

Yasemin Z. Türktorun, S. Franziska C. Wenzel, Julia Mordel,
Sonja Scherer & Holger Horz

Allgemein-wissenschaftliche Methodenkompetenzen erfassen und Fehlkon- zepte aufdecken: Entwicklung und Anwendung ei- nes Wissenstests in der Psychologie

Zusammenfassung

Mit dem Ziel in der Studieneingangsphase allgemein-wissenschaftliche Methodenkompetenzen Psychologiestudierender effektiv zu fördern, wurde für eine Einführungsveranstaltung ein IRT-basierter Wissenstest entwickelt. Dieses Instrument ermöglicht es, Fehlkonzepte in den Themenbereichen „wissenschaftliches Vortragen“, „wissenschaftliches Schreiben“ und „wissenschaftliches Arbeiten“ zu Lehrveranstaltungsbeginn zu identifizieren und deren Abbau im Lauf der Veranstaltung zu erfassen. Der pilotierte Wissenstest wurde in einer Prä-Post-Erhebung mittels zwei Testversionen bei 63 Psychologiestudierenden eingesetzt. Neben einer Lösung sollte pro Aufgabe auch eine Sicherheitseinschätzung zur Korrektheit der Antwort abgegeben werden. Studierende zeigten in allen Themenbereichen einen signifikanten Fähigkeitsanstieg, was als Indikator für den Abbau von Fehlkonzepten gilt. Ein signifikanter Anstieg der Sicherheit konnte für die Themenbereiche wissenschaftliches Vortragen und wissenschaftliches Schreiben ermittelt werden. Im Sinne einer intendierten Reduktion von Heterogenität in der Studieneingangsphase zeigte sich außerdem, dass Studierende mit jeweils niedrigen Ausgangswerten für die Fähigkeit sowie Sicherheit einen höheren Zuwachs aufwiesen als Studierende mit höheren Ausgangswerten. Implikationen werden für Studierende und Lehrende diskutiert.

Schlüsselwörter

Wissenstest; Fehlkonzepte; allgemein-wissenschaftliche Methodenkompetenz; Einführungsveranstaltung; Hochschullehre

Assessing general scientific skills and detecting misconceptions: A knowledge test for educators and students in psychology

Abstract

With the aim of promoting general scientific skills of students in the initial phase of studying psychology, an IRT based test was developed. The instrument allows detecting students' level of misconceptions in the areas of "scientific presentations", "scientific writing", and "academic work" in the beginning and throughout the course. The test was administered to 63 psychology students in a pre-post design. They additionally rated their confidence in solving a task correctly. Results showed that students' knowledge increased significantly in all areas, indicating a reduction of misconceptions. Moreover, a significant increase in task solving certainty was found for scientific presentations and writing. Students with low levels of knowledge and certainty were found to have a greater increase in both of them compared to those with higher levels. This is in line with the aim of reducing heterogeneity in the initial phase of studying. Implications for students and university teachers are discussed.

Keywords

knowledge test; misconceptions; general scientific skills; introductory course; higher education

1 Ausgangslage

Neben dem Aufbau von Fachwissen und Fachkompetenzen ist insbesondere die Stärkung von Methodenkompetenzen in der Studieneingangsphase eine entscheidende Basisqualifikation und ein wichtiges Professionalisierungsmerkmal des Bachelorstudiengangs Psychologie. Angelehnt an die verschiedenen Bereiche der Handlungskompetenz (vgl. z.B. Sonntag, 2000) sollen dabei unter dem Begriff „Methodenkompetenzen“ die Gesamtheit der Grundlagen wissenschaftlichen Arbeitens, also sowohl allgemein-wissenschaftliche Methodenkompetenzen als auch Kompetenzen bezüglich empirischer Forschungsmethoden, verstanden und zusammengefasst werden. Im Rahmen der Einführungsveranstaltung des Moduls „PsyBSc 3: Forschendes Studieren - Präsentieren – Moderieren“ für Bachelorstudierende der Psychologie an der Goethe-Universität Frankfurt sollen daher Methodenkompetenzen mit den besonderen Schwerpunkten „wissenschaftliches Vortragen“, „wissenschaftliches Schreiben“ und „wissenschaftliches Arbeiten“ sowohl theoretisch vermittelt als auch praktisch angewandt und eingeübt werden. Ein konkretes Ziel der Veranstaltung besteht darin, mit Beginn der wissenschaftlichen Bildungskarriere die bisher erworbenen Kenntnisse und Fähigkeiten in den genannten Bereichen der allgemein-wissenschaftlichen Methodenkompetenzen professionell auszubauen, zu fördern und zu vertiefen.

Insbesondere in der Studieneingangsphase und im Bereich der allgemein-wissenschaftlichen Methodenkompetenzen greifen Studierende auf sehr heterogenes und teils unvollständiges Vorwissen oder inkorrekte Vorannahmen zurück (Furnham & Hughes, 2014; Hughes, Lyddy & Lambe, 2013). Diese vorhandene Heterogenität beziehungsweise die daraus häufig resultierende Inkompatibilität des Vorwissens mit neuen Lerninhalten kann für den Wissenserwerb in der Hochschullehre eine Herausforderung darstellen. Im Sinne der Reduktion heterogener Eingangsvoraussetzungen während der Studieneingangsphase und um allen Studierenden die Möglichkeit zu geben, die definierten Lernziele einer Veranstaltung zu erreichen, stellt das Identifizieren und die Auflösung dieser inkorrekten Vorannahmen, der sogenannten „Fehlkonzepten“ (engl. misconceptions) eine notwendige Bedingung dar.

Unter Fehlkonzepten versteht man Vorstellungen und Heuristiken, die sich von den gegenwärtig akzeptierten fachdisziplinären und wissenschaftlichen Konzepten unterscheiden und systematische Denkfehler verursachen (Smith III, diSessa & Roschelle, 1994). Damit sind Fehlkonzepte bei Studierenden insbesondere im Bereich der Psychologie relevant, da hier häufig ein naives und intuitives Alltagswissen mit empirisch fundierten Informationen verschwimmt, sodass eine Unterscheidung schwer fällt (Hughes et al, 2013). Als Beispiel sei hier auf die Taxonomie der Lerntypen (visuelle, auditive, motorische und indifferente Typen) verwiesen. Diese wurden bereits Ende des 19. Jahrhunderts von Charcot vorgeschlagen und werden seitdem in vielen pseudowissenschaftlichen Quellen herangezogen, obwohl das Konzept sich längst als empirisch problematisch herausgestellt hat (vgl. Hasselhorn & Gold, 2017). Den Ursprung solcher Fehlkonzepte führen Studierende laut Befragungen hauptsächlich auf Medienkonsum, persönliche Erlebnisse sowie auf die eigene Lese- und Bildungserfahrung zurück (Furnham & Hughes, 2014; Taylor & Kowalski, 2004). Sobald Fehlkonzepte in Form von inkorrektem Vorwissen in den Wissensstrukturen einer

Person bestehen, interferieren sie mit dem weiteren Wissenserwerb, indem neue Informationen im Sinne der selektiven Wahrnehmung falsch interpretiert beziehungsweise ignoriert werden (Taylor & Kowalski, 2004). Dies wiederum kann langfristig betrachtet den Studienerfolg der Studierenden gefährden (Kuhle, Barber & Bristol, 2009).

Forschungsbefunde legen allerdings nahe, dass durch den Besuch von Einführungsveranstaltungen Fehlkonzepte gezielt adressiert und reduziert werden können (Taylor & Kowalski, 2004; Kowalski & Taylor, 2009). Einführungsveranstaltungen können hier die Möglichkeit bieten, Fehlkonzepte zu identifizieren und durch korrekte Informationen zu ersetzen sowie Studierende darin zu bestärken, die Korrektheit wissenschaftlicher Inhalte zu beurteilen (Landau & Bavaria, 2003; Taylor & Kowalski, 2004). Ein größerer Rückgang von Fehlkonzepten ist vor allem dann zu beobachten, wenn kognitiv aktivierende Lehr-Lern-Formate und -Methoden angewandt werden, die eine eigenständige Auseinandersetzung mit Inhalten ermöglichen und diese in einen wissenschaftlichen Kontext einbetten (Amsel, Johnston, Alvarado, Kettering, Rankin & Ward, 2009; Miller, Wozniak, Rust, Miller & Slezak, 1996; Kowalski & Taylor, 2009). So wurden beispielweise in einer Studie von Kowalski und Taylor (2009) drei unterschiedliche Bedingungen verglichen: Bedingung (1) keine Vorlesung und keine Texte (Kontrollgruppe); Bedingung (2) Vorlesung und Standardtexte; Bedingung (3) Vorlesung und Texte, die vertiefende Argumente gegen bestehende Fehlkonzepte thematisierten. Die Ergebnisse zeigten, dass im Unterschied zur Kontrollgruppe, Fehlkonzepte zu psychologischen Phänomenen bei Studierenden dann abnahmen, wenn sie sowohl durch Texte als auch in der Vorlesung adressiert wurden (Bedingung 1 vs. 2). Der Abbau von Fehlkonzepten war signifikant am größten, wenn das inkorrekte Vorwissen der Studierenden in der Vorlesung und durch Textbearbeitung zunächst angesprochen und anschließend durch die explizite Darbietung korrekter Informationen widerlegt und verworfen wurde (Bedingung 2 vs. 3). Darüber hinaus konnte gezeigt werden, dass neben der Inhaltsvermittlung in Vorlesungen vor allem die aktive Bearbeitung von vertiefenden Texten und die Anregung von Diskussionen einen effektiveren Abbau von Fehlkonzepten herbeiführen konnten (Kowalski & Taylor, 2009).

Auf Basis dieser bisherigen Ergebnisse lässt sich demnach ableiten, dass Fehlkonzepte der Studierenden identifiziert und transparent für alle beteiligten kommuniziert werden müssen, um so die kognitiv aktivierende Auseinandersetzung zu ermöglichen und bestehende Fehlkonzepte durch korrekte Informationen zu ersetzen. Dieser erste Schritt der Identifikation und Kommunikation, das heißt, das Aufdecken und Vergegenwärtigen der Fehlkonzepte, ist eine didaktische Herausforderung, die den Einsatz adäquater Instrumente umfasst. Der vorliegenden Untersuchung liegen demnach zwei zentrale Ziele zugrunde:

1. **Entwicklung eines Wissenstests:** Im Rahmen der genannten Lehrveranstaltung sollte für Bachelorstudierende der Psychologie ein Wissenstest entwickelt werden, der auf die Studieneingangsphase zugeschnitten ist und bestehende Fehlkonzepte in den drei Themenbereichen (1) wissenschaftliches Vortragen, (2) wissenschaftliches Schreiben und (3) wissenschaftliches Arbeiten vor und nach Besuch der Einführungsveranstaltung erfasst.

2. **Anwendung des Wissenstests:** Der entwickelte Wissenstest sollte sich einerseits zur Evaluation des Lehrkonzepts nutzen und andererseits als didaktische Intervention und Aktivierung im Sinne des „Testing Effects“ (Dunn, Saville, Baker & Marek, 2013) anwenden lassen, wodurch zwei Vorteile abgeleitet werden können:
 - (a) *Überprüfung der Wirksamkeit des Lehrkonzepts zum Abbau von Fehlkonzepten.* Hierbei gilt es Aussagen über die Veränderung von Fehlkonzepten innerhalb der als heterogen eingeschätzten Gruppe von Erstsemesterstudierenden zu treffen. Zusätzlich sollte überprüft werden, ob neben einem erwarteten Abbau von Fehlkonzepten, auch Veränderungen in den Sicherheitseinschätzungen der eigenen Antworten durch den Besuch der Lehrveranstaltung zu beobachten sind. Letztere gilt in der Literatur als Indikator für die Stärke der Überzeugung zu vorhandenen Fehlkonzepten (z.B. Taylor & Kowalski, 2004).
 - (b) *Ist-Stand-Erfassung für Lernende und Lehrende.* Anhand der erzielten Ergebnisse aus dem Wissenstest soll Studierenden die Ausprägung ihrer Fehlkonzepte individuell zurückgespielt und dadurch ins Bewusstsein gerufen werden (vgl. Boser et al., 2017). Lehrende hingegen sollen einen Überblick zu bestehenden Fehlkonzepten pro Themenbereich in ihrer Studierendengruppe bekommen. Eine solche Bestandsaufnahme soll Lehrenden die adaptive Gestaltung der weiteren Lehrveranstaltung ermöglichen und den Aufbau wissenschaftlich fundierten Fachwissens fördern.

Im Folgenden soll zunächst das Konzept der Lehrveranstaltung skizziert und anschließend der Wissenstest und die vorgeschaltete Testentwicklung sowie Pilotierung vorgestellt werden. Darauf folgend werden die empirischen Ergebnisse aus der Testanwendung berichtet und gewonnene Erkenntnisse diskutiert. Abschließend erfolgt die Ableitung theoretischer sowie praktischer Implikationen für Lernende und Lehrende und die Präsentation von Ideen für weitere die Lehrveranstaltung begleitende Forschung im Sinne evidenzbasierter Lehre.

2 Didaktische Umsetzung

2.1 Beschreibung des Lehrveranstaltungskonzepts

Im Rahmen des beschriebenen Moduls werden an der Goethe-Universität Frankfurt über drei Semester hinweg verschiedene Lehrveranstaltungen angeboten. Hierbei werden die Studierenden vom *Lernen* zum *Lehren* begleitet und zum weiterführenden Selbststudium angeleitet (vgl. auch Boser, Scherer, Kuchta, Wenzel & Horz, 2017; Höhler et al., 2012; Scherer, Boser & Horz, 2014). Dieses Modul unterstützt das übergeordnete Ziel des Bachelor-Studiengangs in Psychologie basale Methodenkenntnisse zu erwerben und diese auf psychologische Handlungsfelder anzuwenden (vgl. UniReport, 2014). Für den vorliegenden Beitrag sollen lediglich diejenigen Lehrveranstaltungen des Moduls betrachtet werden, die im ersten Semester curricular verankert sind. So startet die Veranstaltung mit einer Vorlesungsphase bestehend aus sieben Sitzungen zu allgemein-wissenschaftlichen Methodenkompetenzen und wird von Blockseminaren, online-Selbstlernmodulen und Präsenz-

Workshops abgelöst. Wie oben beschrieben, lassen sich unter allgemein-wissenschaftlichen Methodenkompetenzen folgende drei Themenbereiche fassen:

1. Der Themenbereich *wissenschaftliches Vortragen* beschäftigt sich mit Inhalten zum Aufbau von Vorträgen und greift hierbei Präsentationsstrukturen und -techniken auf. Darüber hinaus werden neben verbalem Ausdruck als offensichtliches Medium in Präsentationen, die Bedeutung von Körpersprache als nicht-sprachliche Signale thematisiert. Als weitere Inhalte dieses Themenbereichs werden die Entstehung von und der professionelle Umgang mit Nervosität und den damit einhergehenden physischen Reaktionen in Stresssituationen sowie die Gestaltung von Visualisierungsmaterialien auf Basis empirisch gesicherter Befunde gelehrt.
2. Im Themenbereich *wissenschaftliches Schreiben* wird die allgemein anerkannte Gliederung wissenschaftlicher Artikel bzw. empirischer Studien und Kriterien qualitativ hochwertiger empirischer Forschungsarbeiten diskutiert. Außerdem werden nationale und internationale Zitationsrichtlinien für das Fach Psychologie behandelt. Recherchestrategien und Datenbanken sowie Merkmale zur Identifikation von wissenschaftlichen und glaubwürdigen Quellen werden thematisiert.
3. Der Themenbereich *wissenschaftliches Arbeiten* umfasst eine Reihe von theoretischen Konzepten und Modellen zu Lesestrategien und (selbstreguliertem) erfolgreichem Lernen, Stress- und Zeitmanagement zur kompetenten Bewältigung von Studienanforderungen, Zielsetzungsstrategien und Umgang mit Zielkonflikten, aber auch kognitive und emotionale Bedingungen der Motivation und ihre Effekte auf die Lernleistung.

Die einzelnen Themenbereiche sind wiederum in verschiedenen Bestandteilen der Lehrveranstaltung wiederzufinden. So wird beispielsweise das *wissenschaftliche Vortragen* zunächst im Rahmen der Vorlesung theoretisch behandelt. Anschließend werden die Planung und didaktische Gestaltung von aktivierenden Vorträgen sowie Vortragstechniken im Rahmen von online-Selbstlernmodulen selbstreguliert vertieft. Daraufhin wird ein aktivierender Vortrag unter Einhaltung von vermittelten Kriterien innerhalb eines Blockseminars mit anschließendem persönlichem Feedback geübt. Wie aus der Vielfalt an Lehrangeboten hervorgeht, liegt der Lehrveranstaltung ein konstruktivistisches Verständnis von Lehr-Lern-Prozessen zugrunde, was einen geeigneten Ausgangspunkt für die Untersuchung von Fehlkonzepten liefert. So deuten Smith III et al. (1994) darauf hin, dass Fehlkonzepte durch vorausgehende Lernerfahrungen im Vorwissen fest verankert sind. Die Aktivierung bestehender kognitiver Wissensstrukturen und die Rekonstruktion neuer Wissensstrukturen sind wiederum eine wichtige Voraussetzung dafür, um fehlerhafte kognitive Repräsentationen zu verändern und eine Basis für neuen Wissenserwerb zu schaffen (z.B. Smith III et al., 1994; Dole & Sinatra, 1998). Auf diese theoretische Annahme stützt sich das Konzept der Lehrveranstaltung: Es wird eine weitestgehend selbstständige und aktive Auseinandersetzung der Studierenden mit den allgemein-methodischen Lerninhalten gefordert, um eigene Wissensstrukturen aufzubauen beziehungsweise neue Informationen in bestehende kognitive Wissensstrukturen – in das Vorwissen – integrieren zu können (vgl. Scherer et al., 2014; Young & Collin, 2004).

2.2 Entwicklung und Pilotierung eines Wissenstests

2.2.1 Vorüberlegungen

Umfangreiche Recherchen und reifliche Überlegungen führten zu der Entscheidung, ein neues Instrument zur Identifikation von Fehlkonzepten zu entwickeln, welches auch deren Abbau zuverlässig erfassen kann. Die Argumente für die Entwicklung eines neuen Instruments lassen sich wie folgt festhalten:

1. Bisherige Untersuchungen, die Fehlkonzepte bei Psychologiestudierenden erfassen, verwenden überwiegend globale Mythen zu psychologischen Inhalten (z.B. Furnham & Hughes, 2014; Kuhle et al., 2009; Landau & Bavaria, 2003). So kommen ausgewählte Aufgaben aus Werken wie „*Test of Common Beliefs*“ (Vaughan, 1977) oder „*50 Great Popular myths of popular psychology*“ (Lilienfeld, Lynn, Ruscio & Beyerstein, 2010) zum Einsatz. Diese Inhalte passen, wenn überhaupt, nur zu sehr breit angelegten allgemeinen Einführungsveranstaltungen in der Psychologie und weniger zur Vermittlung von Methodenkompetenzen. Da die hier vorgestellte Lehrveranstaltung allerdings ausschließlich allgemein-wissenschaftliche Methodenkompetenzen adressiert, sollten hierfür Aufgaben generiert werden, die gezielt auf diese Lehrinhalte abgestimmt sind.
2. In der Regel werden zur Erfassung von Fehlkonzepten dichotome „Wahr-Falsch“-Aussagen verwendet, wobei alle Aussagen mit „falsch“ zu beantworten sind, um die Aufgabe richtig zu lösen (z.B. Furnham & Hughes, 2014; Taylor & Kowalski, 2004). Diese Art der Fragebogenkonstruktion bringt zentrale Nachteile mit sich, da das angesprochene Antwortmuster (Ablehnung der Aussage als richtige Antwort) für die Versuchspersonen einfach zu durchschauen ist. Um diese Problematik zu umgehen, empfehlen Taylor und Kowalski (2004) die Verwendung von Multiple Choice (MC)-Aufgaben mit mehr als zwei Antwortalternativen. Dieser Implikation zu Folge wurden für die vorliegende Untersuchung sowohl MC-Aufgaben als auch Wahr-Falsch-Aussagen entwickelt, wobei die korrekte Antwort zwischen ‚wahr‘ und ‚falsch‘ variiert wurde.
3. Um nicht nur Aussagen zu anfänglichen Fehlkonzepten der Studierenden treffen zu können, sondern auch deren Entwicklung über den Verlauf der Lehrveranstaltung abbilden zu können, sollte das interessierende Kriterium zu mehreren Zeitpunkten erfasst werden. Betrachtet man geläufige Auswertungsstrategien, die zur Beurteilung der Ausprägung von Fehlkonzepten herangezogen werden, so ist zu beobachten, dass hier der prozentuale Lösungsanteil einer Aufgabe zum ersten sowie zum zweiten Messzeitpunkt ermittelt wird (z.B. Kuhle et al., 2009; Taylor & Kowalski, 2004). Ein höherer Lösungsanteil derselben Aufgabe zum zweiten Messzeitpunkt wird als Indikator für den Abbau von Fehlkonzepten interpretiert. Aus dieser Vorgehensweise ergeben sich zweierlei methodische Probleme: (1) Geht man davon aus, dass es durch den Besuch der Lehrveranstaltung zu einem Wissenszuwachs kommt, so müsste die Beantwortung derselben Aufgaben zum zweiten Messzeitpunkt leichter fallen. Die jeweilige Aufgabenschwierigkeit pro Messzeitpunkt wird allerdings nicht berücksichtigt. (2) Bei Testwiederholung können zudem Erinnerungs- oder Übungeffekte eine Rolle in der Beantwortung von Aufgaben

spielen. Dies würde nicht nur zu einer künstlichen Überschätzung der Test-Reliabilität führen, sondern auch die valide Testwertinterpretation erschweren. Ob ein identisches Antwortverhalten an beiden Messzeitpunkten tatsächlich auf einen Wissenserwerb oder lediglich auf Erinnerungseffekte zurückzuführen ist, wäre also empirisch nicht zu unterscheiden (Schermelleh-Engel & Werner, 2012). Um derart verfälschte Ergebnisse aufgrund der Bearbeitung derselben Aufgaben über mehrere Messzeitpunkte zu vermeiden und zugleich die jeweilige Aufgabenschwierigkeit zu berücksichtigen, sollten die in der vorliegenden Untersuchung individuell bearbeiteten Aufgaben über die Testzeitpunkte variiert werden. Dies wurde durch die Entwicklung und Erprobung der Aufgaben basierend auf der Item-Response-Theorie (IRT; van der Linden, 2016) ermöglicht. Daher wurde das Wissen einer Person zum ersten und zweiten Messzeitpunkt anhand unterschiedlicher Aufgabensets (Testversionen) unter Berücksichtigung der jeweiligen Aufgabenschwierigkeiten bestimmt.

Das interessierende Merkmal „Wissen“ ist eine Personeneigenschaft und wird im Sinne der gängigen Begrifflichkeit in der IRT im Folgenden als „Fähigkeit“ umschrieben und über den Lösungserfolg gegeben der geschätzten Aufgabenschwierigkeit erfasst. Im Folgenden werden die Begriffe Wissen und Fähigkeit synonym verwendet.

2.2.2 Aufgabenentwicklung

Der oben genannten Argumentation folgend, wurden für den zu entwickelnden Wissenstest zur Erfassung der allgemein-wissenschaftlichen Methodenkompetenzen zunächst 53 Aufgaben generiert (vgl. Boser et al., 2017). Ausgangspunkt für die Aufgabenentwicklung war die inhaltliche Ausrichtung der Einführungsveranstaltung in der Psychologie. Diese Inhalte wurden erfahrungsbasiert ausgewählt, so dass sie verbreitete Vorannahmen sowie den vorausgesetzten Wissensstand in Bezug auf wissenschaftliche Methodenkompetenzen in der Studieneingangsphase adressieren. Die Aufgaben wurden demnach so erstellt, dass sie möglichst klar die Vorannahmen in den Bereichen „wissenschaftliches Vortragen“, „wissenschaftliches Schreiben“ und „wissenschaftliches Arbeiten“ operationalisieren und einheitlichen Standards zur Qualitätssicherung bei der Testentwicklung entsprechen (vgl. Osterlind, 2002). Darüber hinaus wurde bereits bei der Aufgabenentwicklung angestrebt, eine möglichst breite Abdeckung der Themenbereiche zu erreichen, indem gezielt Aufgaben unterschiedlicher Schwierigkeit entworfen wurden. Neben MC-Aufgaben mit vier Antwortalternativen wurden auch dichotome Aufgaben eingesetzt, die Aussagen enthalten, die mit „wahr“ oder „falsch“ bewertet werden sollten. In allen konstruierten Aufgaben war immer nur eine Antwortoption korrekt, so dass diese mit einem Punkt bei richtiger Beantwortung bewertet wurden; es wurden keine teilrichtigen Antworten gewertet.

Die Pilotierung der Aufgaben erfolgte an drei Kohorten und über zwei Messzeitpunkte, um eine ausreichend hohe Anzahl an Antworten für jede Aufgabe zu erzielen. Aufgrund der verfügbaren Testzeit von nur 30 Minuten war es nicht möglich, allen Studierenden alle Aufgaben vorzulegen. Daher wurde ein unvollständiges balanciertes Testheftdesign (Frey, Hartig & Rupp, 2009) zur Administration der entwickelten Aufgaben eingesetzt und elf unterschiedliche Testhefte mit jeweils 28 bis 30 Aufgaben zusammengestellt. 286 Personen nahmen insgesamt am ersten und 98 am zweiten Messzeitpunkt der Pilotierung teil. Über

ein Multifacetten-Rasch-Modell wurde kontrolliert, dass die Kohorten sich nicht signifikant im Hinblick auf die Aufgabenbearbeitung unterschieden, was als Voraussetzung für die gemeinsame Analyse der Daten aller Kohorten angenommen werden konnte. Im nächsten Schritt wurden die Aufgaben raschskaliert und hinsichtlich ihrer psychometrischen Güte analysiert. Dabei wurden als Kriterien sowohl die Passung der einzelnen Aufgaben mit dem eindimensionalen Rasch-Modell (Itemfit) als auch die Trennschärfe (Korrelation der Aufgaben mit dem Gesamtergebnis des Tests) betrachtet. Mit Hilfe graphischer Modellgeltungstests wurden unter Nutzung der Daten beider Messzeitpunkte geeignete Aufgaben für die Veränderungsmessung identifiziert. Zur Selektion von Aufgaben für den neuen Wissenstest wurden neben diesen psychometrischen Aspekten auch inhaltliche Kriterien herangezogen. Dabei wurde insbesondere darauf geachtet, dass die drei zu adressierenden Themenbereiche durch die ausgewählten Aufgaben in gleichem Umfang repräsentiert werden. Im Folgenden wird der neue Wissenstest detaillierter vorgestellt.

2.3 Vorstellung des neuen Wissenstests

Der finale Wissenstest besteht aus insgesamt 36 Aufgaben (12 Aufgaben pro Themenbereich) und kann in zwei Testversionen gegliedert werden, die jeweils zu unterschiedlichen Messzeitpunkten eingesetzt werden können, um beispielsweise Veränderung der Fähigkeit über den Verlauf eines Semesters zu messen. Entscheidet man sich für zwei Testversionen, so dienen sechs der Aufgaben als Anker, die in beiden Testversionen identisch sind und so die Verknüpfung der beiden Testversionen ermöglichen. Die Aufgabenanzahl pro Testversion beträgt in diesem Fall 21.

Der gesamte Wissenstest hat eine mittlere Schwierigkeit von 0.01 Logits ($SD = 1.10$) mit einer Range von -2.65 bis 1.77 Logits, wobei negative Werte für leichtere und positive für schwerere Aufgaben sprechen. Mit Blick auf die mittleren Aufgabenschwierigkeiten nach Bereich differenziert zeigt sich, dass die Aufgaben zum Themenbereich „wissenschaftliches Vortragen“ mit einem Mittelwert von -0.69 am leichtesten sind. Die Schwierigkeiten der Aufgaben zum „wissenschaftlichen Schreiben“ liegen im Mittel bei 0.47 und die zum „wissenschaftlichen Arbeiten“ bei 0.26. Die Trennschärfen zu allen Aufgaben liegen im Bereich von .22 bis .57 ($M = .37$, $SD = .08$). Beispielaufgaben pro Themenbereich finden sich in Tabelle 1.

Abbildung 1 stellt die Verteilung der Aufgabenschwierigkeiten des Gesamttests sowie der beiden Testversionen grafisch dar. Die grünen Balken bilden dabei die Anzahl an Aufgaben des jeweiligen Schwierigkeitsbereichs ab, die nur in der ersten Testversion vorgegeben werden. Blaue Balken stehen für Aufgaben die nur in der zweiten Testversion vorgegeben werden. Beide Testversionen werden jeweils durch Anker-Aufgaben (orangefarbene Balken) komplettiert. Beispielsweise kann der Grafik entnommen werden, dass im Schwierigkeitsbereich um -0.5 (zwischen -0.75 und -0.25) zwei Aufgaben in Testversion eins (grün und orange) und vier Aufgaben in Testversion zwei (blau und orange) vorgelegt werden, da jeweils eine dieser Aufgaben als Anker dient. Es wird deutlich, dass sich die Verteilungen der grünen und blauen Balken unterscheiden, da sich die erste Testversion aus etwas leichteren Aufgaben zusammensetzt als Testversion zwei. Dementsprechend liegen die mittleren Aufgabenschwierigkeit der beiden Testversionen bei -0.29 ($Min_1 = -2.65$, $Max_1 = 1.03$, $SD_1 = 1.09$) und 0.24 ($Min_2 = -1.35$, $Max_2 = 1.77$, $SD_2 = 0.92$).

Tab. 1: Beispielaufgaben pro Themenbereich

Themenbereich	Beispielitem
Wissenschaftliches Vortragen	Wenn ich vor meiner Präsentation total aufgeregt bin, sollte ich mich bei meinen Zuhörern da-für nicht entschuldigen, bevor ich anfangе. A <input type="checkbox"/> wahr B <input type="checkbox"/> falsch
Wissenschaftliches Schreiben	Adam ist für seine Hausarbeit zum Thema Depression bei der Recherche von guten Forschungsarbeiten, die er verwenden könnte. Nach anfänglicher Suche merkt er, dass es bereits sehr viele Veröffentlichungen zum Thema Depression gibt und er daher nicht weiß, welche gut bzw. weniger gut sind. Welches Kriterium für gute Forschungsarbeiten können Sie ihm nennen? Gute Forschungsarbeiten sollten... A <input type="checkbox"/> ... nicht mehr als 5 Seiten umfassen. B <input type="checkbox"/> ... in englischer Sprache verfasst sein. C <input type="checkbox"/> ... nicht älter als zwei Jahre sein. D <input type="checkbox"/> ... anerkannte Autoritäten im Feld berücksichtigen.
Wissenschaftliches Arbeiten	Nina bereitet sich auf eine Klausur in Differentieller Psychologie vor. Was ist eine sinnvolle Maßnahme, um Prokrastination vorzubeugen? A <input type="checkbox"/> Nina sollte sich ausreichend Literatur zur Vermeidung von Prokrastination besorgen. B <input type="checkbox"/> Nina sollte Tätigkeiten zusammenfassen und dabei auch unangenehme Aufgaben sofort erledigen. C <input type="checkbox"/> Nina sollte während der Klausurzeit für ausreichend Bewegung und viel Schlaf, sowie für eine gesunde Ernährung sorgen. D <input type="checkbox"/> Nina sollte ihre sozialen Beziehungen für die Lernphase einschränken.

Anmerkungen. Die fett hinterlegten Antwortoptionen sind die korrekten Antworten.

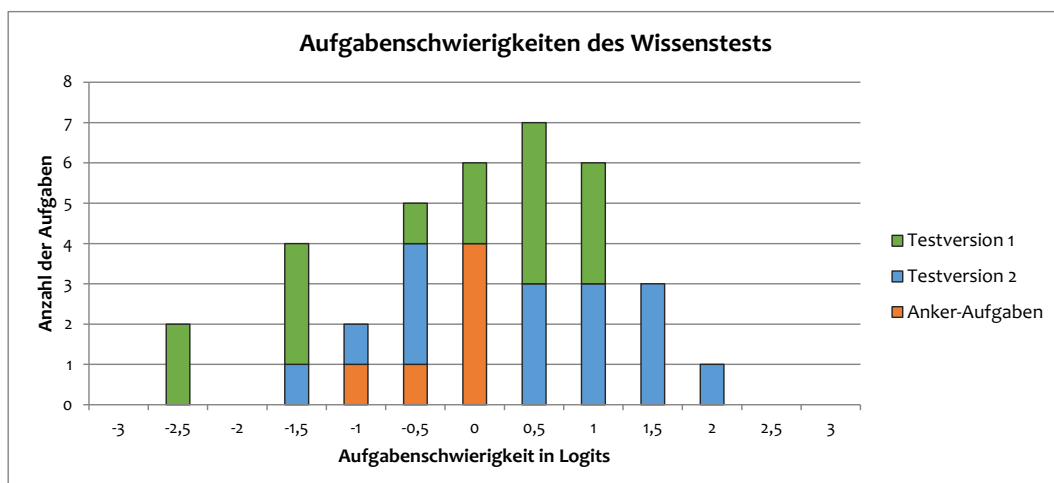


Abb. 1: Schwierigkeit der Aufgaben: Abgetragen ist die Anzahl der Aufgaben im jeweiligen Schwierigkeitsbereich ($\pm 0,25$ Logits).

Der Wissenstest wurde in den vorgestellten Testversionen konzipiert, um zu einem ersten Messzeitpunkt die Fähigkeit der Studierenden vor Beginn der Lehrveranstaltung zu erfassen, wobei hier tendenziell Aufgaben von geringerer Schwierigkeit gewählt wurden. Dagegen wurde zum zweiten Messzeitpunkt die Fähigkeit nach Ende der Lehrveranstaltung erfasst, wobei hier tendenziell schwerere Aufgaben eingesetzt wurden.

3 Anwendung des Wissenstests zur Evaluation des Lehrkonzepts

3.1 Überprüfung der Wirksamkeit des Lehrkonzepts zum Abbau von Fehlkonzepten

Der finale Wissenstest wurde erstmals im Wintersemester 17/18 in der Einführungsveranstaltung des Moduls eingesetzt und sollte in einem ersten Schritt die Wirksamkeit des Lehrkonzepts zum Abbau von Fehlkonzepten überprüfen. Insgesamt haben 63 Erstsemesterstudierende Psychologie (80% Frauen) der Goethe-Universität Frankfurt an der Untersuchung zu beiden Messzeitpunkten teilgenommen. Da eine Sicherheitseinschätzung nicht durch alle Studierenden vorgenommen wurde, beträgt hier die Anzahl gültiger Daten zu beiden Messzeitpunkten $N = 48$. Das Alter in der Stichprobe lag zwischen 17 und 52 Jahren ($M = 23$; $SD = 6.89$). Über alle drei Themenbereiche hinweg zeigte sich deskriptiv sowohl für die Fähigkeit als auch für die Sicherheit ein Anstieg vom ersten (t_1) zum zweiten (t_2) Messzeitpunkt (Abb.2 und 3).

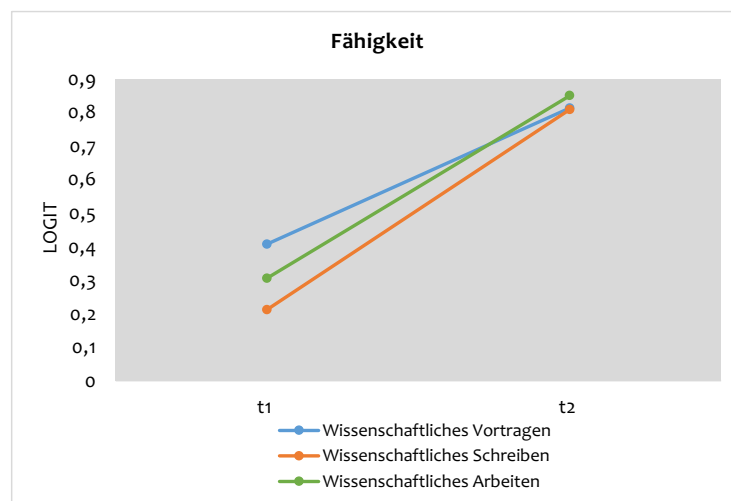


Abb. 2: Fähigkeit pro Themenbereich zu t1 und t2.

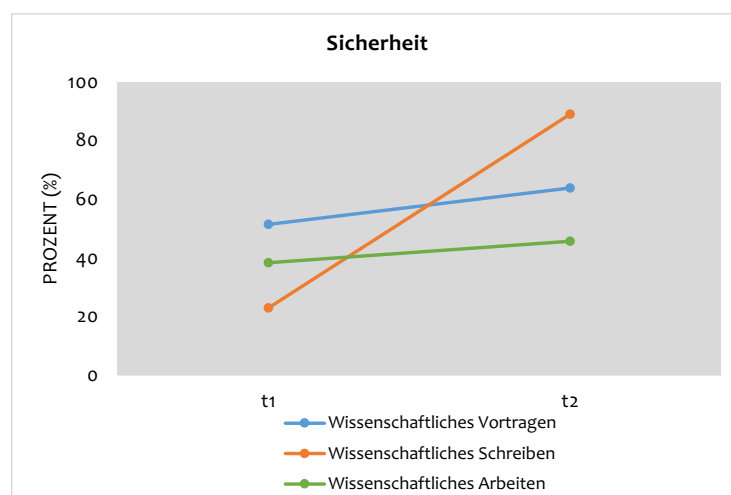


Abb. 3: Sicherheit pro Themenbereich zu t1 und t2.

Die Bedeutsamkeit dieser Veränderung über die Zeit wurde inferenzstatistisch in einer Multivariaten Varianzanalyse (MANOVA) mit Messwiederholung bestätigt. Der Test der Innersubjekteffekte zeigte für den Zeitfaktor einen signifikanten Effekt auf die Veränderung der Fähigkeit, $F(3,60) = 23.14, p < .01, \eta_p^2 = .536$. Für die Veränderung der Sicherheit konnte ebenfalls ein signifikanter Effekt festgehalten werden, $F(3, 45) = 21.86, p < .01, \eta_p^2 = .593$. Die Einzelergebnisse aus den Analysen für die einzelnen Themenbereiche finden sich in Tabelle 2.

Tab. 2: Fähigkeit und Sicherheit pro Themenbereich

Abhängige Variable	Themenbereiche	Prüfgröße F (df)	Effektgröße (η_p^2)
Fähigkeit	Wissenschaftliches Vortragen	$F(1,62) = 38.40^{**}$.382
	Wissenschaftliches Schreiben	$F(1,62) = 71.25^{**}$.535
	Wissenschaftliches Arbeiten	$F(1,62) = 29.87^{**}$.325
Sicherheit	Wissenschaftliches Vortragen	$F(1,47) = 6.04^*$.114
	Wissenschaftliches Schreiben	$F(1,47) = 55.19^{**}$.540
	Wissenschaftliches Arbeiten	$F(1,47) = 2.03$.041

Anmerkungen. MANOVA mit Messwiederholung. Ergebnisdarstellung aus den univariaten Tests. Änderung der Fähigkeit von t1 zu t2 mit $N = 63$. Änderung der Sicherheit von t1 zu t2 mit $N = 48$.

* $p < .05$. ** $p < .01$.

Während sich zu t1 das Ausgangslevel der Fähigkeit für die drei Themenbereiche im Mittel unterschieden hat, gleichen sich diese zu t2 im Mittel an (vgl. Abbildung 2). Hinsichtlich der Veränderung der Sicherheit ist Abbildung 3 zu entnehmen, dass der Zuwachs an Sicherheit im wissenschaftlichen Schreiben im Mittel am größten ist, wobei für die beiden Themenbereiche „wissenschaftliches Vortragen“ und „wissenschaftliches Schreiben“ jeweils ein signifikanter Anstieg von t1 zu t2 festgehalten werden konnte. Eine signifikante Veränderung der Sicherheit für wissenschaftliches Arbeiten konnte nicht nachgewiesen werden.

Die Fähigkeit der Studierenden zu t1 und t2 war unkorreliert, sodass in weiteren Analysen der Einfluss der angenommenen Heterogenität des Wissenstandes zu Studienbeginn und die unterschiedliche Entwicklung dieses Wissenstandes durch den Besuch der Veranstaltung untersucht wurden. Für die Fähigkeit beziehungsweise Sicherheit wurde dazu der Zuwachs, also die Differenz der Mittelwerte zwischen dem ersten und zweiten Messzeitpunkt, betrachtet und in Bezug zur Baseline, also dem Ausgangswert t1, gesetzt (vgl. Tu & Gilthorpe, 2007). Obwohl die so korrelierten Werte mathematisch nicht unabhängig voneinander sind und somit die resultierende Korrelation überschätzt werden kann (Chiolerio, Paradis, Rich & Hanley, 2013; Tu & Gilthorpe, 2007), konnte so exploriert werden, ob Studierende unterschiedlich vom Besuch der Veranstaltung profitierten. Die Ergebnisse zeigten für alle drei Themenbereiche eine signifikant negative Korrelation der Differenzwerte

mit der Baseline (vgl. Tabelle 3). Je niedriger die Fähigkeit beziehungsweise Sicherheit zu t1 war, desto höher fiel die Differenz zwischen t2 und t1 aus, das heißt desto höher war der Zuwachs der Fähigkeit respektive der Sicherheit.

Tab. 3: Korrelationen zu t1 und t2

Themenbereiche	Fähigkeit t1 & Differenz t1 zu t2	Sicherheit t1 & Differenz t1 zu t2
Wissenschaftliches Vortragen	-.613 ^{**}	-.547 ^{**}
Wissenschaftliches Schreiben	-.616 ^{**}	-.457 ^{**}
Wissenschaftliches Arbeiten	-.477 ^{**}	-.490 ^{**}

Anmerkungen. Korrelationen zwischen dem Ausgangsniveau der Fähigkeit bzw. Sicherheit und der Differenz der Mittelwerte t2 abzüglich t1 pro Themenbereich.

^{**} $p < .01$.

4 Fazit und Ausblick

4.1 Der konstruierte Wissenstest

Wie eingangs beschrieben, wurde ein IRT-basierter Wissenstest konstruiert, welcher folgende Ansprüche erfüllt: Erstens werden die Inhalte der betreffenden Lehrveranstaltung in den Aufgaben abgebildet, so dass eine gezielte Ergebnismeldung zu den behandelten Themenbereichen erfolgen kann. Zweitens wurden verschiedene Aufgabentypen kombiniert, wobei die korrekte Beantwortung der wahr-falsch Aussagen nicht nur falsch-Aussagen beinhalteten, um Antworttendenzen zu vermeiden. Drittens wurden zwei Testversionen mit unterschiedlichen Aufgaben erstellt. Für den zweiten Messzeitpunkt wurden Aufgaben ausgewählt, deren mittlere Schwierigkeiten höher waren als zum ersten Messzeitpunkt, um dem erwarteten Wissenszuwachs Rechnung zu tragen. Zudem wurden so Ergebnisverzerrungen durch Erinnerungs- oder Übungseffekte beim zweiten Messzeitpunkt vorgebeugt.

Die methodischen Stärken des Wissenstests betreffen sowohl die Konstruktion, als auch die Anwendung. Zum einen ist der Test mit insgesamt 36 Items (zeit)ökonomisch und lässt sich sowohl als Gesamttest zu einem Messzeitpunkt oder geteilt zu zwei Zeitpunkten einsetzen. Zum anderen wird durch den Einbezug der Aufgabenschwierigkeit ein präziseres Abbilden des Fähigkeitsniveaus ermöglicht, da berücksichtigt wird, dass die Ergebnisse nicht nur von der Fähigkeit der Person, sondern auch durch die Aufgabenschwierigkeit mitbestimmt wird. Eine weitere Stärke des Wissenstests besteht außerdem darin, Studierende mit bestehenden Fehlkonzepten zu konfrontieren und deren Abbau zu initiieren (Biggs & Tang, 2007). Anhand der erzielten Ergebnisse lässt sich dieses Ziel realisieren.

4.2 Diskussion der Ergebnisse

Durch die gezielte Entwicklung eines Wissenstest für den Bereich der allgemein-wissenschaftlichen Methodenkompetenzen sollte zunächst der Wissenserwerb – im Sinne eines Abbaus von Fehlkonzepten – in diesem Bereich erfasst werden. Darüber hinaus sollte eruiert werden, ob eine mögliche und intendierte Veränderung in der Überzeugung, mit der eine Antwortalternative gewählt wird, vorliegt. Um dies abzubilden wurden die Aufgaben immer um eine Einschätzung zur Sicherheit über die Richtigkeit der gegebenen Antwort ergänzt. Die Ergebnisse lassen sich wie folgt zusammenfassen:

1. Der Wissenszuwachs vom ersten zum zweiten Messzeitpunkt stellt einen Indikator für den Abbau von Fehlkonzepten dar. Das bedeutet, dass es einen signifikanten Anstieg in der Fähigkeit pro Themenbereich gab.
2. Die Ergebnisse zeigen einen bedeutsamen Anstieg der Sicherheit für die Themenbereiche „wissenschaftliches Vortragen“ und „wissenschaftliches Schreiben“. Durch den Wissenserwerb und ihre gezielte Anwendung in den verschiedenen Lehrformaten kann für diese beiden Themenbereiche abgeleitet werden, dass die Studierenden darin bestärkt wurden, wissenschaftsbezogene Aussagen subjektiv selbstsicherer zu beurteilen. Im Themenbereich „wissenschaftliches Arbeiten“ konnte ein solcher Anstieg in der Sicherheit nicht festgestellt werden. Dies könnte darauf zurückzuführen sein, dass in den Veranstaltungsformaten jenseits der Vorlesung diese Inhalte nicht praktisch erprobt wurden.
3. Anhand der Ergebnisse können außerdem Aussagen über die unterschiedliche Entwicklung des Wissensstandes getroffen werden. Dies ist in der vorliegenden Stichprobe besonders relevant, da heterogene Studieneingangsvoraussetzungen den Wissenszuwachs behindern können. Die Ergebnisse zeigten, dass Studierende mit einem niedrigen Ausgangswert einen höheren Zuwachs an Fähigkeit beziehungsweise Sicherheit aufweisen als diejenigen, die bereits mit einer hohen Fähigkeit beziehungsweise Sicherheit zu Studienbeginn starten. Dieses Ergebnis spricht dafür, dass ein zentrales Ziel der Studieneingangsphase, das heißt die Angleichung von Eingangsvoraussetzungen durch das Identifizieren und Aufdecken von Fehlkonzepten beziehungsweise deren kognitiv aktivierende Bearbeitung wie im beschriebenen Modul erreicht werden konnte.

4.3 Ist-Stand-Erfassung für Lernende und Lehrende

Nach erfolgreicher Teilnahme am Wissenstest sollen künftig Studierende anhand ihres Versuchspersonen-Codes ihr Ergebnis zu den einzelnen Messzeitpunkten zurückgemeldet bekommen. In diesem Rahmen erhalten sie Informationen darüber,

- wie viele von den vorgelegten Aufgaben sie richtig beantwortet haben und die prozentuale Angabe dazu (Gesamtergebnis),
- bei wie vielen falsch gelösten Aufgaben sie angegeben hatten, sicher zu sein (Fehleinschätzung: Fehlkonzepte),
- bei wie vielen richtig gelösten Aufgaben sie angegeben hatten, unsicher zu sein (Fehleinschätzung: Unterschätzung der eigenen Fähigkeit),

- bei wie vielen Aufgaben die Sicherheitseinschätzung mit der Beantwortung der Frage übereinstimmte (korrekte Einschätzung der eigenen Fähigkeit; falsch und unsicher bzw. richtig und sicher).

Durch dieses individuelle Feedback zu den Ergebnissen beider Messzeitpunkte können Studierende sowohl ihren Leistungsstand als auch das Ausmaß ihrer Fehlkonzepte mitgeteilt bekommen. Dies ermöglicht Studierenden nicht nur ihren Leistungsstand zu erkennen, sondern auch ihren Wissenszuwachs zu überprüfen.

Lehrende können anhand der Testergebnisse ableiten, in welchem Ausmaß Fehlkonzepte bei den Studierenden in ihrer Lehrveranstaltung vorliegen und in welchen Themenbereichen diese besonders ausgeprägt sind. Durch die Abfrage der Sicherheit können Lehrende analysieren, wie Fähigkeit und Sicherheit in den einzelnen Themenbereichen zusammenhängen. Auf Basis entsprechender Befunde gilt dann im Sinne adaptiver Lehre zu entscheiden, ob und welche Teile der Lehrveranstaltung optimiert oder ergänzt werden können, um eine Auseinandersetzung mit Themenbereichen, die mit Fehlkonzepten behaftet sind, stärker zu fördern oder die Sicherheit in diesen Themenbereichen zu stärken.

4.4 Ausblick

Auch wenn der Wissenstest für die frühzeitige Aufdeckung der Fehlkonzepte und vor allem für die Praxisforschung einen erheblichen Beitrag leisten kann, besteht eine erkennbare Limitation darin, dass diese Art der Testentwicklung methodisch und zeitlich aufwendig ist. Damit möglichst viele Dozierende von diesem bewährten Instrument profitieren können, wird in Erwägung gezogen, den Wissenstest künftig für andere Standorte oder fachnahe Disziplinen zugänglich zu machen. Außerdem können so weitere Synergien in der Weiterentwicklung des Tests geschaffen werden, indem beispielsweise für bestehende und neue Themenbereiche weitere Aufgaben generiert und bei den wachsenden Stichproben erprobt werden können.

Im Sinne der Weiterentwicklung der Feedback-Funktion des Wissenstests, ist es vorgesehen, dass Studierende nicht nur ein deskriptives Feedback zu ihrem Leistungsstand bekommen. Für besonders stark mit Fehlkonzepten behaftete Themenbereiche sollen zusätzliche Literaturempfehlungen zur Verfügung gestellt und Workshops angeboten werden. Das bedeutet, dass diejenigen Studierenden, die im Wissenstest eine geringe Fähigkeit rückgemeldet bekommen haben, speziell auf die entsprechenden Angebote hingewiesen werden. Auch Studierenden, die sich bei der Beantwortung der Fragen in bestimmten Themenbereichen unsicher waren, oder weiterführendes Interesse haben, soll die Teilnahme frei stehen. Dies wäre ein möglicher Ansatz, um Inhalte und Methoden auf diejenigen zuzuschneiden, die die betreffenden Themen bis dahin noch nicht vollständig verstanden oder verinnerlicht haben (vgl. Ramsden, 2003).

In der Vergangenheit wurden bereits auf Basis der Ergebnisse in der Pilotierung und der Nachfrage seitens Studierender Vertiefungs-Workshops in Form von interaktiven halb- oder ganztägigen Präsenzveranstaltungen zu folgenden Themen angeboten:

- Wissenschaftliches Schreiben (fachspezifisch)
- Kreatives Visualisieren (wissenschaftliches Vortragen)
- Körpersprache und Videofeedback (wissenschaftliches Vortragen)

Die Evaluationsergebnisse pro Workshop haben gezeigt, dass diese Angebote sehr gut angenommen wurden und entsprechend der identifizierten Bedarfe zukünftig ausgeweitet werden sollten. Neben didaktischen Modifikationen gibt es zudem weitere Begleitforschung. In diesem Rahmen soll unter anderem untersucht werden, inwiefern Lernmotivation, Selbstwirksamkeit und Zielorientierung mit Fehlkonzepten, dem Erfolg im Wissenstest, aber auch mit Wissenszuwachs zusammenhängen. Jenseits der Gestaltung von Lehrinhalten und -formaten könnten diese personenbezogenen Variablen Aufschluss über Tendenzen im Antwortverhalten im Sinne einer Über- oder Unterschätzung der eigenen Fähigkeiten geben. Langfristig soll schließlich nicht nur der konkrete Wissenserwerb im Hinblick auf allgemein-wissenschaftliche Methodenkompetenzen, sondern auch die Kompetenz im Umgang mit diesen Methoden im weiteren Verlauf des Studiums und damit implizit der generelle Studienerfolg und das persönliche Wachstum der Studierenden gefördert werden.

Literatur

- Amsel, E., Johnston, A., Alvarado, E., Kettering, J., Rankin, L. & Ward, M. (2009). The effect of perspective on misconceptions in psychology. A test of conceptual change theory. *The Journal of Instructional Psychology*, 36, 289–296.
- Biggs, J. B. & Tang, C. (2007). *Teaching for quality learning at university. What the student does* (3rd edition). Maidenhead: McGraw-Hill.
- Boser, J., Scherer, S., Kuchta, K., Wenzel, S. F. C. & Horz, H. (2017). Empirically founded teaching in psychology – An example for the combination of evidence-based teaching and the Scholarship of Teaching and Learning. *Psychology Learning & Teaching*, 16, 261–275.
- Chiolero, A., Paradis, G., Rich, B. & Hanley, J. A. (2013). Assessing the relationship between the baseline value of a continuous variable and subsequent change over time. *Frontiers in Public Health*, 1, 29.
- Dole, J. A. & Sinatra, G. M. (1998). Reconceptualizing change in the cognitive construction of knowledge. *Educational Psychologist*, 33, 109–128.
- Dunn, D. S., Saville, B. K., Baker, S. C. & Marek, P. (2013). Evidence-based teaching: Tools and techniques that promote learning in the psychology classroom. *Australian Journal of Psychology*, 65, 5–13.
- Frey, A., Hartig, J. & Rupp, A. A. (2009). An NCME instructional module on booklet designs in large-scale assessments of student achievement: Theory and practice. *Educational Measurement: Issues and Practice*, 28, 39–53.
- Furnham, A. & Hughes, D. J. (2014). Myths and misconceptions in popular psychology. *Teaching of Psychology*, 41 (3), 256–261.
- Hasselhorn, M. & Gold, A. (2017). *Pädagogische Psychologie. Erfolgreiches Lernen und Lehren* (4., aktualisierte Auflage). Stuttgart: Verlag W. Kohlhammer.
- Höhler, J., Horz, H., Hansen, M., Fabriz, S., Stehle, S., Heckmann, C., Ulrich, I. & Secon, R. (2012). Verbesserte fachliche und soziale Integration von Studienanfängern durch Peer-teaching im Bachelorstudiengang Psychologie. In M. Krämer, S. Dutke & J. Barenberg (Hrsg.), *Psychologiedidaktik und Evaluation IX* (S. 315–321). Aachen: Shaker.

- Hughes, S., Lyddy, F. & Lambe, S. (2013). Misconceptions about psychological science: A review. *Psychology Learning & Teaching*, 12, 20–31.
- Kowalski, P. & Taylor, A. K. (2009). The effect of refuting misconceptions in the introductory psychology class. *Teaching of Psychology*, 36, 153–159.
- Kuhle, B. X., Barber, J. M. & Bristol, A. S. (2009). Predicting students' performance in introductory psychology from their psychology misconceptions. *Journal of Instructional Psychology*, 36, 119–124.
- Landau, J. D. & Bavaria, A. J. (2003). Does deliberate source monitoring reduce students' misconceptions about psychology. *Teaching of Psychology*, 30, 311–314.
- Miller, R. L., Wozniak, W. J., Rust, M. R., Miller, B. R. & Slezak, J. (1996). Counterattitudinal advocacy as a means of enhancing instructional effectiveness: How to teach students what they do not want to know. *Teaching of Psychology*, 23, 215–219.
- Osterlind, S. J. (2002). *Constructing test items. Multiple-choice, constructed-response, performance, and other formats* (2nd ed.). Boston: Kluwer Academic.
- Ramsden, P. (2003). *Learning to teach in Higher Education* (2nd ed.). Hoboken: Taylor and Francis.
- Scherer, S., Boser, J. & Horz, H. (2014). "Starker Start ins Studium": Praxisbericht und Evaluation eines Moduls zur Verbesserung der Studieneingangsphase im Fach Psychologie. In M. Krämer, U. Weger & B. Zupancic (Hrsg.), *Psychologiedidaktik und Evaluation X* (S. 113–121). Aachen: Shaker.
- Schermelleh-Engel, K. & Werner, C. S. (2012). Methoden der Reliabilitätsbestimmung. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (Springer-Lehrbuch, Bd. 37, S. 119–141). Berlin, Heidelberg: Springer Berlin Heidelberg. Scott O.
- Lilienfeld, S. O., Lynn, S.J., Ruscio, J. & Beyerstein, B. L. (2010). *50 great myths of popular psychology. Shattering widespread misconceptions about human behavior* (1. ed.). Malden u.a.: Wiley-Blackwell.
- Smith III, J. P., diSessa, A. A. & Roschelle, J. (1994). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *The Journal of the Learning Sciences*, 3, 115–163.
- Sonntag, K. (2000). „Lebenslanges Lernen“ — Beiträge der Arbeits- und Organisationspsychologie. In W. Lempert (Hrsg.), *Psychologische Theorie, Empirie und Therapie* (Lebenslanges Lernen im Beruf, Bd. 3, S. 111–132). Opladen: Leske + Budrich.
- Taylor, A. K. & Kowalski, P. (2004). Naïve psychological science: The prevalence, strength, and sources of misconceptions. *The Psychological Record*, 54, 15–25.
- Tu, Y.-K. & Gilthorpe, M. S. (2007). Revisiting the relation between change and initial value: a review and evaluation. *Statistics in Medicine*, 26 (2), 443–457.
- UniReport (2014). *Ordnung des Fachbereichs Psychologie und Sportwissenschaften der Johann Wolfgang Goethe-Universität Frankfurt am Main für den Bachelorstudiengang Psychologie vom 2. Juli 2014* (Präsident der Johann Wolfgang Goethe-Universität Frankfurt am Main, Hrsg.). Frankfurt am Main: Goethe-Universität.
- van der Linden, W. J. (Hrsg.). (2016). *Handbook of item response theory (Statistics in the social and behavioral sciences)*. Boca Raton: CRC Press, Taylor & Francis Group.
- Vaughan, E. D. (1977). Misconceptions about psychology among introductory psychology students. *Teaching of Psychology*, 4, 138–141.

Young, R. A. & Collin, A. (2004). Introduction: Constructivism and social constructionism in the career field. *Journal of Vocational Behavior*, 64, 373–388.

Autor/-innen

Yasemin Z. Türkcorun. Goethe-Universität, Abteilung für Pädagogische Psychologie - Psychologie des Lehrens und Lernens im Erwachsenenalter, Frankfurt a.M., Deutschland; Email: tuerkcorun@psych.uni-frankfurt.de

S. Franziska C. Wenzel. Goethe-Universität, Abteilung für Pädagogische Psychologie - Psychologie des Lehrens und Lernens im Erwachsenenalter, Frankfurt a.M., Deutschland; Email: wenzel@psych.uni-frankfurt.de

Julia Mordel. Goethe-Universität, Abteilung für Pädagogische Psychologie - Psychologie des Lehrens und Lernens im Erwachsenenalter, Frankfurt a.M., Deutschland; Email: mordel@psych.uni-frankfurt.de

Dr. Sonja Scherer. Goethe-Universität, Abteilung für Pädagogische Psychologie - Psychologie des Lehrens und Lernens im Erwachsenenalter, Frankfurt a.M., Deutschland; Email: scherer@psych.uni-frankfurt.de

Prof. Dr. Holger Horz. Goethe-Universität, Abteilung für Pädagogische Psychologie - Psychologie des Lehrens und Lernens im Erwachsenenalter, Frankfurt a.M., Deutschland; Email: horz@psych.uni-frankfurt.de



Zitiervorschlag: Türkcorun, Y.Z., Wenzel, S.F.C., Mordel, J., Scherer, S. & Horz, H. (2019). Allgemein-wissenschaftliche Methodenkompetenzen erfassen und Fehlkonzepte aufdecken: Entwicklung und Anwendung eines Wissenstests in der Psychologie. *die hochschullehre*, Jahrgang 5/2019, online unter: www.hochschullehre.org