

Entwicklung und Validierung eines Erwartungs- und Interessenstests (E × I - Test) zur Erkundung studienfachspezifischer Passung in einem Online-Self-Assessment

Belinda Merkle, Moritz Schiltenswolf, Andrea Kiesel, Oliver Dickhäuser

Zusammenfassung: Realistische Erwartungen und Passung zwischen Interessen und Studieninhalten sind zentrale Ansatzpunkte bei der Steuerung von Studienwahlentscheidungen. In einem neu entwickelten fachspezifischen Erwartungs- und Interessenstest (E × I - Test) für Psychologie werden erstmals Erwartungsdiskrepanzen und Interessen kombiniert betrachtet und dementsprechend auch übertroffene oder enttäuschte Erwartungen erfasst und rückgemeldet. Die zu den Studieninhalten des neuen Verfahrens entwickelten Items konnten annähernd perfekt den Studienfachbereichen zugeordnet werden und deckten diese weitgehend vollständig und gleichmäßig ab. 2,033 Studieninteressierte bearbeiteten den E × I - Test im Rahmen eines Online-Self-Assessments und fühlten sich danach informierter als vorher. Insgesamt bewerteten die Studieninteressierten das neue Verfahren positiv und 94% würden es weiterempfehlen. Auf Basis des vorgestellten Verfahrens für das Bachelor-Psychologiestudium könnten weitere E × I - Tests für die Orientierung in andere Studienfächer oder Berufe entwickelt und validiert werden, für welche sowohl spezifische Interessen als auch enttäuschte Erwartungen eine Rolle spielen.

Schlüsselwörter: Erwartungs-Interessenstest, Erwartungsdiskrepanzen, Interessen, Valenz der Erwartungsdiskrepanzen, Studierende-Studienfach Passung, Online-Self-Assessment

Development and validation of an Expectation-Interest Test (E × I - Test) to explore fit for a specific major in an online self-assessment

Summary: Both, realistic expectations and fit between interests and study content are crucial to guide study choice decisions. A newly developed subject-specific Expectation-Interest Test (E × I - Test) for psychology considers, for the first time, expectation discrepancies and interests in combination. Thus, exceeded or disappointed expectations are assessed and reported back. It was shown that the newly developed items which represent study contents can be assigned almost perfectly to the study subject areas and cover them to a large extent completely and evenly. 2,033 prospective students completed the E × I - Test as part of an online self-assessment and felt more informed afterwards than before. Overall, prospective students rated the new procedure positively and 94% would recommend it to others. Based on the presented procedures for the bachelor psychology studies, further E × I - Tests could be developed and validated for orientation to other fields of study or professions, for which both specific interests and disappointed expectations play a role.

Keywords: expectation-interest test, expectation discrepancies, interests, valence of expectation discrepancies, student-study subject fit, online self-assessment

1 Einleitung

Die Hochschullandschaft in Deutschland hat sich in den letzten Jahren stark weiterentwickelt. Es gibt eine immer größere Vielfalt an Studiengängen (Frankenberg, 2008) und eine steigende Zahl an Studienanfänger*innen (Statistisches Bundesamt, 2020). Allerdings entsteht bei Studieninteressierten zunehmend Unsicherheit hinsichtlich der Frage, welches Studienfach das richtige für sie ist (Hasenberg, 2012; Wolter, 2008). Die Unsicherheit bei der Studienwahl ist nicht überraschend, wenn man bedenkt, dass die Entscheidung für einen bestimmten Studiengang einen komplexen Informations- und Reflexionsprozess erfordert.

Gemäß der Person-Environment Fit-Theorie drückt sich Passung zwischen Studierenden und Studienfach unter anderem in der Passung zwischen Interessen und Studieninhalten aus, was wiederum die spätere Studienzufriedenheit beeinflusst (Cable & DeRue, 2002; Westermann & Heise, 2018). Zusätzlich spielt die Realitätsnähe der Erwartungen an das Studienfach eine wichtige Rolle für den Studienerfolg (Hasenberg & Schmidt-Atzert, 2013). Somit müssen Studieninteressierte über ihre eigenen Interessen reflektieren und sich über die Angebote ihres angestrebten Studiengangs informieren, um ein akkurates Bild von sich selbst und von ihrem Studiengang zu gewinnen sowie etwaige falsche Erwartungen zu korrigieren (Hasenberg & Schmidt-Atzert, 2013). Und schließlich müssen sie diese beiden Bilder miteinander vergleichen und abwägen, ob der Studiengang der passende für sie ist. Dass dies häufig misslingt, zeigen die zahlreichen Studienfachwechsel und Studienabbrüche von Studierenden, die zu Beginn des Studiums nicht ausreichend informiert waren, mit den falschen Erwartungen ins Studium gestartet sind oder zu geringes Interesse am Studienfach hatten (Heublein, Hutzsch, Schreiber, Sommer & Besuch, 2010; Schiefele, Streblov & Brinkmann, 2007; Schmidt-Atzert, 2005).

Da Studienabbruch mit persönlichen und gesellschaftlichen Kosten verbunden ist, ist es sinnvoll, Studieninteressierte bei diesem Informations- und Reflexionsprozess zu unterstützen (Frebort & Kubinger, 2008; Hasenberg, 2012). Dabei können Erwartungs- und Interessenstests in Online-Self-Assessments (OSAs) einen großen Beitrag leisten, da diese meist kostenlos, eigenständig und barrierearm bearbeitet werden können und somit für Studieninteressierte leicht zugänglich sind (Hasenberg & Schmidt-Atzert, 2014). Je nach der Phase im Entscheidungsprozess können Studieninteressierte dabei auf allgemeine oder fachspezifische Verfahren zurückgreifen (Heukamp, Putz, Milbradt & Hornke, 2009). Allgemeinere Verfahren liefern eine erste Orientierung, was generell in einem Studium erwartet wird und welche der vielen Studienfächer für einen Studieninteressierten in Frage kommen (z.B. *Orientierungstest*, Ministerium für Wissenschaft; vgl. Hell, 2009). Diese Verfahren basieren häufig auf dem RIASEC-Modell (Holland, 1997), welches Interessen in sechs grundlegenden Bereichen umfasst und damit zur Orientierung in einem breiten Range von Studiengängen genutzt werden kann. Zu diesem Zweck erhalten Studieninteressierte eine Empfehlung, welche der vielen Studiengänge zu ihren Interessen passen. Somit sind diese Verfahren besonders zu Beginn des Entscheidungsprozesses relevant (Hasenberg, 2012; Heukamp et al., 2009). Je konkreter der Studienwunsch ist, desto wichtiger sind fachspezifisch-konfirmierende Verfahren, in denen die Passung für ein konkretes Studienfach überprüft werden kann (z.B. *Online-Self-Assessment des Bachelorstudiengangs Psychologie*, Universität Marburg; vgl. Hasenberg & Schmidt-Atzert, 2014). Ziel dieser Gruppe von Verfahren sollte es sein, Interessen und Erwartungsdiskrepanzen (Abweichung der Erwartung der Teilnehmenden von der Ein-

schätzung der Expert*innen) hinsichtlich studienfachspezifischer Themen und Inhalte adäquat zu messen, darauf basierend den Studieninteressierten ein realistisches Bild von sich selbst und von ihrem Studiengang rückzumelden, um letztlich Studienwahlentscheidungen zu optimieren und Studienerfolg zu fördern.

Bisherige Forschung konnte zeigen, dass ein größeres Ausmaß an intrinsischer Studienwahlmotivation (z.B. Wahl eines Studienfachs aus Interesse) mit höherer Studienmotivation, höherer Studienzufriedenheit und geringeren Studienabbruchintentionen in Verbindung steht (Janke, Messerer, Merkle & Krille, 2021; Kegel et al., 2020). Somit sollte das Interesse einen wichtigen Faktor bei der Studienwahlentscheidung darstellen, um ein erfolgreiches Studium zu ermöglichen. Die Optimierung von Studienwahlentscheidungen hinsichtlich der Interessenkongruenz sollte dem Studienerfolg zuträglich sein. Forschung im Bereich der Erwartungstests zeigte, dass die Bearbeitung eines fachspezifischen Erwartungstests mit anschließender Rückmeldung mit einem höheren Informationsstand und realistischeren Erwartungen korrespondierte (Vent & Erdfelder, 2009), womit schließlich auch größerer Studienerfolg in Verbindung stehen sollte (Burkhardt & Hagemeyer, 2018; Hasenberg & Schmidt-Atzert, 2013). Dennoch gibt es, wie sich nachfolgend zeigen wird, Defizite bestehender Verfahren. Basierend auf wissenschaftlichen Grundlagen adressieren wir diese Mängel und liefern vier neuartige Beiträge zur Erwartungs- und Interessensforschung sowie zur praktischen Anwendung von Erwartungs- und Interesstests bei der Studienorientierung.

Der *erste Beitrag* unserer Arbeit liegt in der Konzeption eines Erwartungs- und Interesstests, welcher Erwartungsdiskrepanzen und Interessen in Kombination misst und rückmeldet. Bisher wurden Erwartungsdiskrepanzen in der Forschung und Praxis von Erwartungstests isoliert betrachtet und somit lediglich falsche Erwartungen korrigiert oder rückgemeldet, zu welchem Ausmaß die Erwartungen der Studieninteressierten mit den Einschätzungen der Expert*innen übereinstimmen (Hasenberg, 2012; Karst, Ertelt, Frey & Dickhäuser, 2017). Mithilfe des hier vorgestellten Verfahrens kann zum ersten Mal gemessen und rückgemeldet werden, ob ein bestimmtes Studienfach die Erwartungen von Studieninteressierten übertrifft oder enttäuscht (Valenz der Erwartungsdiskrepanz). Dies ist relevant, da die Valenz der Erwartungsdiskrepanz sowohl die Studienwahlentscheidung als auch die spätere Studienzufriedenheit beeinflussen sollte (Hasenberg, 2012; Karst et al., 2017).

Der *zweite Beitrag* bezieht sich auf neue Validierungsmethoden für die Items von Erwartungs- und Interesstests hinsichtlich ihrer Struktur, Vollständigkeit und Prototypizität. Mit Hilfe dieser Methoden wird in ersten empirischen Studien überprüft, ob die Items zu Studieninhalten eindeutig den zugehörigen Studienfachbereichen zugeordnet werden können und ob sie diese Studienfachbereiche ausreichend und gleichmäßig vollständig sowie gleichmäßig prototypisch abbilden.

Der *dritte Beitrag* ist konzeptioneller Natur und ergibt sich aus der Hochschulunabhängigkeit dieses fachspezifischen Verfahrens. Bisher sind fachspezifische Verfahren meist hochschulspezifisch angelegt (Hasenberg, 2012). Das bietet die Möglichkeit, auch hochschuleigene Marketinginteressen zu verfolgen, was teilweise auch explizit als Funktion von OSAs begriffen wird (Störk & Mocigemba, 2013). Da das neue Verfahren hochschulunabhängig ist, besteht diese Möglichkeit erst gar nicht. Außerdem ist das neue Verfahren durch seine Hochschulunabhängigkeit besonders geeignet für den großen Anteil an Studieninteressierten, die zunächst ihre allgemeine Passung für ein Fach erkunden wollen, bevor sie sich über standortspezifische Besonderheiten informieren (Hovestadt & Stegelmann, 2011).

Der *vierte Beitrag* dieses Manuskripts liegt schließlich darin, erste empirische Befunde zur Evaluation der Güte des gesamten neuen Verfahrens hinsichtlich der Akzeptanz, der Informiertheit und der Weiterempfehlung durch Studieninteressierte zu liefern. Jeder dieser vier Beiträge wird nachfolgend in je einem Abschnitt erläutert.

1.1 Messung und Rückmeldung von Erwartungsdiskrepanzen und Interessen

Es gibt bisher keine einheitliche Struktur zur Verbindung von Erwartungs- und Interessentests. An einigen Hochschulen werden auf fachspezifischer Ebene nur Erwartungstests durchgeführt und keine Interessentests (z.B. *Fit4TU*, Technische Universität Braunschweig; vgl. Thiele & Kauffeld, 2019), wobei die Interessentests manchmal auf fachunspezifischer Ebene vorgeschaltet sind, um die Auswahl an Studienfächern auf einige passende zu begrenzen (z.B. *Study Finder Interessentest und Erwartungcheck*, Universität des Saarlandes; vgl. Stoll & Spinath, 2015). An anderen Hochschulen werden auf fachspezifischer Ebene sowohl Erwartungen als auch Interessen abgefragt (z.B. *Online Studienwahl Assistent*, Universität Freiburg; vgl. Störk & Mocigemba, 2013), wobei nur beim *Bonner testMaker* (vgl. Sindern & Pietrangeli, 2009) Erwartungen und Interessen auch für dieselben Inhalte abgefragt werden, was jedoch von enormer Wichtigkeit für die Rückmeldung ist und in diesem Zusammenhang später detaillierter erläutert wird. Auch die konkrete Messung der Erwartungen und Interessen unterscheidet sich hinsichtlich der Skaleninstruktionen und Antwortoptionen stark über verschiedene OSAs hinweg. Interesse wird in fachspezifischen OSAs als Interesse an fachspezifischen Themen erfragt (z.B. *Bonner testMaker*) oder es wird nach einer Einschätzung der Wichtigkeit verschiedener Interessensbereiche für ein erfolgreiches Psychologiestudium gefragt (z.B. *Online Studienwahl Assistent*).

Die Verfahren zur Erfassung von Erwartungen sind ebenfalls heterogen. Bei manchen Erwartungchecks werden Aussagen über den jeweiligen Studiengang präsentiert, welche von den Studieninteressierten in einem dichotomen Antwortformat als richtige oder falsche Erwartungen klassifiziert werden sollen (z.B. *Bonner testMaker*). Bei anderen Erwartungstests kann man auf mehrstufigen Antwortformaten angeben, für wie treffend man verschiedene Aussagen über das Studium hält oder inwieweit man diesen zustimmt (z.B. *stimmt gar nicht bis stimmt völlig*, *Mannheimer Informationssystem für Studieninteressierte der Sozialwissenschaften* (MISS), Universität Mannheim; vgl. Vent & Erdfelder, 2009).

Allerdings ist dabei nicht eindeutig, was durch das Maß der Zustimmung zu einer Aussage über Studieninhalte erfasst wird. Es könnte ausdrücken, in welchem zeitlichen Ausmaß man einen bestimmten Inhalt im Studium erwartet, für wie wichtig man einen bestimmten Inhalt im Studienfach hält oder auch wie sicher man sich in dieser Erwartung ist. Studieninteressierte könnten demnach bei einem Inhalt, den sie in einem großen zeitlichen Ausmaß im Studium erwarten, sich aber dessen unsicher sind, eine mittlere Zustimmung angeben und würden damit denselben Wert erhalten wie Expert*innen, die sich sicher sind, dass dieser Inhalt einen durchschnittlichen Anteil am Studium einnimmt. Solche fehlende Eindeutigkeit in der Itemformulierung ist problematisch, da es zu einem unterschiedlichen Verständnis bei verschiedenen Personen führen kann, wodurch die Messungen nicht mehr vergleichbar sind (Jonkisz, Moosbrugger & Brandt, 2012). Eine weitere Schwierigkeit bei der Beantwortung und Interpretation bestehender Erwartungstests ist, dass die Zustimmungswahrscheinlichkeit zu einer Aussage über einen Inhalt nicht nur durch das zeitliche Ausmaß dieses Inhalts

beeinflusst wird, sondern auch durch die Einbettung des Inhalts. Beispielsweise bekommt die Einbettung „ein bestimmter Inhalt steht im Mittelpunkt eines Studiums“ eine geringere Zustimmungswahrscheinlichkeit als „ein bestimmter Inhalt wird im Studium behandelt“ trotz gleichen zeitlichen Ausmaßes im Studium. Um nicht nur falsche Erwartungen aufzuzeigen, sondern gleichzeitig ein möglichst akkurates Bild vom Studium aufzubauen, schlagen wir vor, über Zustimmungsabfragen zu bestehenden Aussagen hinauszugehen und stattdessen explizit die Erwartungen über das zeitliche Ausmaß verschiedener Inhalte in einem Studienfach zu erfassen.

Ähnlich divers wie die Messung von Erwartungen und Interessen gestalten sich auch die Rückmeldungen der Ergebnisse der fachspezifischen Tests. Gemeinsame Grundlage der meisten Rückmeldungen ist der Vergleich der Erwartungen von Studieninteressierten mit den Einschätzungen von Expert*innen (Studierenden und/oder Dozierenden). Bei „Korrekt-/Falsch-Aussagen“ wird rückgemeldet, welche Aussagen richtig eingeschätzt wurden und welche nicht; und gegebenenfalls werden noch weitere Informationen angeboten (z.B. *Study Finder Erwartungscheck*). Zusätzlich dazu geben einige Erwartungstests an, wieviel Prozent der Erwartungen insgesamt zutreffend waren (z.B. *Online Studienwahl Assistent*). Dies sagt zwar etwas über die Informiertheit eines Studieninteressierten aus, bietet den Studieninteressierten per se aber noch keine Information darüber, wie sie ihre Einstellung zum Studium und damit ihre Studienwahlentscheidung ändern sollten. Karst et al. (2017) diskutieren, dass eine Fehleinschätzung des Ausmaßes spezifischer Inhalte, in Abhängigkeit von der Wichtigkeit des betreffenden Inhalts für die Studieninteressierten, differentielle Effekte auf die Einstellung zum Studium haben sollte. Auch Hasenberg (2012) nimmt an, dass übertroffene Erwartungen zu mehr Studienzufriedenheit führen könnten.

Diese theoretischen Überlegungen im Rahmen der Studienwahl lassen sich durch Befunde im Bereich der Person-Environment Fit-Theorie im Bereich beruflicher Interessen stützen, die zeigen konnten, dass ein Mangel an interessanten Inhalten in der Umwelt zwar mit weniger Zufriedenheit einhergeht, ein Überfluss an interessanten Inhalten in der Umwelt aber auch mit mehr Zufriedenheit in Verbindung stehen kann (Wiegand, Drasgow & Rounds, 2021). Dies verdeutlicht, wie wichtig es in einem ersten Schritt ist, bei der Konzeptualisierung von Erwartungstests zwischen diesen beiden Formen von Misfit zu unterscheiden und dadurch Studieninteressierten mit Erwartungsdiskrepanzen zusätzlich rückmelden zu können, ob ihre Erwartungen enttäuscht oder übertroffen werden.

Gegenwärtig sind uns keine Erwartungs- und Interessenstests bekannt, welche dies leisten. Einige Tests, bei denen man graduell Zustimmung oder Erwartung ausdrücken kann, melden Studieninteressierten zumindest zurück, ob laut Expert*inneneinschätzung mehr oder weniger von einem Inhalt im Studium vorkommt als von den Studieninteressierten erwartet. Diese Information über die Erwartungsdiskrepanz allein sagt allerdings nichts darüber aus, ob die Erwartungen der Studieninteressierten dadurch enttäuscht oder übertroffen wurden. Um diese Information zu vermitteln, muss man zusätzlich zur Erwartungsdiskrepanz hinsichtlich bestimmter Inhalte und zu den Interessen an bestimmten Inhalten auch eine Kombination von Erwartungsdiskrepanzen und Interessen rückmelden. Dadurch wird berücksichtigt, dass Erwartungsdiskrepanzen hinsichtlich eines Inhalts, je nach Interesse am jeweiligen Inhalt, von unterschiedlicher Wertigkeit sein können. Diese neue Kombination wird nachfolgend als *Valenz der Erwartungsdiskrepanz* bezeichnet.

Dementsprechend werden im neuen Erwartungs- und Interessenstest drei Kennwerte voneinander abgegrenzt und separat erfasst: Das *Interesse der Studieninteressierten* (an be-

stimmten Studieninhalten); die *Erwartung der Studieninteressierten* (an das zeitliche Ausmaß bestimmter Studieninhalte); die *Expert*inneneinschätzung* (die faktische Realisation des zeitlichen Ausmaßes bestimmter Studieninhalte). Für die Rückmeldung lässt sich dann aus diesen Kennwerten im ersten Schritt die Erwartungsdiskrepanz berechnen, als Differenz der Expert*inneneinschätzung und der persönlichen Erwartung der Studieninteressierten. Die Valenz der Erwartungsdiskrepanz ergibt sich dann im zweiten Schritt aus der Erwartungsdiskrepanz multipliziert mit dem Interesse am jeweiligen Inhalt.

In Tabelle 1 sind Beispiele verschiedener Werte für Interessen, Erwartungsdiskrepanzen und die daraus resultierenden Valenzen der Erwartungsdiskrepanzen dargestellt, mit deren Hilfe wir im Folgenden die Bedeutung der Valenz der Erwartungsdiskrepanz erläutern. Eine negative Valenz der Erwartungsdiskrepanz ergibt sich für enttäuschte Erwartungen, die entstehen, wenn im Studium von einem bestimmten Inhalt, an dem man interessiert ist, weniger als erwartet vorkommt (Mangel, Wiegand et al., 2021; -18, siehe Tabelle 1 C) oder von einem Inhalt, an dem man überhaupt nicht interessiert ist, mehr vorkommt als erwartet (-18, siehe Tabelle 1 G). Dementsprechend sollte eine negative Valenz der Erwartungsdiskrepanz zu einer negativeren Einstellung zum Studienfach führen (geringeres Passungserleben, geringere Studienwahlsicherheit) und letztlich auch zu geringerer Studienzufriedenheit. Eine neutrale Valenz der Erwartungsdiskrepanz entsteht, wenn die Erwartung der Teilnehmenden über das zeitliche Ausmaß eines Inhalts mit dem realen Ausmaß des Inhalts laut Expert*inneneinschätzung übereinstimmt (0, siehe Tabelle 1 B, E, H). Da in diesem Fall keine Erwartungsdiskrepanz hinsichtlich des jeweiligen Inhalts vorliegt, können Erwartungen weder enttäuscht noch übertroffen werden und die Einstellung zum Studium sollte sich dementsprechend nicht ändern, unabhängig vom Interesse am jeweiligen Inhalt. Außerdem kann eine neutrale Valenz der Erwartungsdiskrepanz entstehen, wenn man einem Inhalt neutral beziehungsweise mit mittelmäßigem Interesse gegenübersteht (0, siehe Tabelle 1 D, E, F). In diesem Fall ist weder ein größeres zeitliches Ausmaß noch ein kleineres zeitliches Ausmaß dieses Inhalts relevant für die Einstellung zum Studium beziehungsweise die Studienzufriedenheit. Eine positive Erwartungsdiskrepanz beschreibt übertroffene Erwartungen, welche entweder dadurch entstehen, dass im Studium von einem bestimmten Inhalt, an dem man interessiert ist, mehr als erwartet vorkommt (Überfluss, Wiegand et al., 2021; 18, siehe Tabelle 1 A) oder von einem Inhalt, an dem man überhaupt nicht interessiert ist, weniger vorkommt als erwartet (18, siehe Tabelle 1 I). Dies sollte dazu führen, dass man seine Einstellung gegenüber dem Studienfach positiv ändert (höheres Passungserleben, stärkere Studienwahlsicherheit) sowie zufriedener mit dem Studienfach ist.

Der erste Beitrag der vorliegenden Arbeit besteht demnach darin, einen fachspezifischen Interessens- und Erwartungstest vorzustellen, welcher die Valenz der Erwartungsdiskrepanz erfassen und rückmelden kann, indem er Interesse an Studieninhalten und Erwartungsdiskrepanzen hinsichtlich des zeitlichen Ausmaßes dieser Inhalte in Kombination betrachtet. Aufgrund der Verknüpfung von Erwartungen mit Interessen wird der neu entwickelte Test als Erwartungs- \times Interessentest bezeichnet ($E \times I$ -Test).

Tabelle 1: Beispielwerte für die Valenz der Erwartungsdiskrepanz als Kombination von Interesse und Erwartungsdiskrepanz

Interesse	Erwartungsdiskrepanz		
	mehr als erwartet ($\Delta > 0$ max. 6)	wie erwartet ($\Delta = 0$)	weniger als erwartet ($\Delta < 0$ (min. -6))
hoch (max. +3)	18 (A)	0 (B)	-18 (C)
mittel (0)	0 (D)	0 (E)	0 (F)
niedrig (min. -3)	-18 (G)	0 (H)	18 (I)

Anmerkung: Δ bezeichnet die Differenz zwischen der Expert*inneneinschätzung und der persönlichen Erwartung der Studieninteressierten. Unter Zugrundelegung eines 7-stufigen Antwortformats bei der Erfassung dieser Kennwerte beträgt die maximale Differenz 6 beziehungsweise -6.

1.2 Entwicklung und Validierung von Items

Zusätzlich zur Festlegung auf ein übergeordnetes Konzept zur Erfassung und Rückmeldung von Erwartungen und Interessen stellen die Entwicklung und Validierung der konkreten Items sowie die Erhebung und Validierung der Expert*inneneinschätzung weitere kritische Faktoren bei der Entwicklung von fachspezifischen Erwartungs- und Interessenstests dar. So wie es keine Standardstruktur für Erwartungstests gibt, so gibt es auch kein allgemein anerkanntes Standardvorgehen bei deren Entwicklung und Validierung. Ein möglicher Prozess besteht in der Durchführung von hochschulinternen Workshops mit Studierenden, Dozierenden und Beratern für das jeweilige Fach, bei welchen verschiedene Items generiert werden, von denen dann die relevantesten in den Erwartungstest aufgenommen werden (z.B. Stoll & Spinath, 2015). Vent und Erdfelder (2009) wendeten ein alternatives Verfahren zur Auswahl der Iteminhalte an, bei welchem zunächst basierend auf ausgiebigen Recherchen von Lehrbüchern, Vorlesungsverzeichnissen, Studieneinführungsliteratur und Prüfungsordnungen zentrale Studieninhalte herausgearbeitet und durch benötigte Vorkenntnisse sowie spätere berufliche Möglichkeiten ergänzt wurden. Um auch häufige falsche Erwartungen der Studieninteressierten an den Studiengang zu inkludieren, wurden zusätzlich Studieninteressierte nach ihren Erwartungen gefragt. Schließlich wurden alle gesammelten Inhalte von Lehrenden und Studierenden auf Vollständigkeit überprüft (Vent & Erdfelder, 2009). Zusätzlich dazu wurde die zugrundeliegende Struktur der Erwartungstests überprüft, weil auf Basis dieser Struktur Erwartungstests beim Feedback in Oberkategorien zusammengefasst werden (z.B. Wirtschafts- und Werbepsychologie, Anwendungsfelder, etc. im *Mannheimer Informationssystem für Studieninteressierte der Sozialwissenschaften* (MISS)). Bisher wurde für die Überprüfung dieser Struktur getestet, ob die Antworten der Studieninteressierten sich faktorenanalytisch in die erwarteten Kategorien einteilen (Kossner, 2015). Als letzter wichtiger Schritt wurden Expert*inneneinschätzungen für das Ausmaß der Studieninhalte eingeholt, welche in der Rückmeldung mit der Erwartung der Studieninteressierten kontrastiert werden (vgl. Stoll & Spinath, 2015; Vent & Erdfelder, 2009). All dies zeigt, dass es bereits erste Bestrebungen gibt, eine fundierte Itementwicklung und -validierung zu betreiben. Gleichzeitig weisen die bisherigen Vorgehensweisen noch einige Lücken auf.

Die faktoranalytische Betrachtung von Items (Vent & Erdfelder, 2009) ist kritisch zu bewerten, da sich dadurch Faktoren von Inhalten bilden sollten, die in einem ähnlichen Aus-

maß von Studieninteressierten im Studium erwartet werden. Dies ist unter anderem problematisch, weil die so gebildeten Faktoren die systematischen, gegebenenfalls falschen Vorstellungen der Studieninteressierten widerspiegeln und nicht die Studienrealität. Daher schlagen wir vor, die Struktur der Items mithilfe von Expert*innen zu überprüfen, welche die zu verschiedenen Studieninhalten entwickelten Items den vorher definierten Oberkategorien (Fachbereiche eines Studienfachs, z.B. Klinische Psychologie, Pädagogische Psychologie, etc.) zuordnen sollen. Eine eindeutige Zuordenbarkeit der Items zu jeweils einer Oberkategorie spricht dafür, dass eine Zusammenfassung der betreffenden Items in die jeweilige Oberkategorie angemessen ist und demnach Studieninteressierte dabei unterstützen sollte, ein realistisches Bild vom Studium (gemäß dem Bild von Expert*innen) aufzubauen.

Zusätzlich zur Frage nach der Struktur der Items sollte auch sichergestellt werden, dass die Items das interessierende Merkmal – die Inhalte der Fachbereiche eines Studienfachs – möglichst vollständig abdecken und dass die Items in einem angemessenen Verhältnis zueinander stehen (Hartig, Frey & Jude, 2012). Dies ist nicht nur für die inhaltliche Validität wichtig, sondern auch dafür, dass die Studieninteressierten ein realistisches Bild vom Studienfach erhalten, was besonders bei (ungleich) unvollständiger Repräsentation verschiedener Studienfachbereiche nicht möglich wäre. Demnach sollte auch die Vollständigkeit und Prototypizität der Items überprüft werden.

Nachdem die Items zu den Studieninhalten finalisiert sind, werden die Einschätzungen der Expert*innen zu den betreffenden Inhalten eingeholt und gemittelt. Hier wurde oft mit kleineren Gruppen von Expert*innen gearbeitet, die im Rahmen von Workshops um ihre Einschätzungen gebeten wurden (Stoll & Spinath, 2015). Wir halten es für problematisch, hier mit sehr kleinen Stichproben zu arbeiten, und empfehlen, für präzisere Schätzungen auf möglichst groß angelegte Studierenden- und Dozierendenbefragungen zurückzugreifen (Leonhart, 2017). Allerdings sollte vor einer Zusammenfassung der Studierenden- und Dozierendenratings überprüft werden, ob diese eine homogene Expert*inneneinschätzung bilden oder ob systematische Abweichungen vorliegen. Dozierende unterscheiden sich von Studierenden beispielsweise dahingehend, dass das eigene Studium schon länger zurückliegt als bei den Studierenden, wobei sie dafür längere Erfahrung an der Hochschule in einem konkreten Studienfachbereich vorweisen können.

Angesicht dieser aufgezeigten Schwierigkeiten in der Forschung und Praxis der Entwicklung von OSAs besteht der zweite Beitrag dieser Publikation darin, zusätzliche Methoden zur Entwicklung und Validierung von Items sowie die Erhebung und Validierung der Expert*inneneinschätzung vorzustellen, die bei der Entwicklung neuer Erwartungs- und Interessenstests wichtig sind.

1.3 Hochschulunabhängigkeit

Sowohl das Vorgehen bei der Messung und Rückmeldung von Erwartungsdiskrepanzen und Interessen als auch das Vorgehen bei der Entwicklung und Validierung von Items kann von den bisher existierenden fachspezifischen Verfahren genutzt werden, die von einzelnen Hochschulen speziell für die eigenen Studienfächer erstellt wurden und damit zu den *fach- und hochschulspezifischen* Verfahren zählen (z.B. *MISS*). Diese Verfahren sind gut geeignet für alle Studieninteressierten, die bereits entschieden haben, an welcher Hochschule sie konkret studieren möchten, und überprüfen möchten, welches Fach an der jeweiligen Hochschule

am besten zu ihnen passt. Viele Studierende entscheiden sich allerdings zuerst für ein Studienfach, bevor sie sich auf eine bestimmte Hochschule festlegen (Hovestadt & Stegelmann, 2011). Demnach sind in einem ersten Schritt Verfahren relevant, die dabei helfen, die Frage zu beantworten, ob ein bestimmtes Studienfach unabhängig von einer bestimmten Hochschule zu einer bestimmten Person passt, die sich für ein Studium interessiert (fortan als *studienfachspezifisch, hochschulunabhängig* bezeichnet)¹. Leider mangelt es an solchen Verfahren, obwohl diese zusätzlich den Vorteil haben, dass etwaige Marketinginteressen von spezifischen Hochschulen, welche in OSAs auch verfolgt werden können, in den Hintergrund treten (Störk & Mocigemba, 2013). Dadurch kann ohne Interessenskonflikt adäquat sowohl auf beliebte als auch unbeliebte Studieninhalte eingegangen werden, was wichtig ist, um die Erwartungen von Studieninteressierten möglichst gut an die Realität anzupassen, selbst wenn dies – dem Marketinginteresse entgegenwirkend – Studieninteressierte von einer Bewerbung abhalten könnte.

Der dritte Beitrag dieser Arbeit ist es, einen fachspezifischen Erwartungs- und Interessenstest vorzustellen, der es Studieninteressierten erlaubt, ihre Passung zu einem spezifischen Studienfach unabhängig von etwaigen hochschulspezifischen Besonderheiten zu überprüfen.

2 Entwicklung des Erwartungs- und Interessenstests ($E \times I$ - Test)

Der hier vorgestellte $E \times I$ - Test wurde im Rahmen eines Projekts zur Entwicklung eines neuen Studierendenauswahlverfahrens für Psychologie in Baden-Württemberg (*STAV-Psych BaWü*) für ein OSA speziell für das Bachelor-Psychologiestudium entwickelt (*OSA-Psych*). Das Vorgehen zur Entwicklung und Validierung der Items, der Fokus auf Hochschulunabhängigkeit sowie das Konzept zur Kombination von Erwartungen und Interessen im Assessment und in der Rückmeldung sind allerdings allgemein für die Entwicklung von fachspezifischen Erwartungs- und Interessenstests relevant und können auf andere Studienfächer übertragen werden.

2.1 Literaturrecherche zur Auswahl der Inhalte

Den ersten Schritt im Prozess der Entwicklung der Items stellte eine systematische Recherche zu den Inhalten des jeweiligen Studienfachs dar, in diesem Fall den Inhalten des Bachelor-Psychologiestudiums unabhängig von der spezifischen Hochschule. Um zu gewährleisten, dass alle zentralen Inhaltsbereiche des Bachelor-Psychologiestudiums abgedeckt werden, legten wir als Basis für die Auswahl unserer Items das Rahmencurriculum zugrunde, das für dieses Studienfach von der entsprechenden Fachgesellschaft (Deutsche Gesellschaft für Psychologie, DGPs) empfohlen wird. Daraus extrahierten wir die Studienfachbereiche, die allen Hochschulen mit DGPs Gütesiegel gemeinsam sind, und fassten die zugehörigen Studieninhalte zusammen. Dabei nutzten wir hochschulunabhängige Quellen (z.B. Lehrbücher) sowie

1 Hasenberg (2012) stellte zudem bereits fest, dass bisher keine studienfachspezifischen und hochschulübergreifenden Verfahren existieren. Solche Verfahren wären in einem zweiten Schritt hilfreich, um die Frage zu beantworten, an welcher Hochschule ein ausgewähltes Studienfach studiert werden sollte (nach Konfirmierung des Studienfachs, zur Orientierung hin zu einem Hochschulstandort).

hochschulspezifische Quellen (z.B. Modulkataloge und Items bestehender Erwartungs- und Interessenstests). Zusätzlich sammelten wir bei der Recherche Informationen zu sonstigen Studieninhalten, wie Arbeitsweisen, die übergeordnet für mehrere Studienfachbereiche der Psychologie wichtig sind, und zu Inhalten, die Studieninteressierte fälschlicherweise im Psychologiestudium erwarten. Basierend darauf erstellten wir für alle Studienfachbereiche jeweils vier Items, um systematisch und gleichmäßig alle zentralen Inhaltsbereiche des Studiums abzudecken.

Diese systematisch erstellten Items ergänzten wir durch eine flexible Anzahl an Items zu den übergeordneten Arbeitsweisen und falschen Erwartungen, welche später in der Rückmeldung unter der Kategorie „Sonstige Studieninhalte“ und „Häufige Irrtümer“ zusammengefasst wurden.

2.2 Zuordenbarkeit der Items zu Oberkategorien

Zur Testung der Zuordenbarkeit der Items zu den erwarteten Studienfachbereichen wurden Bewertungen von zwei Expert*innen (Personen mit abgeschlossenem Bachelor-Psychologiestudium aus zwei verschiedenen deutschen Hochschulen) eingeholt. Die Expert*innen sollten alle in randomisierter Reihenfolge präsentierten Items jeweils genau einer Oberkategorie (einem Studienfachbereich) zuordnen. Im Anschluss daran wurde die Beurteilerübereinstimmung zwischen ihren Zuordnungen untereinander sowie im Vergleich zu unserer erwarteten Zuordnung mittels Cohens Kappa bestimmt. Die Beurteilerübereinstimmung zwischen den beiden Ratern lag bei $Kappa = .762$ ($p < .001$), 95% KI [.639, .885], was für eine substantielle Übereinstimmung spricht (Landis & Koch, 1977). Die Beurteilerübereinstimmung zwischen den jeweiligen Ratern und der erwarteten Zuordnung lag für beide Rater bei $Kappa = .848$ ($p < .001$), 95% KI [.744, .952], was gemäß den Konventionen von Landis und Koch (1977) einer fast perfekten Übereinstimmung entspricht. Im Anschluss wurden mit beiden Ratern etwaige Differenzen bei der Zuordnung besprochen und darauf basierend zwei Kategorien zusammengeführt, die sich nicht klar trennen ließen, sowie bestehende Items trennschärfer formuliert. Nach diesen Schritten können alle Items genau einem Studienfachbereich zugeordnet werden. Dies ist zum einen wichtig, damit Studieninteressierte durch das in Studienfachbereiche gruppierte Feedback der Items ein adäquates Bild von der Struktur des Studiums aufbauen können. Zum anderen ist es eine wichtige Grundlage, um im nächsten Schritt testen zu können, ob die Inhalte verschiedener Fachbereiche eines Studienfachs ausreichend und gleichmäßig vollständig sowie gleichmäßig prototypisch abgedeckt werden.

2.3 Vollständigkeit und Prototypizität der Items

Um die Vollständigkeit und Prototypizität der Items zu überprüfen, wurden im Rahmen eines Vortests 19 Personen² mit einem abgeschlossenem Bachelor-Psychologiestudium von sieben verschiedenen Hochschulen befragt.

2 Davon mussten zwei Personen ausgeschlossen werden, weil sie die Instruktionen missverstanden hatten und dadurch Antworten gegeben hatten, die logisch inkonsistent waren. Zwei weitere Personen wiesen jeweils auf zwei Variablen offensichtliche Tippfehler auf. Da die restlichen Daten dieser Perso-

Vollständigkeit. Um die Gesamtabdeckung der Items für einen bestimmten Studienfachbereich X zu erfassen, konnten die Absolvent*innen im Bereich von 0% bis 100% angeben, wieviel Prozent der Zeit, die man im Studium für Studienfachbereich X (z.B. Sozialpsychologie) hat, durch alle vier dargebotenen Studieninhalte gemeinsam abgedeckt werden.

Prototypizität. Die Prototypizität der Items für den jeweiligen Studienfachbereich wurde erfasst, indem gefragt wurde, wieviel Prozent der Zeit, die man im Studium für Studienfachbereich X hat, man sich mit den jeweiligen Studieninhalten beschäftigt (z.B. wie sich Einstellungen gegenüber anderen bilden und verändern können). Dabei wurde angemerkt, dass die Prozentzahlen der vier Items pro Studienfachbereich sich nicht zu 100% aufsummieren müssen. Sie können sowohl weniger als 100% ergeben – wenn einzelne Studieninhalte fehlen – oder auch mehr als 100% – wenn einzelne Studieninhalte sich überlappen und dadurch doppelt vorkommen.

Fehlende Inhalte. Zusätzlich dazu wurde bei Personen, die weniger als 100% bei der Gesamtabdeckung angaben, erfragt, welcher zentrale Studieninhalt des Studienfachbereichs X nicht durch die vier obenstehenden Studieninhalte erfasst wurde.

Ergebnis Vollständigkeit. Um zu testen, ob die verschiedenen Studienfachbereiche hinreichend vollständig abgedeckt wurden und ob sie sich in der Vollständigkeit ihrer Abdeckung voneinander unterscheiden, wurde eine ANOVA mit Messwiederholung mit Greenhouse-Geisser-Korrektur durchgeführt, in welcher die Vollständigkeitsangaben in Prozent zwischen den verschiedenen Studienfachbereichen verglichen wurden. Es zeigte sich, dass sich die Vollständigkeitsangaben nicht statistisch signifikant zwischen den einzelnen Studienfachbereichen unterschieden, $F(4.47, 62.62) = 1.60$, *n.s.* Zusätzlich zeigte sich deskriptiv in den *Vollständigkeitsratings über alle Studienfachbereiche hinweg*, dass alle Studienfachbereiche durchschnittlich zu mehr als 90% abgedeckt wurden ($M = 93.33\%$, $SD = 3.83\%$, Range: 87.94% – 99.12%). Das deutet darauf hin, dass die Inhalte aller Studienfachbereiche in ausreichendem Maße und ungefähr gleichmäßig vollständig abgedeckt wurden.

Ergebnis Prototypizität. Um zu testen, ob die Items für verschiedene Studienfachbereiche sich in der Prototypizität ihrer Formulierung unterscheiden, betrachteten wir den Mittelwert der jeweils zu einem Studienfachbereich gehörenden Itemprototypizitätsangaben. Eine ANOVA mit Messwiederholung mit Greenhouse-Geisser-Korrektur zeigte, dass sich die Mittelwerte der Itemprototypizitätsangaben pro Studienfachbereich statistisch signifikant voneinander unterschieden, $F(3.47, 52.06) = 2.82$, $p < .05$, partielles $\eta^2 = .158$. Deskriptive Statistiken zeigten zusätzlich, dass der Mittelwert der Itemprototypizitätsangaben (über alle Items hinweg) 27.04% betrug ($SD = 2.20\%$, Range: 24.07% (Biopsychologie) bis 31.62% (Statistik)). Das bedeutet, dass die Items insgesamt etwas zu allgemein formuliert wurden (Optimum bei vier Items pro Kategorie wären 25%) und konkret zum Beispiel die Items für den Studienfachbereich Biopsychologie tendenziell weniger prototypisch formuliert wurden als die Items für Statistik. Basierend auf diesen Ergebnissen wurde überprüft, ob Formulierungen bei den über-/unterrepräsentierten Fächern schwächer/stärker prototypisch formuliert werden konnten. Ziel war es, Überlappungen zwischen Iteminhalten zu verringern, eine gleichmäßigere Prototypizität der Items verschiedener Studienfächer zu erreichen und damit eine verzerrende Darstellung zu vermeiden.

nen logisch konsistent waren, wurden nur die Einzelangaben in Missings umcodiert, was bei einzelnen Analysen zu einer reduzierten Stichprobe führte.

Schließlich wurde die Verständlichkeit der Items für die Zielgruppe des E × I - Tests durch zwei Schüler*innen getestet. Dafür wurden den Schüler*innen die angepassten Items zur Beantwortung vorgelegt. Anschließend wurden sie im Interview zu Verständnisschwierigkeiten befragt und betreffende Fachbegriffe abgeändert. Der finale hochschulunabhängige Itemkatalog für das Bachelor-Psychologiestudium bestehend aus 61 Items (je 4 Items für 12 Studienfachbereiche, 7 Items zu Sonstigen Studieninhalten und 6 Items zu Häufigen Irrtümern) kann im elektronischen Supplementmaterial A eingesehen werden.

2.4 Erhebung und Validierung der Expert*inneneinschätzung

Für die Gewinnung der Expert*inneneinschätzung hinsichtlich der finalen Items wurde eine Mail an alle Professor*innen für Psychologie deutschlandweit über den DGPs-Hauptverteiler gesendet. Außerdem wurden Masterstudierende der Universität Mannheim und Freiburg befragt, welche zuvor das Bachelor-Psychologiestudium an verschiedenen Hochschulen erfolgreich abgeschlossen hatten. Insgesamt nahmen 149 Personen an der Befragung teil, davon 76 Studierende und 73 Dozierende (Akademische Mitarbeitende, Doktoranden, Postdoktoranden und Professor*innen), von insgesamt acht verschiedenen Hochschulen.

*Expert*inneneinschätzung.* Zur Erfassung des realen Ausmaßes der Inhalte wurden die Expert*innen gefragt, in welchem zeitlichen Ausmaß sich Studierende im Bachelor-Psychologiestudium mit den folgenden Studieninhalten und Arbeitsweisen beschäftigen, und im Anschluss die ausgewählten Items präsentiert. Die Expert*innen konnten dazu auf einer Skala von 1 (*gar nicht/in einem sehr geringen zeitlichen Ausmaß*) bis hin zu 7 (*in einem sehr großen zeitlichen Ausmaß*) ihre Einschätzung abgeben.

*Ergebnisse Expert*inneneinschätzung.* Deskriptive Statistiken zu den Ratings der Expert*innen sowie Übereinstimmungsmaße sind im elektronischen Supplementmaterial B zu finden. Die t-Tests zeigten einen statistisch signifikanten Unterschied zwischen den Studierenden- und Dozierendenangaben bei fünf der 15 untersuchten Dimensionen der Studienfachbereiche, übergeordneten Arbeitsweisen und falschen Erwartungen. Der Betrag der Effektstärken zeigte im Median kleine Effekte ($|d_{Md}| = 0.15$, $|d_R| = 0.00 - 0.63$). Die größten Unterschiede zeigten sich bei der Einschätzung des Ausmaßes, in dem das Bearbeiten von wissenschaftlichen Texten eine zentrale Arbeitsweise im Studium darstellt, während die kleinste Effektstärke bei der Einschätzung des zeitlichen Ausmaßes der Forschungsmethoden vorlag. Auf Basis dieser Ergebnisse wurden die Expert*inneneinschätzung dieser beiden Gruppen für die Rückmeldung zusammengefasst.

3 Beschreibung Assessment und Rückmeldung im E × I - Test

Der neu entwickelte Itemkatalog sowie die Expert*inneneinschätzungen können nun in ein Assessment- und Rückmeldungskonzept eingefügt werden, das ebenfalls neu erarbeitet wurde, um Interessen und Erwartungen kombiniert erfassen und rückmelden zu können. Das Assessment im E × I - Test besteht aus zwei Bestandteilen: einer Skala zur Erfassung der Erwartungen und einer Skala zur Erfassung der Interessen. Der Itemstamm für die Erfassung der Interessen lautet: „Wie sehr interessierst du dich dafür...“ [Items aus dem Itemkatalog,

z.B. ... wie sich Einstellungen gegenüber anderen bilden und verändern können.] Dazu passend kann auf einer siebenstufigen Likert-Skala von -3 (*überhaupt kein Interesse*) bis hin zu $+3$ (*sehr starkes Interesse*) graduell das persönliche Interesse angegeben werden³. Der Itemstamm für die Erfassung der Erwartungen lautet: „In welchem zeitlichen Ausmaß erwartest du, dich in deinem [Studiengang] damit zu beschäftigen, ...“ [Items aus dem Itemkatalog, z.B. ... wie sich Einstellungen gegenüber anderen bilden und verändern können.] und kann mit einer siebenstufigen Likert-Skala von 1 (*gar nicht/in einem sehr geringen zeitlichen Ausmaß*) über 4 (*in einem durchschnittlichen zeitlichen Ausmaß*) bis hin zu 7 (*in einem sehr großen zeitlichen Ausmaß*) beantwortet werden. Die beiden Skalen und die Items wurden so formuliert, dass sie es erlauben, Interessen und Erwartungen hinsichtlich derselben Items kombiniert abzufragen. Die Formulierung zur Erfassung der Erwartungen zielt außerdem klar auf das zeitliche Ausmaß eines spezifischen Inhalts im Studium ab und kann so weder von den Studieninteressierten noch von den Expert*innen mit einer Sicherheit bei der Beantwortung verwechselt werden.

Das Feedback im $E \times I$ -Test gibt Studieninteressierten Rückmeldung zu drei relevanten Indikatoren für deren Studienentscheidung: Interesse, Erwartungsdiskrepanz und Valenz der Erwartungsdiskrepanz. Die Rückmeldung zum Interesse verdeutlicht Studieninteressierten, wie gut die Inhalte des Studienfachs zu ihren Bedürfnissen passen, wobei Interesse an Studieninhalten zunächst per se positiv zu bewerten ist („Super, dass dich dieser Inhalt interessiert“), während überhaupt kein Interesse negativ ist („Schade, dass dich dieser Inhalt nicht interessiert“). Die Erwartungsdiskrepanzen geben an, wie akkurat ihr Bild vom jeweiligen Studienfach ist und ob weiterer Informationsbedarf besteht oder nicht. Die Valenz der Erwartungsdiskrepanz unterstützt Studieninteressierte dabei einzuschätzen, inwiefern sie ihre bisherige Einstellung zum Studium in positiver oder negativer Weise ändern sollten. Beispielsweise könnten Studieninteressierte, die sich für die Ursachen psychischer Störungen sehr stark interessieren, allerdings ein größeres zeitliches Ausmaß davon erwartet hätten, als von Expert*innen eingeschätzt wurde, folgende Rückmeldung erhalten: „Super, dass dich dieser Inhalt sehr stark interessiert. Leider wird dieser Inhalt in einem kleineren Ausmaß vorkommen, als du es erwartet hast.“ Diese negative Erwartungsdiskrepanz – die Erwartungen werden enttäuscht – sollte somit zu einer negativeren Einstellung zum Studium führen (z.B. geringere erlebte Passung, geringere Studienwahrscheinlichkeit). Wenn Studieninteressierte sich aber dafür interessieren, wie psychische Störungen beschrieben und klassifiziert werden können, und sie davon nur ein geringeres zeitliches Ausmaß erwartet hätten, verglichen mit der höheren Einschätzung der Expert*innen, erhalten sie folgende Rückmeldung: „Super, dass dich dieser Inhalt sehr stark interessiert. Erfreulicherweise wird dieser Inhalt in einem größeren Ausmaß vorkommen, als du es erwartet hast.“ Diese positive Erwartungsdiskrepanz – die Erwartungen werden übertroffen – sollte somit zu einer positiveren Einstellung zum Studium führen (z.B. höhere erlebte Passung, stärkere Studienwahrscheinlichkeit).

3 Die Entscheidung zwischen uni- und bipolaren Skalen kann nicht pauschal getroffen werden, sondern muss unter Berücksichtigung der zu erfassenden Konstrukte getroffen werden (Jonkisz, Moosbrugger & Brandt, 2012). Wir nehmen an, dass es Inhalte gibt, denen aufgrund eines starken Interesses ein positiver Wert beigemessen wird, sowie Inhalte, an denen überhaupt kein Interesse besteht, weswegen diese negativ bewertet werden. Um dieser Annahme methodisch gerecht zu werden, haben wir uns für ein bipolares Vorgehen bei Erfassung und Rückmeldung des Interesses entschieden.

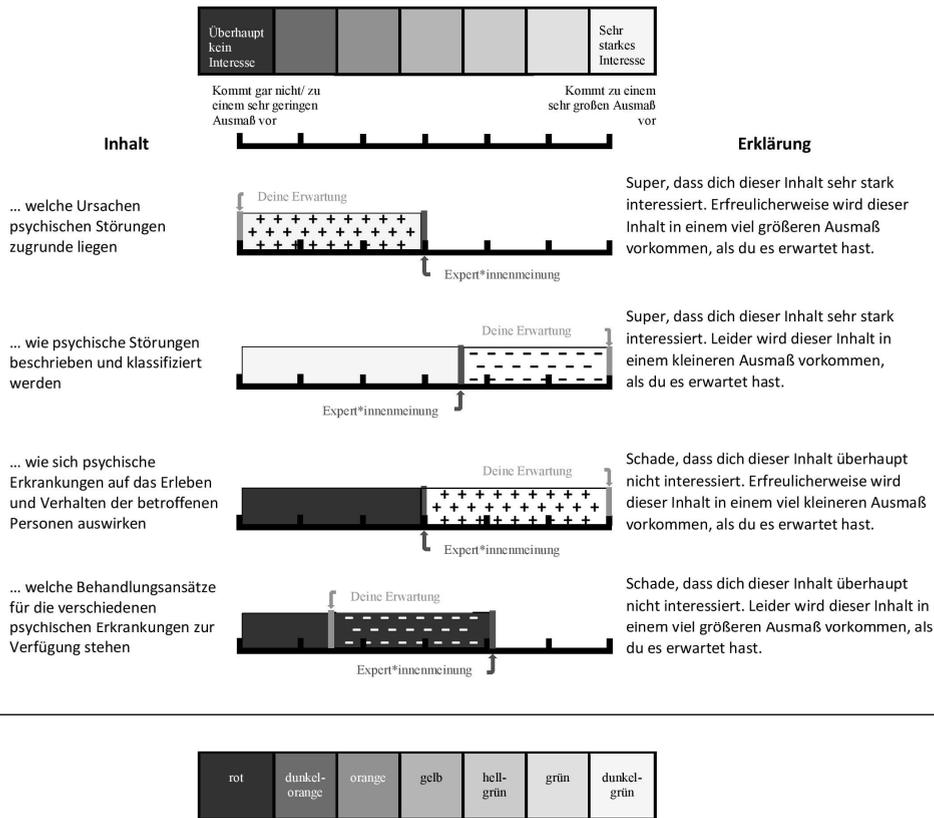


Abbildung 1: Ausschnitt eines Feedbacks zum Interesse, zur Erwartungsdiskrepanz und zur Valenz der Erwartungsdiskrepanz aus dem Erwartungs- und Interessenstest (E × I - Test). Farbe des Balkens: Interesse der Nutzer*innen für einen bestimmten Inhalt. Füllung des Balkensbereichs „+“: positive Valenz der Erwartungsdiskrepanz (Erwartungen übertroffen); „-“ negative Valenz der Erwartungsdiskrepanz (Erwartungen enttäuscht); „°“ Neutrale Valenz der Erwartungsdiskrepanz.

Diese drei Indikatoren werden zunächst sowohl graphisch als auch verbal auf Studienfachbereichsebene gemittelt rückgemeldet, um Studieninteressierten einen Überblick zu geben und die Struktur des Studiums zu verdeutlichen. Mit Klicken auf die Studienfachbereiche wird jeweils auch spezifisch auf Ebene einzelner Items eine Rückmeldung gegeben, um falsche Erwartungen an spezifische Inhalte konkret zu korrigieren. Ein beispielhafter Ausschnitt aus einem Feedback ist in Abbildung 1 dargestellt (für die originale farbige Darstellung siehe das elektronische Supplementmaterial C). Die Farbe des Balkens zeigt an, wie sehr sich die Nutzer*innen für einen bestimmten Inhalt interessieren: Von Rot (*überhaupt kein Interesse*) über Gelb (*mittelmäßiges Interesse*) bis Grün (*sehr starkes Interesse*). Diese Farbcodierung wurde gewählt, da mehr Interesse an Studieninhalten unabhängig von den Erwartungen per se besser ist als weniger Interesse und „Grün“ als Farbe eher positive Assoziationen mit sich zieht, während „Rot“ eher eine Warn- und Signalfarbe darstellt. Die Erwartung der Studieninteressierten hinsichtlich des zeitlichen Ausmaßes beziehungsweise die Expert*inneneinschätzung werden auf der x-Achse abgetragen, von links (*kommt gar nicht/zu einem sehr*

geringen Ausmaß vor) bis rechts (kommt zu einem sehr großen Ausmaß vor). Dadurch wird die Differenz zwischen der Erwartung der Studieninteressierten und der Expert*inneneinschätzung veranschaulicht. Die Füllung des Balkenbereichs dieser Differenz zeigt zusätzlich an, inwiefern Erwartungen übertroffen (positive Valenz der Erwartungsdiskrepanz symbolisiert durch „+“) beziehungsweise enttäuscht (negative Valenz der Erwartungsdiskrepanz symbolisiert durch „-“) wurden, oder ob die Erwartungsdiskrepanz aufgrund eines mittelmäßigen Interesses neutral zu bewerten ist („o“).

Zum Schluss folgt eine verbale Rückmeldung auf gesamter Studienfachebene (über die einzelnen Items und Studienfachbereiche hinweg) bestehend aus dem Gesamtinteresse an allen realen Inhalten (durchschnittliches Interesse der Person, die den $E \times I$ - Test bearbeitet hat, an allen Studieninhalten und Arbeitsweisen exklusive Häufige Irrtümer), dem Gesamtbetrag der Erwartungsdiskrepanz (durchschnittlicher Betrag der Erwartungsdiskrepanzwerte der Person, die den $E \times I$ - Test bearbeitet hat) sowie der Gesamtvalenz der Erwartungsdiskrepanz (durchschnittliche Valenz der Erwartungsdiskrepanzwerte der Person, die den $E \times I$ - Test bearbeitet hat), um den Studieninteressierten für ihre Studienwahlentscheidung eine Zusammenfassung der drei genannten Indikatoren zu liefern. Die exakten Berechnungsformeln und zugehörigen verbalen Feedbacktexte befinden sich im elektronischen Supplementmaterial D. Der gesamte $E \times I$ - Test für Psychologie kann selbstgesteuert, ohne zusätzliche Beratung, kostenlos und online an einem Computer bearbeitet werden, da die Erfassung, Auswertung und Rückmeldung automatisch erfolgen. Die Bearbeitungszeit des $E \times I$ - Tests für Psychologie ist für ungefähr 20 Minuten angesetzt.

4 Erste Befunde zur Evaluation des Erwartungs- und Interessenstests ($E \times I$ - Tests)

Der $E \times I$ - Test wurde sowohl spezifisch als auch allgemein im Rahmen der Bearbeitung des gesamten OSA-Psych evaluiert, welches neben dem $E \times I$ - Test noch einen Fähigkeitstest sowie Informationen zum Bachelor-Psychologiestudium zur Verfügung stellt.

Innerhalb des ersten halben Jahres nach Freishaltung im Februar 2020 bearbeiteten 2023⁴ Studieninteressierte das OSA-Psych, nahmen an der freiwilligen Evaluation teil und stimmten dabei der Speicherung ihrer Daten zu (81.96% weiblich, 17.48% männlich, 0.56% divers, $M_{\text{age}} = 20.0$ Jahre, Range: 16 – 61). 93.29% der Teilnehmenden wollten noch im selben Jahr mit einem Studium beginnen, 5.88% im darauffolgenden Jahr und die restlichen 0.83% in den darauffolgenden Jahren. 79.61% der Studieninteressierten waren sehr sicher, dass sie sich auf das Bachelor-Psychologiestudium bewerben möchten, 18.22% tendierten dazu, sich für ein Bachelor-Psychologiestudium zu bewerben, waren sich aber weniger sicher, 1.60% waren unsicher und 0.57% tendierten sogar eher dazu, sich nicht für ein Bachelor-Psychologiestudium zu bewerben. Somit waren Studieninteressierte aus verschiedenen Phasen im Entscheidungsprozess für ein Bachelor-Psychologiestudium vertreten.

Akzeptanz. In Anlehnung an die Items von Schueller (2011) konnten die Studieninteressierten direkt im Anschluss an die Bearbeitung des $E \times I$ - Tests bewerten, ob sie vom $E \times I$ -

4 Da es in der freiwilligen Befragung für einzelne Items keinen Antwortzwang gab, variierte die Stichprobengröße pro Item zwischen $n = 1876$ und $n = 1985$.

Test profitiert haben, ob ihnen der E × I - Test gefallen hat und wie schwierig sie ihn fanden. Hinsichtlich der Rückmeldung wurden die Studieninteressierten zunächst allgemein gefragt, wie verständlich und übersichtlich sie die Rückmeldung im E × I - Test empfanden. Im Anschluss konnten sie konkret hinsichtlich der neuen Features bewerten, ob sie das Gesamtfeedback als sinnvolle Ergänzung zum inhaltspezifischen Feedback empfanden und ob sie die Rückmeldung der positiven und negativen Erwartungsdiskrepanz als sinnvolle Ergänzung zur separaten Rückmeldung von Interessen und Erwartungen empfanden. Alle Aussagen wurden auf einer Skala von 1 (*trifft überhaupt nicht zu*) bis 7 (*trifft vollkommen zu*) bewertet.

Informiertheit. Sowohl vor als auch nach der Bearbeitung des OSA-Psych wurden die Studieninteressierten gefragt, wie gut sie sich hinsichtlich der Inhalte und Arbeitsweisen im Bachelor-Psychologiestudium informiert fühlen, was sie auf einer Skala von 1 (*trifft überhaupt nicht zu*) bis 7 (*trifft vollkommen zu*) beantworten konnten.

Weiterempfehlung. Außerdem wurden die Studieninteressierten nach der Bearbeitung des gesamten OSA-Psych gefragt, ob sie dieses weiterempfehlen würden, was sie bejahen oder verneinen konnten.

Wir testeten alle Akzeptanzratings gegen den Mittelpunkt der Skala als Maß für ein durchschnittliches Rating. Alle folgenden Aspekte des Akzeptanzratings lagen signifikant über dem Skalenmittelwert ($p < .001$, $d_{Md} = 1.86$). Konkret zeigt sich in den deskriptiven Daten, dass insgesamt 83.64% der Nutzer*innen bei der Frage nach dem Profitieren eine Antwortoption über der Mittelkategorie ankreuzten und somit angaben, vom E × I - Test profitiert zu haben, 90.39% hat der E × I - Test gefallen und 92.29% der Nutzer*innen empfanden ihn als nicht schwierig. Das neu entwickelte Feedback empfanden 96.73% der Nutzer*innen als verständlich und 93.47% als übersichtlich. Das Gesamtfeedback empfanden 90.60% der Nutzer*innen als sinnvolle Ergänzung zum inhaltspezifischen Feedback. 92.86% der Nutzer*innen empfanden die neu ergänzte Rückmeldung der Valenz der Erwartungsdiskrepanz als hilfreiche Ergänzung zur separaten Rückmeldung von Interessen und Erwartungen.

Hinsichtlich der Wissensänderung zeigte sich, dass die selbst eingeschätzte Informiertheit der Studieninteressierten über die Inhalte und Arbeitsweisen im Bachelor-Psychologiestudium von vor der Bearbeitung des OSA-Psych ($M = 5.26$, $SD = 1.24$) im Vergleich zu nach der Bearbeitung des OSA-Psych ($M = 5.81$, $SD = .95$) signifikant zunahm, $t(1911) = -18.46$, $p < .001$, $d = .42$.

Bei bisherigen OSAs lag die Weiterempfehlungsrate bei 76% (Sonnleitner, Kubinger & Frebort, 2009) bis hin zu 90% (was-studiere ich.de). Die Weiterempfehlungsrate für das gesamte OSA-Psych lag signifikant über der 90-Prozent-Rate, $t(1984) = 7.52$, $p < .001$, $d = .17$, da insgesamt 94.01% der Teilnehmenden angaben, dass sie die Teilnahme am OSA-Psych weiterempfehlen würden. Ein Ausblick für eine Erweiterung der Evaluation auf Studienerfolgsindikatoren (Wohlbefinden, Leistung, Studienabbruch) wird in der Diskussion gegeben.

5 Diskussion

Mit dem E × I - Test stellen wir einen neuen Erwartungs- und Interessenstest zur Verfügung, welcher in dreierlei Hinsicht innovativ ist. Der E × I - Test ermöglicht erstmals eine Erfassung und Rückmeldung von Erwartungsdiskrepanzen und Interessen in Kombination und

kann dementsprechend auch messen und rückmelden, ob die Erwartungen an ein Studienfach übertroffen oder enttäuscht wurden, was für die Studienwahlentscheidung relevant sein sollte (Hasenberg, 2012; Karst et al., 2017). Außerdem wurde bei der Entwicklung und Validierung der Items für den $E \times I$ -Test mithilfe neuer Verfahren sichergestellt, dass die neuen Items die Struktur des Studienfachs angemessen widerspiegeln und ausreichend vollständig und annähernd gleichmäßig abdecken, wodurch ein möglichst akkurates Bild vom Studium aufgebaut werden sollte. Schließlich wurde der fachspezifische $E \times I$ -Test für das Bachelor-Psychologiestudium hochschulunabhängig konzipiert und ist damit im Gegensatz zu hochschulspezifischen Tests besser geeignet für den großen Anteil an Studieninteressierten, die sich zuerst für ein Fach und im Anschluss für eine spezifische Hochschule entscheiden (Hovestadt & Stegelmann, 2011). Erste Evaluationsergebnisse zeigen, dass Studieninteressierte das neue Verfahren akzeptieren. Durch die Bearbeitung steigt die Informiertheit der Studieninteressierten und die Weiterempfehlungsraten liegen über den Weiterempfehlungsraten bekannter Verfahren. Da fehlende Informiertheit, zu geringes Interesse und falsche Erwartungen an das Studienfach zentrale Gründe für Studienabbruch darstellen (Heublein et al., 2010; Schiefele et al., 2007; Schmidt-Atzert, 2005), birgt der $E \times I$ -Test somit großes Potential, Studienerfolg langfristig zu fördern.

5.1 Limitationen und zukünftige Forschung

Unsere Ergebnisse sind dergestalt limitiert, dass beim Itementwicklungs- und Itemvalidierungsprozess hinsichtlich der Zuordenbarkeit, Vollständigkeit und Prototypizität der Items jeweils kleine Expert*innenstichproben eingeholt wurden und die auf Basis dieser Ergebnisse angepassten Items nicht mithilfe einer zweiten Expert*innenstichprobe überprüft wurden. Trotz dieser Einschränkungen ist darauf hinzuweisen, dass die begangenen Schritte zur Validierung der Items bereits innovative Wege darstellen, die über bisherige Validierungsprozesse hinausgehen. In Zukunft wäre allerdings eine zweite Validierungsrunde mit einer größeren Expert*innenstichprobe anzustreben.

Zusätzlich dazu bleibt auch bei den Expert*inneneinschätzungen hinsichtlich des zeitlichen Ausmaßes der Studieninhalte offen, ob trotz der kleinen Unterschiede eine der beiden Expert*innengruppen (Studierende vs. Dozierende) eine relevantere Einschätzung für die Studieninteressierten liefert. Man könnte argumentieren, dass es die Einschätzung der Dozierenden ist, da diese mit ihrer langjährigen Erfahrung die Prozesse an der Hochschule besser kennen als die Studierenden. Dem steht entgegen, dass Studierende noch einen besseren Überblick über das Gesamtstudium haben, da bei ihnen das Studium noch nicht so lange zurückliegt wie bei den Dozierenden. Dafür sollte in zukünftigen Studien geprüft werden, ob es Unterschiede in der Vorhersagekraft der berechneten Erwartungsdiskrepanz für späteren Studienerfolg gibt, in Abhängigkeit davon, welche Referenzgruppe (Studierende vs. Dozierende) für die Berechnung verwendet wird. Zusätzlich dazu schwanken die Einschätzungen auch zwischen verschiedenen Hochschulen aufgrund von unterschiedlichen Schwerpunktsetzungen. Dadurch, dass die Expert*innenmeinung aus dem Mittel der Einschätzung von Expert*innen verschiedener Hochschulen bestehen, sollten allerdings standortspezifische Besonderheiten nicht stark ins Gewicht fallen, und eine Hochschulunabhängigkeit sollte gewährleistet sein. Eine alternative Möglichkeit wäre es, gemäß der laut DGPs vorgeschlagenen ECTS zu gewichten. Eine solche Gewichtung nach ECTS zeigte allerdings in unserer Stich-

probe deutliche Abweichungen von einer Gewichtung basierend auf den Expert*inneneinschätzungen hinsichtlich des zeitlichen Ausmaßes der Inhalte. Demnach gilt es zunächst in zukünftiger Forschung zu klären, woher diese Unterschiede kommen und welcher Indikator Studienerfolg besser vorhersagen kann. Außerdem könnte zukünftige Forschung noch berücksichtigen, ob ein Inhaltsbereich fakultativ oder obligatorisch ist, Gegenstand einer Prüfung ist oder wie die Inhalte in der Abschlussnote gewichtet sind. Bis dahin bleiben wir bei den Expert*inneneinschätzungen, da wir davon ausgehen, dass diese näher an der späteren subjektiven Wahrnehmung der Nutzer*innen sein werden und damit prädiktiver für Studienzufriedenheit als das vollkommen objektive Maß der ECTS-Punkte oder andere Kriterien.

Der Fokus auf den zeitlichen Umfang verschiedener Studieninhalte bei der Erfassung von Erwartungen bedingt allerdings, dass keine Erwartungen hinsichtlich etwaigen Anforderungen im Studium oder späteren Berufsmöglichkeiten erfasst werden, wie in anderen Erwartungschecks möglich (z.B. *Study Finder Erwartungscheck*). Um auch diese wichtigen Informationen zu geben, empfehlen wir zusätzlich zum E × I - Test Fähigkeitstests einzusetzen, damit Studieninteressierte abgleichen können, ob sie den Anforderungen des Studiums gewachsen sind, sowie Informationen zu den beruflichen Perspektiven nach dem Studium anzubieten, wie es auch im OSA-Psych gehandhabt wird. Schließlich sei darauf hingewiesen, dass die Evaluation des E × I - Tests hinsichtlich der Informiertheit und Weiterempfehlungsrate im Rahmen der Bearbeitung des gesamten OSA-Psych stattfand. Da im OSA-Psych weitere Informationen zur optimalen Nutzung zur Verfügung gestellt werden (Fähigkeitstest und Informationen zum Studium), können die Ergebnisse zur Informiertheit und Weiterempfehlungsrate nicht ausschließlich auf den E × I - Test allein zurückgeführt werden. Es gibt aber gute Gründe, den E × I - Test als den maßgeblichen Treiber dieser Effekte zu interpretieren. Zum einen schnitt der E × I - Test bei den Akzeptanzratings, die spezifisch sowohl für den E × I - Test als auch für den Fähigkeitstest erfasst wurden, besser ab als der Fähigkeitstest. Zum anderen musste der E × I - Test bearbeitet werden, bevor die Studieninteressierten ihr Teilnahmezertifikat erhielten und an der Nachbefragung teilnehmen konnten. Das Anschauen der Informationsseiten hingegen war freiwillig und stand damit weniger im Fokus der Studieninteressierten, wie sich an den offenen Kommentaren der Studieninteressierten ablesen ließ, welche sich überwiegend auf die beiden Testverfahren bezogen. Dementsprechend ist es plausibel, die positiven Ergebnisse hinsichtlich der Steigerung der Informiertheit sowie der sehr hohen Weiterempfehlungsrate maßgeblich auf den E × I - Test zurückzuführen.

Zukünftige Studien sollten dennoch für eine weiterführende Evaluation die Ergebnisse zur Informiertheit und Weiterempfehlungsrate spezifisch für den E × I - Test replizieren. Außerdem sollte getestet werden, inwiefern die Rückmeldung des Interesses und der Valenz der Erwartungsdiskrepanz die Einstellung der Studieninteressierten gegenüber dem Studium ändert (z.B. Passung, Studienwahlsicherheit) sowie die letztliche Studienwahlentscheidung beeinflusst. Schließlich sollte überprüft werden, ob die Indikatoren des E × I - Tests (Interessen sowie enttäuschte beziehungsweise übertroffene Erwartungen) in erwarteter Weise mit Studienerfolg in Form von Wohlbefinden und Leistung im Studium in Zusammenhang stehen. Dabei wäre es lohnend, den neuen Indikator der Valenz der Erwartungsdiskrepanz hinsichtlich seiner prädiktiven Validität mit anderen möglichen Indikatoren zu vergleichen (z.B. der Diskrepanz zwischen den Interessen und Erwartungen der Studieninteressierten oder der Diskrepanz zwischen den Interessen und der Expert*inneneinschätzung der faktischen Realisation der Studieninhalte). Die Untersuchung dieser zentralen theoretischen wie praktischen Fragestellungen wird durch den neuen E × I - Test erst möglich gemacht.

5.2 Praktische Implikationen

Der $E \times I$ - Test wird bereits von zahlreichen Studieninteressierten für das Bachelor-Psychologiestudium im Rahmen des OSA-Psych bearbeitet (siehe <https://www.osa-psych.de/>).

Basierend auf den ersten Evaluationsbefunden werden wir den $E \times I$ - Test für das Bachelor-Psychologiestudium bundesweit weiter bekannt machen. Dafür wird unter anderem Search Engine Optimization eingesetzt, in bundesweiten Studienorientierungsmaterialien (u.a. auch auf der Website der Deutschen Gesellschaft für Psychologie) über das Verfahren informiert sowie eine sinnvolle Verzahnung mit anderen OSAs (z.B. was-studiere-ich.de) angestrebt.

Zum anderen können auf Basis des hier vorgestellten $E \times I$ - Tests für das Bachelor-Psychologiestudium sowie der hier vorgestellten Validierungsprozesse in Zukunft auch weitere $E \times I$ - Tests entwickelt werden. Darunter könnten zunächst $E \times I$ - Tests für die Orientierung in andere Studienfächer fallen oder in einem breiteren Kontext auch $E \times I$ - Tests für weitere Karriereentscheidungen. Allerdings gibt es erste Hinweise, dass sich Prädiktoren für Studienabbruch in verschiedenen Studienfächern (z.B. Sozialwissenschaften vs. MINT Fächern) unterschiedlich auswirken können (Fischer et al., 2020). Demnach sollte durch zukünftige Forschung geklärt werden, welche Mechanismen diese Unterschiede zwischen den Studienfächern erklären können, um das Verfahren noch zielgerichteter einsetzen zu können.

Die Entwicklung von hochschulunabhängigen Verfahren ist dabei für alle Inhalte erstrebenswert, die einem Studienfach über verschiedene Hochschulen hinweg gemein sind. Im Anschluss an hochschulunabhängige Verfahren kann je nach Bedarf zusätzlich noch auf hochschulspezifische OSAs verwiesen werden, um bei der Entscheidung für einen konkreten Studienort auch standortspezifische Unterschiede in den Blick zu nehmen.

5.3 Schlussfolgerung

Insgesamt steht mit dem $E \times I$ - Test ein neues Verfahren zur Verfügung, das Erwartungen und Interessen zum ersten Mal in Kombination betrachtet und damit auch enttäuschte und übertroffene Erwartungen erfassen und rückmelden kann. Items für die Inhalte des Bachelor-Psychologiestudiums wurden unabhängig von hochschulspezifischen Besonderheiten entwickelt und validiert. Items für andere Studienfächer oder Berufe könnten in Zukunft auf Basis des vorgestellten Verfahrens systematisch entwickelt und validiert werden.

Supplementmaterial

Elektronisches Supplementmaterial kann unter folgendem Link abgerufen werden:

https://osf.io/6pk9b/?view_only=ed95334e5a28403381fcbf619753a2e6

oder

https://is.gd/emali_test

Literatur

- Burkhardt, B. & Hagemeyer, C. (2018). *OSA- wie „ohne sichere Aussagekraft“? Ein systematisches-Review zur prädiktiven Aussagekraft von Online-Self-Assessment für den Studienerfolg*. Technische Universität Dresden, Dresden. Verfügbar unter: <https://nbn-resolving.org/urn:nbn:de:bsz:14-qucosa-234589>
- Cable, D. M. & DeRue, D. S. (2002). The convergent and discriminant validity of subjective fit perceptions. *Journal of Applied Psychology*, 87(5), 875–884. <https://doi.org/10.1037/0021-9010.87.5.875>
- Fischer, V., Walpuski, M., Lang, M., Letzner, M., Manzel, S., Motté, P., ... Leutner, D. (2020). Was beeinflusst die Entscheidung zum Studienabbruch? Längsschnittliche Analysen zum Zusammenspiel von Studienzufriedenheit, Fachwissen und Abbruchintention in den Fächern Chemie, Ingenieur- und Sozialwissenschaften. *Zeitschrift für empirische Hochschulforschung*, 4(1), 5580. <https://doi.org/10.3224/zehf.v4i1.05>
- Frankenberg, P. (2008). Profil und Passung – Hochschulpolitische Thesen zur Studierendenauswahl. In H. Schneider, K. Haase, H.-J. Didi & C. Heine (Hrsg.), *Profil und Passung. Studierendenauswahl in einem differenzierten Hochschulsystem* (HIS:Forum Hochschule 14/2008) (S. 9–15). Hannover: HIS. Verfügbar unter: https://www.dzhw.eu/pdf/pub_fh/fh-200814.pdf
- Frebort, M. & Kubinger, K. D. (2008). Qualitätsansprüche an ein Self-Assessment zur Studienwahlberatung: Der Wiener Ansatz. In H. Schuler & B. Hell (Hrsg.), *Studierendenauswahl und Studienentscheidung* (S. 95–101). Göttingen: Hogrefe.
- Hartig, J., Frey, A. & Jude, N. (2012). Validität. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (2. Auflage) (S. 143–171). Berlin: Springer. https://doi.org/10.1007/978-3-642-20072-4_7
- Hasenberg, S. (2012). *Zur prädiktiven Validität von Self-Assessments für die Studienzufriedenheit*. Dissertation, Philipps-Universität Marburg. Verfügbar unter: <https://archiv.ub.uni-marburg.de/diss/z2013/0050/>
- Hasenberg, S. & Schmidt-Atzert, L. (2013). Die Rolle von Erwartungen zu Studienbeginn: Wie bedeutsam sind realistische Erwartungen über Studieninhalte und Studienaufbau für die Studienzufriedenheit? *Zeitschrift für Pädagogische Psychologie*, 27(1–2), 87–93. <https://doi.org/10.1024/1010-0652/a000091>
- Hasenberg, S. & Schmidt-Atzert, L. (2014). Internetbasierte Selbsttests zur Studienorientierung. *Beiträge zur Hochschulforschung*, 36(1), 8–28.
- Hell, B. (2009). Selbsttests zur Studienorientierung: Nützliche Vielfalt oder unnützer Wildwuchs? In G. Rudinger & K. Hörsch (Hrsg.), *Self-Assessment an Hochschulen: Von der Studienfachwahl zur Profilbildung* (Bd. 4) (S. 9–19). Göttingen: V&R unipress.
- Heublein, U., Hutzsch, C., Schreiber, J., Sommer, D. & Besuch, G. (2010). *Ursachen des Studienabbruchs in Bachelor- und in herkömmlichen Studiengängen: Ergebnisse einer bundesweiten Befragung von Exmatrikulierten des Studienjahres 2007/08* (HIS:Forum Hochschule 2/2010). Hannover: HIS. Verfügbar unter: https://www.dzhw.eu/pdf/21/studienabbruch_ursachen.pdf
- Heukamp, V., Putz, D., Milbradt, A. & Hornke, L. F. (2009). Internetbasierte Self-Assessments zur Unterstützung der Studienentscheidung. *Zeitschrift für Beratung und Studium*, 4(1), 2–8.
- Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3. Auflage). Odessa, USA: Psychological Assessment Resources.
- Hovestadt, G. & Stegelmann, P. (2011). *EDU-CON Studie 2011 Internetnutzung für Studieninformation und Studienentscheidung*. Verfügbar unter: http://edu-con.de/download_files/11-12-20-Aufsatz_final_kurz.pdf
- Janke, S., Messerer, L., Merkle, B. & Krille, C. (2021). STUWA: Ein multifaktorielles Inventar zur Erfassung von Studienwahlmotivation. *Zeitschrift für Pädagogische Psychologie*, 1-17. <https://doi.org/10.1024/1010-0652/a000298>

- Jonkisz, E., Moosbrugger, H. & Brandt, H. (2012). Planung und Entwicklung von Tests und Fragebogen. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (2. Auflage) (S. 27–74). Berlin: Springer. https://doi.org/10.1007/978-3-642-20072-4_3
- Karst, K., Ertelt, B.-J., Frey, A. & Dickhäuser, O. (2017). Studienorientierung durch Self-Assessments: Veränderung von Einstellungen zum Studienfach während der Bearbeitung eines Selbsttests. *Journal für Bildungsforschung Online*, 9, 205–227. <https://doi.org/10.25656/01:14935>
- Kegel, L. S., Schnettler, T., Scheunemann, A., Bülke, L., Thies, D. O., Dresel, M., Fries, S., Leutner, D., Wirth, J. & Grunschel, C. (2020). Unterschiedlich motiviert für das Studium: Motivationale Profile von Studierenden und ihre Zusammenhänge mit demografischen Merkmalen, Lernverhalten und Befinden. *Zeitschrift für empirische Hochschulforschung*, 4(1), 81–105. <https://doi.org/10.3224/zehf.v4i1.06>
- Kossner, C. (2015). *Das Mannheimer Informationssystem für Studieninteressierte der Sozialwissenschaften 2.0* (Unveröffentlichte Masterarbeit). Universität Mannheim, Mannheim.
- Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. <https://doi.org/10.2307/2529310>
- Leonhart, R. (2017). *Lehrbuch Statistik: Einstieg und Vertiefung* (4. Auflage). Bern: Hogrefe.
- Schiefele, U., Streblov, L. & Brinkmann, J. (2007). Aussteigen oder Durchhalten. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 39(3), 127–140. <https://doi.org/10.1026/0049-8637.39.3.127>
- Schmidt-Atzert, L. (2005). Prädiktion von Studienerfolg bei Psychologiestudenten. *Psychologische Rundschau*, 56(2), 131–133. <https://doi.org/10.1026/0033-3042.56.2.131>
- Schueller, S. M. (2011). *Creating a recommendation framework for positive psychology exercises: The Netflix model of positive psychology* (Dissertation). Verfügbar unter: <https://repository.upenn.edu/edissertations/392>
- Sindern, E. & Pietrangeli, S. (2009). Validitätsforschung im Zuge der Entwicklung fachspezifischer Online Self-Assessments an der Rheinischen Friedrich-Wilhelms-Universität Bonn. In G. Rudinger, K. Hörsch & T. Krüger (Hrsg.), *Forschung und Beratung – Das Zentrum für Evaluation und Methoden* (S. 55–66). Göttingen: V&R unipress.
- Sonnleitner, P., Kubinger, K. D. & Frebort, M. (2009). Das Wiener Self-Assessment Psychologie mit seinen Verfahren der experimentalpsychologischen Verhaltensdiagnostik zur Messung studienfachübergreifender Soft-Skills. In G. Rudinger & K. Hörsch (Hrsg.), *Self-Assessment an Hochschulen. Von der Studienfachwahl zur Profilbildung* (Bd. 4) (S. 63–72). Göttingen: V&R unipress.
- Statistisches Bundesamt. (2020, Dezember 11). *Anzahl der Studierenden an Hochschulen in Deutschland in den Wintersemestern von 2002/2003 bis 2020/2021* [Graph]. Statista. Verfügbar unter <https://de.statista.com/statistik/daten/studie/221/umfrage/anzahl-der-studenten-an-deutschen-hochschulen/>
- Stoll, G. & Spinath, F. (2015). Unterstützen Self-Assessments die Studienfachwahl? Erfahrungen und Befunde aus dem Projekt Study-Finder. In A. Hanft, O. Zawacki-Richter & W. B. Gierke (Hrsg.), *Herausforderung Heterogenität beim Übergang in die Hochschule* (S. 113–131). Münster: Waxmann.
- Störk, L. & Mocigemba, D. (2013). Kommunizieren statt Testen. Die Online-Studienwahl-Assistenten der Universität Freiburg. In C. Bremer & D. Krömker (Hrsg.), *E-Learning zwischen Vision und Alltag: Zum Stand der Dinge* (S. 230–240). Münster: Waxmann. <https://doi.org/10.25656/01:10739>
- Thiele, L. & Kauffeld, S. (2019). Online Self-Assessments zur Studien- und Universitätswahl. In S. Kauffeld & D. Spurk (Hrsg.), *Handbuch Karriere und Laufbahnmanagement* (S. 109–132). Berlin: Springer. https://doi.org/10.1007/978-3-662-48750-1_4
- Vent, S. & Erdfelder, E. (2009). Das Mannheimer Informationssystem für Studieninteressierte der Sozialwissenschaften: Führt ein Erwartungstest als Self-Assessment tatsächlich zu korrekteren Erwartungen bei Studienanfängern? In G. Rudinger & K. Hörsch (Hrsg.), *Self-Assessment an Hoch-*

- schulen: Von der Studienfachwahl zur Profilbildung* (Bd. 4) (S. 99–110). Göttingen: V&R uni-press.
- Westermann, R. & Heise, E. (2018). Studienzufriedenheit. In D. H. Rost, J. R. Sparfeldt & S. R. Buch (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (5. Auflage) (S. 818–825). Weinheim: Beltz.
- Wiegand, J. P., Drasgow, F. & Rounds, J. (2021). Misfit matters: A re-examination of interest fit and job satisfaction. *Journal of Vocational Behavior*, 125, 1–17. <https://doi.org/10.1016/j.jvb.2020.103524>
- Wolter, A. (2008). Hochschulzugang im differenzierten Hochschulsystem – Von der Studienberechtigung zur individuellen Kompetenzmessung? In C. Heine, H.-J. Didi, K. Haase & H. Schneider (Hrsg.), *Profil und Passung Studierendenauswahl in einem differenzierten Hochschulsystem* (HIS: Forum Hochschule 14/2008) (S. 16–34). Hannover: HIS. Verfügbar unter: https://www.dzhw.eu/pdf/pub_fh/fh-200814.pdf

Kontakt

Belinda Merkle · Prof. Dr. Oliver Dickhäuser
Universität Mannheim
Lehrstuhl für Pädagogische Psychologie
A5, 6
68159 Mannheim
E-Mail: bmerkle@mail.uni-mannheim.de
E-Mail: oliver.dickhaeuser@uni-mannheim.de

Moritz Schiltenwolf
Eberhard Karls Universität Tübingen
Allgemeine Psychologie, Psychologisches Institut
Schleichstraße 4
72076 Tübingen
E-Mail: Moritz.Schiltenwolf@uni-tuebingen.de

Prof. Dr. Andrea Kiesel
Albert-Ludwigs-Universität Freiburg
Institut für Psychologie, Allgemeine Psychologie
Engelbergerstraße 41
79085 Freiburg
E-Mail: Andrea.Kiesel@psychologie.uni-freiburg.de