

Kompetenzmodellierungen zur wirtschaftlichen Fachkompetenz angehender Bankkaufleute

KURZFASSUNG: Bisher liegen nur wenige Arbeiten zu Strukturen und Niveaus beruflicher Kompetenzen vor. Insbesondere mangelt es an Untersuchungen, die Veränderungen von Kompetenzen im Verlauf der Berufsausbildung beleuchten. Daher werden, anknüpfend an die in der Literatur vorfindbaren Kompetenzmodellierungen für die gewerblich-technische und kaufmännische Berufsausbildung, Längsschnittdaten zur wirtschaftlichen Fachkompetenz angehender Bankkaufleute (n = 452) mittels IRT-basierter Analysen und Strukturgleichungsmodellen ausgewertet. Die Zunahme von sowohl allgemein- als auch bankwirtschaftlicher Kompetenz im Ausbildungsverlauf führt zu einer geringeren Besetzung des untersten von vier Kompetenzniveaus, wohingegen die beiden oberen Niveaus Zuwächse verzeichnen. Bezüglich der Strukturierung wirtschaftlicher Fachkompetenz erweist sich neben der Unterscheidung allgemein- und bankwirtschaftlicher Kompetenz eine Dimensionierung nach Fallsituationen, die Anforderungen nach typischen Situationen des Bankalltags bündeln, als empirisch am tragfähigsten. Im Verlauf der Ausbildung werden die Zusammenhänge zwischen den Kompetenzdimensionen stärker, sodass die einzelnen Auszubildenden zum Ende der Ausbildung relativ ähnliche Leistungen in den Teilbereichen zeigen.

ABSTRACT: So far, only a few studies have been conducted on dimensions and levels of vocational competence. In particular, there is a lack of studies on changes of competences during vocational education. Therefore, taking up existing competence models for industrial-technical and commercial vocational education, longitudinal data on the economic competence of prospective bankers (n = 452) are analysed using IRT-based methods and structural equation modelling. The growth of both general and banking economic competence leads to a smaller number of trainees on the lowest of four competence levels, whereas the two upper levels record gains. In addition to the distinction of general and banking economic competence a differentiation according to situations, which bundle up the requirements of typical everyday bank situations, proves empirically viable. Correlations increase between the competence dimensions during the training, so that each trainee shows relatively similar performance in the subdimensions at the end of the training.

Einleitung

Untersuchungen zur Kompetenzmodellierung sind in der bundesdeutschen Berufsbildung noch nicht sehr zahlreich. Im Wesentlichen kann hier auf die Arbeiten der Gruppen um Nickolaus (GEISSEL 2008; GSCHWENDTNER 2008; NICKOLAUS/GSCHWENDTNER/GEISSEL 2008), Seeber (SEEGER 2008) und Winther (WINTHER/ACHTENHAGEN 2009) verwiesen werden. Wenngleich der Kompetenzbegriff von diesen Arbeitsgruppen unterschiedlich scharf umrissen wird, lassen sich die operationalisierten Konstrukte unter ein gegenwärtig verbreitetes Verständnis von Kompetenz als kontextspezifische kognitive Leistungsdispositionen (KLIEME/LEUTNER 2006, S. 4) subsumieren, wie es auch im Folgenden zugrunde gelegt wird¹.

1 Bei diesem engen Begriffsverständnis entspricht Kompetenz dem klassischen Fähigkeits- oder Fertigkeitensbegriff bezogen auf berufsspezifische Klassen von Anforderungssituationen. Legt man

Während die Forschungsgruppe um Nikolaus Kompetenzmodellierungen für die Grundbildung in zwei gewerblich-technischen Ausbildungsberufen präsentiert (Kfz-Mechatroniker, Elektroniker), fokussieren Seeber und Winther kaufmännische Bildungsgänge (Büro-, Industriekaufleute). Hieran anschließend werden im vorliegenden Beitrag Ergebnisse aus einer Studie mit angehenden Bankkaufleuten berichtet, in deren Rahmen wirtschaftliche Fachkompetenz mehrmals im Verlauf der 2–3jährigen Ausbildung getestet wurde (Beginn, Mitte, Ende). Neben der Analyse von Veränderungen in allgemein- sowie bankwirtschaftlicher Kompetenz und der Ermittlung von Kompetenzniveaus wird die Dimensionalität der erhobenen wirtschaftlichen Kompetenzen exploriert.

Struktur beruflicher Kompetenzen

Die Strukturierung von Kompetenzen orientiert sich an typischen Anforderungen, die zu bewältigen sind (HARTIG/KLIEME 2006, S. 131). Nach welchen Kriterien Anforderungen klassifiziert und Aufgabenbündel zu Teilkompetenzen geschnürt werden, obliegt den Intentionen und (theoretisch begründeten) Erwägungen der Forschenden. Inwiefern die solchermaßen getroffene Zuordnung der Aufgaben zu postulierten (latenten) Konstrukten empirisch tragfähig ist, lässt sich anhand von statistischen Modellvergleichen ermitteln. So basierte die Konstruktion von Kompetenztests für angehende Kfz-Mechatroniker und Elektroniker für Energie- und Gebäudetechnik am Ende der beruflichen Grundbildung auf einem kognitionstheoretischen Gedächtnismodell von FORTMÜLLER (1997), der zwischen deklarativem und prozeduralem Wissen unterscheidet (GEISSEL 2008; GSCHWENDTNER 2008). Gemäß den Ergebnissen von Dimensionsanalysen nach der Item-Response-Theorie konstituieren diese Wissensarten jedoch erwartungswidrig keine unterscheidbaren Fähigkeiten. Vielmehr weisen die berichteten Deviance-Werte und die hohe Korrelation der Dimensionen deklaratives und prozedurales Wissen ($r = .97$, messfehlerbereinigt) bei den Kfz-Mechatronikern darauf hin, dass die Wissensarten Bestandteil einer übergreifenden Leistungsdisposition sind und somit diese Differenzierung für berufsspezifische Modellierungen wenig hilfreich ist (GSCHWENDTNER 2008). Ebenso ließ sich die Bewältigung von Items zu fahrzeugtechnischen und elektrotechnischen Ausbildungsinhalten, aus denen der Test im Ausbildungsberuf Kfz-Mechatroniker zum Ende der einjährigen Grundbildung bestand, nicht auf unterschiedliche Fähigkeiten zurückführen ($r = .90$, messfehlerbereinigt). Da eine explorative Faktorenanalyse für den Pretest eine stabile mehrdimensionale Struktur erbracht hatte, reduzierte sich vermutlich die anfängliche Mehrdimensionalität im Verlauf der Ausbildung (ebd., S. 105f.).

Im kaufmännischen Bereich ergaben sich abweichende Befunde. Die von SEEBER (2008) vorgelegten Analysen basieren auf dem Fachleistungstest für Bürokaufleute aus der Studie ‚Untersuchungen von Leistungen, Motivation und Einstellungen der Schülerinnen und Schüler in den Abschlussklassen der Berufsschulen – ULME III‘ (LEHMANN/SEEGER 2007). Ziel dieser Studie war es, Lernergebnisse in 17 Ausbildungsberufen am Ende der beruflichen Ausbildung zu erheben, wobei vor allem die im

stattdessen ein breiteres Kompetenzverständnis zugrunde, wie es in den Bestimmungen der Kultusministerkonferenz (KMK 1996, 2007) oder von WEINERT (2001) zum Ausdruck kommt, umfasst Kompetenz über die Befähigung zur Anforderungsbewältigung hinaus zusätzliche Komponenten wie die Bereitschaft zur Handlung und Werterwägungen (vgl. STRAKA/MACKE 2010).

schulischen Kontext erworbenen beruflichen Fähigkeiten bzw. Fachkompetenzen im Vordergrund standen. Bei den angehenden Bürokaufleuten erwiesen sich fachwissenschaftliche Kategorien (‚volkswirtschaftliche Aspekte‘, ‚betriebswirtschaftliche Organisation und Leistungsprozesse‘, ‚rechtliche Aspekte‘, ‚Rechnungswesen/Controlling‘) als eine Grundlage dimensional unterscheidbarer Teilkompetenzen.

WINTHER und ACHTENHAGEN (2009) prüften an einer Stichprobe angehender Industriekaufleute ein ‚Kompetenzstrukturmodell für die kaufmännische Bildung‘, das eine Ebene mit domänenverbundenen (Economic Literacy, Economic Numeracy) und domänenspezifischen (Geschäftsvorfälle) *Inhalten* sowie eine Ebene von verstehens- und handlungsbasierten *Zugriffen* unterscheidet. Dimensionsanalysen der computergestützt erhobenen Daten bestätigten die Differenzierung von latenten Fähigkeiten je nach Geschäftsvorfall (Einkauf, Vertrieb, Arbeitsvorbereitung) und im Hinblick auf eine Trennung von verstehensbasierter und handlungsbasierter Kompetenz.

Somit gelangen im gewerblich-technischen Bereich am Ende der Grundbildung weder Differenzierungen nach Wissensarten (deklarativ vs. prozedural) noch nach Inhaltsbereichen (fahrzeugtechnisch vs. elektrotechnisch). Vielmehr ergaben sich Hinweise auf Veränderungen hin zu einer Verschmelzung von anfänglich unterscheidbaren Fähigkeitsdimensionen. Im kaufmännischen Bereich ließen sich hingegen berufliche Kompetenzen am Ende der Ausbildung nach inhaltlichen Kriterien klassifizieren (fachwissenschaftliche Kategorien, Geschäftsvorfälle).

Kompetenzniveaus

Die Unterscheidung von Kompetenzniveaus dient der einfacheren Kommunikation und Veranschaulichung der erfassten quantitativen Leistungsdimensionen (HARTIG/KLIEME 2006). Es werden Aussagen angestrebt, über welche spezifischen Kompetenzen Auszubildende verfügen bzw. welche fachbezogenen Leistungsanforderungen sie zu bewältigen in der Lage sind. Da nicht für jeden Punkt der Skala spezifische Anforderungen bestimmt werden können, wird das Skalenkontinuum in Abschnitte unterteilt, die dann als Kompetenzniveaus oder Kompetenzstufen bezeichnet werden.

Die von den oben genannten Arbeitsgruppen vorgelegten Modelle differenzieren durchgängig vier Kompetenzniveaus und weisen trotz unterschiedlicher Berufe, verschiedener Systematik bei der Niveauermittlung und jeweils anderen Beschreibungsmerkmalen gewisse Gemeinsamkeiten auf, was eine zusammenfassende Beschreibung ermöglicht: Die Niveaus I und II umfassen Aufgaben, die bewältigt werden können durch direkten Abruf beruflichen Basiswissens und/oder durch Verarbeitung von Informationen, die in der Aufgabenbeschreibung oder weiteren Arbeitsunterlagen leicht auffindbar und durch die Berufsausbildung bereits vertraut sind, wobei die Verarbeitung dieser Informationen geringen kognitiven Aufwand, etwa den Einsatz routinierter Prozeduren, erfordert. Während auf Niveau I häufiger ein Alltagsverständnis zur Aufgabenlösung ausreicht, sind die zu verarbeitenden Informationen auf Niveau II berufsspezifischer, teilweise in den Arbeitsunterlagen weniger leicht zu finden, implizit kodiert und/oder es müssen mehr Informationen verarbeitet werden. Ab Niveau III gelingt die Bewältigung von Aufgaben, zu deren Lösung eine Verknüpfung beruflicher Konzepte/Begriffe, oft unter Berücksichtigung situationsspezifischer Randbedingungen, zu bewerten oder zu modellieren ist, was ein tieferes Verständnis der thematisierten beruflichen Sachverhalte voraussetzt.

Während die Inhalte auf Niveau III den Auszubildenden vertraut sind, unterscheiden sich Aufgaben des Niveaus IV vor allem durch die geringere Bekanntheit und den größeren Umfang der in den entsprechenden Aufgaben formulierten Anforderungen. Neben den beschriebenen Gemeinsamkeiten ist die prozentuale Verteilung der Auszubildenden auf den Kompetenzniveaus ähnlich (Tab. 1).

Tab. 1: Verteilung auf die Niveaus

		<Niveau I	Niveau I	Niveau II	Niveau III	Niveau IV
Kfz-Mechatroniker*	Vollzeit	–	57 %	42 %	1 %	0 %
	Teilzeit	–	13 %	80 %	6 %	1 %
Elektroniker*	Vollzeit	–	61 %	33 %	6 %	0 %
	Teilzeit	–	38 %	35 %	26 %	1 %
Bürokaufleute*		–	13 %	60 %	22 %	5 %
Industrie-kaufleute**	Verstehen	36 %	43 %	21 %	0 %	0 %
	Handeln	3 %	20 %	53 %	23 %	2 %

* Item-Lösungswahrscheinlichkeit 65 %; ** Item-Lösungswahrscheinlichkeit 50 %

Es ergeben sich überwiegend rechtsschiefe Verteilungen. Eine Ausnahme bildet die Dimension der handlungsbasierten Kompetenz bei Industriekaufleuten, deren Verteilung symmetrisch ausfällt und in der die Auszubildenden vermehrt ein höheres Niveau aufweisen, was möglicherweise den Vorteil einer nicht überwiegend schulisch vermittelten handlungsbasierten Kompetenz in der dualen Berufsausbildung widerspiegelt (WINTHER/ACHTENHAGEN 2009). Niveau IV wird nur selten erreicht, es ist überwiegend minimal (1–2 %) oder überhaupt nicht besetzt, die größte Häufigkeit bei zukünftigen Bürokaufleuten beträgt gerade einmal 5 %, wobei dieser vergleichsweise hohe Wert vermutlich auf das Verfahren zur Bestimmung der Schwellen zurückzuführen ist. Bei der von SEEGER (2008) verwendeten, an BEATON und ALLEN (1992) angelehnten Vorgehensweise werden die Schwellen nicht anhand einer Regressionsgleichung mit den schwierigkeitsbestimmenden Aufgabenmerkmalen festgelegt (HARTIG 2007), sondern relativ willkürlich gesetzt.

Ziel der Untersuchung

Vor dem Hintergrund der geschilderten Befundlage für andere Ausbildungsberufe werden weitergehende Analysen zur wirtschaftlichen Fachkompetenz angehender Bankkaufleute durchgeführt. Hierfür wird wirtschaftliche Fachkompetenz in eine allgemeinwirtschaftliche und eine bankwirtschaftliche Dimension differenziert. Allgemeinwirtschaftliche Kompetenz im Sinne einer ‚wirtschaftskundlichen Grundbildung‘ (BECK/KRUMM 1999) ermöglicht ökonomiespezifische Denkleistungen auf der Grundlage wirtschaftskundlichen Wissens. Bankwirtschaftliche Kompetenz umfasst kognitive Leistungsdispositionen, die erforderlich sind zur Bewältigung domänenspezifischer beruflicher Anforderungen, wie sie sich in typischen bankwirtschaftlichen Situationen stellen (z.B. Führen von Konten, Anlageberatung). Ausgehend von dieser

Dimensionierung erfolgt eine post hoc geführte Untersuchung der Veränderungen, Niveaus und Strukturen wirtschaftlicher Fachkompetenz von Auszubildenden zum/zur Bankkaufmann/-frau.

Anlage der Untersuchung

Ablauf und Stichprobe

Die Erhebung der für die folgenden Auswertungen herangezogenen Daten erfolgte im Rahmen eines umfangreicheren DFG-Projekts² in den Jahren 2006–2009. In seinem Verlauf wurden neben demographischen Angaben und kognitiven Eingangsbedingungen zwei- bzw. dreimal während der 2–3-jährigen Ausbildung allgemein- und bankwirtschaftliche Kompetenzen erhoben (Abb. 1).

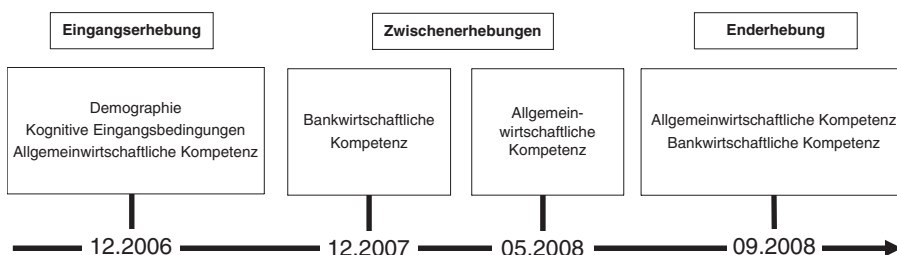


Abb. 1: Untersuchungsdesign³

Die Analysestichprobe umfasst Daten von 452 Auszubildenden in 22 Klassen aus 9 Schulen (Baden-Württemberg, Bremen, Niedersachsen, Rheinland-Pfalz), die sich auf Anfrage bereit erklärt hatten, an der Studie teilzunehmen. Es beteiligten sich 59 % weibliche und 41 % männliche Auszubildende, was gut mit Angaben des Bundesinstituts für Berufsbildung (BIBB 2008) übereinstimmt, nach denen im Jahr 2006 der Anteil weiblicher Personen bei den neu abgeschlossenen Verträgen für die Ausbildung zum/zur Bankkaufmann/-frau 58,3 % betrug. Die Auszubildenden waren zum Zeitpunkt der ersten Befragung auf Jahresbasis durchschnittlich 19,4 Jahre alt (BIBB 2008: 19,7 Jahre)⁴. 72 % der Auszubildenden begannen die Ausbildung nach Erwerb der (Fach)Hochschulreife (BIBB 2008: 61,4 %), 22 % besaßen einen Realschulabschluss (BIBB 2008: 27,5 %) und 6 % hatten sonstige Schulabschlüsse (erweiterter Realschulabschluss, Abschluss der Höheren Handelsschule, theoretischer Teil der Fachhochschulreife) (BIBB 2008: sonstige Abschlüsse oder keine Angaben 11,1 %). Abgesehen von einigen Abweichungen bezüglich der Schulabschlüsse (Realschule –5,5 %, Abitur/Fachabitur +10,6 %, Sonstiges –5,1 %) ist die

2 Projekt: Kontextuelle und personale Bedingungen von domänenspezifischer beruflicher Fachkompetenz bei angehenden Bankkaufleuten (DFG: STR 266/18-1/2)

3 Dargestellt ist das Design für Auszubildende mit 2,5-jähriger Ausbildungsdauer. Bei den Auszubildenden mit 2- und 3-jähriger Vertragslaufzeit erfolgten die letzten Erhebungen zeitlich verschoben.

4 Tag und Monat des Geburtsdatums wurden nicht erhoben, um Anonymität zu gewährleisten

Stichprobe im Bezug auf die betrachteten demographischen Variablen weitgehend repräsentativ.

Instrumente

Zur Kontrolle kognitiver Eingangsbedingungen wurden die kognitive Grundfähigkeit (WONDERLIC 1992, $c_{\alpha} = .73$), Mathematische Kompetenz (LEHMANN et al. 2004, $c_{\alpha} = .65$) und das Leseverständnis (MACGINITIE et al. 2002, $c_{\alpha} = .54$ ⁵) erhoben. Die Erfassung wirtschaftlicher Fachkompetenz erfolgte entsprechend der Differenzierung in eine allgemeinwirtschaftliche und eine bankwirtschaftliche Dimension über zwei Verfahren.

Allgemeinwirtschaftliche Kompetenz

Allgemeinwirtschaftliche Kompetenz wurde mit dem standardisierten ‚Wirtschaftskundlichen Bildungs-Test‘ (WBT, BECK/KRUMM, 1999) operationalisiert. Dieser Test liegt in zwei Formen (A, B) mit 46 Multiple-Choice-Items vor, von denen jeweils eine zu Beginn und im Verlauf der Ausbildung vorgegeben wurde. Die beiden originalen Paralleltestformen sind über 15 gemeinsame Ankeritems miteinander verknüpft⁶. Am Ende der Ausbildung kam eine 23-Item-Variante mit ausgewählten Aufgaben aus beiden Testformen zum Einsatz, wobei drei unterschiedliche Versionen (a, b, c) verwendet wurden, die neben 19 gemeinsamen jeweils 4 spezifische Items enthielten. Die Bewältigung der Testaufgaben erfordert in Anlehnung an BLOOM (1956) die kognitiven Prozesse des Kennens, Verstehens, Anwendens, Analysierens und Bewertens in den Themenbereichen ökonomische Grundlagen (z.B. Knappheit), Makroökonomie (z.B. Inflation und Deflation), Mikroökonomie (z.B. Markt und Preis) und internationale Beziehungen (z.B. Zahlungsbilanz und Devisenkurse).

Bankwirtschaftliche Kompetenz

Die in der Zwischen- und Abschlusserhebung eingesetzten Tests zur bankwirtschaftlichen Kompetenz wurden, anknüpfend an frühere Arbeiten der Arbeitsgruppe (BÖHNER 2005; BÖHNER/STRAKA 2005), neu entwickelt⁷ und umfassen je fünf Fälle, welche sich auf bankwirtschaftliche Situationen beziehen:

- *Zwischenerhebung*: Kontoeröffnung, Kontoführung, Kontoführung im Todesfall, Anlageberatung, Privatkredit
- *Abschlusserhebung*: Kontoführung und Zahlungsverkehr, Kontoführung im Todesfall, Anlageberatung, Privatkredit, Baufinanzierung

In beiden Versionen sind 22 (Teil)Aufgaben mit gleicher Formulierung enthalten (Ankeritems). Die bankwirtschaftlichen Situationen sollten den Auszubildenden zu den jeweiligen Testzeitpunkten aus der betrieblichen Ausbildungspraxis bekannt sein und die dazu gestellten Aufgabeninhalte wurden mit dem zeitlichen Verlauf

5 Die Übertragung des Instruments ins Deutsche und die Anpassung an die Zielgruppe erfolgte unter der Leitung von Dipl.-Gwl. Tobias Gschwendtner (vgl. GSCHWENDTNER, in Vorbereitung).

6 Aufgrund teilweise veralteter Aufgabenstellungen wurde der Test zur wirtschaftskundlichen Grundbildung geringfügig modifiziert, sodass die Testversionen A und B für die Analysen jeweils 45 Aufgaben und 14 gemeinsame Ankeritems umfassten.

7 Die Testentwicklung erfolgte durch Dipl.-Hdl. Gritt Fehring.

des Ausbildungsrahmenplans und des Rahmenlehrplans abgestimmt, sodass curriculare Validität gewährleistet ist. Den verschiedenen Fällen zugeordnet sind Multiple-Choice-, Zuordnungs-, Mehrfachwahl- und offene (Teil)Aufgaben (Zwischen-erhebung: 43⁸, Abschlusserhebung: 48), mit denen die fachbezogenen Kompetenzen der Auszubildenden erfasst werden, welche für die Bearbeitung der in den Fällen gestellten bankwirtschaftlichen Problemstellungen erforderlich sind.

Entsprechend der Skalenstatistiken (Tab. 2) sind die Testdaten für die weiteren Analysen gut geeignet.

Tab. 2: Kennwerte der Tests

Test	n	MW*	SD	Reliabilität (innere Konsistenz)
Allgemeinwirtschaftliche Kompetenz I	410	55 %	11,5	.67
Allgemeinwirtschaftliche Kompetenz II	404	60 %	11,8	.72
Allgemeinwirtschaftliche Kompetenz III	394	63 %	14,6	.60 ^a /.62 ^b /.63 ^c
Bankfachkompetenz II	412	53 %	11,5	.66
Bankfachkompetenz III	402	61 %	14,2	.81

* Durchschnittlicher Prozentsatz richtig gelöster Aufgaben

Befunde

Zur Untersuchung von Kompetenzstrukturen und Kompetenzniveaus wird in der Berufsbildungsforschung testtheoretisch überwiegend das Rasch-Modell aus der Gruppe der Item-Response-Theorien zugrunde gelegt. Während die Ermittlung von Kompetenzniveaus mit dem Ziel einer kriteriumsorientierten Interpretation die Zugrundelegung einer Messtheorie wie das Rasch-Modell erfordert, können Dimensionsanalysen zur Kompetenzstruktur auch durchgeführt werden mittels Strukturgleichungsmodellen, die aussagekräftigere Modellstatistiken erbringen (SEEBER 2008) und weitergehende Hinweise über die in den Daten enthaltenen Zusammenhänge liefern. Somit erscheint eine Kombination beider Methoden vorteilhaft. Zunächst wurde anhand von Mixed-Rasch-Modellen geprüft, inwiefern die Annahme des Rasch-Modells für die Kompetenztests zur allgemeinen und bankwirtschaftlichen Kompetenz haltbar ist.

Modellgeltung

Das Mixed-Rasch-Modell stellt eine Kombination aus latenter Klassenanalyse und Raschmodellierung dar (GOLLWITZER 2007). Mithilfe der latenten Klassenanalyse können die Itemantworten dahingehend untersucht werden, ob die Teilnehmer in gleicher Weise auf die Items reagieren, oder ob die auftretenden Antwortmuster

8 Für die Analysen wurden drei Items wegen negativer Trennschärfen ausgeschlossen.

wahrscheinlicher sind unter Annahme der Existenz mehrerer Klassen bzw. Gruppen, die sich im Hinblick auf die Bearbeitung der Items unterscheiden. Ursachen für das Vorliegen solcher Gruppenunterschiede können verschiedene Strategien bei der Itembearbeitung sein. Ist die Annahme der Ein-Klassen-Lösung nicht zu verwerfen, spricht dies dafür, dass der Bearbeitung der Items eine einheitliche Fähigkeit bzw. ein einheitliches Fähigkeitsbündel zugrunde liegt, was eine Voraussetzung für die Gültigkeit des Rasch-Modells darstellt. Beim Mixed-Rasch-Modell wird neben der Prüfung auf eine oder mehrere Klassen gleichzeitig die Annahme getestet, dass in den Gruppen das Rasch-Modell gilt.

Für die Entscheidung, ob ein Ein-Klassen-Rasch-Modell angenommen werden kann oder möglicherweise Mehr-Klassen-Rasch-Modelle angemessener sind, dienen inferenzstatistische Maße der Modellgüte sowie sogenannte Informationskriterien. Auf der Grundlage einer Simulationsstudie empfiehlt VON DAVIER (1997), zur inferenzstatistischen Modellabsicherung mittels Bootstrap-Verfahren die χ^2 - und die Cressie-Read-Statistik zu verwenden.

Sind mehrere Klassenlösungen inferenzstatistisch haltbar, wird zum Modellvergleich zurückgegriffen auf die Likelihood der Daten, die zum Ausdruck bringt, wie gut es gelingt, mit den (geschätzten bzw. errechneten) Modellparametern genau jene Daten (Antwortmuster) zu erhalten, die man tatsächlich beobachtet hat. Da sich die Likelihood der Daten einerseits mit zunehmender Stichprobengröße ungünstig entwickelt und weil sie andererseits für mehrklassige Modelle trivialerweise besser ausfällt, obwohl diese weniger sparsam und somit aus wissenschaftstheoretischer Perspektive ungünstiger sind (Parsimonitätsprinzip), berücksichtigen Informationskriterien als Gütemaße der Modellanpassung von (Mixed-)Rasch- und Strukturgleichungsmodellen zusätzlich die Anzahl der Modellparameter und die Stichprobengröße. Während das AIC (Akaike Information Criterion) lediglich eine höhere Anzahl von Modellparametern (Dimensionen) „bestraft“, bevorzugt das CAIC (Consistent AIC) sparsamere Modelle noch stärker und berücksichtigt zusätzlich die Stichprobengröße. Somit empfiehlt sich das CAIC für Modellvergleiche, wenn eine möglichst sparsame Modellierung angestrebt wird (z.B. Annahme des Ein-Klassen-Rasch-Modells), das AIC hingegen ist sensitiver im Hinblick auf eine bessere Modellanpassung und erlaubt, die Angemessenheit einer (theoretisch) begründeten Modelldifferenzierung empirisch zu prüfen (z.B. Annahmen zur Dimensionalität). Generell sind solche Modelle zu präferieren, die zum einen inferenzstatistisch nicht signifikant werden und zum anderen möglichst niedrige Werte in dem jeweils verwendeten Informationskriterium (AIC, CAIC) aufweisen.

Für die Tests zu allgemein- und bankwirtschaftlichen Kompetenzen (AWK, BWK) ist gemäß den in Tabelle 3 angeführten p-Werten (p_{χ^2} , p_{CR}) und angesichts der jeweils niedrigeren CAIC-Werte die Annahme des Ein-Klassen-Rasch-Modells sowohl zur Mitte (II) als auch am Ende (III) der Ausbildung haltbar⁹.

Wenngleich die Ein-Klassen-Modelle bis auf eine Ausnahme (BWK II: $p_{CR} = .04$) nicht signifikant und somit inferenzstatistisch unkritisch sind, fallen die entsprechenden Werte für die Zwei-Klassen-Modelle teilweise deutlich besser aus und zudem treten dort keine kritischen Q-Index-Werte auf. Beim Q-Index handelt es sich um einen Trennschärfeindex, der unter der Annahme, dass das Rasch-Modell gilt, die

9 Da es sich bei den Tests zur Allgemeinwirtschaftlichen Kompetenz I und II um Parallelformen handelt, werden im Folgenden nur die Ergebnisse für den Test AWK II berichtet.

Tab. 3: Mixed-Rasch-Modelle der Kompetenztests¹⁰

	Modell	Inferenzstatistik ¹¹		CAIC	Q-Index
AWK II	Ein-Klassen-Modell	$p_{\chi^2} = .12$	$p_{CR} = .06$	20 660	4*, 1**
	Zwei-Klassen-Modell	$p_{\chi^2} = .39$	$p_{CR} = .21$	21 139	
AWK III	Ein-Klassen-Modell	$p_{\chi^2} = .30$	$p_{CR} = .11$	8 722	2*, 1**
	Zwei-Klassen-Modell	$p_{\chi^2} = .88$	$p_{CR} = .94$	8 867	
BWK II	Ein-Klassen-Modell	$p_{\chi^2} = .12$	$p_{CR} = .04$	18 946	6*, 2**
	Zwei-Klassen-Modell	$p_{\chi^2} = .29$	$p_{CR} = .17$	19 322	
BWK III	Ein-Klassen-Modell	$p_{\chi^2} = .19$	$p_{CR} = .08$	21 591	5*, 8**
	Zwei-Klassen-Modell	$p_{\chi^2} = .15$	$p_{CR} = .08$	22 533	

* Anzahl Items mit Q-Index $p < .10$ (zweiseitig),

** Anzahl Items mit Q-Index $p < .02$ (zweiseitig)

Wahrscheinlichkeit des beobachteten Antwortmusters mit der Wahrscheinlichkeit von Antwortmustern bei perfekt positiver und negativer Trennschärfe vergleicht (BÜHNER 2011, S. 543). Die Berechnungen ergeben beispielsweise für 6 Items des BWK II auffällige ($p < .10$, zweiseitig) und für 2 Items signifikante ($p < .02$, zweiseitig) Abweichungen von einer günstigen Trennschärfe im Sinne des Rasch-Modells. Besonders augenfällig sind die Werte für die allgemeinwirtschaftliche Kompetenz am Ende der Ausbildung, denen zufolge ein Zwei-Klassen-Rasch-Modell deutlich besser auf die Daten passt (AWK III, $p_{\chi^2} = .88$, $p_{CR} = .94$). Bei näherer Betrachtung dieser Zwei-Klassen-Lösung erweist sich, dass Auszubildende der prozentual größeren Gruppe (63%) im Durchschnitt 13.3 Aufgaben richtig lösen, Auszubildende der kleineren Gruppe (37%) hingegen 9.8 Aufgaben (Effektstärke des Mittelwertunterschieds: $d = 1.4^{12}$). Die prozentualen Anteile sind in den 22 einbezogenen Schulklassen weitgehend ähnlich. Auszubildende der leistungsstärkeren Klasse haben höhere Werte in kognitiver Grundfähigkeit ($d = 0.69$), im Leseverständnis ($d = 0.46$), in mathematischen Fähigkeiten ($d = 0.59$) und in der bankwirtschaftlichen Fachkompetenz am Ende der Ausbildung ($d = 0.81$). Es sind vermehrt männliche Auszubildende (Kontingenzkoeffizient: $.17$, $p < .01$) und Personen mit Hochschulreife ($.26$, $p < .000$). Bei Betrachtung der Items, deren Schwierigkeitsparameter in beiden Gruppen deutlich voneinander abweichen ($\delta_{K1} - \delta_{K2} > 1.90$), ist zu erkennen, dass Aufgaben, deren Lösung genaues und sorgfältiges Lesen erfordert, wenn sich beispielsweise Antwortoptionen nur in Nuancen und im Grad ihrer Richtigkeit unterscheiden, der leistungsstärkeren Gruppe relativ zu den anderen Items schwerer fallen, wohingegen Items, die in erster Linie schlussfolgerndes Denken mit leicht

10 Es wurden mit der Software WINIMIRA (VON DAVIER 2001) jeweils Modelle mit einer bis fünf Klassen berechnet. Die Drei- bis Fünf-Klassen-Modelle sind im Folgenden nicht aufgeführt, da bereits die Ein-Klassen-Modelle unkritische inferenzstatistische und die jeweils niedrigsten CAIC-Werte aufweisen.

11 Ergebnisse der Bootstrap-Prozedur mit 100 samples

12 Nach COHEN (1988, S. 25) gelten Effektstärken ab 0,5 als „mittel“ und ab 0,8 als „hoch“.

verfügbaren Informationen verlangen, für sie deutlich leichter sind. Einige Aufgaben, die Begriffswissen testen, werden von beiden Gruppen ähnlich häufig gelöst, sodass diese Items für die leistungsstärkere Klasse verhältnismäßig schwerer sind.

Die angeführten Phänomene verweisen auf Unterschiede bei der Aufgabenbearbeitung in Abhängigkeit von allgemeinen Komponenten der kognitiven Leistungsfähigkeit. Die verschiedenen Herangehensweisen betreffen jedoch nicht spezifisch wirtschaftliche Kompetenzen, weshalb eine Differenzierung in zwei Klassen für die nachfolgenden Analysen wenig Erkenntnisgewinn verspricht. Daher wird die Annahme des Ein-Klassen-Rasch-Modells, welche bereits durch die in Tabelle 3 aufgeführten inferenzstatistischen Werte und die Ausprägungen des Informationskriteriums nahe gelegt wurde, aufrecht erhalten.

Skalierung

Zur weiteren Beschreibung wurden die Tests nach dem eindimensionalen Rasch-Modell mit der Software ConQuest (WU/ADAMS/HALDANE 2005) skaliert. Durch die Logit-Transformation können die solcherart ermittelten Parameter der Personenfähigkeit und Itemschwierigkeit prinzipiell von plus bis minus unendlich reichen, in der Regel liegen sie zwischen plus und minus drei (Abb. 2).

Ein Vorteil der Rasch-Modellierung ist die Verortung von Personenfähigkeiten und Itemschwierigkeiten auf einer gemeinsamen Skala, wobei der Mittelwert der Itemschwierigkeiten im vorliegenden Fall den Nullpunkt markiert. Während die ‚X‘ in Abbildung 2 die Verteilungen der Personenparameter veranschaulichen, zeigen die Aufgabennummern (rechts von den X) die Position der jeweiligen Items. Eine Person, deren Fähigkeit (θ) einer bestimmten Itemschwierigkeit (δ) entspricht, löst das jeweilige Item mit einer Wahrscheinlichkeit von 50 %. Je höher die Differenz zwischen Personen- und Itemparameter ausfällt, umso größer ist die Wahrscheinlichkeit, dass eine Person das Item richtig löst. Die Aufgaben 7, 8, 18, 36 (AWK II) sowie 21 und 26 (AWK III) sind demnach leicht und werden von der Mehrzahl der Auszubildenden zutreffend bearbeitet, die Aufgaben 23 und 25 (AWK II) sind hingegen schwer. Die Trennschärfen der Items sind bis auf zwei Ausnahmen signifikant von Null verschieden und durchweg positiv. An den Weighted-Mean-Square-(MNSQ)-Werten ist ablesbar, inwiefern bei Gültigkeit des Rasch-Modells erwartete und empirisch beobachtete Antwortmusterhäufigkeiten eines Items voneinander abweichen, wobei der Erwartungswert für einen idealen Fit bei Eins liegt. Da die zugehörige Signifikanzprüfung nicht zuletzt von der Stichprobengröße abhängt, werden häufig die Absolutwerte zur Beurteilung des Item-Fits herangezogen. Für beide Tests liegen sie innerhalb des Bereichs von 0,8–1,2, was für Multiple-Choice-Tests akzeptable Werte bedeutet (BOND/Fox 2007, S. 243). Die WLE-Reliabilitäten entsprechen den klassischen internen Konsistenzen (Tab. 2) und sind im Test zur allgemeinwirtschaftlichen Kompetenz III (AWK III) niedriger als zuvor (AWK II), da in der Enderhebung Testversionen mit halbiertem Itemanzahl eingesetzt wurden¹³.

13 Da jeweils vier Items rotiert zum Einsatz kamen, wurden im AWK III insgesamt 31 unterschiedliche Aufgaben bearbeitet (vgl. Abb. 2).

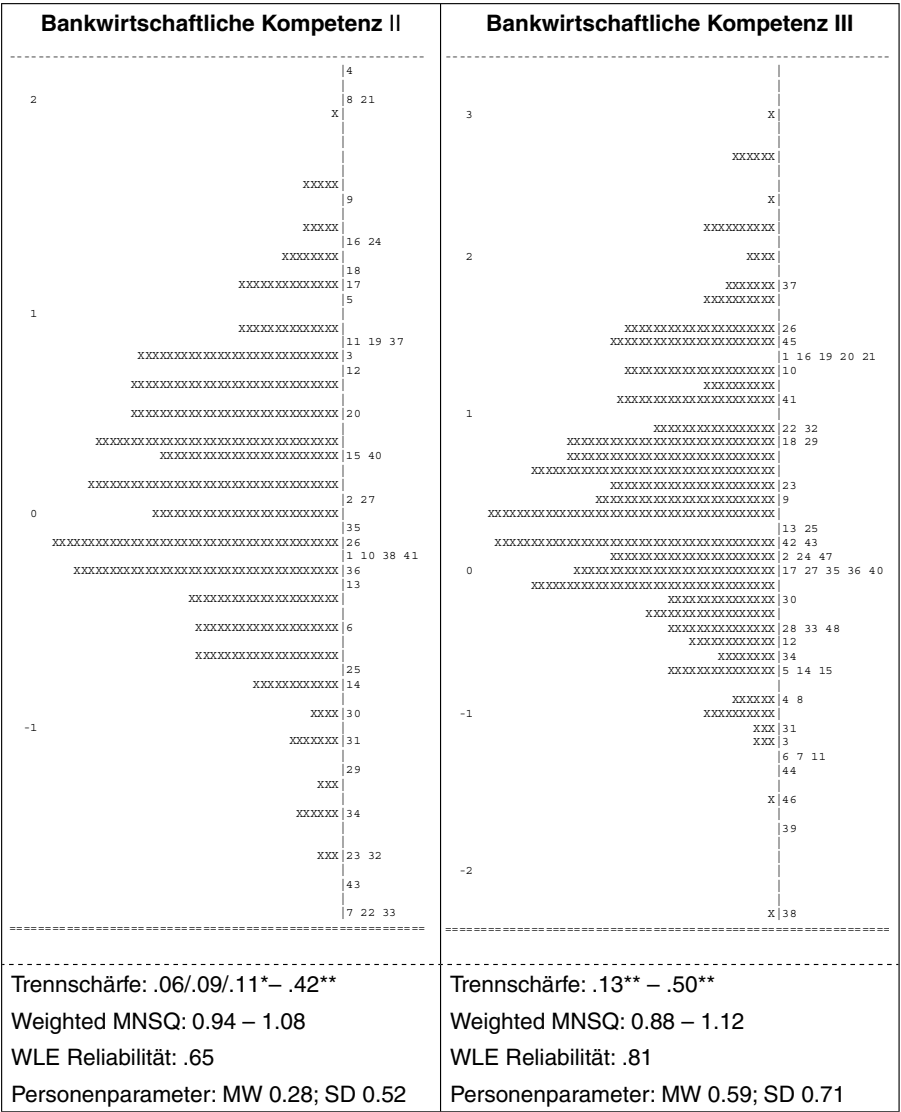
Allgemeinwirtschaftliche Kompetenz II	Allgemeinwirtschaftliche Kompetenz III
<pre> 25 23 XXXX XXXXXXXXXX 2 XXXXXXXXXX XXXXXXXXXX XXXXXX XXXXXXXXXXXXXXXXXX 38 43 3 22 5 26 XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXX 1 XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXX 4 20 33 XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXX 10 46 27 41 XXXXXXXXXXXXXXXXXXXX 28 14 31 XXXXXXXXXXXXXXXXXXXX 6 24 34 39 0 XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXX 17 42 44 40 XXXXXXXXXXXXXXXXXXXX 21 XXXXXXXXXXXXXXXXXXXX 1 16 XXXXXXXXXXXX 29 37 XXXXXX 19 45 X 35 -1 XXXXXX 11 30 X 13 15 XXXX 12 9 7 8 18 36 ===== </pre>	<pre> X X 3 XXXXXX X XXXXXXXXXX 2 XXXXXXXXXXXX XXXX XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXX 1 30 20 XXXXXX 17 19 24 1 23 15 XXXXXXXXXXXXXXXXXXXX 18 XXXXXXXXXXXXXXXXXXXX 31 XXXXXXXXXXXXXXXXXXXX 8 27 XXXXXXXXXXXXXXXXXXXX 12 XXXXXXXXXXXXXXXXXXXX 4 XXXXXXXXXXXXXXXXXXXX 11 2 29 28 0 XXXXXXXXXXXXXXXXXXXX 22 XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXX 5 XXXXXXXXXXXXXXXXXXXX 7 XXXX 13 X 6 XXXXXXXXXXXX X 14 10 XXX XXX 9 16 25 3 XXXX 21 26 ===== </pre>
Trennschärfe: .10* – .47** Weighted MNSQ: 0.90 – 1.07 WLE Reliabilität: .71 Personenparameter: MW 0.64; SD 0.57	Trennschärfe: .02/.17/.17* – .50** Weighted MNSQ: 0.89 – 1.13 WLE Reliabilität: .64 Personenparameter: MW 0.72; SD 0.62

* $p < .05$, ** $p < .01$

Abb. 2: Skalierung allgemeinwirtschaftlicher Kompetenz (WLE-Werte)

Auch für die bankwirtschaftliche Kompetenz (Abb. 3) liegen die Personenparameter im Durchschnitt über dem mittleren Itemparameter (MW 0.28 bzw. 0.59). Die Trennschärfen zweier Items zur bankwirtschaftlichen Kompetenz III sind nicht signifikant von Null verschieden. Der Item-Fit liegt innerhalb der konventionellen Grenzen (0.8–1.2), die Reliabilitäten ähneln wiederum den klassischen internen Konsisten-

zen (Tab. 2). Nicht alle Abschnitte der Skalen sind mit Items besetzt, sodass die Leistungsfähigkeit der Auszubildenden nicht in allen Bereichen fein unterschieden werden kann.



* p < .05; ** p < .01

Abb. 3: Skalierung bankwirtschaftlicher Kompetenz (WLE-Werte)

Fazit

Die Tests zur wirtschaftlichen Fachkompetenz lassen sich gemäß dem eindimensionalen Rasch-Modell skalieren. Dabei weisen die einbezogenen Items keine allzu groben Verstöße gegen die Modellannahmen und überwiegend zweckmäßige Statistiken auf. Somit sind die Voraussetzungen für die folgenden Analysen zu Veränderungen der Kompetenzen im Ausbildungsverlauf und zu den Kompetenzniveaus erfüllt.

Veränderungen der Kompetenzen im Ausbildungsverlauf

Um Veränderungen über die Zeit ermitteln zu können, wurden die Tests über die enthaltenen Ankeritems miteinander verknüpft. Anhand des Imports der Ankeritemparameter eines Zeitpunkts in die Skalierung eines anderen Zeitpunkts lassen sich durch die Verortung von Item- und Personenparametern auf einer gemeinsamen Skala Veränderungen in den Personenfähigkeiten nachweisen. Da bei Gültigkeit des Rasch-Modells die Itemparameter unabhängig sind von der Stichprobe, an der sie ermittelt werden, sollten die Parameter der Ankeritems zu beiden Zeitpunkten annähernd identisch ausfallen. Prüft man diese Annahme mit einer Prozedur nach WRIGHT und MASTERS (1982, S. 114f.), so ist sie für einen erheblichen Teil der Ankeritems nicht erfüllt (Tab. 4).

Tab. 4: Stabilität der Ankeritemparameter

Testpaarung	Enthaltene Ankeritems	Stabile Itemparameter
AWK I – AWK II	14	3
AWK I – AWK III	16	5
AWK II – AWK III	14	3
BWK II – BWK III	22	13

Die geringe Anzahl stabiler Itemparameter ist insbesondere für die nicht curricular geregelte allgemeinwirtschaftliche Kompetenz letztlich nicht überraschend, denn die Aufgabenanforderungen sind relativ heterogen und beziehen sich auf verschiedene Inhaltsbereiche, d.h. beim gemessenen Kompetenzkonstrukt handelt sich eher nicht um ein stabil homogenes Fähigkeitsbündel, wenngleich es sich zu den jeweiligen Zeitpunkten eindimensional skalieren lässt. Vielmehr deutet dies darauf hin, dass sich nicht nur quantitative, sondern auch qualitative Veränderungen in der Kompetenzstruktur vollziehen (vgl. MUSEKAMP/SPÖTTL/BECKER 2010). Beispielsweise könnten einige Items anfangs unlösbar sein, weil später gängige Fachbegriffe zunächst nicht bekannt sind. Am Ende der Ausbildung fallen diese Aufgaben möglicherweise sehr leicht im Vergleich zu anderen, zu deren Lösung man zwar ‚nur‘ ein Alltagsverständnis benötigt, die aber eine Verarbeitung und Verknüpfung mehrerer Informationen erfordern und daher zu beiden Testzeitpunkten mittlere Schwierigkeiten aufweisen. Auch Erinnerungs- und Übungseffekte spielen vermutlich eine Rolle: Wurden Inhalte im Berufsschulunterricht oder im Ausbildungsbetrieb gerade behandelt, fallen sie im aktuellen Test leichter, zu einem späteren Zeitpunkt ist das Erlernte möglicherweise bereits wieder vergessen oder wird überlagert, sodass das Item verhältnismäßig schwieriger wird. Beispielsweise ist für die Auszubildenden das korrekte Bilden

eines Buchungssatzes am Ende der Ausbildung schwieriger ($\delta = -0.9$) als in der Zwischenerhebung ($\delta = -1.8$). Des Weiteren können Aufgaben, die regelmäßig in Ausbildungsbetrieb oder Berufsschule zu bearbeiten sind, mit der Zeit verhältnismäßig leichter werden im Vergleich zu anfangs ähnlich schweren Aufgaben, die jedoch im Verlauf der Ausbildung seltener auftreten etc. Stabilität der Itemparameter ist demnach nicht prinzipiell zu erwarten. Dies widerspricht jedoch dem Gebot der Stichprobenunabhängigkeit, wenn man sie dahingehend versteht, dass das ermittelte Modell zumindest für die Population der angehenden Bankkaufleute gelten sollte – tatsächlich gilt es bestenfalls für den jeweiligen Messzeitpunkt. Selbst diese Annahme ist nicht unzweifelhaft, denn variierende Schwerpunktsetzungen bei den zu vermittelnden Inhalten könnten etwa je nach Berufsschulklasse oder Ausbildungsbetrieb zu unterschiedlichen Itemparametern führen. Da ein Modellvergleich auf Klassen- bzw. Betriebsebene aufgrund der geringen Anzahl an Auszubildenden für die vorliegende Studie nicht möglich ist, erscheint zukünftig eine Erfassung der Intensität der behandelten Ausbildungsinhalte angebracht, um Gruppen ähnlich ausgebildeter Personen zu identifizieren und zu prüfen, inwiefern ein Modell in den verschiedenen Gruppen gleichermaßen gilt (DIF-Analyse).

Verankert man ungeachtet der skizzierten Problematik die Testzeitpunkte, indem man die verbleibenden stabilen Itemparameter in die Skalierung anderer Testzeitpunkte importiert und solchermaßen einen einheitlichen Maßstab bildet, ergeben sich nachvollziehbare Veränderungen. Zur Veranschaulichung wurden für die jeweils erste Testung die Logit-Werte der Personenfähigkeiten auf Skalenwerte mit dem Mittelwert 100 und einer Standardabweichung von 25 transformiert (Abb. 4).

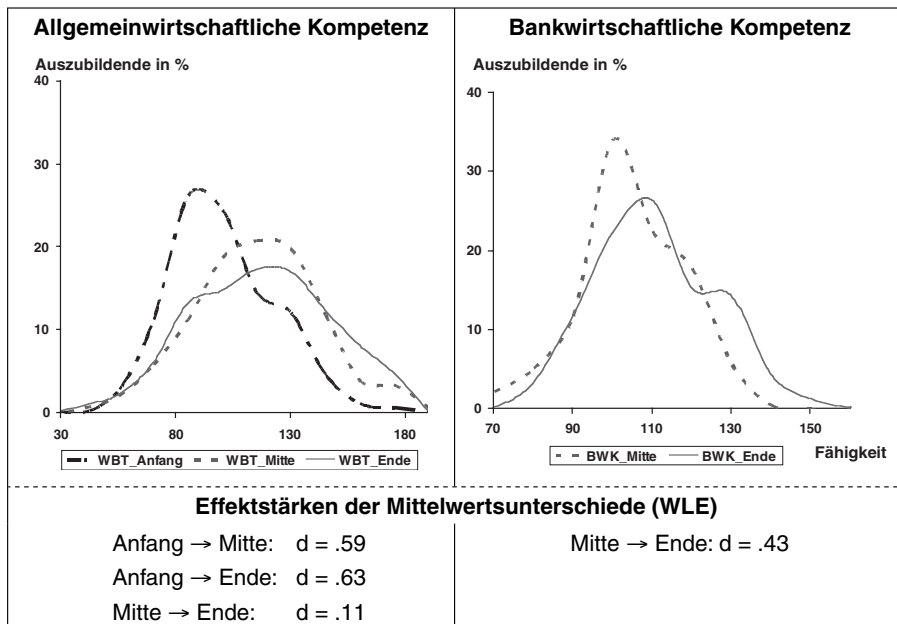


Abb. 4: Veränderungen der Kompetenz

Bei der bankwirtschaftlichen Kompetenz zeigt sich von der Zwischen- zur Enderhebung der Ausbildung ein Kompetenzzuwachs, wenngleich die Effektstärke nicht groß ist ($d < 0.5$). Demnach werden die Aufgaben zum Ende der Ausbildung geringfügig besser bewältigt. Der in diesem Zeitraum verhältnismäßig stärkere Zuwachs der bankwirtschaftlichen Kompetenz im Vergleich zur allgemeinwirtschaftlichen Kompetenz ($d = .43$ vs. $d = .11$), deren Aufgaben nicht umfassend curricular abgesichert sind, unterstreicht die Bedeutung und Effektivität einer gezielten Vermittlung und Aneignung von Ausbildungsinhalten. Da die Verteilungen im Verlauf der Ausbildung flacher und breiter werden, wachsen die Leistungsunterschiede zwischen den Auszubildenden. Während die Standardabweichung für die bankwirtschaftliche Kompetenz von 25 (BWK II) auf 29 (BWK III) zunimmt, steigt sie in der allgemeinwirtschaftlichen Kompetenz von 25 (AWK I) über 28 (AWK II) auf 34 (AWK III).

Kompetenzniveaus

Die Vorgehensweise zur Unterteilung des Skalenkontinuums kann auf unterschiedlichen Wegen erfolgen. Für die vorliegende Untersuchung wurde ein Verfahren angelehnt an BEATON und ALLEN (1992) gewählt, da im Vorfeld keine Erwägungen über schwierigkeitsbestimmende Merkmale getroffen worden waren, wie es das stärker theoriegeleitete Vorgehen nach HARTIG (2007) erfordert. Für die Bestimmung und Beschreibung von Niveaus werden daher nicht die Gesamtheit der Testaufgaben, sondern charakteristische Aufgaben herangezogen, die von Auszubildenden des entsprechenden Niveaus mit einer hinreichenden Wahrscheinlichkeit (65%) und von Auszubildenden an der unteren Schwelle des nächst niedrigeren Niveaus mit deutlich geringerer Wahrscheinlichkeit ($\leq 50\%$) gelöst werden. Die Setzung der Kategoriengrenzen erfolgte nach Inspektion der in einer Rangreihe angeordneten Itemparameter, wobei gleiche Abstände zwischen den mittleren Klassengrenzen angestrebt wurden ([95 – 115 – 135]). Um eine hinreichende Zahl an Markieritems zu gewährleisten, wurde für die Ermittlung der Kompetenzniveaus neben dem Test zur bankwirtschaftlichen Kompetenz III der Test zur allgemeinwirtschaftlichen Kompetenz II herangezogen, da er im Gegensatz zur am Ende eingesetzten Kurzversion mehr Items enthält und die Kompetenzentwicklung zwischen diesen Zeitpunkten nur geringe Effektstärke aufwies (s.o.).

Die Verteilung der angehenden Bankkaufleute auf die Kompetenzniveaus (Abb. 5) ähnelt der von Bürokaufleuten zum Ende der Ausbildung (Niveau I: 13%, II: 60%, III: 22%, IV: 5%, SEEGER 2008). Von der Zwischen- zur Enderhebung ergeben sich in der bankwirtschaftlichen Kompetenz besonders auf dem untersten Niveau deutliche Veränderungen, der Prozentsatz von Auszubildenden halbiert sich in etwa (29% vs. 15%). Während Niveau II, das ungefähr die Hälfte der Auszubildenden umfasst, prozentual nahezu unverändert bleibt, verzeichnen Niveau III (24% vs. 30%) und IV (0% vs. 7%) erkennbare Zuwächse. Ähnliche Tendenzen lassen sich für die allgemeinwirtschaftliche Kompetenz, hier jedoch von der Eingangs- zur Zwischenenerhebung, nachweisen.

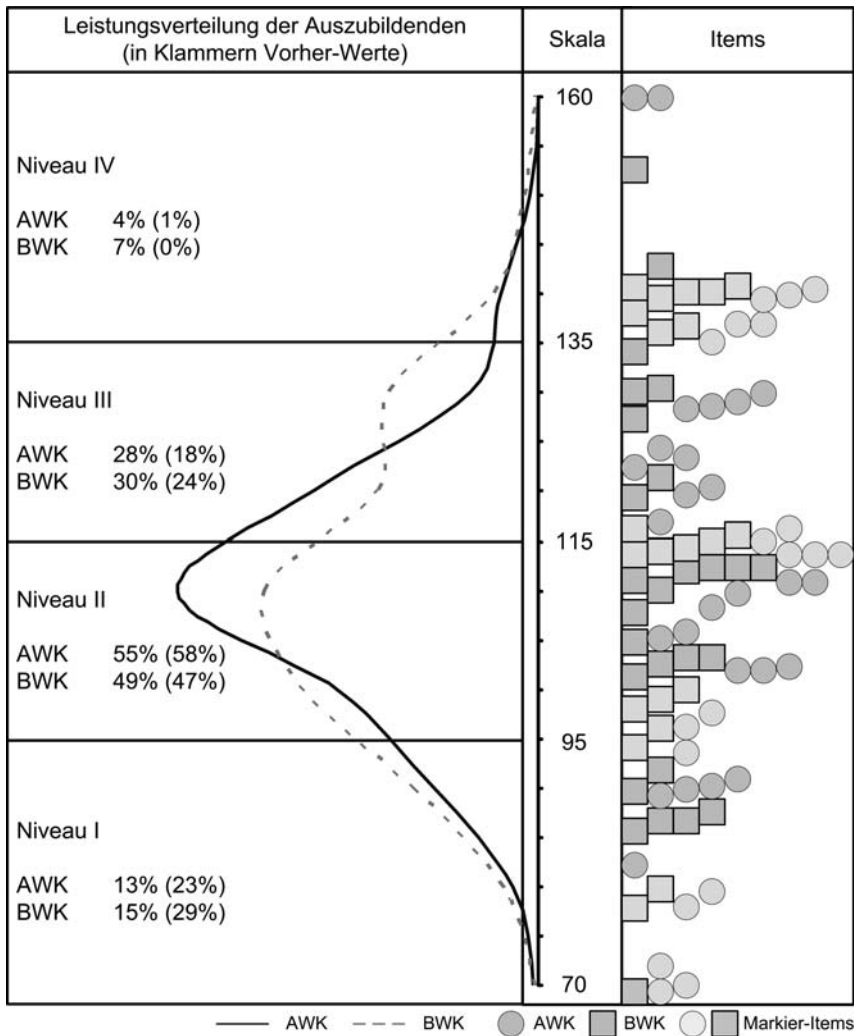


Abb. 5: Niveaus der Bankfachkompetenz

Beschreibung der Kompetenzniveaus

Niveau I

AWK: Auszubildende kennen grundlegende wirtschaftliche Prinzipien (z.B. Wettbewerb als Prinzip der Marktwirtschaft) und Begrifflichkeiten (z.B. Gewinne sind Erlöse minus Kosten)

BWK: Auszubildende verfügen über Begriffswissen (z.B. Ertragswert ist der kapitalisierte Jahresbruttoertrag einer Immobilie) und kennen rechtliche Regeln (z.B. Widerrufsfrist von Darlehensverträgen) aus jüngst behandelten Lernfeldern (z.B. Privatkredit, Baufinanzierung).

Niveau II:

AWK: Auszubildende können die Effekte einfacher wirtschaftlicher Mechanismen bestimmen (geldpolitische Maßnahmen, Wettbewerb) und kennen wirtschaftliche Fachbegriffe (z. B. Realeinkommen).

BWK: Auszubildende kennen wirtschaftliche und rechtliche Aspekte von gängigen Wertpapieren und Finanzprodukten (Scheck, Festgeldkonto). Sie sind in der Lage, einfache Berechnungen durchzuführen und Buchungssätze zu bilden.

Niveau III:

AWK: Auszubildenden gelingt die sichere Bestimmung des jeweils richtig beschriebenen Effekts von wirtschaftlichen Mechanismen, die mehrere Wirkungen haben können (Arbeitsteilung, Steuern, Geldangebot).

BWK: Auszubildende verfügen über Detailwissen zu wirtschaftlichen und rechtlichen Aspekten in häufig bearbeiteten Bereichen (Finanzprodukte: Wandelanleihen, Aktien, Festgeldkonto; Grundschuld).

Niveau IV:

AWK: Auszubildende kennen spezielle (volks)wirtschaftliche Begriffe (Bruttosozialprodukt, Opportunitätskosten), sie können die richtigen Folgen von volkswirtschaftlichen Maßnahmen bei Vorgabe ähnlicher Distraktoren sicher bestimmen (Einfuhrzölle und Handelsbilanzdefizit, Inflationsrate).

BWK: Auszubildende sind zu umfangreicheren und schwierigeren Berechnungen in der Lage (z. B. anbietbarer Höchstzinssatz). Sie verfügen über sicheres Detailwissen z. B. von Finanzprodukten (Bundesanleihen, Schuldverschreibung), sodass ihnen auch bei ähnlichen Distraktoren die Bestimmung der richtigen Lösungen gelingt.

Kompetenzdimensionen

Im Vorfeld der Untersuchung wurde eine Differenzierung wirtschaftlicher Fachkompetenz in eine allgemeinwirtschaftliche und eine bankwirtschaftliche Dimension postuliert. Gemäß der oben angeführten Analysen lassen sich diese Teilkompetenzen bei Zugrundelegung konventioneller Kriterien jeweils eindimensional skalieren. Fraglich bleibt dabei einerseits, ob nicht die Annahme einer übergreifenden wirtschaftlichen Leistungsdisposition ausreicht. Andererseits wäre denkbar, dass eine noch feinere Unterteilung vor allem der bankwirtschaftlichen Kompetenz, etwa nach fachwissenschaftlichen Kategorien (vgl. SEEBER 2008) oder Fallsituationen/Geschäftsvorfällen (vgl. WINTHER/ACHTENHAGEN 2009), adäquater ist. Dies wird im Folgenden geprüft, wobei die ursprünglich nicht vorgesehenen Feindifferenzierungen zu einer teilweise geringen Itemzahl pro Dimension führen, sodass die Ergebnisse explorativen Charakter aufweisen.

Dimensionsanalysen nach Aspekten

Die von SEEBER (2008) für angehende Bürokaufleute empirisch bestätigte Differenzierung nach den fachwissenschaftlichen Kategorien ‚volkswirtschaftliche Aspekte‘, ‚betriebswirtschaftliche Organisation und Leistungsprozesse‘, ‚rechtliche Aspekte‘ und ‚Rechnungswesen/Controlling‘ ist auf die vorliegende Studie ansatzweise übertragbar. Die Abgrenzung eines volkswirtschaftlichen (im Folgenden: allge-

meinwirtschaftliche Kompetenz) und Rechnungswesen-Aspekts bereitet wenig Probleme. Eine Trennung von (betriebs)wirtschaftlichen und rechtlichen Aspekten ist insbesondere bei Aufgaben, die auf konkrete Finanzprodukte abzielen, zwar grundsätzlich möglich, diese Kategorien werden aber nicht getrennt vermittelt, sondern in den Lehrwerken gemeinsam behandelt, weshalb fraglich ist, inwiefern sich diese Differenzierung empirisch abbilden lässt. Daher wird das vierdimensionale Modell zum einen gegen ein dreidimensionales Modell getestet, das die rechtliche und wirtschaftliche Dimension zusammenfasst, und zum anderen gegen ein zweidimensionales Modell, in dem, wie ursprünglich vorgesehen, lediglich die allgemeinwirtschaftliche und bankwirtschaftliche Dimension unterschieden werden. Darüber hinaus wird das eindimensionale Generalfaktor-Modell berechnet (Abb. 6).

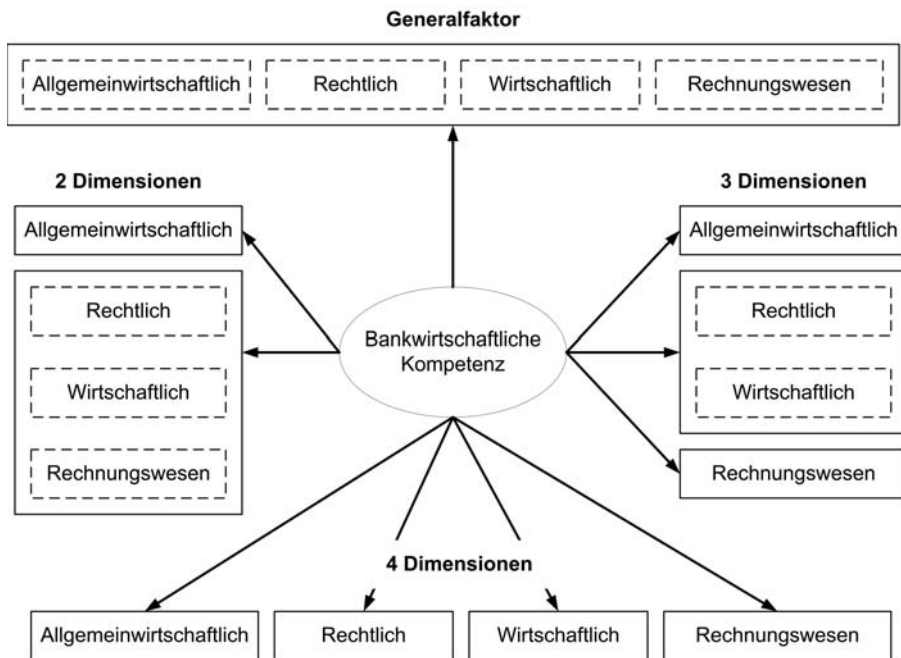


Abb. 6: Modelle der Dimensionsanalyse nach Aspekten

Zwischenerhebung

Entsprechend den in Tabelle 5 aufgeführten Deviance-Werten weisen die höherdimensionalen Modelle jeweils signifikant geringere Abweichungswerte und somit einen besseren Modell-Fit auf.

Die Trennung der beiden Aspekte rechtlich und wirtschaftlich verbessert den Fit nur geringfügig ($p < .05$). Dies spiegelt sich in der höchsten Korrelation dieser Dimensionen wider ($r = .77$, messfehlerbereinigt, Tab. 6).

In Ergänzung zu diesen Analysen, wie sie auch von den anfangs angeführten Arbeitsgruppen vorgelegt wurden, ist zusätzlich zu untersuchen, ob die vorgenommene Dimensionierung die Daten insgesamt gut erklärt. Dies ist mithilfe von Strukturgleichungsmodellen möglich, die im Gegensatz zu den IRT-basierten Analysen

Tab. 5: Vergleich der Modelle nach Aspekten in der Zwischenerhebung

Modell	Deviance	Parameter	Änderung Dev	Änderung Par	p-Niveau
4-dimensional	37 679	95	10	4	p < .05
3-dimensional	37 689	91	37	3	p < .000
2-dimensional	37 726	88	43	2	p < .000
1-dimensional	37 799	86			

Monte Carlo, 2000 nodes, n = 440

Tab. 6: Korrelationen der Aspekte (messfehlerbereinigt) in der Zwischenerhebung

	(1)	(2)	(3)	(4)	EAP/PV Rel.
Allgemeinwirtschaftlich	1.00	.69	.60	.22	.73
Wirtschaftlich	–	1.00	.77	.33	.64
Rechtlich	–	–	1.00	.34	.59
Rechnungswesen	–	–	–	1.00	.35

nicht nur Aussagen darüber zulassen, welches der Modelle die Daten vergleichsweise am besten erklärt, sondern auch Güteindizes für die grundsätzliche Passung des Modells im Bezug auf die empirischen Daten liefern. Zum anderen lässt sich anhand der AIC-Werte abschätzen, inwiefern die nachgewiesenen signifikanten Abnahmen der Deviance (Tab. 5) einem jeweils bedeutsam besseren Modell-Fit in den Strukturgleichungsmodellen entsprechen.

Beim Vergleich der vier Dimensionslösungen mittels Strukturgleichungsmodellen wird die Angemessenheit einer Differenzierung nach vier Aspekten aufgrund des geringsten AIC-Wertes bestätigt (Tab. 7). Allerdings unterschreiten die CFI-Werte (.49–.56) selbst für das bestangepasste Modell den konventionellen Cut-Off-Wert von CFI $\approx$.95 erheblich (BÜHNER 2011, S. 427) und dokumentieren somit einen insgesamt schlechten Modell-Fit.

Tab. 7: Modellstatistiken der Strukturgleichungsmodelle zu Aspekten in der Zwischenerhebung

Modell	AIC	χ^2/df	p-Wert	CFI	SRMR	RMSEA
4-dimensional	4 670	1.2	p < .000	.56	.05	.03
3-dimensional	4 672	1.2	p < .000	.56	.05	.03
2-dimensional	4 902	1.3	p < .000	.53	.05	.03
1-dimensional	4 976	1.3	p < .000	.49	.05	.03

n = 368¹⁴

Demnach stehen die Items über die durch die latenten Konstrukte erfasste Varianz hinaus in Beziehung miteinander, was darauf hinweist, dass möglicherweise zu den

14 Um aussagekräftige Güteindizes zu erhalten, wurden zur Berechnung der Strukturgleichungsmodelle Fälle mit fehlenden Werten ausgeschlossen.

fachwissenschaftlichen Aspekten quer liegende Dimensionen gemeinsame Varianz erzeugen. Bevor dem nachgegangen wird, erfolgt zunächst die Prüfung, ob sich die ermittelten Befunde für den Zwischentest im Bezug auf den Endtest bestätigen lassen.

Enderhebung

Während die Trennung von allgemeinwirtschaftlichen und bankwirtschaftlichen Aspekten wiederum hochsignifikant ausfällt, ist die weitergehende Differenzierung mit dem Aspekt des Rechnungswesens als dreidimensionale Lösung nicht nachweisbar besser (Tab. 8). Hingegen führt die Unterscheidung aller vier Aspekte wiederum zu einem signifikant geringeren Deviance-Wert.

Tab. 8: Vergleich der Modelle nach Aspekten in der Enderhebung

Modell	Deviance	Parameter	Änderung Dev	Änderung Par	p-Niveau
4-dimensional	31 912	89	15	4	$p < .01$
3-dimensional	31 927	85	5	3	$p = .17$
2-dimensional	31 932	82	79	2	$p < .000$
1-dimensional	32 011	80			

Monte Carlo, 2000 nodes, $n = 402$

Die sehr hohe Korrelation der rechtlichen und wirtschaftlichen Dimension ($r = .95$, messfehlerbereinigt, Tab. 9) lässt eine Unterscheidung dieser Aspekte trotz der hochsignifikant geringeren Deviance-Werte fragwürdig erscheinen.

Tab. 9: Korrelationen der Aspekte (messfehlerbereinigt) in der Enderhebung

	(1)	(2)	(3)	(4)	EAP/PV Rel.
Allgemeinwirtschaftlich	1.00	.65	.76	.60	.67
Wirtschaftlich	–	1.00	.95	.76	.75
Rechtlich	–	–	1.00	.83	.80
Rechnungswesen	–	–	–	1.00	.56

Diese Einschätzung wird bei Betrachtung des Informationskriteriums der Strukturgleichungsmodelle bestätigt, da das vierdimensionale gegenüber dem dreidimensionalen Modell einen höheren AIC-Wert aufweist (Tab. 10) und somit die empirischen Daten nicht bedeutsam besser erklärt.

Statistisch betrachtet verschmelzen die rechtliche und wirtschaftliche Dimension im Lauf der Ausbildung, was für eine zunehmende Verknüpfung der zugrunde liegenden kognitiven Strukturen spricht. Während die Unterscheidung im Zwischentest angesichts einer Korrelation von $r = .77$ (Tab. 6) und des geringfügig geringeren AIC-Wertes für das vierdimensionale Modell (Tab. 7) begründbar ist, erscheint sie für den Endtest kaum haltbar. Die CFI-Werte der Strukturgleichungsmodelle fallen am Ende zwar besser aus als zur Mitte der Ausbildung (54–.64 vs. .49–.56), jedoch sind sie weit von akzeptablen Werten entfernt. Alternativ wird im Folgenden geprüft, ob eine Unterscheidung nach Fallsituationen einen besseren Modell-Fit erbringt.

Tab. 10: Modellstatistiken der Strukturgleichungsmodelle zur Dimensionierung nach Aspekten t_3

Modell	AIC	χ^2/df	p-Wert	CFI	SRMR	RMSEA
4-dimensional	3 294	1.4	p < .000	.64	.06	.03
3-dimensional	3 291	1.4	p < .000	.64	.06	.03
2-dimensional	3 296	1.4	p < .000	.64	.06	.03
1-dimensional	3 513	1.5	p < .000	.54	.06	.04

n = 350

Fallsituationen

Eine Differenzierung nach Fallsituationen impliziert entgegen einer an fachwissenschaftlichen Aspekten angelehnten Kompetenzstruktur die Annahme, dass die kontextspezifischen kognitiven Leistungsdispositionen sich nach typischen Situationen ausbilden, wie es ansatzweise im Lernfeldgedanken zum Ausdruck kommt. Aufgrund der Konzeption der beiden bankwirtschaftlichen Kompetenztests zur Zwischen- und Enderhebung, die aus fünf situativen Einbettungen mit darauf bezogenen Aufgaben bestehen, konnte diese Annahme für die bankwirtschaftliche Kompetenz geprüft werden. Die Fallsituationen lassen sich thematisch Lernfeldern zuordnen, wenngleich sie breit angelegt sind und vereinzelt Aufgaben aus anderen Lernfeldern enthalten.

Zwischenerhebung

Im Test zur bankwirtschaftlichen Kompetenz zur Zwischenerhebung (BWK III) sind die Fallsituationen ‚Kontoeröffnung‘ (Lernfeld 2), ‚Anlageberatung‘ (Lernfeld 4), ‚Kontoführung‘ (Lernfeld 2), ‚Kontoführung im Todesfall‘ (Lernfeld 2) und ‚Privatkredit‘ (Lernfeld 7) enthalten.

Nachfolgend wird das eindimensionale Generalfaktormodell gegen das dreidimensionale Lernfeldmodell und das fünfdimensionale Fallsituationsmodell getestet (Abb. 7).

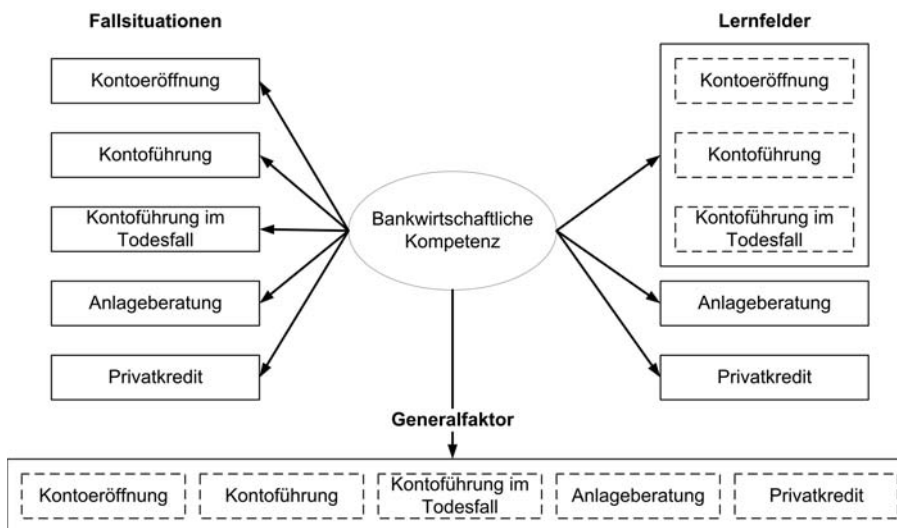


Abb. 7: Modelle der Dimensionsanalyse nach Fallsituationen in der Zwischenerhebung

Den geringsten Deviance-Wert und somit die beste Anpassung weist das fünfdimensionale Modell mit der Differenzierung nach den einzelnen Fallsituationen auf, wobei die Verbesserung in der Modellanpassung gegenüber dem dreidimensionalen Lernfeldmodell hochsignifikant ist (Tab. 11).

Tab. 11: Vergleich der Modelle nach Fallsituationen in der Zwischenerhebung

Modell	Deviance	Parameter	Änderung Dev	Änderung Par	p-Niveau
5-dimensional	17 700	55	29	9	$p < .000$
3-dimensional	17 729	46	19	5	$p < .01$
1-dimensional	17 748	41			

Monte-Carlo, 2000 nodes, $n = 404$

Das dreidimensionale Lernfeldmodell passt zwar besser als das eindimensionale Generalfaktor-Modell, doch zeigen die Korrelationen (Tab. 12), dass die höchsten Zusammenhänge erwartungswidrig nicht zwischen Fallsituationen bestehen, die demselben Lernfeld zuzuordnen sind.

Tab 12: Korrelationen der Fallsituationen (messfehlerbereinigt) in der Zwischenerhebung

	(1)	(2)	(3)	(4)	(5)	EAP/PV Rel.
Kontoeröffnung (LF 2)	1.00	.66	.76	.81	.81	.51
Kontoführung (LF 2)	–	1.00	.72	.66	.67	.57
Kontoführung im Todesfall (LF 2)	–	–	1.00	.70	.81	.60
Anlageberatung (LF 4)	–	–	–	1.00	.51	.53
Privatkredit (LF 7)	–	–	–	–	1.00	.51

Möglicherweise ist die relativ niedrige Korrelation zwischen den Fallsituationen Kontoeröffnung und Kontoführung ($r = .66$), die demselben Lernfeld zuzuordnen sind, auch darauf zurückzuführen, dass die situative Einbettung zur Kontoführung sich auf einen Vorfall mit Geschäftskunden bezieht, wohingegen die anderen Fallsituationen Vorfälle mit Privatkunden thematisieren.

Beim Vergleich der drei Dimensionslösungen mithilfe von Strukturgleichungsmodellen ist die Differenzierung nach Lernfeldern zwar dem eindimensionalen Modell vorzuziehen, doch angesichts des niedrigsten AIC-Wertes entspricht die fünfdimensionale Lösung nach Fallsituationen am ehesten den empirischen Daten (Tab. 13).

Tab. 13: Modellstatistiken der Strukturgleichungsmodelle nach Fallsituationen in der Zwischenerhebung

Modell	AIC	χ^2/df	p-Wert	CFI	SRMR	RMSEA
5-dimensional	1 082	1.2	$p < .000$	.77	.05	.02
3-dimensional	1 087	1.2	$p < .000$	.75	.05	.03
1-dimensional	1 095	1.3	$p < .000$	.73	.05	.03

$n = 368$

Enderhebung

Im Test zum Ende der Ausbildung sind die Fallsituationen ‚Kontoführung und Zahlungsverkehr‘ (Lernfeld 2), ‚Anlageberatung‘ (Lernfeld 4), ‚Kontoführung im Todesfall‘ (Lernfeld 2), ‚Privatkredit‘ (Lernfeld 7) und ‚Baufinanzierung‘ (Lernfeld 11) enthalten. Getestet wird das eindimensionale Generalmodell gegen das vierdimensionale Lernfeldmodell und das fünfdimensionale Fallsituationsmodell (Abb. 8).

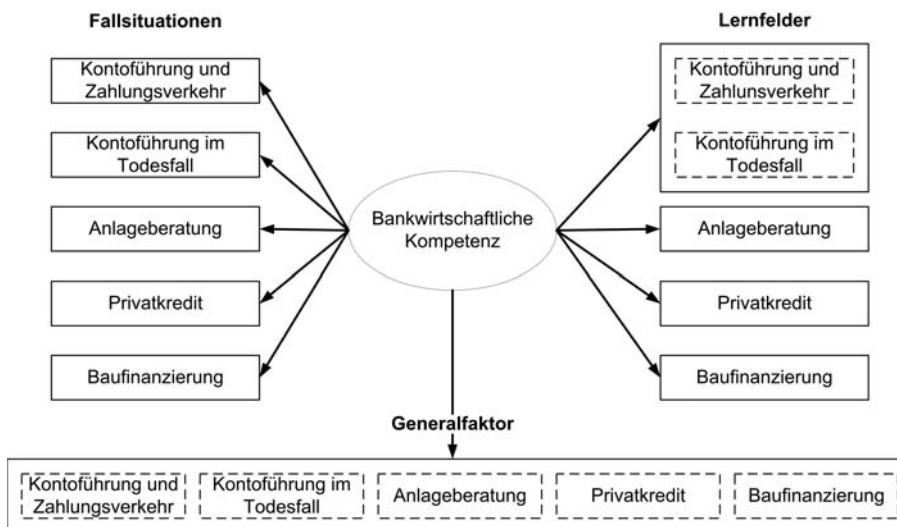


Abb. 8: Modelle der Dimensionsanalyse nach Fallsituationen in der Enderhebung

Es ergeben sich jeweils hochsignifikante Verbesserungen der Deviance-Werte für die höherdimensionalen Modelle (Tab. 14).

Tab. 14: Vergleich der Modelle nach Situationen vs. Lernfeldern t_3

Modell	Deviance	Parameter	Änderung Dev	Änderung Par	p-Niveau
5-dimensional	21 719	63	26	5	$p < .000$
4-dimensional	21 745	58	58	9	$p < .000$
1-dimensional	21 803	49			

Monte Carlo, 2000 nodes, $n = 402$

Wie in den Analysen zum Zwischentest fallen die Korrelationen zwischen Fallsituationen unterschiedlicher Lernfelder teilweise höher aus als zwischen den Fallsituationen ‚Kontoführung und Zahlungsverkehr‘ und ‚Kontoführung im Todesfall‘, die beide dem Lernfeld 2 zuzuordnen sind (Tab. 15).

Tab. 15: Korrelationen der Fallsituationen (messfehlerbereinigt) t_3

	(1)	(2)	(3)	(4)	(5)	EAP/PV Rel.
Kontoführung/Zahlungsverkehr (LF 2)	1.00	.72	.83	.88	.75	.76
Kontoführung im Todesfall (LF 2)	–	1.00	.93	.83	.58	.74
Anlageberatung (LF 4)	–	–	1.00	.93	.73	.80
Privatkredit (LF 7)	–	–	–	1.00	.89	.83
Baufinanzierung (LF 11)	–	–	–	–	1.00	.68

Die latenten Konstrukte korrelieren teilweise sehr hoch miteinander ($r > .90$). Dies spricht trotz der besseren Deviance-Werte, deren signifikante Abnahme nicht zuletzt von der Stichprobengröße abhängt, tendenziell für eine Zusammenfassung der betroffenen Dimensionen. Das anhand von Strukturgleichungsmodellen ermittelte Informationskriterium (AIC, Tab. 16) zeigt jedoch, dass sich die Differenzierung nach Fallsituationen empirisch abbilden lässt und den Modell-Fit durchgängig verbessert.

Tab. 16: Modellstatistiken der Strukturgleichungsmodelle nach Fallsituationen in der End-erhebung

Modell	AIC	χ^2/df	p-Wert	CFI	SRMR	RMSEA
5-dimensional	1 683	1.4	$p < .000$	.77	.05	.03
4-dimensional	1 705	1.4	$p < .000$	.75	.05	.03
1-dimensional	1 754	1.4	$p < .000$	.72	.06	.04

$n = 350$

Die hochsignifikanten p- und die nicht befriedigenden CFI-Werte deuten auch an dieser Stelle darauf hin, dass eine Mehrheit der Items verschiedener Fallsituationen untereinander über die durch die latenten Konstrukte erfasste Varianz hinaus in Beziehung zueinander stehen, d.h. eine Dimensionierung allein nach Fallsituationen erklärt die Zusammenhänge zwischen den Items nicht in befriedigendem Maße.

Fazit

Aus den Indizes für die Anpassungsgüte der Strukturgleichungsmodelle ist ersichtlich, dass das Fallsituationsmodell besser auf die Daten passt als die Dimensionierung nach fachwissenschaftlichen Aspekten. Dennoch erklärt auch diese Strukturierung die Daten nicht in zufriedenstellendem Maße. Die gemäß Lernfeldzuordnung zusammengehörigen Fallsituationen korrelieren nicht am höchsten miteinander, was tendenziell gegen eine ausschließlich an Lernfeldern orientierte Dimensionierung von bankwirtschaftlicher Kompetenz spricht. Die durchschnittliche Korrelation zwischen den Teilkompetenzen nimmt im Verlauf der Ausbildung zu, sodass sich die Leistung der einzelnen Auszubildenden in den verschiedenen Bereichen angleicht.

Diskussion

In der vorliegenden Arbeit werden Veränderungen, Niveaus und die Struktur wirtschaftlicher Fachkompetenz von Auszubildenden zum/zur Bankkaufmann/-frau untersucht. Sowohl allgemeinwirtschaftliche als auch bankwirtschaftliche Kompetenzen nehmen im Laufe der Berufsausbildung zu, wobei in der zweiten Ausbildungshälfte vor allem die curricular abgesicherte bankwirtschaftliche Kompetenz erkennbar wächst. Dies betrifft nicht nur die Breite (zusätzliche Ausbildungsinhalte), sondern auch das Kompetenzniveau, da am Ende der Ausbildung insbesondere auf dem untersten von vier Niveaus weniger Auszubildende verortet werden und die oberen Niveaus, welche ein erfolgreiches Agieren bei höheren kognitiven Anforderungen markieren, von mehr Auszubildenden erreicht werden als zuvor. Dies steht im Gegensatz zu Ergebnissen einer Querschnittstudie mit Auszubildenden des gewerblich-technischen Bereichs, in der keine Veränderung des Kompetenzniveaus zwischen dem zweiten und dritten Ausbildungsjahr ermittelt wurde (RAUNER/HEINEMANN/PIENING/MAURER 2009). Erklärt wird der fehlende Fortschritt mit dem Ausruhen der Auszubildenden nach der Zwischenprüfung, da die Enderhebung bereits 8–9 Monate vor der Abschlussprüfung stattfand (ebd., S. 119). Allerdings fußte die Veränderungsmessung in dieser Untersuchung auf nur jeweils zwei offenen Aufgaben¹⁵, wobei die angegebene Messgenauigkeit des herangezogenen Kompetenzkonstrukts, die anhand der Rater-gestützten Einschätzungen von 40 Kompetenzkriterien ermittelt wurde, sehr beeindruckend ausfällt ($\alpha = .97$).

Ein erheblicher Teil der Aufgaben zur allgemein- und bankwirtschaftlichen Kompetenz, die in den Tests unterschiedlicher Zeitpunkte gleichermaßen enthalten sind, um Kompetenzveränderungen ermitteln zu können, weist instabile Itemparameter auf. Dies lässt es sinnvoll erscheinen, Kompetenzentwicklungsmodelle für abgrenzbare Bereiche zu konzipieren, wobei Vergessenseffekte bzw. Lernzuwächse nicht zuletzt in Abhängigkeit von unterschiedlichen Schwerpunktsetzungen bei der Vermittlung von Inhalten zu berücksichtigen sind.

Die im Vorfeld getroffene Unterscheidung in allgemeinwirtschaftliche und bankwirtschaftliche Kompetenz, anstelle der Annahme einer übergreifenden wirtschaftlichen Leistungsdisposition, wird in Modellvergleichen empirisch bestätigt. Darüber hinaus wurden Modelle geprüft, die weitergehende Differenzierungen wirtschaftlicher Fachkompetenz vorsehen. Die hierzu durchgeführten post-hoc Analysen erfolgten auf der Grundlage von Instrumenten, welche unter anderer Zielsetzung entwickelt worden waren, sodass die nachträglich bestimmten Dimensionen teilweise nur gering mit Items besetzt waren. Die Ergebnisse sind daher vorsichtig zu interpretieren. Dennoch ist es gelungen, vorliegende Befunde für den kaufmännischen Bereich zu bestätigen, gemäß denen Teilkompetenzen nach fachwissenschaftlichen Kategorien unterschieden werden können (SEEBER 2008). Allerdings ist die anhand von Strukturgleichungsmodellen ermittelte Modellanpassung für diese Dimensionierung insgesamt unbefriedigend. Demgegenüber erbringt eine Strukturierung bankwirtschaftlicher Kompetenz nach Fallsituationen, die Anforderungen nach typischen Situationen des Bankalltags bündeln, einen besseren Modell-Fit. Dies entspricht einer Strukturierung von kaufmännischen Kompetenzen nach Geschäftsvorfällen (WINTHER/ACHTENHAGEN 2009) und lässt den

15 Eine kritische Auseinandersetzung mit dieser Operationalisierung von Kompetenz führen GSCHWENDTNER, ABELE und NICKOLAUS (2009).

Schluss zu, dass Modellen beruflicher Kompetenz Arbeits- bzw. Geschäftsprozesse zugrunde zu legen sind. Zudem erweist sich, dass eine Dimensionierung allein nach Lernfeldern, die nicht (zusätzlich) nach Arbeits- bzw. Geschäftsprozessen differenziert, das Gefüge bankwirtschaftlicher Kompetenz nicht optimal abbildet.

Einschränkend ist anzumerken, dass auch eine Strukturierung nach Fallsituationen die empirisch vorfindbaren Zusammenhänge zwischen den Aufgabenantworten nicht in befriedigendem Maße erklärt. Zukünftig ist daher zu prüfen, inwiefern Modelle, die eine kombinierte Dimensionierung nach gleichzeitig fachwissenschaftlichen Aspekten und Arbeits-/Geschäftsprozessen vorsehen, die Daten möglicherweise besser erklären als die bisherigen separat geführten Analysen (Within-Item-Multidimensionality). Dies setzt jedoch die Entwicklung passender Kompetenztests voraus, in denen die inhaltlichen Aspekte in den einzelnen Fallsituationen gleichermaßen repräsentiert sind. Aufgrund der Hinweise auf Bearbeitungsunterschiede in Abhängigkeit von unspezifischeren kognitiven Leistungsdispositionen, die bei der Berechnung von Mixed-Rasch-Modellen zutage treten, sind außerdem Modelle zu spezifizieren, welche solche Grundfähigkeiten als Generalfaktoren berücksichtigen.

Die im Verlauf der Ausbildung zunehmend stärkeren Zusammenhänge zwischen den Kompetenzdimensionen bestätigen Ergebnissen aus dem gewerblich-technischen Bereich (GSCHWENDTNER 2008). Dies kann verschiedene Ursachen haben, wobei sicherlich wünschenswert wäre, wenn die wachsenden Zusammenhänge zwischen den fachwissenschaftlichen Aspekten eine zunehmend stabile Verknüpfung kognitiver Konzepte widerspiegelt. Die ebenfalls stärker werdenden Korrelationen zwischen den Fallsituationen sprechen tendenziell gegen eine Spezialisierung der Auszubildenden auf bestimmte Gebiete, sondern deuten an, dass sich im Laufe der Zeit ein einheitliches individuelles Kompetenzniveau über die verschiedenen Situationen hinweg ausbildet.

Literatur

- BEATON, A. E. & ALLEN, N. L. (1992). Interpreting scales through scale anchoring. *Journal of Educational Statistics*, 17 (2), 191–204.
- BECK, K. & KRUMM, V. (1999). *Wirtschaftskundlicher Bildungstest*. Göttingen: Hogrefe.
- BIBB (2008). Datenblatt 691009 Bankkaufmann/-kauffrau. <http://berufe.bibb-service.de/Z/B/30/69100900.pdf>, abgerufen am 19.10.2010.
- BLOOM, B. (1956). *Taxonomy of educational objectives: Handbook I, The cognitive domain*. New York: David McKay & Co.
- BÖHNER, M. M. (2005). *Bankwirtschaftliche Kompetenz. Theoretische Grundlegung, Entwicklung und Validierung eines standardisierten Erfassungsinstruments*. Hamburg: Dr. Kova.
- BÖHNER, M. & STRAKA, G. A. (2005). Bankwirtschaftliche Kompetenz – Konzept und standardisierte Erfassung. In *bwp@*, Ausgabe Nr. 8.
- BOND, T. G. & FOX, C. M. (2007). *Applying The Rasch Model. Fundamental Measurement in the Human Sciences (Second Edition)*. Mahwah, New Jersey, London: Lawrence Erlbaum Associates.
- BÜHNER, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion (3. Auflage)*. München u. a.: Pearson Studium.
- COHEN, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Earlbaum Associates.
- DAVIER, VON M. (1997). Bootstrapping Goodness-of-Fit Statistics for Sparse Categorical Data – Results of a Monte Carlo Study. *Methods of Psychological Research Online*, Vol. 2 (2), 29–48.
- DAVIER, VON M. (2001). WINMIRA 2001.0012. <http://winmira.von-davier.de>

- 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: Manz-Verlag Schulbuch.
- GEISSEL, B. (2008). Ein Kompetenzmodell für die elektrotechnische Grundbildung: Kriteriumsorientierte Interpretation von Leistungsdaten. In R. Nickolaus & H. Schanz (Hrsg.), *Didaktik der gewerblich-technischen Berufsbildung. Diskussion Berufsbildung*, Band 9, 121–141. Baltmannsweiler: Schneider.
- GOLLWITZER, M. (2007). Latent-Class-Analysis. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion*. Heidelberg: Springer Medizin.
- GSCHWENDTNER, T. (2008). Ein Kompetenzmodell für die kraftfahrzeugtechnische Grundbildung; In R. Nickolaus & H. Schanz (Hrsg.), *Didaktik der gewerblich-technischen Berufsbildung. Diskussion Berufsbildung*, Band 9, 103–119. Baltmannsweiler: Schneider.
- GSCHWENDTNER, T. (in Vorbereitung): Lesekompetenzförderung in Benachteiligtenklassen der beruflichen Bildung. Eine empirische Untersuchung zur praktischen Bedeutsamkeit von reciprocal teaching. Aachen: Shaker, Stuttgart, Univ., Diss. (Stuttgarter Beiträge zur Berufs- und Wirtschaftspädagogik).
- GSCHWENDTNER, T., ABELE, S. & NICKOLAUS, R. (2009). Computersimulierte Arbeitsproben: Eine Validierungsstudie am Beispiel der Fehlerdiagnoseleistungen von Kfz-Mechatronikern.
- HARTIG, J. (2007). Skalierung und Definition von Kompetenzniveaus; In B. Beck & K. Klieme (Hrsg.), *Sprachliche Kompetenzen. Konzepte und Messung. DESI-Ergebnisse Band 1*, 83–99. Beltz: Weinheim.
- HARTIG, J. & KLIEME, E. (2006). Kompetenz und Kompetenzdiagnostik. In K. Schweizer (Hrsg.), *Leistung und Leistungsdiagnostik*, 127–143. Heidelberg: Springer Medizin.
- KLIEME, E. & LEUTNER, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Überarbeitete Fassung des Antrags an die DFG auf Einrichtung eines Schwerpunktprogramms. <http://kompetenzmodelle.dipf.de/pdf/rahmenantrag> (Zugriff am 19.10.2010).
- KULTUSMINISTERKONFERENZ (1996, 2007). Handreichungen für die Erarbeitung von Rahmenplänen der KMK für den berufsbezogenen Unterricht in der Berufsschule und ihre Abstimmung mit Ausbildungsordnungen des Bundes für anerkannte Ausbildungsberufe. Sekretariat der ständigen Konferenz der Kultusminister der Länder in der BRD.
- LEHMANN, R. H., HUNGER, S., IVANOV, S., GÄNSFUSS, R. & HOFFMANN, E. (2004). LAU 11. Aspekte der Lernausgangslage und der Lernentwicklung Klassenstufe 11. Hamburg.
- LEHMANN, R. H. & SEEBER, S. (2007). ULME III. Untersuchung von Leistungen, Motivation und Einstellungen der Schülerinnen und Schüler in den Abschlussklassen der Berufsschulen. Behörde für Bildung und Sport der freien und Hansestadt Hamburg, Hamburger Institut für berufliche Bildung (HIBB) (Hrsg.).
- MACGINITIE, W. H., MACGINITIE, R. K., MARIA, K. & DREYER, L. G. (2002). *Gates-MacGinitie Reading Tests, 4th Edition (GMRT-4)*. Technical Report for Forms S & T. Itasca, IL: River-side Publishing.
- MUSEKAMP, F., SPÖTTL, G. & BECKER, M. (2010). Schriftliche Arbeitsaufträge zur Erfassung von Differenzen in der Expertise von Facharbeitern und Auszubildenden. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 106 (3), 336–360.
- NICKOLAUS, R., GSCHWENDTNER, T. & GEISSEL, B. (2008). Entwicklung und Modellierung beruflicher Fachkompetenz in der gewerblich-technischen Grundbildung. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 104 (1), 48–73.
- RAUNER, F., HEINEMANN, L., PIENING, D. & MAURER, A. (2009). Die Ergebnisse der Pilotphase. In F. Rauner et al., *Messen beruflicher Kompetenzen. Ergebnisse KOMET 2008*, Band II, 107–159. Berlin: LIT.
- SEEBER, S. (2008). Ansätze zur Modellierung beruflicher Fachkompetenz in kaufmännischen Ausbildungsberufen. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 104 (1), 74–97.

- STRAKA, G. A. & MACKE, G. (2010). Kompetenz – nur eine „kontextspezifische kognitive Leistungsdisposition“?. Zeitschrift für Berufs- und Wirtschaftspädagogik, 106 (3), 444–451.
- WEINERT, F. E. (2001). Concept of competence: A conceptual clarification. In D. S. RYCHEN & L. H. SALGANIK (Hrsg.), Defining and selecting key competencies. Seattle: Hogrefe & Huber, 45–65.
- WINTHER, E. & ACHTENHAGEN, F. (2009). Skalen und Stufen kaufmännischer Kompetenz. Zeitschrift für Berufs- und Wirtschaftspädagogik, 105 (4), 521–556.
- WONDERLIC, E. F. (1992). Manual of the Wonderlic Personnel Test & Scholastic Level Exam II. Wonderlic Personnel Test. Inc.
- WU, M., ADAMS, R. & HALDANE, S. (2005). ConQuest: Generalised Item Response Modelling Software. Australian Council for Educational Research, University of California, Berkeley.

Anschrift der Autoren:

Johannes Rosendahl, Institut Technik und Bildung, Universität Bremen, johannes.rosendahl@uni-bremen.de
Gerald A. Straka, Institut Technik und Bildung, Universität Bremen, straka@uni-bremen.de