

**bwp@ Österreich Spezial 3 | September 2021**

**Beiträge zum**

**14. Österreichischen Wirtschaftspädagogik-Kongress  
am 9. April 2021 in Graz**

Hrsg. v. **Peter Slepcevic-Zach & Susanne Kamsker**

**Shireen SARWARI**

(Wirtschaftsuniversität Wien)

**Wirtschaftswissen junger Erwachsener in Österreich  
Entwicklung eines Testinstruments zu Themen in den  
Wirtschaftsberichten von Tageszeitungen**

Online unter:

[http://www.bwpat.de/wipaed-at3/sarwari\\_wipaed-at\\_2021.pdf](http://www.bwpat.de/wipaed-at3/sarwari_wipaed-at_2021.pdf)

www.bwpat.de | ISSN 1618-8543 | **bwp@** 2001–2021

**Open Access publiziert mit Unterstützung der Universität Graz**

**bwp@**

**www.bwpat.de**



Herausgeber von **bwp@** : Karin Büchter, Franz Gramlinger, H.-Hugo Kremer, Nicole Naeve-Stoß, Karl Wilbers & Lars Windelband

**Berufs- und Wirtschaftspädagogik - online**

## **Wirtschaftswissen junger Erwachsener in Österreich. Entwicklung eines Testinstruments zu Themen in den Wirtschaftsberichten von Tageszeitungen**

---

### **Abstract**

Die letzte empirische Untersuchung zur Messung des Wirtschaftswissens von österreichischen Maturantinnen und Maturanten liegt bereits mehr als ein Jahrzehnt zurück und beschränkte sich auf den allgemeinbildenden gymnasialen Schultyp (vgl. Brandlmaier et al. 2006). Derzeit fehlt es an aktuellen Befunden zum vorhandenen Wirtschaftswissen von jungen Erwachsenen. Vor diesem Hintergrund besteht die Zielsetzung des vorliegenden Forschungsvorhabens darin, ein Testinstrument zur Messung des Wirtschaftswissens junger Erwachsener, die sich in der Übergangsphase zwischen sekundärer und tertiärer Ausbildung befinden, zu entwickeln. Die Inhaltsbereiche der Testfragen wurden anhand einer Medienanalyse zu den Themen in den Wirtschaftsberichten österreichischer Tageszeitungen bestimmt. Das nach den Ansprüchen wissenschaftlicher Testtheorie entwickelte Testinstrument wurde einer Experten/innenvalidierung unterzogen und im Rahmen eines ersten Testlaufs (N=388) empirisch erprobt. In diesem Beitrag werden die methodischen Entwicklungsschritte skizziert und ausgewählte Ergebnisse aus dem ersten Testlauf vorgestellt. Die Ergebnisse der Itemanalyse nach klassischer sowie probabilistischer Testtheorie geben Aufschlüsse darüber, ob eine angemessene Passung der Testaufgaben zum Konstrukt Wirtschaftswissen vorliegt. Die Ergebnisse zeigen, dass der Großteil der Items insgesamt eine angemessene psychometrische Eignung aufweist. Die faktorenanalytische Dimensionsanalyse weist auf eine angestrebte eindimensionale Lösung hin.

---

## **Economic knowledge of young adults in Austria Development of an instrument on main topics of economic coverage in daily newspapers**

---

The last empirical study to measure the economic knowledge of students at the upper secondary level has been released over ten years ago (vgl. Brandlmaier et al. 2006). There is a lack of empirical findings on current economic knowledge of young adults. The main aim of the study is to develop an instrument for a standardized assessment of economic knowledge of young adults, who are between completing school leaving examination and just before entering the tertiary educational system in Austria. The content of the items has been determined by a media analysis on main topics of economic coverage in daily newspapers. The development of the instrument is based on test theory. Content validity has been evaluated by experts before the instrument was tested empirically for the first time (N=388). This paper gives an overview of the development of the instrument and selected results of the empirical evaluation of the first version of the instrument according to both classical test theory and item response theory. The results indicate that most of the items show an adequate psychometric quality and the factor analysis indicates a unidimensional solution.

## **1 Problemhintergrund und Zielsetzung**

Die im österreichischen Schuldenreport jüngst veröffentlichten Zahlen sprechen für sich: Rund ein Viertel aller Klientinnen und Klienten der Schuldnerberatung sind unter 30 Jahre alt. „Sie haben also schon in jungen Jahren so viele Schulden angehäuft, dass sie Schwierigkeiten bei der Rückzahlung haben.“ (Schuldnerberatung 2021, 25). Die steigende Zahl junger Erwachsener, die vermehrt in Zahlungsschwierigkeiten geraten, ist ein Indikator dafür, dass deren individuelle wirtschaftliche und finanzielle Herausforderungen einer wachsenden Aufmerksamkeit bedürfen. Um solche Herausforderungen verantwortungsvoll beurteilen und bewältigen zu können, ist es unerlässlich über ein grundlegendes Verständnis für wirtschaftliche Zusammenhänge und Sachverhalte zu verfügen (vgl. Geise 2001, Kaminski/Eggert/Burkhard 2008, Seeber/Retzmann/Remmele 2012).

Am Ende der Sekundarstufe II befinden sich junge Erwachsene an einer Schnittstelle zwischen sekundärer und tertiärer Ausbildung bzw. dem Einstieg in das Berufsleben und sind daher für bildungspolitische Überlegungen besonders relevant. Der Beginn dieses neuen Lebensabschnittes geht mit zunehmender finanzieller Autonomie einher, die auch bei der Schuldenprävention eine wichtige Rolle spielt. Junge Erwachsene sind erstmals mit umfassenden wirtschaftlichen Entscheidungssituationen konfrontiert, wie z. B. die Finanzierung der ersten eigenen Wohnung, die Beantragung eines Stipendiums, die steuerliche Veranlagung oder das erste Arbeitsverhältnis. Als wahlberechtigte Bürgerinnen und Bürger ist zudem ein Grundverständnis für das politische und wirtschaftliche Tagesgeschehen von Relevanz (vgl. Dubs 2014).

Aus bisherigen empirischen Untersuchungen kann abgeleitet werden, dass insbesondere das familiäre Umfeld aber auch Peers eine entscheidende Rolle beim wirtschaftlichen bzw. finanziellen Wissenserwerb junger Erwachsener spielen. Die PISA-Erhebung liefert dazu eine interessante Befundlage: Jene Befragte, die sich regelmäßig mit einem Elternteil über Geldfragen austauschen, weisen ein deutlich höheres Finanzwissen auf (vgl. OECD 2017). Problematisch ist allerdings, dass viele Erwachsene zum Teil erhebliche Lücken im Wirtschafts- bzw. Finanzwissen aufweisen (vgl. z. B. Silgoner/Greimel-Fuhrmann/Weber 2015) und somit selbst nicht in der Lage sind, sich mit ihren Kindern über wirtschaftliche Fragen kundig auszutauschen (vgl. Greimel-Fuhrmann 2018). Es ist daher plausibel anzunehmen, dass das im Rahmen der schulischen Bildungskarriere angeeignete Wirtschaftswissen maßgeblichen Einfluss auf die Fähigkeit hat, wirtschaftlich und finanziell verantwortungsvolle Entscheidungen zu treffen. Es besteht weitgehend Konsens darüber, dass die Grundlage für das Verstehen wirtschaftlicher Zusammenhänge und Sachverhalte in der Schulbildung als Teil der Allgemeinbildung verankert werden muss, damit eine möglichst breite Schicht der Bevölkerung erreicht werden kann (vgl. Greimel-Fuhrmann 2013, Kaminski/Eggert/Burkhard 2008, Dubs 2014).

Um eine Aussage über den vorliegenden Stand des Wirtschaftswissens am Ende eines definierten Bildungsabschnittes treffen zu können, werden in der empirischen Sozialforschung traditionell wissenschaftlich fundierte Messinstrumente eingesetzt. Seitens der österreichischen Bildungspolitik liegt jedoch eine systematische und einheitliche Lernstandserhebung des Wirtschaftswissens derzeit außer Reichweite. Bislang liegen in Österreich zudem nur wenig aktuelle Befunde zum tatsächlich vorhandenen Wirtschaftswissen von jungen Erwachsenen am Ende der Sekundarstufe II vor. Die letzte empirische Untersuchung bei österreichischen Maturantinnen und Maturanten liegt bereits mehr als ein Jahrzehnt zurück und beschränkte sich auf den Allgemeinbildenden gymnasialen Schultyp (vgl. Brandlmaier et al. 2006).

Ziel des vorliegenden Beitrags ist, einen ersten Einblick über die Entwicklung und Erprobung eines Testinstruments zur Messung des Wirtschaftswissens von jungen Erwachsenen zu geben. Zunächst wird ein kurzer Einblick über die Inhaltsbestimmung der Testaufgaben gegeben sowie die Vorgehensweise bei der Entwicklung der Items skizziert. Im letzten Abschnitt werden ausgewählte empirische Ergebnisse aus dem ersten Testlauf (Pretest) erläutert sowie die wichtigsten Ergebnisse der Itemrevison zusammengefasst.

## **1.1 Projektzielsetzung und Zielgruppe**

Die Zielsetzung des vorliegenden Forschungsprojekts umfasst in erster Linie die Entwicklung und empirische Erprobung eines wissenschaftlich fundierten und standardisierten Testinstruments, um das vorhandene Wirtschaftswissen junger Erwachsener zu messen. Durch den erstmaligen Einsatz des Testinstruments bei einer kleineren Stichprobe soll die psychometrische Eignung der Testitems anhand statistischer Kennwerte systematisch überprüft werden. Im zweiten Testlauf soll das Erhebungsinstrument österreichweit eingesetzt werden mit der erweiterten Zielsetzung erste Erkenntnisse zum vorliegenden Stand des Wirtschaftswissens junger Erwachsener zu gewinnen. Die Zielgruppe des Testinstruments umfasst junge Erwachsene, die die Sekundarstufe II in Österreich absolviert haben und sich zum Testzeitpunkt zwischen der Bewerbungs- und Aufnahmephase für ein Studium an einer österreichischen Universität, Fachhochschule oder Pädagogischen Hochschule befinden. Nach österreichischem Bildungssystem schließt die Zielgruppe insbesondere Absolventinnen und Absolventen Allgemeinbildender höherer Schulen (AHS) sowie Berufsbildender höherer Schulen (BHS) mit ein. Darüber hinaus umfasst die Zielgruppe junge Erwachsene, die eine Berufsreife- oder Studienberechtigungsprüfung absolviert haben, welche grundsätzlich zur Aufnahme in ein Studium an einer österreichischen Hochschule berechtigt.

Der Zugang zur Zielgruppe wird dadurch begründet, dass sich der Großteil der österreichischen Maturantinnen und Maturanten (rund 72 %) innerhalb von drei Jahren nach Absolvierung der Matura für ein Studium an einer österreichischen Hochschule entscheidet, wobei die Übertrittsrate bei Absolventinnen und Absolventen der Allgemeinbildenden höheren Schulen mit rund 89% besonders hoch ist. Bei den kaufmännischen Berufsbildenden höheren Schulen liegt die Übertrittsrate innerhalb der ersten drei Jahre nach der Matura bei rund 63% (vgl. Statistik Austria 2020). In der folgenden Abbildung werden die kumulierten Übertrittsdaten von der Matura ins Hochschulsystem nach Schultyp illustriert:

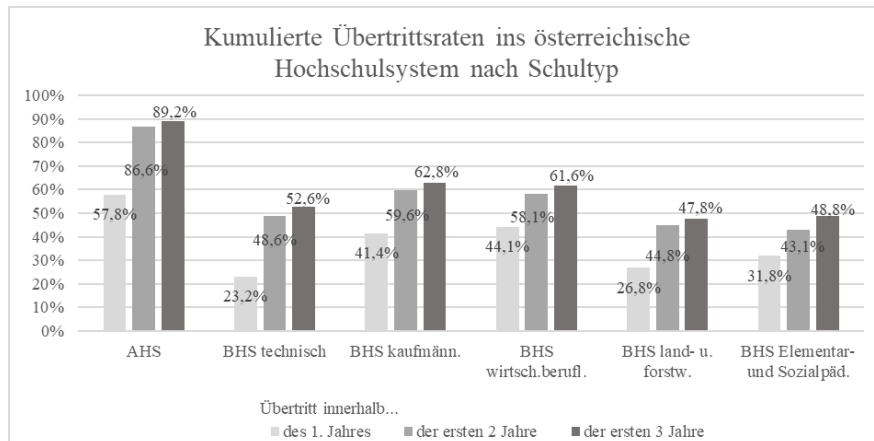


Abbildung 1: Übertrittsraten Hochschulsystem, Abbildung in Anlehnung an Statistik Austria (2020), Zahlen entsprechen der Maturajahrgänge 2012/13 bis 2017/18, für die bis zum Studienjahr 2018/19 Inskriptionsdaten vorlagen.

Eine Untersuchung in der Phase zwischen Studienbewerbung und Studienbeginn ist aufgrund organisatorischer Überlegungen besser geeignet als eine Erhebung direkt an den Schulen. Im letzten Schuljahr liegt das Hauptaugenmerk auf die Vorbereitung der Schülerinnen und Schüler auf die Matura, wobei die Unterrichtszeit aufgrund des verkürzten Schuljahres besonders knapp ist. Aus diesem Grund ist die Bereitschaft der Schulen, an einer Studie teilzunehmen, erfahrungsgemäß als eher gering einzuschätzen. Zudem kann davon ausgegangen werden, dass das Wissensniveau hinsichtlich wirtschaftlicher Inhalte in der Phase zwischen Studienbewerbung und Studienbeginn weitgehend jenem bei Absolvierung der Matura entspricht.

## 1.2 Testinstrumente zur Messung von Wirtschaftswissen

Das fachspezifische Wissen stellt eine unverzichtbare Voraussetzung für die erfolgreiche Bewältigung von Entscheidungssituationen eines bestimmten Bereiches dar (vgl. Weinert 1999). Die meisten empirischen Erhebungen im Forschungsbereich fokussieren sich aus diesem Grund auf die kognitiven Facetten (Wissen, Fähigkeiten, Fertigkeiten) der Wirtschaftskompetenz. Dies hat nicht zuletzt den Vorteil, dass durch eine Eingrenzung des zu messenden Konstrukts die Komplexität der Diagnose- und Messinstrumente reduziert und in der Regel eine schärfere Aussage getroffen werden kann (vgl. Loerwald/Schnell 2016).

Auf dem Niveau der Sekundarstufe II gibt es bereits eine Reihe an deutschsprachigen etablierten Testinstrumenten zur Erhebung von Wirtschaftswissen. Aus unserer Sicht ergibt sich dennoch ein weiterer Forschungsbedarf. Der Großteil der etablierten deutschsprachigen Testinstrumente setzt beispielsweise auf ein reines Single-Choice-Aufgabenformat, welches aus didaktischen Überlegungen eine Reihe von Nachteilen mit sich bringt (vgl. Loerwald/Schnell 2014). Viele Testinstrumente beziehen sich auf einen bestimmten Schultyp bzw. Schulstufe (vgl. z. B. Würth/Klein 2001, Brandlmaier et al. 2006, Lehmann/Seeber 2009), die mit der Zielgruppe des vorliegenden Projekts nicht kompatibel ist. Der Wirtschaftskundliche Bildungstest (Übersetzung des Test of Economic Literacy in dritter Version, vgl. Walstad/Rebeck/Butters 2013) bildet beispielsweise nur volkswirtschaftliche Inhalte ab, was Beck (2000)

zumindest im Hinblick auf die Bereiche Rechnungswesen und Marketing für diskussionswürdig erachtet. Das von Schumann und Eberle (2014) entwickelte OEKOMA-Testinstrument, welches in der Deutschschweiz bei über 2.000 Maturandinnen und Maturanden von Berufsmaturitätsschulen und Allgemeinbildenden höheren Schulen zwischen 2010 und 2012 eingesetzt wurde, umfasst sowohl volkswirtschaftliche als auch betriebswirtschaftliche Inhalte sowie Aufgaben aus dem Bereich Rechnungswesen. Um die Eignung des OEKOMA-Instruments für Österreich zu untersuchen, wurden ausgewählte Testaufgaben bei einer Stichprobe von rund 280 BHS Schülerinnen und Schülern der 5. Klassen im Rahmen einer klein angelegten Studie des Wipäd-Instituts der WU Wien zwischen September und Oktober 2017 eingesetzt. Wie sich herausgestellt hat, ist der Einsatz der Items im leseintensiven Zeitungsartikel-Format für Österreich wenig geeignet. Das OEKOMA-Testinstrument wurde im Jahr 2009 konzipiert, daher weisen ein Großteil der fiktiven Zeitungsartikel einen mangelnden Aktualitätsbezug auf. Ein Großteil der eingesetzten Items musste aufgrund landesspezifischer Eigenheiten überarbeitet werden, wodurch die Originalität des Testinstruments im Grunde verloren ging. Die Testaufgaben weisen zudem eine mittelmäßige psychometrische Qualität auf.

## **2 Entwicklung des Testinstruments**

Um ein nach den Vorgaben der wissenschaftlichen Testkonstruktion konformes Testinstrument entwickeln zu können, werden im Rahmen dieses Projekts mehrere Schritte gesetzt. Diese umfassen zunächst die Inhaltsbestimmung der Testaufgaben zum Wirtschaftswissen anhand einer Medienanalyse (vgl. Früh 2015, Merten 1995, Schumann/Eberle/Oepke 2010) sowie die Exploration der Zielgruppe des Testinstruments anhand problemzentrierter Interviews (vgl. Witzel 2000, Witzel/Reiter 2012, Kuckartz 2016). Die Erkenntnisse aus diesen Voruntersuchungen tragen wesentlich zur Entwicklung der Testitems bei. Das nach den Vorgaben der wissenschaftlichen Testkonstruktion entwickelte Testinstrument wurde bereits bei einem ersten Testlauf (Pretest) eingesetzt und erstmals empirisch erprobt (vgl. Bühner 2013, Lienert/Raatz 1998). Anhand der Erkenntnisse aus dem Pretest wird das Testinstrument überarbeitet und schließlich in einem zweiten Testlauf österreichweit eingesetzt.

### **2.1 Medienanalyse zur Bestimmung der Inhalte der Testaufgaben**

#### *2.1.1 Zielsetzung und methodische Vorgehensweise*

Um die Inhalte für die Testaufgaben zum Wirtschaftswissen zu bestimmen, wurde eine Medienanalyse der Wirtschaftsberichte ausgewählter österreichischer Tageszeitungen durchgeführt. Die Einschränkung auf die Analyse von Tageszeitungen wird dadurch begründet, dass die Gesamtreichweite von Tageszeitungen (Print- und E-Paper-Format) in Österreich höher ist als jene der Magazine, Fachzeitschriften oder auch Zeitungs-Apps (vgl. Media Analyse 2020). Zudem decken sich die Inhalte zwischen dem Print- und E-Format der analysierten Tageszeitungen Großteils. Die Unveränderlichkeit der analysierten Inhalte ist ein wichtiges Kriterium, um die Ergebnisse der Inhaltsanalyse auch noch zu einem späteren Zeitpunkt nachvollziehen zu können (vgl. Früh 2015).



Die Medienanalyse basiert auf der Überlegung, über welche wirtschaftlichen Begriffe und Sachverhalte man als kundige/r Zeitungleser/in Bescheid wissen sollte, um die Wirtschaftsberichte in Tageszeitungen verstehen zu können. Anhand einer themenbezogenen Inhaltsanalyse wurden zentrale wirtschaftliche Begriffe bzw. Sachverhalte systematisch identifiziert und in einem Kategoriensystem zusammengefasst (vgl. Früh 2015, Merten 1995). Das Kategoriensystem wurde in Anlehnung an die im Schweizer OEKOMA-Projekt durchgeführte Medienanalyse deduktiv abgeleitet, wobei länderspezifische und sprachliche Unterschiede in den Kategorien angepasst wurden (vgl. Schumann/Eberle/Oepke 2010). In einem Erhebungszeitraum zwischen Februar 2018 und Februar 2019 wurden anhand einer geklumpten Auswahl nach wöchentlichem Rotationsprinzip (vgl. Früh 2015) Wirtschaftsberichte von sechs ausgewählten österreichischen Tageszeitungen analysiert (Die Presse, Der Standard, Kurier, Kleine Zeitung, Kronen Zeitung, Gratiszeitung Heute). Die Stichprobe umfasst rund 750 analysierte Zeitungsartikel, woraus rund 10 Tausend wirtschaftliche Begriffe bzw. Sachverhalte identifiziert werden konnten.

### 2.1.2 Ergebnisse im Überblick

In der folgenden Abbildung werden die quantitativen Ergebnisse der Medienanalyse (Häufigkeitsanalyse absolut und prozentuell) sowie die approximativ errechneten Richtgrößen für die zu erstellenden Testaufgaben zum Wirtschaftswissen dargestellt:

Haupt- und Unterkategorien		Verteilung absolut	Verteilung in Prozent		Richtgrößen
Codes	<b>Volkswirtschaft</b>	Summe	Summe	VW/BW	30 Items
A1	Grundlagen	701	6,9%	13,0%	3
A2	Mikroökonomie	1055	10,4%	19,6%	3
A3	Makroökonomie	2650	26,2%	49,2%	7
A4	Internationale Wirtschaftsbeziehungen	976	9,7%	18,1%	3
<b>Summe Volkswirtschaft</b>		<b>5382</b>	<b>53,3%</b>	<b>100,0%</b>	<b>16</b>
Haupt- und Unterkategorien		Verteilung absolut	Verteilung in Prozent		Richtgrößen
Codes	<b>Betriebswirtschaft</b>	Summe	Summe	VW/BW	30 Items
B1	Grundlagen	1929	19,1%	40,9%	6
B2	Strategische Führung	560	5,5%	11,9%	1
B3	Organisation, Marketing, Personal, Produktion	1252	12,4%	26,5%	5
B4	Finanz- und Rechnungswesen	979	9,7%	20,7%	2
<b>Summe Betriebswirtschaft</b>		<b>4720</b>	<b>46,7%</b>	<b>100,0%</b>	<b>14</b>
<b>Gesamtsumme</b>		<b>10102</b>			<b>30</b>

Abbildung 2: Ergebnisse der Medienanalyse, Häufigkeitsverteilung der Hauptkategorien eigene Darstellung (in Anlehnung an Schumann/Eberle/Oepke 2010)

Das Kategoriensystem untergliedert sich zunächst in die zwei thematischen Großbereiche Betriebswirtschaft und Volkswirtschaft mit jeweils vier Hauptkategorien. Jede Hauptkategorie umfasst drei bis acht Subkategorien, die hier aus Platzgründen nicht dargestellt sind. Bei den in den Wirtschaftsberichten identifizierten Begriffen handelt es sich beispielsweise um Begriffe wie Bilanz oder Begriffskombinationen wie Europäische Zentralbank. Der Begriff Bilanz wird beispielsweise der Kategorie B4 Finanz- und Rechnungswesen zugeordnet. Die prozentuelle Gesamtverteilung untermauert ein insgesamt relativ ausgeglichenes Verhältnis zwischen den

Bereichen Volks- und Betriebswirtschaft (53,3% zu 46,7%), über die in den Wirtschaftsberichten der Tageszeitungen berichtet wird. Das Ergebnis zeigt, dass insbesondere dem Verständnis für gesamtwirtschaftliche (26,2% der Gesamtnennungen) sowie grundlegende betriebswirtschaftliche Sachverhalte (19,1%) eine hohe Bedeutung zukommt. Aus dem Kategoriensystem lässt sich eine thematische Orientierungshilfe für die zu erstellenden Testaufgaben zum Wirtschaftswissen ableiten. Damit sollte ein im ersten Schritt inhaltlich valider Rahmen für das Testinstrument geschaffen werden.

## **2.2 Festlegung des Testformats und der Aufgabenmerkmale**

Bevor die Testitems inhaltlich konzipiert wurden, musste zunächst das Test- bzw. Aufgabenformat festgelegt werden, welches sich zunächst an vergleichbaren, etablierten Testinstrumenten orientierte (vgl. z. B. Rumpold 2018, Walstaad et al. 2013). Das Testinstrument soll aus einem Leistungstest mit Testaufgaben zum Wirtschaftswissen und einem Kontextfragebogen bestehen, welcher sowohl soziodemografische Merkmale als auch weitere Kompetenzfacetten wie z. B. das wirtschaftliche Interesse erfassen soll. Die Befragungsdauer soll 45 Minuten nicht übersteigen, da sowohl der erste Testlauf als auch der österreichweite zweite Testlauf covid-bedingt gänzlich online stattfinden muss.

Um das Wirtschaftswissen zu erheben wird ein geschlossenes Aufgabenformat bevorzugt, mit dem grundsätzlich eine kurze Bearbeitungs- bzw. Lösungszeit einhergeht. Die Aufgaben wurden in einem echten Multiple-Choice-Aufgabenformat konzipiert. Jede Aufgabe wurde mit vier Antwortmöglichkeiten versehen, wovon eine oder mehrere richtig sein können. Bei diesem Aufgabenformat bedarf es auch keiner zeitintensiven Testinstruktion, was besonders bei Online-Befragungen eine wichtige Rolle spielt. Zudem können geschlossene Aufgabenformate effizienter ausgewertet werden, erhöhen die Auswertungsobjektivität im Gegensatz zu offenen Aufgaben und können somit bei einer großen Anzahl an Testteilnehmerinnen und -teilnehmern forschungsökonomisch eingesetzt werden. Beim echten Multiple-Choice-Format ist die Formulierung der Distraktoren in der Regel zwar anspruchsvoller, es ermöglicht jedoch auch das Abbilden umfassenderer Inhaltsbereiche, da jede Antwortalternative einzeln als richtig oder falsch bewertet werden kann. Zudem reduziert das Multiple-Choice-Format grundsätzlich die Rategefahr (vgl. Bühner 2013, Loerwald/Schnell 2014, Lienert/Raatz 1998). Die Testitems im Kontextfragebogen sind eine Kombination aus überwiegend geschlossenen Fragen, wie z. B. die Frage nach dem zuletzt besuchten Schultyp sowie Fragen im Rating-Format, wie z. B. Items zum wirtschaftlichen Interesse.

Im Hinblick auf eine fachdidaktisch fundierte Erstellung der Testaufgaben zum Wirtschaftswissen wurden unterschiedliche kognitive Anforderungsniveaus berücksichtigt (vgl. Anderson/Krathwohl 2001, Krathwohl 2002) sowie weitere schwierigkeitsbestimmende Merkmale wie z. B. ein zielgruppenadäquater Sprachstil (vgl. Schumann/Eberle 2011). Kompliziert formulierte bzw. verschachtelte Sätze wurden weitestgehend vermieden. Zudem wurde ein besonderes Augenmerk daraufgelegt, dass die Testaufgaben fachlich einwandfrei, klar formuliert und disjunkt sind (vgl. Loerwald/Schnell 2014, Bühner 2013).



## 2.3 Inhaltliche Validierung

Um den Ansprüchen der wissenschaftlichen Testkonstruktion zu entsprechen, muss ein Testinstrument bestimmte Gütekriterien erfüllen. Eines der Hauptgütekriterien ist die Validität, die angibt, ob ein Test das misst, was er behauptet zu messen (vgl. Bühner 2013). Eine hohe Validität gibt Aufschluss darüber, ob das Testinstrument seinem Zweck gerecht wird und die Testergebnisse eine gültige Aussage über das zu messende Persönlichkeitsmerkmal ermöglichen (vgl. Bühner 2013). Für die inhaltliche Validierung werden im vorliegenden Projekt zwei qualitative Schritte eingesetzt. Für die Überprüfung des Begriffsverständnisses der Testaufgaben wurden anhand der Methode „Lauten Denkens“ (vgl. Schnell 2016, Loerwald/Schnell 2014, Moosbrugger/Kelava 2012a) drei Schülerinnen und Schüler aus zwei Maturaklassen Berufsbildender höherer Schulen befragt. Die Schülerinnen und Schüler wurden gebeten, die Items laut vorzulesen, in eigenen Worten zusammenzufassen und ihre Lösungswege offen zu legen. Die daraus gewonnenen Erkenntnisse geben einen ersten Aufschluss darüber, ob die Testaufgaben für die Zielgruppe inhaltlich angemessen sind. Insgesamt fanden die Schülerinnen und Schüler die Formulierungen verständlich, was zu nachvollziehbaren Gedankengängen bei der Lösung der Testaufgaben führte. Als weiteren Schritt wurde eine Validierung des Testinstruments durch zehn Expertinnen und Experten (Mitarbeiterinnen und Mitarbeiter des Lehrstuhls und Wirtschaftslehrkräfte) durchgeführt. Die Items wurden hinsichtlich der fachlichen Korrektheit, einer zielgruppenadäquaten Formulierung, der inhaltlichen Relevanz und des Schwierigkeitsniveaus beurteilt. Die Expertinnen und Experten wurden zudem gebeten etwaige Verbesserungsvorschläge bei den Testaufgaben anzumerken. Im Zuge dieses Prozesses wurden insbesondere die Testaufgaben zum Wirtschaftswissen überarbeitet und dadurch sprachlich klarer sowie aussagekräftiger formuliert und inhaltlich dem Testkonstrukt entsprechend abgerundet. Die Expertinnen und Experten beurteilten das Testinstrument insgesamt als inhaltlich relevant und hinsichtlich des Schwierigkeitsgrades für die Zielgruppe als angemessen.

### 2.3.1 Das Testinstrument in erster Version

Nach der umfangreichen qualitativen Beurteilung wurde das Testinstrument in seiner Erstversion fertig gestellt, um dieses in einem ersten Testlauf einzusetzen:

Bereich	Anzahl Items	Quellen
Wirtschaftswissen	30	Eigenentwicklung
Selbsteinschätzung zum Wirtschaftswissen	8	Wirtschaftswissen in der Sek. I (Rumpold 2018)
wirtschaftliches Interesse	8	Wirtschaftswissen in der Sek. I (Rumpold 2018)
Einstellungen zur Wirtschaft	9	Wirtschaftswissen in der Sek. I (Rumpold 2018)
Soziodemografische Merkmale	12	Eigenentwicklung

Abbildung 3: Anzahl der Items und Quellen, eigene Darstellung

In der Abbildung 4 ist eine Multiple-Choice-Aufgabe zum Wirtschaftswissen illustriert. Diese ist thematisch der Kategorie betriebswirtschaftliche Grundlagen zugeordnet. In den Antwortalternativen werden grundlegende rechtliche Rahmenbedingungen im Zusammenhang mit

Unternehmen in unserer Wirtschaft thematisiert, wie z. B., dass es auch kleine bzw. Ein-Personen-Unternehmen gibt (Antwortalternative 1) oder, dass nicht alle Unternehmen zwangsläufig auf Gewinn ausgerichtet sein müssen (Antwortalternative 3). Die Antwortalternative zwei ist fachlich korrekt.

*Mehrfachantworten möglich!*

Jedes Unternehmen in unserer Wirtschaft...

- hat mindestens eine angestellte Person und eine Vorgesetzte bzw. einen Vorgesetzten.
- unterliegt einem Marktrisiko.**
- ist darauf ausgerichtet, einen möglichst hohen buchhalterischen Gewinn zu erzielen.
- benötigt ein bestimmtes Mindestkapital, das auf ein Bankkonto eingezahlt werden muss.

Abbildung 4: Beispiel-Testaufgabe zum Wirtschaftswissen, eigene Abbildung

## 2.4 Beurteilung der psychometrischen Eignung

Neben der inhaltlichen Beurteilung wurde im nächsten Schritt die psychometrische Eignung der Items anhand statistischer Kennwerte überprüft. In diesem Zusammenhang spielt das Gütekriterium der Konstruktvalidität eine entscheidende Rolle, welche ausdrückt, ob ein Test tatsächlich jene Fähigkeit misst, die er vorgibt zu messen (vgl. Bühner 2013). Gemeint ist, ob z. B. von den Testergebnissen tatsächlich auf das latente Konstrukt Wirtschaftswissen geschlossen werden kann. Die Konstruktvalidität ist zwar eines der wichtigsten wissenschaftlichen Gütekriterien, aber auch jenes, das am Schwierigsten zu messen ist (vgl. Loerwald/Schnell 2016). Im vorliegenden Projekt erfolgt empirisch eine Annäherung über Itemanalysen nach der klassischen sowie probabilistischen Testtheorie, welche im nächsten Abschnitt auszugswise vorgestellt werden.

## 3 Ausgewählte Ergebnisse des ersten Testlaufs

### 3.1 Erhebungszeitraum und Stichprobe

Um statistische Itemkennwerte berechnen und analysieren zu können wurde das Testinstrument erstmals im Rahmen einer Online-Befragung eingesetzt. Beim ersten Testlauf konnten auch wichtige organisatorische sowie technische Erfahrungswerte gesammelt werden, wie z. B. die Teilnahmebereitschaft der kontaktierten Bildungseinrichtungen oder die Rücklauf- und Abbruchquote. Der Pretest erfolgte zwischen Mai und Oktober 2020 und umfasst eine bereinigte Stichprobe von 388 Personen (convenience sample). Bei den Befragten handelt es sich größtenteils um Studienbewerberinnen und Studienbewerber der Wirtschaftsuniversität Wien. Zudem umfasst die Stichprobe Studienanfängerinnen und Studienanfänger der Anglistik der Universität Wien sowie zwei wirtschaftliche Fachhochschulstudiengänge. Bei der Organisation der Stichprobe wurde zwar darauf geachtet, über unterschiedliche Studienrichtungen und Bundesländer zu streuen, aber inmitten der ersten Welle der weltweiten covid-19-Pandemie

ging die Organisation auch mit diversen organisatorischen Hürden einher. Eine Teilnahme an der Online-Befragung war schließlich vom Interesse bzw. der Kooperationsbereitschaft der kontaktierten Bildungseinrichtungen abhängig.

### 3.2 Ausgewählte Stichprobencharakteristika

Die folgenden Abbildungen zeigen einen ersten Einblick in die Stichprobencharakteristika der befragten Personen aus dem ersten Testlauf:

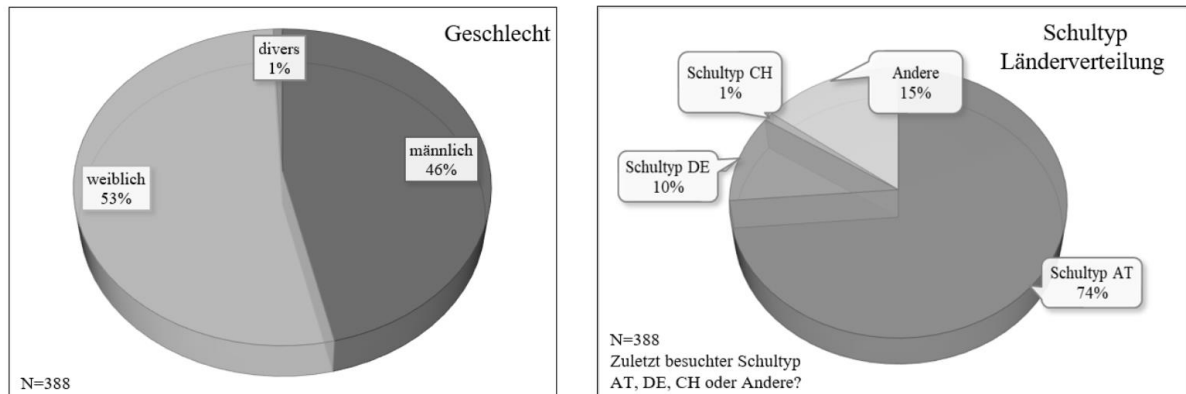


Abbildung 5: Stichprobenverteilung nach Geschlecht und nach zuletzt besuchtem Schultyp nach Länder verteilt, eigene Darstellung

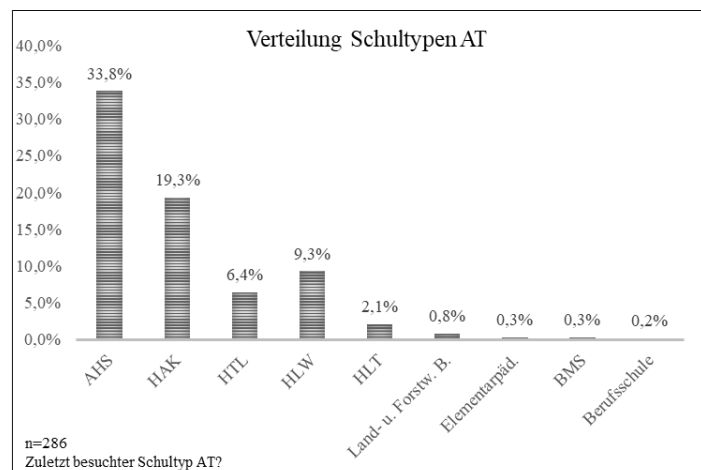


Abbildung 6: Stichprobenverteilung nach zuletzt besuchtem Schultyp in Österreich, eigene Darstellung

Die Verteilung der männlichen bzw. weiblichen Befragten innerhalb der Stichprobe ist recht ausgewogen. Ungefähr zwei Drittel der Befragten haben die universitäre Reife in Österreich erlangt. Die Verteilung nach den absolvierten Schultypen innerhalb Österreichs (n=286) zeigt ein durchaus vergleichbares Bild zu den bereits dargestellten kumulierten Übertrittsraten in den Hochschulsektor. Ein Großteil der Befragten hat die AHS Oberstufe absolviert (33,8%), gefolgt von der Handelsakademie (19,2%).

### 3.3 Itemanalyse nach klassischer Testtheorie

Vorrangiges Ziel der vorliegenden Itemanalysen ist, jene Testaufgaben zum Wirtschaftswissen auszulesen, die eine sowohl psychometrisch valide als auch inhaltskonforme Erfassung des Wirtschaftswissens ermöglichen. Jene Testaufgaben, die diesen Kriterien nicht entsprechen, sollen folglich inhaltlich adaptiert oder als letzte Konsequenz verworfen werden. Erste Erkenntnisse dazu liefert die Itemanalyse nach der klassischen Testtheorie (vgl. Moosbrugger et al. 2012b), welche mit der Software SPSS durchgeführt wurde. Die Lösungsraten der Testaufgaben zum Wirtschaftswissen wurden über zwei Bewertungszugänge ermittelt. Bei der Bewertung zu Teilpunkten können je Antwortalternative 0,25 Punkte erzielt werden. Die Bewertung zu Vollpunkten beruht auf dem Alles-oder-nichts-System, in dem pro Aufgabe maximal ein Punkt erzielt werden kann, falls alle vier Antwortalternativen richtig bewertet werden. Insgesamt konnten 30 Punkte erzielt werden. Die Gesamtlösungsraten der zwei Bewertungssysteme sind in der folgenden Abbildung veranschaulicht:

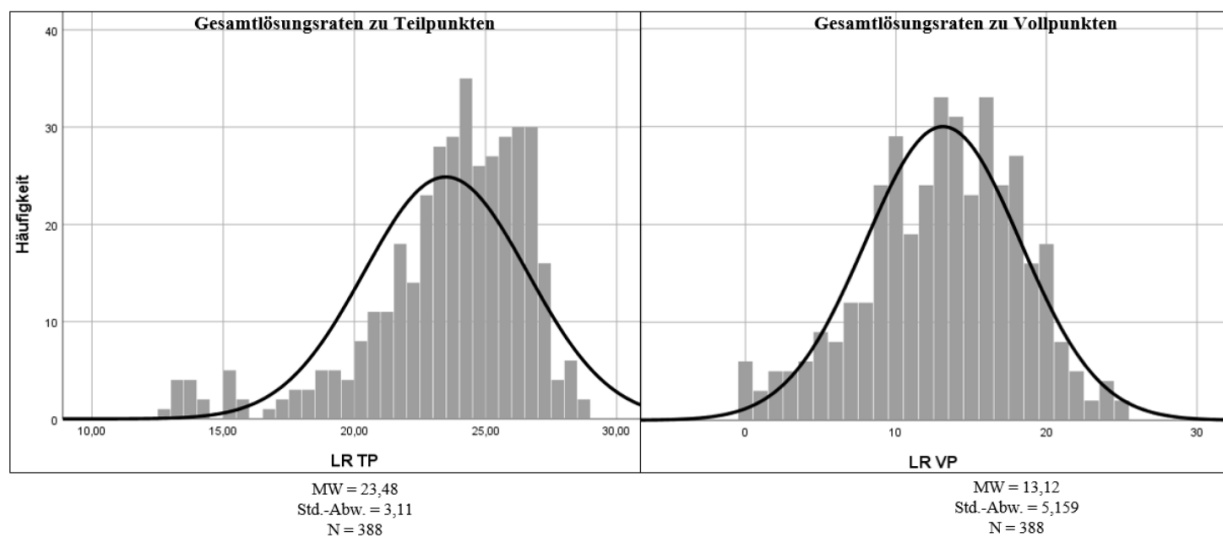


Abbildung 7: Lösungsraten zu Teil- und Vollpunkten, eigene Darstellung

Im linken Histogramm ist eine deutlich linksschiefe Verteilung zu Gunsten höherer Lösungsraten zu erkennen mit einer relativ hohen mittleren Lösungsrate von 23 Punkten und einer Standardabweichung von rund drei Punkten. Die Verteilung der Gesamtlösungsraten zu Vollpunkten zeigt ein durchaus anderes Bild. Mit einer mittleren Lösungsrate von 13 Punkten, also weniger als die Hälfte der Gesamtpunkte, fallen die Lösungsraten in diesem Bewertungssystem wie erwartet deutlich geringer aus. Unterschiedliche Bewertungssysteme können zu differenzierten Analyseergebnissen führen. Aus diesem Grund wurden die Itemkennwerte auf Ebene der einzelnen Antwortalternativen (zu Teilpunkten) und auf aggregierter Ebene zu Teil- bzw. Vollpunkten in die Analyse miteinbezogen, um sich ein möglichst umfassendes Bild über die psychometrische Eignung der Testaufgaben machen zu können. Aus Platzgründen bzw. zur Wahrung der Übersichtlichkeit werden in der folgenden Tabelle lediglich die aggregierten Itemkennwerte zu Teilpunkten pro Aufgabe dargestellt.

Tabelle 1: Ergebnisse der Itemanalyse (zu Teilpunkten) nach klassischer Testtheorie

Inhaltsbereiche	Itemnr.	N	Min	Max	MW	Std.- Abweichung	Schwierigkeit	Trenn- schärfe	Cronbachs Alpha	Ladungen	Varianz- aufklärung
Unternehmen als Akteure der Wirtschaft	B01	388	0,00	1	0,726	0,263	72,6%	0,552	0,874	0,617	22,1%
Unternehmerische Rahmenbedingungen	B02	388	0,00	1	0,688	0,254	68,8%	0,360		0,405	
Unternehmensverpflichtungen	B03	388	0,25	1	0,758	0,207	75,8%	0,390		0,424	
Unternehmensgewinne	B04	388	0,00	1	0,655	0,189	65,5%	0,190		0,236	
Stakeholder von Unternehmen	B05	388	0,00	1	0,832	0,209	83,2%	0,560		0,635	
Management	B06	388	0,00	1	0,816	0,218	81,6%	0,402		0,471	
Konzerne/Unternehmenszusammenschluss	B07	388	0,00	1	0,798	0,199	79,8%	0,193		0,228	
Marketingaktivitäten	B08	388	0,00	1	0,802	0,233	80,2%	0,482		0,572	
Marken	B09	388	0,25	1	0,829	0,213	82,9%	0,512		0,581	
Arbeitsverträge (Arbeitgebersicht)	B10	388	0,00	1	0,789	0,225	78,9%	0,426		0,488	
Industriebetriebe	B11	388	0,00	1	0,869	0,167	86,9%	0,499		0,571	
Finanzwesen Unternehmen	B12	388	0,00	1	0,863	0,199	86,3%	0,465		0,538	
Wertpapiere und Risiko	B13	388	0,00	1	0,864	0,207	86,4%	0,589		0,655	
ökonomisches Denken/Prinzip	V01	388	0,00	1	0,648	0,236	64,8%	0,350		0,395	
Produktionsfaktoren	V02	388	0,00	1	0,727	0,232	72,7%	0,323		0,373	
Wirtschaftskreislauf	V03	388	0,00	1	0,841	0,223	84,1%	0,466		0,546	
Marktwirtschaft	V04	388	0,00	1	0,902	0,232	90,2%	0,469		0,534	
Wirtschaftssektoren	V05	388	0,25	1	0,852	0,216	85,2%	0,482	0,548		
Angebot und Nachfrage	V06	388	0,00	1	0,738	0,299	73,8%	0,235	0,266		
Angebotsmonopole	V07	388	0,00	1	0,670	0,219	67,0%	0,345	0,390		
Wirtschaftswachstum/BIP	V08	388	0,00	1	0,849	0,214	84,9%	0,503	0,567		
Inflation	V09	388	0,00	1	0,687	0,234	68,7%	0,240	0,281		
Arbeitsverträge (aus Sicht privater Haushalte)	V10	388	0,25	1	0,733	0,196	73,3%	0,100	0,114		
Staatshaushalt	V11	388	0,00	1	0,861	0,196	86,1%	0,453	0,522		
Wirtschaftspolitik	V12	388	0,00	1	0,858	0,218	85,8%	0,615	0,689		
Zinsen und Spareinlagen	V13	388	0,00	1	0,852	0,221	85,2%	0,434	0,484		
Kredite (aus Sicht privater Haushalte)	V14	388	0,00	1	0,719	0,235	71,9%	0,400	0,467		
Geldanlage und Risiko	V15	388	0,00	1	0,805	0,248	80,5%	0,518	0,565		
Wechselkurse	V16	388	0,00	1	0,794	0,267	79,4%	0,342	0,378		
Europäische Union	V17	388	0,00	1	0,657	0,246	65,7%	0,262	0,292		

Die Tabelle zeigt zunächst die Inhaltsbereiche der einzelnen Aufgaben sowie die **Minima** und **Maxima** der erreichten Punkte. Beim Großteil der Items wird idealerweise das gesamte Punktespektrum (0 bis 1) voll ausgeschöpft. Beispielsweise wurde bei dem Item B03 (Unternehmensverpflichtungen) oder V10 (Arbeitsverträge) über die Gesamtstichprobe mindestens eine Antwortalternative richtig beantwortet, was ein erster Hinweis dafür ist, dass die psychometrische Eignung unzureichend ist. Die aus den Mittelwerten errechneten **Schwierigkeitswerte** drücken den Anteil der Befragten aus, die eine Aufgabe richtig gelöst haben und stellen somit auch die Lösungsraten (zu Teilpunkten) dar. Schwierigkeitswerte von unter 25% bzw. über 85% werden als unangemessen eingestuft, wobei in vergleichbaren Studien auch Grenzwerte von 5% bzw. 95% zur Anwendung kommen (vgl. z. B. Itzlinger-Bruneforth et al. 2016, Oberrauch 2019). Die Schwierigkeitswerte, welche im Mittel rund 78,3% betragen, liegen mit einer Reichweite von 68,8% bis 90,2% überwiegend in der oberen Hälfte des Punktespektrums. Der hohe Schwierigkeitsbereich ist demnach eher unterrepräsentiert. Beispielsweise sind die Items V04

(Marktwirtschaft) und V12 (Wirtschaftspolitik) tendenziell als zu leicht (>85%) einzustufen. Die Detailanalyse auf Basis der einzelnen Antwortalternativen zeigt jedoch, dass einige Antwortalternativen durchaus hohe Schwierigkeiten aufweisen, was in den aggregierten Itemkennwerten zu Teilpunkten nicht ersichtlich ist. Die **Standardabweichung** zeigt, ob eine angemessene Streuung der Punktwerte vorliegt. Eine mittlere Standardabweichung von 0,23 Punkten mit einer Gesamtreichweite von rund 0,17 bis 0,30 ist als eher niedrig einzuschätzen, was auch den insgesamt relativ hohen Lösungsraten geschuldet ist.

Ein weiterer Indikator für die psychometrische Eignung der Testaufgaben ist die **Trennschärfe**, welche angibt, wie gut ein Item zwischen leistungsstarken bzw. leistungsschwachen Befragten unterscheiden kann. Eine Trennschärfe nahe null würde bedeuten, dass es irrelevant ist, ob ein Befragter tatsächlich Wirtschaftswissen aufweist, um das Item richtig zu lösen (Bühner 2013). In vergleichbaren Studien wird in der Regel ein unterer Grenzwert von 0,2 angeführt (vgl. Itzlinger-Bruneforth et al. 2016, Walstad/Rebeck/Butters 2013). Die Trennschärfen liegen im Mittel bei einem zufriedenstellenden Wert von 0,41 mit einer Gesamtreichweite von 0,10 bis 0,62. Die Items V10 (Arbeitsverträge), B04 (Unternehmensgewinne) sowie B07 (Unternehmenszusammenschlüsse) erweisen sich als zu wenig trennscharf.

**Cronbachs Alpha** gibt Aufschluss darüber, wie gut die Testaufgaben miteinander korrelieren. Ein erzielter Wert von 0,874 über alle 30 Testaufgaben kann als sehr zufriedenstellend beurteilt werden. Cronbachs Alpha ist aber kein Indikator dafür, ob die Testaufgaben zu einer eindimensionalen Gesamtskala zusammengefasst werden können, da die Items auch korrelierte Messfehler besitzen können (vgl. Bühner 2013). Daher wurde im nächsten Schritt eine **explorative Faktorenanalyse** nach der Hauptkomponentenmethode durchgeführt, in der zunächst die Eignung der vorliegenden Daten anhand des KMO-Koeffizienten und des Bartlett-Tests verifiziert wurde. Die Dimensionalität der Testaufgaben wurde mittels MAP-Test sowie einem grafischen Scree-Test bestimmt, wobei grundsätzlich eine eindimensionale Lösung angestrebt wird. Das Ergebnis zeigt, dass eine eindimensionale Lösung rund 22% der Gesamtvarianz aufklärt. Bei einer zweidimensionalen Lösung könnte die erklärte Gesamtvarianz zwar auf 29% gesteigert werden, jedoch ist die Dominanz des ersten Faktors in diesem Fall hervorzuheben. Ein solches Ergebnis ist durchaus vergleichbar mit ähnlichen Studien zur Messung des Wirtschaftswissens bzw. der Wirtschaftskompetenz (vgl. z. B. Rumpold 2018, Oberrauch 2019). Aus der Tabelle 1 sind zudem die Faktorenladungen der einzelnen Items zu entnehmen, wobei hier ein Wert von größer 0,2 als angemessen gilt. Die Detailergebnisse pro Antwortalternative legen teilweise höhere Faktorenladungen offen, die im Zuge der Überarbeitung der Items zu berücksichtigen sind.

### 3.4 Itemanalyse nach probabilistischer Testtheorie

Einen tiefergehenden Einblick über die psychometrische Eignung der Testaufgaben ergibt sich durch die Itemanalyse nach der probabilistischen Testtheorie. Eine wesentliche Voraussetzung für die Anwendung des Rasch-Modells ist die Eindimensionalität. Das bedeutet, dass alle Items die gleiche latente Fähigkeit – also das Wirtschaftswissen – messen. Vorteile vom Rasch-Modell sind einerseits die Messung der Aufgabenschwierigkeit und Personenfähigkeit auf einer



gemeinsamen Skala und andererseits die Vergleichbarkeit der Analyseergebnisse mit ähnlichen Studien (Moosbrugger et al. 2012b). Die Itemanalyse wurde mit der Software R und dem Zusatzpaket eRM durchgeführt. Das Rasch-Modell folgt einer logistischen Regressionsgleichung (Maximum-Likelihood Ratio), die für die Aufgabenantworten Wahrscheinlichkeitswerte berechnet. Für die Analyseebene zu Teilpunkten war das berechnete Rasch-Modell jedoch nicht aussagekräftig, weil bei manchen Items bestimmte Teilpunktwerte nicht in ausreichender Fallzahl vorhanden sind. Die Werte können daher nicht direkt mit der Itemanalyse nach klassischer Testtheorie verglichen werden. Nichtsdestotrotz liefern die Itemkennwerte wichtige Hinweise für die weitere Vorgehensweise bei der Itemrevision. Die 30 Testaufgaben wurden daher zunächst in ein dichotomes Antwortformat umgewandelt (1= Aufgaben vollständig richtig gelöst, 0= mindestens eine Antwortalternative falsch gelöst). Die über das dichotome Rasch-Modell geschätzten Aufgabenschwierigkeiten und Personenfähigkeiten sind auf der **Person-Item-Map** in Abbildung 8 dargestellt:

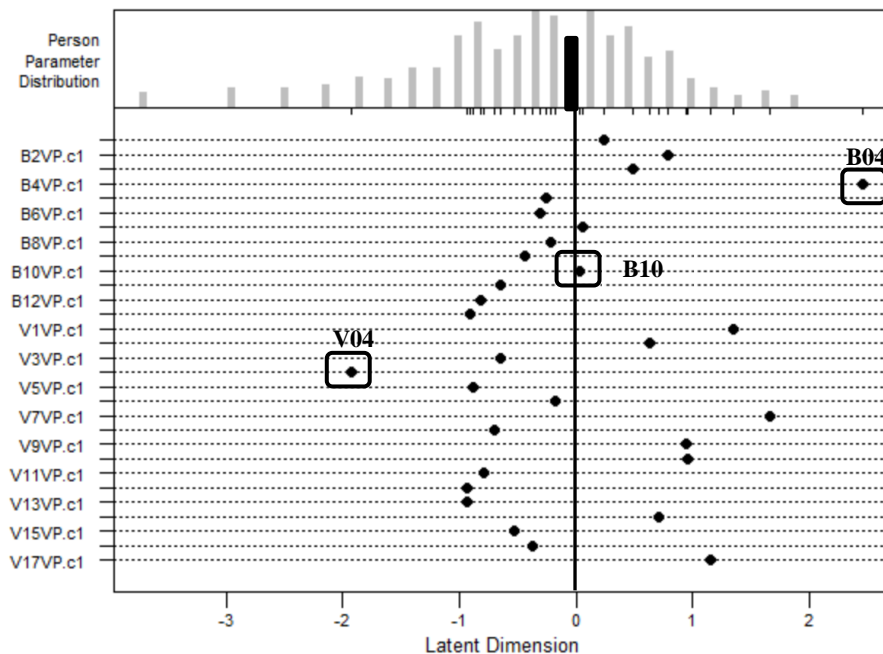


Abbildung 8: Person-Item-Map der Aufgabenschwierigkeiten und Personenfähigkeiten, eigene Abbildung

Die Aufgabenschwierigkeiten sind in den Punkten dargestellt. Die Testaufgaben mit Schwierigkeiten im Minusbereich sind als tendenziell leichter und Werte im Plusbereich als schwieriger einzustufen. Der Großteil der Items ist grundsätzlich zufriedenstellend gestreut in einem angemessenen Bereich von -2 bis +2. Lediglich die Items B04 (Unternehmensgewinne) und V04 (Marktwirtschaft) – welche auch nach der Itemanalyse nach klassischer Testtheorie unzureichende psychometrische Eignung aufweisen – stellen Ausreißer dar. Die Personenfähigkeiten sind in den Balken ersichtlich. Beispielsweise lösen die Befragten, die in der Grafik im schwarzen Balken hervorgehoben sind, die Testaufgabe B10 (Arbeitsverträge) mit ca. 50%iger Wahrscheinlichkeit. Alle Aufgaben links davon fallen diesen Befragten leichter, die Aufgaben rechts fallen ihnen schwerer. Bei den Personenfähigkeiten ist ein Überhang im linken Bereich

zu erkennen, was bedeutet, dass verhältnismäßig mehr Personen mit geringerer Fähigkeit als der durchschnittlichen Aufgabenschwierigkeit vorhanden sind.

Um zu testen, ob alle 30 Testaufgaben grundsätzlich zu den Annahmen des Rasch-Modells passen, wurde ein **Andersen Likelihood-Ratio Test** durchgeführt. Den restriktiven Anforderungen des Rasch-Modells konnte in diesem Zusammenhang nicht Folge geleistet werden, da der Test bei 30 Items signifikant ausgefallen ist, was aufgrund der Breite der Testaufgaben durchaus zu erwarten war. Aus dem **grafischen Modelltest** mit 95%-Konfidenzintervall der Schwierigkeitsparameter und dem Median der Lösungsraten als Splitkriterium (nicht abgebildet) geht jedoch hervor, dass der Großteil der Items eine durchaus passable Modellpassung aufweist. Nur bei den Testaufgaben B04 und V04 konnten merkliche Abweichungen festgestellt werden.

Des Weiteren wurde überprüft, inwieweit die geschätzten Aufgabenparameter in die theoretische Modellvorhersage passen. Hierfür wurden **Modell-Fitwerte** berechnet und im Detail analysiert (Infit MSQ, Outfit MSQ, Discrimination). Insgesamt liegen die mittleren Infit- bzw. Outfitwerte innerhalb der üblichen Grenzwerte von 0,6 bis 1,5. Auch die mittlere Trennschärfe (0,304) ist durchaus zufriedenstellend. Lediglich die Aufgaben B04, B07, V07, V09 sowie V10 weisen zu niedrige Trennschärfen ( $<0,2$ ) auf. Dieses Ergebnis deckt sich größtenteils mit den Ergebnissen nach der klassischen Testtheorie.

### 3.5 Ergebnisse aus der Überarbeitung des Testinstruments

Die aus den Itemanalysen gewonnenen statistischen Kennwerte sollen dabei helfen redundante oder schlecht formulierte Items aufzuspüren. Der Ausschluss einer Aufgabe aufgrund einer schlechten psychometrischen Eignung darf jedoch nicht zu Lasten der Inhaltsbereiche des Wirtschaftswissens gehen. Manchmal ist auch nur eine Umformulierung einer Antwortalternative sinnvoll statt ein Item gänzlich auszuschließen. Die Testaufgaben zum Wirtschaftswissen decken naturgemäß ein sehr breites Inhaltsspektrum ab, daher müssen auch ausreichend heterogene Items enthalten sein. Ein Item soll daher nicht nur aufgrund einer schlechten Trennschärfe entfernt werden. Die Schwierigkeitswerte müssen zudem in Relation zur Pretest-Stichprobe interpretiert werden, da es sich beim Großteil der Befragten um Studienbewerberinnen und -bewerber wirtschaftlicher Studienrichtungen handelt und davon auszugehen ist, dass die Befragten ein höheres wirtschaftliches Interesse aufweisen als jene nicht-wirtschaftlicher Studienrichtungen. Das finale Testinstrument in zweiter Version umfasst schließlich 27 Testaufgaben zum Wirtschaftswissen. Insgesamt wurden vier Testaufgaben verworfen, da diese sowohl inhaltlich redundant als auch wenig zufriedenstellende psychometrische Kennwerte aufweisen. Zum Beispiel wurde die Aufgabe V10 (Arbeitsverträge) entfernt, da ähnliche Inhaltsbereiche auch im Item B10 vorkommen. Ebenso wurde die Aufgabe V12 (Wirtschaftspolitik) verworfen, die als zu leicht und auch teilweise redundant einzustufen ist. Bei ein paar Items wurden einzelne Antwortalternativen zugunsten einer trennschärferen Formulierung etwas umformuliert. Ein Item, welches den Inhaltsbereich Europäische Zentralpolitik und Geldpolitik umfasst, wurde aus inhaltlichen Gründen ergänzt. Die Items zum wirtschaftlichen Interesse, den Einstellungen zur Wirtschaft und zur Selbsteinschätzung zum Wirtschaftswissen

wurden unverändert übernommen. Bei den soziodemografischen Fragen wurden marginale Änderungen vorgenommen. Zum Beispiel wurde die Filterfrage zur Hauptstudienrichtung leicht angepasst.

## 4 Zusammenfassung und Ausblick

Ziel dieses Beitrags ist einen ersten Einblick über die Entwicklung eines standardisierten wissenschaftlichen Testinstruments zur Messung des Wirtschaftswissens von jungen Erwachsenen, die sich zum Testzeitpunkt in der Phase zwischen Studienbewerbung und Studienbeginn befinden, zu geben. Die erste Version des Testinstruments wurde im Mai 2020 fertig gestellt und im Rahmen eines Pretests bei rund 388 jungen Erwachsenen erstmals eingesetzt.

Der vorliegende Beitrag konnte folgende Ergebnisse aus dem Pretest aufzeigen: Die inhaltliche Validität und didaktische Eignung der selbst entwickelten Testaufgaben zum Wirtschaftswissen wurde von Expertinnen und Experten umfangreich beurteilt. Die Testaufgaben wurden als inhaltlich relevant und hinsichtlich des Schwierigkeitsgrades für die Zielgruppe als geeignet eingestuft. Die Angemessenheit der Aufgaben für die Zielgruppe wurde zudem anhand der Methode lauten Denkens mit drei Maturantinnen bzw. Maturanten untersucht. Dabei konnten Verständnisschwierigkeiten bei der Formulierung der Aufgaben sowie deren Gedankengänge bei der Lösung der Aufgaben erkundet werden.

Die psychometrische Eignung des Testinstruments wurde anhand von Itemanalysen sowohl nach klassischer als auch probabilistischer Testtheorie beurteilt. Die Analyseergebnisse gaben einen aufschlussreichen Einblick über die Passung der Testaufgaben zum latenten Konstrukt Wirtschaftswissen und führten schließlich zu einer gezielten Revision der Items. Nach den Kriterien der klassischen Testtheorie ist der hohe Schwierigkeitsbereich der Testaufgaben eher unterrepräsentiert. Die Detailanalyse auf Basis der einzelnen Antwortalternativen zeigt jedoch, dass einige Antwortalternativen durchaus höhere Schwierigkeiten aufweisen, was bei der Überarbeitung der Items entsprechend berücksichtigt wurde. Hinsichtlich der internen Konsistenz des Testinstruments weisen die Items größtenteils mittlere Trennschärfen auf und ein zufriedenstellendes Cronbachs Alpha von 0,874. Einzelne Items erfüllen den geforderten Grenzwert von 0,2 zwar nicht, aber aufgrund des breiten Inhaltsspektrums der Testaufgaben zum Wirtschaftswissen müssen auch ausreichend heterogene Items enthalten sein, was ggf. zu Lasten einer schlechteren Trennschärfe geht. Die faktorenanalytische Dimensionsanalyse weist auf eine angestrebte eindimensionale Lösung hin. Die damit aufgeklärte Gesamtvarianz mit rund 22% fällt zwar nicht sonderlich hoch aus, ist aber durchaus vergleichbar mit ähnlichen Studien zur Messung des Wirtschaftswissens.

Die nach der probabilistischen Testtheorie vorgenommene Itemanalyse anhand eines dichotomen Rasch-Modells bestätigt größtenteils die Schlussfolgerungen aus der Analyse nach klassischer Testtheorie. Die geschätzten Aufgabenschwierigkeiten streuen mit zwei Ausnahmen über einen angemessenen Schwierigkeitsbereich. Die Passung der erhobenen Daten zum Rasch-Modell wurde sowohl anhand eines Anderson Likelihood-Ratio Tests als auch anhand eines

grafischen Modelltests durchgeführt. Der grafische Modelltest zeigt, dass bis auf zwei Ausreißer ein wesentlicher Teil der Items eine passable Modellpassung aufweist, wenngleich im Rahmen der zweiten Testphase noch eine umfangreichere Modelltestung durchgeführt werden muss, um aussagekräftigere Ergebnisse über die Dimensionalität der Items zu erhalten. Des Weiteren wurde überprüft inwieweit die einzelnen geschätzten Aufgabenparameter in die theoretische Modellvorhersage passen. Im Mittel liegen die Infitwerte (0,982) bzw. die Outfitwerte (1,071) innerhalb der üblichen Grenzwerte und sind somit akzeptabel. Auch die mittlere Trennschärfe (0,304) ist durchaus zufriedenstellend.

Die Überarbeitung der Testaufgaben wurde auf Basis der statistischen Kennwerte aus den Itemanalysen sowie inhaltlicher Überlegungen durchgeführt. Das Ergebnis der Itemrevision zeigt, dass insgesamt vier Items aufgrund einer unzureichenden psychometrischen Eignung ausgeschlossen wurden. Bei den restlichen Items wurden kleinere Änderungen einzelner Antwortalternativen zu Gunsten einer disjunkteren Formulierung vorgenommen. Das Testinstrument in der zweiten Version umfasst schließlich 27 Aufgaben zum Wirtschaftswissen, jeweils 8 Items zum wirtschaftlichen Interesse sowie den Einstellungen zur Wirtschaft, 9 Items zur Selbsteinschätzung und 13 Fragen zu soziodemografischen Merkmalen und zur Mediennutzung.

Eine Einschränkung der Aussagekraft der vorliegenden Ergebnisse ergibt sich insbesondere durch die Anwendung eines dichotomen Rasch-Modells und der damit einhergehenden eingeschränkten Vergleichbarkeit mit den Ergebnissen nach klassischer Testtheorie. Zudem wurde auf die Aufdeckung etwaiger Messinvarianzen nach Subgruppen gänzlich verzichtet. Im Rahmen der bevorstehenden zweiten Testphase wird eine differenziertere Analyse nach verschiedenen soziodemografischen Merkmalen angestrebt, z. B. eine Aufdeckung etwaiger Geschlechtereffekte oder Effekte innerhalb unterschiedlicher Schultypen. Des Weiteren ergeben sich Limitationen aus der online durchgeführten Befragung. Erstens kann die Testumgebung nicht kontrolliert werden und damit, ob die Befragten Hilfsmittel für die Beantwortung der Testaufgaben verwenden (obwohl ausdrücklich darauf hingewiesen wird keine Hilfsmittel zu verwenden). Zweitens stellt sich die Ziehung einer Zufallsstichprobe bei Online-Erhebungen üblicherweise als relativ schwierig heraus. Es kann nicht kontrolliert werden, wer letztendlich den Befragungslink anklickt, daher spielen Faktoren wie z. B. Incentives für eine Teilnahme oder ein gewisses Interesse an wirtschaftlichen Fragen wahrscheinlich eine Rolle. Drittens ist die Rücklauf- bzw. Abbruchquote bei Online-Befragungen üblicherweise problematisch. Im Rahmen des vorliegenden Pretests wurde eine Rücklaufquote von rund 13% erreicht und die Abbruchquote lag bei rund 75%.

Die Entwicklung eines standardisierten Testinstruments nach Vorgaben der wissenschaftlichen Testkonstruktion stellt zweifelsohne ein komplexes und aufwendiges Unterfangen dar. Die Items müssen nicht nur fachdidaktischen Ansprüchen genügen, sondern auch einer umfassenden empirischen Testung Stand halten. Zudem bedarf ein entsprechendes Messinstrument in der Regel vieler Überarbeitungsschritte und mehrerer empirischer Testdurchläufe, die zeitlich sehr aufwendig sind. Mit der überarbeiteten zweiten Version des Testinstruments werden aktuell bereits österreichweit weitere Daten erhoben, die über die in diesem Beitrag dargestellten Verfahren hinaus analysiert werden. Angesichts der Relevanz der Wirtschaftsbildung in

Österreich soll mit einem nach wissenschaftlichen Kriterien bewährtem Testinstrument ein elementarer Beitrag zur Messung des Wirtschaftswissens von jungen Erwachsenen nach Abschluss der Sekundarstufe II geleistet werden.

## Literatur

Anderson, Lorin W./Krathwohl, David R. (2001): A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy. New York.

Beck, Klaus (2000): Wirtschaftskundliches Wissen und Denken – zur Bestimmung und Erfassung ökonomischer Kompetenz. In: Euler, Dieter/Twardy, Martin/Jongebloed, Hans-Carl (Hrsg.): Sozialökonomische Theorie – sozialökonomisches Handeln: Konturen und Perspektiven der Wirtschafts- und Sozialpädagogik. Kiel, 211-229.

Brandlmaier, Elke et al. (2006): Ökonomische Bildung von Schüler/innen Allgemeinbildender Höherer Schulen. Modellentwicklung, Entwicklung eines Messinstruments, ausgewählte Ergebnisse. Wien.

Bühner, Markus (2013): Einführung in die Test- und Fragebogenkonstruktion. München.

Dubs, Rolf (2014): Ökonomische Allgemeinbildung in der Sekundarstufe II. In: Retzmann, Thomas (Hrsg.): Ökonomische Allgemeinbildung in der Sekundarstufe II. Konzepte, Analysen und empirische Befunde, Schwalbach, 13-25.

Früh, Werner (2015): Inhaltsanalyse. Theorie und Praxis. Konstanz u. a.

Greimel-Fuhrmann, Bettina (2018): Ist es (un)möglich, finanziell gebildet zu sein? In: bwp@Spezial AT-1: Wirtschaftspädagogische Forschung und Impulse für die Wirtschaftsdidaktik – Beiträge zum 12. Österreichischen Wirtschaftspädagogikkongress, 1-16. Online: [http://www.bwpat.de/wipaed-at1/greimel-fuhrmann\\_wipaed-at\\_2018.pdf](http://www.bwpat.de/wipaed-at1/greimel-fuhrmann_wipaed-at_2018.pdf) (20.07.2021).

Geise, Wolfgang (2001): Ökonomische Bildung zur Bewältigung von Lebenssituationen. In: Geise, Wolfgang (Hrsg.): Ökonomische Bildung zur Bewältigung von Lebenssituationen. Bergisch Gladbach, III-VII.

Greimel-Fuhrmann, Bettina (2013): "Don't Know Much about Economics and Business." Economic und Financial Literacy als wesentliche, jedoch vernachlässigte Bildungsziele. In: WissenPlus, 32, H. 4, I-VIII.

Itzlinger-Bruneforth, Ursula/Kuhn, Jörg-Tobias /Kiefer, Thomas (2016): Testkonstruktion. In: Breit, Simone/Schreiner, Claudia (Hrsg.): Large-Scale Assessment mit R. Methodische Grundlagen der österreichischen Bildungsstandardüberprüfung. Wien.

Kaminski, Hans/Eggert, Katrin/Burkhard, Karl-Josef (2008): Konzeption für die ökonomische Bildung als Allgemeinbildung von der Primarstufe bis zur Sekundarstufe II. Online: [https://bankenverband.de/media/files/Konzeption\\_fuer\\_die\\_oekonomische\\_Bildung.pdf](https://bankenverband.de/media/files/Konzeption_fuer_die_oekonomische_Bildung.pdf) (10.03.2021).

Krathwohl, David R. (2002): A Revision of Bloom's Taxonomy: An Overview. In: Theory into Practice, 41, H. 4, 212-218.

Kuckartz, Udo (2016): Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung. Weinheim u. a.

Lehmann, Rainer/Seeber, Susan (2009): ULME III. Untersuchung von Leistungen, Motivation und Einstellungen der Schülerinnen und Schüler in den Abschlussklassen der Berufsschulen. Hamburg.

Lienert, Gustav A./Raatz, Ulrich (1998): Testaufbau und Testanalyse. Weinheim.

Loerwald, Dirk/Schnell, Christina (2014): Zum Einsatz von Tests im Wirtschaftsunterricht. Hilfestellungen zur Konstruktion, Auswahl und Verwendung standardisierter Testverfahren. In: Unterricht Wirtschaft + Politik, 4, H. 2, 44-50.

Loerwald, Dirk/Schnell, Christina (2016): Diagnostik im Dilemma zwischen fachdidaktischen Ansprüchen und empirischen Anforderungen. Zur (vermeintlichen) Trivialität von Testitems. In: Zeitschrift für Didaktik der Gesellschaftswissenschaften, 7, H. 1, 57-73.

Media Analyse (2020): Verein Arbeitsgemeinschaft Media - Analysen (VMA). Media Analyse 2020. Online: <https://www.media-analyse.at/table/3527> (05.05.2021).

Merten, Klaus (1995): Inhaltsanalyse. Einführung in Theorie, Methode und Praxis. Opladen.

Moosbrugger, Helfried/Kelava, Augustin (2012a): Qualitätsanforderungen an einen psychologischen Test (Testgütekriterien). In: Moosbrugger, Helfried/Kelava, Augustin (Hrsg.): Testtheorie und Fragebogenkonstruktion. Berlin, 7-26.

Moosbrugger, Helfried/Schermelleh-Engel, Karin/Gäde, Jana C./Kelava, Augustin (2012b): Testtheorien im Überblick. In: Moosbrugger, Helfried/Kelava, Augustin (Hrsg.): Testtheorie und Fragebogenkonstruktion. Berlin, 251-273.

OECD (2017): PISA 2015 Results (Volume IV): Students' Financial Literacy, PISA.

Oberrauch, Luis (2019): Ökonomische Kompetenzen von Heranwachsenden: Entwicklung und Validierung eines Testinstruments. In: Zeitschrift für ökonomische Bildung, 2019/08, 60-98.

Rumpold, Herwig (2018): Ökonomisches Wissen in der Sekundarstufe I: Entwicklung eines Testinstruments für Schüler/innen der achten Schulstufe. Wirtschaftsuniversität Wien.

Schnell, Christina (2016): „Lautes Denken“ als qualitative Methode zur Untersuchung der Validität von Testitems. Erkenntnisse einer Studie zur Diagnose des ökonomischen Fachwissens von Schülerinnen und Schülern der Sekundarstufe I. In: Zeitschrift für ökonomische Bildung, 5, 26-49.

Schuldnerberatung (2021): Schuldenreport 2021. Online: <https://www.schuldnerberatung.at/fachpublikum/schuldenreport.php> (12.05.2021).

Schumann, Stephan/Eberle, Franz/Oepke, Maren (2010): Inhaltsauswahl für den Test zur Erfassung ökonomischen Wissens und Könnens im Projekt „Ökonomische Kompetenzen von Maturandinnen und Maturanden“ (OEKOMA). Online: [http://www.ife.uzh.ch/dam/jcr:00000000-5b32-1166-0000-000009b14c87/Bericht\\_Inhaltsauswahl\\_OEKOMA.pdf](http://www.ife.uzh.ch/dam/jcr:00000000-5b32-1166-0000-000009b14c87/Bericht_Inhaltsauswahl_OEKOMA.pdf) (13.05.2021).



Schumann, Stephan/Eberle, Franz (2011): Bedeutung und Verwendung schwierigkeits-bestimmender Aufgabenmerkmale für die Erfassung ökonomischer und beruflicher Kompetenzen. In Fasshauer, Uwe/Fürstenau, Bärbel/Wuttke, Eveline (Hrsg.): Grundlagenforschung zum Dualen System und Kompetenzentwicklung in der Lehrerbildung. Opladen, 77-89.

Schumann, Stephan/Eberle, Franz (2014): Ökonomische Kompetenzen von Lernenden am Ende der Sekundarstufe II. In: Zeitschrift für Erziehungswissenschaft, 17/1, 103-126.

Seeber, Günther/Retzmann, Thomas/Remmele, Bernd (2012): Bildungsstandards der ökonomischen Allgemeinbildung. Kompetenzmodell - Aufgaben - Handlungsempfehlungen. Schwalbach.

Silgoner, Maria/Greimel-Fuhrmann, Bettina/Weber, Rosa (2015): Financial Literacy Gaps of the Austrian Population. In: Monetary Policy & the Economy, 2, 35-51.

Statistik Austria (2020): Bildung in Zahlen 2018/19. Schlüsselindikatoren und Analysen. Online:

[http://www.statistik.at/web\\_de/services/publikationen/5/index.html?includePage=detailedView&sectionName=Bildung&pubId=461](http://www.statistik.at/web_de/services/publikationen/5/index.html?includePage=detailedView&sectionName=Bildung&pubId=461) (13.05.2021).

Walstad, William B./Rebeck, Ken/Butters, Roger (2013): Test of Economic Literacy. Fourth Edition. Examiner's Manual. New York.

Weinert, Franz E. (1999): Konzepte der Kompetenz. Paris.

Witzel, Andreas/Reiter, Herwig (2012): The Problem-centered Interview. Principles and Practice. London.

Würth, Reinhold/Klein, Hans-Joachim (2001): Wirtschaftswissen Jugendlicher in Baden-Württemberg. Eine empirische Analyse. Künzelsau.

## Zitieren dieses Beitrags

---

Sarwari, S. (2021): Wirtschaftswissen junger Erwachsener in Österreich. Entwicklung eines Testinstruments zu Themen in den Wirtschaftsberichten von Tageszeitungen. In: *bwp@ Spezial AT-3*: Beiträge zum 14. Österreichischen Wirtschaftspädagogik-Kongress, 1-20. Online: [http://www.bwpat.de/wipaed-at3/sarwari\\_wipaed-at\\_2021.pdf](http://www.bwpat.de/wipaed-at3/sarwari_wipaed-at_2021.pdf) (13.09.2021).

## Die Autorin

---



### SHIREEN SARWARI MA MSc

Wirtschaftsuniversität Wien, Institut für Wirtschaftspädagogik  
Welthandelsplatz 2, 1020 Wien

[Shireen.Sarwari@wu.ac.at](mailto:Shireen.Sarwari@wu.ac.at)

<https://www.wu.ac.at/wipaed>