung des Fachwissens (CK) sowie videobasierte Instrumente zur Erfassung der aktionsbezogenen (AK) und reflexiven Kompetenz (RK) im Fach Wirtschaft (s. Steckbrief ARKOM); Instrumente zur Erfassung von PCK, CK, AK, RK in Mathematik bei (angehenden) Lehrkräften der Fächer Mathematik und Wirtschaft für domänenvergleichende Analysen
Prinzipiell auch geeignet für: Weitere Instrumente zur Erfassung von fachbezogenen oder generischen Kompetenzen
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für die Gruppenebene; Vergleiche des Kompetenzstands zwischen den einzelnen Phasen der Lehrbildung (Studierende, Referendar*innen, Lehrkräfte).
Prinzipiell auch geeignet für: Internationale Ebene; im zeitlichen Verlauf
Nicht geeignet für: —

EINSATZSZENARIEN

Erprobt für: Siehe zuvor genannte Forschungs- und Lehrprojekte
Prinzipiell auch geeignet für: Erfassung der Wissensentwicklung in allen Lehrbildungsphasen; Vergleiche mit weiteren Schulfächern
Nicht geeignet für: Eingangsdiagnostik

ERLAUBTE KONSEQUENZEN

Erprobt für: Einschätzung des Wissensstands im Lehramtsstudium/Referendariat/Schule
Prinzipiell auch geeignet für: Einschätzung der Wissensentwicklung über die verschiedenen Phasen hinweg; statistische Vorhersage des Wissensstandes zu einem späteren Zeitpunkt
Nicht geeignet für: Eingangsdiagnostik

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Das Instrument wird künftig auch im Transferprojekt „Förderung fachspezifischer Kompetenzen bei Lehramtsstudierenden der Fächer Mathematik und Wirtschaftswissenschaften mittels videobasierter ELMaWi-Tools (ELMaWi-Transfer)" eingesetzt.

WEITERE WICHTIGE ANGABEN

Testaufbau im Detail

Fachdidaktisches Wissen im kaufmännisch-verwaltenden Bereich		
Praktische Anforderungen	Kognitive Prozesse	
	Anwenden & Analysieren	Kreieren
Unterrichtsplanung	5 Items (MC)	6 Items (offen)
Reaktion auf Schüleräußerungen	3 Items (MC)	3 Items (offen)
… in den Kerngebieten „Verkauf", „Beschaffung", „Gesamtwirtschaft" bezogen auf zu fördernde Schülerkompetenzen (z. B. Entscheidungen erwägen, Konflikte bewältigen)		

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Kuhn, C. (2014). *Fachdidaktisches Wissen von Lehrkräften im kaufmännisch-verwaltenden Bereich. Modellbasierte Testentwicklung und Validierung* (Empirische Berufsbildungs- und Hochschulforschung, Bd. 2). Landau: Verlag Empirische Pädagogik.

Kuhn, C.; Alonzo, A. C. & Zlatkin-Troitschanskaia, O. (2016). Evaluating the pedagogical content knowledge of pre- and in-service teachers of business and economics to ensure quality of classroom practice in vocational education and training. *Empirical Research in Vocational Education and Training, 8*(5). DOI: 10.1186/s40461-016-0031-2

Kuhn, C. & Brückner, S. (2013). Analyse des fachdidaktischen Wissens von (angehenden) Lehrkräften in der kaufmännisch-verwaltenden Bildung mit der Methode des lauten Denkens. *bwp@ Berufs- und Wirtschaftspädagogik – online, 24*, S. 1–20. Abgerufen unter http://www.bwpat.de/ausgabe24/kuhn_brueckner_bwpat24.pdf

WEITERFÜHRENDE LITERATUR

Kuhn, C. & Happ, R. (2017). Effekte von curricularen Veränderungen im Zuge der Bologna Reform auf das fachdidaktische Wissen von Studierenden der Wirtschaftspädagogik. *Zeitschrift für Berufs- und Wirtschaftspädagogik, 113*(4), S. 542–568.

Kuhn, C.; Happ, R.; Zlatkin-Troitschanskaia, O.; Beck, K.; Förster, M. & Preuße, D. (2014). Kompetenzentwicklung angehender Lehrkräfte im kaufmännisch-verwaltenden Bereich – Erfassung und Zusammenhänge von Fachwissen und fachdidaktischem Wissen. In E. Winther & M. Prenzel (Hrsg.), Perspektiven der empirischen Berufsbildungsforschung: Kompetenz und Professionalisierung [Sonderheft]. *Zeitschrift für Erziehungswissenschaft, 17*(1), S. 149–167. DOI: 10.1007/s11618-013-0456-3

EXMO: Kompetenzen von Biologielehrkräften im Vermitteln und Beurteilen von Experimenten

KURZFASSUNG

Name des Instruments: Test zu Lehrkompetenzen im experimentellen Biologieunterricht

Fachdomäne: Lehramt Biologie

Erfasste Kompetenzen: Analyse-, Planungs- und Beurteilungskompetenzen für den Biologieunterricht der Sekundarstufe I (Unterricht analysieren und planen, Experimente von Schüler*innen beurteilen)

Zielgruppe: Biologie-Lehramtsstudierende

Testart: Leistungstest; offenes Antwortformat

Modalität: Paper-Pencil-Test

Dauer: 120-140 Minuten (60-70 Minuten: Unterrichtsanalyse- und Planungskompetenzen; 60-70 Minuten: Beurteilungskompetenz)

Testaufbau: „Vermittlungskompetenzen zum Experimentieren im experimentellen Biologieunterricht": Unterricht analysieren: 6 Items; Unterricht planen: 6 Items, Fachwissen im Experimentieren: 6 Items. „Beurteilungskompetenz": Beurteilung der Hypothesenbildung, der Planung von Experimenten und Auswertung von Daten: 7 offene Aufgaben, 22 Bewertungsitems. Jeweils Aufgaben aus den Lehr-Lern-Kontexten: Samenkeimung (5./6. Klasse), Photosynthese (7./8. Klasse), Enzymatik (9./10. Klasse)

Allgemeiner Testzweck: Erfassen des Kompetenzstands bzw. der Kompetenzentwicklung mit Blick auf Analyse-, Planungs- und Beurteilungskompetenzen

Konkrete Einsatzszenarien: Grundlagenforschung; Evaluation von Lehrveranstaltungen, Beurteilung von Studien- bzw. Arbeitsleistungen

Nicht geeignet für: Schulfächer, in denen nicht naturwissenschaftlich experimentiert wird

Hinweise für praktische Anwendung: Zur Auswertung der offenen Aufgaben wird ein Kodiermanual bereitgestellt. Pro Item werden 0-1 oder 0-2 Punkte vergeben.

Bisheriger Einsatz in Projekten: Vermittlungs- und Beurteilungskompetenzen zum Experimentieren: Modellierung, Validierung und Messinstrumententwicklung (EXMO)

Kontakt / Speicherort:
Prof. Dr. Marcus Hammann
Westfälische Wilhelms-Universität Münster,
Zentrum für Didaktik der Biologie
Postanschrift: Schlossplatz 34, 48143 Münster
Tel.: 0251-83-39390 / 31362
Fax: 0251-83-31330
E-Mail: hammann.m@uni-muenster.de
Webseite: http://www.uni-muenster.de/Biologie.Didaktik/

Prof. Dr. Susanne Bögeholz
Georg-August-Universität Göttingen,
Albrecht-von-Haller-Institut für Pflanzenwissenschaften
Postanschrift: Waldweg 26, 37073 Göttingen
Tel.: 0551-399314
E-Mail: sboegeh@gwdg.de
Webseite: https://www.uni-goettingen.de/de/137982.html

Prof. Dr. Claus H. Carstensen
Otto-Friedrich-Universität Bamberg
Postanschrift: Wilhelmsplatz 3, R. 125, 96047 Bamberg
Tel.: 0951-863-3447
Fax: 0951-863-3547
E-Mail: claus.carstensen@uni-bamberg.de
Webseite: https://www.uni-bamberg.de/psymethodenbf/mitarbeiterinnen-und-mitarbeiter/prof-dr-claus-h-carstensen/

ALLGEMEINE INFORMATIONEN

Name des Instruments: Test zu Lehrkompetenzen für den experimentellen Biologieunterricht
Bisheriger Einsatz in Projekten: Vermittlungs- und Beurteilungskompetenzen zum Experimentieren: Modellierung, Validierung und Messinstrumententwicklung (EXMO)

Kontakt / Speicherort:
Prof. Dr. Marcus Hammann
Westfälische Wilhelms-Universität Münster,
Zentrum für Didaktik der Biologie
Postanschrift: Schlossplatz 34, 48143 Münster
Tel.: 0251-83-39390 / 31362
Fax: 0251-83-31330
E-Mail: hammann.m@uni-muenster.de
Webseite: http://www.uni-muenster.de/Biologie.Didaktik/

Prof. Dr. Susanne Bögeholz
Georg-August-Universität Göttingen,
Albrecht-von-Haller-Institut für Pflanzenwissenschaften
Postanschrift: Waldweg 26, 37073 Göttingen
Tel.: 0551-399314
E-Mail: sboegeh@gwdg.de
Webseite: https://www.uni-goettingen.de/de/137982.html

Prof. Dr. Claus H. Carstensen
Otto-Friedrich-Universität Bamberg
Postanschrift: Wilhelmsplatz 3, R.12596047 Bamberg
Tel.: 0951-863-3447
Fax: 0951-863-3547
E-Mail: claus.carstensen@uni-bamberg.de
Webseite: https://www.uni-bamberg.de/psymethodenbf/mitarbeiterinnen-und-mitarbeiter/prof-dr-claus-h-carstensen/

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenz: Unterrichtsanalyse und -planungskompetenzen sowie Beurteilungskompetenzen für den Biologieunterricht der Sekundarstufe I (Unterricht analysieren und planen, Experimente von Schüler*innen beurteilen)
Theoriemodell: Das Kompetenzmodell „Vermittlungs- und Beurteilungskompetenzen zum Experimentieren im Biologieunterricht" basiert auf den drei Phasen des SDDS-Modells von Klahr (2000; Suche im Hypothesenraum, Suche im Experimentierraum und Analyse von Daten; Hammann, 2004). Hammann, Phan, Ehmer & Bayrhuber (2006) beschreiben Defizite von Schüler*innen beim Experimentieren für diese drei Phasen. Für die Lehrerbildung beziehen Hasse et al. (2014) Schülerkompetenzen aus diesen drei Experimentierphasen ein, um Instrumente zur Erfassung von Vermittlungs- und Beurteilungskompetenzen bei Lehrkräften zu entwickeln.
Testart: Leistungstest; offenes Antwortformat
Modalität: Paper-Pencil-Test
Testaufbau:
Itempool: Der Test „Vermittlungskompetenzen zum Experimentieren im experimentellen Biologieunterricht" besteht aus 3 schriftlichen Vignetten und den beiden Teilskalen „experimentellen Biologieunterricht analysieren" (6 Items) und „experimentellen Biologieunterricht planen" (6 Items) sowie einer Skala „Fachwissen zum Experimentieren" (6 Items). Der Test zur „Beurteilungskompetenz" besteht aus 7 Aufgaben zu den 3 Aufgabenkontexten für die Klassen 5-10 des Biologieunterrichts (3 zur Samenkeimung, 2 zur Photosynthese, 2 zur Enzymatik) und den Teilkompetenzen „Beurteilungsfähigkeit der Hypothesenbildung" (6 Items), „Beurteilungsfähigkeit der Planung von Experimenten" (8 Items) und die „Auswertung von Daten" (8 Items). Die 22 Items werden mit 0-1 bzw. 0–2 Punkten gescort.

TESTGÜTEKRITERIEN

Reliabilität: Die EAP/PV-Reliabilitätskoeffizienten betragen 0.69 für die Skala „experimentellen Biologieunterricht planen", 0.73 für die Skala „experimentellen Biologieunterricht analysieren" und 0.72 für die Skala „Fachwissen zum Experimentieren". Für den Gesamttest zu Vermittlungskompetenzen liegt die Interrater-Übereinstimmung bei einem Cohens Kappa von 0.82. Für „Beurteilungskompetenz" liegt die Reliabilität bei 0.65 (WLE) bzw. 0.63 (EAP/PV). Die Qualität des Scorings von vier Kodierern wurde über ein vierfach-Scoring von 10 % der Daten kontrolliert. Krippendorffs α liegt für 18 der 22 Items zwischen .61 und .84.
Validität Testinhalt: Die curriculare Validität, die mittels einer Expertenbefragung mit Biologie-Fachdidaktiker*innen und in der Lehrerbildung aktiven Biologielehrkräften zu „Beurteilungskompetenzen" betrachtet wurde, ist durch die Auswahl der lehrplanrelevanten Themen „Enzymatik", „Photosynthese" und „Samenkeimung" gewährleistet.
Validität Antwortprozesse: Die Items der Skalen „experimentellen Biologieunterricht planen", „experimentellen Biologieunterricht analysieren" und „Beurteilungskompetenz" erfassen die kognitiven Aspekte nach den drei Phasen des SDDS-Modells von Klahr (2000). „Aufstellen bzw. Beurteilen naturwissenschaftlicher Hypothesen", „Planung bzw. Beurteilen naturwissenschaftlicher Experimente" und „Auswertung bzw. Beurteilen der Auswertung naturwissenschaftlicher Experimente", die in Studien mittels Lauten Denkens erfasst wurden.
Validität interner Testaufbau: Nach IRT-Modellvergleichen weist ein zweidimensionales Modell mit den beiden Dimensionen „Fachwissen zum Experimentieren" und „experimentellen Biologieunterricht planen und analysieren" (zusammengefasst) die beste Passung auf. Bislang wurde „Beurteilungskompetenz" als eine Dimension modelliert. Mehrdimensionale Modellierungen, die auf Vermittlungs- und Beurteilungskompetenzen gemeinsam fokussieren, stehen noch aus.
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Als externe Variablen wurden zusammen mit Vermittlungskompetenzen die folgenden Kurzskalen eingesetzt: Eine Kurzskala des Testinstruments *Kowadis* (erfasst das Wissen über das Experimentieren, Hartmann et al., 2015) und eine Skala aus dem Berliner Intelligenz Struktur Test (Jäger et al., 1997) zur Erfassung der Fähigkeit zum verbalen Schlussfolgern. Gemeinsam mit Beurteilungskompetenz wurden relevante spezifische Facetten von Selbstwirksamkeitserwartungen zum Unterrichten von Biologie (Mahler, 2014) und diagnostische Kompetenzen (Dübbelde, 2013) erhoben. Zusammenhänge mit den externen Variablen sind weitgehend theoriekonform.
Validität Konsequenzen der Testergebnisse: keine Angaben
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 120-140 Minuten (60-70 Minuten: Unterrichtsanalyse- und Planungskompetenzen; 60-70 Minuten: Beurteilungskompetenz)
Benötigtes Material: Testhefte, Kodiermanual
Besonderheiten: Zur Auswertung der offenen Aufgaben wird ein Kodiermanual bereitgestellt. Pro Item werden 0-1 oder 0–2 Punkte vergeben.

Praxisbeispiel: Die Skalen zu Vermittlungskompetenzen wurden bei 525 Biologie-Lehramtsstudierenden und die Skala zur Beurteilungskompetenz bei 500 Biologie-Lehramtsstudierenden an 19 deutschen Hochschulen erprobt. Die Teilnehmer*innen studierten Lehramt an Gymnasien, Lehramt an Realschulen bzw. Lehramt an einer anderen Schulform (Grundschule, Berufsschule, Hauptschule, Sonderpädagogik). Die Erkenntnisse aus EXMO flossen in die Gestaltung des Lehramtsstudiums für das Fach Biologie ein. Konkret wurde die Lehre des lehramtsbezogenen Bachelor und Master of Education an einer Universität durch eine ergänzende Ausgestaltungsvariante für das biologiedidaktische Forschungspraktikum mit dem Schwerpunkt der Förderung von Experimentierkompetenzen als fachdidaktische Aufgabe weiterentwickelt. An einer anderen Universität wurden Teile verschiedener Pflichtveranstaltungen der Studiengänge Bachelor und Master of Education auf Lehrkompetenzen im experimentellen Biologieunterricht ausgerichtet.

FACHDOMÄNE

Erprobt für: Lehramt Biologie
Prinzipiell auch geeignet für: Lehramt Chemie, Physik Geografie (physisch)
Nicht geeignet für: Schulfächer, in denen nicht naturwissenschaftlich experimentiert wird

ZIELGRUPPE

Erprobt für: Biologie-Lehramtsstudierende
Prinzipiell auch geeignet für: Referendar*innen
Nicht geeignet für: Keine Einschränkung

ALLGEMEINER TESTZWECK

Erprobt für: Erfassung des Kompetenzstands
Prinzipiell auch geeignet für: Erfassung der Kompetenzentwicklung nach Prüfung von Änderungssensitivität
Nicht geeignet für: Keine Einschränkung

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Fachwissenstest zum Experimentieren als zusätzliche Variable, Skalen zur Erfassung von Selbstwirksamkeitserwartungen für das Unterrichten von Biologie, Skalen zur Erfassung diagnostischer Kompetenzen
Prinzipiell auch geeignet für: Der Test kann mit anderen Instrumenten kombiniert werden.
Nicht geeignet für: Keine Einschränkung

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Gruppen- und Studiengangsebene (Lehramtsstudium für das Unterrichtsfach Biologie)
Prinzipiell auch geeignet für: Hochschulebene
Nicht geeignet für: Individualdiagnostik, Landes- und Bundesebene, internationale Ebene

EINSATZSZENARIEN

Erprobt für: Die Erkenntnisse aus EXMO flossen in die Gestaltung des Lehramtsstudiums für das Fach Biologie in die Lehre des Bachelor und Master of Education an zwei Universitäten ein.
Prinzipiell auch geeignet für: Evaluation von Lehrveranstaltungen und zur Beurteilung von Studien- bzw. Arbeitsleistungen, Evaluation von Modulen bzw. Modulteilen der Lehramtsausbildung, die sich der Förderung der entsprechenden Kompetenzen widmen.
Nicht geeignet für: Der Test widmet sich speziellen Aspekten der Lehramtsausbildung, wodurch seine Einsatzszenarien eingeschränkt sind. Ein Einsatz mit zusätzlichen Maßen wird empfohlen.

ERLAUBTE KONSEQUENZEN

Erprobt für: Kompetenzerfassung
Prinzipiell auch geeignet für: Evaluation von Modulen bzw. Modulbestandteilen
Nicht geeignet für: Keine Einschränkung

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Die Tests sollten auf Veränderungssensitivität geprüft werden, um sie zukünftig für die Evaluation von Modulen bzw. Modulbestandteilen der Lehramtsausbildung zu nutzen.
Die Instrumente sind auf Anfrage bei den Testentwickler*innen zum Einsatz für die validierten Testzwecke verfügbar.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Bögeholz, S.*; Joachim, C.*; Hasse, S. & Hammann, M. (2016). Kompetenzen von (angehenden) Biologielehrkräften zur Beurteilung von Experimentierkompetenzen. *Unterrichtswissenschaft, 44*(1), S. 40-54. [*gemeinsame Erstautorenschaft]
Hasse, S.; Joachim, C.; Bögeholz, S. & Hammann, M. (2014). Assessing Teaching and Assessment Competences of Biology Teacher Trainees: Lessons from Item Development. *International Journal of Education in Mathematics, Science and Technology, 2*(3), S. 191-205.

WEITERFÜHRENDE LITERATUR

Hammann, M. (2004). Kompetenzentwicklungsmodelle: Merkmale und ihre Bedeutung – dargestellt anhand von Kompetenzen beim Experimentieren. *MNU – Der mathematisch-naturwissenschaftliche Unterricht, 57*(4), S. 196-203.
Hammann, M.; Phan, T. T. M.; Ehmer, M. & Bayrhuber, H. (2006). Fehlerfrei Experimentieren. *MNU – Der mathematisch-naturwissenschaftliche Unterricht, 59*(5), S. 292-299.
Klahr, D. (2000). *Exploring Science: The Cognition and Development of Discovery Processes.* Cambridge: MIT Press.

HEED: Unternehmensgründungstest

KURZFASSUNG

Name des Instruments: Test zur Unternehmensgründung

Fachdomäne: Unternehmensgründung

Erfasste Kompetenzen: Fachwissen zur Unternehmensgründung

Zielgruppe: Studierende aller Fachrichtungen

Testart: Leistungstest; Multiple Choice

Modalität: Paper-Pencil-Test

Dauer: 45 Minuten

Testaufbau: 40 Items; 10 Inhaltsbereiche: Gründungsmarketing, Gründungsbezogene Personalwirtschaft, Gründungsrelevante Rechtsformen, Schutzrechte, Innovations- und Technologiemanagement, Rechnungswesen/Gründungsbezogenes Controlling, Gründungsfinanzierung, Strategisches Management und Geschäftsmodell, Businessplanung, Gründungsbezogenes Projektmanagement

Versionen: Zwei Teilversionen (Version A und B)

Allgemeiner Testzweck: Erfassen des Wissensstands

Konkrete Einsatzszenarien: Lehr-Lern-Tool, Kursvergleiche, Hochschulvergleiche

Nicht geeignet für: —

Hinweise für praktische Anwendung: Eine Instruktion, Lösungsheft zur Testauswertung sowie weitere Materialien für die Testdurchführung werden bereitgestellt. Aus allen korrekt gelösten Aufgaben wird ein Summenscore gebildet. Auswertung einzelner inhaltlicher Teilbereiche oder Wissensniveaus möglich.

Bisheriger Einsatz in Projekten: HEED: Higher Entrepreneurship Education Diagnostics

Kontakt / Speicherort:
Prof. Dr. Michael Schefczyk
Lehrstuhl für Entrepreneurship und Innovation,
Fakultät Wirtschaftswissenschaften, TU Dresden
E-Mail: info@gruenderlehrstuhl.de
Webseite: https://tu-dresden.de/bu/wirtschaft/lei/die-professur/beschaeftigte/schefczyk_profil

ALLGEMEINE INFORMATIONEN

Name des Instruments: Test zur Unternehmensgründung
Bisheriger Einsatz in Projekten: HEED: Higher Entrepreneurship Education Diagnostics
Kontakt / Speicherort:
Prof. Dr. Michael Schefczyk
Lehrstuhl für Entrepreneurship und Innovation,
Fakultät Wirtschaftswissenschaften, TU Dresden
E-Mail: info@gruenderlehrstuhl.de
Webseite: https://tu-dresden.de/bu/wirtschaft/lei/die-professur/beschaeftigte/schefczyk_profil

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Fachwissen zur Unternehmensgründung
Theoriemodell: HEED Kompetenzmodell mit 2 Niveaus (1. Niveau: Kenntnisse, d. h. Erinnern und Verstehen; 2. Niveau: Wissenstransformation, d. h. Anwenden, Analysieren und Bewerten)
Testart: Leistungstest; Multiple Choice
Modalität: Paper-Pencil-Test
Testaufbau:
Itempool: 40 Items. 10 Inhaltsbereiche: Gründungsmarketing, Gründungsbezogene Personalwirtschaft, Gründungsrelevante Rechtsformen, Schutzrechte, Innovations- und Technologiemanagement, Rechnungswesen/Gründungsbezogenes Controlling, Gründungsfinanzierung, Strategisches Management und Geschäftsmodell, Businessplanung, Gründungsbezogenes Projektmanagement. Die Inhaltsbereiche stammen aus dem HEED-Kompetenzmodell zu unternehmerischer Kompetenz.
Versionen: Zwei parallele Versionen (Version A und B)

TESTGÜTEKRITERIEN

Reliabilität: keine Angaben
Validität: keine Angaben
Validität Antwortprozesse: keine Angaben
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 45 Minuten
Benötigtes Material: Testform, Lösungsform, Beiheft
Besonderheiten: Für die Testdurchführung und -auswertung werden eine Instruktion und ein Lösungsheft bereitgestellt. Alle Antwortmöglichkeiten müssen korrekt beantwortet werden. Die richtig gelösten Aufgaben werden addiert. Auswertung einzelner inhaltlicher Teilbereiche oder Kompetenzniveaus möglich. Folgende weiteren Möglichkeiten stehen für die Testauswertung zur Verfügung: 1. Addition aller richtig gelösten N1-Aufgaben (Summenscore Kompetenzniveau 1). 2. Addition aller richtig gelösten N2-Aufgaben (Summenscore Kompetenzniveau 2). 3. Addition aller richtig gelösten Aufgaben eines Inhaltsbereiches (Summenscore der Inhaltsbereiche 1-10).
Praxisbeispiel: Der Test wurde mit unterschiedlichen Studierendengruppen (insgesamt N = 203 Personen) mit zwei parallelen Testversionen (Version A und B) erprobt. Mit dieser Stichprobe wurde eine erste Validierung vorgenommen und die Items überprüft und weitreichend überarbeitet. Diese müssen mit einer neuen Stichprobe überprüft werden.

FACHDOMÄNE

Erprobt für: Unternehmensgründung
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

ZIELGRUPPE

Erprobt für: Studierende unterschiedlicher Fachrichtungen an zwei Universitäten, in unterschiedlichen Qualifikationsstufen, technologie- und wissensbasierte Gründungen
Prinzipiell auch geeignet für: Studierende aller Fachrichtungen
Nicht geeignet für: Im Hochschulbereich stehen technologieorientierte und wissensbasierte Unternehmensgründungen im Fokus. Entsprechend sind einzelne Fachkompetenzen, z. B. Schutzrechte, für Low-tech-Gründer*innen weniger relevant.

ALLGEMEINER TESTZWECK

Erprobt für: Erfassen des Wissensstands
Prinzipiell auch geeignet für: Erfassen der Wissensentwicklung, wiederholte Testung
Nicht geeignet für: —

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Einen soziodemografischen Teil
Prinzipiell auch geeignet für: Weitere Maße zur Kompetenzerfassung
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für die Kursebene.
Prinzipiell auch geeignet für: Einzelperson im zeitlichen Verlauf, Hochschulebene
Nicht geeignet für: —

EINSATZSZENARIEN

Erprobt für: Kursvergleiche
Prinzipiell auch geeignet für: Lehr-Lern-Tool, Hochschulvergleiche
Nicht geeignet für: —

ERLAUBTE KONSEQUENZEN

Erprobt für: Beratung von Teilnehmenden
Prinzipiell auch geeignet für: Bisher liegen keine Validierungsbefunde zu den überarbeiteten Items vor, sodass noch keine Konsequenzen abgeleitet werden können.
Nicht geeignet für: —

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Nach ersten Validierungsstudien wurden die Items überarbeitet, sodass die überarbeiteten Aufgaben an einer neuen Stichprobe überprüft werden müssen, bevor zuverlässige und gültige Aussagen möglich sind. Das überarbeitete Testinstrument ist auf Anfrage bei den Testentwickler*innen erhältlich.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Weinberger, E. & Schefczyk, M. (2015). *HEED – Higher Entrepreneurship Education Diagnostics. Teilprojekt: TIP – The Instructor Perspective. Entwicklung eines Kompetenzmodells zur unternehmerischen Kompetenz.* Abschlussbericht. Dresden: Technische Universität Dresden.

WEITERFÜHRENDE LITERATUR

Braukmann, U. & Schneider, D. (2007). Didaktische Zielklassen der Entwicklung unternehmerischer Kompetenz. In R. Bader; G. Keiser & T. Unger (Hrsg.), *Entwicklung unternehmerischer Kompetenz in der Berufsbildung. Hintergründe, Ziele und Prozesse berufspädagogischen Handelns* (S. 157–180). Bielefeld: W. Bertelsmann Verlag.
Haase, H. & Lautenschläger, A. (2010). The 'Teachability Dilemma' of entrepreneurship. *International Entrepreneurship and Management Journal, 7*(2), S. 145-162. DOI: 10.1007/s11365-010-0150-3

KOM@ING KAT: Kategorisierungssysteme

KURZFASSUNG

Name des Instruments: Kategorisierungssysteme (KAT)

Fachdomäne: Elektrotechnik Grundlagen, mathematische Anteile

Erfasste Kompetenzen: Kompetenzen in der inhaltlichen Kategorisierung von Lösungen in Ingenieurwissenschaften / Grundlagenaufgaben der Elektrotechnik

Zielgruppe: Ingenieurstudierende in Grundlagenveranstaltungen

Testart: Klausur unter Prüfungsbedingungen

Modalität: Paper-Pencil-Test

Dauer: Klausur (120 Minuten) plus inhaltliche Kategorisierung

Testaufbau: Klausur mit offenen und geschlossenen Items, Kategorienraster zur Verdeutlichung mathematischer Anteile bei der Bearbeitung der Aufgaben aus der Elektrotechnik (mit Teilkompetenzen wie der Unterteilung von Formelaufstellen und Berechnen einer Größe mit der aufgestellten Formel)

Allgemeiner Testzweck: Hilfsinstrument zur inhaltlichen Analyse von schriftlichen Antworten zu Grundlagenaufgaben aus der Elektrotechnik unter Klausurbedingungen; quantitative Sicherung der Ergebnisse anderer Studien

Konkrete Einsatzszenarien: Inhaltliche Analysen von (Teil-)Kompetenzen aus Klausurbearbeitungen

Nicht geeignet für: Reine Mathematik

Hinweise für praktische Anwendung: Bestandteil eines Sets aus vier Instrumenten, die gemeinsam eingesetzt werden und sich gegenseitig ergänzen (KOM@ING Experteninterview [PARI], Studierende-Expert-Lösung [SEL], Niedrig Inferente Analysen [NIA]).
Im Projekt wurde auf eine bereits geschriebene Klausur zurückgegriffen. Durch die Kategorisierung konnten – über die ursprüngliche Korrektur hinaus – weitere differenziertere inhaltliche Analysen der Kompetenzen der Studierenden vorgenommen werden.

Bisheriger Einsatz in Projekten: KOM@ING-Paderborn

Kontakt / Speicherort:
Dr. Jörg Kortemeyer
TU Clausthal
E-Mail: joerg.kortemeyer@tu-clausthal.de

ALLGEMEINE INFORMATIONEN

Name des Instruments: Kategorisierungssysteme (KAT)
Bisheriger Einsatz in Projekten: KOM@ING-Paderborn
Kontakt / Speicherort:
E-Mail: joerg.kortemeyer@tu-clausthal.de

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Kompetenzen in der inhaltlichen Kategorisierung von Lösungen in Ingenieurwissenschaften / Grundlagenaufgaben der Elektrotechnik
Theoriemodell: Quantitative, evaluierende Inhaltsanalysen
Testart: Klausur unter Prüfungsbedingungen
Modalität: Paper-Pencil-Test
Testaufbau:
Aufgrund des Rückgriffs auf bereits vorhandene Klausurdaten lagen hier übliche Klausurbedingungen in einer zweistündigen Klausur ohne Hilfsmittel vor.

TESTGÜTEKRITERIEN

Reliabilität: Durch Korrektur der Klausur sowie Interrater-Reliabilität (über 80 %)
Validität Testinhalt: Verwendung von realen Aufgaben aus einer Klausur
Validität Antwortprozesse: Mittels Durchführung von Studien zu Bearbeitungsprozessen geprüft
Validität interner Testaufbau: keine Angaben
Validität Zusammenhänge mit anderen Kompetenzindikatoren: keine Angaben
Validität Konsequenzen der Testergebnisse: keine Angaben
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: Klausur (120 Minuten) plus inhaltliche Kategorisierung
Benötigtes Material: Klausuraufgaben und Antworten der Studierenden unter Klausurbedingungen
Besonderheiten: —
Praxisbeispiel: Das Instrument wurde zur Analyse von Antworten der Studierenden aus einer Klausur eingesetzt, welche im Jahr 2012 an der Universität Paderborn zu der Veranstaltung Grundlagen der Elektrotechnik B (GET-B) geschrieben wurde und an der 90 Studierende teilgenommen haben. Die Ergebnisse der Analysen der Klausur wurden an die Lehrenden zurückgemeldet. Deren über die (PARI) Experteninterviews erfassten Erwartungen unterschieden sich häufig von den tatsächlichen Fehlern der Studierenden. Dabei war ein entscheidender Punkt in den Analysen die Trennung von mathematikbezogenen Teilkompetenzen, wie z. B. vom Aufstellen einer Formel und Ermitteln eines Ergebnisses mithilfe dieser Formel. Die Studierenden konnten entweder beides oder keine dieser beiden Kompetenzen ausführen: Wer also die Formel richtig aufstellte, konnte i. d. R. auch mit ihr das richtige Ergebnis bestimmen. Die Klausuren selbst wurden anonymisiert analysiert, sodass kein Einzelfeedback an die Teilnehmer*innen gegeben werden konnte. Änderungen an der Klausurpraxis wurden vorgenommen, z. B. hinsichtlich des Umfangs und der Art von vorgegebenen Informationen in den Aufgaben. Änderungen, die sich in den Curricula ergeben haben, sind nicht bekannt.

FACHDOMÄNE

Erprobt für: Mathematische Anforderungen in Elektrotechnik-Grundlagen
Prinzipiell auch geeignet für: Mathematische Anforderungen in allgemeinen Ingenieuraufgaben beliebiger Semester
Nicht geeignet für: Innermathematische Aufgaben

ZIELGRUPPE

Erprobt für: Lehrende und Studierende der Ingenieurwissenschaften in Grundlagenveranstaltungen
Prinzipiell auch geeignet für: Ingenieurstudierende verschiedener Semester
Nicht geeignet für: Reine Mathematik

ALLGEMEINER TESTZWECK

Erprobt für: Quantitative Sicherung der Ergebnisse anderer Studien zum Kompetenzstand, quantitatives Feedback zu mathematischen Anforderungen
Prinzipiell auch geeignet für: Einteilung eines Lösungsspektrums zum Standard-Setting
Nicht geeignet für: Analysen qualitativer Lösungsprozesse

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Experteninterviews, Studierende-Expert-Lösungen und NIAs
Prinzipiell auch geeignet für: Tests zu mathematischen Kompetenzen wie der LiMSt (Lernstrategien in mathematikhaltigen Studienfächern); andere Kompetenzmaße
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für Einzelpersonen (parallele Bearbeitung).
Prinzipiell auch geeignet für: Gruppenebene, Hochschulebene
Nicht geeignet für: Landes- und Bundesebene, internationale Ebene

EINSATZSZENARIEN

Erprobt für: Analysen von mathematischen (Teil-)Kompetenzen in Klausurbearbeitungen
Prinzipiell auch geeignet für: Analyse von Antworten unterschiedlicher Prüfungen (z. B. auch mündlich)
Nicht geeignet für: —

ERLAUBTE KONSEQUENZEN

Erprobt für: Optimierung von Muster-Klausurlösungen für Studierende
Prinzipiell auch geeignet für: Kompetenzorientierte Umgestaltung von Eingangsphasen in den Ingenieurstudiengängen
Nicht geeignet für: —

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Das Instrument wurde im Rahmen der Dissertation von Jörg Kortemeyer eingesetzt, welche unter go.upb.de/kortemeyer abrufbar ist. Es besteht Kontakt zu weiteren Institutionen, die die Instrumente einsetzen möchten. Es soll zur Optimierung von Muster-Aufgabenlösungen für Studierende eingesetzt werden.

WEITERE WICHTIGE ANGABEN

Das Instrument ist im Kapitel 3 der Dissertation von Jörg Kortemeyer näher beschrieben und auf Anfrage verfügbar.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Biehler, R., Kortemeyer, J. & Schaper, N. (2015). Conceptualizing and studying students' processes of solving typical problems in introductory engineering courses requiring mathematical competences. In K. Krainer & N. Vondrová (Hrsg.), *CERME9: Proceedings of the Ninth Congress of the European Society for Research in Mathematics Education* (S. 2060–2066). Prague: Charles University.
Kortemeyer, J. & Biehler, R. (2017). The interface between mathematics and engineering – problem solving processes for an exercise on oscillating circuits using ordinary differential equations. In *Proceedings of the Tenth Conference of the European Society for Research in Mathematics Education.*
Kortemeyer, J. (2018). *Mathematische Kompetenzen in ingenieurwissenschaftlichen Grundlagenveranstaltungen: Normative und empirische Analysen zu exemplarischen Klausuraufgaben aus dem ersten Studienjahr in der Elektrotechnik* (Dissertation).

KOM@ING NIA: Niedrig Inferente Analysen

KURZFASSUNG

Name des Instruments: Niedrig Inferente Analysen (NIA)

Fachdomäne: Elektrotechnik Grundlagen

Erfasste Kompetenzen: Denkprozesse von Studierenden bei der Bearbeitung von Grundlagenaufgaben in mathematischen Anwendungsfächern

Zielgruppe: Ingenieurstudierende in Grundlagenveranstaltungen

Testart: Kooperativer Leistungstest, Lautes Denken mit videografierten Bearbeitungsprozessen und retrospektivem Interview

Modalität: Paper-Pencil-Test

Dauer: 60-90 Minuten

Testaufbau: Videografieren von Bearbeitungsprozessen und Interview

Allgemeiner Testzweck: Hilfsinstrument zur Identifizierung von Denkprozessen und Kompetenzen von Studierenden bei der Bearbeitung von Ingenieursaufgaben

Konkrete Einsatzszenarien: Lehr-Lern-Tool zum Erfassen der Denkprozesse von Studierenden bei der Bearbeitung von Grundlagenaufgaben sowie Vergleiche dieser Denkprozesse

Nicht geeignet für: Reine Mathematik

Hinweise für praktische Anwendung: Bestandteil eines Sets aus vier Instrumenten, die gemeinsam eingesetzt werden und sich gegenseitig ergänzen (KOM@ING Experteninterview [PARI], Studierende-Expert-Lösung [SEL], Kategorisierungssysteme [KAT]).

Bisheriger Einsatz in Projekten: KOM@ING-Paderborn

Kontakt / Speicherort:
Dr. Jörg Kortemeyer
TU Clausthal
E-Mail: joerg.kortemeyer@tu-clausthal.de

ALLGEMEINE INFORMATIONEN

Name des Instruments: Niedrig Inferente Analysen (NIA)
Bisheriger Einsatz in Projekten: KOM@ING
Kontakt / Speicherort:
Dr. Jörg Kortemeyer, TU Clausthal
E-Mail: joerg.kortemeyer@tu-clausthal.de

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Denkprozesse von Studierenden bei der Bearbeitung von Grundlagenaufgaben in mathematischen Anwendungsfächern
Theoriemodell: Qualitative Inhaltsanalyse
Testart: Kooperativer Leistungstest, Lautes Denken bei videografierten Bearbeitungsprozessen und retrospektivem Interview
Modalität: Paper-Pencil-Test
Testaufbau:
Videografieren von Bearbeitungsprozessen und Interview

TESTGÜTEKRITERIEN

Reliabilität: Analyse der Bearbeitungsprozesse durch weitere Elektrotechnik-Experten; hohe Übereinstimmung
Validität Testinhalt: Vergleich mit Curricula verschiedener Universitäten
Validität Antwortprozesse: Vergleich mit Studierendenlösungen
Validität interner Testaufbau: keine Angaben
Validität Zusammenhänge mit anderen Kompetenzindikatoren: keine Angaben
Validität Konsequenzen der Testergebnisse: keine Angaben
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 60-90 Minuten
Benötigtes Material: Aufgabenstellung, zwei Kameras oder eine Kamera und ein Visualizer
Besonderheiten: Bestandteil eines Sets aus vier Instrumenten, die gemeinsam eingesetzt werden und sich gegenseitig ergänzen (KOM@ING Experteninterview [PARI], Studierende-Expert-Lösung [SEL], Kategorisierungssysteme [KAT]).

Praxisbeispiel: Die Niedrig Inferenten Analysen wurden auf Bearbeitungsprozesse von Studierendenpaaren angewendet, welche an den Universitäten Paderborn und Hannover 2013 gefilmt wurden. Zu diesen Bearbeitungsprozessen wurden Transskripte erstellt, welche mit Methoden der qualitativen Inhaltsanalyse untersucht wurden. Bei den Videostudien wurden Paare von Studierenden aus dem zweiten Semester gebeten, eine Aufgabe mittels lautem Denken zu lösen. Von den Anwendern gab es das Feedback, dass selbst durch die Abstimmung in den Zweier-Teams zusätzliche Kompetenzen gegenseitig deutlich wurden. Änderungen in Studienordnungen und Curricula sind nicht bekannt und erfordern einen hohen Aufwand, da der Kompetenzerwerb parallel in zwei Veranstaltungen (Mathematik für Ingenieure, Grundlagen der Elektrotechnik) stattfindet.

FACHDOMÄNE

Erprobt für: Elektrotechnik-Grundlagen
Prinzipiell auch geeignet für: Allgemeine Ingenieuraufgaben beliebiger Semester
Nicht geeignet für: Innermathematische Aufgaben

ZIELGRUPPE

Erprobt für: Ingenieurstudierende in Grundlagenveranstaltungen
Prinzipiell auch geeignet für: Ingenieurstudierende verschiedener Semester
Nicht geeignet für: Reine Mathematik

ALLGEMEINER TESTZWECK

Erprobt für: Identifizierung von mentalen Prozessen und Kompetenzen von Studierenden bei der Bearbeitung von Ingenieuraufgaben
Prinzipiell auch geeignet für: Erzeugen einer erweiterten Musterlösung
Nicht geeignet für: —

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Kombination mit Experteninterviews, Studierende-Expert-Lösungen und Kategorisierungen
Prinzipiell auch geeignet für: Tests zu mathematischen Kompetenzen wie der LiMSt; andere Kompetenzmaße
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für mehrfache Durchführung bezogen auf dieselbe Aufgabe in Gruppen.
Prinzipiell auch geeignet für: Analyse von Antworten der Studierenden
Nicht geeignet für: Individualdiagnostik, Hochschulebene, internationale Ebene

EINSATZSZENARIEN

Erprobt für: Lehr-Lern-Tool, um das Denken von Studierenden während der Bearbeitung von Grundlagenaufgaben zu erfassen, Vergleiche des lauten Denkens
Prinzipiell auch geeignet für: Erstellung eines Kompetenzkatalogs für Grundlagenveranstaltungen
Nicht geeignet für: —

ERLAUBTE KONSEQUENZEN

Erprobt für: Optimierung von Rückmeldungen an Studierende
Prinzipiell auch geeignet für: Kompetenzorientierte Umgestaltung von Eingangsphasen von Ingenieurstudiengängen
Nicht geeignet für: —

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Das Instrument wurde im Rahmen der Dissertation von Jörg Kortemeyer eingesetzt, welche unter go.upb.de/kortemeyer abrufbar ist. Es besteht Kontakt zu weiteren Institutionen, die die Instrumente einsetzen möchten. Es soll zur Optimierung von Muster-Aufgabenlösungen für Studierende eingesetzt werden.

WEITERE WICHTIGE ANGABEN

Das Instrument ist im Kapitel 3 der Dissertation von Jörg Kortemeyer näher beschrieben.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Kortemeyer, J.; Biehler, R. & Schaper, N. (2014). *Hilft der sogenannte Modellierungskreislauf Lösungsprozesse bei ingenieurwissenschaftlichen Anwendungsaufgaben besser zu verstehen?* Dortmund: Universitätsbibliothek Dortmund.

Kortemeyer, J. & Biehler, R. (2017). The interface between mathematics and engineering – problem solving processes for an exercise on oscillating circuits using ordinary differential equations. In *Proceedings of the Tenth Conference of the European Society for Research in Mathematics Education.*

Kortemeyer, J. (2018). *Mathematische Kompetenzen in ingenieurwissenschaftlichen Grundlagenveranstaltungen: Normative und empirische Analysen zu exemplarischen Klausuraufgaben aus dem ersten Studienjahr in der Elektrotechnik* (Dissertation).

WEITERFÜHRENDE LITERATUR

Biehler, R., Kortemeyer, J. & Schaper, N. (2015). Conceptualizing and studying students' processes of solving typical problems in introductory engineering courses requiring mathematical competences. In K. Krainer & N. Vondrová (Hrsg.), *CERME9: Proceedings of the Ninth Congress of the European Society for Research in Mathematics Education* (S. 2060–2066). Prague: Charles University.

KOM@ING SEL: Studierende-Expert-Lösung

KURZFASSUNG

Name des Instruments: Studierende-Expert-Lösung (SEL)

Fachdomäne: Elektrotechnikaufgaben in Grundlagenveranstaltungen

Erfasste Kompetenzen: Normative Lösungskompetenz zu mathematischen Grundlagenaufgaben in der Elektrotechnik

Zielgruppe: Ingenieurstudierende in Grundlagenveranstaltungen

Testart: Normative Aufgabenlösung durch Expert*innen

Modalität: Paper-Pencil-Test

Dauer: 10 Stunden (inkl. der Einarbeitung in die vorgegebene Lösung der Studierenden, Interview, erstellen der Expertenlösung)

Testaufbau: Zwei Teile: „Aufgabenlösung", angereichert um ausführlichere Anmerkungen in mathematischer und elektrotechnischer Sicht; „Kognitive Ressourcen" für zusätzliche Anmerkungen zur Aufgabenlösung

Allgemeiner Testzweck: Hilfsinstrument zum Erstellen einer Musterlösung; Definition einer Norm unter Verwendung von speziellen Experteninterviews; Schaffung einer Analysegrundlage für qualitative und quantitative Studien (NIAs; KAT)

Konkrete Einsatzszenarien: Lösung zu vier Aufgaben in der Dissertation von Jörg Kortemeyer als Basis für weitere qualitative Analysen (NIAs) und quantitative Analysen (Kategorisierungen)

Nicht geeignet für: Reine Mathematik

Hinweise für praktische Anwendung: Bestandteil eines Sets aus vier Instrumenten, die gemeinsam eingesetzt werden und sich gegenseitig ergänzen (Experteninterview [PARI], Kategorisierungssysteme [KAT], Niedrig Inferente Analysen [NIA]).

Bisheriger Einsatz in Projekten: KOM@ING-Paderborn

Kontakt / Speicherort:
Dr. Jörg Kortemeyer
TU Clausthal
E-Mail: joerg.kortemeyer@tu-clausthal.de

ALLGEMEINE INFORMATIONEN

Name des Instruments: Studierende-Expert-Lösung (SEL)
Bisheriger Einsatz in Projekten: KOM@ING-Paderborn
Kontakt / Speicherort:
Dr. Jörg Kortemeyer
TU Clausthal
E-Mail: joerg.kortemeyer@tu-clausthal.de

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Normative Lösungskompetenz zu mathematischen Grundlagenaufgaben in der Elektrotechnik
Theoriemodell: Normative mathematikdidaktische Aufgabenanalysen
Testart: Normative Aufgabenlösung
Modalität: Paper-Pencil-Test
Testaufbau:
Zwei Teile: „Aufgabenlösung", angereichert um ausführlichere Anmerkungen in mathematischer und elektrotechnischer Sicht; „Kognitive Ressourcen" für zusätzliche Anmerkungen zur Aufgabenlösung

TESTGÜTEKRITERIEN

Reliabilität: keine Angaben
Validität Testinhalt: Vergleich mit Curricula verschiedener Universitäten und Analysen durch weitere Experten
Validität Antwortprozesse: Vergleich mit Studierendenlösungen
Validität interner Testaufbau: keine Angaben
Validität Zusammenhänge mit anderen Kompetenzindikatoren: keine Angaben
Validität Konsequenzen der Testergebnisse: keine Angaben
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 10 Stunden (inkl. Ausarbeiten der Kurzlösung, welche dann über die Experteninterviews um Punkte wie Kompetenzerwartungen ergänzt wurde)
Benötigtes Material: „Musterlösung" zu den Aufgaben (z. B. für Korrektoren), Experteninterview zur Erfassung der Kompetenzen
Besonderheiten: Bestandteil eines Sets aus vier Instrumenten, die gemeinsam eingesetzt werden und sich gegenseitig ergänzen (Experteninterview [PARI], Kategorisierungssysteme [KAT], Niedrig Inferente Analysen [NIA]).

Praxisbeispiel: Das Instrument wurde zur Analyse von Aufgabenantworten aus Videostudien sowie Klausuren entwickelt, um eine Interpretationsgrundlage zu erarbeiten. Das Instrument wurde seit 2013 in Studien an der Universität Paderborn sowie der Leibniz Universität Hannover zur Analyse von Bearbeitungsprozessen und -ergebnissen eingesetzt. Der Einsatz liefert umfassende Informationen, die – neben einer Aufgabenbearbeitung, welche in den Musterlösungen häufig nur stichpunktartig dargestellt war – von den Bearbeiter*innen berücksichtigt werden sollten. Diese schließen bezogen auf die Elektrotechnik viele physikalische Überlegungen sowie Überlegungen zur Verringerung des mathematischen Aufwands mit ein. Konsequenz der Ergebnisse ist ein stärkerer Fokus der Lehrenden auf die Erwartungen, die an Studierende bei Aufgabenbearbeitung gestellt werden, sowie auf implizite Kompetenzerwartungen, welche sich im Rahmen der Erstellung der Studierende-Expert-Lösung gezeigt haben.

FACHDOMÄNE

Erprobt für: Elektrotechnik Grundlagen
Prinzipiell auch geeignet für: Allgemeine Ingenieuraufgaben beliebiger Semester
Nicht geeignet für: Innermathematische Aufgaben

ZIELGRUPPE

Erprobt für: Ingenieurstudierende in Grundlagenveranstaltungen
Prinzipiell auch geeignet für: Ingenieurstudierende verschiedener Semester
Nicht geeignet für: Reine Mathematik

ALLGEMEINER TESTZWECK

Erprobt für: Hilfsinstrument zum Erstellen einer Musterlösung; Definition einer Norm unter Verwendung von speziellen Experteninterviews; Schaffung einer Analysegrundlage für qualitative und quantitative Studien. Das Instrument beschreibt die Schnittstelle zwischen den Kompetenzbereichen Mathematik und Elektrotechnik näher und berücksichtigt insbesondere Kompetenzerwartungen der Lehrenden.
Prinzipiell auch geeignet für: Erzeugung einer erweiterten Musterlösung für Studierende
Nicht geeignet für: —

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Kombination mit (PARI) Experteninterviews, NIAs und Kategorisierungen
Prinzipiell auch geeignet für: Tests zu mathematischen Kompetenzen wie der LiMSt (Lernstrategien in mathematikhaltigen Studienfächern), andere Kompetenzmaße
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für parallele Erstellung durch mehrere Expert*innen unterschiedlicher Institutionen
Prinzipiell auch geeignet für: Bewertungsgrundlage für Klausuren
Nicht geeignet für: —

EINSATZSZENARIEN

Erprobt für: Lösung zu vier Aufgaben in der Dissertation von Jörg Kortemeyer als Basis für weitere qualitative Analysen (NIAs) und quantitative Analysen (Kategorisierungen)
Prinzipiell auch geeignet für: Sämtliche Aufgaben, bei denen Schnittstellen zwischen Mathematik und anderen Fächern auftreten
Nicht geeignet für: —

ERLAUBTE KONSEQUENZEN

Erprobt für: Optimierung von Muster-Klausurlösungen für Studierende
Prinzipiell auch geeignet für: Kompetenzorientierte Umgestaltung von Eingangsphasen von Ingenieurstudiengängen
Nicht geeignet für: —

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Das Instrument wurde im Rahmen der Dissertation von Jörg Kortemeyer eingesetzt, welche unter go.upb.de/kortemeyer abrufbar ist. Es besteht Kontakt zu anderen Institutionen, die die Instrumente einsetzen möchten. Es soll zur Optimierung von Aufgabenlösungen, welche an Studierende gegeben werden, eingesetzt werden. Der Test kann mit den Beschreibungen in der oben genannten Dissertation durchgeführt werden.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Kortemeyer, J., Biehler, R. & Schaper, N. (2014). *Hilft der sogenannte Modellierungskreislauf Lösungsprozesse bei ingenieurwissenschaftlichen Anwendungsaufgaben besser zu verstehen?* Dortmund: Universitätsbibliothek Dortmund.
Kortemeyer, J. & Biehler, R. (2017). The interface between mathematics and engineering – problem solving processes for an exercise on oscillating circuits using ordinary differential equations. In: *Proceedings of the Tenth Conference of the European Society for Research in Mathematics Education.*
Kortemeyer, J. (2018). *Mathematische Kompetenzen in ingenieurwissenschaftlichen Grundlagenveranstaltungen: Normative und empirische Analysen zu exemplarischen Klausuraufgaben aus dem ersten Studienjahr in der Elektrotechnik* (Dissertation).

WEITERFÜHRENDE LITERATUR

Biehler, R., Kortemeyer, J. & Schaper, N. (2015). Conceptualizing and studying students' processes of solving typical problems in introductory engineering courses requiring mathematical competences. In K. Krainer & N. Vondrová (Hrsg.), *CERME9: Proceedings of the Ninth Congress of the European Society for Research in Mathematics Education* (S. 2060–2066). Prague: Charles University.

KOM@ING: Experteninterview PARI

KURZFASSUNG

Name des Instruments: KOM@ING Experteninterview (PARI)

Fachdomäne: Mathematik in der Elektrotechnik

Erfasste Kompetenzen: Kompetenzerwartungen und normative Anforderungen von Lehrenden/anderen Expert*innen zu Kompetenzen von Studierenden in Mathematik in der Elektrotechnik

Zielgruppe: Lehrende/andere Experten der Mathematik in der Elektrotechnik (in Grundlagenveranstaltungen)

Testart: Teilstrukturiertes Interview

Modalität: Zwei-Personen-Interview

Dauer: 30-60 Minuten

Testaufbau: Drei Phasen: Aufgabenbearbeitung, Rekapitulation (fünf Fragen pro Aufgabenschritt), didaktische Rekonstruktion (drei Fragen pro Aufgabenschritt, zwei übergeordnete Fragen)

Allgemeiner Testzweck: Hilfsinstrument zur Optimierung von Testaufgaben: Erfassen von Kompetenzerwartungen und normative Anforderungen bei Expert*innen/Lehrenden

Konkrete Einsatzszenarien: Unterstützung bei der Erstellung der Aufgaben und Musterlösungen für Studierende und Anreicherung von Hinweisen von Expert*innen

Nicht geeignet für: Reine Mathematik

Hinweise für praktische Anwendung: Bestandteil eines Sets aus vier Instrumenten, die gemeinsam eingesetzt werden und sich gegenseitig ergänzen (Kategorisierungssysteme [KAT], Studierende-Expert-Lösung [SEL], Niedrig Inferente Analysen [NIA]).

Bisheriger Einsatz in Projekten: KOM@ING-Paderborn

Kontakt / Speicherort:
Dr. Jörg Kortemeyer
TU Clausthal
E-Mail: joerg.kortemeyer@tu-clausthal.de

ALLGEMEINE INFORMATIONEN

Name des Instruments: KOM@ING Experteninterview (PARI)
Bisheriger Einsatz in Projekten: KOM@ING-Paderborn
Kontakt / Speicherort:
Dr. Jörg Kortemeyer
TU Clausthal
E-Mail: joerg.kortemeyer@tu-clausthal.de

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Kompetenzerwartungen und normative Anforderungen von Lehrenden/anderen Expert*innen zu Kompetenzen von Studierenden in Mathematik in der Elektrotechnik
Theoriemodell: PARI-Methodik (PARI: Precursor-Action-Response-Interpretation; nach Means et al.)
Testart: Teilstrukturiertes Interview
Modalität: Zwei-Personen-Interview
Testaufbau:
Drei Phasen: Aufgabenbearbeitung, Rekapitulation (fünf Fragen pro Aufgabenschritt), didaktische Rekonstruktion (drei Fragen pro Aufgabenschritt, zwei übergeordnete Fragen)

TESTGÜTEKRITERIEN

Reliabilität: Abgleich mit Musterlösungen aus anderen Studien; hohe Übereinstimmung
Validität Testinhalt: Entsprechend curricularer Vergleiche mit anderen Universitäten sowie Klausuraufgaben repräsentativ
Validität Antwortprozesse: Vergleich mit Musterlösungen aus anderen Studien
Validität interner Testaufbau: keine Angaben
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Bestandteil eines Sets aus vier Instrumenten, die gemeinsam eingesetzt werden und sich gegenseitig ergänzen (Kategorisierungssysteme [KAT], Studierende-Expert-Lösung [SEL], Niedrig Inferente Analysen [NIA]).
Validität Konsequenzen der Testergebnisse: Konsequenzen für die Erstellung der Aufgaben und Musterlösungen für Studierende
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 30-60 Minuten
Benötigtes Material: Interviewleitfaden, Videokamera, Visualizer, Aufgabentext
Besonderheiten: —
Praxisbeispiel: Das Instrument wurde 2013 an der Universität Paderborn für Interviews mit wissenschaftlichen Mitarbeiter*innen vom Institut für Leistungselektronik und Antriebstechnik (LEA) eingesetzt.

Hierzu wurden die Mitarbeiter*innen gebeten, Aufgaben aus der Grundlagenveranstaltung GET-B unter einem Visualizer zu lösen und dabei laut zu denken, wobei der Visualizer von oben die Mitschrift filmte und eine Tonaufnahme durchführte. Mithilfe einer Videokamera wurden Gesten erfasst. In einer zweiten Phase wurden mithilfe des Instruments Fragen gestellt, um den Lösungsprozess der Teilnehmer*innen besser zu verstehen. In der dritten Phase wurden die Teilnehmer*innen gebeten, die Aufgaben didaktisch zu rekonstruieren; hierbei wurden z. B. Motive für das Stellen der Aufgabe und Lernziele oder typische Fehler von Zweitsemesterstudierenden, an die sich die Aufgaben richteten, abgefragt. Für das Interview wurde ein teilstrukturierter Interviewleitfaden (PARI-like) entwickelt, welcher für die folgenden Interviews noch ergänzt wurde. Die teilnehmenden Expert*innen zogen aus den Interviews Konsequenzen für ihre Lehre. Unterschiede in der Interpretation ergaben sich daraus, dass die Experten in unterschiedlichen Rollen auftraten, als Aufgabensteller*in und Koordinator*in des Übungsbetriebs (dadurch genauere Kenntnis über Defizite der Studierenden) oder als Mitarbeiter*in in der Veranstaltung (dadurch mehr Kritik an der Aufgabenstellung).

FACHDOMÄNE

Erprobt für: Elektrotechnikaufgaben in Grundlagenveranstaltungen
Prinzipiell auch geeignet für: Allgemeine Ingenieuraufgaben in Veranstaltungen beliebiger Semester
Nicht geeignet für: Innermathematische Aufgaben

ZIELGRUPPE

Erprobt für: Ingenieurstudierende in Grundlagenveranstaltungen
Prinzipiell auch geeignet für: Ingenieurstudierende verschiedener Semester
Nicht geeignet für: Reine Mathematik

ALLGEMEINER TESTZWECK

Erprobt für: Erfassen von Kompetenzerwartungen und normativen Anforderungen
Prinzipiell auch geeignet für: Ergänzungen von Musterlösungen für Studierende
Nicht geeignet für: Reine Mathematik ohne Anwendungsbezüge

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Kombination mit Studierende-Expert-Lösungen, NIAs und Kategorisierungen
Prinzipiell auch geeignet für: Vorlesungsunterlagen und Lernhinweise für Studierende
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für Einzelinterviews, parallele Interviews mit mehreren Experten unterschiedlicher Institutionen.
Prinzipiell auch geeignet für: Vergleich mit analogen Interviews mit Studierenden
Nicht geeignet für: Hochschulvergleiche, Länder- und Bundesebene, internationale Ebene

EINSATZSZENARIEN

Erprobt für: Unterstützung bei der Erstellung und Anreicherung von Musterlösungen mit Hinweisen von Experten
Prinzipiell auch geeignet für: Verfassen von Lehrbüchern, die auch auf kognitive Ressourcen Bezug nehmen
Nicht geeignet für: Reine Mathematik

ERLAUBTE KONSEQUENZEN

Erprobt für: Optimierung von Muster-Klausurlösungen für Studierende
Prinzipiell auch geeignet für: Kompetenzorientierte Umgestaltung von Eingangsphasen in den Ingenieurstudiengängen
Nicht geeignet für: —

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Das Instrument wurde im Rahmen der Dissertation von Jörg Kortemeyer eingesetzt, welche unter go.upb.de/kortemeyer abrufbar ist. Es besteht bereits Kontakt zu anderen Institutionen, die die Instrumente einsetzen möchten. Es soll zur Optimierung von Muster-Aufgabenlösungen für Studierende eingesetzt werden.

WEITERE WICHTIGE ANGABEN

Das Instrument ist im Kapitel 3 der Dissertation von Jörg Kortemeyer näher beschrieben.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Kortemeyer, J., Biehler, R. & Schaper, N. (2014). *Hilft der sogenannte Modellierungskreislauf Lösungsprozesse bei ingenieurwissenschaftlichen Anwendungsaufgaben besser zu verstehen?* Dortmund: Universitätsbibliothek Dortmund.

Kortemeyer, J. & Biehler, R. (2017). The interface between mathematics and engineering – problem solving processes for an exercise on oscillating circuits using ordinary differential equations. In *Proceedings of the Tenth Conference of the European Society for Research in Mathematics Education.*

Kortemeyer, J. (2018). *Mathematische Kompetenzen in ingenieurwissenschaftlichen Grundlagenveranstaltungen: Normative und empirische Analysen zu exemplarischen Klausuraufgaben aus dem ersten Studienjahr in der Elektrotechnik* (Dissertation).

WEITERFÜHRENDE LITERATUR

Hall, E. P.; Gott, S. P. & Pokorny, R. A. (1995). *A procedural guide to cognitive task analysis: The PARI methodology* (No. AL/HR-TR-1995-0108). Brooks Air Force Base, Texas: Armstrong Laboratory, Human Resources Directorate.

Biehler, R., Kortemeyer, J. & Schaper, N. (2015). Conceptualizing and studying students' processes of solving typical problems in introductory engineering courses requiring mathematical competences. In K. Krainer & N. Vondrová (Hrsg.), *CERME9: Proceedings of the Ninth Congress of the European Society for Research in Mathematics Education* (S. 2060–2066). Prague: Charles University.

KoMeWP: Fachwissen und Fachdidaktisches Wissen in Rechnungswesen

KURZFASSUNG

Name des Instruments: Test zur Erfassung des fachwissenschaftlichen (FW) und fachdidaktischen Wissens (FDW) im Rechnungswesen

Fachdomäne: Rechnungswesen, Wirtschaftspädagogik

Erfasste Kompetenzen: Fachwissenschaftliches und fachdidaktisches Wissen im Rechnungswesen

Zielgruppe: Studierende der Wirtschaftspädagogik

Testart: Leistungstest; Multiple Choice und offenes Antwortformat

Modalität: Paper-Pencil-Test

Dauer: 75 Minuten; Teil zum fachwissenschaftlichen und fachdidaktischen Wissen 40 Minuten

Testaufbau: Insgesamt 187 Items und 3 Szenarien: Fragebogen zur Selbsteinschätzung des FW und FDW, Test zum FW und FDW im Rechnungswesen, Test zur Erfassung von fachdidaktischen Entscheidungen, Fragebogen zu lehr-lern-prozessbezogenen Ansichten, Test zur Erfassung des studienbezogenen Lernverhaltens, Fragebogen zum schulischen und beruflichen Werdegang

Itempool: 49 Items; FW: Lerninhaltsbereiche: Aufgaben und Rechtsgrundlagen, System der Doppik, Beschaffungs- und Absatzprozesse. FDW: Didaktische Inhaltsbereiche: Wissen über Schülerkognitionen, Wissen über das Zugänglichmachen von Inhalten, Wissen über das Potenzial von Aufgaben. *Testversionen*: Der Itempool ist aufgeteilt nach einem Multi-Matrix-Design: 7 Itemcluster mit je 7 Items. Jedes Testheft enthält 4 Cluster = 28 Items.

Allgemeiner Testzweck: Erfassen des Wissensstands und der Entwicklung im Querschnitt; Evaluation eines Trainingskonzepts

Konkrete Einsatzszenarien: Lehr-Lern-Tool

Nicht geeignet für: —

Hinweise für praktische Anwendung: Der Test liegt im Multi-Matrix-Design vor und muss entsprechend der Lösungscodierung mit IRT-Analyse ausgewertet werden.

Bisheriger Einsatz in Projekten: Modellierung und Erfassung fachwissenschaftlicher und fachdidaktischer Kompetenzen im wirtschaftspädagogischen Studium (KoMeWP)

Kontakt / Speicherort:
Prof. Dr. Jürgen Seifried
Lehrstuhl für Wirtschaftspädagogik – Berufliches Lehren und Lernen
L4,1, 68161 Mannheim
E-Mail: seifried@bwl.uni-mannheim.de
Webseite: https://www.bwl.uni-mannheim.de/seifried/

ALLGEMEINE INFORMATIONEN

Name des Instruments: Test zur Erfassung des fachwissenschaftlichen (FW) und fachdidaktischen Wissens (FDW) im Rechnungswesen
Bisheriger Einsatz in Projekten: Modellierung und Erfassung fachwissenschaftlicher und fachdidaktischer Kompetenzen im wirtschaftspädagogischen Studium (KoMeWP)
Kontakt / Speicherort:
Prof. Dr. Jürgen Seifried
Lehrstuhl für Wirtschaftspädagogik – Berufliches Lehren und Lernen
L4,1, 68161 Mannheim
E-Mail: seifried@bwl.uni-mannheim.de
Webseite: https://www.bwl.uni-mannheim.de/seifried/

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Fachwissenschaftliches und fachdidaktisches Wissen im Rechnungswesen
Theoriemodell: Strukturmodell der professionellen Kompetenz von Lehrenden nach Weinert (2001); Operationalisierung des professionellen Wissens nach Shulman (1986)
Testart: 49 Items; 23 Multiple-Choice Fragen; 13 Fragen mit offenem Antwortformat
Modalität: Paper-Pencil-Test
Testaufbau:
Itempool:
187 Items und 3 Szenarien: Fragebogen zur Selbsteinschätzung des FW und FDW (Bearbeitungszeit: 5 Minuten; 33 Items), Test zum FW und FDW im Rechnungswesen (Bearbeitungszeit: 40 Minuten; 28 Items), Test zur Erfassung von fachdidaktischen Entscheidungen (Bearbeitungszeit: 7 Minuten; 3 Szenarien), Fragebogen zu lehr-lern-prozessbezogenen Ansichten (Bearbeitungszeit: 8 Minuten; 51 Items), Test zur Erfassung des studienbezogenen Lernverhaltens (Bearbeitungszeit: 10 Minuten; 75 Items), Fragebogen zum schulischen und beruflichen Werdegang (Bearbeitungszeit: 5 Minuten). FW: Lerninhaltsbereiche: Aufgaben und Rechtsgrundlagen; System der Doppik; Beschaffungs- und Absatzprozesse; FDW: Wissen über Schülerkognitionen; Wissen über das Zugänglichmachen von Inhalten; Wissen über das Potenzial von Aufgaben

Testversionen: Der Itempool ist aufgeteilt nach einem Multi-Matrix-Design: 7 Itemcluster mit je 7 Items (Youden Square design). Jedes Testeft enthält 4 Cluster = 28 Items.

TESTGÜTEKRITERIEN

Reabilität: EAP/PV-Reliabilitäten für das FW: .79, für das FDW: .73; MNSQ-Werte (*infit*) zwischen 0.93 ≤ MNSQ ≤ 1.08 ; Testgüte (Itemschwierigkeit, Trennschärfen, interne Konsistenzen, MNSQ-Werte etc.) Berger et al. (2013, 2015); Fritsch et al. (2015); Schnick-Vollmer et al. (2015).
Validität Testinhalt: Dokumentenanalyse und Curriculaanalysen (Schulbücher des Rechnungswesens; [Rahmen-]Lehrpläne); Expertenbefragungen (Berger et al., 2013). Der Test deckt zentrale Lerninhalte der Domäne ab.
Validität Antwortprozesse: Die Laut-Denken-Studie verdeutlicht fachdidaktische Argumentationsmuster.
Validität interner Testaufbau: Ein zweidimensionales Modell (FW vs. FDW, Bouley et al., 2015; Fritsch et al., 2015) bildet die Datenstruktur am besten ab. Eine Unterscheidung verschiedener fachdidaktischer Subdimensionen führt zu keiner Verbesserung des Modellfits.
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Theoriekonforme Zusammenhänge zeigen sich in gleicher Weise für eine deutsche und eine österreichische Stichprobe (zur Konstruktvalidität, Bouley et al., 2015; zur Kriteriumsvalidität, Fritsch et al., 2015). Diskriminante und konvergente Validität wurden mit anderen Maßen kognitiver Fähigkeiten geprüft (Schulnoten, Studiennoten); Mittelwertunterschiede zwischen geprüften Gruppen sind erwartungskonform (Vergleich Deutschland – Österreich).
Validität Konsequenzen der Testergebnisse: keine Angaben
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 75 Minuten; Teil zum FW und FDW 40 Minuten
Benötigtes Material: Testbögen
Besonderheiten: Der Test liegt im Multi-Matrix-Design vor und muss entsprechend der Lösungscodierung mit IRT-Analyse ausgewertet werden.
Praxisbeispiel: Erhebungen mit Studierenden der Wirtschaftspädagogik in Deutschland (Bachelor N = 590 und Master N = 552 an 24 Hochschulstandorten) und Österreich (N = 249 Studierende an 4 Hochschulstandorten). Neben anderen Instrumenten wurde ein Teil der in KoMeWP entwickelten Items (24 bzw. 17 Items) von 2014-2016 genutzt, um den Wissenszuwachs angehender Lehrkräfte durch Interventionen überprüfen zu können (Interventionsziel: Unterstützung beim lernförderlichen Umgang mit Fehlern im Rechnungswesen). Zielgruppe waren zunächst 226 Referendar*innen für berufsbildende Lehrkräfte (Wirtschaftspädagogik) und zu einem späteren Zeitpunkt 31 bzw. 37 Studierende der Wirtschaftspädagogik. Die Interventionen zeigen erwartungskonforme Effekte; insbesondere fachdidaktische Kompetenzen konnten verbessert werden.

FACHDOMÄNE

Erprobt für: Rechnungswesen; Wirtschaftspädagogik
Prinzipiell auch geeignet für: Gymnasium mit dem Fach Wirtschaft
Nicht geeignet für: —

ZIELGRUPPE

Erprobt für: Studierende der Wirtschaftspädagogik in Deutschland und Österreich; Referendar*innen für berufsbildende Lehrkräfte
Prinzipiell auch geeignet für: Lehramtsstudierende für Gymnasium mit dem Fach Wirtschaft
Nicht geeignet für: —

ALLGEMEINER TESTZWECK

Erprobt für: Erfassung des Kompetenzstands und der Entwicklung im Querschnitt
Prinzipiell auch geeignet für: Erfassung der Kompetenzentwicklung
Nicht geeignet für: —

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Noten
Prinzipiell auch geeignet für: Andere Maße zur Kompetenzerfassung
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für die Kursebene, Hochschulebene, Deutschland und Österreich
Prinzipiell auch geeignet für: —
Nicht geeignet für: Individualdiagnostik, Landes- und Bundesebene, internationale Ebene

EINSATZSZENARIEN

Erprobt für: Lehr-Lern-Tool, Evaluation eines Trainingskonzepts
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

ERLAUBTE KONSEQUENZEN

Erprobt für: Verbesserung der Lehre
Prinzipiell auch geeignet für: —
Nicht geeignet für: Individualdiagnostik

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Der Test wird von der Arbeitsgruppe Seifried und Wuttke weiterentwickelt und eingesetzt. Der Test kann unter der oben genannten Kontaktadresse angefordert werden.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Berger, S.; Bouley, F.; Fritsch, S.; Krille, C.; Seifried, J. & Wuttke, E. (2015). Fachwissen und fachdidaktisches Wissen im wirtschaftspädagogischen Studium. Entwicklung eines Testinstruments und erste empirische Befunde. In B. Koch-Priewe; A. Köker; J. Seifried & E. Wuttke (Hrsg.), *Kompetenzerwerb an Hochschulen. Modellierung und Messung. Zur Professionalisierung angehender Lehrerinnen und Lehrer sowie frühpädagogischer Fachkräfte* (S. 105–125). Bad Heilbrunn: Klinkhardt.

Fritsch, S.; Berger, S.; Seifried, J.; Bouley, F.; Wuttke, E.; Schnick-Vollmer, K. & Schmitz, B. (2015). The impact of university teacher training on prospective teachers' CK and PCK – A comparison between Austria and Germany. In O. Zlatkin-Troitschanskaia & R. Shavelson (Hrsg.), *Special Issue on Assessment of Domain-specific Professional Competencies, Empirical Research in Vocational Education and Training (ERVET), 7*(4). DOI 10.1186/s40461-015-0014-8. Download unter http://www.ervet-journal.com/content/7/1/4

Bouley, F.; Wuttke, E.; Schnick-Vollmer, K.; Schmitz, B.; Berger, S.; Fritsch, S. & Seifried J. (2015). Professional Competence of Prospective Teachers in Business and Economics Education – Evaluation of a competence model using structural equation modelling. *Peabody Journal of Education, 90*(4), S. 491-502.

WEITERFÜHRENDE LITERATUR

Bouley, F.; Berger, S.; Fritsch, S.; Wuttke, E.; Seifried, J.; Schnick-Vollmer, K. & Schmitz, B. (2015). Der Einfluss von universitären und außeruniversitären Lerngelegenheiten auf das Fachwissen und fachdidaktische Wissen von angehenden Lehrkräften an kaufmännisch-berufsbildenden Schulen. *Zeitschrift für Pädagogik*, Beiheft 61, S. 100–115.

Krille, C.; Salge, B.; Wuttke, E. & Seifried, J. (2017). Evaluation of a training programme to improve the Professional Error Competence of prospective teachers. In E. Wuttke & J. Seifried (Hrsg.), *Professional Error Competence of Preservice Teachers: Evaluation and Support* (S. 75–98). Cham: Springer.

Schnick-Vollmer, K.; Berger, S.; Bouley, F.; Fritsch, S.; Schmitz, B.; Seifried, J. & Wuttke, E. (2015). Modeling the competencies of prospective business and economics teachers. Professional knowledge in accounting. *Zeitschrift für Psychologie, 223*(1), S. 24-30.

KomPaed PAnTHoa: Praxisanforderungen von Hochschulabsolvent*innen

KURZFASSUNG

Name des Instruments: Erfassung praxisbezogener Anforderungen und Tätigkeiten von Hochschulabsolvent*innen (PAnTHoa)

Fachdomäne: Hochschulabsolvent*innen aller Fachrichtungen

Erfasste Kompetenzen: Praxisbezogene Anforderungen und Tätigkeiten von Hochschulabsolvent*innen

Zielgruppe: Hochschulabsolvent*innen in Deutschland

Testart: Fragebogen mit Likert-Skalen

Modalität: Paper-Pencil-Test und Online-Fragebogen

Dauer: 7 Minuten

Testaufbau: 41 Items; 8 Anforderungs-/Tätigkeitsbereiche: 1. Planung und Organisation von Arbeitsprozessen (8 Items), 2. Förderung anderer Personen (3 Items), 3. Führung und Leitung (5 Items), 4. Selbstständiges Handeln bei herausfordernden Aufgaben (9 Items), 5. Informationsaufbereitung (6 Items), 6. Umgang mit Zahlen (3 Items), 7. Fremdsprachige Kommunikation (4 Items), 8. Körperliches Arbeiten (3 Items)

Allgemeiner Testzweck: Erfassen der Häufigkeit wahrgenommener Anforderungen und ausgeführter Tätigkeiten im Berufsalltag

Konkrete Einsatzszenarien: Hochschulevaluation, Akkreditierung, Curriculumsentwicklung

Nicht geeignet für: Studierende, Individualdiagnostik

Hinweise für praktische Anwendung: —

Bisheriger Einsatz in Projekten: Tätigkeitsbezogene Kompetenzen in pädagogischen Handlungsfeldern (KomPaed)

Kontakt / Speicherort:
Prof. Dr. Edith Braun
E-Mail: edith.braun@uni-giessen.de

Dr. Julia-Carolin Brachem
E-Mail: julia-carolin.brachem@uni-vechta.de

Deutsches Zentrum für Hochschul- und Wissenschaftsforschung GmbH (Speicherort)
Forschungsdatenzentrum für Hochschul- und Wissenschaftsforschung
Postanschrift: Lange Laube 12, 30159 Hannover
E-Mail: dataservice@dzhw.eu
DOI: https://dx.doi.org/10.21249/DZHW:cmp2014:1.0.0

ALLGEMEINE INFORMATIONEN

Name des Instruments: Erfassung praxisbezogener Anforderungen und Tätigkeiten von Hochschulabsolvent*innen (PAnTHoa)
Bisheriger Einsatz in Projekten: KomPaed
Kontakt / Speicherort:
Prof. Dr. Edith Braun
E-Mail: edith.braun@uni-giessen.de
Dr. Julia-Carolin Brachem
E-Mail: julia-carolin.brachem@uni-vechta.de
Deutsches Zentrum für Hochschul- und Wissenschaftsforschung GmbH (Speicherort)
Forschungsdatenzentrum für Hochschul- und Wissenschaftsforschung
Postanschrift: Lange Laube 12, 30159 Hannover
E-Mail: dataservice@dzhw.eu
DOI: https://dx.doi.org/10.21249/DZHW:cmp2014:1.0.0

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Praxisbezogene Anforderungen und Tätigkeiten von Hochschulabsolvent*innen
Theoriemodell: Job Requirements Approach, angepasst an das deutsche Hochschulsystem
Testart: Fragebogen mit Likert-Skalen
Modalität: Paper-Pencil-Test und Online-Fragebogen
Testaufbau:
Itempool: 41 Items; 8 Anforderungs-/Tätigkeitsbereiche: 1. Planung und Organisation von Arbeitsprozessen (8 Items), 2. Förderung anderer Personen (3 Items), 3. Führung und Leitung (5 Items), 4. Selbstständiges Handeln bei herausfordernden Aufgaben (9 Items), 5. Informationsaufbereitung (6 Items), 6. Umgang mit Zahlen (3 Items), 7. Fremdsprachige Kommunikation (4 Items), 8. Körperliches Arbeiten (3 Items)

TESTGÜTEKRITERIEN

Reliabilität: Interne Konsistenz bei Hochschulabsolvent*innen der Abschlussjahrgänge 2001, 2005 und 2009: Planung und Organisation von Arbeitsprozessen (α = .75), Förderung anderer Personen (α = .81), Führung und Leitung (α = .72), Selbstständiges Handeln bei herausfordernden Aufgaben (α = .79), Informationsaufbereitung (α = .79), Umgang mit Zahlen (α = .62), Fremdsprachige Kommunikation (α = .74), Körperliches Arbeiten (α = .77).
Validität Testinhalt: Die Iteminhalte wurden auf Basis einer Literaturrecherche und Arbeitgeberinterviews entwickelt.
Validität Antwortprozesse: Zur Prüfung wurden kognitive Interviews durchgeführt.
Validität interner Testaufbau: Eine konfirmatorische faktorenanalytische Überprüfung des Gesamtmodells zeigt erwartungskonforme Ergebnisse (RMSEA: 0,07, SRMR: 0,08, CFI: 0,74).
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Keine Angaben
Validität Konsequenzen der Testergebnisse: keine Angaben
Testfairness: keine Angaben

FACHDOMÄNE

Erprobt für: Für Hochschulabsolvent*innen aller Studienfächergruppen (Sprach- und Kulturwissenschaften: N = 2.102; Sport: N = 53; Rechts-, Wirtschafts- und Sozialwissenschaften: N = 2.826; Mathematik, Naturwissenschaften: N = 1.926; Humanmedizin/Gesundheitswissenschaften: N = 570; Veterinärmedizin: N = 141; Agrar-, Forst- und Ernährungswissenschaften: N = 464; Ingenieurwissenschaften: N = 1.891; Kunst, Kunstwissenschaft: N = 281) und für Berufsbereiche (Militär: N = 6; Land-, Forst- und Tierwirtschaft und Gartenbau: N = 191; Rohstoffgewinnung, Produktion und Fertigung: N = 1.070; Bau, Architektur, Vermessung und Gebäudetechnik: N = 479; Naturwissenschaft, Geografie und Informatik: N = 832; Verkehr, Logistik, Schutz und Sicherheit: N = 122; Kaufmännische Dienstleistungen, Warenhandel, Vertrieb, Hotel und Tourismus: N = 284; Unternehmensorganisation, Buchhaltung, Recht und Verwaltung: N = 2.205)
Prinzipiell auch geeignet für: —
Nicht geeignet für: Studierende

ZIELGRUPPE

Erprobt für: Hochschulabsolvent*innen in Deutschland
Prinzipiell auch geeignet für: —
Nicht geeignet für: Studierende

ALLGEMEINER TESTZWECK

Erprobt für: Erfassen der Häufigkeit wahrgenommener Anforderungen und ausgeführter Tätigkeiten im Berufsalltag
Prinzipiell auch geeignet für: Ableitung relevanter überfachlicher Schlüsselkompetenzen von Hochschulabsolvent*innen für den Arbeitsmarkt
Nicht geeignet für: —

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: —
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 7 Minuten
Benötigtes Material: —
Besonderheiten: —
Praxisbeispiel: Der Fragebogen wurde im Rahmen der deutschlandweiten Absolventenbefragung des Deutschen Zentrums für Hochschul- und Wissenschaftsforschung (DZHW) eingesetzt (Querschnitts-Online-Befragung der Absolventenjahrgänge 2001, 2005 und 2009 im Jahr 2013/2014). Der Fragebogen erlaubt Rückschlüsse auf Anforderungen und Tätigkeiten, die Hochschulabsolvent*innen auf dem Arbeitsmarkt erwarten. Der Fragebogen wurde zudem in einer Online-Vertiefungsbefragung des DZHW-Absolventenpanels 2005 und in einer weiterentwickelten Form im Rahmen der NEPS-Studie „Hochschulstudium und Übergang in den Beruf" eingesetzt (Deutsches Zentrum für Hochschul- und Wissenschaftsforschung 2017a, 2017b; Leibniz-Institut für Bildungsverläufe e. V. 2018).

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Gruppenebene
Prinzipiell auch geeignet für: —
Nicht geeignet für: Individualdiagnostik, internationale Ebene

EINSATZSZENARIEN

Erprobt für: Hochschulevaluation, Akkreditierung, Curriculumsentwicklung
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

ERLAUBTE KONSEQUENZEN

Erprobt für: Curriculumsentwicklung
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Der PAnTHoa-Fragebogen steht Interessierten online über die Evaluations- und Umfragesoftware „EvaSys" zur Verfügung (https://www.evasys.de/aktuelles/detail/panthoa-fragebogen-329.html). Die den Auswertungen zugrundeliegenden Daten werden durch das Forschungsdatenzentrum für Hochschul- und Wissenschaftsforschung als Scientific Use File zur wissenschaftlichen Nutzung zur Verfügung gestellt (DOI: https://dx.doi.org/10.21249/DZHW:cmp2014:1.0.0).

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Braun, E. & Brachem, J. C. (2018). Erfassung praxisbezogener Anforderungen und Tätigkeiten von Hochschulabsolventinnen/-absolventen (PAnTHoa). *Zeitschrift für Hochschulentwicklung 13*(1), S. 209–232.

Braun, E. & Brachem, J.-C. (2017). The labour market's requirement profiles for higher education graduates. In E. Kyndt; V. Donche; K. Trigwell & S. Lindblom-Ylänne (Hrsg.), *Higher Education Transitions. Theory and Research* (S. 219–237). London: Routledge.

Braun, E. & Brachem, J.-C. (2015). Requirements higher education graduates meet on the labor market. *Peabody Journal of Education. 90*(4), S. 574–595.

WEITERFÜHRENDE LITERATUR

Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (2017a). DZHW-Absolventenpanel. https://www.dzhw.eu/forschung/projekt?pr_id=467 (Aufruf am 20.11.2018).

Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (2017b). Projekt „Tätigkeitsanforderungen Hochqualifizierter" (JobReq). https://www.dzhw.eu/forschung/projekt?pr_id=596 (Aufruf am 20.11.2018).

Leibniz-Institut für Bildungsverläufe e. V. (2018). Hochschulstudium und Übergang in den Beruf. https://www.neps-studie.de/Studien/Hochschulstudiumund%C3%9CbergangindenBeruf.aspx (Aufruf am 20.11.2018).

KomPrü-KomFäh: Performanzbasierte Messung kommunikativer Fähigkeiten von Studierenden

KURZFASSUNG

Name des Instruments: Performanzbasierte Messung kommunikativer Fähigkeiten von Studierenden

Fachdomäne: Hochschulabsolvent*innen aller Fachrichtungen

Erfasste Kompetenzen: Strategische und verständigungsorientierte Kommunikationsfähigkeiten

Zielgruppe: Studierende aller Fachrichtungen in Deutschland

Testart: Standardisierte performanzbasierte Rollenspiele mit geschlossenen Beobachtungsbogen

Modalität: Rollenspiele, Papier-und-Bleistift-Beobachtungsbogen

Dauer: 45 Minuten

Testaufbau: 20 Rollenspiele; je 5 Rollenspiele für strategische und verständigungsorientierte Kommunikationsfähigkeiten; jedes Rollenspiel in einer Version für Unternehmens- oder Bildungseinrichtungskontext

Allgemeiner Testzweck: Erfassung von kommunikativen Fähigkeiten über theoretisch begründete Beobachtungskriterien

Konkrete Einsatzszenarien: Kompetenzorientierte Prüfung, Modul- oder Studienabschluss, Trainings

Nicht geeignet für: —

Hinweise für praktische Anwendung: Hoher Schulungsaufwand für Gesprächspartner*innen und für Beobachter*innen, die mit dem theoretischen Rahmen vertraut sein müssen

Bisheriger Einsatz: In Lehrveranstaltungen für Studierende und Lehrende (Hochschuldidaktik) als Training; im Projekt mit 488 Studierenden als performanzbasierter Test; auch als hochschuldidaktisches Training mit Lehrenden und in Lehrveranstaltungen mit Studierenden

Kontakt / Speicherort:
Prof. Dr. Edith Braun
Institut für Erziehungswissenschaft
Justus-Liebig-Universität Gießen
Postanschrift: Bismarckstraße 37, 35390 Gießen
E-Mail: edith.braun@uni-giessen.de

ALLGEMEINE INFORMATIONEN

Name des Instruments: Performanzbasierte Messung kommunikativer Fähigkeiten von Studierenden
Bisheriger Einsatz in Projekten: KomPrü
Kontakt / Speicherort:
Prof. Dr. Edith Braun
Institut für Erziehungswissenschaft
Justus-Liebig-Universität Gießen
Postanschrift: Bismarckstraße 37, 35390 Gießen
E-Mail: edith.braun@uni-giessen.de

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Strategische und verständigungsorientierte Kommunikationsfähigkeiten
Theoriemodell: Strategische und verständigungsorientierte Kommunikation nach Habermas
Testart: Performanzbasierte Rollenspiele mit standardisierter Instruktion und Beobachtungsbogen
Modalität: Rollenspiele, Papier-und-Bleistift-Beobachtungsbogen
Testaufbau:
Itempool: 20 Rollenspiele; je 5 Rollenspiele für strategische und 5 für verständigungsorientierte Kommunikationsfähigkeiten, jeweils für den Kontext Unternehmen oder Bildungseinrichtung. Jedes Rollenspiel besteht aus einer Instruktion für die Testperson und einem standardisierten Beobachtungsbogen. Je Testdurchführung wird ein*e geschulte*r Gesprächspartner*in und ein*e geschulte*r Beobachter*in benötigt, die mit dem theoretischen Konstrukt und den Rollenspielen vertraut sind (entspricht dem Testsetting von mündlichen Prüfungen mit Prüfer*in und Protokollant*in). Die Testpersonen erhalten pro Rollenspiel 5 Minuten Vorbereitungszeit und bis zu 10 Minuten für das Ausspielen der Situation.
Versionen: Unterschiedliche Kombinationen von je 4 Rollenspielen

TESTGÜTEKRITERIEN

Reliabilität: Rho für interkorrelierte Faktoren (Chao, 2016) = .80 für strategische Kommunikation und .84 für die verständigungsorientierte Kommunikation
Validität Testinhalt: Die Situationen der Rollenspiele wurden mittels einer Absolventenbefragung (N = 10.000) ermittelt. Alle Rollenspiele wurden in einer Expertenrunde diskutiert. Den Rollenspielen wurde hohe Authentizität zugesprochen.
Validität interner Testaufbau: Eine konfirmatorische Faktorenanalyse bestärkt hypothesenkonform ein zweidimensionales Modell ($\chi 2$ = 103,5; df = 53; $p < 0{,}001$; CFI = 0,95; RMSEA = 0,04).

Validität Zusammenhänge mit anderen Kompetenzindikatoren: Strategische Kommunikation korreliert mit para- und nonverbaler Kommunikation (r = 0,60). Für verständigungsorientierte Kommunikation ergibt sich ein Zusammenhang von r = 0,57.
Validität Konsequenzen der Testergebnisse: Die Studierenden berichten eine hohe Authentizität der Rollenspiele für berufliche Situationen.
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 45 Minuten für vier Rollenspiele
Benötigtes Material: Instruktion und Beobachtungsbogen
Besonderheiten: Geschultes Personal (Gesprächspartner*in, Beobachter*in); hoher Aufwand in der Testdurchführung für hohe Augenscheinvalidität
Praxisbeispiel: Der Fragebogen wurde bei 488 Studierenden an 10 Hochschulen deutschlandweit eingesetzt. Die Studierenden waren im Vorfeld nicht vertraut mit den Rollenspielen und haben die Rollenspiele gut durchführen können. Die Rollenspiele werden in der Lehramtsbildung an der Freien Universität Berlin, an der Pädagogischen Hochschule Luzern und an der Universität Gießen verwendet. Im hochschuldidaktischen Bereich werden die Rollenspiele als Training eingesetzt.

FACHDOMÄNE

Erprobt für: Hochschulabsolvent*innen aller Studienfächergruppen, Schwerpunkt Lehramt und Wirtschaftswissenschaften
Prinzipiell auch geeignet für: Dozierende in hochschuldidaktischen Kursen
Nicht geeignet für: —

ZIELGRUPPE

Erprobt für: Studierende aller Fächergruppen in Deutschland
Prinzipiell auch geeignet für: Lehrende und Hochschulabsolvent*innen, Studierende, Absolventinnen und Lehrende in anderen Ländern (wegen Kulturabhängigkeit erneute Prüfung im Zielland notwendig)
Nicht geeignet für: —

ALLGEMEINER TESTZWECK

Erprobt für: Erfassen durch Beobachtung von strategischer und verständigungsorientierter Kommunikation
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: In Teilstichproben wurden für Validierungszwecke weitere Skalen eingesetzt:
Skalen der nonverbalen Kommunikation: Spitzberg, B. H. & Adams, T. W. (2007). *CSRS, the Conversational Skills Rating Scale: An Instructional Assessment of Interpersonal Competence.* NCA, National Communication Association.
Kurzversion des WiWiKom Tests: siehe WiWiKom-Steckbrief (Seite 109)
Skalen zu pädagogischem Unterrichtswissen: König, J. & Blömeke, S. (2010). Pädagogisches Unterrichtswissen: (PUW). *Dokumentation der Kurzfassung des TEDS-M Testinstruments zur Kompetenzmessung in der ersten Phase der Lehrerausbildung.* Humboldt-Universität zu Berlin.
Prinzipiell auch geeignet für: Weitere Maße der Kompetenzmessung
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Einzeldiagnostik
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

EINSATZSZENARIEN

Erprobt für: Trainings ohne Bewertung
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

ERLAUBTE KONSEQUENZEN

Erprobt für: —
Prinzipiell auch geeignet für: Feedback zum Kompetenzstand
Nicht geeignet für: —

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Aktuell besteht ein Austausch zum Legacy-Projekt in Großbritannien..

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Braun, E.; Athanassiou, G.; Pollerhof, K. & Schwabe, U. (2018). Wie lassen sich kommunikative Kompetenzen messen? – Konzeption einer kompetenzorientierten Prüfung kommunikativer Fähigkeiten. *Beiträge zur Hochschulforschung 40*(3), 34–55.
Braun E., Schwabe U., Klein D. (2018). Performance-Based Tests: Using Role Plays to Assess Communication Skills. In: McGrath S., Mulder M., Papier J., Suart R. (eds) *Handbook of Vocational Education and Training.* Cham: Springer.

Ko-NaMa Assessment: Kompetenzen für das Nachhaltigkeitsmanagement

KURZFASSUNG

Name des Instruments: Ko-NaMa Assessment

Fachdomäne: Betriebliches Nachhaltigkeitsmanagement, Betriebswirtschaftslehre

Erfasste Kompetenzen: Kompetenzen im Nachhaltigkeitsmanagement

Zielgruppe: Bachelorstudierende (ab 4. Semester) und Masterstudierende der Betriebswirtschaftslehre/Wirtschaftswissenschaften; Studierende mit betriebswirtschaftlichem Grundlagenwissen

Testart: Leistungstest (Wissenstest, Unternehmenssimulation) und Fragebogen zu soziobiografischen Merkmalen sowie zu Interessen, Lerngelegenheiten und Einstellungen zur Nachhaltigkeit.

Modalität: Computergestützt

Dauer: 90 Minuten in der Minimalversion, 120 bis 150 Minuten in erweiterter Version (modularer Aufbau ermöglicht Erhöhung der Messpräzision über Testzeitverlängerung der einzelnen Testkomponenten). Digitale Testhefte können zweckspezifisch zusammengestellt werden.

Testaufbau: Das Assessment besteht aus 3 Testmodulen: Modul I besteht aus 3 Wissenstests (184 Items). In Modul II wird zur Messung der Anwendung unterschiedlicher Wissensarten (deklarativ, prozedural, schematisch und strategisch) eine Unternehmenssimulation präsentiert (13 Betriebssituationen mit 72 Items für das Modellunternehmen). Modul III besteht aus einem Fragebogen zur Erfassung soziobiografischer Merkmale sowie von Interessen, Lerngelegenheiten und Einstellungen.

Allgemeiner Testzweck: Messung des Kompetenzniveaus und der Kompetenzstruktur im Bereich des Nachhaltigkeitsmanagements; Messung von Kompetenzentwicklung

Konkrete Einsatzszenarien: Kompetenzmessung, Evaluation von Lernergebnissen

Nicht geeignet für: Unternehmens-/Wirtschaftskontexte; Studierende in den ersten drei Fachsemestern

Hinweise für praktische Anwendung: Als Testumgebung wird das Lernmanagementsystem ILIAS als eine Open Source Software verwendet. Alle Testkomponenten einschließlich Fragebogen sind dort implementiert. Die Auswertung der deklarativen Testkomponenten erfolgt weitestgehend automatisiert. Die Simulation enthält neben geschlossenen auch offen gestaltete Items (z. B. zu Entscheidungsbegründungen), die manuell gescort werden müssen. Ein Handbuch für das Rating ist verfügbar.

Bisheriger Einsatz in Projekten: Ko-NaMa – Simulationsbasierte Messung und Validierung eines Kompetenzmodells für das Nachhaltigkeitsmanagement

Kontakt / Speicherort:
Prof. Dr. Susan Seeber
Professur für Wirtschaftspädagogik und Personalentwicklung
Postanschrift: Platz der Göttinger Sieben 5, 37073 Göttingen
Webseite: https://www.uni-goettingen.de/de/531774.html

Speicherort: Institut für Qualitätsentwicklung im Bildungswesen (IQB)

ALLGEMEINE INFORMATIONEN

Name des Instruments: Ko-NaMa Assessment
Bisheriger Einsatz in Projekten: Ko-NaMa – Simulationsbasierte Messung und Validierung eines Kompetenzmodells für das Nachhaltigkeitsmanagement
Kontakt / Speicherort:
Prof. Dr. Susan Seeber
Professur für Wirtschaftspädagogik und Personalentwicklung
Postanschrift: Platz der Göttinger Sieben 5, 37073 Göttingen
Webseite: https://www.uni-goettingen.de/de/531774.html
Speicherort: Institut für Qualitätsentwicklung im Bildungswesen (IQB)

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Kompetenzen im Nachhaltigkeitsmanagement
Theoriemodell: Das theoretische Kompetenzmodell weist eine vierdimensionale Struktur auf: (1) deklaratives Wissen über Nachhaltigkeit aus einer allgemein gesellschaftlichen Perspektive, (2) deklaratives Wissen in der „klassischen" Betriebswirtschaftslehre, (3) deklaratives Wissen über Nachhaltigkeitsmanagement und (4) die Fähigkeit, in unternehmerischen Handlungssituationen Strategien und Begründungen für spezifische Optionen im Sinne der Nachhaltigkeit zu generieren, die sich aus ökonomischen, ökologischen und sozialen Überlegungen ableiten.
Testart: Leistungstest (Wissenstest, Unternehmenssimulation) und Fragebogen zu soziobiografischen Merkmalen sowie zu Interessen, Lerngelegenheiten und Einstellungen zur Nachhaltigkeit.
Modalität: Computergestützt

Testaufbau:
Das Assessment besteht aus 3 Testmodulen (Gesamttestzeit 90 Minuten): Das erste Testmodul umfasst 3 Wissenstests mit einem Itempool von 80 Items zum deklarativen Wissen über Betriebswirtschaftslehre, 53 Items zum deklarativen Wissen über Nachhaltigkeit aus gesamtgesellschaftlicher Perspektive und 51 Items zum deklarativen Wissen über Nachhaltigkeitsmanagement und Nachhaltigkeitscontrolling. Die Items werden zufällig präsentiert und die Bearbeitung auf jeweils 10 Minuten pro Wissenstest beschränkt. Bei einer 10-minütigen Testzeit absolvierten die Studienteilnehmer*innen ca. 22 bis 24 Items pro Wissenstest.
Das zweite Testmodul besteht aus einer Unternehmenssimulation zur Messung der Anwendung betriebswirtschaftlichen und nachhaltigkeitsbezogenen Wissens in verschiedenen betrieblichen Handlungssituationen. Die Unternehmenssimulation umfasst 13 betriebliche Situationen, die zentrale Funktionsbereiche beinhalten (u.a. Unternehmensstrategie, Beschaffung, Produktion, Absatz, Logistik, Personal). Jede Situation umfasst mehrere Items, insgesamt beinhaltet das zweite Testmodul 72 Items. Die Situationen werden über Videovignetten, Audioclips und betrieblichen Schriftwechsel (z. B. E-Mails) simuliert. In einem Balanced-Booklet-Design wurden die Situationen so angeordnet, dass die Teilnehmer*innen jeweils nur drei bis vier der Situationen bearbeiten mussten. Für dieses Testmodul wurde eine Bearbeitungszeit von 45 Minuten angesetzt.
Das dritte Testmodul besteht aus einem Fragebogen zur Erfassung soziobiografischer Merkmale sowie von Interessen, Lerngelegenheiten und Einstellungen. Die Testzeit beträgt ca. 15 Minuten.

TESTGÜTEKRITERIEN

Reliabilität: Die psychometrische Qualität der einzelnen Tests hängt von der Testzeit ab. Zufriedenstellende Messqualität wird ab 20 Minuten Testzeit für jede der deklarativen Wissenskomponenten im Testmodul 1 erreicht. Bei der Unternehmenssimulation liegen nur Erfahrungen mit einer 45-minütigen Testzeit vor. Aufgrund der Heterogenität der Items und der begrenzten Itemzahl, die Probanden in 45 min bearbeiten können, wird lediglich eine zufriedenstellende Reliablität erreicht, die Aussagen für Gruppen zulässt. Eine 45-minütige Testzeit ist nicht für eine Individualdiagnostik im Bereich der Handlungskompetenzen geeignet, für diesen Zweck wird eine Testzeit von 60 bis 80 Minuten empfohlen.
In den Testkomponenten von Modul 1 (Messung des deklarativen Wissens) wurden bei jeweils 10-minütiger Testzeit folgende EAP/PV-Reliabilitäten erreicht: (1) Deklaratives Wissen über Nachhaltigkeit aus gesamtgesellschaftlicher Perspektive: .59, (2) Deklaratives betriebswirtschaftliches Wissen: .51 und (3) Deklaratives Wissen über das betriebliche Nachhaltigkeitsmanagement und Nachhaltigkeitscontrolling: .57.
Bei 45-minütiger Testzeit im Testmodul 2 (Unternehmenssimulation) wird eine EAP/PV-Reliabilität von .55 erreicht. Dabei wurden durchschnittlich 3-4 von insgesamt 13 Situationen bearbeitet.

Validität Testinhalt: Die Hochschullehrerbefragung zur curricularen Validität der Testaufgaben macht deutlich, dass die Studierenden gut bis sehr gut mit den betriebswirtschaftlichen Inhalten des Tests vertraut sind. Im betrieblichen Nachhaltigkeitsmanagement wurden Differenzen in der curricularen Validität festgestellt. Ein wesentlicher Grund liegt in der curricularen Implementierung, da Angebote zur Förderung von Kompetenzen des Nachhaltigkeitsmanagements bisher an den meisten Standorten nicht in Grundlagenmodule integriert, sondern im Wahlbereich verankert sind. Zudem ist das Nachhaltigkeitsmanagement ein junges Teilgebiet der Betriebswirtschaftslehre. Standardisierungen (Entwicklung von Lehrbüchern, inhaltliche Abstimmungen) befinden sich erst in der Entwicklung.
Validität Antwortprozesse: Nicht bestimmt
Validität interner Testaufbau: Nicht bestimmt
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Nicht bestimmt
Validität Konsequenzen der Testergebnisse: Nicht bestimmt
Testfairness: Nicht bestimmt

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 90-150 Minuten
Benötigtes Material: Computer oder Tablets, Taschenrechner, Kopfhörer, Funknetzwerk mit mobilen Servern und zwei Hochleistungs-Routern (für Probandenzahlen > 30)
Besonderheiten: Die Auswertung der erhobenen Daten ist weitestgehend automatisiert bei den deklarativen Komponenten. Die Simulation enthält geschlossene und offene Antwortformate. Die offenen Antwortformate beziehen sich vor allem auf Entscheidungsbegründungen, die manuell zu scoren sind. Ein Handbuch für das Rating ist verfügbar. Die Auswertungsdauer richtet sich entsprechend nach der Anzahl der Testteilnehmer/-innen.
Praxisbeispiel: Das Verfahren wurde bei 872 Bachelor-Studierenden verschiedener wirtschaftswissenschaftlicher Studiengänge eingesetzt.

FACHDOMÄNE

Erprobt für: Betriebliches Nachhaltigkeitsmanagement
Prinzipiell auch geeignet für: Betriebswirtschaftslehre
Nicht geeignet für: Unternehmens- / wirtschaftsfremde Kontexte

ZIELGRUPPE

Erprobt für: Bachelorstudierende (ab 4. Semester) und Masterstudierende der Betriebswirtschaftslehre/Wirtschaftswissenschaften/Wirtschaftspädagogik
Prinzipiell auch geeignet für: Studierende mit betriebswirtschaftlichem Grundlagenwissen (z. B. BWL als Nebenfach)
Nicht geeignet für: Studierende im 1.-3. Fachsemester und Proband*innen ohne wirtschaftswissenschaftliche Studienfachrichtung

ALLGEMEINER TESTZWECK

Erprobt für: Messung des Kompetenzniveaus und der Kompetenzstruktur sowie zur Ermittlung von Erklärungsfaktoren für die Kompetenzausprägung
Prinzipiell auch geeignet für: Messung der Kompetenzentwicklung
Nicht geeignet für: —

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Motivation, Einstellungen, Überzeugungen zur Nachhaltigkeit, sowie zu soziobiografischen Daten. Fragebögen zu Motivation, Einstellungen, Überzeugungen zur Nachhaltigkeit sowie zu sozio-biografischen Daten sind im Instrument enthalten.
Prinzipiell auch geeignet für: Keine Einschränkung
Nicht geeignet für: Keine Einschränkung

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für die Kurs- und Hochschulebene.
Prinzipiell auch geeignet für: Individualdiagnostik bei Verlängerung der Mindesttestzeit von 90 auf 120 bis 150 Minuten mit folgender Teststruktur: 20 Minuten pro deklarativer Wissenstestkomponente (gesamt 60 Minuten) und 60 bis 90 Minuten simulative Komponente (d.h. 4 bzw. 5 statt nur 3 Szenarien); Internationale Ebene (erfordert curriculare Prüfungen); längsschnittliche Untersuchungen
Nicht geeignet für: —

EINSATZSZENARIEN

Erprobt für: Kompetenzmessung, Evaluation von Lernergebnissen
Prinzipiell auch geeignet für: Adaptation als Prüfungsinstrument möglich
Nicht geeignet für: Zulassungsverfahren an Hochschulen

ERLAUBTE KONSEQUENZEN

Erprobt für: —
Prinzipiell auch geeignet für: Kompetenztest im Nachhaltigkeitsmanagement, wenn eine gute Übereinstimmung von Lerninhalten an der jeweiligen Hochschule mit dem Test gegeben ist.
Nicht geeignet für: —

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Der Test wird inhaltlich nicht weiterentwickelt, ebenso wenig aktualisiert oder technologisch angepasst. Ein regelmäßiger Einsatz wird im Rahmen eines Moduls zum Nachhaltigkeitsmanagement an der Wirtschaftswissenschaftlichen Fakultät sowie im Rahmen der ZESS-Kurse der Georg-August-Universität Göttingen erfolgen.
Das Instrument ist verfügbar unter:
Speicherort: Institut für Qualitätsentwicklung im Bildungswesen (IQB)

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

WEITERFÜHRENDE LITERATUR

Anke, J. M. (2019). *IT-gestützte Lern- und Assessmentmodule für nachhaltiges Wirtschaften. Empirische Erkenntnisse und Gestaltungsansätze zum Einsatz IT-gestützter Lern- und Assessmentmodule.* Göttingen: Cuvillier Verlag. Dissertation.

Michaelis, C., Aichele, C., Hartig, J., Seeber, S., Dierkes, S., Schumann, M., Anke J. M., Siepelmeyer, D., & Repp, A. (2020). Impact of Affective-Motivational Dispositions on Competence in Sustainability Management. In O. Zlatkin-Troitschanskaia, H. A. Pant, M. Toepper, & C. Lautenbach (Eds.), *Student Learning in German Higher Education. Innovative Measurement Approaches and Research Results.* Springer: Wiesbaden.

Seeber, S., Michaelis, C., Repp, A., Hartig J., Aichele, C., Schuman, M., Anke, J. M., Dierkes, S. & Siepelmeyer, D. (2019). Assessment of Competences in Sustainability Management: Analyses to the Construct Dimensionality. In *Zeitschrift für Pädagogische Psychologie, 33* (2), 1-11.

KUI: Kompetenzfacetten für das Unterrichten in Informatik

KURZFASSUNG

Name des Instruments: Messung von Kompetenzfacetten für das Unterrichten in Informatik

Fachdomäne: Informatikdidaktik

Erfasste Kompetenzen: Fachdidaktische Kompetenzen, Überzeugungen, motivationale Orientierungen

Zielgruppe: Studierende des Informatiklehramts und Informatiklehrer*innen

Testart: Leistungstest (Multiple-Choice) und Fragebogen

Modalität: Paper-Pencil-Test, Online-Test

Dauer: 45 Minuten

Testaufbau: Teil 1: Didaktische Inhalte und Unterrichtsphasen (36 Items); Teil 2: Professionelle Überzeugungen (34 Items) und motivationalen Orientierungen (13 Items)

Allgemeiner Testzweck: Erfassung des Kompetenzstands der Fachdidaktik Informatik, der Ausprägungen der Überzeugungen und motivationalen Orientierungen

Konkrete Einsatzszenarien: Im Lehramtsstudiengang Informatik, an Schulen

Nicht geeignet für: Testung von Fachwissen

Hinweise für praktische Anwendung: Das Instrument ist auf Deutsch und Englisch verfügbar. Das Instrument bedarf weiterer Entwicklung. Einsatz- und Analysemöglichkeiten müssen mit dem Projektteam abgestimmt werden.

Bisheriger Einsatz in Projekten: KUI (Kompetenzen für das Unterrichten in Informatik)

Kontakt / Speicherort:
E-Mail Niclas Schaper: niclas.schaper@upb.de
E-Mail Peter Hubwieser: peter.hubwieser@tum.de
E-Mail Marc Berges: marc.berges@tum.de
E-Mail Elena Bender: elena.bender@upb.de

ALLGEMEINE INFORMATIONEN

Name des Instruments: Messung von Kompetenzfacetten für das Unterrichten in Informatik
Bisheriger Einsatz in Projekten: KUI (Kompetenzen für das Unterrichten in Informatik)
Kontakt / Speicherort:
E-Mail Niclas Schaper: niclas.schaper@upb.de
E-Mail Peter Hubwieser: peter.hubwieser@tum.de
E-Mail Marc Berges: marc.berges@tum.de
E-Mail Elena Bender: elena.bender@upb.de

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Fachdidaktische Kompetenz, Überzeugungen, motivationale Orientierungen
Theoriemodell: Kompetenzstrukturmodell für das Unterrichten in Informatik (Bender et al., 2015) auf Grundlage des Modells professioneller Handlungskompetenz (Baumert & Kunter, 2011)
Testart: Leistungstest (Multiple-Choice) und Fragebogen
Modalität: Paper-Pencil-Test, Online-Test
Testaufbau:
2 Test-Teile: Teil 1 zur Erfassung der fachdidaktischen Kompetenzen: 6 Skalen mit 36 Items zur Erfassung unterschiedlicher Facetten didaktischer Inhalte (Lerninhalte; Methoden und Sozialformen; Schülerkognition; Zielsetzung von Unterrichtseinheiten) und Unterrichtsphasen (Planung und Gestaltung von Lernsituationen; Reaktion in Unterrichtssituationen; Evaluierung von Unterrichtsverläufen). Teil 2 zur Erfassung der professionellen Überzeugungen: 2 Skalen zur Erfassung der professionellen Überzeugungen (34 Items) und zur Erfassung der motivationalen Orientierungen (13 Items)

TESTGÜTEKRITERIEN

Reliabilität:
Teil 1 noch unzureichende Reliabilität; Teil 2 zufriedenstellende bis ausreichende Reliabilität der einzelnen Skalen.
Validität Testinhalt: Die Testitems bilden das theoretische Kompetenzstrukturmodell ab.
Validität Kognitive Testbearbeitung: Ergebnisse aus Pretests zur Prüfung von Verständlichkeit und Relevanz führten zu Überarbeitungen und einer Verbesserung der Ergebnisse. Die neu überarbeiteten Items sind noch kognitiv zu validieren.
Validität interner Testaufbau: keine Angaben
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Valide Aussagen zu den Zusammenhängen mit anderen Kompetenzindikatoren sind noch nicht möglich
Validität Konsequenzen der Testergebnisse: Nicht bestimmt
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 45 Minuten
Benötigtes Material: Testitems, Fragebogen
Besonderheiten: Das Instrument ist auf Deutsch und Englisch verfügbar. Das Instrument bedarf weiterer Entwicklung. Einsatz- und Analysemöglichkeiten müssen mit dem Projektteam abgestimmt werden.
Praxisbeispiel: Eingesetzt und erprobt bei einer Stichprobe von 155 (angehenden) Informatiklehrkräften (davon N = 123 Lehrpersonen, N = 20 Lehramtsstudierende und N = 12 Referendar*innen) deutschlandweit. Zudem wurde das Kompetenzstrukturmodell in Lehrveranstaltungen zur Didaktik der Informatik für Gymnasium und berufliche Bildung in Bayern und Österreich eingesetzt.

FACHDOMÄNE

Erprobt für: Informatikdidaktik
Prinzipiell auch geeignet für: —
Nicht geeignet für: Erfassung von Fachwissen

ZIELGRUPPE

Erprobt für: Informatiklehrkräfte und Informatiklehramtsstudierende
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

ALLGEMEINER TESTZWECK

Erprobt für: Erfassung des Kompetenzstands der Fachdidaktik Informatik, Ausprägungen der Überzeugungen und motivationalen Orientierungen
Prinzipiell auch geeignet für: —
Nicht geeignet für: Erfassung von Kompetenzentwicklung

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: —
Prinzipiell auch geeignet für: Andere Maße der Kompetenzerfassung
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für Individualdiagnostik, Landes- und Bundesebene. Internationale Ebene: Der Test wurde in Michigan (USA) in einer Pilotstudie eingesetzt.
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

EINSATZSZENARIEN

Erprobt für: Erfassung von Kompetenzstand der Fachdidaktik in Informatik in einzelnen Bundesländern. Unterscheidung von Kompetenzausprägungen nach Lehrerfahrung
Prinzipiell auch geeignet für: —
Nicht geeignet für: Messung des Fachwissens

ERLAUBTE KONSEQUENZEN

Erprobt für: Aussagen über die Struktur der Überzeugungen und motivationalen Orientierungen und deren Zusammenhänge untereinander möglich
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Die Skalen zur Erfassung der professionellen Überzeugungen (34 Items) und zur Erfassung der motivationalen Orientierungen (13 Items) sind verfügbar (Bender, Schaper & Seifert, 2018).

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Bender, E.; Hubwieser, P.; Schaper, N.; Margaritis, M.; Berges, M.; Ohrndorf, L.; Magenheim, J. & Schubert, S. (2015). Towards a competency model for teaching computer science. *Peabody Journal of Education, 90*(4), S. 519-532.
Bender, E.; Schaper, N. & Seifert, A. (2018). Professionelle Überzeugungen und motivationale Orientierungen von Informatiklehrkräften. *Journal for Educational Research Online. 10*(1), S. 70–99.

WEITERFÜHRENDE LITERATUR

ACM/IEEE = The Joint Task Force on Computing Curricula, Association for Computing Machinery, IEEE Computer Society. (2013). Computer Science Curricula 2013. Curriculum Guidelines for Undergraduate Degree Programs in Computer Science. Zugriff am 12.06.2014. Verfügbar unter: https://www.acm.org/binaries/content/assets/education/cs2013_web_final.pdf
Baumert, J. & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In M. Kunter; J. Baumert; W. Blum; U. Klusmann; S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften – Ergebnisse des Forschungsprogramms COACTIV* (S. 29–54). Münster: Waxmann.

LeScEd: Forschungskompetenz in Bildungswissenschaften

KURZFASSUNG

Name des Instruments: Erfassung der Forschungskompetenz in bildungswissenschaftlichen Studiengängen

Fachdomäne: Bildungswissenschaftliche Studiengänge, insbesondere Lehramtsstudierende, Erziehungswissenschaften, Frühe Bildung

Erfasste Kompetenzen: Teilkomponenten der Research Literacy: Information Literacy, Statistical Literacy, Critical Thinking

Zielgruppe: Studierende bildungswissenschaftlicher Studiengänge

Testart: Leistungstest; Multiple-Choice-Test

Modalität: Paper-Pencil-Test

Dauer: 30 Minuten pro Teilkomponente; 30, 60 oder 90 Minuten je nach Testzusammenstellung

Testaufbau: Für jede der drei Teilkomponenten wurden Itempools entwickelt, aus denen flexibel anwendungsspezifisch Testversionen generiert werden können.

Allgemeiner Testzweck: Besonders für den wiederholten Einsatz in Evaluationsstudien geeignet, da parallele Testversionen flexibel erstellt werden können.

Konkrete Einsatzszenarien: Studienerfolgsvorhersage, Embedded Assessment, Hochschulevaluation, Kursvergleiche, Lernerfolgskontrolle

Nicht geeignet für: Testung auf Individualebene (da keine adaptive Erhebungsprozedur entwickelt wurde)

Hinweise für praktische Anwendung: Die Zuverlässigkeit der Skalen hängt von der Item-Zusammensetzung ab; die Eignung hängt von der Passung für den beabsichtigten Testzweck ab. Belastbare Aussagen zu verschiedenen Anwendungsszenarien für den Praxiseinsatz sind nur eingeschränkt möglich. Die Antworten werden mittels IRT-Analysen ausgewertet. Zur Auswertung sind Beispielwerte für unterschiedliche Studiengänge verfügbar.

Bisheriger Einsatz in Projekten: LeScEd

Kontakt / Speicherort:
Dr. Jana Groß Ophoff
E-Mail: jana.grossophoff@ph-freiburg.de

ALLGEMEINE INFORMATIONEN

Name des Instruments: Erfassung der Forschungskompetenz in bildungswissenschaftlichen Studiengängen
Bisheriger Einsatz in Projekten: LeScEd
Kontakt / Speicherort:
Dr. Jana Groß Ophoff
E-Mail: jana.grossophoff@ph-freiburg.de

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Teilkomponenten der Research Literacy: Information Literacy, Statistical Literacy, Critical Thinking
Theoriemodell: Research Literacy nach den Modellen von Davidson (2013), Davies (1999), Lambert (2009) und Pedaste et al. (2015)
Testart: Leistungstest; Multiple-Choice-Test
Modalität: Paper-Pencil-Test
Testaufbau:
Itempool: Für jede der drei Teilkomponenten wurden Itempools entwickelt.
Informationskompetenz (23 Items): Fähigkeit, Informationsbedarf zu erkennen, adäquate Forschungsfragen zu formulieren sowie gezielt und reflektiert zu recherchieren
Statistische Kompetenz (68 Items): Fähigkeit, statistische Informationen aus Tabellen, Grafiken oder Ergebnisdarstellungen systematisch analysieren und evaluieren zu können
Forschungsbezogenes kritisches Denken (91 Items): Fähigkeit zur Interpretation von Evidenz bzw. zur Ableitung und Anwendung angemessener Schlussfolgerungen bzw. kritische Reflexion von Ergebnisinterpretationen, wie sie typischerweise in der Diskussion von Forschungsartikeln anzutreffen sind
Testversionen: Hieraus können flexibel anwendungsspezifisch Testversionen generiert werden.

TESTGÜTEKRITERIEN

Reliabilität: Die Zuverlässigkeit der Skalen hängt von der Item-Zusammensetzung ab. Beim konkreten Einsatz des Instruments ist die Reliabilität bei sehr guten und sehr schlechten Testergebnissen unzureichend.
Validität Testinhalt: In Expertenbefragungen geprüft
Validität Antwortprozesse: In kognitive Pretests geprüft
Validität interner Testaufbau: Drei Dimensionen des Konstrukts in IRT-Analyse bestätigt
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Zusammenhänge mit Motivation wurden geprüft: Lernorientierung als positiver Prädiktor für die selbsteingeschätzte Forschungskompetenz und Arbeitsvermeidung als negativer Prädiktor für die Leistung im Kompetenztest; keine Effekte für die Annäherungsleistungszielorientierung; ein positiver Effekt der Vermeidungsleistungszielorientierung auf die Entwicklung von bildungswissenschaftlicher Forschungskompetenz.
Validität Konsequenzen der Testergebnisse: keine Angaben
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 30 Minuten pro Teilkomponente; 30, 60 oder 90 Minuten je nach Testzusammenstellung
Benötigtes Material: Papierausdrucke, Stifte
Besonderheiten: Die Ergebnisse werden mittels IRT-Analysen ausgewertet. Zur Auswertung sind Vergleichswerte für unterschiedliche Studiengänge verfügbar.
Praxisbeispiel: In bildungswissenschaftlichen Studiengängen können Lehrveranstaltungen und Module, die dem Erwerb forschungsmethodischer Kompetenzen dienen, mittels Prä-Post-Erhebungen evaluiert werden. Weiterhin können Eingangsvoraussetzungen für forschungsmethodische Studiengänge vergleichend (z. B. verschiedene Studiengänge oder Kohorten) ermittelt werden oder zur Planung von curricularen Schwerpunktsetzungen genutzt werden.

FACHDOMÄNE

Erprobt für: Bildungswissenschaftliche Studiengänge, insbesondere Lehramtsstudierende, Erziehungswissenschaften, Frühe Bildung, Gesundheitspädagogik
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

ZIELGRUPPE

Erprobt für: Studierende bildungswissenschaftlicher Studiengänge
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

ALLGEMEINER TESTZWECK

Erprobt für: Besonders für den wiederholten Einsatz in Evaluationsstudien geeignet, da parallele Testversionen flexibel erstellt werden können
Prinzipiell auch geeignet für: Studienerfolgsvorhersage, Embedded Assessment, Hochschulevaluation, Kursvergleiche, Lernerfolgskontrolle
Nicht geeignet für: —

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Insbesondere ein gemeinsamer Einsatz mit Motivations- und Interessenskalen ist empfehlenswert.
Prinzipiell auch geeignet für: Andere Maße zur Kompetenzerfassung
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für die Kursebene, Hochschulebene.
Prinzipiell auch geeignet für: —
Nicht geeignet für: Individualdiagnostik, internationale Ebene

EINSATZSZENARIEN

Erprobt für: Studienerfolgsvorhersage, Embedded Assessment, Hochschulevaluation, Kursvergleiche, Lernerfolgskontrolle
Prinzipiell auch geeignet für: —
Nicht geeignet für: Individualdiagnostik, internationale Ebene

ERLAUBTE KONSEQUENZEN

Erprobt für: Forschungskompetenzen können gezielt für die Planung von Lehrcurricula sowie begleitend und abschließend für Evaluationszwecke erhoben werden.
Prinzipiell auch geeignet für: Landes- und Bundesebene
Nicht geeignet für: Individualdiagnostik, internationale Ebene

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Beim konkreten Einsatz des Instruments ist die Reliabilität bei sehr guten und sehr schlechten Testergebnissen bisher noch unzureichend. Die Items und Itempools sind validiert. Die daraus zusammengestellten Skalen müssen je nach Zusammensetzung allerdings in ihrer Zuverlässigkeit und Passung für den jeweiligen Testzweck geprüft werden. Bitte stimmen Sie sich dazu mit den Testentwickler*innen ab. Das Instrument ist auf Anfrage verfügbar.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Groß Ophoff, J.; Schladitz, S.; Leuders, J.; Leuders, T. & Wirtz, M. (2015). Assessing the development of Educational Research Literacy. The effect of courses on research methods in studies of Educational Science. *Peabody Journal of Education 90*(4), S. 560–573.
Schladitz, S.; Groß Ophoff, J. & Wirtz, M. (2015). Konstruktvalidierung eines Tests zur Messung bildungswissenschaftlicher Forschungskompetenz. In S. Blömeke & O. Zlatkin-Troitschanskaia (Hrsg.), Kompetenzen von Studierenden. *Zeitschrift für Pädagogik*, Beiheft 61, S. 167–184.
Winter-Hölzl, A.; Wäschle, K.; Wittwer, J.; Watermann, R. & Nückles, M. (2015). Entwicklung und Validierung eines Tests zur Erfassung des Genrewissens Studierender und Promovierender der Bildungswissenschaften. In S. Blömeke & O. Zlatkin-Troitschanskaia (Hrsg.), Kompetenzen von Studierenden. *Zeitschrift für Pädagogik*, Beiheft 61, S. 185–202.

WEITERFÜHRENDE LITERATUR

Davidson, K. (2013). Teachers' reported utilization of reading disabilities research. *Alberta Journal of Educational Research, 59*(3), S. 487–502.
Davies, P. (1999). What is evidence-based education? *British Journal of Educational Studies, 47*(2), S. 108–121.
Lambert, C. (2009). Pedagogies of participation in higher education: A case for research-based learning. *Pedagogy, Culture & Society, 17*(3), S. 295–309.
Pedaste, M., Mäeots, M., Siiman, L. A., De Jong, T., Van Riesen, S. A., Kamp, E. T. & Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational Research Review, 14*, S. 47–61.

MultiTex MDC: Verständnis mehrerer Dokumente

KURZFASSUNG

Name des Instruments: MDC-Test

Fachdomäne: Alle Studienfächer

Erfasste Kompetenzen: Multiple Document Comprehension (MDC)

Zielgruppe: Alle deutschen Studierenden

Testart: Leistungstest: Single Choice & Complex Multiple Choice

Modalität: Computergestützt

Dauer: 1 Stunde (für 2 Units)

Testaufbau: 67 Items; 5 Units mit 4 verschiedenen Anforderungen

Allgemeiner Testzweck: Einmalige Erfassung des Kompetenzstands; Kompetenzentwicklung durch wiederholte Testung

Konkrete Einsatzszenarien: Kompetenzstand bei Studierendengruppen; Ableitung von Förderpotenzialen; Evaluation von Maßnahmen zur Förderung von MDC

Nicht geeignet für: —

Hinweise für praktische Anwendung: Entwickelt für den Einsatz bei Studierenden; erfolgreich bei Studierenden eingesetzt; Einzel- oder Gruppentestungen sind möglich.

Bisheriger Einsatz in Projekten: Computerbasierte Messung des Textverstehens multipler Dokumente (MDC) bei Studierenden an vier deutschen Universitäten im Rahmen des MultiTex-Projektes: http://multitex.bildungsforschung.uni-bamberg.de

Kontakt / Speicherort:
Cornelia Schoor
Otto-Friedrich-Universität Bamberg,
Lehrstuhl für Empirische Bildungsforschung
Postanschrift: Markusplatz 3, 96045 Bamberg
E-Mail: cornelia.schoor@uni-bamberg.de

ALLGEMEINE INFORMATIONEN

Name des Instruments: MDC-Test
Bisheriger Einsatz in Projekten: Computerbasierte Messung des Textverstehens multipler Dokumente (MDC) bei Studierenden an zwei deutschen Universitäten im Rahmen des MultiTex-Projektes: http://multitex.bildungsforschung.uni-bamberg.de/
Kontakt / Speicherort:
Cornelia Schoor
Otto-Friedrich-Universität Bamberg,
Lehrstuhl für Empirische Bildungsforschung
Postanschrift: Markusplatz 3, 96045 Bamberg
E-Mail: cornelia.schoor@uni-bamberg.de

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Multiple Document Comprehension (MDC)
Theoriemodell: Documents Model Framework (Rouet & Britt, 2012), Wineburg (1991), MD-TRACE (Rouet & Britt, 2011)
Testart: Leistungstest; Single Choice & Complex Multiple Choice
Modalität: Computergestützt
Testaufbau:
Itempool: 67 Items; 5 Units mit 4 verschiedenen Anforderungen:
1. Inhalte über Texte hinweg vergleichen, 2. Inhalte über Texte hinweg miteinander verknüpfen, 3. Quellen einschätzen und vergleichen (source-source links), 4. Inhalte als von Quellen generiert repräsentieren und Quelle-Inhalt-Kombinationen vergleichen (Source-Content-Links vergleichen)
2 aus 5 Units müssen bearbeitet werden; verschiedene Testversionen sind durch verschiedene Kombinationen der Units möglich.

TESTGÜTEKRITERIEN

Reliabilität: EAP-Reliabilität = .69 bzw. WLE-Reliabilität = .67; Infit-Werte der Items: 0. 91–1.10; Outfit-Werte: 0. 74–1.14
Validität Testinhalt: Die Aufgaben wurden anhand der theoretischen Ansätze des Documents Model Framework (Britt & Rouet, 2012) sowie der Strategien von Wineburg (1991) entwickelt.
Validität Antwortprozesse: Die bisherige Auswertung der Verlaufsdaten (sog. Logdaten) stimmt mit theoretisch angenommenen Antwortprozessen überein.
Validität interner Testaufbau: Dimensionsanalysen zufolge bilden die Items theoriekonform ein eindimensionales Fähigkeitskonstrukt ab.
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Die Abiturnote korrelierte positiv signifikant mit dem MDC-Testwert ($r = -.44$, $p < .001$). Masterstudierende erreichten signifikant höhere MDC-Testwerte als Bachelorstudierende.
Validität Konsequenzen der Testergebnisse: Ein Zusammenhang von MDC mit dem Studienerfolg wird angenommen, muss aber noch überprüft werden.
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 1 Stunde (für 2 Units)
Benötigtes Material: Computer
Besonderheiten: Jede*r Teilnehmer*in erhält aus zeitlichen Gründen 2 von 5 Units.
Bisher erfolgter Einsatz: Zwei Studien (N = 310; N = 508) wurden durchgeführt, in denen Studierende an zwei Universitäten den Test an Computern bearbeiteten, um ihre MDC-Fähigkeiten zu überprüfen. In der zweiten Studie wurden zusätzlich Tests zum Arbeitsgedächtnis sowie zur Lesekompetenz und Lesegeschwindigkeit administriert.
Praxisbeispiel: Der Test kann für eine Erhebung im laufenden Semester an einer deutschen Universität zur Erfassung des Kompetenzstands in MDC von Studierenden eingesetzt werden. Darauf aufbauend könnten Förderpotenziale abgeleitet und Fördermaterialien entwickelt werden, um Studierende in ihrer MDC zu unterstützen. Eine individuelle Rückmeldung an die Studierenden ist möglich. Die Auswertung des Anwenderfeedbacks steht noch aus.

FACHDOMÄNE

Erprobt für: Sozial- und Geisteswissenschaften
Prinzipiell auch geeignet für: Alle Fächer
Nicht geeignet für: Keine Einschränkung

ZIELGRUPPE

Erprobt für: Bachelorstudierende und Masterstudierende zwischen 17 und 42 Jahren; zwei Bundesländer (Bayern und Hessen)
Prinzipiell auch geeignet für: Alle deutschen Studierenden
Nicht geeignet für: —

ALLGEMEINER TESTZWECK

Erprobt für: Erfassen des Textverstehens multipler Dokumente (MDC) bei Studierenden; einmalige Testung
Prinzipiell auch geeignet für: Kompetenzentwicklung bei wiederholter Testung
Nicht geeignet für: —

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Epistemische Überzeugungen, Zielorientierungen, Mental load, Mental effort, Task model, Lesehäufigkeiten und -gewohnheiten, Basale Computer-Fähigkeiten, Arbeitsgedächtnis, Abschlussnoten, Interesse, Selbsteingeschätztes Vorwissen, Lesekompetenz und -geschwindigkeit, Einstellungen zum Lesen
Prinzipiell auch geeignet für: Keine Einschränkung
Nicht geeignet für: Keine Einschränkung

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für Vergleiche innerhalb der getesteten Gruppe der Studierenden
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

EINSATZSZENARIEN

Erprobt für: Kompetenzstand von Studierendengruppen zu Forschungszwecken
Prinzipiell auch geeignet für: Ableitung von Förderpotenzialen; Evaluation von Maßnahmen zur Förderung von MDC
Nicht geeignet für: —

ERLAUBTE KONSEQUENZEN

Erprobt für: —
Prinzipiell auch geeignet für: Ableitung von Fördermaßnahmen
Nicht geeignet für: —

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Der Test wird weiter validiert und weiterentwickelt. Eine aktuelle Version des Instruments mit entsprechender Dokumentation wird zur Veröffentlichung vorbereitet.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Schoor, C., Hahnel, C., Artelt, C., Reimann, D., Kröhne, U., & Goldhammer, F. (2020). Entwicklung und Skalierung eines Tests zur Erfassung des Verständnisses multipler Dokumente von Studierenden [Developing and scaling a test of multiple document comprehension in university students]. *Diagnostica*. doi: 10.1026/0012-1924/a000231

Hahnel, C., Kroehne, U., Goldhammer, F., Schoor, C., Mahlow, N., & Artelt, C. (2019). Validating process variables of sourcing in an assessment of multiple document comprehension. *British Journal of Educational Psychology, 89*(3), 524-537. doi: 10.1111/bjep.12278

Hahnel, C., Schoor, C., Kröhne, U., Goldhammer, F., Mahlow, N., & Artelt, C. (2019). The role of cognitive load for university students' comprehension of multiple documents. *Zeitschrift für Pädagogische Psychologie, 33*(2), 105-118. doi: 10.1024/1010-0652/a000238

WEITERFÜHRENDE LITERATUR

Britt, A. M. & Rouet, J.-F. (2012). Learning with multiple documents: Component skills and their acquisition. In J. R. Kirby & M. J. Lawson (Hrsg.), *Enhancing the quality of learning: Dispositions, instruction, and learning processes* (S. 276–314). New York: Cambridge University Press.

Rouet, J. & Britt, M. A. (2012). Relevance processes in multiple document comprehension. In M. T. McCrudden; J. P. Magliano & G. Schraw (Hrsg.), *Text relevance and learning from text* (S. 19–52). Charlotte, NC: Information Age.

Wineburg, S. S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology, 83*(1), S. 73-87. DOI: 10.1037/0022-0663.83.1.73

M³K: Medienpädagogische Kompetenz

KURZFASSUNG

Name des Instruments: Instrument zur Messung der medienpädagogischen Kompetenz

Fachdomäne: Lehramtsstudierende aller Fächer und Schulformen

Erfasste Kompetenzen: Medienpädagogische Kompetenz in der Ausdifferenzierung von Mediendidaktik, Medienerziehung, medienbezogene Schulentwicklung, medienbezogene Einstellungen, medienbezogene Selbstwirksamkeitserwartungen, medientechnisches Wissen

Zielgruppe: Lehramtsstudierende höherer Semester

Testart: Leistungstest; Multiple-Choice-Test

Modalität: Paper-Pencil-Test

Dauer: 60 Minuten

Testaufbau: 102 Items: 3 Teilskalen zu medienpädagogischen Kompetenzen und zu medienpädagogischen Überzeugungen (Mediendidaktik, Medienerziehung, medienbezogene Schulentwicklung), 1 Subskala zur medienpädagogischen Selbstwirksamkeit und 1 Skala zum medientechnischen Wissen

Allgemeiner Testzweck: Erfassen des Stands und der Entwicklung der medienpädagogischen Kompetenz (als generische fachübergreifende Kompetenz)

Konkrete Einsatzszenarien: Untersuchungen zum Kompetenzerwerb von Lehramtsstudierenden/ Anwärter*innen in der ersten und zweiten Phase der Lehrerbildung; Evaluation von medienpädagogischen Lehrangeboten

Nicht geeignet für: Allgemeine Medienkompetenz; unterrichtsfachspezifische Aspekte von medienpädagogischer Kompetenz

Hinweise für praktische Anwendung: Entwickelt für den Einsatz bei Lehramtsstudierenden oder Lehramtsanwärter*innen; insbesondere für die Evaluation von Maßnahmen zur Förderung von medienpädagogischen Kompetenzen

Bisheriger Einsatz in Projekten: M³K: Modellierung und Messung medienpädagogischer Kompetenz

Kontakt / Speicherort:
Prof. Dr. Bardo Herzig
Universität Paderborn, Institut für Erziehungswissenschaften
Projektbericht hinterlegt im Katalog der Technischen Informationsbibliothek (TIB)-Hannover.
E-Mail: bardo.herzig@upb.de
Postanschrift: Warburger Str. 100, 33098 Paderborn

ALLGEMEINE INFORMATIONEN

Name des Instruments: Instrument zur Messung der medienpädagogischen Kompetenz
Bisheriger Einsatz in Projekten: M³K: Modellierung und Messung medienpädagogischer Kompetenz
Kontakt / Speicherort:
Prof. Dr. Bardo Herzig
Universität Paderborn, Institut für Erziehungswissenschaften
Projektbericht hinterlegt im Katalog der Technischen Informationsbibliothek (TIB)-Hannover.
E-Mail: bardo.herzig@upb.de
Postanschrift: Warburger Str. 100, 33098 Paderborn

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenz: Mediendidaktik, Medienerziehung, medienbezogene Schulentwicklung, medienbezogene Einstellungen, medienbezogene Selbstwirksamkeitserwartungen, medientechnisches Wissen
Theoriemodell: Konzepte zur medienpädagogischen Kompetenz (Blömeke, 2000; Gysbers, 2008; Tulodziecki, Herzig & Grafe, 2010); TPACK-Modell (Mishra & Koehler, 2006)
Testart: Leistungstest; Multiple-Choice-Format
Modalität: Paper-Pencil-Test
Testaufbau:
102 Items; 3 Teildimensionen mit den Subskalen mediendidaktische Kompetenz (16 Items), medienerzieherische Kompetenz (14 Items), medienbezogene Schulentwicklungskompetenz (10 Items), medientechnisches Wissen (26 Items), medienbezogene Selbstwirksamkeit (18 Items), medienbezogene Überzeugungen (18 Items)

TESTGÜTEKRITERIEN

Reliabilität: Interne Konsistenz (Cronbachs α) für Stichprobe mit expliziten Lerngelegenheiten: a = .58 Mediendidaktik, a = .59 Medienerziehung, a = .55 medienbezogene Schulentwicklung
Validität Testinhalt: Die Testinhalte wurden auf der Basis eines strukturierten theoriegleitet entwickelten Modells abgeleitet und in Experteninterviews empirisch validiert.
Validität Antwortprozesse: In einzelnen explorativen kognitiven Pretests betrachtet
Validität interner Testaufbau: Prüfung des Modells mit drei Dimensionen (mediendidaktische Kompetenz, medienerzieherische Kompetenz, medienbezogene Schulentwicklungskompetenz) gegen eindimensionales Modell zeigt eine bessere Passung.
Validität Zusammenhänge mit anderen Kompetenzindikatoren: keine Angaben
Validität Konsequenzen der Testergebnisse: keine Angaben
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 60 Minuten bei Einsatz aller Subskalen
Benötigtes Material: Kopiervorlagen der Testitems
Besonderheiten: Keine
Praxisbeispiel: Die Skalen wurden an einer Stichprobe von 912 Lehramtsstudierenden an unterschiedlichen Universitäten in Deutschland erprobt. Das Instrument kann zur Evaluation der Wirksamkeit von Lerngelegenheiten zum Erwerb medienpädagogischer Kompetenz eingesetzt werden; hierzu wurden Erhebungen an 3 Universitäten in medienpädagogischen Kursen durchgeführt. Das Instrument wurde zudem ins US-Amerikanische übersetzt und in den USA eingesetzt (N = 109, 7 Universitäten und Colleges).

FACHDOMÄNE

Erprobt für: Lehramtsstudierende aller Fächer und Schulformen
Prinzipiell auch geeignet für: Lehramtsanwärter*innen
Nicht geeignet für: —

ZIELGRUPPE

Erprobt für: Lehramtsstudierende höherer Semester
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

ALLGEMEINER TESTZWECK

Erprobt für: Erfassen des Kompetenzstands und der Kompetenzentwicklung; Evaluation von Lehr-Lern-Kursen
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: —
Prinzipiell auch geeignet für: Weitere Maße der Kompetenzmessung
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für einzelne Studierende im zeitlichen Verlauf, Kursebene, Hochschulebene.
Prinzipiell auch geeignet für: Internationale Ebene: Erste Adaption für die USA wird validiert.
Nicht geeignet für: —

EINSATZSZENARIEN

Erprobt für: Überprüfung der Wirkung von medienpädagogischen Lerngelegenheiten im Lehramtsstudium, Evaluation von medienpädagogischen Lehrangeboten
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

ERLAUBTE KONSEQUENZEN

Erprobt für: Beurteilung der Studien- bzw. Arbeitsleistung; Identifikation von Ansatzpunkten zur Verbesserung medienpädagogischer Lerngelegenheiten
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Bei Interesse an dem Instrument wenden Sie sich an Prof. Herzig: bardo.herzig@upb.de.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Herzig, B.; Martin, A.; Schaper, N. & Ossenschmidt, D. (2015). Modellierung und Messung medienpädagogischer Kompetenz – Grundlagen und erste Ergebnisse. In B. Koch-Priewe; A. Köker; J. Seifried & E. Wuttke (Hrsg.), *Kompetenzerwerb an Hochschulen: Modellierung und Messung. Zur Professionalisierung angehender Lehrerinnen und Lehrer sowie frühpädagogischer Fachkräfte* (S. 153–176). Bad Heilbrunn: Julius Klinkhardt.

Herzig, B.; Martin, A. (2018). Lehrerbildung in der digitalen Welt – konzeptionelle und empirische Aspekte. In J. Knopf; S. Ladel & A. Weinberger (Hrsg.), *Digitalisierung und Bildung* (S. 89–113). Wiesbaden: Springer VS.

Tiede, J.; Grafe, S. & Hobbs, R. (2015). Pedagogical Media Competences of Pre-Service Teachers in Germany and the United States of America: A Comparative Analysis of Theory and Practice. In O. Zlatkin-Troitschanskaia & S. Blömeke (Hrsg.), Modeling and Measuring Competencies in Higher Education. A Special Issue on the Assessment of Competencies in Higher Education. *Peabody Journal of Education*, *90*(4), S. 533-545. DOI: http://dx.doi.org/10.1080/0161956X.2015.1068083

WEITERFÜHRENDE LITERATUR

Herzig, B.; Schaper, N.; Breiter, A.; Hartig, J.; Grafe, S.; Martin, A.; Christoph, G.; Ossenschmidt, D.; Brüggemann, M. & Langen, T. (2013). M³K – Modeling and Measuring Pedagogical Media Competencies of Pre-Service Teachers. In S. Blömeke & O. Zlatkin-Troitschanskaia (Hrsg.), *The German funding initiative „Modeling and Measuring Competencies in Higher Education": 23 research projects on engineering, economics and social sciences, education and generic skills of higher education students* (S. 55–57) (KoKoHs Working Papers, 3). Berlin & Mainz: Humboldt University & Johannes Gutenberg University.

Tiede, J. & Grafe, S. (2015). Media Pedagogy in German and U.S. Teacher Education. *Comunicar*, *49*(14), S. 19-29. DOI: http://dx.doi.org/10.3916/C49-2016-02

Blömeke, S. (2000). *Medienpädagogische Kompetenz*. München: kopäd.

MokoMasch ADP: Advanced Design Project Skills im Maschinenbau

KURZFASSUNG

Name des Instruments: Advanced Design Project Skills (ADP)

Fachdomäne: Chemie / Verfahrenstechnik

Erfasste Kompetenzen: Kompetenzen von Studierenden bei interdisziplinären Design-Projekten

Zielgruppe: Masterstudierende, Berufstätige

Testart: Leistungstest; Ratings und offene Fragen

Modalität: Paper-Pencil-Test

Dauer: 60 Minuten

Testaufbau: 36 Items; 2 Teilskalen mit 16 bzw. 20 Items mit 3 bzw. 4 Facetten

Allgemeiner Testzweck: Einmalige Testung sozialer Kompetenzen für Designaufträge

Konkrete Einsatzszenarien: Lehr-Lern-Tool

Nicht geeignet für: Studienzulassung, Studienwahl/-empfehlung, Evaluation von Veranstaltungen, Längsschnittuntersuchung

Hinweise für praktische Anwendung: Der Test orientiert sich an dem Modell für Team Member Effectiveness von Loughry et al. (2007) und ist für eine konkrete Erhebungssituation adaptiert. Ein Teil der Aufgaben erfasst inhaltsbezogene Kompetenzen der übergeordneten technischen Konstruktionsaufgabe.

Bisheriger Einsatz in Projekten: MokoMasch „Modellierung von Kompetenzen von Studierenden des Maschinenbaus"

Kontakt / Speicherort:
Oliver Watteler
Gesis, Leibniz-Institut für Sozialwissenschaften
Postanschrift: Unter Sachsenhausen 6-8, 50667 Köln
E-Mail: oliver.watteler@gesis.org
Webseite: https://www.gesis.org/institut/mitarbeiterverzeichnis/

ALLGEMEINE INFORMATIONEN

Name des Instruments: Advanced Design Project Skills (ADP)
Bisheriger Einsatz in Projekten: MokoMasch „Modellierung von Kompetenzen von Studierenden des Maschinenbaus"
Kontakt / Speicherort:
Oliver Watteler
Gesis, Leibniz-Institut für Sozialwissenschaften
Postanschrift: Unter Sachsenhausen 6-8, 50667 Köln
Univ.-Prof. Dr. Augustin Kelava
Methodenzentrum / Methods Center
Wirtschafts- und Sozialwissenschaften /
Faculty of Economics and Social Sciences
Eberhard Karls Universität Tübingen
E-Mail: augustin.kelava@uni-tuebingen.de

DAS INSTRUMENT

INHALT UND STRUKTUR

Theoriemodell: Drei Kompetenzbereiche beim Design von Anlagen (z. B. Bio-Ethanol) werden unterschieden: „Allgemeine"/Professionelle Kompetenzen, Soziale Kompetenzen und Technisch-spezifische Kompetenzen. „Allgemeine Kompetenzen" beinhalten abstrakte fachwissenschaftliche Fähigkeiten zur Integration von technischen Lösungsmöglichkeiten bzw. Alternativen. Bei „Soziale Kompetenzen / Team Effectiveness" wurde insbesondere auf die Fähigkeit zur Abstimmung im Team (Loughry, Ohland & Moore, 2007) zurückgegriffen. „Spezifische Kompetenzen" umfassen aufgabenbezogene Aspekte.
Testart: Leistungstest; Ratings und offene Fragen
Modalität: Paper-Pencil-Test
Testaufbau:
Itempool: 36 Items: 2 Teilskalen mit 3 Facetten für Teil A (16 Items) bzw. 4 Facetten für Teil B (20 Items). Teil A dient der Erfassung allgemeiner Kompetenzfacetten: Teamkompetenzen (7 Items), Fähigkeiten der Evaluation von Lösungsmöglichkeiten (3 Items) und das Erkennen von komplexen Anlagendesign-Problemen (6 Items). Teil B dient der Erfassung spezieller technisch-fachwissenschaftlicher Kompetenzen: Allgemeines technisches Verständnis (im Kontext der zu lösenden Anlagenproblematik; 5 Items), Prozesswissen (5 Items), fundiertes Fachwissen (5 Items) und spezielles vertieftes Wissen (5 Items).

TESTGÜTEKRITERIEN

Reliabilität: Cronbachs α der Skalen: .61-.63
Validität Testinhalt: keine Angaben
Validität Antwortprozesse: keine Angaben
Validität interner Testaufbau: Zum englischen Test (CATME) s. Loughry, Ohland & Moore (2007)

Validität Zusammenhänge mit anderen Kompetenzindikatoren: Erwartungskonforme Zusammenhänge mit sozialen Kompetenzen (ISK, Kanning, 2009) und Ten-Item-Personality-Inventory (TIPI, Gosling, Rentfrow & Swann, 2003; deutsche Version: Muck, Hell & Gosling, 2007)
Validität Konsequenzen der Testergebnisse: keine Angaben
Testfairness: Keine Geschlechtsdifferenzen

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 60 Minuten
Benötigtes Material: Fragebogen und Papier
Besonderheiten: Der Test orientiert sich am Modell für Team Member Effectiveness von Loughry et al. (2007) und ist für eine konkrete Erhebungssituation adaptiert. Ein Teil der Aufgaben erfasst inhaltsbezogene Kompetenzen der übergeordneten technischen Konstruktionsaufgabe.
Praxisbeispiel: Das Instrument wurde als Begleitinstrument im Advanced Design Project an 3 Institutionen eingesetzt (N = 63) – eine universitäre Lehrveranstaltung, in der eine spätere berufliche Situation simuliert wird, in der Studierende eine Anlage z. B. zur Herstellung von Bio-Ethanol konstruieren sollen. Die Teilaufgaben der Gesamtaufgabe werden innerhalb des Teams verteilt und bearbeitet. Zur erfolgreichen Bearbeitung muss die Leistung der Mitglieder integriert werden (fachliche Leistung). Nur im Falle eines gelungenen Teamworks wird die Aufgabe bestanden. Hinsichtlich der Vorhersagekraft bzw. erklärenden Kraft des Instruments zeigt sich ein Zusammenhang zwischen Leistung und Ergebnis im Team, insbesondere bei den erfassten Soft Skills wie Kompetenzen zur Aneignung und Anwendung des Wissens in der Gruppensituation. Ein kompetenzorientiertes Lehr-Lern-Instrument wird in der Lehrveranstaltung entwickelt.

FACHDOMÄNE

Erprobt für: Chemie / Verfahrenstechnik
Prinzipiell auch geeignet für: Ingenieurwissenschaften; Anlagendesign
Nicht geeignet für: —

ZIELGRUPPE

Erprobt für: Masterstudierende
Prinzipiell auch geeignet für: Berufstätige
Nicht geeignet für: —

ALLGEMEINER TESTZWECK

Erprobt für: Erfassen sozialer Kompetenzen für Designaufträge
Prinzipiell auch geeignet für: Einmalige Testung; wiederholte Testung
Nicht geeignet für: Längsschnittuntersuchung

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Noten, fachwissenschaftliche (Kurz-)Tests, Persönlichkeitsfaktoren
Prinzipiell auch geeignet für: Validierung (selbstberichtsbezogen) von subjektiven Maßen (Soft Skills in Gruppensituationen)
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für Einzelpersonen (nach einer validierten Messmodellstruktur)
Prinzipiell auch geeignet für: Gruppenebene (wenn das Messinstrument angepasst wird, da es aktuell im Round-Robin-Design vorliegt)
Nicht geeignet für: Hochschulebene, Länder- und Bundesebene, internationale Ebene

EINSATZSZENARIEN

Erprobt für: Lehr-Lern-Tool
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

ERLAUBTE KONSEQUENZEN

Erprobt für: Prüfungsnoten bzw. Beurteilung der Studien- oder Arbeitsleistung
Prinzipiell auch geeignet für: —
Nicht geeignet für: Studienzulassung, Studienwahl/-empfehlung, Evaluation von Veranstaltungen

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Der Test wurde einmalig eingesetzt und nach Projektende nicht weiter verwendet. Zusammenhänge zu etablierten Instrumenten zur Erfassung der sozialen Kompetenz und Persönlichkeitsfaktoren wurden untersucht. Das Instrument wird für zukünftige Design-Projekte als Assessment-Tool zur Verfügung gestellt (in Form der Publikation von Anders, Pinkelman, Hampe & Kelava, 2015). Die Nutzung wird über den Kontakt zum Projektleiter organisiert. Die Anwendung ist kostenlos. Etwaige Items müssen dem Anwendungskontext allerdings sorgfältig angepasst werden.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Anders, B.; Pinkelman, R. J.; Hampe, M. J. & Kelava, A. (2015). Development, assessment, and comparison of social, technical, and general (professional) competencies in a university engineering advanced design project – a case study. In G. Spöttl & F. Musekamp (Hrsg.), *Kompetenz im Studium und in der Arbeitswelt. Nationale und internationale Ansätze zur Erfassung von Ingenieurkompetenzen* (S. 217-238). Frankfurt am Main: Peter Lang.

WEITERFÜHRENDE LITERATUR

Kanning, U. P. (2009). *Inventar Sozialer Kompetenzen (ISK).* Göttingen: Hogrefe.
Loughry, M. L.; Ohland, M. W. & Moore, D. D. (2007). Development of a theory-based assessment of team member effectiveness. *Educational and Psychological Measurement, 67*, S. 505.
Muck, P. M.; Hell, B. & Gosling, S. D. (2007). Construct validation of a short Five-Factor Model instrument: A self-peer study on the German adaptation of the Ten-item Personality Inventory (TIPI-G). *European Journal of Personality Assessment, 23*(3), S. 166-175.

PlanvoLL-D: Didaktische Adaptivität von Lehrkräften

KURZFASSUNG

Name des Instruments: Didaktische Adaptivität von Lehrkräften

Fachdomäne: Alle Lehramtsfächer

Erfasste Kompetenzen: Didaktische Adaptivität als Aspekt der Planungskompetenz angehender Lehrkräfte (= fachunspezifische Passung von Aufgaben für die geplante Unterrichtsstunde an die kognitiven Voraussetzungen einer Lerngruppe)

Zielgruppe: Lehrkräfte in der zweiten Ausbildungsphase, Studierende im Praxissemester

Testart: Leistungstest; inhaltliche Analyse von schriftlichen Unterrichtsplanungen der angehenden Lehrkräfte

Modalität: Rekonstruktion und Quantifizierung von Planungsentscheidungen aus schriftlichen Unterrichtsplanungen

Dauer: Abhängig vom Umfang der schriftlichen Unterrichtsplanung, z. B. 15–20 Minuten für eine 15 Seiten umfassende Planung

Testaufbau: 11 Items; 4 Bereiche: Deskriptive Ebene: Beschreibung der Situation (Lerngruppe), Beschreibung der Aufgabenstellung/en; analytische Ebene: Verknüpfung von Voraussetzungen der Lerngruppe und Aufgabenstellung/en, Verknüpfung der didaktischen Adaptivität mit weiteren Planungselementen mit je 2 bis 4 Items. 3 optionale Items zur didaktischen Adaptivität für Lernende mit besonderem Lernstatus

Allgemeiner Testzweck: Erfassen eines Teilaspekts von Planungskompetenz zu einem oder mehreren Zeitpunkten

Konkrete Einsatzszenarien: Analyse von Kompetenzentwicklung, Einflüssen von Lerngelegenheiten, unterschiedlichen Verläufen bei Teilgruppen, z. B. für kompetenzorientierte Optimierung der Lehre

Nicht geeignet für: —

Hinweise für praktische Anwendung: Die Reliabilität der Bewertung hängt von der Qualität der schriftlichen Unterrichtsplanung ab (z. B. Vollständigkeit der Aussagen). Die Auswertung wird von geschulten Raterinnen und Ratern durchgeführt und anschließend mit IRT-Methoden analysiert.

Bisheriger Einsatz in Projekten:
PlanvoLL: https://www.hf.uni-koeln.de/36932
PlanvoLL-D: https://www.hf.uni-koeln.de/37700

Kontakt / Speicherort:
Prof. Dr. Johannes König
Empirische Schulforschung, quantitative Methoden, Department Erziehungs- und Sozialwissenschaften, Humanwissenschaftliche Fakultät, Universität zu Köln
Webseite: https://www.hf.uni-koeln.de/33049
E-Mail: johannes.koenig@uni-koeln.de

ALLGEMEINE INFORMATIONEN

Name des Instruments: Didaktische Adaptivität von Lehrkräften
Bisheriger Einsatz in Projekten:
PlanvoLL: https://www.hf.uni-koeln.de/36932
PlanvoLL-D: https://www.hf.uni-koeln.de/37700
Kontakt / Speicherort:
Prof. Dr. Johannes König
Empirische Schulforschung, quantitative Methoden, Department Erziehungs- und Sozialwissenschaften, Humanwissenschaftliche Fakultät, Universität zu Köln
Webseite: https://www.hf.uni-koeln.de/33049
E-Mail: johannes.koenig@uni-koeln.de

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Didaktische Adaptivität als Aspekt der Planungskompetenz angehender Lehrkräfte = fachunspezifische Passung von Aufgaben für die geplante Unterrichtsstunde an die kognitiven Voraussetzungen einer Lerngruppe
Theoriemodell: Didaktische Adaptivität als situationsspezifische Fähigkeit der Wahrnehmung, Interpretation und Entscheidungsfindung (König et al., 2017; in review)
Testart: Leistungstest; schriftliche Unterrichtsplanungen der angehenden Lehrkräfte werden inhaltlich analysiert.
Modalität: Rekonstruktion und Quantifizierung von Planungsentscheidungen aus schriftlichen Unterrichtsplanungen
Testaufbau:
Itempool: 11 Items; 4 Bereiche: Deskriptive Ebene: Beschreibung der Situation (Lerngruppe), Beschreibung der Aufgabenstellung/en; analytische Ebene: Verknüpfung von Voraussetzungen der Lerngruppe und Aufgabenstellung/en, Verknüpfung der didaktischen Adaptivität mit weiteren Planungselementen mit je 2 bis 4 Items. 3 optionale Items zur didaktischen Adaptivität für Lernende mit besonderem Lernstatus

TESTGÜTEKRITERIEN

Reliabilität: Interrater-Reliabilität (Cohens Kappa für zwei geschulte Rater) im Mittel 0,83 (0,67 bis 0,93)
Validität Testinhalt: Zuwachs in der didaktischen Adaptivität zwischen erster und letzter Lehrprobe des Vorbereitungsdienstes/Referendariats ($d > 0{,}8$)
Validität Antwortprozesse: Ein Experten-Review der Kategorien stützte eine theoriekonforme Konstruktion (Buchholtz & König, 2015).
Validität interner Testaufbau: Die IRT-Skalierung zeigt eine EAP-Reliabilität von 0,703 (Theta-Varianz 1,631). Die Items verteilen sich hinreichend über das Fähigkeitsspektrum. Die Diskriminationswerte liegen bei > 0,3. Die gewichteten Abweichungsquadrate liegen zwischen 0,80 und 1,20.
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Validität wurde über die Abiturnote als Indikator für kognitive Grundfertigkeiten, die Note des Ersten Staatsexamens als Indikator für

Studienerfolg und über pädagogische Überzeugungen zum Lehren und Lernen geprüft. Übereinstimmungsvalidität mit pädagogischem Wissen zum Umgang mit Heterogenität.
Validität Konsequenzen der Testergebnisse: Der Prädiktor Planungskompetenz ist statistisch signifikant bei einer Regression mit der Skala Binnendifferenzierung als abhängige Variable: Referendarinnen und Referendare mit höheren Fähigkeitswerten in der Planungskompetenz erhielten von ihren unterrichteten Schülerinnen und Schülern eher eine Zuschreibung, dass sie binnendifferenziert nach Fähigkeiten unterrichten. Dieser Befund zeigt sich auch an anderen Stichproben sowie über Selbsteinschätzungen von Referendarinnen und Referendaren.
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES
Dauer: Abhängig vom Umfang der schriftlichen Unterrichtsplanung
Benötigtes Material: Schriftliche Planungen, Kodierleitfaden
Besonderheiten: Die Reliabilität der Bewertung hängt von der Qualität der schriftlichen Unterrichtsplanung ab (z. B. Vollständigkeit der Aussagen). Die Auswertung wird von geschulten Raterinnen und Ratern durchgeführt und anschließend mit IRT-Methoden statistisch analysiert.
Praxisbeispiel: König, Buchholtz & Dohmen (2015) bewerten zwei unterschiedliche schriftliche Planungen beispielhaft; der Bewertungsvorgang ist auf andere schriftliche Planungen übertragbar.

FACHDOMÄNE
Erprobt für: Alle Lehramtsfächer
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

ZIELGRUPPE
Erprobt für: Lehrkräfte in der zweiten Ausbildungsphase
Prinzipiell auch geeignet für: Studierende im Praxissemester
Nicht geeignet für: —

ALLGEMEINER TESTZWECK
Erprobt für: Erfassen eines Teilaspekts von Planungskompetenz zu einem oder mehreren Zeitpunkten
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN
Erprobt für: Einsatz mit fachwissenschaftlichen, fachdidaktischen und pädagogischen Wissenstests; Abiturnote, Note des Ersten/Zweiten Staatsexamens; pädagogische Überzeugungen zum Lehren und Lernen; Aspekte der selbst eingeschätzten Unterrichtsqualität
Prinzipiell auch geeignet für: Weitere Maße der Kompetenzerfassung
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN
Erprobt für: Ergebnisse sind statistisch aussagekräftig für Kohorten im zeitlichen Verlauf; Ausbildungsgang- und Ausbildungsinstitutionen, Bundeslandebene
Prinzipiell auch geeignet für: Individualdiagnostische Zwecke, Bundesebene, internationale Ebene (nach Prüfung von kultureller und sprachlicher Vergleichbarkeit)
Nicht geeignet für: —

EINSATZSZENARIEN
Erprobt für: Analyse von Kompetenzentwicklung, Einflüssen von Lerngelegenheiten, unterschiedlichen Verläufen bei Teilgruppen, z. B. für kompetenzorientierte Optimierung der Lehre
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

ERLAUBTE KONSEQUENZEN
Erprobt für: —
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ
Das Erhebungsinstrument ist in Form der Analysekriterien vollständig veröffentlicht worden (König, Buchholtz & Dohmen, 2015), sodass eine Nutzung durch interessierte Wissenschaftler*innen möglich ist. Das Instrument wurde im Projekt PlanvoLL-D auch für heterogene und inklusive Lerngruppen (gemeinsamer Unterricht) weiterentwickelt und um drei optionale Items ergänzt. Zum anderen wird die Adaptivität/Planungskompetenz für das Fach Deutsch fachdidaktisch ausgeweitet. Aktuell wird analog zum entwickelten Vorgehen die Dimension Strukturierung als Aspekt der Planungskompetenz erarbeitet und auf bestehende Datensätze aus PlanvoLL und PlanvoLL-D angewendet. Mittelfristig ist die Entwicklung eines standardisierten vignettenbasierten Tests zur Erfassung der didaktischen Adaptivität als Aspekt der Planungskompetenz vorgesehen.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN
Buchholtz, C. & König, J. (2015). Erfassung von Planungskompetenz im Praxissemester. *Journal für LehrerInnenbildung*, *15*(1), S. 39-45.
König, J.; Buchholtz, C. & Dohmen, D. (2015). Analyse von schriftlichen Unterrichtsplanungen: Empirische Befunde zur didaktischen Adaptivität als Aspekt der Planungskompetenz angehender Lehrkräfte. *Zeitschrift für Erziehungswissenschaft*, *18*(2), S. 375-404. https://link.springer.com/article/10.1007%2Fs11618-015-0625-7

WEITERFÜHRENDE LITERATUR
König, J.; Bremerich-Vos, A.; Buchholtz, C.; Lammerding, S.; Strauß, S.; Fladung, I. & Schleiffer, C. (2017). Die Bedeutung des Professionswissens von Referendarinnen und Referendaren mit Fach Deutsch für ihre Planungskompetenz (PlanvoLL-D). In S. Wernke & K. Zierer (Hrsg.), *Die Unterrichtsplanung: Ein in Vergessenheit geratener Kompetenzbereich?! Status Quo und Perspektiven aus Sicht der empirischen Forschung* (S. 121–133). Bad Heilbrunn: Klinkhardt.

PlanvoLL-D GPK: Pädagogisches Wissen

KURZFASSUNG

Name des Instruments: Pädagogisches Wissen (General Pedagogical Knowledge, GPK)

Fachdomäne: Pädagogisches Wissen in verschiedenen Fachdomänen

Erfasste Kompetenzen: Pädagogisches Wissen von angehenden und berufstätigen Lehrerinnen und Lehrern

Zielgruppe: Lehramtsstudierende zu Beginn der Lehramtsausbildung, während der Lehramtsausbildung; Referendarinnen und Referendare; berufstätige Lehrkräfte

Testart: Leistungstest; Kurzantworten und Multiple-Choice-Items

Modalität: Paper-Pencil-Test; auch Online-Einsatz möglich

Dauer: 20 Minuten (Kurzversion)

Testaufbau: 18 Items (veröffentlichte Kurzversion für die erste Lehrerbildungsphase). Testdesign-Matrix mit 4 inhaltlichen Dimensionen: Pädagogisches Wissen über Umgang mit Heterogenität, Strukturierung von Unterricht, Klassenführung/Motivierung, Leistungsbeurteilung und 3 Dimensionen kognitiver Anforderungen (Erinnern, Verstehen/Analysieren, Kreieren)

Allgemeiner Testzweck: Erfassen von Kompetenzstand und -entwicklung des pädagogischen Wissens während und nach der Lehramtsausbildung

Konkrete Einsatzszenarien: Hochschulvergleiche und -evaluation, Zusammenhänge mit externen Kriterien

Nicht geeignet für: —

Hinweise für praktische Anwendung: Einheitliche Durchführung unter Beobachtung beim Paper-Pencil-Test bzw. beim Online-Test ist erforderlich. Die Antworten auf die offenen Testfragen werden anhand einer Anleitung kodiert. Die Antworten werden IRT-skaliert ausgewertet. Die Ergebnisse können in Form eines Gesamtscores, Teilergebnissen einzelner Skalen (bei Einsatz der Langversion) oder als Kompetenzniveaus zurückgemeldet werden (Klemenz & König, 2019).

Bisheriger Einsatz in Projekten: CME, DIDAKTUM, EMW, LEK, LEK-R, LtP, PKE, PlanvoLL-D, SKILL, TEDS-M, TEDS-Follow Up, TEDS-Unterricht, TEDS-Validierung, ZuS-QS

Kontakt / Speicherort:
Prof. Dr. Johannes König
Empirische Schulforschung, quantitative Methoden, Department Erziehungs- und Sozialwissenschaften, Humanwissenschaftliche Fakultät, Universität zu Köln
Webseite: https://www.hf.uni-koeln.de/33049
Forschungsprojekte der Professur Empirische Schulforschung der Universität zu Köln: https://www.hf.uni-koeln.de/32432
E-Mail: johannes.koenig@uni-koeln.de

ALLGEMEINE INFORMATIONEN

Name des Instruments: Pädagogisches Wissen (General Pedagogical Knowledge, GPK)
Bisheriger Einsatz in Projekten: CME, DIDAKTUM, EMW, LEK, LEK-R, LtP, PKE, PlanvoLL-D, SKILL, TEDS-M, TEDS-Follow Up, TEDS-Unterricht, TEDS-Validierung, ZuS-QS
Kontakt / Speicherort:
Prof. Dr. Johannes König
Empirische Schulforschung, quantitative Methoden, Department Erziehungs- und Sozialwissenschaften, Humanwissenschaftliche Fakultät, Universität zu Köln
Webseite: https://www.hf.uni-koeln.de/33049
Forschungsprojekte der Professur Empirische Schulforschung der Universität zu Köln: https://www.hf.uni-koeln.de/32432
E-Mail: johannes.koenig@uni-koeln.de

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Pädagogisches Wissen von angehenden und berufstätigen Lehrerinnen und Lehrern
Theoriemodell: Pädagogisches Wissen als kognitive Facette der professionellen Kompetenz von (angehenden) Lehrkräften
Testart: Leistungstest; offene und geschlossene Item-Formate: Kurzantworten und Multiple-Choice-Items
Modalität: Paper-Pencil-Test; auch Online-Einsatz möglich
Testaufbau:
Itempool: 18 Items. Testdesign-Matrix mit 4 inhaltlichen Dimensionen: Pädagogisches Wissen über Umgang mit Heterogenität, Strukturierung von Unterricht, Klassenführung/Motivierung, Leistungsbeurteilung) und 3 Dimensionen kognitiver Anforderungen (Erinnern, Verstehen/Analysieren, Kreieren)
Testversionen: Veröffentlichte Kurzversion für die erste Lehrerbildungsphase (König & Blömeke, 2010). Auf Anfrage verfügbar ist eine Kurzversion für die zweite Lehrerbildungsphase bzw. für berufstätige Lehrkräfte.

TESTGÜTEKRITERIEN

Reliabilität: EAP-Reliabilität .76 bzw. .88 (Gesamtscore der Kurzversion für die erste Lehrerbildungsphase)
Validität Testinhalt: Theoriekonforme Befunde beim Vergleich von angehenden Lehrkräften (Studierende, Referendar*innen) in unterschiedlichen Ausbildungsstadien (EMW, LEK, LtP, PKE).
Validität Antwortprozesse: keine Angaben
Validität interner Testaufbau: Die theoretisch angenommene Aufteilung der inhaltlichen und der kognitiven Anforderungsdimensionen wurde statistisch bestätigt (TEDS-M, EMW, LEK).
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Pädagogisches Wissen korreliert schwächer mit Fachwissen als mit fachdidaktischem Wissen bei angehenden Lehrkräften der Fächer Mathematik, Deutsch und Englisch (König et al., 2018); pädagogisches Wissen korreliert schwach mit der Abiturnote; theoriekonforme Korrelationen mit anderen Facetten professioneller Kompetenz (fachdidaktisches Wissen, Fachwissen; TEDS-M, PKE, PlanvoLL-D, ZuS-QS) und Korrelationen mit video-basierten Assessments (CME, TEDS-Follow Up); theoriekonforme Zusammenhänge mit besuchten Lerngelegenheiten und Motivation (LEK, EMW, DIDAKTUM, CME) (s. Überblicksdarstellung in König, 2014).
Validität Konsequenzen der Testergebnisse: Prognostische Validität zeigt sich bei berufstätigen Lehrkräften in Form positiver Korrelationen zwischen dem Testergebnis und Schülereinschätzungen zur Unterrichtsqualität dieser Lehrkräfte (LEK-R, CME, ZuS-QS)
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 20 Minuten (Kurzversion)
Benötigtes Material: Stifte und Testhefte bzw. Computer/Laptops
Besonderheiten: Einheitliche Durchführung unter Beobachtung beim Paper-Pencil-Test bzw. beim Online-Test erforderlich; Antworten auf die offenen Testfragen werden anhand einer Anleitung kodiert. Die Antwortkodierung (einschließlich Doppelkodierung von 20 % aller Fragebögen) muss angeleitet und von geschulten Raterinnen und Ratern durchgeführt werden. Die Antworten werden umkodiert und IRT-skaliert ausgewertet (wenn im Rotationsdesign eingesetzt). Die Ergebnisse können in Form eines Gesamtscores, Teilergebnissen einzelner Skalen (bei Einsatz der Langversion) oder als Kompetenzniveaus zurückgemeldet werden (Klemenz & König, 2019).
Praxisbeispiel: Die bereits veröffentlichte Kurzversion (König & Blömeke, 2010) wird für praktische Zwecke genutzt, z. B. in der Qualitätsoffensive Lehrerbildung zur Evaluation eines neu entwickelten Lehr-Lern-Moduls (König et al., 2018). In verschiedenen Projekten (ZuS-QS, LEK, EMW) wurden differenzierte Analysen zum Zusammenhang des pädagogischen Wissens angehender Lehrkräfte während ihrer Ausbildung mit spezifischen Merkmalen ihrer Lerngelegenheiten durchgeführt. In der Studie *Learning to Practice – Das Praxissemester auf dem Prüfstand* (LtP) wurde pädagogisches Wissen in einem Pre-/Post-Test erfasst, um Lernstände von Lehramtsstudierenden während des Praxissemesters in NRW an drei Standorten zu evaluieren. Schulpraktische Tätigkeiten wie auch mentorielle Unterstützung können insbesondere den Zuwachs im handlungsnahen pädagogischen Wissen (Anforderung „Kreieren") stützen (König et al., in Druck).

FACHDOMÄNE

Erprobt für: Pädagogisches Wissen in verschiedenen Fachdomänen
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

ZIELGRUPPE

Erprobt für: Lehramtsstudierende zu Beginn der Lehramtsausbildung, während der Lehramtsausbildung; Referendarinnen und Referendare; berufstätige Lehrkräfte
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

ALLGEMEINER TESTZWECK

Erprobt für: Erfassen von Kompetenzstand und -entwicklung des pädagogischen Wissens während und nach der Lehramtsausbildung
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Instrumente zum fachdidaktischen Wissen, „pädagogischen Wissen für inklusiven Unterricht" (König et al., 2017)
Prinzipiell auch geeignet für: Weitere Maße der Kompetenzentwicklung
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für Kursebene, Hochschulebene, Landes- und Bundesebene, internationale Ebene
Prinzipiell auch geeignet für: Individualdiagnostik
Nicht geeignet für: —

EINSATZSZENARIEN

Erprobt für: Hochschulvergleiche, Hochschulevaluation; Zusammenhänge mit externen Kriterien; Vorhersage von beruflichen Erfolgskriterien (z. B. Unterrichtsqualität); Bildung von Kompetenzstufenniveaus
Prinzipiell auch geeignet für: Individuelle Entwicklungsverläufe
Nicht geeignet für: —

ERLAUBTE KONSEQUENZEN

Erprobt für: —
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

In 2019 und Folgejahren: Einsatz des Tests in verschiedenen empirischen Untersuchungen; seit 2010 ist Kurzversion für interessierte Forscher*innen verfügbar; weitere Versionen (u. a. Langversion des Tests) werden zur Veröffentlichung vorbereitet.

WEITERE WICHTIGE ANGABEN

Projekte, in denen das Instrument eingesetzt wird/wurde:
CME – Classroom Management Expertise; **DIDAKTUM** – Didaktisches Wissen und berufliche Motivation von Lehramtsstudierenden; **EMW** – Entwicklung von berufsspezifischer Motivation und pädagogischem Wissen in der Lehrerausbildung; **LEK** – Längsschnittliche Erhebung pädagogischer Kompetenzen von Lehramtsstudierenden; **LEK-R** – Längsschnittliche Erhebung pädagogischer Kompetenzen von Lehramtsstudierenden und Referendar*innen; **LtP** – Learning to Practice – Das Praxissemester auf dem Prüfstand; **PKE** – Professionelle Kompetenz von Englischlehrkräften: Fachdidaktisches Wissen angehender Englischlehrkräfte – Konzeption, Messung, Validierung; PlanvoLL-D – Die Bedeutung des professionellen Wissens angehender Deutschlehrkräfte für ihre Planung von Unterricht; **SKILL** – Studie zur Kompetenzentwicklung in der Lehrerinnen- und Lehrerausbildung; **TEDS-M** – Teacher Education and Development Study; **TEDS-FU** – Teacher Education and Development Study Follow up; **TEDS-Unterricht** – Teacher Education and Development Study – Unterricht; **TEDS-Validierung** – Teacher Education and Development Study – Validierung; **ZuS-QS** – Zukunftsstrategie Lehrer*innenbildung (ZuS), Teilprojekt Qualitätssicherung

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

König, J. & Blömeke, S. (2009). Pädagogisches Wissen von angehenden Lehrkräften: Erfassung und Struktur von Ergebnissen der fachübergreifenden Lehrerausbildung. *Zeitschrift für Erziehungswissenschaft*, *12*(3), 499-527. https://doi.org/10.1007/s11618-009-0085-z
König, J. & Blömeke, S. (2010). *Pädagogisches Unterrichtswissen (PUW). Dokumentation der Kurzfassung des TEDS-M-Testinstruments zur Kompetenzmessung in der ersten Phase der Lehrerausbildung.* Berlin: Humboldt-Universität.
König, J.; Blömeke, S.; Paine, L.; Schmidt, B. & Hsieh, F-J. (2011). General Pedagogical Knowledge of Future Middle School Teachers. On the Complex Ecology of Teacher Education in the United States, Germany, and Taiwan. *Journal of Teacher Education*, *62*(2), 188-201.

WEITERFÜHRENDE LITERATUR

Klemenz, S. & König, J. (2019). Modellierung von Kompetenzniveaus im pädagogischen Wissen bei angehenden Lehrkräften: Zur kriterialen Beschreibung von Lernergebnissen der fächerübergreifenden Lehramtsausbildung. *Zeitschrift für Pädagogik*, *65*(3), 355-377.
König, J.; Doll, J.; Buchholtz, N.; Förster, S.; Kaspar, K.; Rühl, A.-M.; Strauß, S.; Bremerich-Vos, A.; Fladung, I. & Kaiser, G. (2018). Pädagogisches Wissen versus fachdidaktisches Wissen? Struktur des professionellen Wissens bei angehenden Deutsch-, Englisch- und Mathematiklehrkräften im Studium. *Zeitschrift für Erziehungswissenschaft*, *21*(3), S. 611-648. https://doi.org/10.1007/s11618-017-0765-z
König, J.; Gerhard, K.; Melzer, C.; Rühl, A.-M.; Zenner, J. & Kaspar, K. (2017). Erfassung von pädagogischem Wissen für inklusiven Unterricht bei angehenden Lehrkräften: Testkonstruktion und Validierung. *Unterrichtswissenschaft*, *45*(4), S. 223-242.

Profile-P PCK: Mechanik von Physiklehrkräften

KURZFASSUNG

Name des Instruments: Test zur Erfassung fachdidaktischen Wissens – Physik (Profile-P-PCK)

Fachdomäne: Lehramt Physik – Mechanik

Erfasste Kompetenzen: Physikdidaktisches Wissen in der Mechanik

Zielgruppe: Lehramtsstudierende im Bachelor/Master, Lehrkräfte, Referendarinnen und Referendare

Testart: Leistungstest; Multiple Choice und offene Items

Modalität: Paper-Pencil-Test

Dauer: 65 Minuten

Testaufbau: 43 Items; 4 Wissensdimensionen: Instruktionsstrategien, Schülervorstellungen, Experimente, Fachdidaktische Konzepte

Allgemeiner Testzweck: Erfassen von fachdidaktischem Wissen und seiner Entwicklung; wiederholte Testung; Individualdiagnostik

Konkrete Einsatzszenarien: Hochschulevaluation, Bedarfsplanung Unterricht, Kursvergleiche, Studienerfolgsvorhersage, Lehr-Lern-Tool, Hochschulvergleiche, Akkreditierung

Nicht geeignet für: Internationale Ebene, Zwischen- und Abschlussprüfungen, Rankings, Einschätzung vorausgegangener Bildungsstufen, Berufseignung, Vorhersage der Leistungsentwicklung, Studienzulassung, Studienwahlempfehlung, Prüfungs- oder Abschlussnote, Berufsempfehlung

Hinweise für praktische Anwendung: Durchführungszeit ist verbindlich (Speedkomponente), Einsatz gemäß Testleitermanual. Bei der Interpretation der Testergebnisse sollte berücksichtigt werden, dass der inhaltliche Schwerpunkt auf dem Bereich Mechanik liegt, der in den meisten Lehramtsstudiengängen Physik nur in den ersten Semestern explizit in Lehrveranstaltungen behandelt wird.

Bisheriger Einsatz in Projekten: Profile-P (DaWis)

Kontakt / Speicherort:
Christoph Vogelsang
Zentrum für Bildungsforschung und Lehrerbildung – PLAZ Professional School, Universität Paderborn
Postanschrift: Warburger Str. 100, 33098 Paderborn
E-Mail: christoph.vogelsang@upb.de

ALLGEMEINE INFORMATIONEN

Name des Instruments: Test zur Erfassung fachdidaktischen Wissens – Physik (Profile-P-PCK)
Bisheriger Einsatz in Projekten: Profile-P, Profile-P+
Kontakt / Speicherort:
Christoph Vogelsang
E-Mail: christoph.vogelsang@upb.de
Webseite: Nach Projektabschluss ist das Instrument zugänglich über https://www.forschungsdaten-bildung.de

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Physikdidaktisches Wissen in der Mechanik
Theoriemodell: Modell professioneller Handlungskompetenz angehender Physiklehrkräfte und Modell physikdidaktischen Wissens (Gramzow, Riese & Reinhold, 2013) und zugehöriges Item-Entwicklungsmodell, siehe Abschnitt „Weitere wichtige Angaben"
Testart: Leistungstest; Multiple Choice und offene Items
Modalität: Paper-Pencil-Test
Testaufbau:
43 Items (23 offen, 20 geschlossen). 4 Wissensdimensionen: Instruktionsstrategien (9 Items), Schülervorstellungen (12 Items), Experimente (6 Items), Fachdidaktische Konzepte (16 Items)

TESTGÜTEKRITERIEN

Reliabilität: Psychometrische Anforderungen an Schwierigkeit, Trennschärfen und Reliabilität werden erfüllt: EAP-Relibialität des Gesamtscores: 0.84 (Pilotstudie); 0.79 (Profile-P-Hauptstudie); Interkoderreliabilität bei offenen Items: k = 0.75 bis 0.93 (Profile-P-Hauptstudie)
Validität Testinhalt: Itemerstellung anhand einschlägiger Literatur; Validierung durch Expertenbefragung zur curricularen Passung an vier deutschen Universitätsstandorten
Validität Antwortprozesse: Überprüfung durch lautes Denken mit 15 Proband*innen; zur korrekten Beantwortung der Items greifen die Proband*innen; erwartungskonform auf fachdidaktische Wissensbestände zurück.
Validität interner Testaufbau: Überprüfung der Itemzuordnung im Modell durch Expertenbefragung; Ein vierdimensionales Modell weist theoriekonform die beste Passung an die empirischen Daten auf (Pilotstudie).
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Erwartungskonforme positive Korrelationen zu Credit Points im Bereich Fachdidaktik und der Physiknote im Abitur; weitere Validierung in Profile-P+
Validität Konsequenzen der Testergebnisse: keine Angaben
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 65 Minuten
Benötigtes Material: Testhefte und Stifte
Besonderheiten: Die Zeit ist verbindlich (Speedkomponente); Testeinsatz gemäß Testleitermanual. Bei der Interpretation der Testergebnisse sollte berücksichtigt werden, dass der inhaltliche Schwerpunkt auf dem Bereich Mechanik liegt, der in den meisten Lehramtsstudiengängen Physik nur in den ersten Semestern explizit in Lehrveranstaltungen behandelt wird.
Praxisbeispiel: Der Test wurde in den Projekten Profile-P und Profile-P+ eingesetzt. In Profile-P wurde er im Querschnitt bei N = 294 Lehramtsstudierenden mit Fach Physik in drei Bundesländern (hoher Anteil von Studienanfänger*innen) eingesetzt. Der Einsatz diente der Evaluation der Wirkung des Physiklehramtsstudiums und der Aufklärung verschiedener Einflussfaktoren (z. B. erworbene Credit Points in der Fachdidaktik, Abiturnote, Studienfortschritt) auf den Kompetenzstand im physikdidaktischen Wissen. In Profile-P+ wurde das Instrument weiterentwickelt und anschließend in einer Längsschnitterhebung zu drei Zeitpunkten im Bachelorstudium und bei Masterstudierenden vor und nach dem Praxissemester eingesetzt. Die Erhebungen werden sowohl in Deutschland als auch in Österreich durchgeführt. Die Untersuchungen zielen darauf, die Kompetenzentwicklung nachzuzeichnen und Zusammenhänge zur Entwicklung des fachdidaktischen Wissens in Physik sowie der Performanz in Erklärsituationen, bei der Unterrichtsplanung und der Unterrichtsreflexion zu analysieren.

FACHDOMÄNE

Erprobt für: Lehramt Physik – Mechanik
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

ZIELGRUPPE

Erprobt für: Bachelor- und Masterstudierende Lehramt Physik
Prinzipiell auch geeignet für: Lehrkräfte, Referendar*innen
Nicht geeignet für: —

ALLGEMEINER TESTZWECK

Erprobt für: Erfassen von Kompetenzstand und -entwicklung, wiederholte Testung
Prinzipiell auch geeignet für: Individualdiagnostik, Evaluation von Lehr-Lern-Angeboten
Nicht geeignet für: —

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Noten, Fachwissen Physik, bildungswissenschaftliches Wissen, Performanztests (Erklären, Planung, Reflexion)
Prinzipiell auch geeignet für: Weitere Maße der Kompetenzerfassung
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für die Hochschulebene und Kursebene
Prinzipiell auch geeignet für: Einzelpersonen im zeitlichen Verlauf, Landes- und Bundesebene
Nicht geeignet für: Internationale Ebene

EINSATZSZENARIEN

Erprobt für: Hochschulevaluation, Bedarfsplanung Unterricht, Kursvergleiche
Prinzipiell auch geeignet für: Studienerfolgsvorhersage, Lehr-Lern-Tool, Hochschulvergleiche, Akkreditierung
Nicht geeignet für: Zwischen- und Abschlussprüfungen, Rankings, Einschätzung vorausgegangener Bildungsstufen, Berufseignung, Vorhersage der Leistungsentwicklung

ERLAUBTE KONSEQUENZEN

Erprobt für: Kursevaluation
Prinzipiell auch geeignet für: Beurteilung der Studien- bzw. Arbeitsleistung, Kursempfehlung
Nicht geeignet für: Studienzulassung, Studienwahlempfehlung, Prüfungs- oder Abschlussnote, Berufsempfehlung

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

In Profile-P+ wird das Instrument zur Analyse des Zusammenhangs von physikdidaktischem Wissen und der Performanz von Lehramtsstudierenden in den Standardsituationen erklären, Unterricht planen und Unterricht reflektieren sowie zur Analyse der Wissensentwicklung primär im Bachelorstudium eingesetzt.
Anfragen zum Einsatz des Instruments für die Evaluation von Studienprogrammen für das Lehramt im Fach Physik liegen vor.
Das Instrument ist veröffentlicht in Gramzow (2015), in Profile-P+ wurden allerdings weitere Verbesserungen an der Kodierung vorgenommen.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

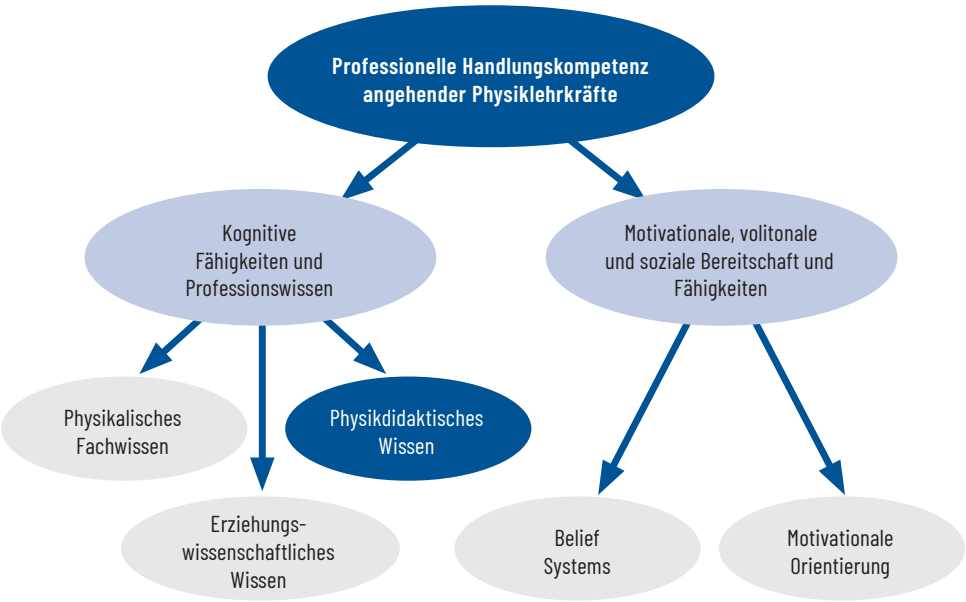
Gramzow, Y.; Riese, J. & Reinhold, P. (2013). Modellierung fachdidaktischen Wissens angehender Physiklehrkräfte. *Zeitschrift für Didaktik der Naturwissenschaften, 19*, S. 7-30.
Gramzow, Y.; Riese, J. & Reinhold, P. (2014). Prospective physics teachers' pedagogical content knowledge – Validating a test instrument by using a think aloud study. In C. P. Constantinou; N. Papadouris & A. Hadjigeorgiou (Hrsg.), *E-Book Proceedings of the ESERA 2013 Conference: Science Education Research For Evidence-based Teaching and Coherence in Learning* (S. 2242-2249). Nicosia, Zypern: European Science Education Research Association.
Gramzow, Y. (2015). *Fachdidaktisches Wissen von Lehramtsstudierenden im Fach Physik: Modellierung und Testkonstruktion.* Logos Berlin.

WEITERFÜHRENDE LITERATUR

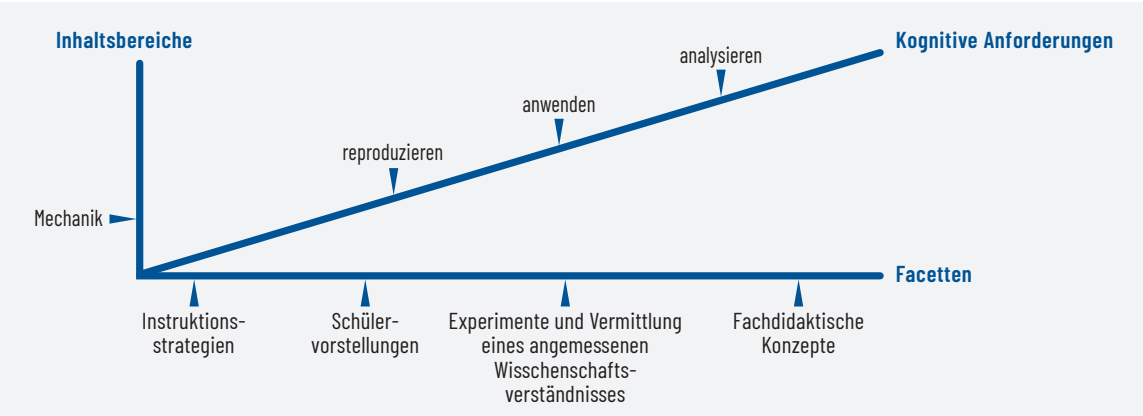
Riese, J.; Gramzow, Y. & Reinhold, P. (2017). Das fachdidaktische Wissen von Anfängern und Fortgeschrittenen im Lehramtsstudiengang Physik. *Zeitschrift für Didaktik der Naturwissenschaften, 23*, S. 99-112. DOI: 10.1007/s40573-017-0059-2

WEITERE WICHTIGE ANGABEN

Modell professioneller Handlungskompetenz angehender Physiklehrkräfte (Gramzow, Riese & Reinhold, 2013)



Modell physikdidaktischen Wissens (Gramzow, Riese & Reinhold, 2013)



Profile-P CK: Mechanik von Physiklehrkräften

KURZFASSUNG

Name des Instruments: Test zur Erfassung fachlichen Wissens – Physik (Profile-P-CK)

Fachdomäne: Lehramt Physik – Mechanik

Erfasste Kompetenzen: Physikalisches Wissen in der Mechanik

Zielgruppe: Studierende Lehramt und Fach Physik, Lehrkräfte, Referendar*innen

Testart: Leistungstest; Multiple-Choice-Test & Single Select

Modalität: Paper-Pencil-Test

Dauer: 50 Minuten

Testaufbau: 48 Items; 3 Teilskalen: Schulisches Wissen (20 Items), Vertieftes Wissen (14 Items), Universitäres Wissen (14 Items)

Allgemeiner Testzweck: Erfassen von Fachwissen und -entwicklung, wiederholte Testung

Konkrete Einsatzszenarien: Hochschulevaluation, Bedarfsplanung Unterricht, Kursvergleiche, Studienerfolgsvorhersage, Lehr-Lern-Tool, Akkreditierung

Nicht geeignet für: Internationale Ebene, Zwischen- und Abschlussprüfungen, Rankings, Einschätzung vorausgegangener Bildungsstufen, Berufseignung, Studienzulassung, Studienwahlempfehlung, Prüfungs- oder Abschlussnote, Berufsempfehlung

Hinweise für praktische Anwendung: Die Testzeit ist verbindlich festgelegt. Testeinsatzinteressierte sollten sich an die Koordination des Verbundprojekts Profile-P+ wenden: christoph.vogelsang@upb.de

Bisheriger Einsatz in Projekten: Profile-P; Profile-P+

Kontakt / Speicherort:
Christoph Vogelsang
Zentrum für Bildungsforschung und Lehrerbildung
PLAZ Professional School, Universität Paderborn
Postanschrift: Warburger Str. 100, 33098 Paderborn
E-Mail: christoph.vogelsang@upb.de

ALLGEMEINE INFORMATIONEN

Name des Instruments: Test zur Erfassung fachlichen Wissens – Physik (Profile-P-CK)
Bisheriger Einsatz in Projekten: Profile-P; Profile-P+
Kontakt / Speicherort:
Christoph Vogelsang
Zentrum für Bildungsforschung und Lehrerbildung
PLAZ Professional School, Universität Paderborn
Postanschrift: Warburger Str. 100, 33098 Paderborn
E-Mail: christoph.vogelsang@upb.de

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Physikalisches Wissen in der Mechanik
Theoriemodell: Modell professioneller Handlungskompetenz angehender Physiklehrkräfte (Riese & Reinhold, 2012); Inhaltsbereich Mechanik für die Facetten Schulisches Wissen, Vertieftes Schulwissen und Universitäres Wissen
Testart: Leistungstest; Multiple-Choice-Test & Single Select
Modalität: Paper-Pencil-Test
Testaufbau:
Itempool: 48 Items; 3 Teilskalen: Schulisches Wissen (20 Items), Vertieftes Wissen (14 Items) Universitäres Wissen (14 Items)

TESTGÜTEKRITERIEN

Reliabilität: EAP/PV-Reliabilität = .84, Varianz = 64.7 nach Pilotanalysen
Validität Testinhalt: Lehrbuchanalysen für Schulen und Universitäten, Analysen der Vorlesungsinhalte von Einführungsvorlesungen in die Physik, Zuordnung von Aufgaben zu Fachstufen, Diskussion der Zuordnung mit Fachphysikern und Fachdidaktikern, Inhaltsvalidierung mittels Konsenskodierung durch Lehramtsstudierende höherer Semester, Interviewstudie mit Physiklehrkräften zur Überprüfung der Berufsrelevanz
Validität Antwortprozesse: Think-Aloud-Studie zur Überprüfung der Lösungsprozesse der Fachwissensfacetten durchgeführt.
Validität interner Testaufbau: Die Dimensionalität im IRT-Modell überprüft: Die Daten passten theoriekonform signifikant besser zum dreidimensionalen als zum eindimensionalen Modell.
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Hohe Korrelationen zu einem Test für mathematische Fähigkeiten, allerdings von Schulwissen zu universitärem Wissen abnehmend. Etwas geringere Korrelationen zur Abiturgesamtnote sowie zur letzten Mathematik- und Physiknote in der Schulzeit. Sehr geringe signifikante Korrelationen zur letzten Schulnote im Fach Deutsch. Positive Korrelationen zu fachdidaktischem physikalischen Wissen und zur Fähigkeit im Physikerklären
Validität Konsequenzen der Testergebnisse: keine Angaben
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 50 Minuten
Benötigtes Material: Testhefte und Stift
Besonderheiten: Die Testzeit ist verbindlich festgelegt; maximal 48 Punkte erreichbar.
Praxisbeispiel: Der Test wurde in den Projekten Profile-P bzw. Profile-P+ eingesetzt. In Profile-P wurde er im Querschnitt bei N = 294 Lehramtstudierenden mit Fach Physik in drei Bundesländern (hoher Anteil von Studienanfänger*innen) eingesetzt. Der Einsatz diente der Evaluation der Wirkung des Physiklehramtsstudiums und der Aufklärung verschiedener Einflussfaktoren (z. B. erworbene Credit Points in der Fachdidaktik, Abiturnote, Studienfortschritt) auf den Kompetenzstand. In Profile-P+ wurde das Instrument weiterentwickelt und anschließend in einer Längsschnitterhebung zu drei Zeitpunkten im Bachelorstudium und bei Masterstudierenden vor und nach dem Praxissemester eingesetzt. Die Untersuchungen zielen darauf, die Fachwissensentwicklung nachzuzeigen und Zusammenhänge zur Entwicklung des fachdidaktischen Wissens in Physik sowie der Performanz in Erklärsituationen, bei der Unterrichtsplanung und der Unterrichtsreflexion zu analysieren.

FACHDOMÄNE

Erprobt für: Lehramt Physik – Mechanik
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

ZIELGRUPPE

Erprobt für: Studierende Lehramt und Fach Physik
Prinzipiell auch geeignet für: Lehrkräfte, Referendar*innen
Nicht geeignet für: —

ALLGEMEINER TESTZWECK

Erprobt für: Erfassung von Fachwissenstand und -entwicklung, wiederholte Testung
Prinzipiell auch geeignet für: Evaluationen von Lehr-Lern-Angeboten
Nicht geeignet für: —

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Noten, fachdidaktisches Wissen Physik, Performanztests (Erklären, Planung, Reflexion)
Prinzipiell auch geeignet für: Weitere Maße der Kompetenzmessung
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind für die Hochschulebene statistisch aussagekräftig und eingeschränkt auf Kursebene
Prinzipiell auch geeignet für: —
Nicht geeignet für: Internationale Ebene

EINSATZSZENARIEN

Erprobt für: Hochschulevaluation, Bedarfsplanung Unterricht, bedingt Kursvergleiche
Prinzipiell auch geeignet für: Studienerfolgsvorhersage, Lehr-Lern-Tool, Akkreditierung
Nicht geeignet für: Zwischen- und Abschlussprüfungen, Rankings, Einschätzung vorausgegangener Bildungsstufen, Berufseignung

ERLAUBTE KONSEQUENZEN

Erprobt für: —
Prinzipiell auch geeignet für: Beurteilung der Studien- bzw. Arbeitsleistung
Nicht geeignet für: Studienzulassung, Studienwahlempfehlung, Prüfungs- oder Abschlussnote, Berufsempfehlung

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

In Profile-P+ das Instrument zur Analyse des Zusammenhangs von physikalischem Fachwissen und der Performanz von Lehramtsstudierenden in den Standardsituationen erklären, Unterricht planen und Unterricht reflektieren sowie zur Analyse der Wissensentwicklung primär im Bachelorstudium eingesetzt.

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Riese; J.; Kulgemeyer, C.; Zander, S.; Borowski, A.; Fischer, H.; Gramzow, Y.; Reinhold, P.; Schecker, H. & Tomczyszyn, E. (2015). Modellierung und Messung des Professionswissens in der Lehramtsausbildung Physik. In S. Blömeke & O. Zlatkin-Troitschanskaia (Hrsg.), Kompetenzen von Studierenden. *Zeitschrift für Pädagogik*, Beiheft 61, S. 55–79. Weinheim: Beltz.

Enkrott, P.; Buschhüter, D. & Borowski, A. (2018). Modeling and Development of Professional Content Knowledge of Pre-Service Physics Teachers. In C. Maurer (Hrsg.), *Qualitätsvoller Chemie- und Physikunterricht – normative und empirische Dimensionen. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Regensburg 2017* (S. 896–899). Regensburg: Universität Regensburg.

Profile-P DET: Dialogischer Erklärtest für Physiklehrkräfte

KURZFASSUNG

Name des Instruments: DET: Dialogischer Erklärtest

Fachdomäne: Lehramt Physik (Mechanik) sowie andere naturwissenschaftliche Domänen

Erfasste Kompetenzen: Performanzbasierte Erklärfähigkeit im Fach Physik

Zielgruppe: Angehende und berufstätige Lehrkräfte der Physik

Testart: Leistungstest; Videoanalyse von Erklärungen

Modalität: Performanztest

Dauer: 10 Minuten pro Test plus 10 Minuten Vorbereitungszeit

Testaufbau: Drei Themen zur Erklärung; trainierte Schüler*innen/ Lehramtsstudierende (Sachunterricht) stellen teilstandardisierte Fragen

Allgemeiner Testzweck: Erfassung des Zusammenhangs Performanz und Wissen, Evaluation von Interventionen

Konkrete Einsatzszenarien: Entwicklung von handlungsnahen Kompetenzen im Lehramt

Nicht geeignet für: Schüler*innen, Zulassungen zum Lehramt, als Prüfungsersatz

Hinweise für praktische Anwendung: Training der Schüler*innen erforderlich, die die Adressaten für die Erklärungen spielen

Bisheriger Einsatz in Projekten: ProfiLe-P (EWis)

Kontakt / Speicherort:
PD Dr. Christoph Kulgemeyer
Universität Bremen, Institut für Didaktik der Naturwissenschaften
Postanschrift: Universität Bremen, Otto-Hahn-Allee 1, 28334 Bremen
E-Mail: kulgemeyer@physik.uni-bremen.de

ALLGEMEINE INFORMATIONEN

Name des Instruments: DET: Dialogischer Erklärtest
Bisheriger Einsatz in Projekten: ProfiLe-P (EWis)
Kontakt / Speicherort:
PD Dr. Christoph Kulgemeyer
Universität Bremen, Institut für Didaktik der Naturwissenschaften
Postanschrift: Otto-Hahn-Allee 1, 28334 Bremen
E-Mail: kulgemeyer@physik.uni-bremen.de

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Performanzbasierte Erklärfähigkeit im Fach Physik
Theoriemodell: Modell dialogischen Erklärens mit Vermittlungsabsicht (Kulgemeyer & Schecker, 2012; Wittwer & Renkl, 2008)
Testart: Leistungstest; Videoanalyse von Erklärungen
Modalität: Performanztest
Testaufbau:
Itempool: 3 Themen sind einer Person zu erklären; trainierte Schüler*innen stellen teilstandardisierte Fragen. Auswertung mithilfe kriterienbasierter Videoanalyse hinsichtlich Sachgerechtheit und Adressatengemäßheit der Erklärung (12 Auswertekategorien).

TESTGÜTEKRITERIEN

Reliabilität: Cronbachs α = .772 (Gesamtscore), Kodiererübereinstimmung: PÜ = 0.91
Validität Testinhalt: Kategorien aus Literatur zur Erklärqualität; Themen aus der Mechanik entsprechen Themen der Mittelstufenphysik.
Validität Antwortprozesse: Intendierte Prozesse untersucht durch Stimulated Recall Interviews.
Validität interner Testaufbau: Eindimensionales Raschmodell beschreibt die Daten theoriekonform.
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Ein nomologisches Netzwerk ist theoretisch formuliert und überprüft: Positive mittlere Korrelationen zu Fachwissen, zu fachdidaktischem Wissen, zur Selbstwirksamkeitsüberzeugung im Erklären in Physik und zu konstruktivistischen epistemologischen Überzeugungen.
Validität Konsequenzen der Testergebnisse: keine Angaben
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 10 Minuten pro Test
Benötigtes Material: Materialsatz mit Abbildungen, Szenarienbeschreibungen, Papier, Stifte, Videokamera
Besonderheiten: 10 Minuten Vorbereitungszeit pro Proband zur Vorbereitung auf die Erklärung. 10 Minuten Erklärungszeit mit trainiertem Schüler bzw. trainierter Schülerin. Auswertung der Videos nach Kodiermanual, 50 Minuten Auswertungszeit pro Video, Scorebildung durch Summe auftretender Kategorien
Praxisbeispiel: Der Test wurde an fünf deutschen Universitäten eingesetzt, um den Einfluss universitär vermittelten professionellen Wissens auf Handlungsqualität in einer ausgewählten Unterrichtssituation, dem Erklären, zu evaluieren. Dazu haben angehende Lehrkräfte der Physik die Tests zum Fachwissen und zum fachdidaktischen Wissen ausgefüllt, die curricular valide zum Inhalt des Lehramtsstudiums sind. Weiterhin haben sie in den Erklärtests agiert. Zusammenhänge wurden mit Pfadmodellen analysiert; das fachdidaktische Wissen mediiert den Einfluss des Fachwissens auf Erklärqualität. Die Schüler*innen, die die Adressaten für die Erklärungen spielen, wurden sorgfältig geschult. Dabei wurde Videofeedback eingesetzt und im Training wurden Erklärungen nach den Vorgaben eingeübt, videobasiert reflektiert und weiter verfeinert.

FACHDOMÄNE

Erprobt für: Lehramt Physik (Mechanik)
Prinzipiell auch geeignet für: Andere naturwissenschaftliche Domänen
Nicht geeignet für: —

ZIELGRUPPE

Erprobt für: Angehende Lehrkräfte Physik
Prinzipiell auch geeignet für: Lehrkräfte, Referendar*innen
Nicht geeignet für: Schüler*innen

ALLGEMEINER TESTZWECK

Erprobt für: Zusammenhangsanalysen von Performanz und Wissen
Prinzipiell auch geeignet für: Evaluation von Interventionen
Nicht geeignet für: Zulassungen zum Lehramt, Eignungsprüfung von Quereinsteiger*innen

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Wissenstest zu Fachwissen und fachdidaktischem Wissen, Skalen zu Einstellungen
Prinzipiell auch geeignet für: Andere Einsätze nach Rücksprache mit den Testentwickler*innen
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für Gruppenvergleiche.
Prinzipiell auch geeignet für: Individualdiagnostik
Nicht geeignet für: —

EINSATZSZENARIEN

Erprobt für: Entwicklungen von handlungsnahen Kompetenzen im Lehramt
Prinzipiell auch geeignet für: Andere Einsätze nach Rücksprache mit den Testentwickler*innen
Nicht geeignet für: Prüfungsersatz

ERLAUBTE KONSEQUENZEN

Erprobt für: —
Prinzipiell auch geeignet für: Die Zulässigkeit von möglichen Konsequenten muss zunächst in der Validierung geprüft werden.
Nicht geeignet für: —

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Unter https://www.explainingscience.de findet sich eine digitale Version des Tests mit einem simulierten Dialog, der von Kulgemeyer und Bartels erstellt wurde.

WEITERE WICHTIGE ANGABEN

Interessenten an der Nutzung können Kontakt mit Priv.-Doz. Dr. Christoph Kulgemeyer, kulgemeyer@physik.uni-bremen.de, aufnehmen.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Kulgemeyer, C. & Tomczyszyn, E. (2015). Physik erklären – Messung der Erklärensfähigkeit angehender Physiklehrkräfte in einer simulierten Unterrichtssituation. *Zeitschrift für Didaktik der Naturwissenschaften, 21*(1), S. 111-126.
Kulgemeyer, C. & Riese, J. (2018). From professional knowledge to professional performance: The impact of CK and PCK on teaching quality in explaining situations. *Journal of Research in Science Teaching,* S. 1-26. DOI: 10.1002/tea.21457
Bartels, H. & Kulgemeyer, C. (2018). Explaining physics: an online test for self-assessment and instructor training. *European Journal of Physics, 40*(1). DOI: 10.1088/1361-6404/aaeb5e

WEITERFÜHRENDE LITERATUR

Kulgemeyer, C. & Schecker, H. (2012). Physikalische Kommunikationskompetenz – Empirische Validierung eines normativen Modells. *Zeitschrift für Didaktik der Naturwissenschaften, 18*, S. 29–54.
Wittwer, J. & Renkl, A. (2008). Why instructional explanations often do not work: A framework for understanding the effectiveness of instructional explanations. *Educational Psychologist, 43*(1), S. 49–64.

Pro-KomMa: KomMa-Tests zu fachlichem und pädagogischem Wissen (MCK, MPCK und GPK) von frühpädagogischen Fachkräften in Mathematik

KURZFASSUNG

Name des Instruments: KomMa-Test – Test zur Erfassung professioneller Kompetenz von (angehenden) frühpädagogischen Fachkräften im Bereich Mathematik

Fachdomäne: Angehende frühpädagogische Fachkräfte (sowohl mit angestrebtem Fachschulabschluss als auch Fachhochschulabschluss)

Erfasste Kompetenzen: Mathematisches Fachwissen (MCK), mathematikdidaktisches Wissen (MPCK), allgemein-pädagogisches Wissen (GPK)

Zielgruppe: Beginn und Ende der Ausbildung (Fachschule und Fachhochschule), alle Bundesländer in Deutschland; Berufseinsteiger*innen (max. 5 Jahre nach Abschluss der Ausbildung)

Testart: Leistungstest; Multiple-Choice-Items und Items mit offenen Antworten

Modalität: Paper-Pencil-Test (empfohlen), auch Online-Testung

Dauer: 60 Minuten

Testaufbau: 70 Items; 3 Teiltests: KomMa-MCK-Test (24 Items), KomMa-MPCK-Test (28 Items), KomMa-GPK-Test (18 Items)

Allgemeiner Testzweck: Erfassen des Wissensstands, Entwicklung des professionellen Wissens im Bereich Mathematik

Konkrete Einsatzszenarien: Nur für Forschungszwecke

Nicht geeignet für: Individualdiagnostik

Hinweise für praktische Anwendung: Nur für Forschungszwecke, keine Individualdiagnostik

Bisheriger Einsatz in Projekten:
KomMa: https://www.projekte.hu-berlin.de/de/pro-komma/projekt-komma
Pro-KomMa: https://www.projekte.hu-berlin.de/de/pro-komma

Kontakt / Speicherort:
Dr. Lars Jenßen
Humboldt-Universität zu Berlin
Institut für Erziehungswissenschaften, Abteilung Grundschulpädagogik
Postanschrift: Unter den Linden 6, 10099 Berlin
E-Mail: lars.jenssen@hu-berlin.de

ALLGEMEINE INFORMATIONEN

Name des Instruments: KomMa-Test – Test zur Erfassung professioneller Kompetenz von (angehenden) frühpädagogischen Fachkräften im Bereich Mathematik
Bisheriger Einsatz in Projekten:
KomMa:
https://www.projekte.hu-berlin.de/de/pro-komma/projekt-komma
Pro-KomMa: https://www.projekte.hu-berlin.de/de/pro-komma
Kontakt / Speicherort:
Dr. Lars Jenßen
Humboldt-Universität zu Berlin
Institut für Erziehungswissenschaften, Abteilung Grundschulpädagogik
Postanschrift: Unter den Linden 6, 10099 Berlin
E-Mail: lars.jenssen@hu-berlin.de

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Mathematisches Fachwissen (MCK), mathematikdidaktisches Wissen (MPCK), allgemein-pädagogisches Wissen (GPK)
Theoriemodell: Differenzierung der Wissensfacetten des professionellen Wissens in (mathematisches) Fachwissen, (mathematik-) didaktisches Wissen und allgemein-pädagogisches Wissen nach Shulman (1986) und Ball & Bass (2002); Kompetenz als Kontinuum (professionelles Wissen als Disposition professioneller Kompetenz) nach Blömeke, Gustafsson & Shavelson (2015)
Testart: Leistungstest; Multiple-Choice-Items und Items mit offenen Antworten
Modalität: Paper-Pencil-Test (empfohlen), auch Online-Testung
Testaufbau:
Itempool: 70 Items in den 3 Subtests: KomMa-MCK-Test: 24 Items zu den Inhalten "Zahlen, Mengen und Operationen", "Form, Raum und Veränderung", "Größen, Messen und Relationen" und "Daten, Kombinatorik und Zufall" mit den Prozessen "Modellieren", "Problemlösen", "Argumentieren", "Kommunizieren", "Darstellen" und "Muster und Strukturen erkennen"; KomMa-MPCK-Test: 28 Items zu den Dimensionen "Entwicklung und Diagnose mathematischer Fähigkeiten bei Kindern" (17 Items) und "Gestaltung mathematischer Lernumgebungen" (11 Items); KomMa-GPK-Test: 18 Items zu den Inhalten "Bildungstheoretische Grundlagen" (5 Items), "Psychologische Grundlagen" (6 Items) und "Didaktisch-methodische Grundlagen" (7 Items)

TESTGÜTEKRITERIEN

Reliabilität: Reliabilitätskoeffizienten für 2PL-Modelle für die Stichprobe der angehenden frühpädagogischen Fachkräfte (N = 1.851): KomMa-MCK-Test: Rel = 0,88; KomMa-MPCK-Test: Rel = 0,87; KomMa-GPK-Test: Rel = 0,68
Validität Testinhalt: Ein Experten-Panel (N = 12) stützt die Validität des Testinhalts (Jenßen, Dunekacke & Blömeke, 2015)
Validität Antwortprozesse: Analyse der Testbearbeitungsprozesse mittels der Methode des lauten Denkens mit 18 angehenden frühpädagogischen Fachkräften (Jenßen, Dunekacke & Blömeke, 2015).
Validität interner Testaufbau: Mehrere Studien mit verschiedenen Populationen zeigen jeweils eine theoriekonforme empirische Differenzierbarkeit von MCK, MPCK und GPK (Blömeke et al., 2015).
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Mehrere Studien mit verschiedenen Populationen bestätigen die theoretisch angenommenen Zusammenhänge mit affektiv-motivationalen Dispositionen, allgemein-kognitiven Fähigkeiten, handlungsnahen Konstrukten und Performanz (Jenßen, Dunekacke, Eid & Blömeke, 2015; Dunekacke, Jenßen, Eilerts & Blömeke, 2016; Blömeke, Jenßen, Dunekacke, Suhl, Grassmann & Wedekind, 2015; Blömeke, Jenßen, Grassmann, Dunekacke & Wedekind, 2017).
Validität Konsequenzen der Testergebnisse: keine Angaben
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 60 Minuten
Benötigtes Material: Testbögen und Stift oder PC/Laptop/Tablet
Besonderheiten: Nicht zur Individualdiagnostik geeignet
Praxisbeispiel: Der Paper-Pencil-Test kann z. B. in Seminargruppen oder im Klassenverband eingesetzt werden. Bei größeren Stichproben und dem Einsatz weiterer Verfahren, z. B. Intelligenztests oder Fragebögen zu Lerngelegenheiten, wird die Anwendung im Multi-Matrix-Design empfohlen. Hierzu kann bei der Projektkoordination die Verteilung der Items für derartige Anwendungen erfragt werden. Die Testinhalte beziehen sich explizit auf Ausbildungsinhalte, die an Fachschulen und Fachhochschulen in allen Bundesländern Deutschlands unterrichtet werden. Bisher wurden die Tests bei ca. 1.800 angehenden frühpädagogischen Fachkräften und bei ca. 250 frühpädagogischen Fachkräften in der Berufseinstiegsphase eingesetzt. Zudem wird der Test von der Abteilung Elementarbereich des Deutschen Zentrums für Lehrerbildung (DZLM) für die Evaluation von Fortbildungen eingesetzt.

FACHDOMÄNE

Erprobt für: Frühpädagogik Mathematik
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

ZIELGRUPPE

Erprobt für: Angehende frühpädagogische Fachkräfte (sowohl mit angestrebtem Fachschulabschluss als auch Fachhochschulabschluss) zu Beginn und Ende der Ausbildung (Fachschule und Fachhochschule), alle Bundesländer in Deutschland
Prinzipiell auch geeignet für: Berufseinsteiger*innen (max. 5 Jahre nach Abschluss der Ausbildung)
Nicht geeignet für: Frühpädagogische Fachkräfte mit längerer Berufserfahrung

ALLGEMEINER TESTZWECK

Erprobt für: Erfassen des Wissensstands, Entwicklung des professionellen Wissens im Bereich Mathematik
Prinzipiell auch geeignet für: Evaluation von Fortbildungen
Nicht geeignet für: Individualdiagnostik

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Empfohlen werden Fragebögen zur Erfassung von mathematikbezogenen Überzeugungen, Videotests (z. B. Tests zur mathematikbezogenen Situationswahrnehmung und Handlungsplanung von Dunekacke et al., 2016), Intelligenztests
Prinzipiell auch geeignet für: Weitere Maße der Kompetenzmessung
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für die Ausbildungsebene (Fachschule vs. Fachhochschule und Beginn vs. Ende)
Prinzipiell auch geeignet für: —
Nicht geeignet für: Individualdiagnostik, Länder- und Bundesebene, internationale Vergleiche

EINSATZSZENARIEN

Erprobt für: Forschungszwecke
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

ERLAUBTE KONSEQUENZEN

Erprobt für: —
Prinzipiell auch geeignet für: —
Nicht geeignet für: Individualdiagnostik

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Bei Interesse am Test kontaktieren Sie bitte Dr. Lars Jenßen (lars.jenssen@hu-berlin.de). Zur Nutzung reichen Sie bitte eine Kurzbeschreibung ein, wofür Sie den Test verwenden möchten, und geben Sie stets den Verweis auf grundlegende Projektpublikationen (Blömeke et al., 2015) mit an. Die Projektkoordinator*innen behalten sich in Einzelfällen vor, dass die Nutzung erst nach Beendigung der Qualifikationsarbeiten im Projekt Pro-KomMa erlaubt wird.

WEITERE WICHTIGE ANGABEN

Zu beachten ist, dass das Feld der frühen mathematischen Bildung stetig weiter erforscht wird und sich pädagogische Konzepte in der Kita-Praxis ändern können. Auch die Testaufgaben unterliegen diesen Änderungen in der Praxis.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Blömeke, S.; Jenßen, L.; Dunekacke, S.; Suhl, U.; Grassmann, M. & Wedekind, H. (2015). Leistungstests zur Messung der professionellen Kompetenz frühpädagogischer Fachkräfte. *Zeitschrift für Pädagogische Psychologie, 29*(3-4), S. 177-191.

Jenßen, L.; Dunekacke, S. & Blömeke, S. (2015). Qualitätssicherung in der Kompetenzforschung: Empfehlungen für den Nachweis von Validität in Testentwicklung und Veröffentlichungspraxis. *Zeitschrift für Pädagogik*, Beiheft 61, S. 11-31.

WEITERFÜHRENDE LITERATUR

Jenßen, L.; Dunekacke, S.; Eid, M. & Blömeke, S. (2015). The relationship of mathematical competence and mathematics anxiety – An application of latent state-trait theory. *Zeitschrift für Psychologie, 223*(1), S. 31-39.

Blömeke, S.; Jenßen, L.; Grassmann, M.; Dunekacke, S. & Wedekind, H. (2017). Process mediates structure: Relation of preschool teacher education and preschool teachers' knowledge. *Journal of Educational Psychology, 109*, S. 338-354.

Dunekacke, S.; Jenßen, L.; Eilerts, K. & Blömeke, S. (2016). Epistemological beliefs of prospective preschool teachers and their relation to knowledge, perception, and planning abilities in the field of mathematics: A process model. *ZDM, 48*(1-2), S. 125–137.

PRO-SRL-EVA – PRO-SRL: Mein Lerntagebuch

KURZFASSUNG

Name des Instruments: Mein Lerntagebuch

Fachdomäne: Alle Studiengänge

Erfasste Kompetenzen: Kompetenzen des selbstregulierten Lernens (SRL) (Motivation, Selbstwirksamkeit, Planung, Zielsetzung, Reflexion) und Stimmung

Zielgruppe: Studierende während der Prüfungsvorbereitung, Schüler*innen ab der Sekundarstufe während der Prüfungsvorbereitung

Testart: Tagebuch

Modalität: Smartphone-App (online)

Dauer: 5 Minuten pro Ausfülleinheit

Testaufbau: 20 Items; Kombination aus geschlossenen und offenen Fragen

Allgemeiner Testzweck: Erfassen der SRL-Kompetenzen im Zeitverlauf innerhalb von 2 Wochen oder auch über einen anderen Zeitraum

Konkrete Einsatzszenarien: Lehr-Lern-Tool, Zwischenprüfungen, Abschlussprüfungen, Kursvergleiche, Hochschulvergleiche

Nicht geeignet für: Grundschüler*innen, Einmaltestungen, internationale Ebene, Zulassungsauswahl, Studienerfolgsvorhersage, Studienabbruchvorhersage, Hochschulevaluation, Akkreditierung, Rankings, Einschätzung vorausgegangener Bildungsstufen, Berufseignung, Vorhersage Leistungsentwicklung

Hinweise für praktische Anwendung: Für jede Person müssen ein Teilnehmercode und ein Passwort generiert werden. Die Auswertung erfolgt über Zeitreihenanalysen. Die resultierenden Daten müssen vor statistischen Analysen erst aufbereitet werden.

Bisheriger Einsatz in Projekten: PRO-SRL-EVA

Kontakt / Speicherort:
E-Mail: schmitz@psychologie.tu-darmstadt.de
Webseite: https://www.paedpsy.psychologie.tu-darmstadt.de/team_paedpsy/aktuelles_team_paedpsy/schmitz/schmitz.de.jsp

ALLGEMEINE INFORMATIONEN

Name des Instruments: Mein Lerntagebuch
Bisheriger Einsatz in Projekten: PRO-SRL-EVA
Kontakt / Speicherort:
E-Mail: schmitz@psychologie.tu-darmstadt.de
Webseite: https://www.paedpsy.psychologie.tu-darmstadt.de/team_paedpsy/aktuelles_team_paedpsy/schmitz/schmitz.de.jsp

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Kompetenzen des selbstregulierten Lernens (SRL) (Motivation, Selbstwirksamkeit, Planung, Zielsetzung, Reflexion) und Stimmung
Theoriemodell: SRL-Modell nach Schmitz & Wiese (2006)
Testart: Tagebuch
Modalität: Smartphone-App (online)
Testaufbau:
Itempool: 20 Items, geschlossene und offene Fragen

TESTGÜTEKRITERIEN

Reliabilität: keine Angaben
Validität Testinhalt: keine Angaben
Validität Antwortprozesse: keine Angaben
Validität interner Testaufbau: Laut Faktoranalysen sind die erfassten Variablen ein Teil von selbstreguliertem Lernen als übergeordnetem Faktor.
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Signifikante Korrelationen zu anderen SRL-Instrumenten (z. B. mit dem SRL-Fragebogen nach Bellhäuser, Roth & Schmitz, 2015; $r = .68^{**}$, SJT zur Erfassung der Motivationsregulationsstrategien; $r = .54^{**}$)
Validität Konsequenzen der Testergebnisse: keine Angaben
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 5 Minuten pro Ausfülleinheit
Benötigtes Material: Smartphone
Besonderheiten: Für jede Person müssen ein Teilnehmercode und ein Passwort generiert werden. Die Eingaben werden über Zeitreihenanalysen ausgewertet. Die resultierenden Daten müssen vor statistischen Analysen erst aufbereitet werden.
Praxisbeispiel: Das Lerntagebuch wurde in seiner ersten Version im Jahr 2016 mit Psychologie-Studierenden (N = 48) und 2017 überarbeitet bei weiteren Studierenden (N = 75) der TU Darmstadt im Zeitraum von je 14 Tagen getestet. Die Studierenden erhielten individuelles Feedback, z. B. zu ihrer Motivationsentwicklung im Zeitverlauf und zum Ausmaß der persönlichen Zielerreichung. 2018 wurde das Tagebuch erneut überarbeitet und weitere Variablen wurden integriert (Motivationsregulationskompetenzen und geplanter vs. tatsächlicher Einsatz von Lernstrategien). Eingesetzt wird es derzeit im Rahmen des Projekts PRO-SRL-EVA bei Psychologie- und Mathematik-Studierenden der TU Darmstadt, der Universität Augsburg und der Universität Wien. Praktische Konsequenzen werden nach Abschluss der Evaluation abgewogen.

FACHDOMÄNE

Erprobt für: Mathematik und Psychologie
Prinzipiell auch geeignet für: Alle Studiengänge
Nicht geeignet für: —

ZIELGRUPPE

Erprobt für: Studierende während der Prüfungsvorbereitung
Prinzipiell auch geeignet für: Schüler*innen ab Sekundarstufe während der Prüfungsvorbereitung
Nicht geeignet für: Grundschüler*innen

ALLGEMEINER TESTZWECK

Erprobt für: Erfassen der SRL-Kompetenzen im Zeitverlauf innerhalb von 2 Wochen
Prinzipiell auch geeignet für: Andere Zeiträume
Nicht geeignet für: Einmaltestungen

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Situational Judgment Test zum Erfassen der Motivationsregulationskompetenz, SRL-QuAK
Prinzipiell auch geeignet für: Alle SRL-Variablen
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für die Kursebene, Hochschulebene
Prinzipiell auch geeignet für: Einzelpersonen
Nicht geeignet für: Länder- und Bundesebene, internationale Ebene

EINSATZSZENARIEN

Erprobt für: Lehr-Lern-Tool, Zwischenprüfungen, Abschlussprüfungen, Kursvergleiche, Vergleiche des Lernverhaltens zwischen verschiedenen Gruppen Studierender (z. B. Regelzeit- vs. Langzeitstudierende)
Prinzipiell auch geeignet für: Hochschulvergleiche
Nicht geeignet für: Zulassungsauswahl, Studienerfolgsvorhersage, Studienabbruchvorhersage, Hochschulevaluation, Akkreditierung, Rankings, Einschätzung vorausgegangener Bildungsstufen, Berufseignung, Vorhersage Leistungsentwicklung

ERLAUBTE KONSEQUENZEN

Erprobt für: —
Prinzipiell auch geeignet für: Verbesserung der Lernqualität
Nicht geeignet für: —

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Perspektivisch kann die App allen interessierten Studierenden kostenlos zur Verfügung gestellt werden, damit sie den eigenen Lernprozess und -erfolg besser überwachen können. Allerdings ist hierfür eine regelmäßige Wartung durch eine*n Informatiker*in sicherzustellen.

WEITERE WICHTIGE ANGABEN

Das Verfahren ist nicht sprachneutral, da die Items zur Erfassung der SRL-Variablen auf Deutsch sind. Eine Ausweitung auf andere Sprachen setzt einen Übersetzungs- und erneuten Validierungsprozess voraus.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Schmitz, B. & Wiese, B. S. (2006). New perspectives for the evaluation of training sessions in self-regulated learning: Time-series analyses of diary data. *Contemporary Educational Psychology*, *31*(1), S. 64-96. http://dx.doi.org/10.1016/j.cedpsych.2005.02.002

PRO-SRL-EVA – PRO-SRL: Situational Judgement Test zur Motivationsregulation

KURZFASSUNG

Name des Instruments: Situational Judgement Test zur Erfassung des konditionalen Strategiewissens zur Motivationsregulation

Fachdomäne: Alle Studienfächer

Erfasste Kompetenzen: Konditionales Strategiewissen zur Motivationsregulation (Wissen über [un]geeignete Strategien zur Motivationsregulation je nach Situationen und Anforderungen)

Zielgruppe: Studierende im deutschsprachigen Raum

Testart: Leistungstest; Situational Judgement Test

Modalität: Paper-Pencil-Test oder Online-Test

Dauer: 15 Minuten

Testaufbau: Der Test besteht aus 8 (Kurzversion: 5) standardisierten Situationsbeschreibungen zu motivationalen Problemen vor oder während des Lernens (Lernsituationen „Schreiben einer wissenschaftlichen Arbeit" und „Vorbereitung auf eine Prüfung") mit je 9 Items zu möglichen Strategien, deren Eignung auf einer 6-stufigen Likert-Skala einzuschätzen ist.

Allgemeiner Testzweck: Erfassen des Kompetenzstands und der Kompetenzentwicklung (z. B. nach Intervention)

Konkrete Einsatzszenarien: Studienerfolgsvorhersage, Diagnose bei Lernproblematiken, Studienleistungsentwicklung, Studienabbruchvorhersage

Nicht geeignet für: Zwischenprüfungen, Abschlussprüfungen, Akkreditierung, Nutzung außerhalb der Hochschule

Hinweise für praktische Anwendung: Gruppentestung möglich; Scoring entsprechend Urteilen von Expert*innen über die Differenz zwischen geeigneten und ungeeigneten Strategien je Situation und Anforderung. Das Instrument ist in Steuer, Engelschalk, Eckerlein und Dresel (2019) ausführlich dargestellt. Interessierte können Kontakt mit den Autor*innen aufnehmen.

Bisheriger Einsatz in Projekten:
PRO-SRL: https://www.kompetenzen-im-hochschulsektor.de/pro-srl/
PRO-SRL-EVA: https://www.uni-augsburg.de/de/fakultaet/philsoz/fakultat/psychologie/forschung/pro-srl-eva/

Kontakt / Speicherort:
Prof. Dr. Markus Dresel
Lehrstuhl für Psychologie, Universität Augsburg
E-Mail: markus.dresel@phil.uni-augsburg.de

ALLGEMEINE INFORMATIONEN

Name des Instruments: Situational Judgement Test zur Erfassung des konditionalen Strategiewissens zur Motivationsregulation
Bisheriger Einsatz in Projekten:
PRO-SRL: https://www.kompetenzen-im-hochschulsektor.de/pro-srl/
PRO-SRL-EVA: https://www.uni-augsburg.de/de/fakultaet/philsoz/fakultat/psychologie/forschung/pro-srl-eva/
Kontakt / Speicherort:
Prof. Dr. Markus Dresel
Lehrstuhl für Psychologie, Universität Augsburg
E-Mail: markus.dresel@phil.uni-augsburg.de

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Konditionales Strategiewissen zur Motivationsregulation
Theoriemodell: Heuristisches Modell des SRL von Dresel et al. (2015) sowie Vorannahmen zur Passung von spezifischen Strategien zu verschiedenen Situationen (Engelschalk, Steuer & Dresel, 2015, 2016)
Testart: Leistungstest; Situational Judgement Test
Modalität: Paper-Pencil-Test oder Online-Test
Testaufbau:
Der Test besteht aus 8 (Kurzversion: 5) standardisierten Situationsbeschreibungen zu den motivationalen Problemen (geringe Erfolgserwartung zur Bearbeitung der Lernaufgabe oder niedriger subjektiver Wert der Lernaufgabe) vor oder während des Lernens (in den beiden Lernsituationen „Schreiben einer wissenschaftlichen Arbeit" und „Vorbereitung auf eine Prüfung"). Zu jeder Vignette werden verschiedene Strategien zur Regulation der eigenen Motivation präsentiert, deren Eignung zur Lösung des dargebotenen Motivationsproblems die Testteilnehmer*innen beurteilen sollen. Die Beurteilung der Eignung erfolgt anhand von Likert-Skalen von 1 (gar nicht geeignet) bis 6 (völlig geeignet). Aus dem Vergleich mit Expert*innenurteilen zur Eignung der Strategien wird anschließend eine Einschätzung des konditionalen Strategiewissens der Testteilnehmer*innen abgeleitet. Dazu werden paarweise die Strategien, welche die Expert*innen eindeutig als für die gegebene Situation geeignet einschätzten, mit denjenigen Strategien verglichen, die sie eindeutig als ungeeignet beurteilten. Die so resultierenden Paardifferenzen werden zu einem Score zusammengefasst, der das konditionale Strategiewissen der Testteilnehmer*innen abbilden soll.
Itempool: 8 Vignetten mit jeweils 9 Items mit unterschiedlichen Strategien, die auf ihre Eignung zu beurteilen sind.
Testversionen: 1 Originalversion (8 Vignetten), 1 Kurzversion (5 Vignetten)

TESTGÜTEKRITERIEN

Reliabilität: Cronbachs α ≥ .92 (Bäulke, Eckerlein & Dresel, 2018; Steuer, Engelschalk, Eckerlein & Dresel, 2019)
Validität Testinhalt: Als motivationale Problemsituationen wurden Lernsituationen ausgewählt („Schreiben einer wissenschaftlichen Arbeit", „Vorbereitung auf eine Prüfung"), die laut den separaten Expertenbefragungen (Dresel et al., 2015; Steuer et al., 2015) in vielen Studiengängen eine hohe Bedeutung haben und große Anforderungen an die Selbstregulation des Lernens stellen. Innerhalb dieser beiden Lernsituationen wurden gemäß einer 2x2-Systematik motivationale Probleme konstruiert und präsentiert („geringe Erfolgserwartung" vs. „geringer subjektiver Wert" x „Problem des Anfangens in der präaktionalen Phase" vs. „Problem des Weitermachens in der aktionalen Phase") – basierend auf Standardmodellen der Motivationspsychologie und eigenen Forschungen, die darauf verweisen, dass Studierende in ihrer Motivationsregulation genau zwischen diesen Aspekten unterscheiden (Engelschalk et al., 2015, 2016).
Validität Antwortprozesse: keine Angaben
Validität interner Testaufbau: Die Vignetten reflektieren unterschiedliche motivationale Problemsituationen, um durch eine theoretisch begründete und repräsentative Stichprobe an Situationen die Validität des Tests zu sichern. Das mit dem Verfahren gemessene Strategiewissen zur Passung von Strategien zu Situationen ist jedoch als übergreifender Faktor zu verstehen.
Validität Zusammenhänge mit anderen Kompetenzindikatoren: In mehreren Studien (Bäulke et al., 2018; Steuer et al., 2019) fanden sich hypothesenkonforme kovergente Zusammenhänge mit anderen Variablen: Quantität/Umfang der Motivationsregulation (r = .46), Effektivität der Motivationsregulation (r = .36 bis r = .51), Anstrengung im Studium (r = .50), Prokrastination (r = -.33), Intention zum Studienabbruch (r = -.23)
Validität Konsequenzen der Testergebnisse: keine Angaben
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 15 Minuten
Benötigtes Material: Testheft bzw. Onlineplattform zur Testadministration
Besonderheiten: Gruppentestung möglich; Auswertung anhand Paardifferenzen
Praxisbeispiel: Das Instrument wurde in zwei Studien (N=135 und N=131) erfolgreich als Indikator für die Wirksamkeit eines Trainingsansatzes zur Verbesserung der Motivationsregulation eingesetzt. Erfasst wurden Änderungen im Strategiewissen zur Motivationsregulation zwischen einem Prä- und einem Posttest zu Beginn und zum Ende des Semesters. Das Instrument konnte die Veränderungen durch das Training zur Motivationsregulation abbilden und verschiedene abhängige Variablen (wie Anstrengung) vorhersagen. Die Teilnehmer*innen erhielten eine Ergebnisrückmeldung, die sie zur Selbsteinschätzung und Reflexion ihres Motivationsregulationsverhaltens nutzen konnten. Seit der ersten Entwicklung wurde – u. a. durch Feedback von Studierenden – der Test um drei Situationen gekürzt und liegt auch in einer Kurzversion mit fünf Vignetten vor. Das Instrument wurde zudem an drei Universitäten (N = 450) eingesetzt, um Unterschiede im konditionalen Strategiewissen zwischen Studierenden unterschiedlicher Fachsemester und Studiengänge abzubilden.

FACHDOMÄNE

Erprobt für: Mathematik, Lehramt, Wirtschaftswissenschaften, Psychologie
Prinzipiell auch geeignet für: Alle sonstigen Studienfächer
Nicht geeignet für: Nutzung außerhalb der Hochschule

ZIELGRUPPE

Erprobt für: Bachelor- und Masterstudierende unterschiedlicher Semester in Deutschland (Bayern, Hessen) und Österreich (Wien)
Prinzipiell auch geeignet für: Alle sonstigen Studierenden
Nicht geeignet für: Nutzung außerhalb der Hochschule

ALLGEMEINER TESTZWECK

Erprobt für: Erfassung des Kompetenzstands und der Kompetenzentwicklung (z. B. nach Intervention)
Prinzipiell auch geeignet für: Evaluation von Trainingsmaßnahmen
Nicht geeignet für: —

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Quantität und Qualität der Strategienutzung, Noten, Intelligenz
Prinzipiell auch geeignet für: Allgemeine Einschätzung des Motivationsregulationsverhaltens
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für Einzelpersonen im zeitlichen Verlauf und auf Kursebene
Prinzipiell auch geeignet für: Studiengangs- und Hochschulebene, internationale Ebene
Nicht geeignet für: —

EINSATZSZENARIEN

Erprobt für: Studienerfolgsvorhersage, Diagnose bei Lernproblemen
Prinzipiell auch geeignet für: Vorhersage der Studienleistungsentwicklung, Studienabbruchvorhersage
Nicht geeignet für: Zwischenprüfungen, Abschlussprüfungen, Akkreditierung

ERLAUBTE KONSEQUENZEN

Erprobt für: Ermittlung des Interventionsbedarfs, Beurteilung der motivationalen Anteile des Lernverhaltens
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Das Instrument eignet sich zur Erfassung von Strategiewissen zur Motivationsregulation und kann zur einmaligen Testung oder mehrmaligen Messung von Veränderungen, z. B. vor und nach einer Intervention, eingesetzt werden. Das Instrument wird weiter genutzt im Projekt „ProkRASt – Prokrastination als Risikofaktor für den Abbruch des Studiums: Eine motivations- und handlungsregulatorische Perspektive und soll in einem nachfolgenden Transferprojekt genutzt werden". An dem Instrument Interessierte können Kontakt mit den Autor*innen aufnehmen.

WEITERE WICHTIGE ANGABEN

Das Instrument sollte mit zwei weiteren Instrumenten zur Qualität und Quantität des Strategieeinsatzes eingesetzt werden (Engelschalk, Steuer & Dresel, 2017). Diese wurden ebenfalls im Rahmen des Projekts PRO-SRL bzw. PRO-SRL-EVA entwickelt. In Kombination bilden die drei Instrumente die Motivationsregulation von Studierenden ganzheitlich ab.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Steuer, G.; Engelschalk, T.; Eckerlein, N. & Dresel, M. (2019). Assessment and relationships of conditional motivational regulation strategy knowledge as an aspect of undergraduates' self-regulated learning competencies. *Zeitschrift für Pädagogische Psychologie, 33*, 95–104. DOI:10.1024/1010-0652/a000237

Engelschalk, T.; Steuer, G. & Dresel, M. (2016). Effectiveness of motivational regulation: Dependence on specific motivational problems. *Learning and Individual Differences*, *52*, S. 72–78. DOI: 10.1016/j.lindif.2016.10.011

Engelschalk, T.; Steuer, G. & Dresel, M. (2017). Quantity and quality of motivational regulation among university students. *Educational Psychology*, *37*, S. 1154–1170. DOI: 10.1080/01443410.2017.1322177

WEITERFÜHRENDE LITERATUR

Dresel, M., Schmitz, B., Schober, B., Spiel, S., Ziegler, A., Engelschalk, T., Jöstl, G., Klug, J., Roth, A., Wimmer, B., & Steuer, G. (2015). Competencies for successful self-regulated learning in higher education: Structural model and indications drawn from expert interviews. *Studies in Higher Education*, 40, 454–470. DOI:10.1080/03075079.2015.1004236

Bäulke, L.; Eckerlein, N. & Dresel, M. (2018). Interrelations between motivational regulation, procrastination and college dropout intentions. *Unterrichtswissenschaft*, *46*(4), S. 461-479. DOI: 10.1007/s42010-018-0029-5

Engelschalk, T.; Steuer, G. & Dresel, M. (2015). Wie spezifisch regulieren Studierende ihre Motivation bei unterschiedlichen Anlässen? Ergebnisse einer Interviewstudie. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *47*, S. 14–23. DOI: 10.1026/0049-8637/a000120

PRO-SRL-EVA – PRO-SRL Fragebogen: Selbstreguliertes Lernen – Wissen und Anwendung

KURZFASSUNG

Name des Instruments: Self-Regulated Learning – Questionnaire for Action and Knowledge (SRL-QuAK); SRL-QuAKs (Kurzversion)

Fachdomäne: Alle Studiengänge

Erfasste Kompetenzen: Selbstreguliertes Lernen (Wissen und Anwendung)

Zielgruppe: Studierende an Hochschulen

Testart: Situationsspezifischer Fragebogen (7-stufige Likert-Skala)

Modalität: Online-Test

Dauer: 120 Minuten (Langversion), 15 Minuten (Kurzversion)

Testaufbau: Teilskalen: Metakognition, Kognition, Frustration, Langeweile. 366 Items (Langversion), 96 Items (Kurzversion)

Allgemeiner Testzweck: Erfassen von Diskrepanzen zwischen SRL-Wissen und -Anwendung sowie die Gründe für die Nichtanwendung günstiger SRL-Strategien; Erfassen von Kompetenzstand und -entwicklung; wiederholter Test nach 1 Semester; längsschnittliche Erhebung zum Erfassen der langfristigen Kompetenzentwicklung

Konkrete Einsatzszenarien: Überprüfung von Interventionen/Trainings zur Förderung von SRL(-Teilen); Vorhersage von Kurs- und Prüfungsleistungen; Vorhersage des Studienerfolgs, Vorhersage der Leistungsentwicklung, Hochschulvergleiche

Nicht geeignet für: Nutzung außerhalb der Hochschule; internationale Ebene

Hinweise für praktische Anwendung: SRL-QuAKs kann über verschiedene Online-Fragebogenprogramme (unipark, soscisurvey, limesurvey etc.) eingesetzt werden.

Bisheriger Einsatz in Projekten: PRO-SRL, PRO-SRL-EVA

Kontakt / Speicherort:
Univ.-Prof. Dipl.-Psych. Dr. Barbara Schober
Institut für Angewandte Psychologie: Arbeit, Bildung, Wirtschaft
Postanschrift: 1010 Wien, Universitätsstraße 7 (NIG)
E-Mail: barbara.schober@univie.ac.at
Webseite: https://bildung-psy.univie.ac.at/forschung/aktuelle-forschungsthemen/selbstreguliertes-lernen/
Speicherort: Voraussichtlich FDZ des IQB Berlin

ALLGEMEINE INFORMATIONEN

Name des Instruments: Self-Regulated Learning – Questionnaire for Action and Knowledge (SRL-QuAK); SRL-QuAKs (Kurzversion)
Bisheriger Einsatz in Projekten: PRO-SRL, PRO-SRL-EVA
Kontakt / Speicherort:
Univ.-Prof. Dipl.-Psych. Dr. Barbara Schober
Institut für Angewandte Psychologie: Arbeit, Bildung, Wirtschaft
Postanschrift: 1010 Wien, Universitätsstraße 7 (NIG)
E-Mail: barbara.schober@univie.ac.at
Webseite: https://bildung-psy.univie.ac.at/forschung/aktuelle-forschungsthemen/selbstreguliertes-lernen/
Speicherort: Voraussichtlich FDZ des IQB Berlin

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Selbstreguliertes Lernen (Wissen und Anwendung)
Theoriemodell: SRL-Modell von Dresel et al. (2015), basierend auf Zimmermans (2000) zyklischem Modell selbstregulierten Lernens
Testart: Situationsspezifischer Fragebogen (7-stufige Likert-Skala)
Modalität: Online-Test
Testaufbau:
Der Fragebogen besteht aus Skalen zu Metakognition, Kognition, Frustration und Langeweile, die jeweils nochmals in Subskalen mit Fokus auf Wissen und Anwendung aufgeteilt sind. Zu Beginn jeder Skala wird eine kurze Vignette vorgegeben, in der die Studierenden gebeten werden, sich in eine Lernsituation (z.B. den Besuch einer Vorlesung, die Erstellung einer wissenschaftlichen Arbeit oder die Vorbereitung auf eine größere Prüfung) hineinzuversetzen. Es folgt eine kurze Beschreibung der formalen Herausforderungen der Situation (z.B. Aufbereitung einer großen Stoffmenge mit Fokus auf Verständnis). Im nächsten Schritt werden für die Abfrage des deklarativen Strategiewissens mögliche Strategien bzgl. deren Eignung für diese Situation erfragt (z.B. „Ich strukturiere den zu lernenden Stoff.") und auch, inwiefern diese angewendet werden. Für die Messung des prozeduralen Wissens werden nun für die einzelnen Strategien konkrete Umsetzungsmöglichkeiten dargestellt (z.B. „Ich benutze Mind Maps, Poster o.ä., um den Stoff zu visualisieren."). Die Teilnehmer*innen sollen wiederum angeben, ob sie die genannten Strategien für günstig halten und ob sie diese anwenden.
Itempool: 366 Items. Teilskalen: mit jeweils Subskalen, die auf Wissen oder Anwendung fokussieren.
Testversionen: 1 Langversion, 1 Kurzversion (96 Items)

TESTGÜTEKRITERIEN

Reliabilität: Cronbachs α: .78-.94.
Validität Testinhalt: Als relevante Lernsituationen wurden jene ausgewählt, die über viele Studiengänge hinweg eine hohe Bedeutung haben und große Anforderungen an die Selbstregulation des Lernens stellen (z.B. „Schreiben einer wissenschaftlichen Arbeit", „Vorbereitung auf eine größere Prüfung"). Dies wurde durch separate Expert*innenbefragungen in der ersten Projektphase von PRO-SRL sichergestellt (Dresel et al., 2015; Steuer et al., 2015). Auf Basis von Interviews und Vorstudien mit Lehrenden und Studierenden in den ersten Projektphasen (Foerst, Klug, Schober & Spiel, 2017), ergab sich zudem die hohe Relevanz der Differenzierung von Wissen über Strategien und deren tatsächliche Anwendung.
Validität Antwortprozesse: keine Angaben
Validität interner Testaufbau: Die vorgegebenen Vignetten beziehen sich auf SRL-relevante Herausforderungssituationen mit sehr spezifischen Anforderungen. Diese wurden in Vorstudien als solche durch Studierende und Expert*innen identifiziert, ebenso wie die für diese Situationen passenden Strategien. Die Validität ergibt sich daher aus der Identifikation von relevanten SRL-Herausforderungen, den dazu „passenden" sinnvollen Vorgehensweisen und Strategien - durch Literatur und Expert*innen - im Sinne eines Kriteriums und dann in der Folge der Messung von Wissen und Anwendung bei den Studierenden. Es wird dabei von einer Situationsspezifität ausgegangen, die in Summe aber wieder einen Score erlaubt, der die Identifikation und Anwendung von passenden Strategien insgesamt abbildet.
Validität Zusammenhänge mit anderen Kompetenzindikatoren: In mehreren Studien fanden sich hypothesenkonforme konvergente Zusammenhänge mit anderen Variablen: beispielsweise zwischen Kognition Anwendung und Leistung (r = . 24) und Frustration Anwendung und Leistung (r = .29). Innerhalb der Teilskalen zeigten sich konsisten höhere Zusammenhänge mit anderen Variablen (z.B. Fähigkeitsselbstkonzept, Interesse, intrinsische Motivation, Noten im Studium, Organisierten, Kritisch Prüfen).
Validität Konsequenzen der Testergebnisse: keine Angaben
Testfairness: Es zeigten sich keine Geschlechtsunterschiede in den Teilskalen.

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 120 Minuten (Langversion), 15 Minuten (Kurzversion)
Benötigtes Material: PC, Internetzugang
Besonderheiten: SRL-QuAKs kann über verschiedene Online-Fragebogenprogramme (unipark, soscisurvey, limesurvey etc.) eingesetzt werden.
Praxisbeispiel: Die erste Validierungsstudie der Langversion SRL-QuAK wurde 2016 an der Universität Wien in den Studienrichtungen Wirtschaftswissenschaften und Psychologie durchgeführt. Studierende verfügen demnach über ein hohes Wissen bezüglich SRL-Strategien, wenden dieses jedoch kaum an. Ausgehend von den Ergebnissen wurden zwei Interventionsstudien in einem Bachelorarbeitseminar (WS 16/17) und einem Masterarbeitseminar (SS 18) durchgeführt. Ebenso wurde eine effiziente Kurzversion des Fragebogens erstellt (SRL-QuAKs) und im SS 2018 in einer Validierungsstudie bei *N* = 340 Studierenden verschiedener Studienrichtungen an österreichischen Universitäten getestet.

FACHDOMÄNE

Erprobt für: Wirtschaftswissenschaften, Psychologie
Prinzipiell auch geeignet für: Alle Studiengänge (domänenunspezifisch)
Nicht geeignet für: Nutzung außerhalb der Hochschule

ZIELGRUPPE

Erprobt für: Studierende jedes Semesters; Validierung Langversion, *N* = 408 (2016, Universität Wien); Validierung Kurzversion, *N* = 340 (2018, Universitäten in Österreich)
Prinzipiell auch geeignet für: Studierende an Fachhochschulen, Studierende im deutschsprachigen Raum (D-A-CH)
Nicht geeignet für: Nutzung außerhalb der Hochschule

ALLGEMEINER TESTZWECK

Erprobt für: Erfassen von Diskrepanzen zwischen SRL-Wissen und -Anwendung sowie die Gründe für die Nicht-Anwendung günstiger SRL-Strategien; Erfassen von Kompetenzstand und -entwicklung; Testwiederholung nach einem Semester
Prinzipiell auch geeignet für: Längsschnittliche Erhebung zum Erfassen der langfristigen Kompetenzentwicklung
Nicht geeignet für: —

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Intelligenz, Noten, motivationale Orientierung
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für Einzelperson im zeitlichen Verlauf (Anfang und Ende des Semesters), Kursebene.
Prinzipiell auch geeignet für: Hochschulebene, Länder- und Bundesebene, internationale Ebene (D-A-CH)
Nicht geeignet für: Internationale Ebene (nicht deutschsprachiger Raum)

EINSATZSZENARIEN

Erprobt für: Überprüfung von Interventionen/Trainings zur Förderung von SRL(-Teilen); Vorhersage von Kurs- und Prüfungsleistungen
Prinzipiell auch geeignet für: Vorhersage des Studienerfolgs, Vorhersage der Leistungsentwicklung, Hochschulvergleiche
Nicht geeignet für: —

ERLAUBTE KONSEQUENZEN

Erprobt für: —
Prinzipiell auch geeignet für: Beurteilung der Studienleistung
Nicht geeignet für: —

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Denkbar ist der Einsatz des Instruments an Universitäten als Teil von Studierendenevaluationen sowie für Studiengänge, die die Vermittlung von SRL-Kompetenzen curricular verankert haben. Das Instrument eignet sich zudem für die Überprüfung des Erfolgs von SRL-Trainingsstudien. Das Instrument ist auf Anfrage bei den Autor*innen erhältlich.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Dresel, M.; Schmitz, B.; Schober, B.; Spiel, C.; Ziegler, A.; Engelschalk, T.; Jöstl, G.; Klug, J.; Roth, A.; Wimmer, B. & Steuer, G. (2015). Competencies for successful self-regulated learning in higher education: Structural model and indications drawn from expert interviews. *Studies in Higher Education, 40*(3), S. 454–470. DOI:10.1080/03075079.2015.1004236

Foerst, N. M.; Klug, J.; Jöstl, G.; Spiel, C. & Schober, B. (2017). Knowledge vs. Action: Discrepancies in University Students' Knowledge about and Self-Reported Use of Self-Regulated Learning Strategies. *Frontiers in psychology, 8*, S. 1288. DOI:10.3389/fpsyg.2017.01288

Foerst, N. M.; Pfaffel, A.; Klug, J.; Spiel, C. & Schober, B. (2019). SRL to Go? – Promoting SRL via Smartphone-App. *Unterrichtswissenschaft, 47*(3), S. 337-366. DOI:10.1007/s42010-019-00046-7

SOSCIE: Unterrichtsplanung von Geschichtslehrkräften

KURZFASSUNG

Name des Instruments: Testinstrument zur Ermittlung des geschichtsdidaktischen Unterrichtsplanungswissens von angehenden Geschichtslehrer*innen

Fachdomäne: Geschichtsdidaktik

Erfasste Kompetenzen: Geschichtsdidaktisches Wissen zur Unterrichtsplanung

Zielgruppe: Angehende Geschichtslehrkräfte (Studienanfänger*innen im Bachelor bis Absolvent*innen), Referendar*innen; Geschichtslehrkräfte

Testart: Leistungstest (Multiple-Choice-Test, Forced-Choice-Test)

Modalität: Paper-Pencil-Test; Online-Test

Dauer: 60 Minuten

Testaufbau: 165 Items (dichotom) in 4 Inhaltsdimensionen: Ziele & Prinzipien, Phasierung & Strukturierung, Medien, Arbeitsaufträge sowie in 2 kognitiven Anforderungsdimensionen bzw. Wissensarten: Faktual/Konzeptionelles Wissen, Prozedurales Wissen. Zusätzlich werden Studiums- und Sozialdaten (10 Items), Epistemologische Überzeugungen (25 Items, 6-stufige Likert-Skala) und Lerngelegenheiten (60 dichotome Items) erfasst.

Allgemeiner Testzweck: Erfassen des Kompetenzstands und der Kompetenzentwicklung; formatives und summatives Assessment, Evaluation, Small-Scale-Assessment, Large-Scale-Assessment; wiederholte Testungen

Konkrete Einsatzszenarien: Kursvergleiche; Hochschulvergleiche; Output: Studien- und Lehrgestaltung; Lehr-Lern-Tool (Selbstevaluation)

Nicht geeignet für: Schüler*innen; internationale Ebene

Hinweise für praktische Anwendung: —

Bisheriger Einsatz in Projekten: SOSCIE (Future Social Sciences Teacher´s Competencies. Modellierung und Erfassung der fachdidaktischen Kompetenz von künftigen Geschichtslehrern und Geschichtslehrerinnen)

Kontakt / Speicherort:
Prof. Dr. Nicola Brauch
Ruhr-Universität Bochum, Didaktik der Geschichte, GA 5/62
Postanschrift: Universitätsstr. 150, 44801 Bochum
Webseite:
http://www.ruhr-uni-bochum.de/histdidaktik/Forschung/soscie.html
E-Mail: nicola.brauch@ruhr-uni-bochum.de

ALLGEMEINE INFORMATIONEN

Name des Instruments: Testinstrument zur Ermittlung des geschichtsdidaktischen Unterrichtsplanungswissens von angehenden Geschichtslehrer*innen
Bisheriger Einsatz in Projekten: SOSCIE (Future Social Sciences Teacher´s Competencies. Modellierung und Erfassung der fachdidaktischen Kompetenz von künftigen Geschichtslehrern und Geschichtslehrerinnen)
Kontakt / Speicherort:
Prof. Dr. Nicola Brauch
Ruhr-Universität Bochum, Didaktik der Geschichte, GA 5/62
Postanschrift: Universitätsstr. 150, 44801 Bochum
Webseite:
http://www.ruhr-uni-bochum.de/histdidaktik/Forschung/soscie.html
E-Mail: nicola.brauch@ruhr-uni-bochum.de

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Geschichtsdidaktisches Wissen zur Unterrichtsplanung
Theoriemodell: Modell professioneller Handlungskompetenz von Lehrer*innen (Baumert & Kunter, 2006)
Testart: Leistungstest; Multiple-Choice-Test, Forced-Choice-Test
Modalität: Paper-Pencil-Test; Online-Test
Testaufbau:
Itempool: 165 Items (dichotom) in 4 Inhaltsdimensionen: Ziele & Prinzipien, Phasierung & Strukturierung, Medien, Arbeitsaufträge sowie in 2 kognitiven Anforderungsdimensionen bzw. Wissensarten (Faktual/Konzeptionelles Wissen, Prozedurales Wissen). Zusätzlich werden Studiums- und Sozialdaten (10 Items), Epistemologische Überzeugungen (25 Items, 6-stufige Likert-Skala) und Lerngelegenheiten (60 dichotome Items) erfasst. Die detaillierte Struktur s. unten.

TESTGÜTEKRITERIEN

Reliabilität: EAP/PV-Realibilitäten: 0,870 (Gesamtkonstrukt); Wissensbereiche: [1]: 0,711, [2] 0,732, [3] 0,77, [4] 0,739; Wissensarten: [5] 0,777; [6] 0,862
Validität Testinhalt: Systematische Ableitung der Iteminhalte aus geschichtsdidaktischer Theorie zur Unterrichtsplanung; Anpassung der Aufgabenkomplexität an Taxonomie von Anderson & Krathwohl et al. (2014); Expertenworkshops
Validität Antwortprozesse: In CogLabs betrachtet
Validität interner Testaufbau: Die Passung der Daten (Modell-Fit nach Rasch-Skalierung) zeigt signifikant bessere Werte für vier theoretisch angenommene inhaltliche Dimensionen und zwei Wissensarten als für die Modelle mit je einer Dimension.
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Erwartungskonforme kleine bis moderate signifikante Zusammenhänge der Testergebnisse mit Schulnoten und besuchten Lehrangeboten

Validität Konsequenzen der Testergebnisse: Erwartungskonforme bedeutsame Zusammenhänge der Testleistung in inhaltlichen Dimensionen mit der Position der Proband*innen im gesamten Studienverlauf ($F(2,362) = 30,248$, $p < ,001$, $\eta 2 = ,143$). Der Zusammenhang mit B.A.-Noten zeigt moderate prognostische Qualität und die Beurteilungsqualität des Instruments.
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 60 Minuten
Benötigtes Material: Stift und Testheft oder Online-Zugang
Besonderheiten: Die psychometrischen Angaben können statistisch genau aufgeschlüsselt werden.
Praxisbeispiel: Die Pilotierung des Instruments erfolgte im Zeitraum Mai-Juli 2015 an fünf Hochschulen (N = 272 Studierende). Ein systematisches Anwenderfeedback wurde in kognitiven Labs eingeholt. Das überarbeitete Instrument wurde im WS 2015/16 an der Universität Bochum erneut pilotiert (N = 365 Studierende). Untersucht wurde die Wissensausprägung bei Einsteiger*innen im Bachelorstudium im Fach Geschichte in den Integrierten Proseminaren (IPS), bei Einsteiger*innen in der Masterphase des Master of Education („Einführungsseminar") sowie bei fortgeschrittenen Masterstudierenden gegen Ende ihres Studiums („Praxisseminar"). Die Ergebnisse verweisen auf einen signifikanten Wissenszuwachs beim geschichtsdidaktischem Unterrichtsplanungswissen während des Studiums. Auffällig ist die generell schwächere Testleistung im prozeduralen Wissen und dass das Wissen nicht in allen inhaltlichen Dimensionen in gleicher Weise ansteigt.

FACHDOMÄNE

Erprobt für: Geschichtsdidaktik
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

ZIELGRUPPE

Erprobt für: Angehende Geschichtslehrkräfte (Studienanfänger*innen im Bachelor bis Masterabsolvent*innen in integrierten Proseminaren, Einführungsseminaren [M.Ed.] und Praxisseminaren [M.Ed.])
Prinzipiell auch geeignet für: Referendar*innen; Geschichtslehrkräfte
Nicht geeignet für: Schüler*innen

ALLGEMEINER TESTZWECK

Erprobt für: Erfassung des Kompetenzstandes; formatives und summatives Assessment, low-scale Assessment, Evaluation
Prinzipiell auch geeignet für: Erfassung der Kompetenzentwicklung, large-scale Assessment, wiederholte Testungen
Nicht geeignet für: —

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Studiums- und Sozialdaten, epistemologische Überzeugungen, Lerngelegenheiten
Prinzipiell auch geeignet für: Andere Maße der Kompetenzerfassung
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für die Kursebene.
Prinzipiell auch geeignet für: Einzelne Studierende, Referendar*innen, Lehrer*innen, Hochschulebene, Seminare der 2. Phase der Lehramtsausbildung (ZfsL), Bundes- & Länderebene
Nicht geeignet für: —

EINSATZSZENARIEN

Erprobt für: Kursvergleiche; Hochschulvergleiche; Output: Studien- und Lehrgestaltung
Prinzipiell auch geeignet für: Lehr-Lern-Tool (Selbstevaluation)
Nicht geeignet für: —

ERLAUBTE KONSEQUENZEN

Erprobt für: Optimierungen des Lehr-Lernangebots. Die signifikant niedrigere Testleistung im Bereich des prozeduralen Wissens spricht für den Ausbau praktischer Anwendungsmöglichkeiten der theoretisch vermittelten Inhalte sowie den Ausbau des Scaffoldings bei der Lernaufgabenentwicklung.
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Aktuell wird das Instrument in einem Dissertationsprojekt zur Entwicklung des Wissens zur Unterrichtsplanung von angehenden Geschichtslehrkräften (einschließlich Referendar*innen) genutzt und dabei inhaltlich gekürzt. Der Test liegt neben der Papierversion auch in einer Online-Version für verschiedene Probandengruppen (Studierende, Referendar*innen) vor. Der Test ist nach Kontaktaufnahme mit den Autor*innen der Studie (Kontakt s. o.) verfügbar.

WEITERE WICHTIGE ANGABEN

Testaufbau des Instruments

Inhaltliche Dimensionen (Wissensbereiche)	Kognitive Anforderungsdimensionen (Wissensarten)	
	Faktual/Konzeptionelles Wissen (FAK) (gesamt 87 Items)	**Prozedurales Wissen (PR)** (gesamt 78 Items)
Ziele & Prinzipien (insgesamt 54 Items)	50 Items zu: Geschichtsbewusstsein, Geschichtskultur, Gegenwartsbezug, Multiperspektivität, Problemorientierung, Quellenorientierung	4 Items zu: Problemorientierung im Geschichtsunterricht
Phasierung & Strukturierung (insgesamt 41 Items)	20 Items zu: Einstiegsphase, Erarbeitungsphase, Strukturierungskonzepte	21 Items zu: Funktionen verschiedener Einstiege, Zuordnung von Reihentiteln zu Strukturierungskonzepten
Medien (insgesamt 41 Items)	17 Items zu: Textquellen, Bildquellen	24 Items zu: Möglichkeiten des Einsatzes unterschiedlicher Unterrichtsmaterialien in bestimmter Unterrichtsstunde
Arbeitsaufträge (insgesamt 29 Items)	0 Items	29 Items zu: Erschließung von Bildquellen, Arbeitsaufträge zur Förderung von Methoden- oder Urteilskompetenz; Anforderungsbereiche von konkreten Arbeitsaufträgen, Passung von Lernaktivitäten zu Themenformulierung

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Wolf, J., Seiffert, J., Seifert, A., Rothland, M. & Brauch, N. (2018). Das geschichtsdidaktische Planungswissen von angehenden Geschichtslehrer/innen. Entwicklung und Pilotierung eines Testinstruments zur Messung des fachdidaktischen Planungswissens von Lehramtsstudierenden im Fach Geschichte. *Geschichte in Wissenschaft und Unterricht 69*, H. 7/8, S. 373–392.

Wolf, J., Seiffert, J., Seifert, A., Rothland, M. & Brauch, N. (2017). Unterrichtsplanung im Fach Geschichte – Modellierung und Erfassung geschichtsdidaktischen Planungswissens von Lehramtsstudierenden. In Stephan Wernke & Klaus Zierer (Hrsg.), *Die Unterrichtsplanung: Ein in Vergessenheit geratener Kompetenzbereich?! Status Quo und Perspektiven aus Sicht der empirischen Forschung* (S. 62–76). Bad Heilbrunn: Klinkhardt.

WEITERFÜHRENDE LITERATUR

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J. & Wittrock, M. C. (2014). *A taxonomy for learning, teaching, and assessing. A revision of Bloom's.* Harlow, Essex: Pearson.

TEDS-Validierung: P-PID und M-PID als situatives allgemeines pädagogisches Wissen und situatives fachdidaktisches Wissen von Mathematiklehrkräften

KURZFASSUNG

Name des Instruments: Mathematics-related perceptual, interpretation, and decision-making skills (M-PID); Pedagogical perceptual, interpretation, and decision-making skills (P-PID)

Fachdomäne: Mathematik-Lehrkräfte und Mathematik-Lehramtsstudierende

Erfasste Kompetenzen: Situatives pädagogisches Wissen (P-PID), Situatives mathematikdidaktisches Wissen (M-PID)

Zielgruppe: Praktizierende Mathematiklehrkräfte der Sekundarstufe, Mathematik-Lehramtsstudierende im Masterstudium

Testart: Leistungstest; offene Antworten, Bewertungsskala

Modalität: Online-Test mit Videos

Dauer: 60 Minuten

Testaufbau: 42 Items (P-PID), 36 Items (M-PID)

Allgemeiner Testzweck: Erfassen von Kompetenzstand und Kompetenzentwicklung, Evaluation von Fortbildungsmaßnahmen

Konkrete Einsatzszenarien: Gruppen-Vergleiche zur Evaluation

Nicht geeignet für: Fachfremde Lehrkräfte und Nichtlehrkräfte, Einzelfallstudien, Ausleseprozesse

Hinweise für praktische Anwendung: Der Test enthält Videovignetten, die nur einmal angesehen werden können. Neben den automatisch ausgewerteten geschlossenen Items werden die Antworten auf offene Testfragen kodiert (20 % doppelt). Die persönliche Rückmeldung differenziert nach Stärken und Schwächen der Teilnehmer*innen.

Bisheriger Einsatz in Projekten: TEDS-Follow Up, TEDS-Unterricht, TEDS-Validierung, TEDS-East-West

Kontakt / Speicherort:
Prof. Dr. Gabriele Kaiser
Fakultät für Erziehungswissenschaft, Fachbereich 5
Postanschrift: Von-Melle-Park 8, 20146 Hamburg
E-Mail: Gabriele.Kaiser@uni-hamburg.de
Webseiten:
http://www.dr-gabriele-kaiser.de/teds-validierung.html
http://www.dr-gabriele-kaiser.de/research-east-west.html

ALLGEMEINE INFORMATIONEN

Name des Instruments: Mathematics-related perceptual, interpretation, and decision-making skills (M-PID); Pedagogical perceptual, interpretation, and decision-making skills (P-PID)
Bisheriger Einsatz in Projekten: TEDS-Follow Up, TEDS-Unterricht, TEDS-Validierung, TEDS-East-West
Kontakt / Speicherort:
Prof. Dr. Gabriele Kaiser
Fakultät für Erziehungswissenschaft, Fachbereich 5
Postanschrift: Von-Melle-Park 8, 20146 Hamburg
E-Mail: Gabriele.Kaiser@uni-hamburg.de
Webseiten:
http://www.dr-gabriele-kaiser.de/teds-validierung.html
http://www.dr-gabriele-kaiser.de/research-east-west.html

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Situatives pädagogisches Wissen (P-PID), Situatives mathematikdidaktisches Wissen (M-PID)
Theoriemodell: PID-Modell (perception of classroom situations, interpretation, decision making; Blömeke et al., 2015)
Testart: Leistungstest; offene Antworten, Bewertungsskala
Modalität: Online-Test mit Videos
Testaufbau:
Itempool: 42 Items (P-PID) zu den Kategorien Klassenführung, konstruktive Unterstützung und kognitive Aktivierung; 36 Items (M-PID) zum mathematischen und mathematikdidaktischen Wissen

TESTGÜTEKRITERIEN

Reliabilität: Gesamtscore für M-PID: WLE = .67; Gesamtscore für P-PID WLE = .70 (Blömeke et al., 2016 auf Basis der TEDS-FU-Daten); geringe Testlet-Effekte (Blömeke et al., 2015)
Validität Testinhalt: Expertenratings zur Angemessenheit der Videoclips und der offenen Fragen
Validität Antwortprozesse: Zweifaches Expertenrating zur Auswertung der Likert-Skalen (Hoth et al., 2016)
Validität interner Testaufbau: Empirische Trennbarkeit der Skalen Wahrnehmung und Interpretation/Entscheidungen treffen (König et al., 2014; Yang et al., 2018). Die Korrelationen zwischen den einzelnen Kompetenzfacetten wurden überprüft, d. h. M-PID in Zusammenhang mit MCK, MPCK und P-PID in Zusammenhang mit GPK.
Validität Zusammenhänge mit anderen Kompetenzindikatoren: keine Angaben

Validität Konsequenzen der Testergebnisse: M-PID hängt mit fachspezifischer Unterrichtsqualität und kognitiver Aktivierung zusammen (Jentsch et al., under review); Zusammenhang der Kompetenzen der Lehrkräfte mit dem Leistungsfortschritt der Lernenden, mediiert durch Unterrichtsqualität.
Testfairness: DIF-Analysen für Proband*innen mit verschiedenen Professionalisierungsgraden verweisen auf Kulturabhängigkeit im Vergleich ostasiatischer und deutscher Lehrkräfte (Yang et al., 2018).

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 60 Minuten
Benötigtes Material: Computer/Laptop mit Internetverbindung
Besonderheiten: Der Test enthält Videovignetten, die nur einmal angesehen werden können. Neben den automatisch ausgewerteten geschlossenen Items sollten die Antworten auf offene Testfragen kodiert werden (20 % doppelt). Die persönliche Rückmeldung differenziert nach Stärken und Schwächen der Teilnehmer*innen.
Praxisbeispiel: Im Projekt TEDS-Validierung (2016-2019) werden die Instrumente M-PID und P-PID im Rahmen der videobasierten Evaluation von Lehrerkompetenzen eingesetzt. Die Testproband*innen sehen jedes der drei Videos einmal, das jeweils eine Kurzübersicht über eine Mathematikstunde der Sekundarstufe mit ausgewählten spezifischen Szenen zeigt. Dann werden den Proband*innen offene und geschlossene Fragen gestellt. Unter einer mathematikdidaktischen Perspektive beziehen sich die Unterrichtsvideos auf Fehlvorstellungen der Schüler*innen, die wahrzunehmen sind, auf die Interpretation dieser Fehlvorstellungen und die Entwicklung entsprechender Handlungsmöglichkeiten, z. B. als Weiterführung des Unterrichts oder als Feedback für die beteiligten Schüler*innen. Zudem spielen Fragen des mathematischen Umgangs mit mathematischen Problemen eine Rolle (enaktiv-ikonisch-symbolisch) sowie Fragen der curricularen Einbindung der gezeigten unterrichtlichen Handlungen. Auch das kognitive Potenzial von Aufgaben soll analysiert werden.
Unter einer pädagogischen Perspektive werden Fragen des Classroom Managements fokussiert, Formen der Unterrichtsplanung zur Vermeidung von Unterrichtsstörungen, verschiedene Schülerarbeitsformen (kooperative versus konkurrente Arbeit) und deren Weiterentwicklung.

FACHDOMÄNE

Erprobt für: Lehramt Mathematik
Prinzipiell auch geeignet für: —
Nicht geeignet für: Fachfremde Lehrkräfte und Nichtlehrkräfte

ZIELGRUPPE

Erprobt für: Praktizierende Mathematiklehrkräfte der Sekundarstufen
Prinzipiell auch geeignet für: Masterstudierende im Lehramt Mathematik
Nicht geeignet für: Nicht-Zielgruppe

ALLGEMEINER TESTZWECK

Erprobt für: Erfassen des Kompetenzstands
Prinzipiell auch geeignet für: Erfassen der Kompetenzentwicklung, Evaluation von Fortbildungsmaßnahmen
Nicht geeignet für: Andere Testzwecke

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Tests zu CME, GPK, MCK, MPCK; Fehlererkennungstest (M_Speed); Tests zu Überzeugungen (beliefs) der Lehrkräfte; In-vivo-Instrumente zur Erhebung der Unterrichtsqualität durch externe Rater*innen, auch videobasierte Evaluation; Selbsteinschätzungsbogen der Lehrkräfte
Prinzipiell auch geeignet für: Zusätzlicher Einsatz eines Instruments zur Einschätzung der Unterrichtsqualität durch Schüler*innen
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Gruppenebene
Prinzipiell auch geeignet für: Hochschulebene, Länder- und Bundesebene, internationale Ebene
Nicht geeignet für: Einzelfallstudien

EINSATZSZENARIEN

Erprobt für: Die Tests sind international vergleichend einsetzbar, was bereits mit Stichproben von chinesischen Lehrkräften gezeigt wurde.
Prinzipiell auch geeignet für: Gruppen-Vergleiche zur Evaluation von professionstheoretisch orientierten Lehrveranstaltungen; Identifikation der Profile von Lehrkraftgruppen, d. h. Stärken-und-Schwächen-analysen von Gruppen.
Nicht geeignet für: Ausleseprozesse

ERLAUBTE KONSEQUENZEN

Erprobt für: —
Prinzipiell auch geeignet für: Zuverlässigkeit auf Individualebene bisher nicht geprüft
Nicht geeignet für: Auswahlprozesse

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Die Instrumente werden aktuell weiteren Validierungsanalysen unterzogen (z. B. Vorhersage von Schülerleistungen). Internationale Vergleichsstudien mit kulturell vergleichbaren oder kulturell sehr unterschiedlichen Gruppen werden durchgeführt. Unter anderem wurden die Instrumente mit chinesischen Mathematiklehramtsstudierenden in der Masterphase eingesetzt. Es ist auch geplant, sie in der Mathematiklehrerausbildung in Deutschland einzusetzen, voraussichtlich in der Masterphase sowie im Referendariat.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Blömeke, S., Busse, A., Kaiser, G., König, J. & Suhl, U. (2016). The relation between content-specific and general teacher knowledge and skills. *Teaching and Teacher Education*, *56* (May), S. 35-46.

Kaiser, G., Busse, A., Hoth, J., König, J. & Blömeke, S. (2015). About the Complexities of Video-Based Assessments: Theoretical and Methodological Approaches to Overcoming Shortcomings of Research on Teachers' Competence. *International Journal of Science and Mathematics Education*, *13*(3), S. 369-387.

König, J., Blömeke, S., Klein, P., Suhl, U., Busse, A. & Kaiser, G. (2014). Is teachers' general pedagogical knowledge a premise for noticing and interpreting classroom situations? A video-based assessment approach. *Teaching and Teacher Education*, *38*, S. 76-88.

WEITERFÜHRENDE LITERATUR

Blömeke, S., König, J., Suhl, U., Hoth, J. & Döhrmann, M. (2015). Wie situationsbezogen ist die Kompetenz von Lehrkräften? Zur Generalisierbarkeit der Ergebnisse von videobasierten Performanztests. *Zeitschrift für Pädagogik*, *61*(3), S. 310-327.

Hoth, J., Schwarz, B., Kaiser, G., Busse, A., König, J. & Blömeke, S. (2016). Uncovering predictors of disagreement: ensuring the quality of expert ratings. *ZDM Mathematics Education*, *48*(1-2), S. 83-95.

Yang, X., Kaiser, G., König, J. & Blömeke, S. (2018). Professional Noticing of Mathematics Teachers: a Comparative Study between Germany and China. *International Journal of Science and Mathematics Education 17*(5), 943–963. doi.org/10.1007/s10763-018-9907-x

TEDS-Validierung: CME Classroom Management Expertise von Lehrkräften

KURZFASSUNG

Name des Instruments: Classroom Management Expertise (CME)

Fachdomäne: Unterschiedliche Fächer

Erfasste Kompetenzen: Classroom Management Expertise (CME)

Zielgruppe: Bachelor- und Masterstudierende Lehramt, Referendarinnen und Referendare, berufserfahrene Lehrkräfte

Testart: Leistungstest; Multiple-Choice-Test, offene Antworten

Modalität: Online-Test mit Videos

Dauer: 20 Minuten (bei berufstätigen Lehrkräften)

Testaufbau: 24 Testfragen; 4 Video-Clips à 1-2 Minuten mit 19 offenen und 5 geschlossenen Items. Inhaltliche Dimensionen: Gestaltung von Übergängen, Organisation zeitlicher Abläufe, Regelung von Schülerverhalten, Erteilung von Rückmeldungen. Kognitive Prozesse: Genauigkeit der Wahrnehmung und Holistische Wahrnehmung, Interpretation

Allgemeiner Testzweck: Erfassen von Kompetenzniveaus

Konkrete Einsatzszenarien: Vergleiche auf Kursebene und Hochschulebene

Nicht geeignet für: Bisher sind keine konkreten Konsequenzen über die Grundlagenforschung hinaus geprüft worden.

Hinweise für praktische Anwendung: Einsatz als Paper-Pencil-Test bzw. Online-Test. Besonderheiten bei der Kodierung der Antworten bei den offenen Testfragen (20 % sind doppelt zu kodieren). Bisher wird nur ein Gesamtscore zurückgemeldet. Die Rückmeldung von Ergebnissen zu Teilskalen nach inhaltlichen Dimensionen oder kognitiven Prozessen ist aufgrund der knappen Testlänge und geringen Itemzahl nicht möglich.

Bisheriger Einsatz in Projekten: Classroom Management Expertise (CME, König, 2015), LEK-R (König & Kramer, 2016), TEDS-Unterricht, TEDS-Validierung (König et al., in review)

Kontakt / Speicherort:
Prof. Dr. Johannes König
Empirische Schulforschung, quantitative Methoden, Department Erziehungs- und Sozialwissenschaften, Humanwissenschaftliche Fakultät, Universität zu Köln
Webseite: https://www.hf.uni-koeln.de/33049
Projektwebseite CME: https://www.hf.uni-koeln.de/35829
E-Mail: johannes.koenig@uni-koeln.de

ALLGEMEINE INFORMATIONEN

Name des Instruments: Classroom Management Expertise (CME)
Bisheriger Einsatz in Projekten: Classroom Management Expertise (CME, König, 2015), LEK-R (König & Kramer, 2016), TEDS-Unterricht, TEDS-Validierung (König et al., in review)
Kontakt / Speicherort:
Prof. Dr. Johannes König
Empirische Schulforschung, quantitative Methoden, Department Erziehungs- und Sozialwissenschaften, Humanwissenschaftliche Fakultät, Universität zu Köln
Webseite: https://www.hf.uni-koeln.de/33049
Projektwebseite CME: https://www.hf.uni-koeln.de/35829
E-Mail: johannes.koenig@uni-koeln.de

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Classroom Management Expertise (CME)
Theoriemodell: „Kompetenz als Kontinuum"-Modell (Blömeke, Gustafsson & Shavelson, 2015); Facetten „Wahrnehmung" (Perception) und „Interpretation" (Interpretation) (unberücksichtigte Facette: „Entscheidungen treffen" [Decision-Making])
Testart: Leistungstest; Multiple-Choice-Test, offene Antworten
Modalität: Online-Test mit Videos
Testaufbau:
Itempool: 24 Items; 4 Video-Clips zu je 1-2 Minuten mit 19 offenen und 5 geschlossenen Items. Inhaltliche Dimensionen: Gestaltung von Übergängen, Organisation zeitlicher Abläufe, Regelung von Schülerverhalten, Erteilung von Rückmeldungen. Kognitive Prozesse: Genauigkeit der Wahrnehmung und Holistische Wahrnehmung, Interpretation

TESTGÜTEKRITERIEN

Reliabilität: Reliabilität des Gesamtscores für Lehrkräfte: EAP .699, WLE .706, Cronbachs α .700 (König, 2015); angehende Lehrpersonen: EAP .73 (König & Kramer, 2016); Lehramtsstudierende: Cronbachs α .74 (Kramer et al., 2017); Geringe Testlet-Effekte durch die Video-Vignetten (König, 2015; Casale et al., 2017)
Validität Testinhalt: keine Angaben
Validität Antwortprozesse: Statistisch signifikant höhere Korrelation mit der Teilskala Pädagogischen Wissens „Kreieren" als mit den Subskalen „Erinnern", „Verstehen/Analysieren" (König, 2015)
Validität interner Testaufbau: keine Angaben
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Mittelhohe Korrelation mit pädagogischem Wissen (TEDS-M Test, .47***) (König, 2015); statistisch signifikant höhere Korrelation mit der Teilskala pädagogischen Wissens „Klassenführung/Motivierung" als mit den Teilskalen „Strukturierung", „Umgang mit Heterogenität", „Leistungsbeurteilung" (König, 2015)

Validität Konsequenzen der Testergebnisse: Berufstätige Lehrkräfte (durchschnittlich 18 Jahre Lehrerfahrung) schneiden signifikant besser ab als angehende Lehrkräfte (König & Kramer, 2016); CME-Scores bei Referendar*innen sagen Schülerbewertung der Unterrichtsqualität voraus (Allgegenwärtigkeit: $\beta = .47^*$, Regelklarheit: $\beta = .36^*$) (König & Kramer, 2016); CME-Scores bei berufstätigen Mathematiklehrkräften sagen signifikant Leistungszuwachs von Schüler*innen der 7. Klasse zur 8. Klasse voraus (König et al., in review); CME-Scores von Lehrkräften korrelieren negativ mit den Burnout-Skalen (König, 2015) (-.23 für „reduzierte Leistungsfähigkeit" und –.28 für „Depersonalisierung", beide $p \leq .01$) (König & Rothland, 2016). Training in Seminaren zum Thema Klassenführung (video-/transkriptbasiert) fördert den Zuwachs im CME stärker als Seminare zu anderen Themen im Bereich des Unterrichtens (Kramer et al., 2017).
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 20 Minuten (bei berufstätigen Lehrkräften)
Benötigtes Material: Papier, Bleistift, Testbögen, Computer mit Beamer oder Computer/Laptop und Internetverbindung
Besonderheiten: Einsatz als Paper-Pencil-Test oder Online-Test. Besonderheiten bei der Kodierung der Antworten bei den offenen Testfragen (20 % sind doppelt zu kodieren). Bisher wird nur ein Gesamtscore zurückgemeldet. Die Rückmeldung von Ergebnissen zu Teilskalen nach inhaltlichen Dimensionen oder kognitiven Prozessen ist aufgrund der knappen Testlänge und geringen Itemzahl nicht möglich.
Praxisbeispiel: Die Gesamtstichprobe umfasst 744 angehende und berufstätige Lehrkräfte. Das Instrument wurde u. a. in universitären Lehrerbildungsseminaren eingesetzt, um zu prüfen, inwieweit der Erwerb situationsspezifischer Fähigkeiten für die Klassenführung und der Erwerb pädagogischen Wissens bei Lehramtsstudierenden wirksamer durch Unterrichtsvideos als durch Unterrichtstranskripte gefördert werden kann (Kramer et al., 2017). Hierzu wurden Trainings in Form eines video- und eines transkriptgestützten Seminars an der Universität zu Köln umgesetzt, die jeweils durch Seminare ohne spezifische Medienunterstützung desselben Moduls kontrolliert wurden. Alle Gruppen (N = 222 Lehramtsstudierende) nahmen an einer Prä- und Postmessung teil. Studierende der video- und transkriptgestützten Gruppen wiesen im Vergleich zur Kontrollgruppe statistisch signifikante Zuwächse in beiden Testergebnissen mit praktischer Bedeutsamkeit auf. Zwischen den Experimentalgruppen zeigten sich keine Unterschiede im Kompetenzzuwachs. Weiterhin wurde das Instrument bei berufserfahrenen Lehrkräften im Rahmen der CME-Studie eingesetzt (König, 2015); ein positiver Zusammenhang zum pädagogischen Wissen sowie negative Korrelationen zu Burnout-Skalen wurden nachgewiesen (N = 119).

FACHDOMÄNE

Erprobt für: Unterschiedliche Fächer
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

ZIELGRUPPE

Erprobt für: Bachelor- und Masterstudierende Lehramt, Referendarinnen und Referendare, berufserfahrene Lehrkräfte
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

ALLGEMEINER TESTZWECK

Erprobt für: Erfassen von Kompetenzniveaus; Prä-Post-Messungen von Kompetenzveränderung bei Studierenden (über ein Semester)
Prinzipiell auch geeignet für: Evaluation von Trainingsmaßnahmen
Nicht geeignet für: —

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Tests zu pädagogischem, fachlichem und fachdidaktischem Wissen in Mathematik (GPK, MCK, MPCK); situationsspezifische Kompetenzfacetten (M_PID und P_PID); Fehlererkennungstest (M_Speed); Tests zu Überzeugungen (beliefs) der Lehrkräfte; In-vivo-Instrumente zur Erhebung der Unterrichtsqualität durch externe Rater*innen, z. T. auch videobasierte Evaluation; Selbsteinschätzungsbogen der Lehrkräfte sowie Einschätzung der Unterrichtsqualität durch Schüler*innen; Burnout-Skalen; Selbstwirksamkeit
Prinzipiell auch geeignet für: Andere Maße der Kompetenzmessung
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für die Kursebene, Hochschulebene, Landes- und Bundesebene, internationale Ebene (D-A-CH)
Prinzipiell auch geeignet für: —
Nicht geeignet für: Individualdiagnostik, internationale Ebene (nicht deutschsprachiger Raum)

EINSATZSZENARIEN

Erprobt für: Vergleiche auf Kursebene oder Hochschulebene
Prinzipiell auch geeignet für: Bisher sind keine konkreten Einsatzszenarien über die Grundlagenforschung hinaus geprüft worden.
Nicht geeignet für: —

ERLAUBTE KONSEQUENZEN

Erprobt für: —
Prinzipiell auch geeignet für: Bisher sind keine konkreten Konsequenzen über die Grundlagenforschung hinaus geprüft worden.
Nicht geeignet für: —

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Das CME-Instrument wird aktuell in einer Validierungsstudie in den drei deutschsprachigen Ländern Deutschland, Österreich und der Schweiz eingesetzt, um den möglichen Einsatzbereich zu erweitern (König, Kramer et al., 2018).
Ferner wird aktuell eine Erweiterung des CME-Instruments entwickelt, die speziell die bislang unberücksichtigte Facette „Entscheidungen treffen" (Decision-Making) misst.
Zu beiden Nutzungsaspekten werden zeitnah neue Erkenntnisse vorliegen.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Casale, G.; Strauß, S.; Hennemann, T. & König, J. (2016). Wie lässt sich Klassenführungsexpertise messen? Überprüfung eines videobasierten Erhebungsinstruments für Lehrkräfte unter Anwendung der Generalisierbarkeitstheorie. *Empirische Sonderpädagogik, 8*(2), S. 119-139.
König, J. (2015). Measuring Classroom Management Expertise (CME) of Teachers: A Video-Based Assessment Approach and Statistical Results. *Cogent Education, 2*(1), 991178.
König, J.; Kramer, C.; Affolter, B.; Bach, A.; Biedermann, H.; Brühwiler, C.; Greiner, U.; Hollenstein, L. & Katstaller, M. (2018). *CME-Testung in unterschiedlichen Ausbildungskontexten: Ergebnisse aus Messinvarianzanalysen.* [CME-Assessment in different educational contexts: Findings from measurement invariance analyses.] Presentation on the GEBF annual conference, Basel, Switzerland.

WEITERFÜHRENDE LITERATUR

König, J. & Kramer, C. (2016). Teacher professional knowledge and classroom management: On the relation of general pedagogical knowledge (GPK) and classroom management expertise (CME). *ZDM – The International Journal on Mathematics Education, 48*(1), S. 139-151.
König, J. & Rothland, M. (2016). Klassenführungswissen als Ressource der Burnout-Prävention? Zum Nutzen von pädagogisch-psychologischem Wissen im Lehrerberuf. *Unterrichtswissenschaft, 44*(4), S. 425-441.
Kramer, C.; König, J.; Kaiser, G.; Ligtvoet, R. & Blömeke, S. (2017). Der Einsatz von Unterrichtsvideos in der universitären Ausbildung: Zur Wirksamkeit video- und transkriptgestützter Seminare zur Klassenführung auf pädagogisches Wissen und situationsspezifische Fähigkeiten angehender Lehrkräfte. *Zeitschrift für Erziehungswissenschaft, 20* (Supplement 1), S. 137-164.

ValiDiS: Ko-WADiS-Test Wissenschaftliches Denken in Naturwissenschaften

KURZFASSUNG

Name des Instruments: Ko-WADiS-Test

Fachdomäne: Naturwissenschaften; Lehramt Biologie, Chemie, Physik sowie weiterer naturwissenschaftlicher Fächer, z. B. Geologie, Meteorologie; weitere empirisch-wissenschaftliche Fächer, z. B. Psychologie

Erfasste Kompetenzen: Wissenschaftliches Denken/Scientific Reasoning

Zielgruppe: Studierende zwischen Studienbeginn Bachelor und Studienabschluss Master, Abiturient*innen, Referendar*innen, Lehrkräfte im Schuldienst, Promovierende der naturwissenschaftlichen Fachdidaktiken

Testart: Leistungstest; Multiple-Choice-Test im Single-Best-Answer-Format

Modalität: Paper-Pencil-Test

Dauer: 35-45 Minuten (keine Zeitbeschränkung)

Testaufbau: 63 Items; 7 Teilkompetenzen: Naturwissenschaftliche Fragen formulieren, Hypothesen generieren, Untersuchungen planen, Daten auswerten und interpretieren, Zweck naturwissenschaftlicher Modelle bestimmen, Modelle testen, Modelle ändern. Kontexte aus 3 Fachdomänen: Biologie, Chemie, Physik. Multimatrix-Design (Balanced-Incomplete-Blockdesign) mit 9 parallelen Testheftformen mit je 21 Items pro Testheft (7 Items pro Fach, 3 Items pro Teilkompetenz)

Allgemeiner Testzweck: Erfassen von Kompetenzstand und Kompetenzentwicklung; Erfassen der Wirksamkeit von Lerngelegenheiten

Konkrete Einsatzszenarien: Erfassen des Kompetenzstands und der Kompetenzentwicklung, z. B. Jahrgangsvergleich mehrerer Studierendenkohorten; Kursvergleiche und Evaluation von Lehrveranstaltungen

Nicht geeignet für: Individualdiagnostik; nicht empirisch arbeitende Wissenschaftsdisziplinen

Hinweise für praktische Anwendung: Das Verfahren wird im Multimatrix-Design eingesetzt (Balanced-Incomplete-Blockdesign), wodurch designbedingt fehlende Werte entstehen. Es sollte mittels probabilistischen Verfahren ausgewertet werden.

Bisheriger Einsatz in Projekten:
Ko-WADiS:
https://didaktik.physik.fu-berlin.de/projekte/kowadis/index.html
ValiDiS:
https://www2.hu-berlin.de/biologie/biodidaktik/forschung/Validis.html

Kontakt / Speicherort:
Freie Universität Berlin
Didaktik der Biologie
Postanschrift: Schwendenerstraße 1, 14195 Berlin

Humboldt-Universität zu Berlin
Fachdidaktik und Lehr-/Lernforschung Biologie
Postanschrift: Unter den Linden 6, 10099 Berlin
Webseite: https://www2.hu-berlin.de/biologie/biodidaktik/

ALLGEMEINE INFORMATIONEN

Name des Instruments: Ko-WADiS-Test
Bisheriger Einsatz in Projekten:
Ko-WADiS:
https://didaktik.physik.fu-berlin.de/projekte/kowadis/index.html
ValiDiS:
https://www2.hu-berlin.de/biologie/biodidaktik/forschung/Validis.html
Kontakt / Speicherort:
Freie Universität Berlin
Didaktik der Biologie
Postanschrift: Schwendenerstraße 1, 14195 Berlin

Humboldt-Universität zu Berlin
Fachdidaktik und Lehr-/Lernforschung Biologie
Postanschrift: Unter den Linden 6, 10099 Berlin
Webseite: https://www2.hu-berlin.de/biologie/biodidaktik/

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Wissenschaftliches Denken/Scientific Reasoning
Theoriemodell: Wissenschaftliches Denken nach Mayer (2007) und naturwissenschaftliche Modellbildung nach Upmeier zu Belzen und Krüger (2010)
Testart: Leistungstest; Multiple-Choice-Test im Single-Best-Answer-Format
Modalität: Paper-Pencil-Test

Testaufbau:
Itempool: 63 Items; 7 Teilkompetenzen: Naturwissenschaftliche Fragen formulieren, Hypothesen generieren, Untersuchungen planen, Daten auswerten und interpretieren, Zweck naturwissenschaftlicher Modelle bestimmen, Modelle testen, Modelle ändern. Kontexte aus 3 Fachdomänen: Biologie, Chemie, Physik
Testversionen: Multimatrix-Design (Balanced-Incomplete-Blockdesign) mit 9 parallelen Testheftformen mit je 21 Items pro Testheft (7 Items pro Fach, 3 Items pro Teilkompetenz)

TESTGÜTEKRITERIEN

Reliabilität: EAP/PV-Reliabilitäten: .56-.81; Cronbachs α je Testheft: .44-.81
Validität Testinhalt: Die inhaltliche und curriculare Validität der Tests wurden durch Konstruktion nach standardisierten Anleitungen und nach den theoretischen Modellen, Expertenurteilen zu allen Items und Expert*innenratings zu ausgewählten Items belegt.
Validität Antwortprozesse: Die Antwortoptionen zu den Multiple-Choice-Items wurden aus offenen Antworten Studierender und Ergebnissen der Think-Aloud-Studie mit Studierenden bei ausgewählten Items abgeleitet. Eyetracking-Studie mit der Cued-Retrospective Methode anhand einer online-basierten Version.
Validität interner Testaufbau: Die Analysen zur Dimensionalität sprechen für eine eindimensionale Kompetenzstruktur; allerdings sind Fitmaße mehrdimensionaler Modelle nicht wesentlich schlechter als die des eindimensionalen Modells.
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Validitätsanalysen mit bekannten Gruppen wurden für passende Stichproben durchgeführt; theoriekonform positive Zusammenhänge des erfassten Konstrukts zu folgenden anderen Kompetenzindikatoren sind für Teilstichproben gegeben: Kognitive Grundfähigkeit, schlussfolgerndes Denken, komplexes Problemlösen, Lehrerprofessionswissen, Note der Hochschulzugangsberechtigung.
Validität Konsequenzen der Testergebnisse: Kompetenzzuwächse im Regelstudium und bei gezielten Fördermaßnahmen (Interventionen) werden vom Test erfasst.
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 35-45 Minuten (keine Zeitbeschränkung)
Benötigtes Material: Stift, Fragebögen
Besonderheiten: Das Verfahren wird im Multimatrix-Design eingesetzt (Balanced-Incomplete-Blockdesign), wodurch designbedingt fehlende Werte entstehen. Es sollte mittels probabilistischen Verfahren ausgewertet werden.
Praxisbeispiel: Der erste Einsatz der Multiple-Choice-Items fand 2012 an zwei Universitäten statt. Weiterhin wurden Testungen an vier weiteren Universitäten durchgeführt; kleinere Vergleichsstichproben (N < 100) wurden an weiteren Hochschulen erhoben. Von ursprünglich 141 entwickelten Items wurde die Zahl nach psychometrischen und qualitativen Analysen auf 123 Items reduziert. Im weiteren Projektverlauf wurden die 63 Items mit den besten psychometrischen Eigenschaften für ein optimiertes Instrument ausgewählt. Die Items mit biologischen Kontexten wurden zudem in die englische, spanische und griechische Sprache übersetzt und in Australien, Chile, Kanada, Spanien und auf Zypern in Kooperationen eingesetzt.

FACHDOMÄNE

Erprobt für: Lehramt Biologie, Chemie, Physik
Prinzipiell auch geeignet für: Naturwissenschaften, Lehramt weiterer naturwissenschaftlicher Fächer, z. B. Geologie, weitere empirisch-wissenschaftliche Fächer, z. B. Psychologie
Nicht geeignet für: Nicht empirisch arbeitende Wissenschaftsdisziplinen

ZIELGRUPPE

Erprobt für: Bachelor- und Masterstudierende von Studienbeginn bis -abschluss
Prinzipiell auch geeignet für: Abiturient*innen, Referendar*innen, Lehrkräfte im Schuldienst, Promovierende der naturwissenschaftlichen Fachdidaktiken
Nicht geeignet für: —

ALLGEMEINER TESTZWECK

Erprobt für: Erfassen von Kompetenzstand und Kompetenzentwicklung
Prinzipiell auch geeignet für: Erfassen der Wirksamkeit von Lerngelegenheiten
Nicht geeignet für: Individualdiagnostik

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Fachwissenstests, komplexes Problemlösen (Genetics Lab), Videostudie zum Experimentieren, Intelligenz-Struktur-Test I-S-T Screening (Teilskalen verbal, numerisch,figural)
Prinzipiell auch geeignet für: Weitere Maße der Kompetenzmessung
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für Gruppen über 100 Personen, Hochschulebene
Prinzipiell auch geeignet für: Unter Voraussetzungen großer Gruppen auch auf Kursebene einsetzbar, Landes- und Bundesebene, internationale Ebene
Nicht geeignet für: Einzelpersonen (hierzu wird ein computergestützter adaptiver Test mit Ko-WADiS-Items entwickelt).

EINSATZSZENARIEN

Erprobt für: Erfassen des Kompetenzstands und der Kompetenzentwicklung von Studierendenkohorten
Prinzipiell auch geeignet für: Unter Voraussetzungen großer Gruppen Kursvergleiche und Evaluation von Lehrveranstaltungen
Nicht geeignet für: Individualdiagnostik

ERLAUBTE KONSEQUENZEN

Erprobt für: —
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

In der Arbeitsgruppe Physikdidaktik an der Freien Universität Berlin wird derzeit eine computergestützte, adaptive Version des Ko-WADiS-Tests entwickelt. Das Ziel ist es, die Kompetenzmessung bei gleichbleibender Validität noch ökonomischer, reliabler und flexibler, auch zur Individualdiagnostik, zu ermöglichen.
Durch die vorliegenden Übersetzungen und die vergleichenden Untersuchungen zur Validität werden internationale Vergleichsstudien angestoßen.

WEITERE WICHTIGE ANGABEN

Das Instrument wurde mit dem Ziel entwickelt, die Kompetenzentwicklung ganzer Studierendenkohorten reliabel, valide und ökonomisch zu erfassen. Individualdiagnostik ist nicht beabsichtigt und geht mit einem deutlich erhöhten Messfehler einher. Die Reliabilität des Instruments und die Validitätsevidenz stützen die Interpretation der Testwerte als Maß für Kompetenzen im Bereich Scientific Reasoning. Die Längsschnittergebnisse zeigen einen annähernd gleichmäßig linearen Kompetenzzuwachs über das Studium hinweg. Das Ziel einer ökonomischen, hinreichend reliablen und validen Kompetenzmessung im Large-Scale-Einsatz wurde erreicht, und mit dem Instrument kann die Kompetenzentwicklung wissenschaftlicher Denkweisen über ein gesamtes Studium hinweg erfasst werden.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Hartmann, S.; Upmeier zu Belzen, A.; Krüger, D. & Pant, H. A. (2015). Scientific reasoning in higher education: Constructing and evaluating the criterion-related validity of an assessment of preservice science teachers' competencies. *Zeitschrift für Psychologie, 223*, S. 47-53. DOI: 10.1027/2151-2604/a000199

Stiller, J.; Hartmann, S.; Mathesius, S.; Straube, P.; Tiemann, R.; Nordmeier, V. et al. (2016). Assessing scientific reasoning: A comprehensive evaluation of item features that affect item difficulty. *Assessment and Evaluation in Higher Education, 41*, S. 721–732. DOI: 10.1080/02602938.2016.1164830

Krell, M.; Redman, C.; Mathesius, S.; Krüger, D. & van Driel, J. (2018). Assessing Pre-Service Science Teachers' Scientific Reasoning Competencies. *Research in Science Education.* DOI: 10.1007/s11165-018-9780-1

Mathesius, S.; Krell, M; Upmeier zu Belzen, A.; Krüger, D. (2019). Überprüfung eines Tests zum wissenschaftlichen Denken unter Berücksichtigung des Validitätskriteriums relations-to-other-variables. *Zeitschrift für Pädagogik, 65*(4), S. 492-510. DOI: 10.3262/ZP1904492

WEITERFÜHRENDE LITERATUR

Mayer, J. (2007). Erkenntnisgewinnung als wissenschaftliches Problemlösen. In D. Krüger & H. Vogt (Hrsg.), *Theorien in der biologiedidaktischen Forschung*, S. 177-186. Berlin: Springer. DOI: 10.1007/978-3-540-68166-3_16

Upmeier zu Belzen, A. & Krüger, D. (2010). Modellkompetenz im Biologieunterricht. *Zeitschrift für Didaktik der Naturwissenschaften, 16*, S. 41-57.

WiKom-SoWi: Wissenschaftliche Kompetenz in Politikwissenschaften

KURZFASSUNG

Name des Instruments: Test wissenschaftlicher Kompetenzen in Politikwissenschaften

Fachdomäne: Politikwissenschaften

Erfasste Kompetenzen: Wissenschaftliche Kompetenz in Politikwissenschaften

Inhaltsbereiche: Forschungsmethoden, Internationale Beziehungen, Politische Theorie, Vergleichende Regierungslehre

Zielgruppe: Studierende der Politikwissenschaften

Testart: Leistungstest; szenariobasierter Test mit verschiedenen Aufgabenformaten wie Multiple Choice, Wahr/Falsch-Aufgaben, Zuordnungsaufgaben, Freitextaufgaben

Modalität: Paper-Pencil-Test

Dauer: 90 Minuten

Testaufbau: 173 Items in 4 Szenarien zu den Inhaltsbereichen Forschungsmethoden, Internationale Beziehungen, Politische Theorie und vergleichende Regierungslehre.

Allgemeiner Testzweck: Erfassen des Kompetenzstands und der Kompetenzentwicklung von Studierenden der Politikwissenschaften, insbesondere der Fähigkeit, verschiedene kognitive Operationen auf unterschiedliche Inhaltsbereiche anzuwenden.

Konkrete Einsatzszenarien: Erfassung des Kompetenzstands von Studierenden der Politikwissenschaften; Nutzung auch zur formativen und summativen Evaluation bestimmter Lernprozesse möglich.

Nicht geeignet für: —

Hinweise für praktische Anwendung: Derzeit noch keine Aussagen zu psychometrischen Eigenschaften des Tests möglich. Beispielitems und Test sind auf Anfrage bei Testentwicklerinnen und Testentwicklern verfügbar.

Bisheriger Einsatz in Projekten: WiKom-SoWi – Modellierung und Messung wissenschaftlicher Kompetenz in sozialwissenschaftlichen Fächern

Kontakt / Speicherort:
Prof. Dr. Birgit Spinath und Dr. Heike Dietrich
Universität Heidelberg
Webseite:
https://www.psychologie.uni-heidelberg.de/ae/paeps/index.html

Prof. Dr. Robin Stark und Dr. Eric Klopp
Universität des Saarlandes
Webseite: https://www.uni-saarland.de/lehrstuhl/stark.html

ALLGEMEINE INFORMATIONEN

Name des Instruments: Test wissenschaftlicher Kompetenz in Politikwissenschaften
Bisheriger Einsatz in Projekten: WiKom-SoWi – Modellierung und Messung wissenschaftlicher Kompetenz in sozialwissenschaftlichen Fächern
Kontakt / Speicherort:
Prof. Dr. Birgit Spinath und **Dr. Heike Dietrich**
Universität Heidelberg
Webseite:
https://www.psychologie.uni-heidelberg.de/ae/paeps/index.html
Prof. Dr. Robin Stark und **Dr. Eric Klopp**
Universität des Saarlandes
Webseite: https://www.uni-saarland.de/lehrstuhl/stark.html

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Wissenschaftliche Kompetenz in Politikwissenschaften
Theoriemodell: Modell wissenschaftlicher Kompetenzen in sozialwissenschaftlichen Fächern
Testart: Leistungstest; szenariobasierter Test mit verschiedenen Aufgabenformaten wie Multiple Choice, Wahr/Falsch-Aufgaben, Zuordnungsaufgaben, Freitextaufgaben
Modalität: Paper-Pencil-Test
Testaufbau:
Itempool: 173 Items in vier Inhaltsbereichen. Pro Inhaltsbereich ein Szenario mit Einführungstext und anschließenden Aufgaben. Die Szenarien präsentieren Fragen, Probleme und Befunde, wie sie typischerweise vorgefunden werden.
Forschungsmethoden: 4 Aufgaben mit 3 Multiple Choice-Items, 3 wahr-falsch-Items, 28 Zuordnungsitems und 8 Freitextitems
Internationale Beziehungen: 3 Aufgaben mit 8 Multiple Choice-Items, 19 Zuordnungsitems und 2 Freitextitems

Politische Theorie: 5 Aufgaben mit 6 Multiple Choice-Items, 9 wahr-falsch-Items, 28 Zuordnungsitems und 12 Freitextitems
Vergleichende Regierungslehre: 4 Aufgaben mit 6 Multiple Choice-Items, 38 Zuordnungsitems und 3 Freitextitems.

TESTGÜTEKRITERIEN

Noch keine Aussagen zu den psychometrischen Eigenschaften des Tests möglich, da bisher nur Pilotierung mit N = 15 Studierenden der Politikwissenschaften stattgefunden hat.

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 90 Minuten
Benötigtes Material: Testbögen und Stifte
Besonderheiten: Ein Einsatz einzelner Szenarien zum jeweils interessierenden Inhaltsbereich ist möglich.
Praxisbeispiel: Erfassung des Kompetenzstands und der Kompetenzentwicklung von Studierenden der Politikwissenschaften während und nach dem Studium; eine Durchführung in Gruppensettings ist möglich.

FACHDOMÄNE

Erprobt für: Politikwissenschaften
Prinzipiell auch geeignet für: Andere sozialwissenschaftliche Fächer
Nicht geeignet für: Andere Studienfächer

ZIELGRUPPE

Erprobt für: Studierende der Politikwissenschaften
Prinzipiell auch geeignet für: Studierende anderer Fächer (z.B. Sozialwissenschaften), die Veranstaltungen in Politikwissenschaften belegen.
Nicht geeignet für: Studierende anderer Fächer ohne inhaltlichen Bezug zu Soziologie

ALLGEMEINER TESTZWECK

Erprobt für: Erfassen des Kompetenzstands und der Kompetenzentwicklung von Studierenden der Politikwissenschaften, insbesondere der Fähigkeit, verschiedene kognitive Operationen auf unterschiedliche Inhaltsbereiche anzuwenden
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: —
Prinzipiell auch geeignet für: Intelligenz, Studiennoten, Motivation, Interesse, Persönlichkeit sowie andere Maße der Kompetenzerfassung.
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: —
Prinzipiell auch geeignet für: Individual- und Gruppenebene/ Kursebene
Nicht geeignet für: —

EINSATZSZENARIEN

Erprobt für: Erfassung des Kompetenzstands von Studierenden der Politikwissenschaften
Prinzipiell auch geeignet für: Erfassung der Kompetenzentwicklung von Studierenden der Politikwissenschaft; Nutzung auch zur formativen und summativen Evaluation bestimmter Lernprozesse möglich.
Nicht geeignet für: —

ERLAUBTE KONSEQUENZEN

Erprobt für: Nur für Forschungszwecke
Prinzipiell auch geeignet für: Individuelles und Gruppenfeedback
Nicht geeignet für: —

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Das Instrument wurde bisher bei einer kleinen Stichprobe eingesetzt; die Testgüte kann erst nach weiteren Einsätzen bestimmt werden. Beispielitems und Test sind auf Nachfrage und nach Abstimmung des Einsatzzwecks bei den Testentwicklerinnen und Testentwicklern verfügbar.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Dietrich, H.; Zhang, Y.; Klopp, E.; Brünken, R.; Krause, U.-M.; Spinath, F. M.; Stark, R. & Spinath, B. (2015). Scientific Competencies in the Social Sciences. *Psychology Learning and Teaching*, *14*, S. 115-130.

WiKom-SoWi: Forschungskompetenz in Psychologie

KURZFASSUNG

Name des Instruments: Test forschungsmethodischer Kompetenzen in der Psychologie

Fachdomäne: Psychologie

Erfasste Kompetenzen: Forschungsmethodische Kompetenzen: Elementare Statistik, Experiment und Varianzanalyse, Korrelation und Regression, Faktoranalyse, Psychometrie

Zielgruppe: Studierende der Psychologie

Testart: Leistungstest; szenariobasierter Test mit den Antwortformaten Multiple Choice, Wahr/Falsch-Antworten, Freitext

Modalität: Paper-Pencil-Test

Dauer: 120 Minuten

Testaufbau: 120 Items zu 5 Szenarien zur Erfassung der Kompetenzen in den Bereichen „Elementare Statistik", „Experiment und Varianzanalyse", „Korrelation und Regression", „Faktoranalyse", „Psychometrie"

Allgemeiner Testzweck: Erfassen des Kompetenzstands in psychologischen Forschungsmethoden

Konkrete Einsatzszenarien: Erfassung des Kompetenzstands von Studierenden der Psychologie; Nutzung auch zur formativen und summativen Evaluation bestimmter Lernprozesse möglich.

Nicht geeignet für: —

Hinweise für praktische Anwendung: Die Ergebnisse werden kriteriumsorientiert ausgewertet und können individuell rückgemeldet werden, z. B. in Form von Noten.

Bisheriger Einsatz in Projekten: WiKom-Sowi – Modellierung und Messung wissenschaftlicher Kompetenz in sozialwissenschaftlichen Fächern

Kontakt / Speicherort:
Prof. Dr. Robin Stark und Dr. Eric Klopp
Universität des Saarlandes
Webseite: https://www.uni-saarland.de/lehrstuhl/stark.html

Prof. Dr. Birgit Spinath und Dr. Heike Dietrich
Universität Heidelberg
Webseite:
https://www.psychologie.uni-heidelberg.de/ae/paeps/index.html

ALLGEMEINE INFORMATIONEN

Name des Instruments: Test forschungsmethodischer Kompetenzen in der Psychologie
Bisheriger Einsatz in Projekten: WiKom-Sowi – Modellierung und Messung wissenschaftlicher Kompetenz in sozialwissenschaftlichen Fächern
Kontakt / Speicherort:
Prof. Dr. Robin Stark und Dr. Eric Kloppk
Universität des Saarlandes
Webseite: https://www.uni-saarland.de/lehrstuhl/stark.html

Prof. Dr. Birgit Spinath und Dr. Heike Dietrich
Universität Heidelberg
Webseite:
https://www.psychologie.uni-heidelberg.de/ae/paeps/index.html

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Forschungsmethodische Kompetenzen: Elementare Statistik, Experiment und Varianzanalyse, Korrelation und Regression, Faktoranalyse, Psychometrie
Theoriemodell: Modell wissenschaftlicher Kompetenzen in sozialwissenschaftlichen Fächern nach Dietrich et al. (2015)
Testart: Leistungstest; szenariobasierter Test mit den Antwortformaten Multiple Choice, Wahr/Falsch-Antworten, Freitext
Modalität: Paper-Pencil-Test
Testaufbau:
Itempool: 120 Items zu deklarativem Wissen und zur anwendungsorientierten Erfassung der Kompetenzen zu 5 Szenarien, um ein möglichst authentisches Setting zu gewährleisten. Die Szenarien präsentieren Befunde, wie sie in der Forschungspraxis typischerweise vorgefunden werden.
Elementare Statistik: 20 Items zum Umgang mit statistischen Grundbegriffen (Mittelwert, Streuung, Verteilung etc.)
Experiment und Varianzanalyse: 21 Items zur Auswertung psychologischer Experimente und zur Anwendung varianzanalytischer Verfahren
Korrelation und Regression: 22 Items zur Anwendung von Korrelations- und Regressionsanalysen
Faktoranalyse: 25 Items zur Anwendung und Interpretation von Faktorenanalysen
Psychometrie: 31 Items zur Anwendung psychometrischer Verfahren

TESTGÜTEKRITERIEN

Reliabilität: Kriteriumsbezogene Reliabilität (Subkoviak, 1976) für Elementare Statistik (.81), Experiment und Varianzanalyse (.82), Korrelation und Regression (.78), Faktoranalyse (.87) und Psychometrie (.84)
Validität: keine Angaben
Testfairness: keine Angaben

PRAKTISCHER EINSATZ

ALLGEMEINES
Dauer: 120 Minuten
Benötigtes Material: Testbögen und Stifte
Besonderheiten: Der Test ist kriteriumsorientiert konstruiert und ermöglicht eine individuelle Rückmeldung des Kompetenzstandes, z. B. in Form von Noten.
Praxisbeispiel: Bspw. als Evaluationsinstrument in der Methodenausbildung, Einsatz in Gruppensettings möglich

FACHDOMÄNE
Erprobt für: Psychologie
Prinzipiell auch geeignet für: Andere empirisch arbeitende Fächer
Nicht geeignet für: —

ZIELGRUPPE
Erprobt für: Studierende der Psychologie
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

ALLGEMEINER TESTZWECK
Erprobt für: Erfassung des Kompetenzstands in psychologischen Forschungsmethoden
Prinzipiell auch geeignet für: Erfassung der Kompetenzentwicklung in psychologischen Forschungsmethoden
Nicht geeignet für: —

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN
Erprobt für: Intelligenz, Motivation, Interesse, Persönlichkeit, Studiennoten
Prinzipiell auch geeignet für: Andere Maße der Kompetenzmessung
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN
Erprobt für: Ergebnisse sind statistisch aussagekräftig für Einzelpersonen im zeitlichen Verlauf, Gruppen in verschiedenen Studienabschnitten, Gruppen mit verschiedenen Instruktionsmethoden
Prinzipiell auch geeignet für: —
Nicht geeignet für: internationale Ebene

EINSATZSZENARIEN
Erprobt für: Lehr-Lern-Tool, Kursvergleich (sofern das Curriculum den Inhalten des Tests entspricht)
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

ERLAUBTE KONSEQUENZEN
Erprobt für: Nur für Forschungszwecke
Prinzipiell auch geeignet für: Individuelle und Gruppenfeedbacks
Nicht geeignet für: —

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ
Das Instrument wurde bisher bei einer kleinen Stichprobe eingesetzt; die Testgüte kann erst nach weiteren Einsätzen bestimmt werden. Das Instrument ist auf Anfrage bei den Testentwickler*innen verfügbar.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN
Dietrich, D.; Zhang, Y.; Klopp, E.; Brünken, R.; Krause, U.-M.; Spinath, F. M.; Stark, R. & Spinath, B. (2015). Scientific competencies in the social sciences. *Psychology Learning and Teaching*, *14*(2), S. 115-130.

WEITERFÜHRENDE LITERATUR
Subkoviak, M. (1976). Estimating reliability from a single administration of a criterion-referenced test. *Journal of Educational Measurement*, *13*, S. 256-276.

WiKom-SoWi: Wissenschaftliche Kompetenz in Soziologie

KURZFASSUNG

Name des Instruments: Test wissenschaftlicher Kompetenzen in der Soziologie

Fachdomäne: Soziologie

Erfasste Kompetenzen: Inhaltsbereiche: soziologische Theorie, Forschungsmethoden, Mediennutzung, Wandel der Arbeitswelt, Bildungschancen

Zielgruppe: Studierende der Soziologie

Testart: Leistungstest; szenariobasierter Test mit verschiedenen Aufgabenformaten wie Multiple Choice, Wahr/Falsch-Aufgaben, Zuordnungsaufgaben, Freitextaufgaben

Modalität: Paper-Pencil-Test

Dauer: 90 Minuten

Testaufbau: 191 Items in 5 Szenarien zu den Inhaltsbereichen Soziologische Theorie, Forschungsmethoden, Mediennutzung, Wandel der Arbeitswelt und Bildungschancen.

Allgemeiner Testzweck: Erfassung des Kompetenzstands und der Kompetenzentwicklung von Studierenden der Soziologie, insbesondere der Fähigkeit, verschiedene kognitive Operationen auf unterschiedliche Inhaltsbereiche anzuwenden

Konkrete Einsatzszenarien: Erfassung des Kompetenzstands von Studierenden der Soziologie; Nutzung auch zur formativen und summativen Evaluation bestimmter Lernprozesse möglich.

Nicht geeignet für: —

Hinweise für praktische Anwendung: Derzeit noch keine Aussagen zu psychometrischen Eigenschaften des Tests möglich. Beispielitems und Test sind auf Anfrage bei Testentwicklerinnen und Testentwicklern verfügbar.

Bisheriger Einsatz in Projekten: WiKom-SoWi – Modellierung und Messung wissenschaftlicher Kompetenz in sozialwissenschaftlichen Fächern

Kontakt / Speicherort:
Prof. Dr. Birgit Spinath und Dr. Heike Dietrich
Universität Heidelberg
Webseite:
https://www.psychologie.uni-heidelberg.de/ae/paeps/index.html

Prof. Dr. Robin Stark und Dr. Eric Klopp
Universität des Saarlandes
Webseite: https://www.uni-saarland.de/lehrstuhl/stark.html

ALLGEMEINE INFORMATIONEN

Name des Instruments: Test wissenschaftlicher Kompetenz in der Soziologie
Bisheriger Einsatz in Projekten: WiKom-SoWi – Modellierung und Messung wissenschaftlicher Kompetenz in sozialwissenschaftlichen Fächern
Kontakt / Speicherort:
Prof. Dr. Birgit Spinath und Dr. Heike Dietrich
Universität Heidelberg
Webseite:
https://www.psychologie.uni-heidelberg.de/ae/paeps/index.html

Prof. Dr. Robin Stark und Dr. Eric Klopp
Universität des Saarlandes
Webseite: https://www.uni-saarland.de/lehrstuhl/stark.html

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Wissenschaftliche Kompetenz in der Soziologie
Theoriemodell: Modell wissenschaftlicher Kompetenzen in sozialwissenschaftlichen Fächern (Dietrich et al., 2015a; b)
Testart: Leistungstest; szenariobasierter Test mit verschiedenen Aufgabenformaten wie Multiple Choice, Wahr/Falsch-Aufgaben, Zuordnungsaufgaben, Freitextaufgaben
Modalität: Paper-Pencil-Test
Testaufbau:
Itempool: 191 Items in fünf Inhaltsbereichen. Pro Inhaltsbereich ein Szenario mit Einführungstext und anschließenden Aufgaben. Die Szenarien präsentieren Fragen, Probleme und Befunde, wie sie typischerweise vorgefunden werden.
Soziologische Theorie: 4 Aufgaben mit 1 Multiple Choice-Item, 10 wahr-falsch-Items, 27 Zuordnungsitems und 1 Freitextitem
Forschungsmethoden: 3 Aufgaben mit 2 Multiple Choice-Items, 9 wahr-falsch-Items, 21 Zuordnungsitems und 6 Freitextitems

Mediennutzung: 4 Aufgaben mit 2 Multiple Choice-Items, 20 Zuordnungsitems und 6 Freitextitems
Wandel der Arbeitswelt: 4 Aufgaben mit 2 Multiple Choice-Items, 6 wahr-falsch-Items, 27 Zuordnungsitems und 1 Freitextitem
Bildungschancen: 4 Aufgaben mit 1 Multiple Choice-Item, 3 wahr-falsch-Items, 43 Zuordnungsitems und 3 Freitextitems

TESTGÜTEKRITERIEN

Noch keine Aussagen zu den psychometrischen Eigenschaften des Tests möglich, da bisher nur Pilotierung mit N = 11 Studierenden der Soziologie stattgefunden hat.

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 90 Minuten
Benötigtes Material: Testbögen und Stifte
Besonderheiten: Ein Einsatz einzelner Szenarien zum jeweils interessierenden Inhaltsbereich ist möglich.
Praxisbeispiel: Erfassung des Kompetenzstands und der Kompetenzentwicklung von Studierenden der Soziologie, während und nach dem Studium; eine Durchführung in Gruppensettings ist möglich.

FACHDOMÄNE

Erprobt für: Soziologie
Prinzipiell auch geeignet für: Andere sozialwissenschaftliche Fächer
Nicht geeignet für: Andere Studienfächer

ZIELGRUPPE

Erprobt für: Studierende der Soziologie
Prinzipiell auch geeignet für: Studierende anderer Fächer (z. B. Sozialwissenschaften), die Veranstaltungen in Soziologie belegen.
Nicht geeignet für: Studierende anderer Fächer ohne inhaltlichen Bezug zu Soziologie

ALLGEMEINER TESTZWECK

Erprobt für: Erfassen des Kompetenzstands und der Kompetenzentwicklung von Studierenden der Soziologie, insbesondere der Fähigkeit, verschiedene kognitive Operationen auf unterschiedliche Inhaltsbereiche anzuwenden
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: —
Prinzipiell auch geeignet für: Intelligenz, Studiennoten, Motivation, Interesse, Persönlichkeit sowie andere Maße der Kompetenzerfassung.
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: —
Prinzipiell auch geeignet für: Individual- und Gruppenebene/ Kursebene
Nicht geeignet für: —

EINSATZSZENARIEN

Erprobt für: Erfassung des Kompetenzstands von Studierenden der Soziologie
Prinzipiell auch geeignet für: Erfassung der Kompetenzentwicklung von Studierenden der Soziologie;Nutzung auch zur formativen und summativen Evaluation bestimmter Lernprozesse möglich.
Nicht geeignet für: —

ERLAUBTE KONSEQUENZEN

Erprobt für: Nur für Forschungszwecke
Prinzipiell auch geeignet für: Individuelles und Gruppenfeedback
Nicht geeignet für: —

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Das Instrument wurde bisher bei einer kleinen Stichprobe eingesetzt; die Testgüte kann erst nach weiteren Einsätzen bestimmt werden. Beispielitems und Test sind auf Nachfrage und nach Abstimmung des Einsatzzwecks bei den Testentwicklerinnen und Testentwicklern verfügbar.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Dietrich, H.; Zhang, Y.; Klopp, E.; Brünken, R.; Krause, U.-M.; Spinath, F. M.; Stark, R. & Spinath, B. (2015a). Es gibt nichts Praktischeres als solide wissenschaftliche Kompetenzen – Erwerb und Gebrauch wissenschaftlicher Kompetenzen in der Soziologie. *Zeitschrift für Sozialmanagement*, *13*, S. 49-62.
Dietrich, H.; Zhang, Y.; Klopp, E.; Brünken, R.; Krause, U.-M.; Spinath, F. M.; Stark, R. & Spinath, B. (2015b). Scientific Competencies in the Social Sciences. *Psychology Learning and Teaching*, *14*, S. 115-130.

WiWiKom: Wirtschaftswissenschaftliche Fachkompetenz

KURZFASSUNG

Name des Instruments: WiWiKom-Test

Fachdomäne: Wirtschaftswissenschaften

Erfasste Kompetenz: Wirtschaftswissenschaftliche Fachkompetenz

Zielgruppe: Bachelorstudierende im Verlauf und am Ende des Studiums, Masterstudierende zu Beginn des Studiums

Testart: Leistungstest; Multiple-Choice-Aufgaben, Auswahlaufgaben, Zuordnungsaufgaben, Reihenfolgeaufgaben

Modalität: Paper-Pencil-Test, auch als Online-Version verfügbar

Dauer: 30-40 Minuten, je nach Länge des Testheftes

Testaufbau: 204 Items: 60 Items zu Volkswirtschaftslehre (VWL) umfassen die Subskalen Mikroökonomie (30 Items) und Makroökonomie (30 Items); 144 Items zu Betriebswirtschaftslehre (BWL) umfassen die Subskalen Organisation (40 Items), Finanzierung (24 Items), Marketing (40 Items), Personal (20) und Internes Rechnungswesen (20 Items). Testversionen: Langversion im Testheftdesign: 24-30 Items je Fragebogen sowie soziobiografischer Teil; Einzel(sub)skalen; Kurzversion: 36 Items

Allgemeiner Testzweck: Erfassen des Stands der wirtschaftswissenschaftlichen Kompetenz zu einem beliebigen Zeitpunkt im Studium (einmalige Testung, z. B. mit der Kurzversion); Erfassen der Kompetenzentwicklung über mehrere Zeitpunkte im Laufe des Studiums (im Bachelor bis zum Beginn des Masters in Wirtschaftswissenschaften)

Einsatzszenarien: Formatives Assessment als Lehr-Lern-Tool für Studierende und Dozent*innen; Beurteilen des Kompetenzstands als Ergänzung zu Prüfungen (auch als Self-Assessment); Erkennen von Stärken und Defiziten im Studienverlauf zur Empfehlung von Kursen und Förderangeboten; Beobachtung der Kompetenzentwicklung im Studienverlauf (z. B. jährlicher Zuwachs); Erkennen von Wissensdefiziten durch Dozent*innen und adressatengerechte Gestaltung der Lehre; summatives Assessment am Ende eines Studienabschnitts; weitere Einsatzmöglichkeiten: Nach Absprache mit der Projektgruppe

Nicht geeignet für: Studieneingangs- und Studieneignungsdiagnostik; Einsatz bei ausländischen Studierenden ohne Fachkenntnisse in Deutsch

Hinweise für praktische Anwendung: Die Langversion wird im Booklet-Design mit 43 Fragebogenvarianten (je 24-30 Items) eingesetzt. Die Auswertung der Ergebnisse erfordert entsprechend ein statistisches Vorgehen auf Basis der IRT.

Bisheriger Einsatz in Projekten: WiWiKom: Modellierung und Messung wirtschaftswissenschaftlicher Fachkompetenz bei Studierenden bzw. Hochschulabsolvent*innen; Nationales Bildungspanel (NEPS) „Hochschulstudium und Übergang in den Beruf"; PRO-SRL „Produkt- und Prozessorientierte Modellierung und Erfassung von Kompetenzen zum Selbstregulierten Lernen im Tertiären Bereich"; KomPrü – KomFäh „Kompetenzorientierte Prüfung kommunikativer Fähigkeiten"; Schweizerischer Leading House LINCA „Lehr-Lernprozesse im kaufmännischen Bereich"; Einsatz in weiteren Kooperationsprojekten in Japan, Finnland, Russland und Südkorea.

Kontakt / Speicherort:
Prof. Dr. Olga Zlatkin-Troitschanskaia
Johannes Gutenberg-Universität Mainz
Fachbereich 03: Rechts- und Wirtschaftswissenschaften
Lehrstuhl für Wirtschaftspädagogik
Postanschrift: Jakob-Welder-Weg 9, 55128 Mainz
E-Mail: lstroitschanskaia@uni-mainz.de
Webseite: https://www.wiwi-kompetenz.de

ALLGEMEINE INFORMATIONEN

Name des Instruments: WiWiKom-Test
Bisheriger Einsatz in Projekten: WiWiKom: Modellierung und Messung wirtschaftswissenschaftlicher Fachkompetenz bei Studierenden bzw. Hochschulabsolvent*innen; NEPS-Studie „Hochschulstudium und Übergang in den Beruf"; Projekt PRO-SLR; Projekt KomPrü-KomFäh; Projekt LINCA; weitere Projekte in Japan, Finnland, Russland und Südkorea
Kontakt / Speicherort:
Prof. Dr. Olga Zlatkin-Troitschanskaia
Johannes Gutenberg-Universität Mainz
Fachbereich 03: Rechts- und Wirtschaftswissenschaften
Lehrstuhl für Wirtschaftspädagogik
Postanschrift: Jakob-Welder-Weg 9, 55128 Mainz
E-Mail: lstroitschanskaia@uni-mainz.de
Webseite: https://www.wiwi-kompetenz.de

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Wirtschaftswissenschaftliche Fachkompetenz
Theoriemodell: Das Theoriemodell (Zlatkin-Troitschanskaia et al., 2014) unterscheidet 3 Dimensionen: a) Wissensarten: Propositionales, fallbezogenes und strategisches Fachwissen (Shavelson & Ruiz-Primo, 1999); b) Kognitive Stufen: Erinnern und Verstehen, Anwenden und Analysieren, Kreieren und Beurteilen (Taxonomiestufen nach Anderson & Krathwohl, 2001, speziell für ökonomische Kontexte nach Walstad et al., 2007); c) Inhaltliche Einteilung nach Fachgebieten: VWL: Mikro- und Makroökonomie, BWL: Organisation, Finanzierung, Marketing, Personal, Internes Rechnungswesen
Testart: Leistungstest; Multiple-Choice-Aufgaben, Auswahlaufgaben, Zuordnungsaufgaben, Reihenfolgeaufgaben
Modalität: Paper-Pencil-Test, auch als Online-Version verfügbar

Testaufbau

Itempool: 204 Items: 60 Items zu Volkswirtschaftslehre (VWL) wurden aus der vierten Version des US-amerikanischen Test of Understanding College Economics (TUCE IV) adaptiert und umfassen die Subskalen Mikroökonomie (30 Items) und Makroökonomie (30 Items); 144 Items zu Betriebswirtschaftslehre (BWL) stammen aus der deutschen Adaption des spanischsprachigen Examen General para el Egreso de la Licenciatura (EGEL) mit den Teilskalen Organisation (40 Items), Finanzierung (24 Items), Marketing (40 Items), Personal (20) und Internes Rechnungswesen (20 Items)
Testversionen: Einzelne Skalen: s.o.; Kurzversion: 36 Items aus allen Bereichen; Langversion mit 204 Items im Booklet-Design für mehrmalige Gruppentestung: 43 Fragebogenversionen mit je 24-30 Items und soziobiografischem Teil

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 30-40 Mintuten, je nach der Testheftlänge
Benötigtes Material: Testhefte, Stifte, Taschenrechner (für die Items der Teilskalen Finanzierung und Rechnungswesen); Computer beim Einsatz der Onlineversion
Besonderheiten: Feedback: Als Interpretationshilfe wird ein Online-Tool bereitgestellt. Die Studierenden erstellen beim Ausfüllen des Tests einen individuellen Code und können damit ihre pseudonymisierten Ergebnisse online einsehen unter: https://mars.wiwi.hu-berlin.de/mediawiki/feedback/index.php/Hauptseite
Praxisbeispiel: Items aus dem WiWiKom-Test wurden im Nationalen Bildungspanel (NEPS) sowie im Projekt PRO-SRL (Produkt- und Prozessorientierte Modellierung und Erfassung von Kompetenzen zum Selbstregulierten Lernen im Tertiären Bereich) in Österreich und im Projekt KomPrü – KomFäh (Kompetenzorientierte Prüfung kommunikativer Fähigkeiten) eingesetzt. Darüber hinaus wurde der Test in der Schweiz im Leading House LINCA (Lehr-Lernprozesse im kaufmännischen Bereich) genutzt sowie in einem finnischen Projekt an der Universität Turku und in einem russischen Projekt an der Universität Higher School in Economics. Der gesamte Test bzw. Teiltests sind für weitere Länder adaptiert und erprobt: Schweiz (Deutsch), Österreich (Deutsch), Japan (Japanisch), Finnland (Finnisch und Schwedisch), Südkorea (Koreanisch), Russland (Russisch) sowie in den Herkunftsländern der Tests (USA [Englisch] und Mexiko [Spanisch]).

TESTGÜTEKRITERIEN

Reliabilität: Cronbachs α bei allen Skalen zu allen Messzeitpunkten (am Ende des 1., 2., und 3. Studienjahres in Bachelor): 0.6–0.8
Validität Testinhalt: Analyse von Studienordnungen und Modulhandbüchern von 96 Studiengängen in 64 Hochschulen (Curriculaanalysen) sowie Online-Expertenbefragungen bei 78 Professor*innen und Dozent*innen der Wirtschaftswissenschaften an deutschen Hochschulen: Die Testitems decken ein grundlegendes Kerncurriculum der Wirtschaftswissenschaften in Bachelor repräsentativ ab.
Validität Antwortprozesse: Kognitive Interviews mit Studierenden (N = 120) mit der Methode des direkten lauten Denkens während des Bearbeitungsprozesses sowie Interviews mit Experten*innen (N = 32): Bei der Fragenbearbeitung zeigen sich die beabsichtigten wirtschaftswissenschaftlichen Denkprozesse; Aufgabenbearbeitungsstrategien wie Raten weisen keinen statistisch signifikanten Zusammenhang mit erfolgreichen Aufgabenlösungen auf.
Validität interner Testaufbau: In statistischen Analysen (konfirmatorische Faktoranalyse, IRT-Modellierung) konnte die theoretisch angenommene inhaltliche Mehrdimensionalität bestätigt werden; die betriebs- und volkswirtschaftlichen Fachgebiete sind eigenständige Teildimensionen.
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Erwartungskonforme Zusammenhänge, z. B. wirtschaftsbezogenes Vorwissen (besuchte Lerngelegenheiten im Fach Wirtschaft, Abschließen einer kaufmännisch-verwaltenden Ausbildung) zeigen einen positiven Zusammenhang mit dem Testergebnis; ebenso konnten positive Zusammenhänge mit der erfassten kognitiven Leistungsfähigkeit der Proband*innen (Note der Hochschulzugangsberechtigung; Intelligenz) nachgewiesen werden.
Validität Konsequenzen der Testergebnisse: Im Querschnitt erwartungskonforme Differenzen der Testleistung zwischen den Kohorten der Studierenden verschiedener Fachsemester: Höhere Leistung mit zunehmendem Fachsemester; im Längsschnitt im Schnitt erwartungskonforme Kompetenzverläufe auf Gruppenebene, jedoch auch Kohorten- und Kompositionseffekte; verschiedene, auch nicht erwartungskonforme, Verläufe auf Individualebene.
Testfairness: Nach Messinvarianzanalysen (CFA, IRT) funktionieren fast alle Items für verschiedene Studierendengruppen gleich gut; Ausnahmen bilden dabei wenige Items, die einen mittelgroßen Effekt zwischen Teilgruppen aufweisen: Es zeigen sich Effekte des Geschlechts und Migrationshintergrunds: Weibliche Studierende (M[weiblich] = 12.03, SD = 4.196, n = 3438; M[männlich] = 14.28, SD = 4.289, n = 4125; $t(7561) = -22.954$, $p<0.0001$) und Studierende mit nichtdeutscher Muttersprache (M[andere Muttersprache als Deutsch] = 13.34, SD = 4.335, n = 7294; M[Deutsch als Muttersprache] = 9.05, SD = 3.936, n = 228; $t(244{,}527) = -16.379$, $p<0.0001$) schnitten im Durchschnitt schlechter ab. Es ist noch nicht hinreichend geklärt, ob die Ergebnisse auf Testeigenschaften oder tatsächliche Kompetenzunterschiede zurückzuführen sind.

FACHDOMÄNE

Erprobt für: Wirtschaftswissenschaften und Sozialwissenschaften
Prinzipiell auch geeignet für: Alle Studiengänge mit wirtschaftswissenschaftlichen Fachinhalten
Nicht geeignet für: —

ZIELGRUPPE

Erprobt für: Bachelorstudierende im Verlauf und am Ende des Studiums, Masterstudierende zu Beginn des Studiums

Prinzipiell auch geeignet für: Masterstudierende
Nicht geeignet für: Studieneinsteiger*innen, ausländische Studierende ohne Fachkenntnisse in Deutsch

ALLGEMEINER TESTZWECK

Erprobt für: Erfassen des Stands der wirtschaftswissenschaftlichen Kompetenz zu einem beliebigen Zeitpunkt im Studium (einmalige Testung, z. B. mit der Kurzversion aus 36 Items); Erfassen der Kompetenzentwicklung über mehrere Zeitpunkte im Laufe des Studiums (im Bachelor bis zum Beginn des Masters Wirtschaftswissenschaften)
Prinzipiell auch geeignet für: Self-Assessment
Nicht geeignet für: Studieneingangs- und Eignungsdiagnostik

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Noten aus den besuchten wirtschaftswissenschaftlichen Lehrveranstaltungen im Studium (BWL, VWL; Methoden); soziodemografischer Fragebogen; Intelligenz; weitere psychologische Skalen
Prinzipiell auch geeignet für: Weitere Maße der Kompetenzerfassung
Nicht geeignet für: —

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für die Individual- und Gruppenebene, auch im zeitlichen Verlauf, Hochschulebene, Länder- und Bundesebene, internationale Ebene (mit dem gleichen Instrument und Teilskalen)
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

EINSATZSZENARIEN

Erprobt für: Formatives Assessment als Lehr-Lern-Tool für Studierende und Dozent*innen; Beurteilen des Kompetenzstands als Ergänzung zu Prüfungen; Erkennen von Stärken und Defiziten im Studienverlauf zur Empfehlung von Kursen und Förderangeboten; Beobachtung der Kompetenzentwicklung im Studienverlauf (z. B. jährlicher Zuwachs)
Prinzipiell auch geeignet für: Weitere Einsatzmöglichkeiten nach Absprache mit den Testentwickler*innen
Nicht geeignet für: Studieneingangs- und Eignungsdiagnostik

ERLAUBTE KONSEQUENZEN

Erprobt für: Individual- und Gruppenfeedback zum Kompetenzstand und zur Kompetenzentwicklung
Prinzipiell auch geeignet für: Kursempfehlungen (z. B. Wahl[pflicht]kurse); Evaluationen von Lehr-Lern-Maßnahmen
Nicht geeignet für: Auswahlentscheidungen, Beurteilung der Studienleistung

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Der WiWiKom-Test wird während der Projektlaufzeit weiterhin in den jährlichen Erhebungen sowie in den verschiedenen weiteren Forschungsprojekten national und international eingesetzt: Erfassung der Langzeitentwicklung von Studierendenkompetenzen (im Längsschnitt) und weitere internationale Analysen werden 2020 fortgesetzt.
Der Test ist auf Anfrage für den Einsatz an Hochschulen erhältlich.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Brückner, S., & Pellegrino, J. W. (2016). Integrating the Analysis of Mental Operations into Multilevel Models to Validate an Assessment of Higher Education Students' Competency in Business and Economics. *Journal of Educational Measurement. 53*(3), 293–312.
Zlatkin-Troitschanskaia, O., Schmidt, S., Brückner, S., Förster, M., Yamaoka, M., & Asano, T. (2016). Macroeconomic Knowledge of Higher Education Students in Germany and Japan – A Multilevel Analysis of Contextual and Personal Effects. *Assessment & Evaluation in Higher Education, 41*(5), 787–801.
http://dx.doi.org/10.1080/02602938.2016.1162279
Zlatkin-Troitschanskaia, O., Jitomirski, J., Happ, R., Molerov, D., Schlax, J., Kühling-Thees, C., Förster, M., & Brückner, S. (2019). Validating a Test for Measuring Knowledge and Understanding of Economics Among University Students. *Zeitschrift für Pädagogische Psychologie, 33*(2), 119–133.

WEITERFÜHRENDE LITERATUR

Vidal, R.U. (2013). Measurement of learning outcomes in higher education: The case of Ceneval in Mexiko. In S. Blömeke; O. Zlatkin-Troitschanskaia; C. Kuhn & J. Fege (Hrsg.), *Modeling and Measuring Competencies in Higher Education.* Rotterdam: Sense Publishers.
Walstad, W. B. & Rebeck, K. (2008). The Test of Understanding of College Economics. *American Economic Review, 98*, 547–551.
Zlatkin-Troitschanskaia, O.; Förster, M.; Brückner, S. & Happ, R. (2014). Insights From a German Assessment of Business and Economics Competence. In H. Coates (Hrsg.), *Higher Education Learning Outcomes Assessment: International Perspectives* (S. 175–197). Frankfurt am Main: Peter Lang. http://dx.doi.org/10.3726/978-3-653-04632-8

WiWiSET: Test zur ökonomischen Studieneingangskompetenz

KURZFASSUNG

Name des Instruments: WiWiSET-Test

Fachdomäne: Wirtschaftswissenschaften

Erfasste Kompetenzen: Ökonomisches Denken und Verständnis; studienfachrelevantes Vorwissen

Zielgruppe: Bachelorstudierende zu Beginn des Studiums in den Wirtschaftswissenschaften; Bewerber*innen/Studieninteressierte in Wirtschaftswissenschaften

Testart: Leistungstest im Multiple-Choice-Format

Modalität: Paper-Pencil-Test; auch computergestützt

Dauer: 25 Minuten

Testaufbau: Langversion: 45 Items zu ökonomischen Grundlagen; Kurzversion: 25 Items in 3 Dimensionen: Ökonomische Grundlagen (15 Items), Grundlagen der Mikroökonomie (5 Items), Grundlagen der Makroökonomie (5 Items)

Allgemeiner Testzweck: Erfassen der ökonomischen Studieneingangskompetenz zu Studienbeginn einmalig oder im zeitlichen Verlauf (z. B. vor und nach den Vorbereitungskursen), Erfassen von Faktoren zur Prognose des Studienerfolgs bzw. des Studienabbruchs in den Wirtschaftswissenschaften

Konkrete Einsatzszenarien: Eingangsdiagnostik; Supportdiagnostik in der Studieneingangsphase und Studienorientierungsphase; Einschätzung des studienfachrelevanten Vorwissens, Planung und Organisation der Studieneingangsphase und Studienorientierungsphase (z. B. Empfehlungen von Vorbereitungskursen/Tutorien etc.), Studienwahlentscheidungen, Self-Assessment

Nicht geeignet für: Eignungsdiagnostik, Beurteilung der Studienleistung

Hinweise für praktische Anwendung: Einzel- oder Gruppentestungen möglich. Der Test sollte unter Aufsicht durchgeführt werden. Es existieren zwei Versionen des Fragebogens mit umgedrehter Item-Reihenfolge (in deutscher, englischer und japanischer Sprache); die Ergebnisse werden per Onlinesystem rückgemeldet und von den Testteilnehmer*innen mit einem pseudonymisierten, selbst erstellten Code abgerufen. Die Sprachkompetenzen sollten möglichst testbasiert kontrolliert werden.

Bisheriger Einsatz in Projekten:
WiWiSET: www.wiwi-kompetenz.de/wiwiset-2016-2019/

Kontakt / Speicherort:
Prof. Dr. Hans Anand Pant
Postanschrift: Unter den Linden 6, 10099 Berlin
E-Mail: hansanand.pant@hu-berlin.de
Webseite: https://www.wiwi-kompetenz.de/wiwiset-2016-2019/

ALLGEMEINE INFORMATIONEN

Name des Instruments: WiWiSET-Test
Bisheriger Einsatz in Projekten:
WiWiSET: https://www.wiwi-kompetenz.de/wiwiset-2016-2019/
Kontakt / Speicherort:
Prof. Dr. Hans Anand Pant
Postanschrift: Unter den Linden 6, 10099 Berlin
E-Mail: hansanand.pant@hu-berlin.de

DAS INSTRUMENT

INHALT UND STRUKTUR

Erfasste Kompetenzen: Ökonomisches Denken und Verständnis; studienfachrelevantes Vorwissen
Theoriemodell: Zum theoretischen Rahmenmodell zur Erfassung des wirtschaftswissenschaftlichen Fachwissens s. Zlatkin-Troitschanskaia et al. (2014), das zwischen drei kognitionsbezogenen Stufen und den drei Inhaltsdimensionen (Grundlagen der VWL, Mikro- und Makroökonomie) differenziert.
Testart: Leistungstest; Multiple-Choice-Test
Modalität: Paper-Pencil-Test; auch computergestützt

Testaufbau

Itempool: Langversion: 45 Items zu ökonomischen Grundlagen; Kurzversion: 25 Items in 3 Dimensionen: Ökonomische Grundlagen (15 Items), Mikroökonomie (5 Items), Makroökonomie (5 Items).
Die 10 Items zur Mikro- und Makroökonomie wurden aus der vierten Version des US-amerikanischen Test of Understanding College Economics (TUCE IV) adaptiert. Die 45 Items zu ökonomischen Grundlagen wurden aus der vierten Version des US-amerikanischen Test of Economic Literacy (TEL IV) adaptiert.
Versionen: Zwei Versionen mit umgekehrter Item-Reihenfolge. Der Test ist in deutscher, englischer und japanischer Sprache verfügbar.

TESTGÜTEKRITERIEN

Reliabilität: Psychometrische Anforderungen an Schwierigkeit, Trennschärfen und Reliabilität werden erfüllt. Interne Konsistenz bei repräsentativer Befragung mit N = 3.977 Studieneinsteiger*innen: $\alpha = .75$ (25 Items)

Validität Testinhalt: Lehrbuchanalysen, Curriculaanalysen von 96 wirtschaftswissenschaftlichen Studiengängen und preuniversitären schulischen Lehr-Lernangeboten bundesweit sowie Onlinerating mit 78 Expert*innen bestätigen die inhaltliche und curriculare Testvalidität.
Validität Antwortprozesse: Die theoretischen kognitionsbezogenen Annahmen zu den drei Stufen wurden in kognitiven Interviews mittels der Methode des lauten Denkens mit 20 Studierenden der Wirtschaftswissenschaften bestätigt.
Validität interner Testaufbau: Die Prüfung von Messmodellen zur internen Struktur zeigt theoriekonforme Ergebnisse; die konfirmatorischen Faktoranalysen (CFA) bestätigen die Eindimensionalität des Konstrukts (χ^2 (778) = 1198.295; RMSEA = .019 (90% CI .017-.021); SRMR = .029; CFI = .894, AIC = 72684.160).
Validität Zusammenhänge mit anderen Kompetenzindikatoren: Theoriekonforme Zusammenhänge mit anderen Maßen kognitiver Fähigkeiten wie Schulnoten, fluide Intelligenz (erfasst mittels BEFKI 11+ [Berliner Test zur Erfassung fluider und kristalliner Intelligenz]).
Validität Konsequenzen der Testergebnisse: Der WiWiSET-Testwert vom ersten Erhebungszeitpunkt erklärt einen signifikanten Teil der Varianz der Testperformance zum zweiten Erhebungszeitpunkt (nach einem Studienjahr) auch unter Kontrolle von fluider Intelligenz, Note der Hochschulzugangsberechtigung, Migrationshintergrund, Geschlecht und Alter ($F(6,773) = 132.78$, $p<.001$, Adj $R^2 = .5037$). Die theoretischen Annahmen zur prognostischen Validität des WiWiSET-Tests wurden bestätigt.
Testfairness: DIF-Analysen und Messinvarianz-Analysen deuten auf einen Genderbias und einen Sprachbias hin. Beim Testeinsatz sollten die Sprachkompetenzen (möglichst mit einem Sprachtest) kontrolliert werden.

PRAKTISCHER EINSATZ

ALLGEMEINES

Dauer: 25 Minuten
Benötigtes Material: Testhefte (bzw. Computer)
Besonderheiten: Einzel- oder Gruppentestungen möglich. Der Test sollte unter Aufsicht durchgeführt werden. Der Fragebogen kann in zwei Versionen mit umgedrehter Item-Reihenfolge eingesetzt werden; die Ergebnisse werden per Onlinesystem rückgemeldet und von den Testteilnehmer*innen mit einem pseudonymisierten, selbst erstellten Code abgerufen. Der Test ist in deutscher, englischer und japanischer Sprache verfügbar.
Praxisbeispiel: Der Test wurde zu zwei Erhebungszeitpunkten vor Ort unter kontrollierten Bedingungen an Hochschulen eingesetzt (Beginn des WS 2016/17 an 42 Universitäten und Hochschulen mit N = 3.977 Studienanfänger*innen und Beginn des WS 2017/18 an 22 Universitäten und Hochschulen mit N = 2.832 Studierenden am Ende des 2./Beginn des 3. Fachsemesters). Die Testteilnehmer*innen konnten ihre Ergebnisse mittels Online-Rückmeldesystem einsehen.

FACHDOMÄNE

Erprobt für: Wirtschaftswissenschaften
Prinzipiell auch geeignet für: Alle Fächer mit ökonomischen Fachinhalten
Nicht geeignet für: —

ZIELGRUPPE

Erprobt für: Bachelorstudierende zu Beginn des Studiums in den Wirtschaftswissenschaften
Prinzipiell auch geeignet für: Studieninteressent*innen/ Bewerber*innen an Hochschulen
Nicht geeignet für: Studierende in höheren Fachsemestern

ALLGEMEINER TESTZWECK

Erprobt für: Erfassen von ökonomischer Studieneingangskompetenz zu Studienbeginn; Erfassen von Faktoren zur Prognose des Studienerfolgs bzw. des Studienabbruchs
Prinzipiell auch geeignet für: Erfassen des ökonomischen Fachwissens im zeitlichen Verlauf während der Studieneingangsphase (z. B. vor und nach den Vorbereitungskursen)
Nicht geeignet für: Eignungsdiagnostik, Beurteilung der Studienleistung

GEMEINSAMER EINSATZ MIT WEITEREN VERFAHREN UND FRAGEBÖGEN

Erprobt für: Intelligenz (BEFKI 11+), Note der Hochschulzugangsberechtigung
Prinzipiell auch geeignet für: Alle anderen Tests zur Erfassung von kognitiven und nicht kognitiven Kompetenzen von Studierenden
Nicht geeignet für: Keine Einschränkung

EIGNUNG

VERGLEICHSEBENEN

Erprobt für: Ergebnisse sind statistisch aussagekräftig für Individual- und Gruppenebene (auch im zeitlichen Verlauf); institutionelle Vergleichsebene (FH vs. Uni); Hochschulebene; Landes- und Bundesebene; internationale Ebene
Prinzipiell auch geeignet für: —
Nicht geeignet für: —

EINSATZSZENARIEN

Erprobt für: Eingangsdiagnostik; Einschätzung des studienfachrelevanten Vorwissens; Vorhersage des Studienerfolgs, Lernprozessdiagnostik in der Studieneingangsphase
Prinzipiell auch geeignet für: Supportdiagnostik sowie Planung und Organisation der Studieneingangsphase und Studienorientierungsphase (z. B. Empfehlungen von Vorbereitungskursen/Tutorien etc.), Studien- und Berufsberatung; Studienwahlentscheidungen, Self-Assessment
Nicht geeignet für: Eignungsdiagnostik, Beurteilung der Studienleistung

ERLAUBTE KONSEQUENZEN

Erprobt für: Individual- und Gruppenfeedback zum Kompetenzstand und zur Kompetenzentwicklung
Prinzipiell auch geeignet für: Vorhersage des Studienerfolgs bzw. Misserfolgs, Studienwahlempfehlung, Kursempfehlung (z. B. Vorkurse)
Nicht geeignet für: Auswahlentscheidungen, Zulassung, Beurteilung der Studienleistung

WEITERE ANGABEN

WEITERENTWICKLUNG UND KÜNFTIGER EINSATZ

Der WiWiSET-Test ist auf Anfrage bei dem Projektteam erhältlich.

Haben Sie noch Fragen oder benötigen Sie Hilfe, z.B. beim praktischen Einsatz oder bei der Auswertung? Kontaktieren Sie uns gern!

AUSGEWÄHLTE PROJEKTPUBLIKATIONEN

Kühling-Thees, C., Happ, R., Zlatkin-Troitschanskaia, O. & Pant, H.A. (2020). The Impact of Entry Preconditions on Student Dropout and Subject Change in Business and Economics. In O. Zlatkin-Troitschanskaia, H. A. Pant, M. Toepper & C. Lautenbach (Eds.), *Student Learning in German Higher Education Innovative Measurement Approaches and Research Results*. Wiesbaden: Springer.

Pant, H. A.; Zlatkin-Troitschanskaia, O.; Schipolowski, S. & Förster, M. (2016). WiWiSET – Validierung eines Studieneingangstests in der Fachdomäne Wirtschaftswissenschaften: Eine national und international vergleichende Studie an Universitäten und Fachhochschulen. In O. Zlatkin-Troitschanskaia; H. A. Pant; C. Lautenbach & M. Toepper (Hrsg.), *Kompetenzmodelle und Instrumente der Kompetenzerfassung im Hochschulsektor – Validierungen und methodische Innovationen (KoKoHs)*. (KoKoHs Working Papers, 10, S. 70–72). Berlin & Mainz: Humboldt-Universität & Johannes Gutenberg-Universität.

Zlatkin-Troitschanskaia, O. & Kühling-Thees, C. (2019). Study-related Domain-specific and Generic Competencies of Economics Students: Insights from a German-Japanese Study. *Journal of Asia Pacific Studies (Waseda University, Japan), 35*, 285 – 297.

WEITERFÜHRENDE LITERATUR

Schipolowski, S., Wilhelm, O., & Schroeders, U. (in Druck). *Berliner Test zur Erfassung fluider und kristalliner Intelligenz ab der 11. Jahrgangsstufe (BEFKI 11+)*. Göttingen: Hogrefe.

Walstad, W. B.; Rebeck, K. & Butters, R. B. (2013). The Test of Economic Literacy: Development and results. *Journal of Economic Education, 44*(3), 298–309.

Zlatkin-Troitschanskaia, O.; Förster, M.; Brückner, S. & Happ, R. (2014). Insights From a German Assessment of Business and Economics Competence. In H. Coates (Hrsg.), *Higher Education Learning Outcomes Assessment: International Perspectives* (S. 175–197). Frankfurt am Main: Peter Lang. http://dx.doi.org/10.3726/978-3-653-04632-8

Impressum

KoKoHs Assessment-Portfolio: Testverfahren zur Modellierung und Messung generischer und domänenspezifischer Kompetenzen bei Studierenden und Hochschulabsolventen

HerausgeberInnen
Olga Zlatkin-Troitschanskaia, Hans Anand Pant, Marie-Theres Nagel, Dimitri Molerov, Corinna Lautenbach, Miriam Toepper

Grafik und Layout
axeptDESIGN.de

Druck
PRINZ-DRUCK Print Media GmbH & Co KG

Bildnachweis
U1 und U4: © connel_design/Fotolia

1. Auflage: Januar 2020

Nachdruck und Verwendung in elektronischen Systemen – auch auszugsweise – nur mit vorheriger schriftlicher Genehmigung durch die HerausgeberInnen und AutorInnen

KoKoHs