

Wirtschaftswissen in der Sekundarstufe I

Entwicklung eines Erhebungsinstruments für die Zielgruppe von Schüler/inne/n der achten Schulstufe

Herwig Rumpold, Bettina Greimel-Fuhrmann

Institut für Wirtschaftspädagogik, Wirtschaftsuniversität Wien

Zusammenfassung

Während für die Sekundarstufe II zahlreiche erprobte Instrumente zur standardisierten Messung des Wirtschaftswissens existieren, stehen solche für die Sekundarstufe I im deutschsprachigen Raum bislang kaum zur Verfügung. Vor diesem Hintergrund besteht die Zielsetzung des in diesem Beitrag dargestellten Forschungsvorhabens darin, ein Erhebungsinstrument zur standardisierten Messung des Wirtschaftswissens von Schüler/inne/n am Ende der Sekundarstufe I zu entwickeln, das wissenschaftlichen Kriterien entspricht. Die Instrumententwicklung umfasst dabei die theoretisch-konzeptionelle Modellierung des Messgegenstands, die modellgeleitete Erstellung des Erhebungsinstruments sowie dessen empirische Erprobung hinsichtlich der inhaltlichen Qualität und der Erfüllung psychometrischer Anforderungen. In diesem Beitrag werden die methodische Vorgehensweise der Instrumententwicklung beschrieben und die empirischen Ergebnisse aus der Erprobung der ersten Version des Erhebungsinstruments präsentiert. Die Ergebnisse der Itemanalysen sprechen sowohl nach der klassischen Testtheorie als auch nach probabilistischen Verfahren weitgehend für eine aussichtsreiche psychometrische Güte des Instruments.

Abstract

While a number of instruments for standardized assessments of economic knowledge have been developed and scientifically tested for students at the higher secondary level, there is a remarkable lack of such instruments for students at the lower secondary level, particularly within German-speaking countries. Therefore, it is the main aim of the study presented in this paper to develop and test an instrument for eighth-grade students that allows for a standardized assessment of their economic knowledge. The test development process involves the systematic conceptualization of a theoretical framework for the testing object, the model-based generation of the instrument and its items as well as its empirical evaluation. This paper gives an overview of the methodical design of the test development and presents the results of the empirical evaluation of the first version of the instrument. The results largely indicate the content-related adequacy of the test as well as its satisfying psychometric quality according to both classical test theory and item response theory.

1 Problemhintergrund und Zielsetzung

Wirtschaftliche Sachverhalte und Zusammenhänge sind für alle Menschen von zentraler Bedeutung. Bereits im Jugendalter stellt sich eine Vielzahl an ökonomisch geprägten Lebenssituationen, seien es Kauf-, Spar- oder Vorsorgeentscheidungen, Fragen der Berufswahl oder die Ausübung des Wahlrechts. Im Hinblick auf diese vielfältigen Lebenssituationen sind zweifellos planvolle und reflektierte Entscheidungen auf Basis einer fundierten ökonomischen Bildung desiderabel. Ohne ökonomische Bildung sind aktuelle ökonomische, politische wie auch gesellschaftliche Entwicklungen und Herausforderungen, beispielsweise die derzeit viel diskutierte Frage um den Verbleib einzelner Staaten in der Europäischen Union, kaum zu verstehen und daher auch nicht in ihren Auswirkungen nachzuvollziehen (vgl. Aff 2015; Davies 2001; Stern 2002). Ökonomische Bildung stellt demnach einen notwendigen Bestandteil der Allgemeinbildung dar, indem sie über die Förderung ökonomischer Kompetenzen zur Bewältigung ökonomisch geprägter Lebenssituationen beiträgt (vgl. Albers 1988; DeGÖB 2005; Kaminski 2014).

Eine ökonomisch allgemein gebildete Gesellschaft kann indes nicht als selbstverständlich angesehen werden, tatsächlich offenbaren empirische Studien im deutschsprachigen Raum teilweise erhebliche Defizite in den ökonomischen Kompetenzen quer über unterschiedliche Altersgruppen, Bildungsniveaus und soziodemografische Schichten hinweg (vgl. Hartmann/Pühringer 2000; Hummelsheim 2010). Vor allem zum Stand der ökonomischen Kompetenzen von Schüler/inne/n am Ende der Sekundarstufe II liegen ausgiebige empirische Befunde vor, welche auf stellenweise gravierende Bildungsdefizite hindeuten (vgl. Greimel-Fuhrmann 2013; Leinert 2004; Leinert/Wagner 2004).

Das dabei insbesondere in internationalen Studien zumeist eingesetzte Erhebungsinstrument ist der „Test of Economic Literacy“ (TEL), ein standardisierter und normierter Test zur Erfassung des Wirtschaftswissens auf dem Niveau der Reife- und Diplomprüfung (vgl. Walstad et al. 2013). Im deutschsprachigen Raum besteht mit dem „wirtschaftskundlichen Bildungstest“ (WBT) eine von Beck et al. (1998) wissenschaftlich übersetzte Version des Tests, die ebenfalls breite Anwendung findet (vgl. etwa Förster et al. 2012; Katschnig/Hanisch 2005; Müller et al. 2007). Daneben existiert eine Reihe von Studien, in denen über das Wissen hinaus noch weitere Facetten ökonomischer Kompetenzen berücksichtigt sind. Beispielhaft ist etwa die von Würth/Klein (2001) durchgeführte Studie zum Wirtschaftswissen Jugendlicher in Baden-Württemberg, in der auch ökonomische Einstellungen und Interessen untersucht wurden. Vergleichbare Untersuchungen für Österreich liegen unter anderem von

Brandlmaier et al. (2006) und Schmid (2006) vor. In jüngeren Studien eingesetzte Erhebungsinstrumente basieren zumeist auf einem umfassenderen Kompetenzverständnis und versuchen, ökonomische Kompetenzen in ihren Inhalts- und Verhaltensdimensionen präzise zu modellieren. So erfolgt etwa im Projekt „Ökonomische Kompetenzen von Maturandinnen und Maturanden“ (OEKOMA) von Schumann/Eberle (2014) eine Unterscheidung des Wirtschaftswissens in die Inhaltskategorien Volkswirtschaftslehre, Betriebswirtschaftslehre und Rechnungswesen mit entsprechenden Unterkategorien. Macha/Schuhen (2011) nehmen im „Siegener Kompetenzmodell zur ökonomischen Kompetenz“ dagegen eine Differenzierung von verbalen und mathematischen sowie verstehensbasierten und handlungsbasierten ökonomischen Kompetenzen vor.

Aus wissenschaftlicher Perspektive wird hinsichtlich der Messung ökonomischer Kompetenzen jedoch nach wie vor offener Forschungsbedarf konstatiert, der sich sowohl auf inhaltliche und methodische Einschränkungen der eingesetzten Erhebungsinstrumente als auch auf Lücken in der Datenlage bezieht (vgl. Gnan et al. 2007; Retzmann/Frühauf 2014; Schumann/Eberle 2015). Insbesondere Befunde zu ökonomischen Kompetenzen von Schüler/innen in der Sekundarstufe I erweisen sich noch als stark unterrepräsentiert, sodass hier eine der wesentlichen Forschungslücken geortet werden kann (vgl. Greimel-Fuhrmann et al. 2016). Während für die Sekundarstufe II zahlreiche erprobte Instrumente zur standardisierten Messung ökonomischer Kompetenzen vorliegen, stehen solche für die Sekundarstufe I zumindest im deutschsprachigen Raum bislang kaum zur Verfügung. Dies zeigt sich beispielsweise darin, dass der „Test of Economic Knowledge“ (TEK) (vgl. Walstad et al. 2010), das englischsprachige Pendant zum TEL für den mittleren Schulabschluss, noch keine wissenschaftliche Übersetzung für den Einsatz im deutschsprachigen Raum erfahren hat. Mit Ausnahme des „Tests für Wirtschaftswissen in der Sekundarstufe I“ (TWWS-1) (vgl. Nagy et al. 2008) und einem laufenden Projekt zur Entwicklung eines Wirtschaftstests für die achte Schulstufe am Institut für Ökonomische Bildung in Oldenburg (vgl. Loerwald/Schnell 2014) sind den Autor/inn/en zudem keine Initiativen zur originären Instrumententwicklung im deutschsprachigen Raum bekannt.

Die ökonomischen Kompetenzen von Schüler/innen am Ende der Sekundarstufe I sind jedoch nicht nur aus wissenschaftlicher Perspektive von großem Interesse, sodass deren Vernachlässigung in empirischen Studien nur wenig nachvollziehbar ist. Zum einen befinden sich diese Schüler/innen in einem Alter, in dem sie über zunehmende ökonomische Ressourcen verfügen und ihre ökonomischen Entscheidungen auch immer häufiger selbst

treffen, sei es die freie Verfügung über ihnen überlassenes Taschengeld, das erste Jugendkonto bei der Bank oder die baldige Wahlberechtigung im Alter von 16 Jahren. Zum anderen steht für die Schüler/innen am Ende der Sekundarstufe I die Entscheidung über den weiteren Ausbildungsweg bevor, der vielfältige Möglichkeiten bietet und starke Implikationen für die spätere Berufstätigkeit aufweist (vgl. Geise 2001; May 2011). Es erscheint durchaus plausibel, dass das Wirtschaftswissen der Schüler/innen die genannten Entscheidungen mitbestimmt, womit dessen Messung bei dieser Zielgruppe aus wissenschaftlicher Perspektive einen zusätzlichen Erklärungsbeitrag für deren Entscheidungsverhalten erwarten lässt. Zudem könnte der Einsatz eines wissenschaftlich fundierten und erprobten Erhebungsinstrumentes die Kontrolle von Lernfortschritten und davon ausgehend die Steuerung individueller Lernprozesse sowie die Evaluation des Wirtschaftsunterrichts einschließlich der Generierung curricularer Impulse für die ökonomische Bildung in der Sekundarstufe I unterstützen (vgl. Ingenkamp/Lissman 2005; Seeber et al. 2010).

Vor diesem Hintergrund besteht die Zielsetzung des in diesem Beitrag dargestellten Forschungsvorhabens darin, ein Erhebungsinstrument zur standardisierten Messung des Wirtschaftswissens von Schüler/innen am Ende der Sekundarstufe I zu entwickeln, das wissenschaftlichen Standards entspricht. Die Zielgruppe setzt sich im österreichischen Bildungssystem dabei konkret aus den Schüler/innen der achten Schulstufe sowohl der Allgemeinbildenden Höheren Schule als auch der Neuen Mittelschule zusammen. Die Instrumententwicklung schließt die theoretisch-konzeptionelle Modellierung des Messgegenstands, die modellgeleitete Erstellung des Erhebungsinstrumentes sowie dessen empirische Erprobung hinsichtlich der inhaltlichen Angemessenheit und der Erfüllung psychometrischer Anforderungen mit ein. Die erste Version des Erhebungsinstrumentes wurde im Frühjahr 2015 erstellt und erprobt. In diesem Beitrag werden die methodische Vorgehensweise der Instrumententwicklung beschrieben und die empirischen Ergebnisse aus der Erprobung der ersten Instrumentversion dargelegt. Der Beitrag schließt mit einer vorläufigen Beurteilung der Güte dieser ersten Version und einem Ausblick auf die weiteren Entwicklungsschritte.

2 Methodische Vorgehensweise der Instrumententwicklung

2.1 Modellierung des Messgegenstands

2.1.1 Wirtschaftswissen als zentrale Facette ökonomischer Kompetenzen

Die Vielfalt an konzeptionellen Auseinandersetzungen und bestehenden Forschungszugängen erfordert die Eingrenzung und Konkretisierung jener Facetten ökonomischer Kompeten-

zen, die als Bezugsrahmen für die Inhaltsbestimmung des Erhebungsinstruments, die Festlegung der Aufgabenformate und die Generierung der Testitems dienen. Diesbezüglich wird dem Kompetenzverständnis von Weinert (2001) gefolgt, der als Kompetenzen jene Fähigkeiten und Fertigkeiten, motivationalen, volitionalen und sozialen Bereitschaften versteht, die zur Problemlösung in variablen Anforderungssituationen einer bestimmten Domäne erforderlich sind. In der Domäne der ökonomischen Bildung bestehen diese Anforderungssituationen in den verschiedenen ökonomisch geprägten Lebenssituationen, zu deren erfolgreicher Bewältigung die ökonomische Bildung als Zielvorstellung befähigen soll (vgl. Albers 1995; Kaminski et al. 2008).

Die Messung von ökonomischen Kompetenzen nach diesem Verständnis stellt eine hoch anspruchsvolle und komplexe Aufgabe dar, die sowohl didaktischen als auch psychometrischen Anforderungen zu entsprechen hat. Während aus didaktischer Perspektive das zu messende Konstrukt ökonomischer Kompetenzen in seinen Facetten möglichst umfassend abgebildet werden soll, sollen die eingesetzten Erhebungsinstrumente aus psychometrischen Gesichtspunkten den Gütekriterien der Objektivität, Reliabilität und Validität entsprechen. Es ist zu hinterfragen, inwieweit es möglich ist, eine so komplexe Zielvorstellung wie die erfolgreiche Bewältigung ökonomisch geprägter Lebenssituationen in variablen Handlungszusammenhängen mit standardisierten Instrumenten psychometrisch abgesichert zu messen (vgl. Greimel-Fuhrmann 2014; Loerwald/Schnell 2016). Aus diesem Grund wird mit dem in diesem Beitrag dargestellten Forschungsvorhaben auch nicht das Ziel verfolgt, ökonomische Kompetenzen in ihrer Gesamtheit zu erfassen, sondern eine Einschränkung auf das Wirtschaftswissen vorgenommen, welches als zentrale Facette ökonomischer Kompetenzen umfassend konzeptualisiert und weithin anerkannt ist (vgl. Beck 1989; Schumann/Eberle 2014). Zudem stehen zur Messung von Wissen geeignete diagnostische Verfahren aus der pädagogischen Psychologie zur Verfügung, sodass die Möglichkeit zur psychometrischen Absicherung des entwickelten Erhebungsinstruments gegeben ist.

Die pädagogische Psychologie meint mit dem Begriff Wissen die kognitive Repräsentation der Umwelt einer Person. Das Wissen äußert sich demnach in Kenntnissen der Person über die Bereiche, Objekte, Verfahrensweisen, Werte und Normen ihrer Umwelt (vgl. Kluge 1979). Gängig ist die grundlegende Unterscheidung in deklaratives und prozedurales Wissen, wobei ersteres jene Kenntnisse umfasst, die Gegenstand informationsverarbeitender Prozesse sein können, und letzteres die Kenntnisse von Prozeduren zur gezielten Informationsverarbeitung beinhaltet (vgl. Eysenck/Keane 2005; Fortmüller 1997). Während deklarati-

ves Wissen explizit dargelegt werden kann, zeigt sich prozedurales Wissen meist implizit in der erbrachten kognitiven Leistung. Bei der Problemlösung werden schließlich Informationen aus dem deklarativen Wissen mit Informationsverarbeitungsprozeduren aus dem prozeduralen Wissen verknüpft und zur Lösung einer Problemsituation angewandt (vgl. Dörner 1987; Fortmüller 1997).

An dieser grundlegenden Unterscheidung knüpfen die meisten pädagogischen Ansätze zur Bestimmung unterschiedlicher Wissensarten bei der Festlegung von Lernzielen an. Anderson/Krathwohl (2001) differenzieren hier näher zwischen faktischem Wissen, konzeptuellem Wissen, prozeduralem Wissen und meta-kognitivem Wissen. Das kognitive Anspruchsniveau innerhalb einer Wissensart wiederum wird von Krathwohl (2002) in einer Überarbeitung von Blooms Lernzieltaxonomie im kognitiven Bereich (vgl. Bloom et al. 1972) nach steigendem Anspruch in die sechs Niveaustufen Erinnern, Verstehen, Anwenden, Analysieren, Evaluieren und Erschaffen unterteilt. In dem in diesem Beitrag dargestellten Forschungsvorhaben erfolgt eine Eingrenzung auf faktisches und konzeptuelles Wissen, wobei im Hinblick auf das kognitive Anspruchsniveau vorrangig die Niveaustufen Verstehen und Anwenden angezielt werden. Diese Eingrenzung erscheint sinnvoll, da sich übergeordnete Wissensarten und Niveaustufen erst in realitätsgetreuen, authentischen Problemsituationen zeigen und sich daher mit einem standardisierten Erhebungsinstrument nur schwer angemessen erfassen lassen. Zudem scheint bei der Voraussetzung höherer Wissensniveaus die Angemessenheit des Erhebungsinstruments für die Zielgruppe von Schüler/inne/n der achten Schulstufe gefährdet.

2.1.2 Inhaltsbereiche des Wirtschaftswissens

Bei der Bestimmung der Inhaltsbereiche des Wirtschaftswissens im Rahmen des in diesem Beitrag dargestellten Forschungsvorhabens wird versucht, sowohl dem Lehrplan für die Sekundarstufe I zu entsprechen als auch bestehenden Konzeptionalisierungen ökonomischer Bildung im deutschsprachigen Raum gerecht zu werden. Die wirtschaftlichen Inhalte des in Österreich geltenden Lehrplans für die Sekundarstufe I fokussieren dabei sehr stark auf volkswirtschaftliche Grundzusammenhänge, rein betriebswirtschaftliche Überlegungen werden dagegen erst in der Sekundarstufe II aufgegriffen und bei der Inhaltsbestimmung für das Erhebungsinstrument folglich nicht näher berücksichtigt (vgl. BMBF 2015). In bestehenden Konzeptionalisierungen ökonomischer Bildung sind zur inhaltlichen Strukturierung des Wirtschaftswissens unterschiedliche Ansatzpunkte mit spezifischen Perspektiven auf die Domäne auszumachen. Während etwa bei kategorialen Ansätzen zentrale Kategorien des wirt-

schaftlichen Handelns wie Knappheit, Bedürfnisgetriebenheit, Koordination und Allokation im Vordergrund stehen, stellen institutionale Ansätze auf die verschiedenen Wirtschaftssubjekte und deren Zusammenwirken in ihren wirtschaftlichen Aktivitäten ab. Situationale Ansätze hingegen knüpfen an konkreten ökonomischen Lebenssituationen an, in denen die Individuen unterschiedliche ökonomische Rollen einnehmen (vgl. Hedtke 2011; Weber 2005). Zur Inhaltsbestimmung des Erhebungsinstruments werden ausgewählte Aspekte dieser Ansätze integriert, indem wirtschaftliches Handeln auf unterschiedlichen Abstraktionsebenen beleuchtet wird. Mit den hier dargestellten Inhaltsbereichen wird nicht der Anspruch erhoben, das Wirtschaftswissen vollständig und in exakt operationalisierte Bestandteile zerlegt abzubilden. Vielmehr sollen sie eine hinreichende Grundlage zur inhaltlichen Legitimation der verwendeten Items, Skalen und Indikatoren des Erhebungsinstruments liefern.

Die oberste Abstraktionsebene beschreibt die grundlegende Funktion des wirtschaftlichen Handelns, die jeglicher wirtschaftlichen Aktivität immanent ist. Sie ergibt sich aus der Diskrepanz zwischen dem Überfluss unbefriedigter menschlicher Bedürfnisse und der Knappheit der zur Bedürfnisbefriedigung zur Verfügung stehenden Ressourcen. Diese schafft die Notwendigkeit der Koordination und Allokation der gegebenen Ressourcen und führte letztlich von der Selbstversorgung über den Warentausch zu den ausdifferenzierten arbeitsteiligen Strukturen moderner Gesellschaften.

Wirtschaftliche Aktivitäten sind dabei stets in den Kontext der geltenden Wirtschaftsordnung eingebettet, welche die strukturellen Rahmenbedingungen des wirtschaftlichen Handelns vorgibt. Bestehende Ordnungsform ist hierbei jene der sozialen Marktwirtschaft, welche durch die Ordnungselemente der dezentralen Planung, des Privateigentums, des Gewinnprinzips und der Marktpreisbildung bei staatlicher Umverteilung gekennzeichnet ist (vgl. Kaminski 1996). Das wirtschaftstheoretische Fundament dafür bildet ein klassisches bzw. neoklassisches Wirtschaftsverständnis, da dieses die gegebene Wirtschaftsordnung sowie die wirtschaftspolitische Diskussion am besten abbildet. Wenngleich die klassische Wirtschaftstheorie das wirtschaftliche Handeln der Wirtschaftssubjekte sicherlich nicht restlos erklären kann und daher aus didaktischen Überlegungen weitere Ansätze einbezogen werden könnten (vgl. Kaminski 1997; Karpe/Krol 1999), stellt sie die naheliegende wirtschaftstheoretische Grundpositionierung dar.

Im Rahmen der Wirtschaftsordnung finden wirtschaftliche Aktivitäten in Form von Austauschprozessen der Wirtschaftssubjekte mit wechselseitigen Leistungsbeziehungen statt. Als Wirtschaftssubjekte treten hierbei die Haushalte, die Unternehmen, der Staat und das

Ausland auf, ihre wirtschaftlichen Aktivitäten bestehen im Güterkreislauf aus Produktion, Distribution und Konsumtion sowie im Geldkreislauf aus Einkommensentstehung und Einkommensverwendung (vgl. Kaminski 1996). Der Güterkreislauf und der Geldkreislauf greifen insofern ineinander, als die Produktion und die Distribution mit der Einkommensentstehung und die Konsumtion mit der Einkommensverwendung einhergehen (vgl. Steinmann 1997). Diese wirtschaftlichen Zusammenhänge ergeben in ihrer Gesamtheit den Wirtschaftskreislauf, der über die zentrale Ergebnisgröße der gesamtwirtschaftlichen Wertschöpfung einen geeigneten Bezugspunkt bildet.

Innerhalb des Wirtschaftskreislaufs nimmt jedes Wirtschaftssubjekt als Individuum in seinen ökonomisch geprägten Lebenssituationen unterschiedliche Rollen mit spezifischen Handlungszusammenhängen ein. Ausgehend von den maßgeblichen Rollen der Verbraucher/innen, der Erwerbstätigen und der Wirtschaftsbürger/innen können noch stärkere Ausdifferenzierungen vorgenommen werden; so lässt sich die Rolle der Verbraucher/innen etwa in jene der Konsument/inn/en, der Sparer/innen, der Kreditnehmer/innen oder der Versicherungsnehmer/innen mit jeweils korrespondierenden Handlungszusammenhängen untergliedern. Eine solche Differenzierung erscheint sinnvoll, da in diesen Rollen auf unterschiedlichen Märkten und Marktseiten agiert wird, was gänzlich andere Perspektiven erfordert (vgl. Dubs 2014; Retzmann 2005; Seeber et al. 2012). Auf diesem Weg ermöglicht das Rollenkonzept die Strukturierung ökonomisch geprägter Lebenssituationen und deren Einbettung in die Gesamtzusammenhänge des Wirtschaftskreislaufs auf der untersten Abstraktionsebene des wirtschaftlichen Handelns.

Diese vier Abstraktionsebenen von der grundlegenden Wirtschaftsfunktion über die Wirtschaftsordnung und den Wirtschaftskreislauf hin zu konkreten Wirtschaftssituationen schaffen den Bezugsrahmen für die Inhaltsbestimmung des Erhebungsinstruments. Eine trennscharfe Abgrenzung dieser Inhaltsbereiche wird dabei nicht angestrebt und erscheint auch nicht zielführend, da sich in konkreten ökonomischen Lebenssituationen stets auch wirtschaftliche Kreislaufzusammenhänge sowie elementare Kennzeichen der Wirtschaftsordnung und der Wirtschaftsfunktion widerspiegeln und somit die unteren Abstraktionsebenen die übergeordneten Abstraktionsebenen miteinschließen. Zusammengefasst besteht der Messgegenstand des Erhebungsinstruments damit in faktischem und konzeptuellem Wissen über die verschiedenen Abstraktionsebenen wirtschaftlichen Handelns auf den kognitiven Anspruchsniveaus Verstehen und Anwenden. Der Fokus liegt auf Kenntnissen über die zentralen Bestimmungsmerkmale der Wirtschaftsordnung, auf dem Verständnis modellhafter

wirtschaftlicher Ursache-Wirkungs-Zusammenhänge und deren Anwendung auf wirtschaftliche Sachverhalte in der Gestalt ökonomischer Lebenssituationen.

2.2 Erstellung des Erhebungsinstruments

2.2.1 Bestimmung des Test- und Aufgabenformats

Die Zielsetzung des Erhebungsinstruments der standardisierten Messung von Wirtschaftswissen erforderte die Konstruktion eines eindimensionalen Instruments sowie die Wahl eines geeigneten Test- und Aufgabenformats. Die Wissensüberprüfung entspricht dabei dem Testformat eines Leistungstests, in dem über die Beantwortung von Wissensaufgaben das bestmögliche Leistungsverhalten der Testpersonen im Hinblick auf das angezielte Konstrukt des Wirtschaftswissens gefordert und einer standardisierten Beurteilung unterzogen wird. Der Leistungstest wurde als Kombination aus Power- und Speedtest konzipiert, welche durch Testaufgaben unterschiedlichen Schwierigkeitsgrads gekennzeichnet ist, die in einer begrenzten Zeitvorgabe zu bearbeiten sind (vgl. Moosbrugger/Kelava 2012; Schmidt-Atzert/Amelang 2012). Das Erhebungsinstrument wurde auf eine Bearbeitungsdauer von maximal 45 Minuten ausgerichtet, um den Einsatz im Rahmen einer Unterrichtsstunde zu ermöglichen. Eine längere Bearbeitungsdauer erscheint für die Zielgruppe ohnehin nicht angemessen. Als Erhebungsform wurde die Paper-and-Pencil-Form gewählt, da eine computergestützte Erhebung nicht von jeder Schule administriert werden könnte.

Zur objektiven und standardisierten Erfassung des Wirtschaftswissens wurde das geschlossene Aufgabenformat bestimmt. Zwar haben geschlossene Aufgaben den Nachteil, dass die Testpersonen nicht in ihrem eigenen Referenzsystem antworten können und auf die vorgegebenen Antwortmöglichkeiten beschränkt sind, womit höhere kognitive Anspruchsniveaus nur schwer zu erreichen sind. Der Messgegenstand des konzeptionellen Verständnisses und der modellhaften Anwendung von faktischem und konzeptuellem Wirtschaftswissen ist mit geschlossenen Aufgaben aber durchaus erfassbar (vgl. Ingenkamp/Lissman 2005). Zudem ermöglichen geschlossene Aufgaben eine einheitliche und vergleichbare Messung und eine aus forschungsökonomischen Gründen wesentlich effizientere Durchführung und Auswertung der Erhebung (vgl. Schnell et al. 2011). Die Wissensaufgaben wurden im Multiple-Choice-Format mit vier Antwortalternativen und mehreren möglichen korrekten Antworten entwickelt. Im Vergleich zu Single-Choice-Aufgaben, bei denen stets nur eine Antwortalternative korrekt ist, sind diese Aufgaben zwar schwieriger zu konstruieren und statistisch möglicherweise weniger trennscharf, sie reduzieren jedoch die Ratergefahr erheblich und

erlauben es, bedeutend mehr Gedankengänge der Testpersonen zu erfassen, da jede Antwortalternative für sich als richtig oder falsch markiert werden kann (vgl. Bühner 2013; Lienert/Raatz 1994). Damit ist dieses Aufgabenformat jenem von reinen Single-Choice-Aufgaben vor allem aus didaktischen Überlegungen heraus eindeutig überlegen.

2.2.2 Festlegung des Inhaltsniveaus

Da für die Zielgruppe der Schüler/innen am Ende der Sekundarstufe I noch kaum umfangreiche Studien zu deren Wirtschaftswissen vorliegen, war zur Festlegung des inhaltlichen Niveaus und der Komplexität des Erhebungsinstruments eine explorative Vorstudie erforderlich. Nachdem sowohl der Lehrplan als auch gängige Schulbücher für die Sekundarstufe I wirtschaftliche Inhalte größtenteils fragmentarisch behandeln (vgl. Jagsch 2015), wurde zu diesem Zweck über die Analyse der wesentlichen Lehrplan- und Schulbuchinhalte hinaus eine mündliche Befragung von Schüler/innen zur Erfassung ihrer individuellen Vorstellungen von Wirtschaft durchgeführt.

Diese explorative Befragung erfolgte in Form von leitfadengestützten problemzentrierten Interviews (vgl. Witzel 2000; Witzel/Reiter 2012) zu den festgelegten Inhaltsbereichen des Wirtschaftswissens. Den Beginn der Interviews markierten dem idealtypischen Ablaufschema des problemzentrierten Interviews folgend sehr offen gehaltene, erzählgenerierende Einstiegsfragen, zu denen die Schüler/innen völlig unvoreingenommen ihre spontanen Gedanken zum Thema Wirtschaft und der Funktion wirtschaftlichen Handelns mitteilen sollten. Die anschließenden Sondierungsfragen behandelten vor allem die Rollen der unterschiedlichen Akteur/inn/e/n des Wirtschaftskreislaufs, insbesondere der privaten Haushalte, der Unternehmen und des Staates. Bei allen Fragen wurde versucht, durchgehend an die Lebenswelt der Schüler/innen anzuknüpfen, indem die verschiedenen Inhaltsbereiche vornehmlich anhand ökonomisch geprägter Alltagssituationen thematisiert wurden, mit denen die Schüler/innen aus eigener Erfahrung bereits vertraut sein sollten.

Für die explorative Befragung konnten insgesamt 41 Schüler/innen aus sieben Schulen in drei Bundesländern gewonnen werden. Die Interviews waren auf eine Dauer von maximal 20 Minuten angelegt, was einen der Zielgruppe angemessenen Zeitrahmen bildete. Sämtliche Interviews wurden aufgezeichnet und ins Schriftdeutsche geglättet transkribiert. Die Interviewanalyse erfolgte mittels zusammenfassender Inhaltsanalyse nach Mayring (2010), bei der aus den Äußerungen der Schüler/innen durch induktive Kategorienbildung typische Vorstellungskomplexe zum Thema Wirtschaft, der Funktion wirtschaftlichen Handelns und den

Rollen der Wirtschaftsakteur/inn/e/n generiert wurden. Die auf diesem Weg identifizierten Interpretationen und Fehlinterpretationen wirtschaftlicher Sachverhalte und Zusammenhänge lieferten nicht nur einen Indikator für die Bestimmung des inhaltlichen Niveaus und des Komplexitätsgrads des Erhebungsinstruments, sie boten darüber hinaus auch konkrete Anhaltspunkte für die Generierung der Testitems. So konnten etwa auffällige Fehlinterpretationen zur Erstellung plausibler Distraktoren herangezogen werden.

2.2.3 Generierung der Testitems

Nach der Bestimmung des Aufgabenformats und der Festlegung des Inhaltsniveaus erfolgte die Generierung der Items des Erhebungsinstruments, bei der sowohl top-down auf Grundlage der konzeptionellen Auseinandersetzung mit dem Messgegenstand als auch bottom-up basierend auf empirischen Befunden und bestehenden Erhebungsinstrumenten zum Wirtschaftswissen vorgegangen wurde (vgl. Bühner 2013). Als konzeptionelles Fundament wurde dabei die eingangs erfolgte Modellierung des Messgegenstands herangezogen, die über die theoretische Auseinandersetzung mit dem Konstrukt des Wirtschaftswissens führte. Ergänzende empirische Befunde ergaben sich aus der eigenen explorativen Befragung sowie aus der Analyse anschlussfähiger bestehender Studien und Erhebungsinstrumente zum Wirtschaftswissen von Schüler/inne/n.

Von den analysierten Instrumenten sind jedoch nur wenige für die Zielgruppe von Schüler/inne/n der achten Schulstufe konzipiert, die überwiegende Mehrheit der Instrumente zielt auf die Stufe der Reife- und Diplomprüfung ab und ist damit hinsichtlich der behandelten Inhaltsbereiche als auch des Komplexitätsgrads auf einem höheren Niveau angesiedelt. Zudem setzen beinahe alle analysierten Instrumente auf das Single-Choice-Aufgabenformat, von dem in dem in diesem Beitrag dargestellten Forschungsvorhaben aus didaktischen Gründen Abstand genommen wird. Als Anhaltspunkt für die Generierung der eigenen Testitems waren die Items der analysierten Instrumente demnach nur im Falle inhaltlicher Angemessenheit und gegebenenfalls einer entsprechenden Komplexitätsreduktion sowie einer Adaption in das Multiple-Choice-Aufgabenformat geeignet. Aus diesem Grund wurde der Großteil der Testitems schließlich selbst entwickelt.

Aus didaktischer Perspektive wurde bei der Itemgenerierung für das Erhebungsinstrument darauf geachtet, dass trotz notwendiger inhaltlicher Begrenzungen und Vereinfachungen im Sinne einer didaktischen Reduktion stets die fachliche Korrektheit der Items gewahrt bleibt. Missverständliche oder gar irreführende inhaltliche Verkürzungen schmälern die Aussage-

kraft der Ergebnisse und sind daher unbedingt zu vermeiden (vgl. Posch et al. 1994). Neben dieser grundlegenden didaktischen Zielvorgabe wurden bei der Generierung der Testitems die gängigen Empfehlungen zur Frage- und Antwortformulierung von Erhebungsinstrumenten berücksichtigt, um die Eindeutigkeit und Verständlichkeit der Items sicherzustellen (vgl. Bühner 2013; Paier 2010).

2.3 Erprobung des Erhebungsinstruments

2.3.1 Beurteilung der inhaltlichen und didaktischen Eignung

Einen ersten Indikator für die Güte des Erhebungsinstruments lieferte die Beurteilung der inhaltlichen und didaktischen Eignung der Testitems durch Expert/inn/en in der Domäne der ökonomischen Bildung. Diese Vorgehensweise entspricht einer Einschätzung der Inhaltsvalidität des Erhebungsinstruments, welche dann gegeben ist, wenn die Testitems eine repräsentative Auswahl aus dem hypothetischen Itemuniversum des zugrundeliegenden Konstrukts bilden. Die Inhaltsvalidität wird in der Regel nicht quantitativ anhand eines Kennwerts ermittelt, sondern aus qualitativen fachlichen Überlegungen heraus geschlussfolgert (vgl. Moosbrugger/Kelava 2012). Das Konzept zur Bestimmung der Inhaltsvalidität ist mangels geeigneter Anhaltspunkte zur Konkretisierung des hypothetischen Itemuniversums zwar nicht auf alle Tests anwendbar, bei der Validierung von Bildungstests aber durchaus gebräuchlich, da sich hier das Itemuniversum beispielsweise aus Lehrplänen oder Bildungsstandards ableiten lässt (vgl. Loerwald/Schnell 2014; Schmidt-Atzert/Amelang 2012).

Als Expert/inn/en für die Inhaltsvalidierung des Erhebungsinstruments wurden sowohl Mitarbeiter/innen des Lehrstuhls als auch Wirtschaftslehrkräfte einbezogen, um unterschiedliche Zugänge zur Domäne ökonomischer Bildung abzudecken. Die Expert/inn/en wurden gebeten, die Testitems hinsichtlich ihrer inhaltlichen Relevanz und ihrer didaktischen Qualität zu beurteilen und etwaige Verbesserungsvorschläge direkt auf dem Instrumententwurf anzumerken. Zusätzlich sollte eine Gesamtbeurteilung des Erhebungsinstruments vorgenommen werden, bei der zu berücksichtigen war, ob das Instrument die zentralen wirtschaftlichen Themenbereiche abdeckt und der Zielgruppe bezüglich der Inhalte, der Formulierungen und des Schwierigkeitsgrads angemessen ist.

2.3.2 Erkundung der Angemessenheit für die Zielgruppe

Die inhaltliche Beurteilung des Erhebungsinstruments durch Expert/inn/en alleine gibt noch keinen Aufschluss darüber, ob die Testitems von der Zielgruppe tatsächlich verstanden und in der vorgesehenen Art und Weise bearbeitet werden. Auch aus den Testergebnissen lässt

sich dies nicht gesichert ableiten, da das Erhebungsinstrument ausschließlich aus geschlossenen Aufgabenformaten besteht und die Antworten folglich auch gänzlich erraten sein könnten. Aus diesem Grund wurde eine zusätzliche Erkundung der Angemessenheit für die Zielgruppe nach der Methode lauten Denkens vorgenommen, um die Gedankengänge der Testpersonen und damit deren Problemlösungsstrategien bei der Aufgabenbearbeitung aufzudecken (vgl. Loerwald/Schnell 2014; Moosbrugger/Kelava 2012). Die Methode wurde in der Form der Introspektion angewandt, bei der die Testperson ihre Gedanken unmittelbar im Zuge der Aufgabenbearbeitung verbalisiert. Dabei erfolgt weder eine Strukturierung noch eine Interpretation oder Reflexion der eigenen Gedanken, die Testperson richtet ihre bewusste Anstrengung gänzlich auf die bearbeitete Aufgabe. Diese Form der Methode lauten Denkens weist die engste und natürlichste Verbindung zu den Gedankengängen der Testperson auf (vgl. Ericsson/Simon 1993; Konrad 2010).

Als Testpersonen für die Durchführung der Methode konnten vier Schüler/innen der Zielgruppe gewonnen werden, welche alle keine Schwierigkeiten darin hatten, ihre eigenen Gedanken zu artikulieren. Die Schüler/innen wurden bei der Aufgabenbearbeitung gebeten, die Aufgabenstellungen jeweils laut vorzulesen und sodann noch einmal in eigenen Worten zusammenzufassen. Anschließend sollten sie ihre Gedanken bei der Aufgabenbeantwortung, etwa warum sie eine Antwortalternative als richtig oder falsch erachteten, laut äußern, so als würden sie mit sich selbst sprechen. Das laute Denken der Schüler/innen wurde auf Tonband aufgezeichnet, für die Beurteilung der Testitems unmittelbar relevante Äußerungen wurden zudem direkt auf dem Instrumententwurf notiert. Die Ausführungen der Schüler/innen gaben reichen Aufschluss über die Angemessenheit des Erhebungsinstruments für die Zielgruppe, etwa ob die Aufgabenstellungen die intendierten Gedankengänge auslösen, die Antwortalternativen eindeutig und nachvollziehbar sind oder die verwendeten Formulierungen und Begrifflichkeiten verständlich sind.

2.3.3 Überprüfung der Erfüllung psychometrischer Anforderungen

Neben der inhaltlichen Angemessenheit aus der Sicht von Expert/inn/en sowie der Zielgruppe spielt die Erfüllung psychometrischer Anforderungen eine zentrale Rolle für die Güte des Erhebungsinstruments. Ein wesentlicher Schritt im Rahmen der Instrumenterprobung war daher die Itemanalyse, in deren Verlauf psychometrische Itemkennwerte ermittelt und anhand vorgegebener Qualitätsstandards beurteilt wurden. Dem Konzept der klassischen Testtheorie folgend wurden zur Itemanalyse die folgenden statistischen Verfahren eingesetzt (vgl. Bortz/Döring 2006; Bühner 2013; Lienert/Raatz 1994):

- Die Rohwerteverteilung der Items wurde über das Minimum, das Maximum, den Mittelwert und die Standardabweichung der Testwerte ausgedrückt. Bei allen Items wurde eine hinreichende Streuung über den Wertebereich angestrebt.
- Die Itemschwierigkeit wurde als prozentueller Indexwert der Mittelwerte der Testwerte ermittelt. Schwierigkeiten unter 20 % sowie über 80 % des erreichbaren Punktwerts eines Items wurden als unangemessen eingestuft.
- Die Trennschärfe der Items wurde anhand der part-whole-korrigierten Korrelation eines Items zu den übrigen Items des Erhebungsinstruments errechnet. Trennschärfen unter 0,2 wurden als zu gering beurteilt.
- Die Homogenität der Items wurde in Form von Cronbachs Alpha über die Items des Erhebungsinstruments angegeben. Alpha-Werte unter 0,7 wurden als unzureichend bewertet.
- Die Dimensionalität der Items wurde über eine Faktorenanalyse nach der Hauptkomponenten-Methode mit dem MAP-Test sowie dem Scree-Test als Extraktionskriterien überprüft. Hinsichtlich der Anzahl an extrahierten Faktoren wurde eine einfaktorielle Lösung oder zumindest ein varianzstärkster erster Faktor angestrebt.

Über die Verfahren der klassischen Testtheorie hinaus wurde mit dem Rasch-Modell auch eine Itemanalyse nach der probabilistischen Testtheorie vorgenommen. Diese umfasste folgende Analyseschritte (vgl. Strobl 2012):

- Die Schwierigkeitsparameter für die Items und die Fähigkeitsparameter für die Testpersonen wurden über eine Maximum-Likelihood-Schätzung ermittelt und in einer Person-Item-Map dargestellt.
- Die Item-Characteristic-Curves für die Items wurden ermittelt, auf strenge monotone Steigung geprüft und in einem ICC-Plot aneinandergereiht.
- Die Gültigkeit des Modells wurde weiters über einen grafischen Modelltest mit 95 %-Konfidenzintervallen der Schwierigkeitsparameter und den Summenwerten der Testpersonen als Splitkriterium überprüft.

Die Durchführung der Itemanalyse erforderte den Einsatz des Erhebungsinstruments in einer Voruntersuchung. Diese diente dazu, das Instrument einer ersten empirischen Bewährungsprobe auszusetzen und in einer Erstanalyse der Items den dringlichsten Überarbeitungsbedarf zu identifizieren. Um eine aussagekräftige Datenbasis für die Analyse zu generieren, wurde das Erhebungsinstrument in der für den Echteininsatz bestimmten Form und unter vergleichbaren Bedingungen erprobt (vgl. Lienert/Raatz 1994). Der Einsatz erfolgte an insge-

samt 432 Schüler/inne/n aus 16 Schulen in acht Bundesländern. Bei der Stichprobenbildung wurde darauf geachtet, über unterschiedliche Schulen, Schultypen und Bundesländer zu streuen. Innerhalb dieser Rahmenbedingungen folgte die Auswahl einem Convenience Sampling, da die Teilnahme an der Erhebung die Genehmigung der zuständigen Schulbehörden sowie das Einverständnis der Schulen, der Lehrkräfte und der Erziehungsberechtigten der Schüler/innen erforderte und somit einer Vielzahl an unkontrollierbaren Faktoren unterlag. Während der Erhebung waren jeweils ein/e geschulte/r Testleiter/in und nach Möglichkeit auch die Klassenlehrkräfte anwesend, um die Ernsthaftigkeit der Situation zu verdeutlichen. Fragen der Schüler/innen zum Test oder zum Forschungsvorhaben wurden ausschließlich von den Testleiter/inne/n beantwortet.

Die Antworten der Schüler/innen wurden nachträglich manuell am Computer erfasst und aufbereitet. Bei offensichtlich unmotivierter oder willkürlicher Bearbeitung des Erhebungsinstruments wurden die entsprechenden Exemplare schon während der Erhebungssituation ausgeschieden, fehlende oder fehlerhafte Einzelangaben wurden im Zuge der Datenerfassung gesondert codiert und bei der anschließenden Datenauswertung nicht berücksichtigt. Die Datenaufbereitung diente dazu, eine möglichst hohe Datenqualität für die Auswertung und Analyse der Itemkennwerte sicherzustellen.

3 Erste Ergebnisse der Instrumententwicklung

3.1 Erhebungsinstrument der Voruntersuchung

Nach der beschriebenen methodischen Vorgehensweise wurde von April bis Mai 2015 die erste Version des Erhebungsinstruments in Form eines schriftlichen Fragebogens erstellt. Dieser wurde aus mehreren Teilbereichen zusammengesetzt, die neben dem Wirtschaftswissen als Kern noch weitere Facetten ökonomischer Kompetenzen behandeln, welche als mögliche Kontroll- bzw. Moderatorvariablen für die Analyse der Wissensaufgaben dienen. Hierfür wurde von den Testpersonen zusätzlich das selbst eingeschätzte Wirtschaftswissen, das Interesse an Wirtschaft und die Einstellung zur Wirtschaft erfragt. Am Ende des Fragebogens wurden zudem einige soziodemografische Merkmale der Testpersonen erfasst, um eine differenzierte Analyse nach Teilpopulationen zu ermöglichen. Der Inhalt der ersten Version des Erhebungsinstruments stellt sich damit wie folgt dar:

- Das Wirtschaftswissen wurde mithilfe von 25 Items erhoben, bei denen die Schüler/innen Wissensaufgaben zu wirtschaftlichen Sachverhalten und Zusammenhängen beantworten mussten. Die Aufgaben wurden im Multiple-Choice-Format mit je vier Antwortalternativen

und mehreren möglichen korrekten Antworten gestellt, bei denen auch Teilpunkte erreicht werden konnten.

Das Wirtschaftswachstum ist eine wichtige Kennzahl für die wirtschaftliche Entwicklung eines Landes. Welche der folgenden Situationen werden zu einem höheren Wirtschaftswachstum führen? Kreuze diese an! Mehrere Antworten können korrekt sein.

- Die Bevölkerung legt mehr Geld auf das Spargbuch, anstatt es auszugeben.
- Der Staat lässt neue Häuser und Straßen bauen.**
- Der Staat erhöht die Steuern und zahlt damit seine Staatsschulden zurück.
- Die Unternehmen exportieren mehr Produkte in das Ausland.**

Abbildung 1: Beispielitem zum Wirtschaftswissen

- Zur Erfassung des selbst eingeschätzten Wirtschaftswissens wurden acht Items eingesetzt, bei denen die Schüler/innen ihre Kenntnisse über die Wirtschaft selbst bewerten mussten. Dazu kam das Antwortformat einer fünfstufigen Likert-Skala von „nicht genügend“ bis „sehr gut“ zum Einsatz.

Ich kann erklären, warum wir wirtschaften.

Abbildung 2: Beispielitem zum selbst eingeschätzten Wirtschaftswissen

- Das Interesse an Wirtschaft wurde anhand von acht Items erhoben, bei denen die Schüler/innen die Bedeutsamkeit von Wirtschaft sowie die Häufigkeit und die Freude ihrer Auseinandersetzung mit wirtschaftlichen Themenbereichen selbst einschätzen mussten. Als Antwortformat stand eine fünfstufige Likert-Skala von „stimme nicht zu“ bis „stimme sehr zu“ zur Verfügung.

Wirtschaftliche Themen sind für mich wichtig.

Abbildung 3: Beispielitem zum Interesse an Wirtschaft

- Zur Erfassung der Einstellung zur Wirtschaft wurden neun Items verwendet, bei denen die Schüler/innen verschiedene als Gerechtigkeitsurteile über die soziale Marktwirtschaft formulierte Aussagen selbst beurteilen mussten. Die Beantwortung erfolgte auf einer fünfstufigen Likert-Skala von „stimme nicht zu“ bis „stimme sehr zu“.

Es ist gerecht, dass in unserer Wirtschaft alle Steuern zahlen müssen.

Abbildung 4: Beispielitem zur Einstellung zur Wirtschaft

- An soziodemografischen Merkmalen wurden schließlich Alter und Geschlecht der Schüler/innen, der besuchte Schultyp und das Bundesland sowie die zuhause am häufigsten gesprochene Sprache erfragt.

Die Items des Erhebungsinstruments wurden aus bereits genannten Gründen überwiegend selbst entwickelt, eine geringere Anzahl an Items wurde aus bestehenden Instrumenten adaptiert und übernommen. Tabelle 1 stellt den Inhalt und die Quellen des Erhebungsinstruments im Überblick dar. Auf Anfrage stellen die Autor/inn/en dieses Beitrags das Erhebungsinstrument zur Verfügung.

Teilbereich	Anzahl Items	Quelle
Wirtschaftswissen	25	15 Items in Eigenentwicklung 6 Items adaptiert aus Loerwald/Schnell (2014) 4 Items adaptiert aus Walstad et al. (2010)
Selbsteinschätzung des Wirtschaftswissens	8	8 Items in Eigenentwicklung
Interesse an Wirtschaft	8	7 Items in Eigenentwicklung 1 Item adaptiert aus Oberste (2012)
Einstellung zur Wirtschaft	9	4 Items in Eigenentwicklung 3 Items adaptiert aus Oberste (2012) 1 Item adaptiert aus Soper/Walstad (1983) 1 Item adaptiert aus GESIS (2014)

Tabelle 1: Inhalt und Quellen des Erhebungsinstruments

3.2 Itemanalyse der Voruntersuchung

Nach einer vorläufigen qualitativen Beurteilung, im Zuge derer noch geringfügige inhaltliche und sprachliche Anpassungen vorgenommen wurden, wurde die erste Version des Erhebungsinstruments zwischen Juni und Juli 2015 empirisch erprobt. Die inhaltliche Beurteilung durch die Expert/inn/en sowie die Erkundung der Angemessenheit für die Zielgruppe ergaben keinen großen Überarbeitungsbedarf. Aus Sicht der Expert/inn/en wurde das Instrument weitgehend als anspruchsvoll, von den behandelten Inhalten aber als für die Zielgruppe relevant eingestuft. Die Methode lauten Denkens mit der Zielgruppe zeigte, dass die Aufgabenstellungen für die Schüler/innen größtenteils verständlich waren und auch inhaltlich

plausible Überlegungen auslösten. Aufschluss über die psychometrische Güte der Items lieferte schließlich die statistische Itemanalyse, in deren Verlauf sowohl Verfahren der klassischen als auch der probabilistischen Testtheorie eingesetzt wurden. Tabelle 2 bietet eine kurze Stichprobencharakteristik der 432 an der Voruntersuchung teilnehmenden Schüler/innen:

Merkmal		Anzahl	%
Alter	<14	90	21,1
	14	277	64,9
	15	51	11,9
	>15	9	2,1
Geschlecht	Männlich	219	51,3
	Weiblich	208	48,7
Schultyp	AHS	219	50,7
	NMS	213	49,3
Bundesland	Wien	39	9,0
	Burgenland	41	9,5
	Niederösterreich	135	31,3
	Oberösterreich	15	3,5
	Steiermark	63	14,6
	Kärnten	60	13,9
	Tirol	35	8,1
	Vorarlberg	44	10,2
Zuhause am häufigsten gesprochene Sprache	Deutsch	338	79,3
	Andere Sprache	88	20,7

Tabelle 2: Stichprobencharakteristik der Voruntersuchung

3.2.1 Itemanalyse nach der klassischen Testtheorie

Die Testwerte der Wissensaufgaben für die Itemanalyse wurden über ein Bewertungsverfahren ermittelt, nach dem je Aufgabe maximal ein Punkt erzielt werden konnte. Die einzelnen Antwortalternativen gingen dabei gewichtet in die Bewertung ein, sodass auch Teilpunkte erreicht werden konnten. Mit den ermittelten Testwerten wurde sodann mit der Software SPSS die Itemanalyse durchgeführt, deren Ergebnisse in Tabelle 3 dargestellt sind:

Item	Minimum	Maximum	Mittelwert	Standard-abweichung	Schwierigkeit	Trennschärfe	Cronbachs Alpha	Varianz-aufklärung
w1	0,000	1,000	0,176	0,381	17,6%	0,141	0,766	16,3%
w2	0,000	1,000	0,296	0,457	29,6%	0,283		
w3	0,000	1,000	0,498	0,177	49,8%	0,276		
w4	0,000	1,000	0,373	0,484	37,3%	0,162		
w5	0,000	1,000	0,650	0,477	65,0%	0,348		
w6	0,000	1,000	0,384	0,409	38,4%	0,228		
w7	0,000	1,000	0,458	0,399	45,8%	0,278		
w8	0,000	1,000	0,303	0,386	30,3%	0,264		
w9	0,000	1,000	0,433	0,431	43,3%	0,297		
w10	0,000	1,000	0,407	0,492	40,7%	0,270		
w11	0,000	1,000	0,653	0,444	65,3%	0,233		
w12	0,000	1,000	0,622	0,291	62,2%	0,267		
w13	0,000	1,000	0,368	0,425	36,8%	0,312		
w14	0,000	1,000	0,597	0,491	59,7%	0,408		
w15	0,000	1,000	0,624	0,444	62,4%	0,513		
w16	0,000	1,000	0,395	0,440	39,5%	0,355		
w17	0,000	1,000	0,343	0,475	34,3%	0,100		
w18	0,000	1,000	0,338	0,355	33,8%	0,205		
w19	0,000	1,000	0,635	0,450	63,5%	0,470		
w20	0,000	1,000	0,438	0,372	43,8%	0,362		
w21	0,000	1,000	0,447	0,498	44,7%	0,286		
w22	0,000	1,000	0,380	0,486	38,0%	0,425		
w23	0,000	1,000	0,436	0,480	43,6%	0,245		
w24	0,000	1,000	0,442	0,390	44,2%	0,376		
w25	0,000	1,000	0,502	0,458	50,2%	0,385		

Tabelle 3: Ergebnisse der Itemanalyse nach klassischer Testtheorie

Die Minima und Maxima der erreichten Punkte zeigen, dass bei allen Items das gesamte Punktespektrum in den Daten abgebildet wird. Auch die Standardabweichungen von rund 0,4 Punkten zeugen von einer hinreichenden Streuung der Testwerte. Eine Ausnahme bildet das Item w3, das mit 0,177 Punkten eine vergleichsweise geringe Standardabweichung aufweist. Die aus den Mittelwerten der erreichten Punkte abgeleiteten Schwierigkeitswerte der Items reichen von 17,6 % bis 65,3 % und liegen überwiegend in der unteren Hälfte des Punktespektrums, was mittleren bis höheren Itemschwierigkeiten entspricht. Insbesondere das Item w1 erweist sich angesichts des Schwierigkeitswerts von 17,6 % als zu schwer. Die Trennschärfen der Items können mit Werten um 0,3 weitgehend als akzeptabel beurteilt werden. Als zu wenig trennscharf stellen sich jedoch mit Trennschärfen von 0,141, 0,162 und 0,100 die Items w1, w4 und w17 heraus. Die Homogenität der Items stellt sich mit einem Cronbachs Alpha von 0,766 als erfüllt dar, was die Zusammenfassung der Items zu einer gemeinsamen Wissensskala legitimiert. Dies wird durch die faktorenanalytische Überprüfung weiter unterstützt, welche sowohl nach dem Scree-Test als auch nach dem MAP-Test (nicht abgebildet) für eine einfaktorielle Lösung spricht, wenngleich die durch diesen Faktor erklärte Gesamtvarianz von 16,3 % als gering einzustufen ist.

Fasst man die Testwerte der Items zu einer gemeinsamen Wissensskala zusammen, so ergibt sich die in Abbildung 5 dargestellte Verteilung der Summenwerte:

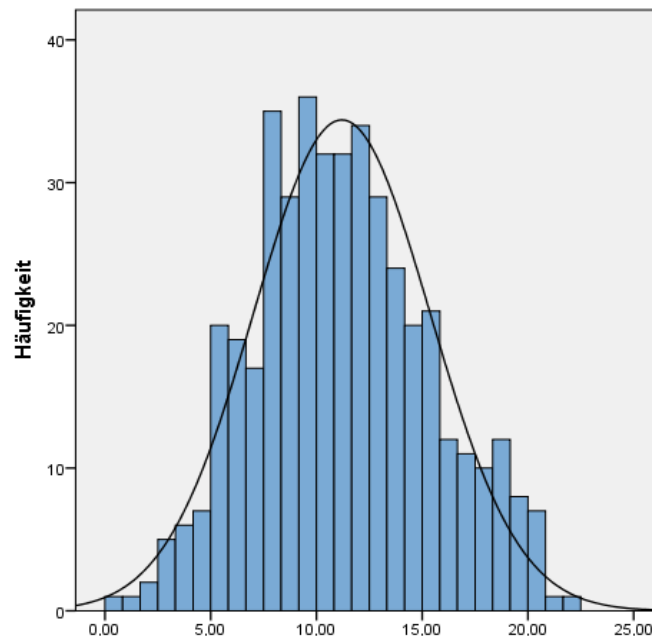


Abbildung 5: Testwertverteilung der gemeinsamen Wissensskala

Das Histogramm belegt eine annähernde Normalverteilung der Summenwerte auf der gemeinsamen Wissensskala, was die Skala als Ausgangspunkt für weiterführende statistische Analysen, beispielsweise in Form von Regressionsanalysen, qualifiziert. Im Mittel wird mit 11,20 von 25 Punkten knapp die Hälfte der erreichbaren Punkteanzahl erzielt, die Standardabweichung beträgt dabei 4,175 Punkte. Das Minimum liegt bei 0, das Maximum bei 21,67 erreichten Punkten. Damit ist eine weitere Vorgabe der Testentwicklung erfüllt, nach der die höchstmögliche Bewertungsstufe unerreicht bleiben sollte.

3.2.2 Itemanalyse nach der probabilistischen Testtheorie

Für die Itemanalyse nach der probabilistischen Testtheorie wurden die zu Teilpunkten ermittelten Testwerte zunächst übernommen und in ein Partial-Credit-Modell nach Rasch überführt. Leider erwies sich dieses Modell als nicht praktikabel, da sich bei den einzelnen Items eine unterschiedliche Anzahl an Bewertungsstufen ergab und bei der begrenzten Stichprobe viele Bewertungsstufen mit einer unzureichenden Anzahl an Testwerten besetzt waren. Aus diesem Grund wurden die Testwerte sodann in ein rein dichotomes Format (0/1) transformiert und mit einem dichotomen Rasch-Modell mit der Software R analysiert. Dies hat zur Folge, dass die Itemschwierigkeiten nicht direkt mit jenen aus der Analyse nach klassischer Testtheorie verglichen werden können, schmälert aber die Aussagekraft der weiteren Analyseschritte nicht. Die über das dichotome Rasch-Modell geschätzten Itemschwierigkeits- und Personenparameter sind auf der Person-Item-Map in Abbildung 6 dargestellt:

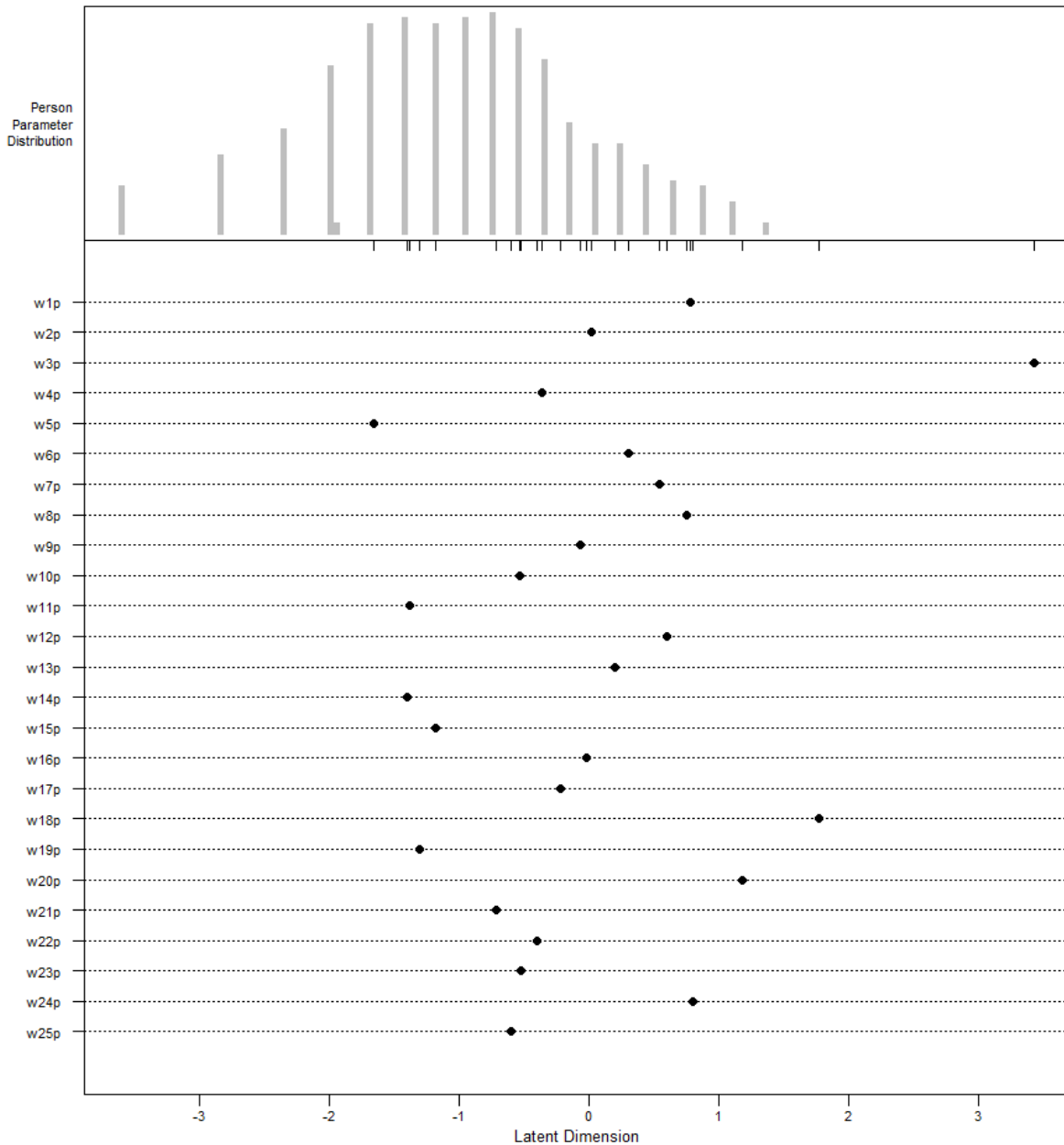


Abbildung 6: Person-Item-Map der Schwierigkeits- und Personenparameter

Aus der Abbildung geht deutlich hervor, dass die Schwierigkeitsparameter zufriedenstellend in einem angemessenen Bereich von -2 bis +2 auf der latenten Dimension streuen. Dies zeugt von unterschiedlichen Itemschwierigkeiten und lässt darauf schließen, dass die latente Dimension mit wenig redundanter Information gut abgedeckt wird. Lediglich das Item w3 weist mit einem Schwierigkeitsparameter von 3,42 eine unangemessene Itemschwierigkeit auf und stellt sich eindeutig als zu schwer heraus. Die Personenparameter sind wie die

Summenwerte auf der gemeinsamen Wissensskala annähernd normalverteilt und erlauben deren Einbezug in weiterführende statistische Analysen.

Zur eingehenderen Analyse der Items wurden über das Rasch-Modell auch die Lösungswahrscheinlichkeiten der Items in Abhängigkeit der Personenparameter ermittelt und als Item-Characteristic-Curves grafisch ausgegeben. Diese sind für alle Items verdichtet in Abbildung 7 abgebildet:

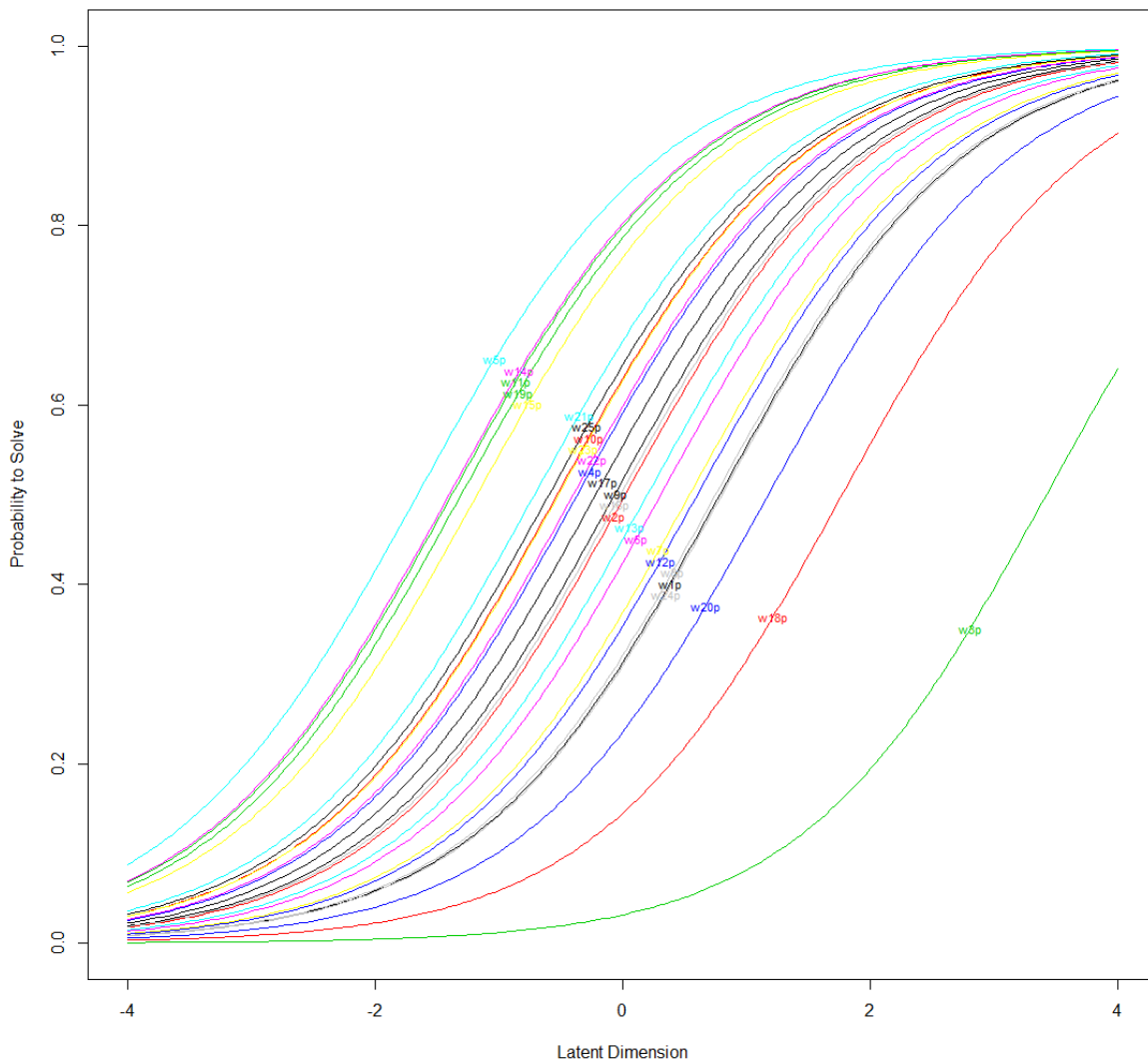


Abbildung 7: Itemlösungswahrscheinlichkeiten nach Personenparameter

Aus der streng monotonen Steigung der Item-Characteristic-Curves geht hervor, dass bei allen Items die Lösungswahrscheinlichkeit mit höherer Personenfähigkeit auf der latenten Dimension ansteigt. Die parallele Verschiebung der Kurven drückt dabei die unterschiedlichen Itemschwierigkeiten von leicht (links) bis schwer (rechts) aus. Auch in dieser Darstel-

lung zeigt sich der überhöhte Schwierigkeitsgrad des Items w3, so ist eine 50%ige Lösungswahrscheinlichkeit auf das Item nur bei sehr hoher Personenfähigkeit erreicht. Weiters kreuzen sich die Kurven der einzelnen Items nicht, womit sich Änderungen der Personenfähigkeit auf der latenten Dimension gleichermaßen auf die Lösungswahrscheinlichkeiten aller Items auswirken. Dies spricht in Summe für die Gültigkeit des Rasch-Modells.

In einem letzten Analyseschritt wurde die Eindimensionalität als zentrale Voraussetzung für die Anwendbarkeit des Rasch-Modells überprüft. Die Prüfung erfolgte über einen grafischen Modellgeltungstest mit 95%-Konfidenzintervallen der Schwierigkeitsparameter und dem Median der Summenwerte als Splittkriterium. Abbildung 8 zeigt die Ergebnisse des Tests:

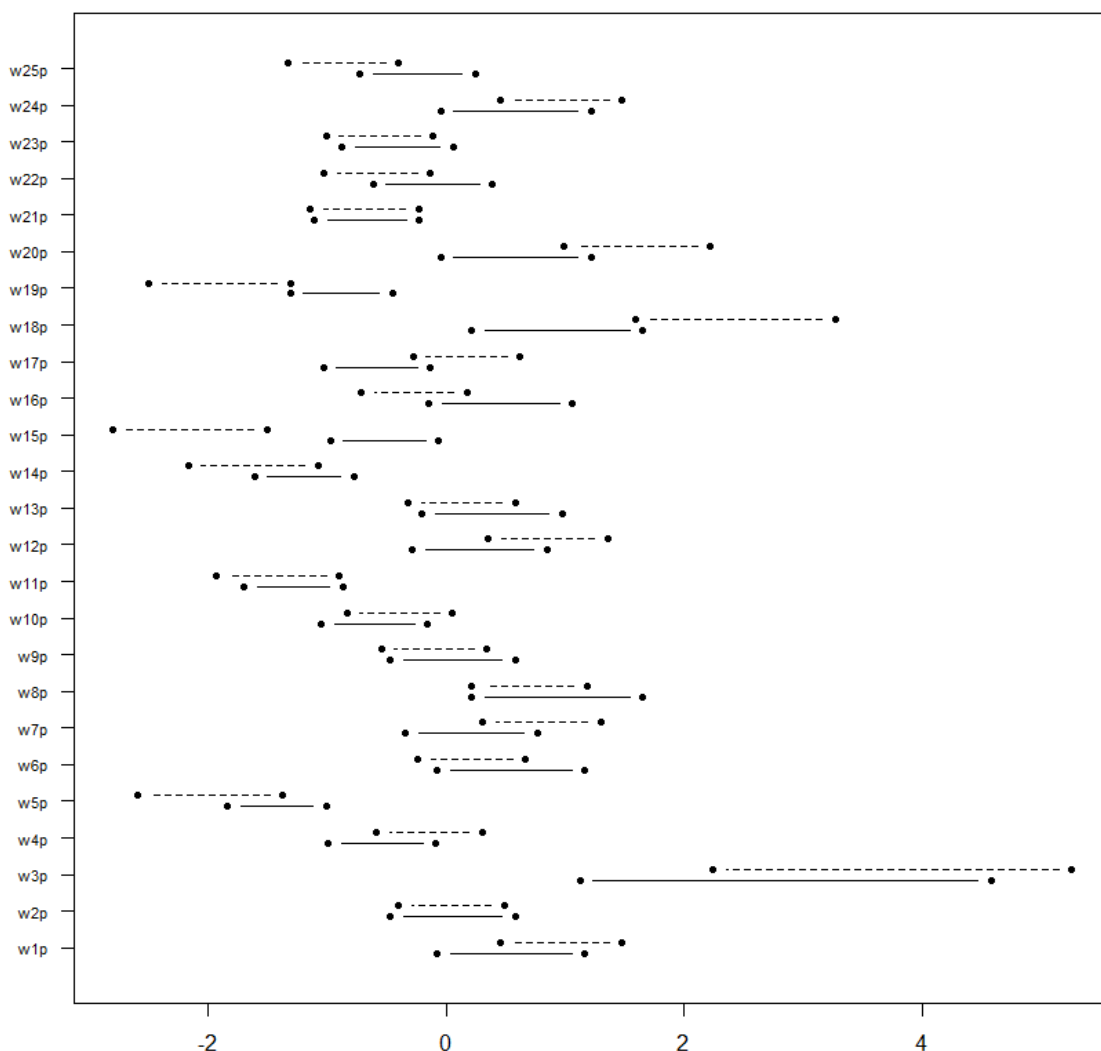


Abbildung 8: Grafischer Modellgeltungstest der Schwierigkeitsparameter

Die unterbrochenen und die durchgezogenen Linien stellen die 95%-Konfidenzintervalle der Schwierigkeitsparameter für die beiden Subgruppen mit hohen und niedrigen Summenwerten dar. Um eine Alphafehler-Kumulierung bei multiplen Vergleichen zu vermeiden, wurde für die Konfidenzintervalle eine Bonferroni-Korrektur vorgenommen. Überlappende Konfidenzintervalle der Schwierigkeitsparameter eines Items deuten darauf hin, dass das Item für beide Subgruppen Ähnliches misst und daher Eindimensionalität angenommen werden kann. Dies ist bei sämtlichen Items mit Ausnahme des Items w15 der Fall, dessen Konfidenzintervalle auffällig voneinander abweichen. Bei diesem Item ist von Differential-Item-Functioning auszugehen, offensichtlich misst das Item neben der latenten Dimension Wirtschaftswissen noch weitere Merkmale, die die beiden Subgruppen ungewollt voneinander unterscheiden. Bei den verbleibenden Items überlappen sich die Konfidenzintervalle jedoch größtenteils deutlich, sodass es legitim erscheint, die Annahme der Eindimensionalität in der Gesamtbetrachtung beizubehalten.

4 Zusammenfassung und Ausblick

Die Zielsetzung des in diesem Beitrag dargestellten Forschungsvorhabens besteht darin, nach wissenschaftlichen Standards ein Erhebungsinstrument zur standardisierten Messung des Wirtschaftswissens von Schüler/inne/n am Ende der Sekundarstufe I zu entwickeln. Eine erste Version des Erhebungsinstruments wurde im Frühjahr 2015 erstellt und zwischen Juni und Juli 2015 empirisch erprobt. Die Ergebnisse der Instrumenterprobung sind dabei erfolgversprechend. Im Hinblick auf die inhaltliche und didaktische Qualität wurde das Instrument von den befragten Expert/inn/en weitgehend als anspruchsvoll, aber als inhaltlich relevant beurteilt. In der Erkundung der Angemessenheit für die Zielgruppe nach der Methode lauten Denkens mit Schüler/inne/n stellte sich heraus, dass die Aufgabenstellungen hinsichtlich der verwendeten Formulierungen verständlich waren und auch inhaltlich plausible Gedankengänge auslösten.

Auch die psychometrische Güte des Erhebungsinstruments kann nach den durchgeführten Itemanalysen als aussichtsreich eingestuft werden. In der Analyse nach klassischer Testtheorie liegen die Itemschwierigkeiten überwiegend im angestrebten Bereich bei zugleich hinreichender Streuung der Testwerte. Nur Items mit sehr geringem Schwierigkeitsgrad erweisen sich als verhältnismäßig unterrepräsentiert, sodass das Instrument bei gering ausgeprägtem Wirtschaftswissen womöglich noch unzureichend differenziert. Die Trennschärfen der Items übertreffen den angezielten Grenzwert zumeist deutlich, nur in wenigen Fällen

wird die geforderte Trennschärfe von 0,2 nicht erreicht. Des Weiteren kann die Itemhomogenität mit einem Cronbachs Alpha von 0,766 als erfüllt betrachtet werden und auch die faktoranalytische Dimensionalitätsprüfung legt die angestrebte einfaktorielle Lösung nahe, wengleich die dadurch erklärte Gesamtvarianz mit 16,3 % nicht beträchtlich ausfällt. Die nach der probabilistischen Testtheorie mittels dichotomem Rasch-Modell ergänzend vorgenommene Analyse unterstützt diese Schlussfolgerungen. Die geschätzten Schwierigkeitsparameter der Items streuen mit einer Ausnahme über einen angemessenen Schwierigkeitsbereich und erweisen sich in ihren Item-Characteristic-Curves als modellkonform. Der grafische Modellgeltungstest der Schwierigkeitsparameter mit dem Median der Summenwerte als Splitkriterium lässt in der Gesamtbetrachtung eine Eindimensionalität der Items annehmen, lediglich bei einem Item ist von ungewolltem Differential-Item-Functioning auszugehen.

Eine Einschränkung der Aussagekraft dieser Ergebnisse ergibt sich vor allem aus der für eine umfassende Instrumenterprobung begrenzten Stichprobengröße von 432 Schüler/inne/n, aufgrund derer etwa das zur Itemanalyse nach der probabilistischen Testtheorie ursprünglich angestrebte Partial-Credit-Modell letztlich nicht angewandt werden konnte. Zudem wurde auf vertiefte Analysen wie eine differenzierte Betrachtung der Messinvarianz nach verschiedenen soziodemografischen Merkmalen oder zusätzliche Modellgeltungstests noch verzichtet. Im Hinblick auf den Status einer Voruntersuchung zur Erprobung der ersten Instrumentversion erscheinen diese Einschränkungen jedoch argumentierbar.

Ausgehend von den empirischen Befunden dieser Voruntersuchung wurde das Instrument bereits überarbeitet und weiterentwickelt. Items, an denen Überarbeitungsbedarf festgestellt wurde, wurden entweder umformuliert oder inhaltlich abgeändert. Nach einer neuerlichen inhaltlichen und didaktischen Beurteilung durch Expert/inn/en und einer weiteren Erkundung der Angemessenheit für die Zielgruppe wurden mit der überarbeiteten zweiten Version des Erhebungsinstruments österreichweit bereits weitere Daten erhoben, die noch eingehend analysiert werden. Im Rahmen der bevorstehenden Itemanalyse wird über die in diesem Beitrag dargestellten Verfahren hinaus eine differenzierte Betrachtung der Itemgüte nach verschiedenen Teilpopulationen erfolgen. Zudem werden die Zusammenhänge zwischen dem Wirtschaftswissen und den über zusätzliche Items erhobenen weiteren Facetten ökonomischer Kompetenzen näher analysiert, die sich im Zuge der Voruntersuchung ebenfalls als reliable Skalen erwiesen. Mit einem nach diesen Kriterien bewährten Instrument könnte ein substantieller Beitrag zur wissenschaftlich fundierten Messung des Wirtschaftswissens in der Sekundarstufe I geleistet werden.

Literaturverzeichnis

- Aff, J. (2015): Plädoyer für mehr politische Bildung durch Aufklärung über ökonomisch-gesellschaftliche Zusammenhänge. In: WissenPlus Wissenschaft, 33 (5), 4-14.
- Albers, H.-J. (1988): Ökonomische Bildung und Allgemeinbildung. In: Bundesfachgruppe für ökonomische Bildung (Hg.): Ökonomische Bildung – Aufgabe für die Zukunft, Bergisch Gladbach, 1-16.
- Albers, H.-J. (1995): Handlungsorientierung und ökonomische Bildung. In: Albers, H.-J. (Hg.): Handlungsorientierung und ökonomische Bildung, Bergisch Gladbach, 1-22.
- Anderson, L. W./Krathwohl, D. R. (2001): A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy, New York.
- Beck, K. (1989): „Ökonomische Bildung“ – Zur Anatomie eines wirtschaftspädagogischen Begriffs. In: Zeitschrift für Berufs- und Wirtschaftspädagogik, 85 (7), 579-596.
- Beck, K./Krumm, V./Dubs, R. (1998): Wirtschaftskundlicher Bildungstest (WBT), Göttingen.
- Bloom, B. S./Engelhart, M. D./Furst, E. J. et al. (1972): Taxonomie von Lernzielen im kognitiven Bereich, Weinheim/Basel.
- BMBF (2015): Lehrplan des Unterrichtsgegenstands Geographie und Wirtschaftskunde. Online: https://www.bmbf.gv.at/schulen/unterricht/lp/ahs9_784.pdf?4dzgm2 (17.05.2016).
- Bortz, J./Döring, N. (2006): Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler, Heidelberg.
- Brandmaier, E./Frank, H./Korunka, C. et al. (2006): Ökonomische Bildung von Schüler/innen Allgemeinbildender Höherer Schulen, Wien.
- Bühner, M. (2013): Einführung in die Test- und Fragebogenkonstruktion, München.
- Davies, P. (2001): Citizenship and Economic Understanding in England. In: Schlösser, H. J. (Hg.): Stand und Entwicklung der ökonomischen Bildung, Bergisch Gladbach, 1-20.
- DeGÖB (2005): Kompetenzen der ökonomischen Bildung für allgemein bildende Schulen und Bildungsstandards für den mittleren Schulabschluss. In: Weitz, B. O. (Hg.): Standards in der ökonomischen Bildung, Bergisch Gladbach, 3-16.
- Dörner, D. (1987): Problemlösen als Informationsverarbeitung, Stuttgart.
- Dubs, R. (2014): Unterrichtsplanung in der Praxis. Ein Handbuch für den Lernbereich Wirtschaft, Stuttgart.
- Ericsson, K. A./Simon, H. A. (1993): Protocol Analysis. Verbal Reports as Data, Cambridge/London.
- Eysenck, M./Keane, M. (2005): Cognitive Psychology. A Student's Handbook, Hove.
- Förster, M./Happ, R./Zlatkin-Troitschanskaia, O. (2012): Valide Erfassung des volkswirtschaftlichen Fachwissens von Studierenden der Wirtschaftswissenschaften und der Wirtschaftspädagogik – eine Untersuchung der diagnostischen Eignung des Wirtschaftskundlichen Bildungstests (WBT). In: bwp@ Berufs- und Wirtschaftspädagogik – online, 12 (22), 1-21. Online: http://www.bwpat.de/ausgabe22/foerster_etal_bwpat22.pdf (07.01.2016).

- Fortmüller, R. (1997): Wissen und Problemlösen. Eine wissenspsychologische Analyse der notwendigen Voraussetzungen für die Bewältigung von (komplexen) Problemen und Konsequenzen für den Unterricht in berufsbildenden Vollzeitschulen, Wien.
- Geise, W. (2001): Vorwort. In: Geise, W. (Hg.): Ökonomische Bildung zur Bewältigung von Lebenssituationen, Bergisch Gladbach, III-VII.
- GESIS (2014): ALLBUS 2014 Fragebogendokumentation. Online: http://www.gesis.org/fileadmin/upload/dienstleistung/daten/umfragedaten/allbus/Fragebogen/ALLBUS_2014.pdf (12.05.2015).
- Gnan, E./Silgoner, M. A./Weber, B. (2007): Volkswirtschafts- und Finanzbildung: Konzepte, Ziele, Messung. In: Geldpolitik & Wirtschaft, 4 (3), 30-52.
- Greimel-Fuhrmann, B. (2013): „Don't Know Much about Economics and Business.“ Economic und Financial Literacy als wesentliche, jedoch vernachlässigte Bildungsziele. In: WissenPlus, 32 (4), I-VIII.
- Greimel-Fuhrmann, B. (2014): Financial Literacy – die schwierige Messung eines „Life Skill“. Methodische und inhaltliche Überlegungen zur Messung des kompetenten Umgangs mit Geld und Finanzen. In: WissenPlus, 33 (5), 48-53.
- Greimel-Fuhrmann, B./Kronberger, R./Rumpold, H. (2016): Befunde und Desiderata zur Wirtschaftsbildung österreichischer Mittelschüler/innen. In: Wirtschaftspolitische Blätter, 21 (1), 249-263.
- Hartmann, C./Pühringer, J. (2000): Ökonomische Bildung in Österreich, Wien.
- Hedtke, R. (2011): Konzepte ökonomischer Bildung, Schwalbach/Ts.
- Hummelsheim, S. (2010): Ökonomische Grundbildung tut not. Empirische Studien weisen auf erhebliche Defizite in der ökonomischen und finanziellen Grundbildung hin. Online: <http://www.die-bonn.de/doks/hummelsheim1001.pdf> (04.08.2014).
- Ingenkamp, K./Lissman, U. (2005): Lehrbuch der Pädagogischen Diagnostik, Weinheim/Basel.
- Jagsch, D. (2015): Finanzielle Allgemeinbildung in der Sekundarstufe I in Österreich – eine Lehrplan- und Lehrbuchanalyse, Bachelorarbeit, Wien, Wirtschaftsuniversität Wien.
- Kaminski, H. (1996): Ökonomische Bildung und Gymnasium. Ziele, Inhalte, Lernkonzepte des Ökonomieunterrichts, Berlin.
- Kaminski, H. (1997): Neue Institutionenökonomik und ökonomische Bildung. In: Kruber, K.-P. (Hg.): Konzeptionelle Ansätze ökonomischer Bildung, Bergisch Gladbach, 129-160.
- Kaminski, H. (2014): Entwicklung und Institutionalisierung der ökonomischen Bildung als Teil der schulischen Allgemeinbildung. In: Unterricht Wirtschaft + Politik, 4 (3), 50-54.
- Kaminski, H./Eggert, K./Burkard, K.-J. (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 (25.08.2014).
- Karpe, J./Krol, G.-J. (1999): Funktionsbedingungen moderner Gesellschaften und Neue Institutionenökonomik als Herausforderungen für die ökonomische Bildung. In: Krol, G.-J./Kruber, K.-P. (Hg.): Die Marktwirtschaft an der Schwelle zum 21. Jahrhundert – Neue Aufgaben für die ökonomische Bildung?, Bergisch Gladbach, 21-48.

- Katschnig, T./Hanisch, G. (2005): Wirtschaftswissen von Maturant/inn/en im internationalen Vergleich. Eine empirische Studie in den Ländern Österreich, Deutschland, Tschechien und Ungarn. Online: http://homepage.univie.ac.at/tamara.katschnig/Publ%20down/Katschnig_Wirtschaftswissen%202005.pdf (07.01.2016).
- Kluwe, R. (1979): Wissen und Denken. Modelle, empirische Befunde und Perspektiven für den Unterricht, Stuttgart.
- Konrad, K. (2010): Lautes Denken. In: Mey, G./Mruck, K. (Hg.): Handbuch Qualitative Forschung in der Psychologie, Wiesbaden, 476-490.
- Krathwohl, D. R. (2002): A Revision of Bloom's Taxonomy: An Overview. In: Theory into Practice, 41 (4), 212-218.
- Leinert, J. (2004): Finanzieller Analphabetismus in Deutschland: Schlechte Voraussetzungen für eigenverantwortliche Vorsorge. Bertelsmann Stiftung Vorsorgestudien 25. Online: https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/GP_Finanzieller_Analphabetismus_in_Deutschland.pdf (27.07.2016).
- Leinert, J./Wagner, G. G. (2004): Konsumentensouveränität auf Vorsorgemärkten eingeschränkt. Mangelnde „Financial Literacy“ in Deutschland. In: Wochenbericht des DIW Berlin, 71 (30), 427-432.
- Lienert, G. A./Ratz, U. (1994): Testaufbau und Testanalyse, Weinheim.
- Loerwald, D./Schnell, C. (2014): Tests als Instrumente zur Individualdiagnostik in der ökonomischen Bildung – Konzeption, Validierung und Auswertung von Testaufgaben für die Sekundarstufe I in Niedersachsen. In: Retzmann, T. (Hg.): Ökonomische Allgemeinbildung in der Sekundarstufe I und Primarstufe. Konzepte, Analysen, Studien und empirische Befunde, Schwalbach/Ts., 294-306.
- Loerwald, D./Schnell, C. (2016): Diagnostik im Dilemma zwischen fachdidaktischen Ansprüchen und empirischen Anforderungen. Zur (vermeintlichen) Trivialität von Testitems. In: Zeitschrift für Didaktik der Gesellschaftswissenschaften, 7 (1).
- Macha, K./Schuhlen, M. (2011): Modellierung ökonomischer Kompetenz in einer Pilotstudie zu ECOS. Siegener Beiträge zur Ökonomischen Bildung 2/2011. Online: http://dokumentix.uni-siegen.de/opus/volltexte/2012/599/pdf/Modellierung_oekonomischer_kompetenz.pdf (27.07.2016).
- May, H. (2011): Ökonomische Bildung als Allgemeinbildung. In: Aus Politik und Zeitgeschichte, 59 (12), 3-9.
- Mayring, P. (2010): Qualitative Inhaltsanalyse: Grundlagen und Techniken, Weinheim.
- Moosbrugger, H./Kelava, A. (2012): Testtheorie und Fragebogenkonstruktion, Berlin.
- Müller, K./Fürstenau, B./Witt, R. (2007): Ökonomische Kompetenz sächsischer Mittelschüler und Gymnasiasten. In: Zeitschrift für Berufs- und Wirtschaftspädagogik, 103 (2), 227-247.
- Nagy, G./Trautwein, U./Jonkmann, K. et al. (2008): Ein Test zur Erfassung von Wirtschaftswissen am Ende der Sekundarstufe I (TWWS-1). Herbsttagung der Sektion Empirische Bildungsforschung der Deutschen Gesellschaft für Erziehungswissenschaft, Kiel.
- Oberste, M. (2012): Economic and Financial Literacy, Attitude & Behavior – The Case of German High School Students, Dissertation, Münster, Westfälische Wilhelms-Universität Münster.

- Paier, D. (2010): Quantitative Sozialforschung. Eine Einführung, Wien.
- Posch, P./Schneider, W./Mann, W. (1994): Unterrichtsplanung. Mit Beispielen für den betriebswirtschaftlichen Unterricht, Wien.
- Retzmann, T. (2005): Nationale Standards für die ökonomische Bildung – Theoretische Grundlagen und offene Forschungsfragen. In: Weitz, B. O. (Hg.): Standards in der ökonomischen Bildung, Bergisch Gladbach, 51-72.
- Retzmann, T./Frühauf, F. (2014): „Financial Fitness for Life“ - Reichweite und Grenzen der US-amerikanischen Testreihe für die finanzielle Allgemeinbildung. In: Retzmann, T. (Hg.): Ökonomische Allgemeinbildung in der Sekundarstufe I und Primarstufe. Konzepte, Analysen, Studien und empirische Befunde, Schwalbach/Ts., 43-56.
- Schmid, K. (2006): Bildung für eine globalisierte Welt. Halten Österreichs Schulen mit der Internationalisierung der österreichischen Wirtschaft Schritt?, Wien.
- Schmidt-Atzert, L./Amelang, M. (2012): Psychologische Diagnostik, Berlin.
- Schnell, R./Hill, P. B./Esser, E. (2011): Methoden der empirischen Sozialforschung, München.
- Schumann, S./Eberle, F. (2014): Ökonomische Kompetenzen von Lernenden am Ende der Sekundarstufe II. In: Zeitschrift für Erziehungswissenschaft, 17 (1), 103-126.
- Schumann, S./Eberle, F. (2015): Editorial. In: Empirische Pädagogik, 29 (1), 5-9.
- Seeber, G./Retzmann, T./Remmele, B. et al. (2012): Bildungsstandards der ökonomischen Allgemeinbildung. Kompetenzmodell – Aufgaben – Handlungsempfehlungen, Schwalbach/Ts.
- Seeber, S./Nickolaus, R./Winther, E. et al. (2010): Kompetenzdiagnostik in der Berufsbildung. Begründung und Ausgestaltung eines Forschungsprogramms. Beilage zu BWP 1/2010. Online: <https://www.bibb.de/veroeffentlichungen/de/publication/download/id/6162> (27.07.2016).
- Soper, J. C./Walstad, W. B. (1983): On Measuring Economic Attitudes. In: Journal of Economic Education, 14 (4), 4-17.
- Steinmann, B. (1997): Das Konzept Qualifizierung für Lebenssituationen im Rahmen der ökonomischen Bildung heute. In: Kruber, K.-P. (Hg.): Konzeptionelle Ansätze ökonomischer Bildung, Bergisch Gladbach, 1-22.
- Stern, G. H. (2002): Economic Literacy Leads to Better Grasp of Public Policy Issues. In: The Region, 16 (4), 2-5.
- Strobl, C. (2012): Das Rasch-Modell. Eine verständliche Einführung für Studium und Praxis, München/Meiring.
- Walstad, W. B./Rebeck, K./Butters, R. B. (2010): Test of Economic Knowledge. Examiner's Manual, New York.
- Walstad, W. B./Rebeck, K./Butters, R. B. (2013): The Test of Economic Literacy: Development and Results. In: Journal of Economic Education, 44 (3), 298-309.
- Weber, B. (2005): Bildungsstandards, Qualifikationserwartungen und Kerncurricula: Stand und Entwicklungsperspektiven der ökonomischen Bildung. In: Weitz, B. O. (Hg.): Standards in der ökonomischen Bildung, Bergisch Gladbach, 17-50.

- Weinert, F. E. (2001): Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In: Weinert, F. E. (Hg.): Leistungsmessungen in Schulen, Weinheim/Basel, 17-32.
- Witzel, A. (2000): Das problemzentrierte Interview. Online: <http://www.qualitative-research.net/index.php/fqs/article/view/1132/2519> (07.07.2016).
- Witzel, A./Reiter, H. (2012): The Problem-centered Interview. Principles and Practice, London.
- Würth, R./Klein, H. J. (2001): Wirtschaftswissen Jugendlicher in Baden-Württemberg. Eine empirische Untersuchung, Künzelsau.