



Evaluation von Online Self-Assessments zur Studienwahl – Hindernisse, Chancen und Praxisbeispiele

Pascale Stephanie Petri¹  und Martin Kersting² 

¹Fakultät FWR Wirtschaftspsychologie, Hochschule für Wirtschaft und Umwelt Nürtingen-Geislingen, Geislingen, Deutschland

²Psychologische Diagnostik, Justus-Liebig-Universität Gießen, Deutschland

Zusammenfassung: Die Studienwahl ist die erste Herausforderung zu Beginn einer akademischen Laufbahn. Ist die Wahl getroffen, müssen sich Studieneinsteigerinnen und Studieneinsteiger in der neuen akademischen Umwelt zurechtfinden und den Anforderungen ihres Studienfachs gerecht werden. In keiner anderen Phase des Studiums brechen so viele Studentinnen und Studenten ab. Zu den Hauptgründen dafür zählen Leistungsdefizite und Motivationsmangel. Diese wiederum rühren nicht selten daher, dass Studieneinsteigerinnen und Studieneinsteiger von den tatsächlichen Studieninhalten überrascht werden. Daraus können Überforderung und Enttäuschung resultieren. Viele Hochschulen versuchen diese Diskrepanz zwischen Studiererwartungen und Studienwirklichkeit zu minimieren, indem sie Studieninteressierten die Möglichkeit bieten, sich im Rahmen von Online Self-Assessments (OSA) vorab intensiv mit ihrem Wunschstudienfach auseinander zu setzen. Bietet eine Hochschule dieses Instrument an, so sollte es evaluiert werden. Je nach Evaluationszeitpunkt und -zweck bieten sich verschiedene Ansätze an. In diesem Beitrag präsentieren wir Beispiele dafür, wie eine Evaluation gestaltet werden kann, wenn – wie in den meisten Fällen – keine direkte Verknüpfung des individuellen OSA-Ergebnisses mit Studienverlaufsdaten möglich ist, und reflektieren Chancen und Hindernisse dieser exemplarischen Ansätze. Die Beispiele stammen aus dem OSA der Justus-Liebig-Universität Gießen (JLU). Zunächst zeigen wir auf Basis der Daten einer Studieneingangsbefragung, wie viele der eingeschriebenen Studienanfängerinnen und -anfänger das OSA genutzt haben und wie sie es bewerten (Studie 1). Anschließend stellen wir detailliertes Feedback der Nutzerinnen und Nutzer vor (Studie 2). Es folgt der Bericht über eine Studie zur Akzeptanz von studienbezogenen Intelligenztests und Persönlichkeitsfragebogen (Studie 3): Verfahren, die typischerweise im Rahmen von OSA eingesetzt werden. Abschließend diskutieren wir, an welche Gelingensbedingungen Evaluationen eines OSA geknüpft sind und welche strategischen Implikationen sich, aus unseren Erkenntnissen, ableiten lassen.

Schlüsselwörter: Akzeptanz, Evaluation, Online Self-Assessments

Online self-assessments to support study choices – obstacles, opportunities and practical examples

Abstract: Choosing a course of study is the first challenge at the beginning of an academic career. Once the choice has been made, first-year students have to find their way in the new academic environment and meet the requirements of their field of study. At no other stage of study do so many students drop out. The main reasons for this include performance deficits and a lack of motivation. These, in turn, often stem from the fact that first-year students are surprised by the actual content of their studies. This can result in an overburdening and disappointing study situation. Many higher education institutions try to minimize this discrepancy between study expectations and study reality by offering prospective students online self-assessments (OSA) as an opportunity to intensively examine their desired field of study in advance. If higher education institutions offer such tools, evaluations should be implemented. Depending on the time point and the purpose of the evaluation, different evaluation approaches can be used. In this paper, we provide and critically reflect examples on how to design evaluations if – as in most cases – a direct link between the used OSA performance with administrative student data is not possible. The examples stem from the OSA at the Justus-Liebig-University Gießen (JLU). Firstly, building on data collected in the context of the annual freshmen survey at JLU, we report how many first-year students claimed to have used an OSA and how they rated it. Secondly, we present detailed user experiences and feedback reported by OSA users directly after using the tool. Thirdly, we illustrate an approach to collect data on user acceptance focused on the psychometric instruments used in OSA: study-related cognitive ability tests as well as questionnaires assessing personality – two examples for instruments typically used in the OSA context. We finally resume the conditions for a successful evaluation and derive recommendations for OSA evaluations.

Keywords: acceptance, evaluation, online self-assessments

Einleitung

Ein Studium beginnen: Diesen großen Schritt unternahmen im Jahr 2021 in Deutschland mehr als 470 000 Menschen. Seit Mitte der 2000er Jahre ist die Anzahl der Studentinnen und Studenten immer größer geworden. Obgleich dieser Trend in Bezug auf die Studienanfängerinnen und Studienanfänger in den letzten vier Jahren leicht rückläufig war (Statistisches Bundesamt, 2022), ist die Anzahl der Studentinnen und Studenten insgesamt weiterhin immens: fast 3 000 000 (Hochschulrektorenkonferenz [HRK], 2021). Zeitgleich zum Aufwuchs an Studentinnen und Studenten, hat sich auch die Hochschullandschaft verändert. Das Angebot an Studienfächern an den staatlichen und privaten Hochschulen und Universitäten ist kaum noch zu überblicken. Laut einem Bericht der HRK (2021) hat sich das Angebotsspektrum in den letzten 15 Jahren auf fast 21 000 (grundständige und weiterführende) Studiengänge nahezu verdoppelt. Die Studienwahl ist dadurch nicht einfacher geworden.

Ein Online Self-Assessment-Angebot (OSA-Angebot) zu etablieren, stellt eine Möglichkeit dar, Studieninteressierte bei der Studienwahl zu unterstützen. Mittlerweile bieten immer mehr Hochschulen in Deutschland solche Online-Tools an (Stoll & Weis, 2022), um die teilweise im Studieneinstieg erlebte Diskrepanz zwischen Studierenerwartungen und Studienwirklichkeit zu minimieren (Hasenberg & Schmidt-Atzert, 2014; Stoll et al., 2022). Obgleich es zahlreiche OSA gibt, sind systematische Evaluationen ebendieser rar. Vor diesem Hintergrund stellen wir dar, welche Evaluationsansätze denkbar sind, und berichten exemplarisch von drei Evaluationsstudien im Rahmen des OSA-Projektes der Justus-Liebig-Universität Gießen (JLU). Abschließend reflektieren wir Möglichkeiten und Grenzen der OSA-Evaluation.

Relevanz und Funktion von Online Self-Assessments

Dass die Studienwahl von großer Bedeutung ist, unterstreichen nicht zuletzt Untersuchungen zu Ursachen von Studienzufriedenheit sowie Abbruch(-intentionen). Im Mittel brechen an deutschen Hochschulen in den grundständigen Studiengängen 27% der Studentinnen und Studenten ab (Heublein, Richter & Schmelzer, 2020); die meisten von ihnen im ersten Studienjahr. Zu den Hauptgründen dafür zählen Leistungsdefizite und Motivationsmangel. Daneben gelten auch Unzufriedenheit mit den Studienbedingungen und den (Vorab-)Informationen zum Studium als häufig genannte Abbruchgründe (Neugebauer, Daniel & Wolter, 2021). Dem Studienerfolg hingegen zuträglich sind unter

anderem die Zufriedenheit mit den Inhalten und den Studienbedingungen (Hiemisch, Westermann & Michael, 2005).

Aus individueller Sicht ist es folglich wünschenswert ein Studienfach zu wählen, bei dem die eigene Studienmotivation hoch ist (Messerer, Karst & Janke, 2023). Darüber hinaus ist es sinnvoll, sich vor der Bewerbung um einen Studienplatz hinreichend über die tatsächlichen Studieninhalte zu informieren (Päßler, 2023) und sich mit den Studienbedingungen an der konkreten Hochschule auseinanderzusetzen. Neben dieser Perspektive (Mikroebene) ist es auch aus der Perspektive der Hochschulen (Mesoebene) wünschenswert, dass Studieninteressierte eine „gute Wahl“ treffen. Hierfür gibt es zahlreiche Gründe. Einer davon ist die an die Anzahl an Absolventinnen und Absolventen gekoppelte Finanzierung der Hochschulen, die durch hohe Abbruchquoten (und eine geringe Anzahl an Absolventinnen und Absolventen) gefährdet wird. Darüber hinaus ist eine geringe Quote an Personen, die ihr Studium abbrechen oder das Fach wechseln, auch gesamtgesellschaftlich (Makroebene) wünschenswert: Angesichts des aktuellen Personal- und Fachkräftemangels (z.B. Gehrlach, Bergen & Eiler, 2022) ist es nicht wünschenswert, Studienzeiten zu verlängern.

Eine Möglichkeit zur Unterstützung Studieninteressierter auf dem Weg zu einer informierten Studienwahl, die von immer mehr Hochschulen in Deutschland genutzt wird, ist das Anbieten von OSA (Hasenberg & Schmidt-Atzert, 2014; Stoll et al., 2022). Im Rahmen von OSA können sich Studieninteressierte anonym und kostenfrei mit ihrem Studienwunsch auseinandersetzen und Fragen klären wie: „Bringe ich die nötigen Voraussetzungen für Erfolg und Zufriedenheit in diesem Studium mit?“ und „Entsprechen die Studieninhalte dem, was ich mir vorstelle?“. Darüber hinaus können mit OSA ggf. vorhandene Defizite im Vorwissen (z.B. mathematische Vorkenntnisse) frühzeitig sichtbar gemacht werden, wodurch die Möglichkeit besteht, diese (z.B. durch den Besuch eines Online-Propädeutikums) aufzuarbeiten, um einen möglichst reibungslosen Einstieg ins Studium zu finden. Des Weiteren konnten Karst, Ertel, Frey & Dickhäuser (2017) zeigen, dass die Teilnahme an einem OSA motivationale Merkmale einer Person (z.B. den intrinsischen Wert) in Bezug auf das Studienfach verändern kann.

Die Teilnahme an einem OSA kann auch dazu führen, von dem ursprünglichen Studienwunsch Abstand zu nehmen und sich nochmal auf die Suche nach alternativen Studienfächern oder Ausbildungsmöglichkeiten zu begeben. Dies ist beispielsweise der Fall, wenn durch das OSA Fehlerwartungen offensichtlich werden und sich zeigt, dass individuelle Vorstellungen und Voraussetzungen wenig zu den Studieninhalten und -bedingungen passen. Zusammengefasst ist es das Ziel von OSA, studieninteressierte Personen auf dem Weg zu einer informierten Studienwahl zu unter-

stützen, um die Selbstselektion (mehr geeignete und weniger ungeeignete Bewerberinnen und Bewerber) zu fördern.

Evaluation von Online Self-Assessments

Die Evaluation eines OSA hinsichtlich dieses Ziels die Selbstselektion zu fördern, wird als summative Evaluation bezeichnet. Wie Weis et al. (2022) in ihrem Beitrag zu Entwicklung, Konzepten und Qualitätsstandards für OSA erläutern, erfolgt die summative OSA-Evaluation nach Abschluss der OSA-Entwicklung, ggf. als Teil einer Rechenschaftslegung. In diesem Zusammenhang kann man OSA im Hochschulkontext angelehnt an die Evaluationsliteratur als Interventionen verstehen. Im Idealfall würde man sensu Wittmann (1990) mit sogenannten Cattellschen Datenboxen arbeiten und mit einer „Experimentalgruppe“ und einer „Kontrollgruppe“ untersuchen, ob sich OSA-Nutzerinnen und -Nutzer von Nicht-Nutzerinnen und -Nutzern in relevanten Kriterienvariablen unterscheiden (Petri, Schütte & Beauducel, 2022). Weiterhin könnte man untersuchen, ob sich nach der Einführung eines OSA die Studienabbruchquote verringert respektive ob sich der Anteil geeigneter Bewerberinnen und Bewerber erhöht (Stoll & Gfrörer, 2023).

Wünschenswert wäre auch, dass OSA-Nutzerinnen und -Nutzer sich besser informiert fühlen, höhere Studienzufriedenheit, bessere Studienleistungen und eine niedrigere Abbruchintention aufweisen. Tatsächlich gibt es unserer Kenntnis nach keine publizierte Untersuchung, die eine umfassende Evaluation in diesem Sinne dokumentiert. Angesichts der großen (und wachsenden) Zahl an OSA in Deutschland, Österreich und der Schweiz, die bereits im Jahr 2022 von Stoll et al. auf über 850 an 75 Standorten beziffert wird (S. 2), mag dies zunächst überraschen. Dass keine umfangreichen summativen OSA-Evaluationen vorliegen, kann in erster Linie daran liegen, dass die OSA-Nutzung in der Regel vor der Immatrikulation und anonym stattfindet. Es ist daher nur begrenzt möglich, nachzuerfolgen, ob OSA-Nutzerinnen und -Nutzer das Studienfach (an der jeweiligen Hochschule), das ihnen empfohlen wurde oder zu dem sie ein OSA bearbeitet haben, wählen und sich letztendlich einschreiben. Noch schwieriger ist es, eine „Kontrollgruppe“ ebenfalls (längsschnittlich) zu untersuchen, um somit auch innerhalb dieser Gruppe Kriteriendaten zu sammeln.

Eine Möglichkeit, von Anfang an eine Basis für eine (längsschnittliche) Evaluation bezüglich derjenigen, die sich für ein Studium entscheiden, zu schaffen, besteht darin, die Teilnahme am OSA zur Pflicht zu erklären. Die Teilnahme ist zwar anonym, man muss aber bei der Imma-

trikulation einen Code eingeben, der die Teilnahme bestätigt. Über diesen Code werden die im OSA erzielten Ergebnisse mit den späteren Studienverlaufdaten verknüpft. An einigen Hochschulen wird dieses Verfahren verpflichtend (für einige Studiengänge) umgesetzt (z.B. RWTH Aachen, Uni Tübingen), an anderen läuft die OSA-Teilnahme auf freiwilliger Basis (z.B. HfWU Nürtingen-Geislingen), die Nutzerinnen und Nutzer werden aber um Generierung und Angabe eines Codes zwecks Datenverknüpfung gebeten. An wieder anderen wird dies mit dem Hinweis auf datenschutzrechtliche Bedenken abgelehnt (alle Angaben: Stand Januar 2023). Erfolgt die Verknüpfung der OSA-Ergebnisse mit den Studienverlaufdaten auf freiwilliger Basis, so kann man zumindest die OSA-Ergebnisse dieser Freiwilligen mit den administrativen Daten längsschnittlich verknüpfen. Dies unterliegt zwar offensichtlich einer Selektivität, erlaubt aber dennoch eine erste Schätzung der Zusammenhänge. So hat man zumindest von allen Freiwilligen die OSA-Ergebnisse und Studienverlaufdaten und somit Daten für längsschnittliche Untersuchungen.

Bereits während des Entwicklungsprozesses und damit als Unterstützung bei der Gestaltung und Optimierung des OSA kann eine formative Evaluation erfolgen. Als Beispiele für Kriterien der formativen Evaluation führen Weis et al. (2022) Akzeptanz, Nutzerfreundlichkeit (Usability), Ökonomie, Nachhaltigkeit, Qualität der Konzeption und der Implementation an. In diesem Beitrag gehen wir auf verschiedene Aspekte der Akzeptanz ein. Konkrete Evaluationsfragen können beispielsweise abzielen auf die „Zufriedenheit mit der Teilnahme am OSA, die wahrgenommene Nützlichkeit bei der Studienwahlentscheidung, Ernsthaftigkeit und Vollständigkeit der Bearbeitung des OSA (Compliance) etc.“ (Weis et al., 2022, S. 211). Bezugnehmend auf das Evaluationsmodell nach Kirkpatrick, übertragen auf den Hochschulkontext (Praslova, 2010), sind die Aspekte der Akzeptanz der ersten der vier Evaluationsebenen – der Ebene der „Reaction“ – zuzuordnen.

Angesichts der Freiwilligkeit der Teilnahme kommt der Akzeptanz der OSA eine besondere Bedeutung zu. Auf der Basis der Forschung zur Personalauswahl (Hausknecht, Day & Thomas, 2004) ist davon auszugehen, dass die Akzeptanz sowohl für die Teilnahme am OSA, als auch für die Beachtung der OSA-Ergebnisse (Feedback) eine hohe Bedeutung hat. Somit ist die Akzeptanz zwar nur einer der von Weis et al. (2022) genannten Aspekte und „nur“ die unterste Ebene im Modell von Kirkpatrick (zitiert nach Praslova, 2010), jedoch sollte ihre Bedeutung nicht unterschätzt werden: Die Erfüllung der Anforderungen der untersten Ebene im Modell von Kirkpatrick ist die notwendige (wenngleich nicht hinlängliche) Voraussetzung für die Erfüllung der weiteren Anforderungen. Wird das OSA von den Nutzerinnen und Nutzern nicht gut akzeptiert, werden

sie das erhaltene Feedback nach der OSA-Bearbeitung weniger wahrscheinlich ernstnehmen bzw. der Empfehlung folgen. Zudem sind (ehemalige) Nutzerinnen und Nutzer „story teller“. Empfehlen sie das bearbeitete OSA nicht weiter oder raten gar von dessen Nutzung ab, kann sich das negativ auf die zukünftigen Nutzungszahlen auswirken.

Um die Akzeptanz eines OSA zu erfassen, müssen die Nutzerinnen und Nutzer das Verfahren nach der Nutzung unter Akzeptanzgesichtspunkten beurteilen. Dabei kann man einerseits ein OSA als ein Verfahren im Gesamten betrachten (ähnlich eines Globalurteils über Assessment Center) oder aber differenziert betrachten, wie die Akzeptanzurteile gegenüber verschiedenen, in OSA häufig genutzten, diagnostischen Testverfahren ausfallen. Hierfür eignen sich sowohl ins OSA integrierte Evaluationsitems, als auch separate Evaluationsstudien bezüglich der verwendeten Testverfahren (respektive Prototypen der jeweiligen Testverfahren).

Im Folgenden möchten wir mit drei Studien Beispiele dafür geben, wie OSA formativ evaluiert werden können, selbst wenn aus formalen (z.B. datenschutzrechtlichen) Gründen keine direkte Verknüpfung des individuellen OSA-Ergebnisses mit Studienverlaufsdaten möglich ist: In Studie 1 zeigen wir auf Basis der Berichte der Studieneingangsbefragung der JLU Gießen, wie viele der Befragten innerhalb der OSA-Fächer angaben, das OSA der JLU genutzt zu haben und als wie hilfreich sie es bewerteten. Daran anschließend werden wir in Studie 2 auf Basis einer OSA-internen Evaluation und in Studie 3 anhand einer separaten Studie weitere Akzeptanzurteile (Studie 2: Globalurteil bezüglich Nützlichkeit; Studie 3: Akzeptanz gegenüber verschiedenen Testverfahren) berichten. Damit soll dieser Artikel – neben den einzelnen Ergebnissen – aufzeigen, welche Art von Evaluation bei einem über viele Jahre laufenden, aber nicht verpflichtenden OSA, ohne Verknüpfung mit administrativen Daten, möglich ist und welche Probleme sich unter diesen Rahmenbedingungen für die Evaluation ergeben.

Fallbeispiel: Das OSA der JLU

Die im Folgenden berichteten Studien 1 und 2 stellen eine Evaluation des OSA der JLU dar. Die JLU hat studieninteressierten Personen von 2013 bis 2020 ein so genanntes konfirmatorisches OSA angeboten. Das bedeutet, dass studieninteressierte Personen für verschiedene (grundständige) Studienfächer prüfen konnten, ob und wie gut ihr Wunschstudium zu ihnen passt. Das Angebot richtete sich also an Personen, die bereits eines oder mehrere konkrete Wunschfächer ins Auge gefasst hatten, mit denen sie sich dann bei der OSA-Nutzung im Detail auseinandersetzen konnten.

Das OSA-Angebot der JLU war entsprechend fachspezifisch organisiert. Orientiert am Bedarf und Interesse der jeweiligen Fächer und Fachbereiche wurden jeweils in enger Zusammenarbeit zwischen dem OSA-Team und Fachvertreterinnen und -vertretern fachspezifische OSA für 14 verschiedene Fächer entwickelt. Tabelle E1 im elektronischen Supplement (ESM) 1 zeigt die im OSA-Angebot der JLU vertretenen Fächer und den jeweiligen Startzeitpunkt des fachspezifischen Angebots.

Die OSA der JLU waren modular aufgebaut. Abbildung E1 im ESM 1 zeigt einen Screenshot der Darstellung des Aufbaus, exemplarisch für das Fach Chemie.

Jedes OSA enthielt neben einem einleitenden Erläuterungstext und allgemeinen Hinweisen zur Struktur sowie Bearbeitungshinweisen folgende Module: Interessensfragebogen, studienbezogene Persönlichkeitsfragebogen und Intelligenztests, Erwartungschecks sowie das Feedbackmodul und das Evaluationsmodul. Die Diagnostik zur Persönlichkeit und Intelligenz konnte man optional hinzuwählen.

Im sogenannten Erwartungscheck wurden Aussagen zum Studium präsentiert, zu denen man angeben sollte, inwiefern man sie für zutreffend hält. Nach jeder Aussage wurde direkt aufgelöst, ob diese Aussage zutrifft oder nicht und der thematisierte Sachverhalt wurde kurz erläutert. Weiterhin gab es Informationsmodule zur JLU und zum jeweiligen Fach.

Falls seitens des jeweiligen Fachbereichs gewünscht, wurden zusätzlich Fachaufgaben angeboten, die einen Eindruck von den Anforderungen zu Studienbeginn bieten. Dies war der Fall für folgende Fächer: Anglistik, Angewandte Theaterwissenschaften, Chemie, Germanistik, Materialwissenschaft, Philosophie, Physik, Romanistik und Slavistik. Die geschätzte Bearbeitungszeit lag jeweils bei 45 bis 60 Minuten für das OSA insgesamt.

Unmittelbar nach der Bearbeitung erhielten die Teilnehmerinnen und Teilnehmer ein individuelles Feedback zu ihren Ergebnissen im Bereich Interessen, Persönlichkeit und Intelligenz (Feedbackmodul).

Methode

Alle in diesem Manuskript präsentierten Analysen wurden mittels SPSS Statistics (Version 21; IBM Corp., 2012) durchgeführt.

Studie 1: Studieneingangsbefragung der JLU

In der ersten Studie beantworten wir auf Basis der Berichte zu den jährlich im Wintersemester von der „Servicestelle Lehrevaluation“ bei Studieneinsteigerinnen und Studi-

eneinsteigern durchgeführten Studieneingangsbefragungen der JLU die folgenden beiden Forschungsfragen:

Forschungsfragen

1. Nutzung: Wie viele Studieneinsteigerinnen und Studieneinsteiger in den OSA-Fächern gaben an, das OSA der JLU zur Studienwahl-Orientierung genutzt zu haben?
2. Nützlichkeit: Als wie hilfreich bei der Studienwahl/Orientierung wurde das jeweilige OSA bewertet?

Stichprobe

Für die Evaluation liegen Daten aus den Wintersemestern 2015/2016 bis 2019/2020 vor. Wir nutzen jeweils nur Angaben aus den Fächern, zu denen zum jeweiligen Zeitpunkt der Studienbewerbung ein OSA zur Verfügung stand. Unsere Evaluation basiert auf den jeweiligen Berichten der Stabstelle Lehre der JLU (z. B. Ehrlich, 2020). In jedem Wintersemester wurden alle Studieneinsteigerinnen und Studieneinsteiger im Rahmen der Studieneinführung zur Teilnahme an der Befragung eingeladen (freiwillige Teilnahme). An der Studieneingangsbefragung nahmen in dem von uns betrachteten Zeitraum pro Studienjahr mindestens $N = 2590$ bis höchstens $N = 3311$ teil. Die Rücklaufquote lag damit im Bereich 54 % bis 62%. Die demografischen Merkmale entsprechen denen „typischer“ Studienanfängerinnen und Studienanfänger, mit einem größeren Anteil von Frauen (65–67%) und einem stark vertretenen Altersbereich zwischen 18 bis 21 Jahren. Details dazu legt Ehrlich (2020) dar.

Die Studieneingangsbefragung umfasst Daten zur Nutzung und Nützlichkeit des OSA für die Fächer Agrarwis-

senschaften, Chemie, Ernährungswissenschaften, Ökotrophologie, Physik, Umweltmanagement und Wirtschaftswissenschaften. Für die übrigen Fächer lagen entweder zu wenige (Materialwissenschaften) oder keine Daten vor (weil das OSA noch nicht eingeführt war). Das Fach Philosophie wurde in der Studieneingangsbefragung nicht gesondert codiert (sondern als „Cluster“ mit anderen Fächern behandelt), sodass hierzu ebenfalls keine spezifischen Daten vorlagen.

Ergebnisse

Tabelle 1 zeigt die Nutzungszahlen und Nützlichkeitsratings (Item der Studieneingangsbefragung: „Welche Informationskanäle/-angebote zur Studienwahl/-orientierung haben Sie genutzt und wie hilfreich waren diese?“) des OSA-Angebotes aus den oben genannten fünf Studieneingangsbefragungen (Wintersemester 2015/2016 bis 2019/2020) für die sieben zuvor genannten OSA-Fächer, für die zum jeweiligen Zeitpunkt der Befragung ein OSA-Angebot bestand.

Nutzung

Es zeigte sich kein fächerübergreifender Nutzungstrend. Die Nutzungsdaten differenzieren über die Zeit hinweg von Fach zu Fach. Erwartungsgemäß ist in den Fächern, die allgemein höheren Zulauf haben, auch die absolute Anzahl der OSA-Nutzerinnen und -Nutzer unter den befragten Personen höher.

Nützlichkeit

Im Analysezeitraum wurde die für die Bewertung der Nützlichkeit verwendete Skala geändert. Die ersten drei

Tabelle 1. Daten aus den Studieneingangsbefragungen der Jahrgänge 15/16 bis 19/20 (Studie 2): Nutzung (absolute Angaben) und wahrgenommene Nützlichkeit des OSA-Angebots der JLU (Studie 1)

Fach	$N_{\text{OSA-Nutzung}}^a$ ($N_{\text{WiSe 15/16}}$)	Bewertung ^b	$N_{\text{OSA-Nutzung}}^a$ ($N_{\text{WiSe 16/17}}$)	Bewertung ^b	$N_{\text{OSA-Nutzung}}^a$ ($N_{\text{WiSe 17/18}}$)	Bewertung ^b	$N_{\text{OSA-Nutzung}}^a$ ($N_{\text{WiSe 18/19}}$)	Bewertung ^b	$N_{\text{OSA-Nutzung}}^a$ ($N_{\text{WiSe 19/20}}$)	Bewertung ^b
Agrarwissenschaften	7 (89)	3.0 (1.5)	21 (95)	3.9 (1.2)	13 (55)	3.1 (1.1)	1 (41)	–	8 (57)	43 %
Chemie	9 (72)	3,6 (–)	13 (96)	3,9 (–)	22 (103)	4,0 (–)	7 (66)	29 %	6 (38)	33 %
Ernährungswissenschaften	22 (81)	3,9 (0,9)	10 (44)	3,8 (0,9)	25 (134)	3,9 (1,1)	7 (73)	57 %	23 (129)	48 %
Ökotrophologie	25 (140)	3,8 (1,0)	38 (148)	3,6 (1,1)	46 (222)	3,8 (1,1)	11 (103)	64 %	26 (144)	54 %
Physik	6 (34)	3,0 (1,4)	11 (34)	2,8 (1,5)	9 (34)	3,9 (1,1)	5 (28)	80 %	9 (25)	44 %
Umweltmanagement	14 (70)	3,6 (1,1)	23 (97)	3,3 (1,1)	25 (106)	3,6 (1,2)	13 (105)	69 %	18 (95)	69 %
Wirtschaftswissenschaften	41 (324)	3,6 (1,2)	87 (433)	3,6 (1,1)	57 (336)	3,6 (0,9)	16 (239)	75 %	10 (121)	80 %

Anmerkungen: WiSe = Wintersemester. Die Daten in der Kategorie „Chemie“ sind die Zusammenfassung für die beiden BSc-Studiengänge Chemie und Lebensmittelchemie, für die ein gemeinsames OSA angeboten wird. ^a Absolute Anzahl OSA-Nutzerinnen und -Nutzer, in Klammern absolute Anzahl Befragter pro Fach. ^b In den Berichten aus den WiSe 15/16 bis 17/18 erfolgte die Beantwortung (angegeben als M (SD)) des Items „Wie hilfreich war das OSA zur Studienwahl/Orientierung?“ von „1“ = „nicht hilfreich“ bis „5“ = „hilfreich“. ^c In den Berichten ab WiSe 18/19 wurde die Bewertung der Nützlichkeit nur noch in drei Kategorien zusammengefasst („sehr“ hilfreich; „indifferent“, „weniger/gar nicht hilfreich“). Angegeben sind für WiSe 18/19 und WiSe 19/20, wie viel Prozent der Nutzerinnen und Nutzer die Bewertung „(sehr) hilfreich“ abgaben.

Analyse-Semester wurde eine fünfstufige Skala (von „1“ = „nicht hilfreich“ bis „5“ = „hilfreich“) genutzt. Die Mittelwerte der Bewertungen liegen – mit Ausnahme der Einschätzung für das Fach Physik im Wintersemester 2016/2017, die allerdings nur auf $n = 11$ Personen beruht – bei allen Fächern kontinuierlich über dem Skalenmittel von 3. Das zeigt allgemein, dass diejenigen, die ein OSA bearbeitet haben, dieses tendenziell als „hilfreich“ für die Studienwahl/Orientierung ansehen.

In den letzten beiden gesichteten Befragungen wurde das Nützlichkeitsrating auf einer dreistufigen Skala erfasst („(sehr) hilfreich“, „indifferent“, „weniger/gar nicht hilfreich“). Hier zeigte sich – außer für die Chemie – dass stets über 40% der OSA-Nutzerinnen und -Nutzer dieses als „(sehr) hilfreich“ bewerteten. Teilweise gaben dies sogar 80% der Befragten an (Physik; Wintersemester 2018/2019).

Studie 2: OSA-interne Evaluation

Mit der zweiten Studie untersuchen wir, wie die OSA-Nutzerinnen und -Nutzer im unmittelbaren Anschluss an die OSA-Bearbeitung die Nutzung im Hinblick auf verschiedene der oben genannten Ziele von OSA einschätzen.

Forschungsfragen

Die für Studie 2 genutzten Evaluationsitems waren angelehnt an die im Rahmen des Netzwerks Online Self-Assessment erarbeiteten Evaluationsitems (vgl. Weis et al., 2022). Mit diesen Fragen erfassen wir: Wie beurteilen OSA-Nutzerinnen und -Nutzer ...

- ... das OSA-Angebot insgesamt (Globalurteil)?
- ... den Wert des OSA-Angebots als Hilfestellung bei der Studienwahl-Entscheidung?

Darüber hinaus untersuchen wir: Inwiefern hat das OSA den Nutzerinnen und Nutzern Hilfe geleistet ...

- ..., einen Überblick über die Voraussetzungen des jeweiligen Studiums zu erhalten?
- ... bei der Einschätzung der eigenen Fähigkeiten?
- ... bei der Einschätzung der eigenen Interessen?

Schließlich prüfen wir:

- Gibt es fachspezifische Unterschiede in den fünf zuvor genannten Bewertungsdimensionen?

Stichprobe

Von 2017 bis 2020 wurden insgesamt $N = 5030$ Zugriffe auf das OSA der JLU registriert, wobei das Angebot von Juni bis September 2018 aufgrund technischer Probleme und von Dezember 2019 bis April 2020 aufgrund der Folgen des Cyber-Angriffs auf die JLU (#JLUoffline) nicht abgerufen werden konnte. Alle Nutzerinnen und Nutzer wurden zum Ende der OSA-Bearbeitung gebeten (freiwillige Basis), das OSA zu evaluieren. Direkt vor den Evaluationsitems wurden auf freiwilliger Basis demografische Daten erhoben. Von der Gesamtgruppe machten viele keine Angaben zur Demographie. Von denen, die diese Angaben machten ($n = 2952$), ist bekannt, dass sie im Mittel $M_{Alter} = 21.03$ Jahre ($SD_{Alter} = 6.7$ Jahre) alt waren. 62.5% von ihnen gaben an, weiblich zu sein.

Die einzelnen Evaluationsfragen wurden von minimal $n_{Evalmin} = 1784$ (35.5% aller Nutzerinnen und Nutzer) bis maximal $n_{Evalmax} = 2448$ Personen (48.7%) beantwortet (siehe Abbildung 1). Diese Evaluationsteilnehmerinnen und -teilnehmer wiesen hinsichtlich der Verteilung auf die Geschlechter und das mittlere Alter nahezu exakt die gleichen Kennwerte auf wie die genannte Subgruppe derjenigen, die Angaben zur Demografie machten.

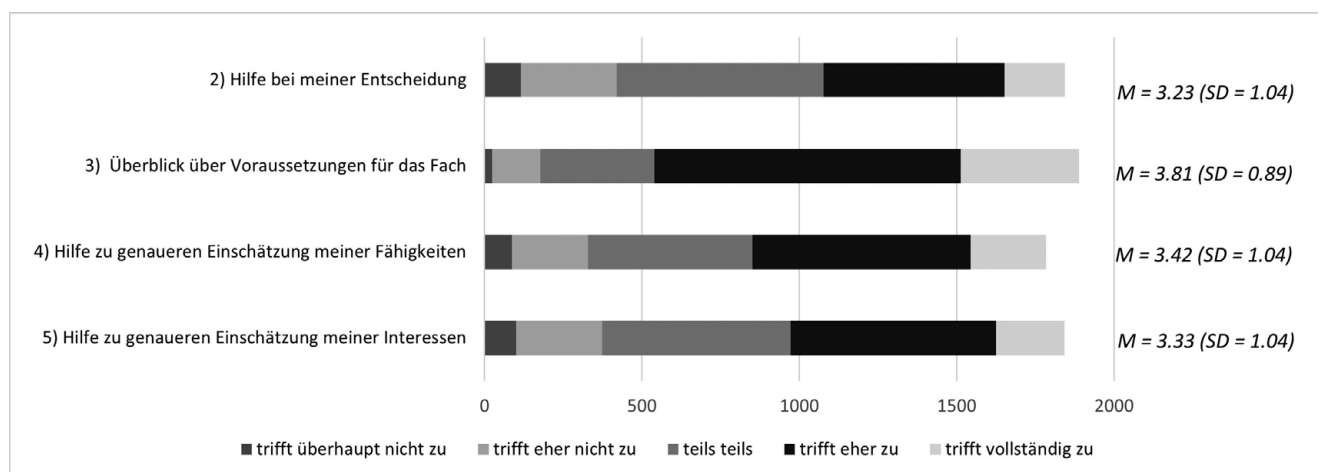


Abbildung 1. Ausgewählte Evaluationsitems im Rahmen des OSA (Studie 2). Antwortformat: Zustimmungsskala von 1 = „trifft überhaupt nicht zu“ bis 5 = „trifft vollständig zu“. $n_2 = 1844$, $n_3 = 1889$, $n_4 = 1784$ bis $n_5 = 1843$ (Studie 2).

Ergebnisse

Befragt danach, welche Schulnote (von „1“ = „sehr gut“ bis „6“ = „ungenügend“) sie dem OSA-Angebot der JLU geben würden (Forschungsfrage 1), bewerteten die Nutzerinnen und Nutzer das Angebot mit $M_{Note} = 2.69$ ($SD_{Note} = 1.0$, $Md = 2$; $n_{Note} = 2448$). Abbildung 1 zeigt exemplarisch weitere Ergebnisse der OSA-internen Evaluation anhand ausgewählter Items, die zur Beantwortung der oben genannten Fragen 2 bis 5 dienten. Knapp 42% aller Evaluationsteilnehmerinnen und -teilnehmer urteilten, dass ihnen die Nutzung des OSA ihnen eine Hilfe bei der Studienwahl-Entscheidung ist (Frage 2). Jeweils mindestens die Hälfte der Befragten bestätigte, dass sie durch die Nutzung des OSA einen guten Überblick über die Voraussetzungen in ihrem Wunschstudium und eine genauere Einschätzung ihrer (studienrelevanten) Fähigkeiten sowie ihrer Interessen erlangt haben (Fragen 3 bis 5; alle prozentualen Angaben beziehen sich auf die Summe der Antworten „trifft eher zu“ und „trifft vollständig zu“). Abbildung 1 gibt die zugehörigen deskriptiven Statistiken zur OSA-internen Evaluation wieder.

Ergänzend wurde mit einer multivariaten Varianzanalyse (MANOVA) der sechsten Frage nachgegangen, ob sich die mittleren Urteile der Nutzerinnen und Nutzer unterscheiden, je nachdem, welches fachspezifische OSA zuvor bearbeitet wurde (Frage 6). Um nur mit ausreichend besetzten Zellen zu arbeiten, berichten wir im Folgenden exemplarisch die Ergebnisse für die drei im OSA der JLU am stärksten nachgefragten Fächer (Physik, Wirtschaftswissenschaften und die Gruppe der Studienfächer des Fachbereichs 09, die ein gemeinsames OSA nutzten). Unter Verwendung der fünf oben genannten Evaluationsitems als abhängige Variablen und der Einteilung, welches der drei fachspezifischen OSA bearbeitet wurde, als fester Gruppenfaktor, ergab sich mit $F(10, 3428) = 2.444$, $p < .007$, Wilks $\lambda = .986$, partielles $\eta^2 = .007$, ein bedeutsamer Unterschied zwischen den Fächern. Exploriert man diesen Effekt weiter mit Post-hoc-Vergleichen, so zeigen sich nur

bei den ersten beiden Evaluationsitems (Globalurteil als Schulnote und OSA als Hilfe bei der Entscheidung) ein bedeutsamer Unterschied: Wie Tabelle 2 zeigt, wurde jeweils das Physik-OSA am besten bewertet.

Studie 3: Akzeptanz gegenüber spezifischen Testverfahren im Rahmen von OSA

Außerhalb der OSA-Plattform wurde über mehrere Kohorten von Studienanfängerinnen und Studienanfänger im Bachelorstudiengang Psychologie an der JLU eine separate Studie durchgeführt, mit der u. a. die Akzeptanz gegenüber verschiedenen, typischerweise in OSA eingesetzten psychometrischen Testverfahren im Detail untersucht werden sollte (Wiesenhütter, 2022).

Im Folgenden berichten wir Ergebnisse, die auf aggregierten Daten, die in den Sommersemestern 2013 bis 2015 sowie den Wintersemestern 2014/2015 bis 2018/2019 gesammelt wurden, beruhen. Psychologie-Studentinnen und -Studenten im dritten Semester bearbeiteten an drei separaten Terminen zunächst zwei Intelligenztest (IST Screening von Liepmann, Beauducel, Brocke & Nettelstroth, 2012, als Maß für die fluide Intelligenz und später den Mehrfachwahl-Wortschatz-Intelligenztest [MWT-B] von Lehl, 2005, als Maß für kristalline Intelligenz) und anschließend einen Persönlichkeitsfragebogen (BFI-K mit 21 Items von Rammstedt & John, 2005). Intelligenz- und Wissenstests sowie Persönlichkeitsfragebogen werden standardmäßig in OSA eingesetzt (Kersting & Marquardt, 2022), sodass die in unserer Studie erzielten Ergebnisse grundsätzlich als Indikatoren für die Akzeptanz gegenüber typischerweise in OSA eingesetzten Testverfahren interpretiert werden können. Die Akzeptanz eines diagnostischen Verfahrens hängt nicht nur von dem Verfahren selbst ab, sondern wird u. a. auch durch Situationsmerkmale beeinflusst. Es spielt z. B. eine Rolle, ob von dem Testergebnis eine bedeutsame Entscheidung abhängt oder nicht. Von da-

Tabelle 2. Deskriptive Statistiken der nach Fächern differenzierten OSA-internen Evaluation (Studie 2)

Evaluationsitem	Physik	Wirtschaftswissenschaften	Fachbereich 09
	M (SD)	M (SD)	M (SD)
Globalurteil Schulnote	2.35 (0.99)	2.56 (0.99)	2.59 (0.96)
Hilfe bei Studienwahlentscheidung	3.48 (0.97)	3.32 (1.0)	3.20 (1.05)
Überblick Voraussetzungen	3.85 (0.93)	3.82 (0.91)	3.82 (0.87)
Genauere Einschätzung Fähigkeiten	3.52 (0.96)	3.48 (1.09)	3.41 (1.03)
Genauere Einschätzung Interessen	3.30 (1.05)	3.44 (1.04)	3.32 (1.04)

Anmerkungen: Fachbereich 09 an der JLU: Agrarwissenschaften, Ernährungswissenschaften, Ökotrophologie und Umweltmanagement mit einem gemeinsamen großen, ausdifferenzierten OSA. Globalurteil auf der Schulnotenskala von „1“ = „sehr gut“ bis „6“ = „ungenügend“. Antwortformat für die anderen Items: Zustimmungsskala von 1 = „trifft überhaupt nicht zu“ bis 5 = „trifft vollständig zu“. $n_{Physik} = 117$, $n_{Wirtschaft} = 324$, $n_{Fachbereich09} = 1280$. $F(10, 3428) = 2.444$, $p < .007$, Wilks $\lambda = .986$, partielles $\eta^2 = .007$. Post-hoc-Vergleiche: Nur bei den ersten beiden Evaluationsitems ergab sich ein bedeutsamer Unterschied: beste Bewertung im Physik-OSA.

her können bereits vorliegende Erkenntnisse zur Akzeptanz von Tests (z. B. Anderson, Salgado & Hülshager, 2010) nicht ohne Weiteres auf den OSA-Kontext übertragen werden.

Direkt nach der Bearbeitung des jeweiligen Verfahrens füllten die Studentinnen und Studenten unter der Bedingung der Anonymität den so genannten „Akzept-Fragebogen“ aus. Das Akzeptanzurteil stellt die Summe der auf das individuelle Erleben und Bewerten diagnostischer Verfahren gerichteten Aspekte dar. Die „Akzept!“-Fragebogenreihe fokussiert ausschließlich solche Aspekte, die unmittelbar mit dem Verfahren zu tun haben. Die Akzeptanz wird dabei verfahrensspezifisch entweder mit dem Akzept-L!-Fragebogen (Denker, Schütte, Kersting, Wepfert & Stegt, 2023; Kersting, 2008) oder den Akzept-P!-Fragebogen (Beermann, Kersting, Stegt & Zimmerhofer, 2013) erfasst. Diese Fragebogen sind auf die jeweilige Verfahrenskategorie (Persönlichkeitsfragebogen oder Leistungstest) zugeschnitten. Beide Fragebogen erfassen mit je vier Items die folgenden drei Dimensionen der Akzeptanz: Kontrollierbarkeit, Messqualität und Augenscheinvalidität. Im Akzept-L! (Einsatz bei Leistungstests) kommt noch die Dimension Belastungsfreiheit (4 Items) hinzu, wohingegen beim Akzept-P! (Einsatz bei Persönlichkeitsfragebogen) noch die Dimensionen Wahrung der Privatsphäre (4 Items) und Antwortfreiheit (3 Items) hinzukommen. Beide Fragebogen sehen zudem ein abschließendes Single-Item vor, mit dem auf der Schulnotenskala ein Akzeptanz-Gesamturteil über das zuvor bearbeitete diagnostische Verfahren (Intelligenztest oder Persönlichkeitsfragebogen) ausgesprochen werden soll. Beim Akzept-L! kommt schließlich noch ein Single-Item zur Erfassung der verfahrensbezogenen Selbsteinschätzung zum Einsatz. Die psychometrische Qualität der beiden Akzept-Fragebogen wurde in verschiedenen Untersuchungen belegt (Beermann et al., 2013; Kersting, 2008).

Forschungsfragen

Untersucht werden sollte: Welche differenzierten Akzeptanzurteile sprechen die getesteten Personen über ...:

1. ... den Test zur fluiden Intelligenz: IST-Screening (Lipmann et al., 2012) aus?
2. ... den Test zur kristallinen Intelligenz: MWT-B (Lehrl, 2005) aus?
3. ... den Persönlichkeitsfragebogen: BFI-K (Rammstedt & John, 2005) aus?

Neben den absoluten Werten interessierte uns der Vergleich der Akzeptanzwerte:

4. Unterscheiden sich die mittleren Akzeptanz-Ratings in den drei Dimensionen Kontrollierbarkeit, Messqualität und Augenscheinvalidität sowie das Akzeptanz-Gesamturteil (Schulnote) zwischen den drei Verfahren?

Gesamtstichprobe

Insgesamt umfasst die Stichprobe $N = 1325$ Personen, wobei nicht alle jedes Verfahren bearbeitet haben¹. Von denen, die Angaben zur Demografie machten, gaben $n = 637$ an, weiblich zu sein und $n = 278$ gaben an, männlich zu sein. Das Alter wurde in Kategorien abgefragt: 20.2% gaben an, 20 Jahre oder jünger zu sein, 28.8% gaben an, 21 oder 22 Jahre alt zu sein. Weitere 8.9% gaben an, 23 oder 24 Jahre alt zu sein und 17% gaben an, älter als 24 Jahre zu sein.

Teilstichprobe

Schränkt man die Stichprobe auf diejenigen ein, die alle drei Verfahren und die zugehörigen Akzept-Fragebogen ausgefüllt haben, so reduziert sich der Stichprobenumfang auf $n = 98$. In dieser Teilstichprobe gaben 67 Personen (68.4%) an, weiblich zu sein und 11 Personen (11.2%) gaben an, männlich zu sein (25 Personen machten keine Angabe). Bezüglich des in Kategorien abgefragten Alters ergab sich folgende Verteilung: 44.9% gaben an, 20 Jahre oder jünger zu sein, 36.4% gaben an, 21 oder 22 Jahre alt zu sein. Weitere 8.2% gaben an, 23 oder 24 Jahre alt zu sein und 9.2% gaben an, älter als 24 Jahre alt zu sein (1 Person machte keine Angabe).

Ergebnisse

Als Globalurteil auf der Schulnotenskala erzielten die drei Verfahren die folgenden Werte:

1. **IST-Screening:** $M = 2.36$ ($SD = 0.65$, $Md = 2$, $n = 601$), dies entspricht einem „gut“.
2. **MWT-B:** $M = 3.10$ ($SD = 0.88$, $Md = 3$, $n = 559$), dies entspricht der Bewertung „befriedigend“.
3. **BFI-K:** $M = 2.69$ ($SD = 0.82$, $Md = 3$, $n = 568$), die Bewertung liegt zwischen „gut“ und „befriedigend“.

Tabelle 3 zeigt die (deskriptiven Statistiken der) Akzeptanz-Ratings in Bezug auf die einzelnen Akzeptanzdimensionen für alle drei Verfahren.

Vergleich IST-Screening, MWT-B und BFI-K (Teilstichprobe)

Um die Akzeptanzurteile direkt zu vergleichen, beschränken wir die Analyse auf die Teilstichprobe von $n = 98$

¹ Die Datenerhebung in Studie 3 fand pro Kohorte an drei verschiedenen Terminen statt. Die Studierenden wurden gebeten, jedes Mal bei der Paper-Pencil-Erhebung einen persönlichen Code zwecks Matchings der Daten anzugeben. Dass nicht zu allen drei Testverfahren Daten aller Studierenden der jeweiligen Kohorte vorliegen, kann drei Gründe haben: 1. Die Teilnahme war stets freiwillig; jede/jeder konnte an jedem Termin auch entscheiden, nicht teilzunehmen. 2. Manche Personen waren ggf. an manchen Terminen nicht vor Ort in der Vorlesung. 3. Leider hat es nicht immer geklappt, dass die Studierenden den Code korrekt angaben, obgleich immer die (gleiche) Anleitung zur Erstellung des individuellen Codes präsentiert wurde.

Tabelle 3. Deskriptive Statistiken der Akzeptanz-Ratings zum IST-Screening, MWT-B und BFI-K (Studie 3)

Akzept-Dimension	Fluide Intelligenz (IST-Screening)		Kristalline Intelligenz (MWT-B)		Persönlichkeit (BFI-K)	
	M (SD)	N	M (SD)	N	M (SD)	N
Kontrollierbarkeit	5.42 (0.69)	615	5.61 (0.65)	559	5.81 (0.38)	577
Messqualität	4.23 (0.72)	614	3.46 (0.87)	556	3.67 (0.85)	577
Augenscheinvalidität	3.28 (0.92)	613	2.45 (0.87)	551	3.53 (0.90)	577
Belastungsfreiheit	4.25 (0.94)	613	4.46 (0.90)	557	–	–
Wahrung der Privatsphäre	–	–	–	–	5.08 (0.96)	577
Antwortfreiheit	–	–	–	–	4.29 (1.04)	576
Globalurteil (Schulnote)	2.36 (0.65)	601	3.09 (0.87)	548	2.69 (0.82)	568

Anmerkungen: Zustimmung zu verschiedenen positiven Aussagen von „1“ = „trifft nicht zu“ bis „6“ = „trifft genau zu“ sowie (letzte Zeile) das Gesamt-Akzeptanzurteil als Schulnote (von „1“ = „sehr gut“ bis „6“ = „ungenügend“).

Personen, die alle drei Instrumente bearbeitet haben (Within-Subjects-Design) sowie auf die drei Akzeptanzdimensionen, die sowohl für Leistungstests, als auch für Persönlichkeitsfragebogen bewertet werden, plus das Globalurteil als Schulnote. Tabelle 4 gibt die zugehörigen deskriptiven Statistiken der Akzeptanz-Ratings für diese Stichprobe wieder. Eine MANOVA mit Messwiederholung (Within-Subject-Faktor *Verfahren* [IST-Screening versus MWT-B versus BFI-K]) in Bezug auf die drei Akzept-Dimensionen, die sowohl der Akzept-L!, als auch der Akzept-P! enthält, plus das Globalurteil, ergab einen signifikanten Effekt: $F(8, 90) = 34.520, p < .001$, partielles $\eta^2 = .754$, Wilks $\lambda = .246$.

Sowohl für die Akzept-Dimension Kontrollierbarkeit² ($F(1.897) = 13.60, p < .001$, partielles $\eta^2 = .123$), als auch für die Akzept-Dimensionen Messqualität ($F(2) = 37.95, p < .001$, partielles $\eta^2 = .281$) und Augenscheinvalidität ($F(2) = 75.53, p < .001$, partielles $\eta^2 = .438$) sowie für das Akzeptanz-Globalurteil ($F(2) = 42.60, p < .001$, partielles $\eta^2 = .305$) ergaben sich signifikante Unterschiede in den mittleren Ratings zwischen den drei Verfahren. Auffällig ist, dass der MWT-B, der am ehesten einem Wissenstest gleicht, in allen drei hier

betrachteten Akzept-Dimensionen am schlechtesten bewertet wurde. Das niedrige Akzeptanz-Rating zum MWT-B bezüglich der Augenscheinvalidität geht mit dem größten der vier oben genannten, hier beobachteten, Effekte einher.

Diskussion

Studie 1

Zusammenfassung

Die Studieneingangsbefragung ist grundsätzlich eine wertvolle Datenquelle. Die Befragung der Studieneinsteigerinnen und Studieneinsteiger in der Studieneinführungswoche, ob sie (unter anderem) das OSA-Angebot genutzt haben, und – falls ja –, als wie hilfreich sie es für die Studienwahl/Orientierung bewerten, kann helfen, einen ersten Eindruck des Bekanntheitsgrades (im Sinne von Nutzung) und der wahrgenommenen Nützlichkeit zu erhalten.

Es zeigte sich hier kein fächerübergreifender Nutzungstrend. Die Nutzungsdaten differenzieren über die Zeit hin-

Tabelle 4. Deskriptive Statistiken der Akzeptanz-Ratings zum IST-Screening, MWT-B und BFI-K (Studie 3; Teilstichprobe)

Akzept-Dimension	Fluide Intelligenz (IST-Screening)		Kristalline Intelligenz (MWT-B)		Persönlichkeit (BFI-K)	
	M (SD)	N	M (SD)	N	M (SD)	N
Kontrollierbarkeit	5.52 (0.62)	615	5.51 (0.80)	559	5.87 (0.33)	577
Messqualität	4.24 (0.76)	614	3.45 (0.78)	556	3.85 (0.70)	577
Augenscheinvalidität	3.17 (0.96)	613	2.41 (0.81)	551	3.67 (0.87)	577
Gesamturteil (Schulnote)	2.33 (0.64)	601	3.16 (0.89)	548	2.47 (0.72)	568

Anmerkungen: $N = 98$. Within-Subject-Design. Zustimmung zu verschiedenen positiven Aussagen von „1“ = „trifft nicht zu“ bis „6“ = „trifft genau zu“ sowie (letzte Zeile) das Akzeptanz-Gesamturteil als Schulnote (von „1“ = „sehr gut“ bis „6“ = „ungenügend“). Sowohl für die Dimension *Kontrollierbarkeit* ($F(1.897) = 13.60, p < .001$, partielles $\eta^2 = .123$), als auch für die *Messqualität* ($F(2) = 37.95, p < .001$, partielles $\eta^2 = .281$) und *Augenscheinvalidität* ($F(2) = 75.53, p < .001$, partielles $\eta^2 = .438$) sowie für das Akzeptanz-Globalurteil ($F(2) = 42.60, p < .001$, partielles $\eta^2 = .305$) ergaben sich signifikante Unterschiede in den mittleren Ratings zwischen den drei Verfahren.

² Da eine Verletzung der Voraussetzung der Sphärizität vorlag, wurde eine Huynh-Feldt-Korrektur der Freiheitsgrade vorgenommen.

weg von Fach zu Fach. Dies spricht nicht dafür, dass das OSA-Angebot kontinuierlich bekannter wird. Die Bewertungen hinsichtlich der Frage, wie nützlich das OSA zur Studienwahl/Orientierung war, fielen eher positiv aus.

Stärken und Limitationen

Bei den Fächern, die generell kleine Startkohorten haben ($n < 100$), sind die Fallzahlen in der Studieneingangsbefragung klein. Folglich sind die Bewertungen wenig belastbar. Obgleich die umfangreiche Studieneingangsbefragung der JLU regelmäßig durchgeführt wird und einen guten Rücklauf aufweist, können so nur sehr wenige OSA-Nutzerinnen und -Nutzer befragt werden. Zum einen liegt das daran, dass im OSA-Angebot der JLU „kleine“ Fächer vertreten waren. Zum anderen aber auch daran, dass nicht alle OSA-Nutzerinnen und -Nutzer sich nach der Nutzung an der JLU bewarben, einen Studienplatz bekamen, sich immatrikulierten und Monate später an der Studieneingangsbefragung teilnahmen, wobei sie sich dann an die OSA-Nutzung erinnern konnten und den Fragebogen der Studieneingangsbefragung auch bis zu den das OSA betreffenden Items ausfüllten.

Dass es relativ große Schwankungen in den Zugriffszahlen geben kann, ist eine Beobachtung, die wir explizit teilen möchten. Zwar kann über verschiedene Erklärungen dafür spekuliert werden; allerdings kann keine davon auf Basis der berichteten Daten be- oder widerlegt werden. Dass wir diese durchaus kritisch zu sehende Herangehensweise zur Klärung der Frage nach der Nutzung und der Nützlichkeit berichten, soll hier im Sinne der „Kopfstandmethode“ vor allem zeigen: Trotz der an der JLU vorhandenen sehr guten Infrastruktur in Bezug auf die Studieneingangsbefragung ist die indirekte Analyse der Nutzungshäufigkeit und Nützlichkeit wesentlich ungenauer, als eine Analyse auf Basis einer direkten (anonymisierten) Verknüpfung der OSA-Daten mit Studienverlaufsdaten.

Angaben zur Nutzung und wahrgenommene Nützlichkeit durch Personen, die nicht an der OSA-durchführenden Hochschule studieren, sind allerdings sowohl in den vorliegenden Daten als auch bei einer verpflichtenden OSA-Nutzung nicht vorhanden. Das bedeutet, dass möglicherweise ein Selektionsbias vorliegt, falls die Entscheidung für eine Bewerbung nicht unabhängig von der wahrgenommenen Nützlichkeit des OSA ist. Zudem wurden die Studieneinsteigerinnen und -einsteiger erst viele Wochen nach der OSA-Nutzung zu dieser befragt. Verzerrungen in der Bewertung der Nützlichkeit sind dadurch möglich.

Die Ergebnisse der Studie können wir nur berichten, aber nicht vergleichend interpretieren. Wie die beobachteten mittleren Ratings der Nützlichkeit zu bewerten sind (Was wäre als „positives Ergebnis“ einzustufen?), hängt vom Referenzrahmen ab. Da es bislang unserer Kenntnis

nach keine vergleichbaren publizierten Evaluationsergebnisse gibt, fällt die Einordnung schwer. Zudem ist weitere Forschung nötig, um der Frage nachzugehen, inwieweit verschiedene OSA-Angebote (verschiedene Fächer/Standorte) vergleichbar sind. Wünschenswert wären größere, hochschulübergreifende Projekte, anhand derer mit jeweils gleichen Messinstrumenten verschiedene Aspekte fächer- und hochschulübergreifend evaluiert werden.

Unser Fazit zu Studie 1 lautet, dass die Nutzung von Daten aus der Studieneingangsbefragung zur OSA-Evaluation mit erheblichen Nachteilen behaftet und daher nur eingeschränkt empfehlenswert ist.

Studie 2

Zusammenfassung

In Bezug auf die Ziele des OSA, studieninteressierten Personen die Möglichkeit zu geben, sich genauer mit den Voraussetzungen des jeweiligen Studiums und ihrer Passung hierzu zu beschäftigen, lässt sich festhalten, dass diese bei einem Großteil der OSA-Nutzerinnen und -Nutzer der JLU erreicht wurden: Die formative Evaluation der wahrgenommenen Nützlichkeit ergab als Globalurteil (Schulnote) für das OSA der JLU im Mittel Schulnoten zwischen „gut“ bis „befriedigend“, mit einem Median von „gut“ (Schulnote 2). Insbesondere mit Blick auf die angebotenen Module ordnen wir diese Bewertung als zufriedenstellend ein. Allerdings kann diese Einordnung kritisch gesehen werden. Mangels gegenwärtiger weiterführender Forschung, in welcher Hinsicht Evaluationen verschiedener OSA vergleichbar sein können, fehlt der Maßstab. Auch liegen unserer Kenntnis nach keine allgemein anerkannten Referenzwerte vor, um OSA-Evaluationsergebnisse in einem größeren Rahmen kriterienorientiert als „gut“ oder „weniger gut“ einzuordnen. Belastbar ist aber der von uns vorgenommene Vergleich der OSA-Angebote für verschiedene Fächer. Solche Vergleiche ermöglichen die Identifikation von quantitativen Bewertungsunterschieden, die dann in qualitativen Folgestudien aufgeklärt werden können, womit die Grundlage für gezielte Optimierungen geschaffen wird.

Stärken und Limitationen

Eine Stärke dieser Studie ist, dass die Evaluation direkt nach der Nutzung des OSA erfolgte. Retrospektive Verzerrungen, wie man sie bei Studie 1 befürchten könnte, sind daher weitestgehend auszuschließen. Zudem wurden über insgesamt mehr als sechs Jahre Evaluationsdaten gesammelt, sodass sich diese Befunde auf eine solide Datenbasis stützen. Hinzu kommt die Differenziertheit der Beurteilung: Die verschiedenen eingangs dargelegten Ziele von OSA wurden – zusätzlich zu einem Globalurteil

(Schulnotenskala) – separat mit Einzelitems erfragt, was ein differenziertes Bild der wahrgenommenen Nützlichkeit zeichnet. Darüber hinaus können Vergleiche zwischen den Beurteilungen einzelner fachspezifischer OSA Hinweise darauf liefern, welche Aspekte der Nützlichkeit positiv bewertet werden und an welchen Stellen ggf. Optimierungsbedarf besteht. Gleichzeitig kann es sinnvoll sein, eine genauere Analyse der Gründe für etwaige Unzufriedenheit der OSA-Nutzerinnen und -Nutzer bei eben diesen zu erfragen. Es bietet sich hierbei an, auch Freitext-Kommentarfelder als Möglichkeit, dem OSA-Team Feedback zu geben, anzubieten und auszuwerten. Zudem kann man Nutzerinnen und Nutzer anfragen, ob sie bereit wären als Interviewpartnerinnen bzw. -partner im Rahmen einer qualitativen Studie zu verschiedenen Aspekten der User Experience befragt zu werden.

Gleichwohl in Studie 2 bereits verschiedene Aspekte erfragt wurden, könnte die Messung detaillierter sein, indem für jeden der genannten Aspekte nicht nur ein Item, sondern mehrere eingesetzt werden. Für die zukünftige Praxis möchten wir anregen, mit Prä-Post-Messungen zu arbeiten, indem man vor der eigentlichen OSA-Bearbeitung z. B. fragt: „Wie gut informiert fühlst du dich in Bezug auf die Voraussetzungen dieses Studiums?“. Nach der OSA-Bearbeitung würde man das Pendant-Item vorgeben und die Differenz der Antworten analysieren.

Studie 3

Zusammenfassung

Es zeigen sich differenzierte Akzeptanz-Ratings gegenüber den drei typischerweise in OSA eingesetzten Testverfahren: Sowohl die beiden Intelligenztests als auch der Persönlichkeitsfragebogen erhielten als Akzeptanz-Globalurteil Schulnoten im Bereich von „gut“ bis „befriedigend“. Zwar sollte ein direkter Vergleich dieser Globalurteile mit Globalurteilen anderer Stichproben in Bezug auf andere Testverfahren nur mit Zurückhaltung erfolgen, dennoch lässt sich zumindest deskriptiv festhalten, dass die hier berichteten Befunde in dieselbe Richtung weisen wie die vorherigen Untersuchungen (Denker et al., 2023; Kersting, 2008). Passend zu früheren Befunden (Beermann et al., 2013) zeigt sich auch in unserer Studie, dass das Globalurteil, eine Verfahrensklasse sei beliebter als die andere, schwer haltbar ist. Vielmehr lohnt eine differenzierte Betrachtung wie sie hier mit Hilfe der Akzept-Fragebogen (Kersting, 2008) erfolgte. Testmodifikationen mit dem Ziel der Akzeptanzsteigerung sollten gemäß den vorliegenden Erkenntnissen bei der Augenscheinvalidität ansetzen. Diese lässt sich beispielsweise dadurch steigern, dass die Aufgaben spätere Studieninhalte aufgreifen (Denker et al., 2023).

Stärken und Limitationen

Eine Stärke dieser Untersuchung ist, dass die Personen nicht nur direkt vor dem Akzeptanz-Rating die jeweiligen Verfahren bearbeiteten (eine zwingende Voraussetzung für solide Akzeptanz-Untersuchungen; vgl. Kersting, 2008), sondern darüber hinaus auch Personen aus der Stichprobe im Within-Subject-Design mindestens zwei oder sogar alle drei Verfahren unter gleichen Durchführungsbedingungen bearbeiteten. Vergleiche, z. B. zwischen den Ratings bezüglich der beiden Intelligenztests, sind durchaus belastbar. Hierbei zeigt der Vergleich der Schulnoten als Globalurteil je Testverfahren, dass der Test auf fluide Intelligenz (IST-Screening, Liepmann et al., 2012) von den getesteten Personen besser bewertet wurde. Über die Ursachen des Akzeptanzvorteils des IST-Screenings können wir hier nur spekulieren. Der Test sieht Items mit verbalem, numerischem und figuralem Inhalt vor – ggf. wird dies im Kontrast zu den „nur“ verbalen Aufgaben des MWT-B (Lehrl, 2005) als positiv erlebt. Möglicherweise sind Wissenstest aber grundsätzlich unbeliebter als Tests zum schlussfolgernden Denken. Um dies weiter zu explorieren, könnte dieser quantitativen Analyse eine qualitative Untersuchung folgen, indem Interviews mit Personen aus der OSA-Zielgruppe geführt werden. Einschränkend muss darauf hingewiesen werden, dass es kein fachspezifischer Wissenstest war – der ggf. eine höhere Akzeptanz findet. Eine weitere Limitation ist, dass exemplarisch eine Gruppe von Psychologie-Studentinnen und -Studenten teilnahmen. Vorsicht ist geboten bei der Generalisierung der Akzeptanzurteile dieser Gruppe auf alle Studentinnen und Studenten. Beispielsweise könnte man vermuten, dass die hier untersuchte Gruppe grundsätzlich durch ihr fachliches Interesse Tests und Fragebogen gegenüber besonders positiv oder besonders kritisch eingestellt ist. Zudem war es nur eine vergleichsweise kleine Teilstichprobe, von der Akzeptanzratings zu allen drei Verfahren vorlag. Welchen Effekt dieses Muster von Missings bewirkt, kann an dieser Stelle nicht aufgeklärt werden. Die Belastbarkeit der Befunde ist folglich eingeschränkt.

Die Erhebung von spezifischen Akzeptanz-Werten können wir insgesamt empfehlen: Die Akzeptanz der eingesetzten Verfahren ist eine notwendige Voraussetzung dafür, dass freiwillige OSA bearbeitet werden.

Gesamt-Fazit: Take-Home-Message zur Planung eigener OSA-Evaluationen

Abschließend möchten wir die berichteten Befunde in der Gesamtschau reflektieren und dabei sowohl inhaltliche als auch strategische Implikationen für die OSA-Evaluation ableiten.

Wer ein OSA aufbaut, sollte sich frühzeitig Gedanken um eine ausführliche Evaluationsstrategie machen (Weis et al., 2022). Die Kombination verschiedener Evaluations-

ansätze angelehnt an den argumentationsbasierten Ansatz der Validierung gemäß den Standards für pädagogisches und psychologisches Testen (vgl. Hartig, Frey, & Jude, 2020) ist empfehlenswert.

Wie Studie 1 zeigte, lässt sich kein fächerübergreifender Nutzungstrend unter den Studieneinsteigerinnen und -einstiegern der JLU in den untersuchten Kohorten erkennen. Beobachtete Schwankungen in den Zugriffszahlen und Nützlichkeitsratings sind auf Basis dieses Datensatzes schwer nachvollziehbar. Zudem unterliegen die Daten potenziell retrospektiver Verzerrung. Ein Teil der OSA-Nutzerinnen und -Nutzer, nämlich diejenigen, die kein Studium an der OSA anbietenden Hochschule aufnehmen, werden nicht erfasst. Obgleich die Daten der Studieneingangsbefragung grundsätzlich sehr wertvoll sind, sehen wir ihren Nutzen als eine (indirekte) Quelle zur OSA-Evaluation kritisch.

Grundsätzlich lässt sich allerdings aus der insgesamt niedrigen Nutzung, die sich in unserer und in anderen Studien (z.B. Päßler, 2023) zeigt, ableiten, dass nicht nur Ressourcen für die Erstellung eines OSA, sondern auch für das Marketing bereitgestellt werden sollten. Marketingmaßnahmen sollten in der Folge wiederum evaluiert werden, um die Ressourcen zukünftig mit Fokus auf die strategisch günstigsten Maßnahmen zu bündeln. Allgemein lässt sich festhalten: OSA-Verantwortliche täten gut daran, das Angebot innerhalb der Zielgruppe bekannt(er) zu machen. Hierfür ist es unter anderem wichtig, Multiplikatorinnen und Multiplikatoren über das Angebot zu informieren, damit diese dann gezielt auf die OSA verweisen können.

Im Zusammenhang mit möglichen Schul-Hochschul-Kooperationen könnte man auch eine große Evaluationsstudie, längsschnittlich über mehrere Jahre, konzipieren, im Rahmen derer die Teilnehmerinnen und Teilnehmer im Schulabschlussjahr OSA bearbeiten und über den Studieneinstieg hinweg wiederholt befragt werden, sowie der Nutzung von objektiven Studienverlaufsdaten zustimmen. Solange solche „vollständigen“ Designs in der Praxis noch nicht umgesetzt werden können, bleibt nur, bei der Interpretation sämtlicher nicht auf diesem Wege gewonnenen Ergebnisse davon auszugehen, dass eine „Range Restriction“ vorliegt. In den hier dargestellten drei Studien haben wir Stichproben aus verschiedenen Populationen (Studienanfängerinnen und -anfänger aller (OSA-)Fächer, OSA-Nutzerinnen und -Nutzer, Studieneinsteigerinnen und -einstieger eines exemplarisch ausgewählten Faches) analysiert, um zumindest in dieser Hinsicht verschiedene Perspektiven einzubeziehen.

Wie am Beispiel der JLU gezeigt, können im Zuge der formativen Evaluation verschiedene Ansätze relativ ökonomisch umgesetzt werden. Evaluationsitems im OSA selbst sind einfach umzusetzen und liefern ein direktes, wertvolles Feedback der Nutzerinnen und Nutzer (Studie

2). Dabei ist allerdings zu beachten, dass die Art des im OSA generierten individuellen Feedbacks an die OSA-Teilnehmerinnen und -Teilnehmer von großer Bedeutung ist (Persike, Zay, Ewert, Wachtel & Teige-Mocigemba, 2022; Petri, Schütte & Beauducel, 2022). Personen, bei denen sich eine geringe Passung zum (Wunsch-)Studium zeigt, erhalten eine Rückmeldung, die Diskrepanzerlebnisse auslösen oder sogar eine Abwehrhaltung anstoßen kann. Dies kann dazu führen, dass die Personen zum Selbstwertschutz das OSA abwerten.

Bei der OSA-internen Evaluation sollte man sich einerseits an der bestehenden Literatur orientieren – sowohl in Bezug auf Evaluationsinstrumente als auch in Bezug auf den Vergleich der eigenen erzielten Erkenntnisse mit denen aus anderen OSA (Studie 1 und 3) – zugleich aber berücksichtigen, dass jedes OSA individuell ist: Die veranschlagte Bearbeitungsdauer und die Inhalte können entscheidend dafür sein, wie Evaluationen und Akzeptanz-Ratings ausfallen (vgl. Studie 2 und 3).

Untersuchungen zur Bekanntheit des Angebots in der Zielgruppe (= studieninteressierte Personen) sind schwierig zu gestalten, da diese Zielgruppe nicht zentral (z.B. in einer Schule) zu erreichen ist. Zugleich wäre es zwecks Steigerung des Bekanntheitsgrades („OSA können nur nützen, wenn sie genutzt werden.“) aber offensichtlich notwendig, bei der Zielgruppe für die OSA-Nutzung zu werben und die Qualität des Angebots zu evaluieren. Evaluationen, die anhand der Daten von Studentinnen und Studenten einer Hochschule vorgenommen werden, sind unbefriedigend, weil hier weder die Personen erfasst werden, die sich gegen ein Studium entscheiden, noch die, die an einer anderen Hochschule studieren.

Als Empfehlung für im Aufbau befindliche OSA möchten wir daher resümieren, dass es wichtig ist, von Beginn an neben der elaborierten Evaluationsstrategie, auch ein Marketing-Konzept zu entwickeln und Zugänge zur Zielgruppe zu sichern. So sollten sich zumindest einige der hier berichteten Hürden bei der OSA-Evaluation, die sich durch ungünstige Rahmenbedingungen für das Sammeln längsschnittlicher und verknüpfter Datensätze ergeben, vermeiden lassen.

Elektronische Supplemente (ESM)

Das elektronische Supplement ist mit der Online-Version dieses Artikels verfügbar unter <https://doi.org/10.1024/1010-0652/a000382>

ESM 1. OSA-Angebot der JLU: Fächer und Angebotsverfügbarkeit (Tabelle E1); Darstellung im OSA zur Erläuterung des modularen Aufbaus, hier am Beispiel des Chemie-OSAs (Abbildung E1).

Literatur

- Anderson, N., Salgado, J. F. & Hülsheger, U. R. (2010). Applicant Reactions in Selection: Comprehensive Meta-Analysis into Reaction Generalization Versus Situational Specificity. *International Journal of Selection and Assessment*, 18, 291 – 304.
- Beermann, D., Kersting, M., Stegt, S. & Zimmerhofer, A. (2013). Vorurteile und Urteile zur Akzeptanz von Persönlichkeitsfragebogen. *PersonalQuarterly*, 65, 41 – 45.
- Denker, M., Schütte, C., Kersting, M., Weppert, D. & Stegt, S. J. (2023). How can applicants' reactions to scholastic aptitude tests be improved? A closer look at specific and general tests. *Frontiers in Education*, 7, 931841. <https://doi.org/10.3389/feduc.2022.931841>
- Ehrlich, S. (2020). *Studieneingangsbefragung WS 2019/20. Auswertung pro Studiengang*. Gießen: Servicestelle Lehrevaluation der Justus-Liebig-Universität. Verfügbar unter: <https://www.uni-giessen.de/de/org/admin/stab/stl/servicestelle/datenportal/datenste>
- Gehrlach, C., Bergen, M. von & Eiler, K. (Hrsg.). (2022). *Zwischen gesellschaftlichem Auftrag und Wettbewerb. Sozialmanagement und Sozialwirtschaft in einem sich wandelnden Umfeld* (Perspektiven Sozialwirtschaft und Sozialmanagement). Wiesbaden: Springer VS. <https://doi.org/10.1007/978-3-658-35381-0>
- Hartig, J., Frey, A. & Jude, N. (2020). Validität von Testwertinterpretationen. In: Moosbrugger, H., Kelava, A. (Hrsg.) *Testtheorie und Fragebogenkonstruktion*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-61532-4_21
- Hasenberg, S. & Schmidt-Atzert, L. (2014). Internetbasierte Selbsttests zur Studienorientierung. *Beiträge zur Hochschulforschung*, 36(1), 8 – 28.
- Hausknecht, J. P., Day, D. V. & Thomas, S. C. (2004). Applicant Reactions to Selection Procedures: An Updated Model and Meta-Analysis. *Personnel Psychology*, 57(3), 639 – 683. <https://doi.org/10.1111/j.1744-6570.2004.00003.x>
- Heublein, U., Richter, J. & Schmelzer, R. (2020). *Die Entwicklung der Studienabbruchquoten in Deutschland*. https://doi.org/10.34878/2020.03.DZHW_BRIEF
- Hiemisch, A., Westermann, R. & Michael, A. (2005). Die Abhängigkeit der Zufriedenheit mit dem Medizinstudium von Studienzielen und ihrer Realisierbarkeit. *Zeitschrift für Psychologie*, 213(2), 97 – 108. <https://doi.org/10.1026/0044-3409.213.2.97>
- Hochschulrektorenkonferenz. (2021). *Statistische Daten zu Studienangeboten an Hochschulen in Deutschland – Studiengänge, Studierende, Absolventinnen und Absolventen. Wintersemester 2021/2022*. Berlin.
- IBM Corp. (2012). *SPSS Statistics for Windows, Version 21.0*. Armonk, NY.
- Kersting, M. (2008). Zur Akzeptanz von Intelligenz- und Leistungstests. *Report Psychologie*, 33, 420 – 433.
- Kersting, M. & Marquardt, C. (2022). Tests, Fragebogen & Co – worauf es bei diagnostischen Verfahren für OSA ankommt. In G. Stoll & S. Weis (Hrsg.), *Online-Self-Assessments zur Studienfachwahl. Entwicklung – Konzepte – Qualitätsstandards* (S. 107 – 129). Berlin: Springer.
- 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 for Educational Research Online*, 9(2), 205 – 227.
- Lehrl, S. (2005). *MWT-B – Mehrfachwahl-Wortschatz-Intelligenztest. 5., unveränderte Auflage*. Balingen: Spitta.
- Liepmann, D., Beauducel, A., Brocke, B. & Nettelstroth, W. (2012). *IST-Screening. Intelligenz-Struktur-Test*. Göttingen: Hogrefe.
- Messerer, L. A. S., Karst, K. & Janke, S. (2023). Choose wisely: intrinsic motivation for enrollment is associated with ongoing intrinsic learning motivation, study success and dropout. *Studies in Higher Education*, 48(1), 137 – 150. <https://doi.org/10.1080/3075079.2022.2121814>
- Neugebauer, M., Daniel, H.-D. & Wolter, A. (2021). *Studienerfolg und Studienabbruch*. Wiesbaden: Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-32892-4>
- Päßler, K. (2023). Die Integration von Online-Self-Assessments in die Studien- und Berufsberatung. In B. Knickrehm, T. Fletemeyer & B.-J. Ertelt (Hrsg.), *Berufliche Orientierung und Beratung. Aktuelle Herausforderungen und digitale Unterstützungsmöglichkeiten* (1. Auflage 2023, S. 359 – 371). Wiesbaden: Springer Fachmedien Wiesbaden GmbH. https://doi.org/10.1007/978-3-658-40601-1_21
- Persike, M., Zay, K., Ewert, M., Wachtel, T. & Teige-Mocigemba, S. (2022). Die Rückmeldung als zentrales Element von Online-Self-Assessments: Funktionen, Gütekriterien und Gestaltungsempfehlungen. In G. Stoll & S. Weis (Hrsg.), *Online-Self-Assessments zur Studienfachwahl. Entwicklung – Konzepte – Qualitätsstandards* (S. 153 – 181). Berlin: Springer. https://doi.org/10.1007/978-3-662-63827-9_7
- Petri, P. S., Schütte, N. & Beauducel, A. (2022). Mit OSA erfasste Personenmerkmale und deren Interaktion mit OSA-Rückmeldungen und Informationselementen. In G. Stoll & S. Weis (Hrsg.), *Online-Self-Assessments zur Studienfachwahl. Entwicklung – Konzepte – Qualitätsstandards*. Berlin: Springer.
- Praslova, L. (2010). Adaptation of Kirkpatrick's four level model of training criteria to assessment of learning outcomes and program evaluation in Higher Education. *Educational Assessment, Evaluation and Accountability*, 22(3), 215 – 225. <https://doi.org/10.1007/s11092-010-9098-7>
- Rammstedt, B. & John, O. P. (2005). Kurzversion des Big Five Inventory (BFI-K). *Diagnostica*, 51(4), 195 – 206. <https://doi.org/10.1026/0012-1924.51.4.195>
- Statistisches Bundesamt. (2022). *Wintersemester 2022/2023: Erstmals seit 15 Jahren weniger Studierende als im Vorjahr. Pressemitteilung Nr. 503*. Verfügbar unter: https://www.destatis.de/DE/Presse/Pressemitteilungen/2022/11/PD22_503_21.html
- Stoll, G. & Weis, S. (Hrsg.). (2022). *Online-Self-Assessments zur Studienfachwahl. Entwicklung – Konzepte – Qualitätsstandards*. Berlin: Springer. <https://doi.org/10.1007/978-3-662-63827-9>
- Stoll, G., Weis, S., Beauducel, A., Becker, A., Brazel, C., Delzepich, R. et al. (2022). Ein Rahmenmodell zu Entwicklung, Wirkmechanismen und Bewertung von Online Self-Assessments. In G. Stoll & S. Weis (Hrsg.), *Online-Self-Assessments zur Studienfachwahl. Entwicklung – Konzepte – Qualitätsstandards*. Berlin: Springer.
- Stoll, G. & Gfrörer, T. (2023, März). Anregung zur Selbstselektion und Steigerung der Grundquote – Welchen Beitrag leisten Online-Self-Assessments im Kontext der Studierendenauswahl? Paper präsentiert auf der 10. GEBF-Tagung [Gesellschaft für Bildungsforschung], Essen, Deutschland.
- Weis, S., Stoll, G., Dietrich, S. & Wachtel, T. (2022). Evaluation von Online-Self-Assessments. In G. Stoll & S. Weis (Hrsg.), *Online-Self-Assessments zur Studienfachwahl. Entwicklung – Konzepte – Qualitätsstandards* (S. 203 – 230). Berlin: Springer. https://doi.org/10.1007/978-3-662-63827-9_9
- Wiesenhütter, E. (2022). *Studienerfolg in der Psychologie: Welche Vorhersagekraft haben die fluide und kristalline Intelligenz sowie die Gewissenhaftigkeit über die Abiturnote hinaus?* Unveröffentlichte Bachelorthesis, Justus-Liebig-Universität Gießen.
- Wittmann, W. W. (1990). Brunswik-Symmetrie und die Konzeption der Fünf-Datenboxen – Ein Rahmenkonzept für umfassende Evaluationsforschung. *Zeitschrift für Pädagogische Psychologie*, 4(4), 241 – 251.

Historie

Manuskript eingereicht: 17.01.2023

Manuskript nach Revision angenommen: 02.04.2024


Onlineveröffentlichung: 19.04.2024

Förderung


Open-Access-Veröffentlichung ermöglicht durch die Hochschule für Wirtschaft und Umwelt Nürtingen-Geislingen.

ORCID

Pascale Stephanie Petri

 <https://orcid.org/0000-0002-4471-2515>

Martin Kersting

 <https://orcid.org/0000-0003-2501-5287>

Pascale Stephanie Petri

Hochschule für Wirtschaft und Umwelt

Parkstraße 4

73312 Geislingen

Deutschland

pascale.petri@hfwu.de