

Metakognitives Strategiewissen in sprachbezogenen Situationen

Interne Struktur und Validität des ScenEx

Jennifer Seeger¹ , Wolfgang Lenhard¹ und Katrin Wisniewski²

¹Institut für Psychologie, Fakultät für Humanwissenschaften, Julius-Maximilians-Universität Würzburg

²Herder-Institut, Philologische Fakultät, Universität Leipzig

Zusammenfassung: Studieren stellt hohe Anforderungen an selbstregulatorische Fähigkeiten und eigenverantwortlichen Umgang mit schwierigen Situationen. Aus den zusätzlichen sprachlichen Barrieren für ausländische Studierende erwachsen spezifische selbstregulatorische Aufgaben, wie der Umgang mit Verständnisproblemen in Vorlesungen. Da hierfür bisher kaum geeignete Erhebungsinstrumente existieren, versucht ScenEx diese Lücke zu schließen. Der Test erfasst das metakognitive Strategiewissen in sprachlich herausfordernden Situationen im Studienalltag. Anhand einer Stichprobe von 290 ausländischen Studierenden im ersten Fachsemester wird die psychometrische Qualität und interne Struktur des Instruments überprüft. ScenEx zeigt eine zufriedenstellende interne Konsistenz und gute Itemfit-Kennwerte, erwartungskonform liegen lokale stochastische Abhängigkeiten der Aufgaben innerhalb der Szenarien vor. Eine konfirmatorische Faktorenanalyse bestätigt die Grobstruktur der Szenarien und des Gesamtscores des Tests. Das Verfahren ist für die weitere Entwicklung der Sprachkompetenz über die anfängliche Sprachfähigkeit hinaus prädiktiv. ScenEx erweist sich insgesamt als ein reliables und valides Instrument zur Erfassung des Strategiewissens in schwierigen Situationen im Studium.

Schlüsselwörter: Selbstregulation, Strategiewissen, Bildungsausländer, Studienerfolg, Deutsch als Fremdsprache

Metacognitive Strategic Knowledge in Language-Related Situations

Abstract: Studying at a university requires a high degree of self-reliance. Self-regulatory skills and the ability to independently handle difficult situations are important for successfully completing a degree. In addition, foreign students are faced with language-based challenges demanding specific self-regulatory skills, such as dealing with comprehension problems in lectures. To date, assessing these skills has been difficult because of a lack of suitable instruments. The ScenEx Questionnaire purports to close this gap by measuring metacognitive strategic knowledge in everyday study situations by posing language-based challenges, presented as scenarios. Using a sample of 290 foreign students in their first semester, we assessed the psychometric quality of ScenEx. ScenEx shows a satisfactory internal consistency and good item fit. A confirmatory factor analysis confirms the structure of the scenarios and the overall score of the test. The procedure predicts the further development of language competence beyond the initial stage. ScenEx is a valid and reliable instrument for assessing strategic knowledge regarding linguistically difficult situations at the university.

Keywords: self-regulation, strategic knowledge, foreign students, academic success

Der Umgang mit herausfordernden Situationen gehört unabweichlich zum Studienalltag von Studierenden. Die Fähigkeit, angemessen und effektiv auf schwierige Situationen zu reagieren, wie beispielsweise eine Vorlesung, der man nicht gut folgen kann oder eine Klausuraufgabe, die man nicht vollständig versteht, ist für den Studienerfolg von großer Bedeutung. Diese Art der Herausforderungen häufen sich besonders für die Gruppe der ausländischen Studierenden, für welche zu inhaltlichen Schwierigkeiten hohe sprachliche Hürden hinzukommen können.

Im internationalen Vergleich stellt sich die Gestaltung des Studiums in Deutschland besonders an den Universitäten als sehr offen und unverschult dar. Gerade ausländische Studierende laufen Gefahr, von der weitgehend eigenständigen Einteilung der Lernzeiten und der freien Kurswahl überfordert zu werden. Ferner ist davon auszugehen, dass insbesondere kommunikative und wissenschaftssprachliche Anforderungen in der Fremdsprache Deutsch ein Hindernis für ein erfolgreiches Studium darstellen können (Bärenfänger, Lange & Möhring, 2016; Heublein, 2015; Stemmer, 2013; Wisniewski, 2018) und zu den im Vergleich zu deutschen Studierenden sehr hohen Abbruchquoten entscheidend beitragen könnten. Zur Bewältigung dieses herausfordernden Studienalltags sind selbstregulatorische Fähigkeiten unabdingbar, um mit den Anforderungen und der Aufgabenfülle Schritt halten zu

dische Studierende laufen Gefahr, von der weitgehend eigenständigen Einteilung der Lernzeiten und der freien Kurswahl überfordert zu werden. Ferner ist davon auszugehen, dass insbesondere kommunikative und wissenschaftssprachliche Anforderungen in der Fremdsprache Deutsch ein Hindernis für ein erfolgreiches Studium darstellen können (Bärenfänger, Lange & Möhring, 2016; Heublein, 2015; Stemmer, 2013; Wisniewski, 2018) und zu den im Vergleich zu deutschen Studierenden sehr hohen Abbruchquoten entscheidend beitragen könnten. Zur Bewältigung dieses herausfordernden Studienalltags sind selbstregulatorische Fähigkeiten unabdingbar, um mit den Anforderungen und der Aufgabenfülle Schritt halten zu

können (Heublein, 2015). Hinsichtlich der studienbezogenen metakognitiven Fähigkeiten bzw. des selbstregulatorischen Verhaltens steht in der psychologischen Forschung vor allem die Erfassung der selbstberichteten Anwendungshäufigkeit von Lernstrategien im Vordergrund (Pintrich, Smith, Garcia & McKeachie, 1993; Wild & Schiefele, 1994). Die selbstberichtete Strategieverwendung wird dabei in kognitive, metakognitive und ressourcenbezogene Stützstrategien unterteilt. Eine hohe Anwendungshäufigkeit von Lernstrategien wird meist als wünschenswert angesehen, auch wenn der Selbstbericht nicht zwangsläufig die tatsächliche Anwendung vorhersagt. Handlungsnahe Erhebungen von Lernstrategien verweisen zudem nicht zwingend auf einen direkten Einfluss von Strategien, sondern kennzeichnen diese als Mediator zwischen emotionalen und motivationalen Variablen und Studienerfolg (Artelt, 1999). Längsschnittlich finden sich relativ geringe Zusammenhänge zwischen der retrospektiv berichteten Anwendungshäufigkeit der Strategien und der tatsächlichen Anwendungshäufigkeit sowie niedrige Zusammenhänge mit dem tatsächlichen Studien- oder Lernerfolg (Artelt, 2000; Leutner & Leopold, 2003; Schiefele, Streblo & Brinkmann, 2007; Waldeyer, Fleischer, Wirth & Leutner, 2017). Teilweise lassen sich Auswirkungen strategischen und selbstregulatorischen Verhaltens auf die Leistungsentwicklung nachweisen, wenn der Lernerfolg differenziell betrachtet wird, beispielsweise durch Differenzierung des Lernerfolges in „Anwendungsfeld“ und „Lernfeld“ bei berufstätigen Studierenden an Fachhochschulen. Boerner, Seeber, Keller und Beinborn (2005) fanden hier einen Zusammenhang der berichteten Anwendungshäufigkeit von Lernstrategien und dem Erfolg im „Lernfeld“, also der nicht praxisbezogenen Verwendung des Erlernten an der Hochschule. Diese gemischten bis schwachen Befunde werfen Fragen bezüglich der Validität der bisherigen Erfassung von Lernstrategieinsatz und der Vorhersage des tatsächlichen Lern- und Studienerfolgs auf. Ein möglicher Erklärungsansatz findet sich in der retrospektiven Befragung, die von den Teilnehmenden ein hohes Maß an Erinnerungsleistung fordert und deshalb leicht verzerrt sein kann (Neuenhaus, Artelt & Schneider, 2017), bei gleichzeitig fehlenden objektiven Vergleichsstandards. Eine möglichst handlungsnahe Erfassung selbstregulatorischer Fähigkeiten ist deshalb indiziert, aber im Falle von Beobachtungsstudien und Lernprotokollen sehr aufwendig. Vielversprechend scheint dagegen die Erfassung metakognitiven Strategiewissens in Form eines Leistungstests, der auf die Einschätzung der Situationsangemessenheit bestimmter Strategien in spezifischen Situationen abzielt (Hasselhorn & Labuhn, 2008; Schneider & Artelt, 2010).

Das Projekt „Sprache und Studienerfolg bei Bildungsausländer/-innen“ (SpraStu) untersucht den Studienerfolg

ausländischer Bachelor-Studierender in Deutschland im Längsschnitt und in einem multifaktoriellen Design (Wisniewski, 2018). Hierzu gehört auch das selbstregulierte Lernen und das angemessene Reagieren auf sprachlich herausfordernde Situationen im Studienalltag, die als Einflussfaktoren auf ein gelingendes Studium ausländischer Studierender angenommen werden. Aufgrund der unklaren Validität bisheriger Erhebungsinstrumente und dem fehlenden sprachlichen Bezug wurde im Rahmen des Projekts ein neues Instrument zur Erfassung von metakognitivem Strategiewissen zur Anpassung an schwierige Situationen im Studienalltag entwickelt (Wisniewski, Parker, Lenhard & Seeger 2019). Zwar existiert in der Fremdsprachenlehr- und -lernforschung eine lange Tradition der Erforschung der Strategieverwendung (Rubin, 1975; Oxford, 1990), und seit einigen Jahren werden zunehmend auch Bezüge zu psychologischen Modellen selbstregulierten Lernens hergestellt (Dörnyei & Ryan, 2015), trotzdem fehlen nach wie vor validierte quantitative Erhebungsinstrumente. Die Pädagogische Psychologie wiederum verfügt über innovative szenariobasierte und validierte Verfahren, die bei der Entwicklung des ScenEx („Scenario-based executive metacognition“) Orientierung boten (Schlagmüller & Schneider, 2007; Neuenhaus et al., 2017), aber inhaltlich nicht spezifisch für Fremdsprachenkontexte sind.

Wir versuchen, diese Lücke durch die Entwicklung des Testinstruments ScenEx zu schließen, um dem sprachlichen Aspekt des Studiums in einer Fremdsprache Rechnung zu tragen und somit neben dem Fokus auf eine handlungsnahe Erfassung ein kontextspezifisches Instrument für ausländische Studierende zu bieten. Das Verfahren orientiert sich in seiner Herangehensweise an Ansätzen zur Erfassung des Strategiewissens im Lesen (WLST 7–12; Schlagmüller & Schneider, 2007), in der Mathematik (Lingel, Götz, Artelt & Schneider, 2014; Lingel, 2016) und dem Englischerwerb (Neuenhaus et al. 2017). Anders als bei Fragebögen zur selbstberichteten Strategieverwendung wird nicht nach der Häufigkeit der Anwendung bestimmter Lernstrategien gefragt, sondern es wird das Wissen über die Effektivität von Strategien in bestimmten sprachrelevanten Lern- und Leistungssituationen erfasst. Auf der Grundlage schwieriger Situationen im Studienalltag sind verschiedene Strategien auf ihre Angemessenheit und Effektivität in dem entsprechenden Szenario zu bewerten und diese Bewertungen werden untereinander paarweise verglichen. Für die Bewertung des Antwortverhaltens werden diese paarweisen Vergleiche mit zuvor erhobenen Ratings von Expertinnen und Experten abgeglichen und ermittelt, ob die Bewertung durch die Testpersonen mit dem Expertenwissen korrespondiert. ScenEx misst also Wissen über die Angemessenheit von Strategien in einer anwendungsbezogenen Perspektive, d.h. metakognitives Strategiewissen in Form eines

Tabelle 1. Die sechs Szenarien des ScenEx

Szenario 1

Sie sitzen in einer Vorlesung. Der Dozent/Die Dozentin spricht sehr schnell und kompliziert, Sie haben Schwierigkeiten ihn/sie zu verstehen. Was tun Sie?

Szenario 2

In einer Vorlesung erklärt der Dozent/die Dozentin etwas. Gleichzeitig versuchen Sie, den Text auf den PowerPoint-Folien zu lesen. Das ist zu viel auf einmal und Sie merken, dass Sie nicht mehr folgen können. Was tun Sie?

Szenario 3

Sie sollen einen Text zur Vorbereitung lesen. Der Text enthält sehr viele Informationen. Sie können nicht erkennen, was wichtig ist und was nicht. Was tun Sie?

Szenario 4

Sie müssen einen Text auf Deutsch für Ihr Studium schreiben, aber Sie haben Probleme, Ihre Gedanken auf Deutsch auszudrücken. Was tun Sie?

Szenario 5

Sie schreiben eine Klausur. Bei einer Multiple-Choice-Frage kommen Sie nicht weiter. Was tun Sie?

Szenario 6

In einer Klausur kommt eine Frage vor, die Sie beantworten sollen. Die Frage ist aber ziemlich lang und kompliziert formuliert. Sie verstehen nicht genau, was Sie machen sollen. Was tun Sie?

Anmerkungen: Szenarien im Wortlaut. Jeweils sieben bis acht Strategien zum Umgang mit der gegebenen Situation.

Wissenstests unter Bezugnahme auf einen qualitativen Standard in Form von Expertenwissen. Dabei wurde folgendermaßen vorgegangen (detaillierte Darstellung der theoretischen Einbettung siehe Wisniewski et al., 2019; Szenarien siehe Tab. 1):

1. Analyse sprachlich herausfordernder Situationen: Basierend v. a. auf der Bedarfsanalyse von Bärenfänger et al. (2016) wurden sechs verschiedene Situationen identifiziert, die für Studierende ohne deutsche Muttersprache besonders herausfordernd sind. Eine solche Situation stellt beispielsweise Szenario 2 dar: „In einer Vorlesung erklärt der Dozent etwas. Gleichzeitig versuchen Sie, den Text auf den PowerPoint-Folien zu lesen. Das ist zu viel auf einmal, und Sie merken, dass Sie nicht mehr folgen können. Was tun Sie?“

2. In einem zweiten Schritt wurden pro Szenario bis zu sieben mögliche, unterschiedlich zielführende Herangehensweisen formuliert. Im obigen Szenario gehören hierzu beispielsweise „Ich schlage Wörter, die ich nicht verstehe, sofort im Wörterbuch oder im Internet nach.“ oder „Ich überlege, welche Inhalte besonders wichtig sind, und konzentriere mich darauf“. Jedes Szenario enthält neben sinnvollen Herangehensweisen auch ungünstige Strategien, die wenig zielführend sind. Der Test als Ganzes ist als Elektronisches Supplement 1 (ESM1) hinterlegt.

3. Zur Gewinnung eines Vergleichsstandards wurden 32 Fachkolleginnen und -kollegen aus dem Bereich der pädagogischen Psychologie und der Fremdsprachdidaktik darum gebeten, jede Strategie auf einer 6-stufigen Likertskala von 1 (*überhaupt nicht hilfreich*) bis 6 (*sehr hilfreich zu bewerten*). Innerhalb jedes Szenarios wurden alle möglichen paarweisen Einzelvergleiche gebildet und jene Kombinationen ausgewählt, bei denen von den Experten

und Expertinnen eine Übereinstimmung von 95% aller Antworten vorlag. Dabei kommt es nicht auf die absolute Höhe des Ratings an, sondern lediglich darauf, dass eine Strategie eindeutig besser/schlechter bewertet wurde als die jeweils andere Strategie.

4. Dieser Vergleichsstandard dient der Bewertung der Antworten durch die Testpersonen. Jeder Einzelvergleich, der mit dem Experten-Vergleichsstandard übereinstimmt, wird mit einem Rohpunkt bewertet.

Die absolute Höhe der Einstufung einer Strategie ist also nicht von Bedeutung und geht nicht in die Analyse ein. Stattdessen wird ein Vergleich der Bewertungen der Strategien innerhalb des Szenarios vorgenommen. Die Berechnung der Gesamtscores erfolgt durch Paarvergleiche der Antworten untereinander. Dies führt zu insgesamt 52 eindeutigen Paarvergleichen, die in die finale Fassung des Instruments übernommen wurden. Die erreichbare Höchstpunktzahl liegt dementsprechend bei 52 Punkten. Innerhalb der Szenarien können 6 bis 11 Punkte erreicht werden. Während der Bearbeitung kann zwischen den Szenarien vor und zurück gewechselt werden, es gibt keine vorgegebene Zeitbegrenzung und die Bearbeitung kann sowohl auf Papier als auch digital stattfinden (Test und Auswertungssyntax in den Sprachen Deutsch und Englisch verfügbar unter: <https://osf.io/u9v5r/> sowie <https://go.uni-wue.de/scenex>).

Untersuchungsziele und Hypothesen

Unsere im Folgenden dargestellte Untersuchung zielt auf die Analyse der Messeigenschaften von ScenEx ab. Da-

mit der ScenEx gewinnbringend eingesetzt werden kann, muss das Verfahren über eine hohe Reliabilität und Homogenität verfügen, verschiedene Personengruppen fair bewerten, die zu messende latente Fähigkeit hinreichend genau abbilden und prognostisch valide sein. Die Konstruktion von ScenEx baut auf der Annahme auf, dass das sprachbezogene metakognitive Strategiewissen eine eindimensionale, latente Fähigkeit darstellt. Diese Grundannahme wird auf verschiedene Weise geprüft. Zum einen erwarten wir eine hohe interne Konsistenz des Fragebogens (Hypothese 1), ausgedrückt durch hohe Homogenitätskennwerte (Rammstedt, 2010). Des Weiteren berichten wir die Ergebnisse der Skalierung mittels eines einparametrischen-logistischen Modells (1-PL-Modell; Moosbrugger, 2012; Strobl, 2015) und überprüfen die Geltung der Rasch-Modellannahmen. So sollen die Items des ScenEx das metakognitive Strategiewissen in verschiedenen Teilstichproben gleichermaßen erfassen und der Gesamtscore sinnvoll interpretiert werden können (Geiser & Eid, 2010; Hypothese 2, Annahme spezifischer Objektivität). Spezifische Objektivität liegt dann vor, wenn die Schätzung der Itemparameter in unterschiedlichen Stichproben zu den gleichen Ergebnissen bzw. zu den gleichen Verhältnissen zwischen den Einzelitems führt.

Das Instrument besteht aus 6 Szenarien, deren Items aus Paarvergleichen der Strategien untereinander gebildet werden. Durch die gemeinsame Zugehörigkeit der Einzelvergleiche zu Szenarien und darüber hinaus durch die konstruktionsbedingte Abhängigkeit bei Paarvergleichen, kann eine Verletzung der lokalen stochastischen Unabhängigkeit in den Rohdaten nicht ausgeschlossen werden. Eine Verletzung der Unabhängigkeitsannahme soll deshalb innerhalb der Szenarien explizit geprüft werden (Rutsch, Vogel, Rehm & Dörfler, 2018; Hypothese 3). Gleichzeitig handelte es sich bei den Szenarien um unterschiedliche Situationen. Ob diese auf latenter Ebene ein eindimensionales Konstrukt ergeben und das Messmodell eine hinreichende Passungsgüte aufweist, wird mittels einer konfirmatorische Faktorenanalyse geprüft (Nachweis der Konstruktvalidität, Hypothese 4).

Das im ScenEx erfasste metakognitive Strategiewissen bildet sich in sprachlich herausfordernden Situationen ab. Zum einen erwarten wir, dass Personen mit guten Sprachkenntnissen gleichzeitig auch höhere metakognitive Fähigkeiten besitzen; einerseits, weil Sprachkompetenzen die Anwendung von Strategien erleichtern und zum anderen, da metakognitives Wissen den Spracherwerb unterstützt. Folglich sollte eine substantielle positive Korrelation zwischen den Sprachkompetenzen und dem metakognitiven Wissen existieren (Hypothese 5a) und gleichzeitig sollte im Längsschnitt das Strategiewissen die Sprachkompetenzen zu T2 (nach einem Studienjahr) über

die Sprachkompetenzen zu T1 (zu Studienbeginn) hinaus vorhersagen (Hypothese 5b).

Methode

Beschreibung der Stichprobe

Verteilt über drei Erhebungszeitpunkte jeweils zum Beginn des Wintersemesters in den Studienjahren 2017, 2018 und 2019 füllten insgesamt $N = 290$ ausländische Studierende im ersten Studiensemester (T1) an den Universitäten Leipzig und Würzburg den Fragebogen ScenEx aus (64 % weiblich, 36 % männlich, 2 % keine Angabe). An jedem der drei Erhebungszeitpunkte nahmen Studierende im ersten Studiensemester aus Bachelor- und Staatsexamensstudiengängen teil. Die Probandinnen und Probanden, die in einem Jahr bereits teilgenommen hatten, wurden für das folgende Jahr erneut kontaktiert und zur Teilnahme gebeten, um den Verlauf der erhobenen Kompetenzen erfassen zu können, sodass im Jahr 2018 die Kohorte aus 2017 und im Jahr 2019 die Kohorte aus 2018 erneut untersucht werden konnten (T2). Für 89 Teilnehmende liegen diese Daten zusätzlich in einer zweiten Messung im Abstand eines Jahres vor, sodass die Kompetenzentwicklung längsschnittlich untersucht werden konnte.

Aus Datenschutzgründen liegen für die Probandinnen und Probanden im ersten Studiensemester lediglich kategoriale Altersangaben vor, wobei sich die mit Abstand größte Gruppe im Altersbereich von 21 bis 25 Jahren befindet (43 %; deskriptive Daten siehe Tabelle 2). Die Studierenden waren in über 70 verschiedene Studienfächer immatrikuliert und wiesen 41 verschiedene Muttersprachen auf. Im Durchschnitt waren die Teilnehmenden zum Beginn der Studie seit 22.53 Monaten in Deutschland ($SD = 27.18$). Von den Teilnehmenden waren $n = 37$ erst zum Zeitpunkt des Studiums nach Deutschland gekommen. Die Rekrutierung der Studierenden erfolgte über zentrale E-Mail-Verteiler und Werbung in sozialen Medien sowie bei Willkommenstagen etc.

Für die Aufnahme eines regulären Studiums an Hochschulen ist im Falle einer ausländischen Hochschulzugangsberechtigung der Nachweis von Sprachfertigkeiten auf dem Niveau C1 im Sinne des gemeinsamen europäischen Referenzrahmens (GER; Europarat, 2001) notwendig. Bei C1 handelt es sich um eine weit fortgeschrittene Sprachkompetenz. Betreffende Personen sollten in der Lage sein, lange, komplexe Texte zu verstehen, sich spontan und fließend auszudrücken, ohne nach Worten suchen zu müssen und sich klar und strukturiert zu komplexen Fragestellungen äußern zu können. Zum sprachlichen

Tabelle 2. Deskriptive Statistiken der Stichprobe

	N	%		N	%
Altersgruppe			Muttersprachen		
16 – 20 Jahre	99	34	Arabisch	44	15
21 – 25 Jahre	124	43	Russisch	43	15
26 – 30 Jahre	46	16	Spanisch	26	9
31 – 35 Jahre	11	4	Bulgarisch	15	5
Studienfächer					
Medizin	32	11			
Informatik	30	10			
Wirtschaftswissenschaften	22	8			
Kommunikationswissenschaften	22	8			

Anmerkungen: Gesamt $N = 290$. Häufigste Studienfächer und Muttersprachen in der Stichprobe.

Hochschulzugang müssen ausländische Studierende einen Sprachtest erbringen, beispielsweise die Deutsche Sprachprüfung für den Hochschulzugang (DSH) oder den TestDaF. Diese und andere Testverfahren fokussieren ein Schwellenniveau rund um das Niveau C1. Die teilnehmenden Personen hatten dieser Vorgaben entsprechend hinreichend sprachliche Kompetenzen zur Beantwortung des eingesetzten Untersuchungsinstrumentes. Der Erwerbsbeginn der deutschen Sprache lag in der untersuchten Stichprobe bei $M = 10.65$ Jahren ($SD = 9.24$) und die Anzahl an Jahren formellen Deutschunterrichts bei $M = 3.82$ ($SD = 3.44$).

Durchführung der Studie und erhobene Instrumente

Die Teilnehmenden des SpraStu-Projekts absolvierten insgesamt drei Termine von je etwa zwei Stunden, in denen sie sechs verschiedene Sprachtests (rezeptiver und produktiver Wortschatz, Hör- und Leseverständnis, TestDaF-Schreibaufgabe und Einstufung der Sprachkompetenzen mit dem onSET) und zwei Fragebögen bearbeiteten (Wisniewski, Möhring, Lenhard & Seeger, 2020). Der hier vorgestellte ScenEx bildete gemeinsam mit dem LIST (Lernstrategien im Studium von Wild & Schiefele, 1994) und einer Selbsteinschätzung zu strategischen Kompetenzen der sprachbezogenen Skalen des Gemeinsamen europäischen Referenzrahmens einen der beiden Fragebögen des Projekts. Im LIST wird danach gefragt, wie oft die Teilnehmenden selbst die vorgeschlagenen Lernstrategien nutzen. Im Gegensatz zur handlungsnahen und leistungsprüfenden Erfassung des ScenEx bildet der LIST also ein subjektives Maß der Selbstauskunft zur Nutzung von Lernstrategien. Der ScenEx wurde in der Regel als erstes Instrument während des zweiten Termins der Pro-

bandinnen und Probanden bearbeitet und nahm rund 20 Minuten in Anspruch. Die Untersuchung fand als parallele, computerbasierte Gruppenuntersuchung zu verschiedenen Terminen in hierfür reservierten, öffentlichen Computerräumen der Universitäten Leipzig und Würzburg statt. Die Teilnehmenden konnten sich frei für die Termine eintragen, die vor Ort von Mitarbeiterinnen und Mitarbeitern sowie Hilfskräften durchgeführt und beaufsichtigt wurden. Für ihre Teilnahme erhielten die Studierenden pro Zeitstunde 10 Euro als Vergütung.

Für die Validierung der Ergebnisse anhand der Sprachfertigkeiten wurde auf den onSET zurückgegriffen, der als ein maßgebliches Instrument in der Attestierung des Sprachlevels für die Hochschulzulassung gilt. Der onSET wurde durch das TestDaF-Institut entwickelt und operationalisiert das Sprachlevel über acht thematisch variierte Lückentexte mit jeweils 20 Leerstellen, wobei in jeder Lücke jeweils nur das Ende eines Wortes ergänzt werden muss. Damit folgt der onSET dem Aufbau eines klassischen C-Tests, für dessen erfolgreiche Bewältigung unterschiedlichste Bereiche der Sprachfähigkeit kombiniert und angewandt werden müssen (Eckes, 2010). Wenngleich für die aktuelle Testversion keine Testgütekriterien vorliegen, konnten frühere Studien von Eckes und Grotjahn (2006) sowie Kaufmann (2016) eine gute Separationsreliabilität von .86 bis .97 zeigen. Folglich gilt der onSET als geeignetes Instrument zur Erfassung der globalen Sprachkompetenz und findet breite Anwendung in Universitäten und Sprachinstituten zur Diagnostik der allgemeinen Sprachfähigkeit (TestDaF-Institut, 2018).

Statistische Auswertung

Die deskriptiven Statistiken sowie Ergebnisse zur internen Konsistenz erfolgten mit Hilfe von R (Version 4.0.2) und

dem Package psych (Version 2.0.9; Revelle, 2020). Das Signifikanzniveau wurde auf $\alpha < .05$ festgelegt. Für die Berechnungen zur Q3-Statistik, zum 1-PL-Modell und zur konfirmatorischen Faktorenanalyse wurden Funktionen der Packages mirt (Chalmers, 2012), TAM (Kiefer, Robitzsch, & Wu, 2020) und eRm (Mair, Hatzinger & Maier, 2020) genutzt. Innerhalb des Fragebogens gab es keine fehlenden Werte, da dieser nur als Ganzes ausgefüllt und nicht weitergeklickt werden konnte, wenn nicht alle Felder bearbeitet waren. Der Umgang mit einzelnen fehlenden Werten innerhalb des Instruments war aus diesem Grund nicht erforderlich. Da für die Erhebung ein Jahr später (T2) jedoch nur noch eine Teilstichprobe zur Verfügung stand, prüften wir einen möglichen systematischen Unterschied zwischen den verbleibenden und den nicht mehr verfügbaren Personen anhand von T-Tests für unabhängige Stichproben. Weder im Ergebnis des LIST zu T1 ($t(290) = .22, p = .828$), im onSET zu T1 ($t(281) = 1.31, p = .192$), noch in der Leistung im ScenEx zu T1 ($t(290) = 1.40, p = .162$) wichen diese Gruppen bedeutend voneinander ab. Es zeigten sich demnach keine Hinweise auf systematischen Drop-Out.

Ergebnisse

Deskriptive Statistiken

Zunächst sind die deskriptiven Statistiken der Instrumente zur Erfassung des metakognitiven Strategiewissens sowie der Sprachkompetenz in Tabelle 3 abgetragen. Es wird eine breite Streuung der Ergebnisse deutlich, wobei die Maximalpunktzahl des ScenEx zum ersten Erhebungszeitpunkt von 14 Personen (5%) und zum zweiten Erhebungszeitpunkt von vier Personen (1%) erreicht wurde. Eingeteilt nach dem GER wurde die Stufe C1 im onSET von $n = 101$ Personen (35%), Stufe B2 von $n = 137$ (47%) und Stufe B1 von $n = 41$ Personen (14%) erreicht. Die restlichen 4% entfallen auf $n = 2$ Personen, die lediglich A2 erreichten und $n = 9$ Personen, für die aufgrund

Tabelle 3. Deskriptive Statistiken der Erhebungsinstrumente

	N	M	SD	Min	Max
ScenEx T1	290	42.67	7.09	8	52
onSET T1	281	113.4	21.26	59	155
ScenEx T2	89	43.36	7.11	14	52
onSET T2	91	120.02	19.93	68	153

Anmerkungen: Max. erreichbare Punktzahl ScenEx: 52; onSET: 160.

verpasster Erhebungstermine kein Ergebnis vorliegt. In Tabelle 4 sind zudem die Interkorrelationen der Instrumente zu Strategiewissen und Sprachkompetenz ersichtlich. Zwischen dem ScenEx und dem onSET zeigen sich durchweg signifikante Korrelationen, während dies für den LIST nicht ersichtlich ist. Der LIST korreliert lediglich innerhalb des Instruments zwischen den Erhebungszeitpunkten und schwach mit dem ScenEx zu T1. Im Gegensatz zum ScenEx zeigt sich hier jedoch kein Zusammenhang mit dem Sprachmaß des onSET.

Homogenität

Die Trennschärfen und Itemschwierigkeiten der Einzelvergleiche über die sechs Szenarien stehen als ESM2 zur Verfügung. Der ScenEx zeigt insgesamt eine gute interne Konsistenz mit einem Cronbachs $\alpha = .87$ ($\Omega = .89$) bei 52 Items, auch die einzelnen Szenarien mit 6 bis 11 Items weisen jeweils ein α von .73 bis .79 und somit für sich allein eine zufriedenstellende interne Konsistenz auf.

Für die Items des ScenEx ergeben sich Itemschwierigkeiten von .58 bis .95, wobei drei hiervon eine Trennschärfe zwischen .1 und .2 zeigen. Alle anderen Trennschärfewerte fallen höher aus. Die Itemschwierigkeiten lassen auf ein Verfahren schließen, das besonders für einen niedrigen bis mittleren Fähigkeitsbereich des metakognitiven Strategiewissens zuverlässig differenziert. Hypothese 1 der hohen internen Konsistenz und guten Homogenitätskennwerten kann somit angenommen werden.

Tabelle 4. Interkorrelationen der Erhebungsinstrumente

	1 ScenEx T1	2 LIST T1	3 onSET T1	4 ScenEx T2	5 LIST T2	6 onSET T2
1		.127* (288)	.371** (281)	.668** (89)	-.054 (89)	.334** (91)
2			-.076 (281)	-.131 (89)	.588** (89)	-.194 (91)
3				.263* (88)	-.139 (88)	.860** (90)
4					-.102 (89)	.348** (88)
5						-.021 (88)

Anmerkungen: N in Klammern. ** = $p < .01$; * = $p < .05$.

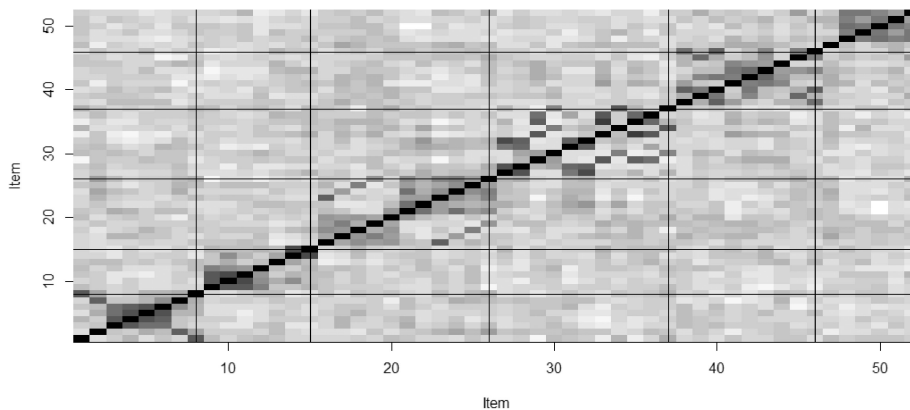


Abbildung 1. Q3-Statistik der Einzeltens, Szenarien durch Linien getrennt.

Skalierung

Die Skalierung von ScenEx erfolgte auf der Basis eines einparametrischen, logistischen Modells (1-PL; Rasch, 1960). Weighted-Likelihood-Estimates (WLE) dienen zur Schätzung der Personenparameter (Warm, 1989; ESM3) Analog zu Cronbachs α weist auch die WLE Reliabilität mit .78 einen befriedigenden bis guten Wert auf. Die Kennwerte des Outfits liegen zwischen .59 und 1.24 ($M = 0.94$, $SD = 0.16$), die Kennwerte des Infits liegen für alle Items im gut differenzierenden Bereich zwischen .89 und 1.17 ($M = 1.00$, $SD = 0.06$).

Mittels des Anderson-Likelihood-Ratio-Tests wurde die spezifische Objektivität des Verfahrens geprüft. Hierzu wurde die Stichprobe am arithmetischen Mittel gesplittet. Es ergab sich keine signifikante Verletzung der spezifischen Objektivität. Die EAP-Reliabilität weist einen guten Wert von .84 auf. Eine spezifisch objektive Messung der latenten Variable durch die Items kann somit angenommen werden (Hypothese 2). Ein Waldtest auf Itemebene mit Teilung der Stichprobe anhand der Geschlechtsvariable zeigte für 46 der 52 Items keine signifikanten Unterschiede. Bei sechs Items wurde das Ergebnis signifikant. Die Ergebnisse von zwei dieser Items waren zugunsten der Frauen ausgeprägt, vier zugunsten der Männer. Die sechs auffälligen Items im Waldtest sind bezüglich der Outfit- sowie Infit-Kennwerte unauffällig.

Überprüfung der lokalen stochastischen Unabhängigkeit

Da sich in ScenEx die Lösungsalternativen auf bestimmte Szenarien beziehen und zusätzlich in den Paarvergleichen eine einzelne Lösungsstrategie u. U. mehrfach berücksichtigt wird, liegt die Verletzung der lokalen stochastischen Unabhängigkeit innerhalb der verschiedenen Szenarien nahe. Zur Prüfung der Abhängigkeit von Items innerhalb eines Szenarios bei gleichzeitiger Unabhängig-

keit von Items anderer Szenarien wurde auf die Q3-Statistik (Yen, 1993) zurückgegriffen. Abbildung 1 bildet diese Statistik grafisch ab. Die Items werden hier ebenfalls paarweise verglichen. Die Grenze eines Szenarios zum nächsten wird hier jeweils durch die horizontalen und vertikalen Linien verdeutlicht. Dunklere Regionen im Plot zeigen Abhängigkeiten an und im vorliegenden Beispiel sind diese insbesondere innerhalb der einzelnen Szenarien sichtbar. Erwartungskonform zeigen die Items lokale Abhängigkeiten innerhalb der Szenarien, während sich keine Hinweise auf Abhängigkeiten zu Items anderer Szenarien ergeben (Hypothese 3).

Konfirmatorische Faktorenanalyse

In einer konfirmatorischen Faktorenanalyse wurde im nächsten Schritt auf der Basis der Ergebnisse der sechs Szenarien ein Messmodell auf einen latenten Faktor gebildet und geprüft, ob die Szenarienscores eine gemeinsame latente Fähigkeit zum sprachbezogenen metakognitiven Wissen formen. Die Fit-Indizes lassen hierbei auf einen nahezu perfekten Modell Fit schließen mit CFI = 1.0, RMSEA = 0.0, SRMR = .02 und AIC = 7 463.51. Die Daten weisen folglich eine erwartungskonform gute Passung zur theoretischen Modellvorstellung auf (Hypothese 4). Das Konstrukt des sprachbezogenen metakognitiven Strategiewissens scheint durch die Szenarien abgebildet zu werden und so der Beantwortung des ScenEx als gemeinsame latente Fähigkeit zugrunde zu liegen.

Kriteriumsvalidität

In einem finalen Schritt wurde die Kriteriumsvalidität anhand des Zusammenhangs und der Vorhersage von Sprachkompetenz durch das metakognitive Strategiewissen geprüft. Die Korrelation der Leistung im ScenEx mit der des onSETs desselben Erhebungszeitpunkts und da-

mit der sprachlichen Kompetenz der ausländischen Studierenden zeigt ein signifikantes Ergebnis ($r(281) = .371$, $p < .01$; Hypothese 5a). Selbiges gilt für den Zusammenhang des ScenEx mit dem onSET ein Jahr später, also nach dem ersten Studienjahr, ($r(91) = .334$, $p < .01$), was einer Varianzaufklärung von 11 % entspricht.

Eine lineare Regression auf das onSET-Ergebnis ein Jahr später zeigt eine prädiktive Relevanz des ScenEx-Ergebnisses über das onSET-Ergebnis zu Erhebungsstart hinaus $b = .11$, $t(89) = 1.89$, $p = .031$. Die Gesamtaufklärung der beiden Prädiktoren an der Varianz des onSET-Ergebnisses ein Jahr später beträgt 75 %. Das durch den ScenEx erfasste metakognitive Wissen klärte inkrementell 4 % der Varianz über den OnSET zu T1 hinaus auf und war prädiktiv für den Zuwachs an Sprachfertigkeiten (Hypothese 5b).

Diskussion

ScenEx ist ein Verfahren zur Erfassung des metakognitiven Strategiewissens in sprachlich schwierigen Situationen für Studierende mit Deutsch als Fremdsprache. Für diese Personengruppe existierten bislang keine spezifischen Verfahren zur Lernregulation im Studium, die den besonderen Bedingungen sprachlich anspruchsvoller akademischer Settings Rechnung tragen (Bärenfänger et al., 2016; Wisniewski et al., 2019). Anders als bei Verfahren zur Selbsteinschätzung strategischen Verhaltens versucht der ScenEx über die Bewertung der konkreten Handlungsalternativen eine direktere Messung des metakognitiven Strategiewissens zu realisieren. Ziel der vorliegenden Studie war die Überprüfung der Testeigenschaften anhand einer größeren Stichprobe hinsichtlich der internen Struktur und Validität des Verfahrens. Die Ergebnisse zeigen eine insgesamt gute Testgüte und weisen damit auf die Eignung des ScenEx zur Erfassung des zugrundeliegenden Konstrukts des metakognitiven Strategiewissens hin.

Das Verfahren weist, wie in Hypothese 1 postuliert, sowohl im Hinblick auf die einzelnen Szenarien als auch im Gesamtverfahren eine sehr gute interne Konsistenz auf. Dies ist umso bemerkenswerter, da Cronbachs α mit der Anzahl an Aufgaben ansteigt und die guten Kennwerte trotz niedriger Itemanzahlen in den einzelnen Szenarien erzielt werden konnten. Somit zeigte sich die interne Konsistenz als erster untersuchter Indikator für die Testgüte des ScenEx erwartungskonform günstig. Die darauf folgende Skalierung zeigte in gleicher Weise gute Reliabilitätskennwerte auf Skalenebene. Mit dem nicht signifikanten Ergebnis des Anderson-Likelihood-Tests zeigt sich eine gleichermaßen gute Passung der Skala für Per-

sonen mit niedriger und hoher Merkmalsausprägung des latenten Konstrukts. Die Annahme des Vorliegens spezifischer Objektivität (Hypothese 2) wird somit nicht zurückgewiesen. Auch liegen die Fit-Indizes der Einzelvergleiche in einem guten Bereich. In einem Waldtest mit dem Geschlecht als Split-Kriterium zeigte sich darauffolgend eine kleine Anzahl auffälliger Aufgaben zugunsten der männlichen Versuchspersonen, die jedoch angesichts der Gesamtzahl an Aufgaben als nicht maßgeblich eingestuft werden kann. Die aufgeführten Resultate lassen auf Items schließen, die homogen die latente Variable des sprachbezogenen metakognitiven Strategiewissens erfassen und den Gesamtscore des ScenEx entsprechend interpretierbar machen.

Die dabei durch die Szenarien gegebene Testlet-Struktur des Verfahrens untersuchten wir mithilfe der Q3-Statistik. Es ergaben sich Abhängigkeiten der Antwortmuster innerhalb, nicht jedoch zwischen den Szenarien, die vermutlich durch die mehrfache Berücksichtigung der einzelnen Antwortalternativen bei den verschiedenen Einzelvergleichen entstanden. Unsere theoretischen Vorannahmen zielten des Weiteren darauf, dass das metakognitive Strategiewissen einen gesamten latenten Faktor bildet. Erwartungskonform fällt die Modellierung einer Einfaktorenlösung mittels konfirmatorischer Faktorenanalysen aus: Die Kennwerte der CFA der Szenarien auf den Gesamtscore zeigen einen nahezu perfekten Modelfit an, womit sich Hypothese 4 der Grobstruktur des Tests bestätigt. Die sechs Szenarien bilden dementsprechend zusammen die latente Variable ab, die durch den Gesamtscore wiedergegeben werden kann (Hypothese 4).

Ein Grund für die nur schwache Korrelation des ScenEx mit dem LIST zu T1 und die fehlende Korrelation zu T2 könnte in den unterschiedlichen Erhebungsmethoden liegen. Der ScenEx stellt einen Leistungstest des strategischen Wissens in schwierigen Studienszenarien dar, wobei im Abgleich mit der Blaupause der Expertinnen und Experten Punkte vergeben werden. Der Test hat also eher den Charakter eines Leistungstests. Die Angaben im LIST bilden eine subjektive Selbsteinschätzung der Häufigkeit der Strategienutzung ab. Die Daten aus solchen Selbstauskünften können verzerrt sein, beispielsweise durch Erinnerungseffekte, selbstwertdienliche Selbsteinschätzungen, Bezugsgruppeneffekte, Tendenz zur Mitte, und soziale Erwünschtheit (Pintrich, 2004). Zudem erfasst der LIST nicht das Verhalten in spezifischen Situationen, sondern fragt retrospektiv nach der Strategienutzung. Im Gegensatz zu herkömmlichen Verfahren wie dem LIST korrelieren die Werte des ScenEx einerseits querschnittlich mit den Sprachtestergebnissen, sagten aber auch Sprachtestergebnisse zu T2 unter Kontrolle der sprachlichen Ausgangsleistung zu T1 vorher. Die ScenEx-Ergebnisse erwiesen sich als prädiktiv für den Zuwachs an

sprachlichen Kompetenzen. Dieser Validitätsbeleg stellt einen klaren Vorteil zu herkömmlichen Verfahren der Strategieerfassung mittels Selbsteinschätzung dar und zeigt die Relevanz einer handlungsnahen Erfassung strategischer Kompetenz. Da metakognitives Wissen gut vermittelbar ist, eröffnet sich hierdurch auch die Möglichkeit für eine Förderung über eine reine Sprachkompetenzvermittlung hinaus.

Limitationen

Eine Einschränkung der Aussagekraft des Verfahrens liegt in einer reduzierten Stichprobengröße zum Zeitpunkt T2. Die Regressionsanalyse zur Vorhersage der Sprachleistung ein Jahr nach Erhebung des ScenEx basierte auf einer kleineren Stichprobe. Dies ist hauptsächlich auf das Kohortensequenzdesign der Ausgangsuntersuchung zurückzuführen, da viele Studierende sich in einer Kohorte befanden, die T2 noch nicht durchlaufen hatte. Es wäre jedoch auch denkbar, dass Studierende mit niedrigeren Sprachleistungen während des ersten Studienjahres tendenziell eher abbrechen. Es kann folglich in Teilen eine Selbstselektion der Studierenden nicht ausgeschlossen werden. Wir begegneten diesem Umstand mit einer Untersuchung eines möglichen Unterschieds dieser beiden Gruppen im Ergebnis des ScenEx und auch der Sprachleistung zu T1, welcher in der Folge ausgeschlossen werden konnte. Die Gefahr, dass es sich beim vorliegenden Befund um ein Artefakt handeln könnte, ist aus den genannten Gründen somit gering.

Wie bei vielen anderen Instrumenten zur Erfassung des metakognitiven Strategiewissens, kann auch im Fall des ScenEx der sog. Common Method Bias nicht ausgeschlossen werden. Untersucht man beispielsweise mittels eines Fragebogens, ob Strategiewissen zum Leseverständnis beiträgt, so liegt eine Konfundierung zwischen der Erfassungsmethode, dem zu erfassenden Konstrukt und der abhängigen Variable vor. Ähnlich verhält es sich auch bei ScenEx. Je schwächer die Sprachleistungen sind, desto schwieriger dürfte es Personen fallen, die Szenarien angemessen zu bewerten. Um dem entgegenzuwirken wurden alle Formulierungen des Instruments von Mitarbeitenden der Fremdsprachendidaktik dahingehend geprüft, ob für eine Person mit einer Hochschulzugangsberechtigung in Deutschland Verständnisprobleme durch linguistische Komplexität der Formulierungen zu erwarten sind. Die verwendete Sprache wurde dementsprechend so vereinfacht, dass Verständnisschwierigkeiten nicht auftreten sollten. Des Weiteren weisen die von der vorliegenden Stichprobe durchschnittlich erreichten Werte im onSET, sowie deren überwiegende Einstufung in die Niveaustufen

C1 und B2 des GER auf eine angemessene Sprachkompetenz der Teilnehmenden hin. Wir gehen davon aus, dass die Reduktion der sprachlichen Komplexität des Instruments mögliche Konfundierungen zumindest reduziert hat, auch wenn schlussendlich Überlappungen der Fähigkeiten nicht völlig ausgeräumt werden können.

Fazit

Insgesamt erwies sich der ScenEx als ein geeignetes Verfahren zur Erfassung studien- bzw. sprachbezogenen metakognitiven Wissens von Studierenden mit nicht deutscher Muttersprache. Sein Einsatz in der Hochschulforschung könnte dazu beitragen, die Bedingung für das Gelingen eines Studiums zielgenauer zu erforschen. Auch auf individueller Ebene ist der Einsatz neben anderen Kriterien der Hochschulzulassung interessant, um mittels spezifischer Beratungsangebote Studierende bei der Aufnahme eines sprachlich herausfordernden Studiums besser zu unterstützen. Über das zentrale Kriterium der Sprachfähigkeit hinaus eröffnet sich somit ein neuer Ansatzpunkt für eine niedrigschwellige und zielgerichtete Förderung des Studienerfolgs ausländischer Studierender. Das Instrument könnte in der Studienvorbereitungsphase dabei helfen, Studierende nicht nur hinsichtlich ihrer Sprachfertigkeiten auf ein Studium vorzubereiten, sondern gezielt bei der Vermittlung von Strategiewissen zu helfen, um kritische Situationen im Studium besser zu meistern und einem Drop-out aus dem Studium protektiv zu begegnen.

Elektronische Supplemente (ESM)

Die elektronischen Supplemente sind mit der Online-Version dieses Artikels verfügbar unter <https://doi.org/10.1026/0012-1924/a000275>

ESM 1. ScenEx im Original.

ESM 2. Tabelle der deskriptiven Statistiken.

ESM 3. Tabelle der Personenparameter.

Literatur

- Artelt, C. (1999). Lernstrategien und Lernerfolg – eine handlungsnaher Studie. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 31, 86–96. <https://doi.org/10.1026/0049-8637.31.2.86>
- Artelt, C. (2000). Wie prädiktiv sind retrospektive Selbstberichte über den Gebrauch von Lernstrategien für strategisches Ler-

- nen? *Zeitschrift für Pädagogische Psychologie*, 14(2/3), 72–84. <https://doi.org/10.1024/1010-0652.14.23.72>
- Bärenfänger, O., Lange, D. & Möhring, J. (2016). *Sprache und Bildungserfolg: sprachliche Anforderungen in der Studieneingangsphase*. Leipzig: Institut für Testforschung und Testentwicklung.
- Boerner, S., Seeber, G., Keller, H. & Beinborn, P. (2005). Lernstrategien und Lernerfolg im Studium. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 37, 17–26. <https://doi.org/10.1026/0049-8637.37.1.17>
- Chalmers, R. P. (2012). mirt: A Multidimensional Item Response Theory Package for the R Environment. *Journal of Statistical Software*, 48(6), 1–29. <https://doi.org/10.18637/jss.v048.i06>
- Dörnyei, Z. & Ryan, S. (2015). *The psychology of the language learner revisited*. London: Routledge.
- Eckes, T. (2010). Der Online-Einstufungstest Deutsch als Fremdsprache (onDaF): Theoretische Grundlagen, Konstruktion und Validierung. In R. Grotjahn (Hrsg.), *Der C-Test: Beiträge aus der aktuellen Forschung* (S. 125–192). Berlin: Peter Lang.
- Eckes, T. & Grotjahn, R. (2006). A closer look at the construct validity of C-tests. *Language Testing*, 23, 290–325. <https://doi.org/10.1191%2F0265532206lt330oa>
- Europarat (2001). *Gemeinsamer europäischer Referenzrahmen für Sprachen: Lernen, lehren, beurteilen*. München: Langenscheidt.
- Geiser, C. & Eid, M. (2010). Item-Response-Theorie. In C. Wolf & H. Best (Hrsg.), *Handbuch der sozialwissenschaftlichen Datenanalyse* (S. 311–332). Wiesbaden: Springer.
- Hasselhorn, M. & Labuhn, A. (2008). Metakognition und selbstreguliertes Lernen. In W. Schneider & M. Hasselhorn (Hrsg.), *Handbuch der Pädagogischen Psychologie* (S. 28–37). Göttingen: Hogrefe.
- Heublein, U. (2015). Von den Schwierigkeiten des Ankommens: Überlegungen zur Studiensituation ausländischer Studierender an den deutschen Hochschulen. *Die Neue Hochschule*, 1, 14–17.
- Kaufmann, N. (2016). Die Vorhersage der Schwierigkeit deutscher C-Test-Texte: Untersuchungen am Beispiel des onDaF. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 21(2), 111–126.
- Kiefer, T., Robitzsch, A. & Wu, M. (2020). *TAM: Test Analyses Modules* [Computer software]. Verfügbar unter: <https://cran.r-project.org/web/packages/TAM/index.html>
- Leutner, D. & Leopold, C. (2003). Selbstreguliertes Lernen als Selbstregulation von Lernstrategien – Ein Trainingsexperiment mit Berufstätigen zum Lernen aus Sachtexten. *Unterrichtswissenschaft*, 31(1), 38–56.
- Lingel, K. (2016). *Metakognitives Wissen Mathematik: Entwicklung und Zusammenhang mit der Mathematikleistung in der Sekundarstufe I*. Dissertation, Julius-Maximilians-Universität Würzburg.
- Lingel, K., Götz, L., Artelt, C. & Schneider, W. (2014). *MAESTRA 5–6 + Mathematisches Strategiewissen für fünfte und sechste Klassen*. Göttingen: Hogrefe.
- Mair, P., Hatzinger, R. & Maier, M. J. (2020). *Extended Rasch Modeling: The eRm Package for the Application of IRT Models in R* [Computer software]. Verfügbar unter: <https://cran.r-project.org/package=eRm>
- Moosbrugger, H. (2012). Item-Response-Theorie (IRT). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 215–259). Heidelberg: Springer.
- Neuenhaus, N., Artelt, C. & Schneider, W. (2017). Lernstrategiewissen im Bereich Englisch. *Diagnostica*, 63, 135–147. <https://doi.org/10.1026/0012-1924/a000171>
- Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. Boston, MA: Heinle & Heinle.
- Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review*, 16, 385–407. <https://doi.org/10.1007/s10648-004-0006-x>
- Pintrich, P. R., Smith, D., Garcia, T. & McKeachie, W. J. (1993). Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychological Measurement*, 53, 801–813. <https://doi.org/10.1177%2F0013164493053003024>
- Rammstedt, B. (2010). Objektivität, Reliabilität, Validität. In C. Wolf & H. Best (Hrsg.), *Handbuch der sozialwissenschaftlichen Datenanalyse* (S. 239–258). Wiesbaden: Springer.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Kopenhagen: Danish Institute for Educational Research.
- Revelle, W. (2020). *psych: Procedures for Personality and Psychological Research* [Computer software]. Verfügbar unter: <https://CRAN.R-project.org/package=psych>
- Rubin, J. (1975). What the "Good Language Learner" Can Teach Us. *TESOL Quarterly*, 9(1), 41–51. <https://doi.org/10.2307/3586011>
- Rutsch, J., Vogel, M., Rehm, M. & Dörfler, T. (2018). Modellierung der Testletstruktur bei vignettenbasierten Testverfahren mit geschlossenem Antwortformat. In J. Rutsch, M. Vogel, M. Rehm, M. Seidenfuß & T. Dörfler (Hrsg.), *Effektive Kompetenzdiagnose in der Lehrerbildung: Professionalisierungsprozesse angehender Lehrkräfte untersuchen* (S. 26–46). Berlin: Springer.
- Schiefele, U., Streblow, L. & Brinkmann, J. (2007). Aussteigen oder Durchhalten. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 39, 127–140. <https://doi.org/10.1026/0049-8637.39.3.127>
- Schlagmüller, M. & Schneider, W. (2007). *Würzburger Lesestrategie-Wissenstest für die Klassen 7–12: Ein Verfahren zur Erfassung metakognitiver Kompetenzen bei der Verarbeitung von Texten*. Göttingen: Hogrefe.
- Schneider, W. & Artelt, C. (2010). Metacognition and mathematics education. *ZDM Mathematics Education*, 42, 149–161. <https://doi.org/10.1007/s11858-010-0240-2>
- Stemmer, P. (2013). *Studien- und Lebenssituation ausländischer Studierender an deutschen Hochschulen: Analyse – Handlungsfelder – strategische Entscheidungsmöglichkeiten*. Baden-Baden: Nomos. <https://doi.org/10.5771/9783845246819>
- Strobl, C. (2015). *Das Rasch-Modell: Eine verständliche Einführung für Studium und Praxis*. München: Rainer Hampp Verlag.
- TestDaF-Institut (2018). *onSET-Handbuch: Planung und Durchführung von Online-Spracheinstufungstests – onSET-Deutsch, onSET-English*. Bochum: TestDaF-Institut.
- Waldeyer, J., Fleischer, J., Wirth, J. & Leutner, D. (2017). Selbstreguliertes Lernen – Ressourcenmanagement im Studium. In C. Maurer (Hrsg.), *Implementation fachdidaktischer Innovation im Spiegel von Forschung und Praxis* (S. 63–66). Regensburg: Universität Regensburg.
- Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, 54, 427–450. <https://doi.org/10.1007/BF02294627>
- Wild, K.-P. & Schiefele, U. (1994). Lernstrategien im Studium: Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 15, 185–200.
- Wisniewski, K. (2018). Sprache und Studienerfolg von Bildungsausländerinnen und -ausländern: Eine Längsschnittstudie an den Universitäten Leipzig und Würzburg. *Informationen Deutsch als Fremdsprache*, 45, 573–597. <https://doi.org/10.1515/infodaf-2018-0074>
- Wisniewski, K., Möhring, J., Lenhard, W. & Seeger, J. (2020). Sprachkompetenzen und Studienerfolg von Bildungsausländer/-innen zu Studienbeginn: Erste Erkenntnisse eines empiri-

schen Längsschnittprojekts. In A. Drackert, M. Mainzer-Murrenhoff, A. Soltyska & A. Timukova (Hrsg.), *Testen bildungssprachlicher Kompetenzen und akademischer Sprachkompetenzen – Synergien zwischen Schule und Hochschule erkennen und nutzen*. Frankfurt a. M.: Language Testing and Assessment.

Wisniewski, K., Parker, M., Lenhard, W. & Seeger, J. (2019). Sprachbezogenes metakognitives Strategiewissen im Studienalltag internationaler Studierender: Der szenariobasierte Fragebogen ScenEx. *Zeitschrift für Fremdsprachenforschung*, 30(1), 57–78.

Yen, W. M. (1993). Scaling performance assessments: Strategies for managing local item dependence. *Journal of educational measurement*, 30, 187–213. <https://doi.org/10.1111/j.1745-3984.1993.tb00423.x>

Historie

Onlineveröffentlichung: 29.07.2021

Danksagung

Wir danken Maria Parker für ihre Beiträge zum ScenEx und für ihre Unterstützung bei der Erstellung des Manuskripts.

Förderung

Das Projekt "Sprache und Studienerfolg bei Bildungsausländer/-innen" wird gefördert durch das Bundesministerium für Bildung und Forschung, Projektlinie „Studienerfolg und Studienabbruch“ (Förderkennzeichen 01PX16020B).

Open Access-Veröffentlichung ermöglicht durch die Julius-Maximilians-Universität Würzburg.

ORCID

Jennifer Seeger

 <https://orcid.org/0000-0002-7624-4453>

M.Sc. Jennifer Seeger

Lehrstuhl für Psychologie IV

Institut für Psychologie

Julius-Maximilians-Universität Würzburg

Wittelsbacherplatz 1

97074 Würzburg

jennifer.seeger@uni-wuerzburg.de