

Modellierung und Messung domänenspezifischer Problemlösekompetenz bei angehenden Industriekaufleuten

Entwicklung eines Testinstruments und erste Befunde zu kognitiven Kompetenz-facetten

KURZFASSUNG: Der kaufmännische Arbeitsalltag erfordert in zunehmenden Maß domänenspezifische Problemlösekompetenz. Diese umfasst nach unserem Verständnis neben domänenspezifischem Wissen auch selbstregulatorische/metakognitive Fähigkeiten sowie nicht-kognitive Facetten (Selbstkonzept und Interesse). Im Rahmen des Verbundprojekts DomPL-IK (Domänenspezifische Problemlösekompetenz bei Industriekaufleuten) haben wir ausgehend von einem mehrdimensionalen Kompetenzmodell problemhaltige Szenarien für die Erfassung domänenspezifischer Problemlösekompetenz von Industriekaufleuten entwickelt und in einer Pilotstudie mit angehenden Industriekaufleuten (n = 86) und IT-Systemkaufleuten (n = 37) erprobt. Der vorliegende Beitrag konzentriert sich auf die Auswertungen im Bereich der Wissensanwendung. Die in einem ersten Schritt vorgenommene feingliedrige inhaltsanalytische Kodierung wurde anschließend zu Partial Credits verdichtet, die dann wiederum für die Fähigkeitsschätzung mittels IRT herangezogen wurden. Das erwartungsgemäß bessere Abschneiden der Industriekaufleute beim Lösen der dargebotenen Problemszenarien (im Vergleich zu IT-Systemkaufleuten) kann – mit der gebotenen Zurückhaltung – als Indikator für die curriculare Validität der Szenarien gedeutet werden. Auf Potenziale zur Verbesserung sowohl der Problemstellungen als auch der Auswertungsmethoden verweisen dagegen insbesondere die stellenweise nicht zufriedenstellenden internen Konsistenzen (Reliabilitätswerte) sowie inhaltliche Schwierigkeiten der Testteilnehmenden bei der Bearbeitung der Szenarien.

ABSTRACT: Daily working routines in offices increasingly require domain specific problem solving competence. According to our understanding, this involves – besides domain specific knowledge – self-regulation and metacognitive skills as well as non-cognitive competence facets (self-concept and interest). In the context of the German project network DomPL-IK (Domänenspezifische Problemlösekompetenz von Industriekaufleuten/Domain specific problem solving competence of industrial management assistants) we developed a multidimensional model of problem solving competence. Based on this model we created scenarios that require industrial management assistants in training to solve typical problems of their trade. We then tested the scenarios in a pilot study with 123 participants (n = 86 industrial management assistants, n = 37 management assistants in informatics). In the article at hand we focus on knowledge application. Problem solutions were analyzed in a fine-grained content analytic approach and then condensed into partial credits. Those were used to estimate competences with item response theory. As expected industrial management assistants score better than management assistants in informatics. This can be cautiously interpreted as an indicator for curricular validity of the scenarios. Moderate internal consistencies (reliabilities) and content-related difficulties in the problem solving process reveal room for improvement for the main study.

1 Problemstellung

Vor dem Hintergrund aktueller und zukünftiger Veränderungen der Arbeitswelt (u. a. neue Technologien, Globalisierung und demographische Entwicklungen, siehe Stock-Homburg 2013) besteht weitgehend Konsens, dass berufliche Anforderun-

gen durch ein hohes Maß an Komplexität gekennzeichnet sind (ACHTENHAGEN 2000; BECKETT & HAGER 2002; FULLER & UNWIN 2003; ROSA 2005; SEMBILL 2008; STENSTRÖM & TYNJÄLÄ 2009). Die Bewältigung von Veränderungen, das Einarbeiten in neue Arbeitsbereiche oder der Umgang mit neuen Technologien treten – in vielen Branchen und Tätigkeitsfeldern – zunehmend an die Stelle repetitiver Routinearbeiten. Von Arbeitnehmern wird die Bewältigung von Arbeitssituationen mit ungewissem Ablauf und Ausgang erwartet (SCHMITT et al. 2003). Derartige, oft als „21st century skills“ (JONASSEN 2010) bezeichnete Kompetenzen sind nicht nur für „Knowledge Worker“ relevant, sondern auch für den produzierenden Sektor. Mitarbeiterinnen und Mitarbeiter werden zunehmend zu Problemlösern.

Trotz weitreichendem Konsens hinsichtlich der Notwendigkeit von Problemlösefähigkeit in beruflichen Domänen ist deren Diagnostik bislang eher unterentwickelt (zu jüngeren Entwicklungen siehe ACHTENHAGEN 2007; ACHTENHAGEN & WINTHER 2010; BAETHGE, ACHTENHAGEN, ARENDS, BABIC & BAETHGE-KINSKY 2006; BECK 2005 und 2009; GSCHWENDTNER, GEISSEL & NICKOLAUS 2010; LEHMANN & SEEGER 2007; NICKOLAUS, ROSENDAHL, GSCHWENDTNER, GEISSEL & STRAKA 2010; SEEGER 2008; WINTHER 2010 sowie ZLATKIN-TROITSCHANSKAIA, BECK, SEMBILL, NICKOLAUS & MULDER 2009). Zwar finden sich in der allgemeinen, d. h. domänenfreien Problemlöseforschung interessante Ansätze (für einen Überblick siehe GREIFF 2012), doch ist eine Übertragung insbesondere aufgrund der Domänenspezifität beruflicher Kompetenzen nicht ohne weiteres möglich. Während die in Large-Scale-Assessments angewendeten, kleinschrittigen Verfahren hinsichtlich der ökologischen Validität kritisch erscheinen, weisen offene, umfangliche Problemszenarien oft eine nicht hinreichenden Testökonomie sowie Reliabilitätsprobleme auf (SEMBILL 1992; WUTTKE & WOLF 2007).

Zentrales Ziel des Verbundprojekts DomPL-IK („Domänenspezifische Problemlösekompetenz von Industriekaufleuten“)¹ ist die Entwicklung eines ökologisch validen und standardisierbaren Testinstrumentes zur Messung domänenspezifischer Problemlösekompetenz von angehenden Industriekaufleuten gegen Ende einer i. d. R. dreijährigen Ausbildung. Die darin enthaltenen Szenarien umfassen sowohl Routinetätigkeiten als auch Elemente von echten Problemen, so dass wir im Folgenden von problemhaltigen Aufgaben sprechen. Als Domäne wurde Controlling als zentraler Bereich kaufmännischer Tätigkeiten gewählt. Im vorliegenden Beitrag berichten wir Daten aus einer Pilotierungsstudie. Es werden folgende Fragen beantwortet:

- (1) Wie stellen sich Handhabbarkeit und Qualität des Testverfahrens dar?
- (2) Lassen sich hinsichtlich der Lösungsqualitäten Unterschiede in Abhängigkeit vom Ausbildungsberuf (Industriekaufleute vs. IT-Systemkaufleute) ausmachen?

Im Folgenden erläutern wir zunächst das im Rahmen des Projekts entwickelte Modell domänenspezifischer Problemlösekompetenz. Danach gehen wir auf die Entwicklung der Problemszenarien, die Datenerhebung und ausgewählte Bereiche der Datenauswertung ein. Es folgen Befunde und Erfahrungen aus einer Pilotstudie. Über die in unserem Ansatz als bedeutsam erachteten nicht-kognitiven Facetten

1 Das durch das Bundesministerium für Bildung und Forschung geförderte Projekt ist eingebettet in die Forschungsinitiative Ascot (Technology-based Assessment of Skills and Competencies in VET/Technologie-orientierte Kompetenzmessung in der Berufsbildung, siehe www.ascot-vet.net). Beteiligt sind die Universitäten Bamberg, Bremen, Frankfurt, Mannheim sowie das Deutsche Institut für Internationale Pädagogische Forschung (DIPF) in Frankfurt (Förderkennzeichen 01DB1119 bis 01DB1123).

wird im vorliegenden Beitrag nicht berichtet, da die Pilotstudie diesbezüglich nur wenig Aufschluss bietet und dies zudem den Umfang des Beitrags sprengen würde.

2 Modellierung domänenspezifischer Problemlösekompetenz

2.1 Problembegriff und Problemlösen

Klassischen Definitionen zufolge beschreibt ein Problem eine Situation, in der ein Individuum ein Ziel hat, ohne im Moment über die Mittel zu verfügen, den als unerwünscht bewerteten Ist-Zustand in einen erwünschten Zielzustand zu überführen (DÖRNER 1987; DUNCKER 1945). Wohldefinierte Probleme sind durch eine klare Ausgangslage, klare Ziele und ein klares Set an erlaubten Operatoren gekennzeichnet. Bei nicht wohldefinierten oder dialektischen Problemen sind diese Merkmale nicht gegeben. Dies trifft auf alltägliche Probleme des (Berufs-)Lebens eher zu als auf Probleme in schulischen Lehr-Lern-Arrangements (DÖRNER 1987; MAYER 2011, S. 112f.). Insbesondere in komplexen Problemen des realen Lebens kommen meist mehrere Barrieren vor, die die Überführung eines unerwünschten Ausgangszustandes in einen erwünschten Zielzustand erschweren (DÖRNER 1987). Die Komplexität von Problemsituationen wird bedingt durch Anzahl und Vernetzung der Variablen, Polytelie (mehrere, teils konfligierende Ziele), Intransparenz bezüglich der Variablen und deren Vernetzung und Wirkungsbeziehungen sowie eigendynamische Entwicklungen der Situation und Zeitdruck (DÖRNER 2000, S. 58ff.; FUNKE 2003, S. 126ff.). Allerdings sind weder der Barrieretyp noch die Komplexität eines Problems objektive Merkmale. Ob eine Situation als komplexes Problem oder als einfache Aufgabe wahrgenommen wird, hängt auch vom Individuum und dessen Vorerfahrungen im betreffenden Realitätsbereich ab (DÖRNER 1987; FUNKE 2003; GREIFF & FUNKE 2010). Ein solcher Realitätsbereich ist gemäß DÖRNER (1987) durch eine Menge möglicher Sachverhalte und eine Menge an Operatoren charakterisiert. Im berufsbildenden Bereich werden die in einem Berufsfeld auffindbaren Bündel zusammengehöriger Realitätsbereiche als Domänen bezeichnet (ACHTENHAGEN 2007, S. 485 sowie dort zitierte Literatur), deren Grenzen von der jeweiligen Berufs- oder Fachgruppe definiert werden (SLOANE 2008, S. 501) und die arbeitspsychologische Konzepte wie *personal initiative*, *organizational citizenship behavior* etc. einschließen (NERDINGER, BLICKLE & SCHAPER 2014; NEUBERGER 2006; SMITH, ORGAN & NEAR 1983). Hinsichtlich der Quantität und Qualität der Vorerfahrungen in einer Domäne können Personen dann auf einem Kontinuum bzw. Entwicklungspfad zwischen Novize und Experte verortet werden (van GOG 2012). Während Novizen in einer Domäne eine domänenspezifische Handlungssituation eher als Problem erleben, kann die gleiche Handlungssituation für Experten in der betreffenden Domäne lediglich eine Aufgabe darstellen.

Problemlöseprozesse werden meist in idealtypischen Schritten beschrieben, wenngleich diese empirisch kaum identifizierbar/trennbar sind, sondern vielmehr iterativ ablaufen (BETSCH, FUNKE & PLESSNER 2011). Übliche Handlungsstationen sind: (1) Analyse und Bewertung der Ausgangslage, (2) Definition und Priorisierung von Handlungszielen, (3) Generierung von alternativen Handlungsplänen, (4) Bewertung von Handlungsalternativen, (5) Entscheidung und Durchführung sowie (6) Handlungskontrolle (z.B. BETSCH, FUNKE & PLESSNER 2011; BRANSFORD & STEIN 1993; DÖRNER 2000; LIPSHITZ & BAR-ILAN 1996; NEWELL & SIMON 1972; SEMBILL 1992; WUTTKE & WOLF 2007). Dieser Prozess ist, im Unterschied zu rein reproduktiven

Vorgängen, mindestens in Teilen konstruktiver Natur (FUNKE 2003). Handlungspläne sind dabei oft mit Unsicherheit verbunden, können ggf. irreversibel sein und somit im ungünstigsten Fall eine Problemlösung sogar verhindern. All dies erfordert eine umfassende Vorausplanung, die auch die Prognose von Neben- und Folgeeffekten einschließt (DÖRNER 1987; SEMBILL 1992). Herausragende Leistungen werden auch auf selbst-regulatorische, metakognitive Fähigkeiten wie die Kontrolle der eigenen Handlungsprozesse (Monitoring) zurückgeführt (GLASER 1994). Problemerkennung, Strategieauswahl und Handlungskontrolle finden sich ähnlich in Intelligenzdefinitionen wieder (EYSENK 1994). Die empirischen Befunde zum Einfluss von Intelligenz auf Problemlösen sind jedoch sehr heterogen und reichen von Nullkorrelationen (bspw. in Studien der Forschergruppen um DÖRNER und SEMBILL) bis hin zu Größenordnungen um .50 (FUNKE 2003). Neben „domänenfreien“ Dispositionen wie Intelligenz und Handlungsregulation spielt insbesondere domänenspezifische Expertise eine zentrale Rolle (NOKES, SCHUNN & CHI 2011). Domänenspezifische Fakten, Konzepte, Prozeduren und Nebenbedingungen helfen, Probleme zu kategorisieren und zu repräsentieren sowie Lösungen zu finden, umzusetzen und zu beurteilen (siehe auch ABERNETHY & CÔTÉ 2012, S. 1235 f.).

Erweitert man die kognitive Perspektive der Informationsverarbeitung und stellt Problemlösen in den Kontext allgemeiner Handlungsmodelle, treten zusätzlich emotionale und motivationale (Bewertungs-)Prozesse auf den Plan (FUNKE 2003; FUNKE & FRENSCH 1995; JONASSEN & HUNG 2012; SEMBILL 1984; 1992). Neben die Zielrelevanz (Wertkomponente) der Problemlösung tritt in geläufigen Motivationsmodellen die Erwartungs-Komponente, d. h. die Einschätzung der prospektiven Bewältigungsmöglichkeiten (SEMBILL 1992). Letztere beschreibt das subjektive Zutrauen einer Person in ihre Fähigkeit, mit der gegebenen Situation fertig werden zu können (DÖRNER 1985; DÖRNER, REITHER & STÄUDEL 1983). Mangelndes Selbstvertrauen kann zu vermindertem Interesse an einem Problembereich sowie zu generellen Ängsten vor der Auseinandersetzung mit Problemen und vor negativen Rückmeldungen führen (BRANSFORD & STEIN 1993).

2.2 Ansätze zur Modellierung von Problemlösekompetenz

In Anlehnung an den Sprachwissenschaftler CHOMSKY (1965) betrachten wir Kompetenz als eine latente, nicht beobachtbare Disposition, welche sich in beobachtbaren Performanzen manifestiert und auf Basis solcher Performanzen attribuiert wird. Trotz begrifflicher Unklarheiten und vieler methodischer Schwierigkeiten wird die Kompetenzdefinition von Weinert als Referenz-Zitat angesehen (KLIEME 2004, S. 13). WEINERT (2001, S. 27 f.) definiert Kompetenz als „die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können“ (eine ähnliche Definition findet sich bei den Autoren des National Center for Research on Evaluation, Standards, and Student Testing (CRESST); HERL et al. 1999, S. 2). Auch im Bereich der Mathematik werden international ähnliche Modelle diskutiert (SCHOENFELD 1992, 2013; VERSCHAFEL, VAN DOOREN & DE SMEDT 2012). Schließlich wurde auch die den PISA-Studien zugrunde liegende Definition von Problemlösen in der jüngeren Vergangenheit um nicht-kognitive Facetten erweitert (OECD 2013).

Auch wenn die Mehrdimensionalität des Kompetenzkonstrukts konsensfähig erscheint, bleibt strittig, inwiefern die verschiedenen Kompetenzbestandteile isoliert voneinander erfasst werden können. Unseres Erachtens erwachsen hierdurch mindestens zwei Gefahren: Erstens besteht eine große Diskrepanz hinsichtlich der „Situationsnähe“ der Erhebungsverfahren. Während die kognitiven Dimensionen anhand konkreter Testaufgaben gemessen werden, werden die nicht-kognitiven Dimensionen häufig mit globalen Selbsteinschätzungen erhoben (z. B.: „Ich bin mir sicher, die Ausbildung gewählt zu haben, die meinen Interessen entspricht.“). Ein erwartbar geringer Zusammenhang wird dann manchmal als Argument für die generell vernachlässigbare Bedeutung nicht-kognitiver Facetten herangezogen. Zweitens werden Befunde zu nicht-kognitiven Dimensionen ungeachtet der i. d. R. zugrunde gelegten mehrdimensionalen Kompetenzstrukturmodelle vergleichsweise selten berichtet. Im Rahmen des DomPL-IK-Projekts wird dagegen eine integrierte Erfassung von Kompetenz angestrebt (SEMBILL, RAUSCH & KÖGLER 2013).

2.3 Modellierung domänenspezifischer Problemlösekompetenz in DomPL-IK

Problemlösekompetenz i. S. des adäquaten Umgangs mit komplexen Realitätsausschnitten erfordert Fachwissen (ABELE 2014; SCHMIDT & HUNTER 2004, 170f.), das für die Bewältigung beruflicher Anforderungen und für beruflichen Erfolg eine notwendige, aber nicht hinreichende Bedingung darstellt. Die in einer Domäne erworbenen wissens- und erfahrungsbasierten Handlungsroutinen sind ein Kernbestandteil von Expertise (ERICSSON 2006; GRUBER, HARTEIS & REHRL 2006, 193), sie müssen angesichts einer sinkenden „Halbwertszeit“ von Wissen jedoch regelmäßig hinterfragt und ggf. modifiziert werden. Hinzu treten auch emotional-motivationale Aspekte sowie die Fähigkeit zur Handlungsregulation (BRANSFORD & STEIN 1993; DÖRNER & KAMINSKI 1988; KÜHL 2001; RAUSCH 2011; RUIZ-PRIMO 2009; SCHUMACHER 2002; SEMBILL 1992, 1995, 2010; WOLF 2003; WUTTKE 1999). Die Gesamtheit dieser Faktoren macht domänenspezifisches Problemlösen aus. Das individuelle Potenzial, domänenspezifische Probleme zu bearbeiten, bezeichnen wir im Folgenden als domänenspezifische Problemlösekompetenz.

Bei der Kompetenzmodellierung unterscheiden wir vier Kompetenzdimensionen: (1) Wissensanwendung, (2) Handlungsregulation, (3) Selbstkonzept und (4) Interesse. Die beiden letztgenannten werden als nicht-kognitive Facetten bezeichnet, wenngleich sie streng genommen auch kognitive Bestandteile enthalten. Die analytische Trennung dient der Operationalisierung. Die Kompetenzmodellierung orientiert sich ferner am Verlauf eines typischen Problemlöseprozesses. Dadurch spannt sich ein zweidimensionaler Raum auf, in dem 13 Kompetenzfacetten verortet sind. Die Modellelemente sind domänenunspezifisch formuliert. Dennoch ist das Modell nicht als Ausdruck einer allgemeinen Problemlösefähigkeit zu sehen; der Domänenbezug entsteht durch die spezifischen Problemszenarien.

Kompetenz- dimensionen	13 Kompetenzfacetten			
(A) Wissens- anwendung (Kognition)	Handlungsbedarfe und Informations- quellen identifizieren	Informationen verarbeiten	Begründete Entscheidung treffen	Entscheidung angemessen kommunizieren
(B) Handlungs- regulation (Metakognition)	Planvolles (strukturiertes) Vorgehen	Persistenz (fokussiertes) Vorgehen		Retrospektive Handlungs- kontrolle
(C) Selbstkonzept (Erwartungen)	Angemessenes Vertrauen in die eigene Kompetenz	Unsicherheits-/ Ungewissheits- toleranz		Angemessenes Vertrauen in die eigene Lösung
(D) Interesse (Valenzen)	Inhaltliches Interesse am Problem	Positive Aktiviertheit		Interesse an Problemfortgang / an Lernchancen

Abb. 1: Kompetenzmodell (ausführlich bei RAUSCH & WUTTKE in Vorb.)

3 Erhebungsmodellierung

3.1 Prinzipielle Anforderungen an eine Kompetenzdiagnostik

Im Kern geht es bei der Kompetenzdiagnostik darum, Handlungssituationen zu gestalten, in denen Teilnehmende eine beobachtbare Performanz (Verhalten) zeigen, von der auf ihre (nicht beobachtbare) Kompetenz (Handlungspotenziale bzw. Dispositionen) geschlossen werden kann. Dieses zunächst einfach anmutende Grundprinzip ist an einige nicht-triviale Voraussetzungen gebunden:

- (1) Alle das Handeln konstituierenden inneren Prozesse und deren Wechselwirkungen sind bekannt. Es liegen einige integrierte Handlungsmodelle vor (z. B. PSI-Theorien von DÖRNER 2008 oder KÜHL 2001, zu Weiterentwicklungen siehe Martens 2012), doch diese sind empirisch noch wenig geprüft. Das in Abbildung 1 dargestellte Kompetenzmodell stellt insofern eine Arbeitshypothese dar.
- (2) Alle das Handeln konstituierenden inneren Prozesse sind im Verhalten äußerlich sichtbar. Gerade Denk- und Erlebensprozesse sind aber kaum sichtbar, sondern müssen – teilweise hoch-inferent – erschlossen werden. Hierzu wurde im Rahmen von DomPL-IK mit den eingebetteten Erlebensstichproben (EES) ein innovatives Verfahren entwickelt (RAUSCH, SEIFRIED & KÄRNER in Vorb.).
- (3) Für die Problemlösung relevantes Verhalten ist als Datum beschreibbar und speicherbar. In einer „idealen“ Messumgebung würden bspw. auch Augenbewegungen erfasst (Eye-tracking), um darauf zu schließen, welche Bildschirminformation gerade fokussiert wird. In einer größeren Auflösung könnte bspw. erfasst werden, welche Bildschirminhalte aufgerufen werden (Speicherung von Logfiles in computerbasierten Umgebungen). Letzteres wird durch die im Rahmen von DomPL-IK entwickelte Testumgebung TeBaDoSLA (Technology Based Domain Specific Learning Assessment) ermöglicht.

- (4) Sämtliche erfassten Verhaltensdaten sind eindeutig interpretierbar. Hiermit ist u. a. die Frage verbunden, wie das mehrmalige Abrufen ein und derselben Informationsquelle zu interpretieren ist (z. B. Gewissenhaftigkeit vs. Unsicherheit vs. mangelnde Merkfähigkeit).
- (5) Das korrekt interpretierte Handeln ist ein valider Indikator für die domänenspezifische Problemlösekompetenz, d. h. für das entsprechende Handeln in realen Arbeitskontexten. Dies betrifft neben der Authentizität der Testsituation auch Fragen der Reaktivität und Reaktanz, welche bei bestimmten Erhebungsverfahren zu erwarten sind.

3.2 Entwicklung domänenspezifischer Problemlöseszenarien

Als Domäne wurden im vorliegenden Fall kaufmännische Entscheidungen aus dem Beschaffungsbereich gewählt, die mit dem Instrumentarium des operativen Controllings bearbeitet werden. Dem Projekt liegt ein breites Controlling-Verständnis zugrunde: Controlling umfasst die bedarfsgerechte Informationsversorgung des Managements zur Vorbereitung betriebswirtschaftlicher Entscheidungen (z. B. REICHMANN 2001), Maßnahmen der operativen Planung und Kontrolle (z. B. HAHN 1987, HAHN & HUNGENBERG 2001) sowie Instrumente der Kosten und Leistungsrechnung. Dieser Inhaltsbereich umfasst sprach- und zahlenbasierte Steuerungs- und Entscheidungsaufgaben im mittleren Management, also in typischen Einsatzbereichen von Industriekaufleuten. Derartige Planungs-, Steuerungs- und Kontrollaufgaben erstrecken sich über (nahezu) alle Lernfelder (vgl. KMK 2002), so dass ein Transfer vom Beschaffungsbereich zu anderen unternehmerischen Funktionen wie Produktion oder Absatz möglich sein sollte.

Zur Entwicklung der Problemlöseszenarien wurden (1) Lehr- und Bildungspläne sowie Ausbildungsordnungen und Schulbücher analysiert, (2) eine Fragebogenstudie zu Anforderungen an ControllerInnen in Kooperation mit dem Lehrstuhl Unternehmensführung & Controlling der Universität Bamberg, (3) Tagebuchstudien zum Problemlösen am Arbeitsplatz im kaufmännischen Bereich (RAUSCH, SCHLEY & WARWAS in Vorb.) sowie (4) eine Interviewstudie mit 17 DomänenexpertenInnen zur Identifikation von Anforderungen und Aufgabenbereichen von Industriekaufleuten im Bereich Controlling durchgeführt (MERKERT 2013; SCHWAB 2012).²

Typische Aufgaben für Auszubildende und junge Fachkräfte im Controlling sind insbesondere solche, die Informationsverarbeitung beinhalten (z. B. Informationsbeschaffung, Systempflege, Ermittlung von Kennzahlen, Anfertigung und Präsentation von Auswertungen). Im Rahmen der Planung und Kontrolle dominiert die Durchführung von Abweichungsanalysen. Koordinationstätigkeiten wurden dagegen selten genannt. Insgesamt haben die typischen Aufgaben primär operativen Charakter, was bei Berufsanfängern plausibel ist. Trotz einer heterogenen Stichprobe ergibt sich ein recht einheitliches Bild. Dies gilt auch für die Qualitätskriterien der Aufga-

2 Die Befragung wurde mit 17 Probanden (5 w, 12 m) in leitfadengestützten Interviews zwischen April und Mai 2012 durchgeführt. Die Experten, überwiegend Projekt-, Team- oder Abteilungsleiter, kommen aus Unternehmen mit zwischen 220 bis 60.000 Mitarbeitern und verfügen über mindestens fünf Jahre Berufserfahrung in der Domäne sowie eine kaufmännische Ausbildung oder einen entsprechenden Studienabschluss. Sie arbeiten in verschiedenen Controllingbereichen (IT-Controlling, Facilities, Produktionscontrolling, im Wesentlichen inklusive Kostenrechnung).

benerledigung. Den befragten Experten sind die Korrektheit und Nachvollziehbarkeit des Arbeitsergebnisses, die empfängergerechte Aufbereitung und eine zügige Bearbeitung am wichtigsten. ERP-Systeme sowie Tabellenkalkulationen werden als zentrale Werkzeuge identifiziert. Als Implikationen für die Gestaltung der Szenarien und der Testumgebung ließ sich festhalten:

- (1) Geeignete Testinhalte sind insbesondere die Abbildung wiederkehrender Controlling-Prozesse und die Darstellung und datenbasierte Interpretation von Entwicklungen und klassischen Reports, die Durchführung und Interpretation von Abweichungsanalysen und die Aufbereitung von Kalkulationen zur Vorbereitung von Managemententscheidungen wie z. B. Make-or-Buy.
- (2) Die praxisnahe Gestaltung der Szenarien erfordert eine typische Büroumgebung inkl. der Einbindung von Materialien und Tools (Tabellenkalkulation, Email, Rechner, Verfahrensrichtlinien, Templates).
- (3) Zeitliche Vorgaben müssen eingehalten werden und eine Plausibilisierung der Ergebnisse ist einzufordern.

Grundlage für die realitätsnahe Umsetzung komplexer Probleme war ein Modellunternehmen, das in Kooperation mit einem mittelständischen Fahrradhersteller und unter Nutzung echter Unternehmensdaten konzipiert wurde. In der Testumgebung werden nach einem einführenden Tutorial drei Problemszenarien von je 30 Minuten präsentiert (siehe Abbildung 2).

Szenarien	(1) Abweichungsanalyse	(2) Lieferantenauswahl (Nutzwertanalyse)	(3) Eigenfertigung oder Fremdfertigung
Inhalt	Bezüglich der Umsatz- bzw. Absatzentwicklung eines Produktes treten signifikante Unterschiede zwischen Ist- und Planwerten auf. Mittels einer Abweichungsanalyse sollen die Ursachen ermittelt werden.	Für eine neue Produktvariante ist ein Lieferant für den Bezug des Fahrradrahmens zu ermitteln. Entsprechende Angebote wurden bereits eingeholt und sollen nun für einen Angebotsvergleich herangezogen werden.	Ein neues standardisiertes Produkt soll in die Produktpalette aufgenommen werden. Die Frage ist: Fremdbezug oder Eigenfertigung?
Problemlöseschritte	(1) Handlungsbedarfe und Informationsquellen identifizieren • Sichtung und Identifikation relevanter Informationen und Verarbeitung der Information • Identifikation zu bewältigender Arbeitsschritte (2) Controllingaufgaben (Informationen verarbeiten)		
	• Abweichungsanalyse auf Basis von Ist- und Sollwerten • Ermittlung von Abweichungsgründen	• Quantitativer Angebotsvergleich (Kalkulation des Einstands-/ Bezugspreises auf Basis von Listeneinkaufspreisen, Rabatte und Skonti) • Qualitativer Angebotsvergleich (auf Basis der Unternehmensziele und einer vorgegebenen Kriterienliste – Einstandspreis, Lieferfrist, Qualität, Lieferzuverlässigkeit, Zahlungsziel) • Nutzwertberechnung	• Ermittlung und Gegenüberstellung der Kosten von Eigenfertigung und Fremdbezug teilmontierter Komponenten • Berücksichtigung rechtlicher Bestimmungen (z. B. Zoll) sowie distributiver und unternehmensspezifischer Ziele (z. B. Abhängigkeit vom Lieferanten, Qualität, Risiken)

Szenarien	(1) Abweichungsanalyse	(2) Lieferantenauswahl (Nutzwertanalyse)	(3) Eigenfertigung oder Fremdfertigung
Problemlöseschritte	(3) Begründete Entscheidung treffen (4) Entscheidungen angemessen kommunizieren		
	• Relevante Abweichungen markieren • Deckungsbeitragsrechnung durchführen • Ergebnisreport	• Entscheidungsgrundlagen erarbeiten sowie Argumente für die getroffene Auswahl anführen • Tabellarische Auflistung des Angebotsvergleichs • Ergebnisreport	• Make-or-buy-Entscheidung begründen • Chancen/Risiken diskutieren • Kostenvergleich: Eigenfertigung vs. Fremdbezug • Ergebnisreport
Ergebniserwartungen	In einem Tabellenkalkulationsblatt ermitteln Probanden die Soll-Ist-Abweichungen, markieren über einem vorgegebenen Grenzwert liegende Abweichungen sowie weitere auffällige Sachverhalte. Sie erläutern schriftlich (per Mail an den Vorgesetzten) die identifizierten Abweichungsursachen sowie deren Konsequenzen für zukünftige Planungen.	In einem Tabellenkalkulationsblatt führen die Probanden eine Nutzwertanalyse durch, indem ein quantitativer und qualitativer Vergleich dreier potentieller Lieferanten vorgenommen wird. Die Probanden begründen schriftlich (per Mail an den Vorgesetzten) ihre Entscheidung für bzw. gegen einen Lieferanten.	In einem Tabellenkalkulationsblatt ermitteln die Probanden die Kosten der Eigenfertigung und des Fremdbezugs von teilmontierten Komponenten. Die Probanden diskutieren schriftlich (per Mail an den Vorgesetzten) die Entscheidung im Hinblick auf mögliche Chancen und Risiken.

Abb. 2: Problemlöseszenarien

3.3 Technologiebasierte Datenerhebung

Die Problemlöseszenarien sind in eine Testplattform eingebunden. Jeder Problemfall beginnt mit der Präsentation des zu lösenden Problems auf Basis von authentischen Materialien (Email mit Informationen und Anhängen von einer Vorgesetzten oder einem Kollegen). In diesem Schritt ist das Problem zu identifizieren und die Ziele der Problembearbeitung sind festzulegen. Die notwendigen Werkzeuge wie Notizblock, Rechner, Tabellenkalkulationsprogramm sowie Arbeitsdokumente sind in der Testumgebung verfügbar. Es werden sowohl relevante als auch irrelevante Informationen zur Verfügung gestellt. Abschließend ist stets die Lösung zu kommunizieren und zu begründen. Neben der Qualität der Problemlösung (siehe Abbildung 2 „Ergebniserwartungen“) werden die erhobenen Selbstberichtsdaten zum situationsspezifischen Interesse und zum problembezogenen Selbstkonzept ausgewertet. Die Berücksichtigung nicht-kognitiver Facetten wird über situierte Nachfragen von Kolleginnen und Kollegen zum emotional-motivationalen Erleben und zum Vertrauen in die eigenen Kompetenzen implementiert. Daraus resultieren die folgenden Datenbestände:

- (1) Für die kognitiven Facetten des Problemlöseprozesses (Wissensanwendung): Qualität der Arbeitsergebnisse (Email, Berechnungen etc.).
- (2) Für die nicht-kognitiven Facetten (Interesse und Selbstkonzept): Eingebettete Erlebensstichproben (RAUSCH, SEIFRIED & KÄRNER in Vorb.).
- (3) Für die Prozessqualität der Problemlösung (Handlungsregulation): Logdaten aus der Testumgebung (geloggt wird jede Eingabe und jeder Mausklick).

Im Folgenden konzentrieren wir uns ausschließlich auf die Qualität der Arbeitsergebnisse, d.h. die Kompetenzdimension Wissensanwendung (vgl. Abbildung 1), da andere Daten aus der hier zugrundeliegenden Pilotstudie noch nicht oder nur rudimentär zur Verfügung stehen.

3.4 Auswertungsmodellierung (Kompetenzdimension Wissensanwendung)

Die Kodierung der vier Facetten im Bereich Wissensanwendung (siehe Abbildung 1) erfolgte in einem zweistufigen Prozess. Zunächst wurden die Antworten sehr feingliedrig im Rahmen von dichotomen Kategorisierungen beschrieben (z. B. „hat unwichtige Dokumente verwendet“: „ja“/„nein“). Für die vier Facetten fielen insgesamt 67 Kategorisierungen an. Aus diesen Kategorisierungen wurden anschließend je Aufgabe und Facette die aufgetretenen Antwortmuster extrahiert, durch Experten geratet bzw. auf verschiedene, qualitativ unterschiedliche Niveaus eingeordnet. Die identifizierten Niveaustufen wurden dann schließlich als Partial Credits im Rahmen des (mehrdimensionalen) Partial-Credit-Modells (MASTERS 1982) interpretiert. Je Facette wurden auf diese Weise drei Partial Credit Items (eines je Szenario) kodiert, die Anzahl der Partial Credits je Item lag dabei zwischen drei und sieben (siehe Tabelle 1). Für den Informationsgehalt, den die Items zur Messung einer Facette bieten, ist dementsprechend zu berücksichtigen, dass ein Partial-Credit-Item mit bspw. fünf Partial Credits vier dichotomen Items entspricht. Das bedeutet, dass zwar lediglich drei Items je Facette vorliegen, diese jedoch eine breite Informationsbasis repräsentieren. So entspricht beispielsweise der Informationsgehalt von Facette 1 dem von 18 dichotomen Items (siehe hierzu die jeweilige Anzahl der Partial Credits in Tabelle 1).

Vorteil des angewandten Verfahrens ist zunächst, dass Erfahrungen mit detaillierten Kodierungen gewonnen werden, die es im Rahmen der Hauptstudie erlauben, die Kodierungen zum Teil automatisiert aus Logging-Daten zu gewinnen. Die Zusammenfassung der Antwortmuster zu einem einzigen Partial Credit-Item hat darüber hinaus den Vorteil, dass lokale Abhängigkeiten, die die Items innerhalb eines Szenarios und einer Aufgabe sonst aufweisen würden, umgangen werden können, so dass die berechneten Leistungsschätzer und ihre Reliabilitäten nicht verzerrt werden (vgl. WANG & WILSON 2005; WAINER, BRADLOW & WANG 2007).

4 Befunde und Erfahrungen aus der Pilotstudie

4.1 Eckdaten der Pilotstudie

Im Frühjahr 2013 wurde nach verschiedenen kleineren Tests mit Studierenden der Wirtschaftspädagogik eine umfangreiche Pilotstudie mit Auszubildenden durchgeführt. Ziele der Pilotstudie waren die Erprobung der entwickelten Problemlöserszenarien, die Sammlung erster Erfahrungen mit den (hier noch papierbasierten) eingebetteten Erlebensstichproben sowie die Generierung von Daten zur Entwicklung des o. a. Auswertungsverfahrens. Die Stichprobe umfasst 123 Personen, davon 86 Industriekaufleute und 37 IT-Systemkaufleute. Letztere wurden als Referenzstichprobe herangezogen. Die Eignung der IT-Systemkaufleute als Referenzgruppe wurde mittels Curriculumanalysen überprüft. Diese verweisen auf eine ausreichende

Ähnlichkeit der Curricula bezüglich der von uns in den Blick genommenen Inhaltsbereiche. Dennoch kann davon ausgegangen werden, dass die IT-Systemkaufleute aufgrund der im Vergleich zu den Industriekaufleuten geringeren (aber dennoch gegebenen) curricularen Passung des Kompetenztests bei diesem schlechter abschneiden als die Industriekaufleute. Die Teilnehmenden sind im Mittel 22 Jahre alt und mehrheitlich männlich (66 m, 57 w). Zum Zeitpunkt der Datenerhebung befanden sich die Probandinnen und Probanden überwiegend im zweiten Ausbildungsjahr (119 von 123 Probanden). Ergänzend zu der Bearbeitung der Problemfälle wurden die Teilnehmenden um Selbstauskünfte zur beruflichen Selbstwirksamkeit sowie zum Ausbildungsinteresse gebeten. Zudem waren ein kognitiver Fähigkeitstest (CFT 20-R in gekürzter Fassung, WEISS 2006) sowie ein inhaltspezifischer Wissenstest (Aufgaben aus den Inhaltsbereichen der Szenarien) zu bearbeiten. Die Testung fand in Klassenverbänden in schultypischen Computerräumen statt.

Die Testung auf domänenspezifische Problemlösekompetenz begann mit der Präsentation des Modellunternehmens anhand eines Videos. Danach wurde die Arbeitsumgebung vorgestellt, die – anders als in der Hauptstudie – noch nicht in einer alle Testelemente umfassenden Umgebung integriert war. In der Pilotstudie wurden die Szenarien noch in einer Windows®-üblichen Ordnerstruktur anhand von PDF-Dokumenten präsentiert. Als Tabellenkalkulation wurde Microsoft Excel® verwendet und die Antwortmails wurden in Microsoft Word® verfasst. Erste Erfahrungen mit den eingebetteten Erlebensstichproben wurden über papierbasierte Kurzbefragungen gewonnen, die mittels durchnummerierter Umschläge während der Szenarien zu bearbeiten waren. Die Vorgehensweise wurde im Vorfeld erläutert. Zudem waren die Bearbeitungszeiten der drei Szenarien aufsteigend (30, 40 und 50 Minuten), während in der späteren Hauptstudie für jedes Szenario 30 Minuten zur Verfügung standen.

4.2 Zentrale Befunde der Pilotstudie

Handhabbarkeit: Die Pilotierung zeigt, dass die Probandinnen und Probanden mit der technischen Umsetzung gut zurechtkamen. In Gruppendiskussionen, die sich an die Testung anschlossen, wurde die Umsetzung der Szenarien als realistisch und authentisch eingeschätzt. Das Embedded Experience Sampling (s. o., nicht Thema dieses Beitrags) wurde von den Probanden als realistisch und dem Büroalltag entsprechend eingeschätzt. Die Beantwortung der Fragen sei ehrlich und spontan erfolgt, ohne „über besonders wünschenswerte“ Antworten nachzudenken.

Qualität des Testinstrumentes/interne Konsistenz: Die für die vier wissensbasierten Facetten analysierten Daten wurden auf Basis eines vierdimensionalen Rasch Modells ausgewertet. Bezüglich der Qualität des Testinstruments lässt sich festhalten, dass die EAP-PV-Reliabilitäten für die vier Facetten (Problemlöseschritte) noch nicht zufriedenstellen. Sie betragen .72 (Handlungsbedarfe und Informationsquellen identifizieren), .67 (Informationen verarbeiten), .55 (Begründete Entscheidungen treffen) sowie .56 (Entscheidungen angemessen kommunizieren).

Itemschwierigkeiten: Tabelle 1 zeigt zur ersten Orientierung die Itemparameter des vierdimensionalen Rasch-Modells für die vier Facetten der Dimension Wissensanwendung (Problemlöseschritte; siehe Abbildung 1). Die Schätzung des Modells (ebenso wie die weiter unten folgenden DIF-Analysen) wurde in R mit Hilfe des TAM-Paketes durchgeführt (KIEFER, ROBITZSCH & WU 2013). Die Werte verdeutlichen,

dass Szenario 1 über alle Facetten hinweg schwieriger ist als Szenario 2 und für die Facetten 1, 2 und 4 auch schwieriger als Szenario 3.

Als eher einfach erweisen sich die Problemlöseschritte 1 (Handlungsbedarfe und Informationsquellen identifizieren) und 4 (Kommunikation der Problemlösungen). Bei der zuletzt genannten Facette gehen in das Scoring u. a. auch Aspekte ein, die Hinweise darauf geben, ob die Probanden in der Lage sind, eine Geschäfts-E-Mail zu schreiben und angemessen mit Vorgesetzten zu kommunizieren. Dies gelingt im vorliegenden Fall überwiegend. Schließlich fällt auf, dass Problemlöseschritt 2 in Szenario 1 die höchste Schwierigkeit aufweist. Die Hürde für die Probanden stellte sich hier gleich zu Beginn, da viele Testteilnehmende an der Unterscheidung zwischen Plankosten und Sollkosten scheiterten. Ein weiterer Problembereich waren für die Probanden zudem die hier notwendigen Berechnungen in Excel®. Szenario 2 weist insgesamt gesehen die geringste Schwierigkeit auf und konnte von den Probanden – abgesehen von Unachtsamkeitsfehlern (Rechenfehler, nicht beachtete Anweisungen) – recht gut gelöst werden. Szenario 3, das gemäß Untersuchungsdesign nicht von allen Probanden bearbeitet wurde, weist dann wieder höhere Schwierigkeiten auf. Viele Probanden waren zwar in der Lage, die notwendigen Informationen zu identifizieren, konnten diese dann aber aufgrund der Informationsvielfalt und Komplexität bei der Problemlösung nicht ausreichend berücksichtigen (Schritt 2: Informationsverarbeitung) bzw. nicht adäquat als Entscheidungsgrundlage heranziehen (Schritt 3). Mit Blick auf die Step-Parameter schließlich ist auffallend, dass diese vielfach nicht gemäß einer zunehmenden Schwierigkeit geordnet sind. Im Rahmen der Pilotierung war dies aufgrund der zum Teil sehr schwach besetzten Kategorien (geringer Stichprobenumfang) jedoch nicht unerwartet.

Tab. 1: Übersicht über die Item-Schwierigkeiten und die Anzahl der Partial Credits

Szenario/Facette	Itemparameter	Anzahl Partial Credits
Abweichungsanalyse		
(1) Handlungsbedarfe und Informationsquellen identifizieren	.5070	7
(2) Informationen verarbeiten	1.3920	5
(3) Begründete Entscheidungen treffen	.8213	6
(4) Entscheidungen angemessen kommunizieren	.4854	4
Lieferantenauswahl		
(1) Handlungsbedarfe und Informationsquellen identifizieren	-.5552	7
(2) Informationen verarbeiten	.5821	4
(3) Begründete Entscheidungen treffen	.3294	7
(4) Entscheidungen angemessen kommunizieren	.1884	4
Make-or-buy		
(1) Handlungsbedarfe und Informationsquellen identifizieren	-.0874	7
(2) Informationen verarbeiten	1.1885	4
(3) Begründete Entscheidungen treffen	1.0045	3
(4) Entscheidungen angemessen kommunizieren	.4002	4

Die latenten Korrelationen zwischen den Problemlöseschritten und der Intelligenz (CFT-Test, EAP-PV-Reliabilität = .82) sind Tabelle 2 zu entnehmen. Auf der Dia-

gonalen sind die Standard-Abweichungen ausgewiesen, die Werte oberhalb der Diagonalen zeigen die latenten Korrelationen zwischen den Konstrukten. Für die Intelligenz zeigen sich Werte zwischen .41 und .07. Insbesondere für Facette 2 (Informationen verarbeiten) zeigt sich ein Zusammenhang mit der Intelligenz. Angesichts des geringen Stichprobenumfangs sind diese Befunde allerdings mit der gebotenen Vorsicht zu interpretieren.

Tab. 2: Latente Korrelationen zwischen Problemlöseschritten und Intelligenz

	(1)	(2)	(3)	(4)	(5)
(1) Handlungsbedarfe	.64	.88	.74	.58	.29
(2) Informationsverarbeitung		.73	.71	.52	.41
(3) Entscheidungen treffen			.41	.33	.11
(4) Entscheidung kommunizieren				.51	.07
(5) Intelligenz (CFT)					.74

Um zu überprüfen, inwieweit die Items bzw. die Problemszenarien für die verschiedene Personengruppen gleich bzw. verschieden schwierig sind, wurden DIF-Analysen für die Variablen Ausbildungsberuf und Geschlecht durchgeführt. Dazu wurde je Facette und DIF-Variable ein eindimensionales Mehrfacetten-Modell berechnet.

Einfluss des Ausbildungsberufs auf die Qualität der Problembearbeitung: Erwartungsgemäß sind alle drei Problemszenarien signifikant leichter für die Industriekaufleute. Tabelle 3 zeigt hierzu die Ergebnisse für die vier Problemlöseschritte. Es werden jeweils die DIF-Parameter (Differenzen zwischen dem erzielten Wert der Referenzgruppe und dem Wert Null, negative Werte weisen auf Vorteile der Gruppe der Industriekaufleute), der Standardfehler sowie der z-Wert und das zugehörige Signifikanzniveau ausgewiesen. Der Unterschied zwischen den Berufen ergibt sich in diesem Fall aus der Verdoppelung des DIF-Parameters, da die DIF-Variable zwei Gruppen von Probanden unterscheidet.

Tab. 3: DIF-Analysen für die Variable Ausbildungsberuf (Industrie- vs. IT-Systemkaufleute)

Facette	DIF-Parameter	SE	z-Wert	p
Handlungsbedarfe/Informationen identifizieren	-.3033	.03879	-7.8192	< .01
Informationen verarbeiten	-.2791	.05709	-4.8887	< .01
Begründete Entscheidungen treffen	-.3389	.04976	-6.8118	< .01
Entscheidungen angemessen kommunizieren	-.1718	.0506	-3.3944	< .01

Anmerkungen:

- (1) DIF-Parameter: Item-Schwierigkeitsunterschied, d.h. um wie viele Logits ist das Item für Industriekaufleute leichter bzw. für IT-Systemkaufleute schwerer. Der Unterschied zwischen Industriekaufleuten und IT-Systemkaufleuten ergibt sich aus der Verdoppelung des DIF-Parameters.
- (2) Standardfehler von (1)
- (3) z-Wert: Prüfwert der Z-Statistik, hier (1)/(2);
- (4) p: Wahrscheinlichkeit entsprechend der z-Statistik, entspricht hier der Wahrscheinlichkeit, mit der der angegebene Schwierigkeitsunterschied nicht von Null abweicht.

Weiterführende Analysen zeigen dann, dass Industriekaufleute über ein signifikant höheres szenariospezifisches Vorwissen verfügen als die IT-Systemkaufleute. Unterschiede bezüglich der beruflichen Selbstwirksamkeit, des Ausbildungsinteresses (jeweils mittels globaler Selbsteinschätzung) sowie der Intelligenz bestehen nicht. Signifikante Unterschiede bezüglich der selbst berichteten Anstrengung bei der Bearbeitung der Szenarien (erfasst über den aus den PISA-Studien bekannten Anstrengungsthermometer) liegen für Szenario 2 und 3 vor (jeweils mit Vorteilen zu Gunsten der Industriekaufleute).

Für die Variable Geschlecht liegen für die Facetten „Handlungsbedarfe & Informationsquellen identifizieren“, „Begründete Entscheidungen treffen“ sowie „Entscheidungen angemessen kommunizieren“ ebenfalls signifikante DIF-Effekte vor. Die Probandinnen schneiden jeweils besser ab als die männlichen Testteilnehmer. Erklärungen, warum Frauen die Problemszenarien besser bearbeiten, könnten in unterschiedlichen Vorwissenbeständen zu finden sein, da Frauen auch beim szenariospezifischen Vorwissenstest signifikant besser abschnitten. Zudem berichten die Frauen durchgängig über ein signifikant höheres Maß an Anstrengung bei der Problembearbeitung (Anstrengungsthermometer) sowie ein signifikant höheres allgemeines Interesse an der Ausbildung. Unterschiede bezüglich der beruflichen Selbstwirksamkeit ließen sich dagegen nicht finden.

5 Diskussion und Ausblick

Im vorliegenden Beitrag berichteten wir über Befunde aus einer Pilotstudie im BMBF-Verbundprojekt DomPL-IK (Domänenspezifische Problemlösekompetenz bei Industriekaufleuten). Ausgehend von der Modellierung von Problemlösekompetenz haben wir die Entwicklung von Problemszenarien und erste Ergebnisse zur Auswertung der kognitiven Facetten aufgezeigt. Die Probandinnen und Probanden kamen prinzipiell gut mit dem offenen Testformat zurecht und konnten die dargebotenen Szenarien in unterschiedlichem Ausmaß bewältigen. Unterschiede konnten bspw. dahingehend festgestellt werden, dass Industriekaufleute die Szenarien mit größerem Erfolg bearbeiteten als die Referenzstichprobe der IT-Systemkaufleute. Dieser Befund ist angesichts der besseren curricularen Passung des Tests für den Beruf der Industriekauffrau/des Industriekaufmanns erwartungskonform und kann – mit der gebotenen Zurückhaltung – als erstes Zeichen für curriculare Validität interpretiert werden. Nicht zu erwarten waren dagegen die vielfältigen Vorteile der Probandinnen im Vergleich zu den männlichen Testteilnehmern. Plausible Ursachen könnten das höhere Vorwissen, das höhere Maß an Anstrengung bei der Problembearbeitung sowie das allgemein höhere Ausbildungsinteresse (jeweils Selbstberichte) seitens der Probandinnen sein. Inwiefern sich dieser Befund auch in der Hauptstudie zeigt, bleibt abzuwarten.

Die Analyse der Itemschwierigkeiten liefert zudem eine Reihe von Hinweisen auf größere fachliche Hürden, die die Mehrzahl der Probanden nicht hinreichend gut meistern konnte. Einige dieser Hürden wurden – soweit dies curricular vertretbar war – etwas entschärft, so dass zwar einerseits weiterhin von Problemhaltigkeit der Aufgaben (zumindest für den Großteil der Teilnehmenden) ausgegangen werden kann, aber andererseits verhindert wird, dass von zu vielen Teilnehmenden keine auswertbaren Daten vorliegen. Ferner wurden die Lösungsversuche in der Tabellenkalkulation (MS Excel®) analysiert, um die funktionalen Anforderungen

an die integrierte Testumgebung zu erfassen. Im Rahmen der Durchführung der Pilotierungsstudie konnten zudem wichtige Informationen zur Optimierung der Testanweisung, der Reihenfolge der Darbietung der Testanforderungen sowie weiteren logistischen und technischen Feinheiten gewonnen werden.

Im Rahmen der Kodierung der Aufgaben wurden darüber hinaus wichtige Erfahrungen gesammelt, auf deren Basis die Aufgaben für die Hauptstudie sowie deren Kodierung verbessert werden konnten. Bezüglich der Facette „Entscheidungen angemessen kommunizieren“ zeigte sich etwa, dass Detailkodierungen weniger zentraler Aspekte (z. B. die Kodierung der Anzahl der Rechtsschreibfehler) zu beachtlichen Abweichungen zwischen Kodierern führten. Für den Haupttest wurden daher sowohl die Items als auch das Kodiererhandbuch nochmals grundlegend überarbeitet, so dass hier bessere Realisierbarkeiten zu erwarten sind. Weiterhin stehen im Rahmen der Hauptstudie zahlreiche weitere Datenbestände aus dem Logging zur Verfügung (etwa welche Informationsquellen tatsächlich genutzt wurden). Dies ermöglicht teilautomatisierte Auswertungen, wodurch sich die teils ernüchternden Reliabilitätswerte der Pilotstudie deutlich verbessern sollten.

Zum aktuellen Zeitpunkt ist die bundesweit durchgeführte Haupterhebung abgeschlossen, und es liegen Daten von knapp 800 Probandinnen und Probanden vor. Befunde, insbesondere auch zu den nicht-kognitiven Facetten, werden in Bälde berichtet.

Literatur

- ABELE, S. (2014). *Modellierung und Entwicklung berufsfachlicher Kompetenz in der gewerblich-technischen Ausbildung*. Stuttgart: Steiner.
- ABERNETHY, B. & CÔTÉ, J. (2012). Expert perceptual and decision-making skills: Effects of structured activities and play. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 1235–1238). New York, NY: Springer.
- ACHTENHAGEN, F. (2000). Lebenslanges Lernen aus der Sicht des Mastery Learning. In F. ACHTENHAGEN & W. LEMPERT (Hrsg.), *Lebenslanges Lernen im Beruf – seine Grundlegung im Kindes- und Jugendalter: Band 4: Formen und Inhalte von Lernprozessen* (S. 123–140). Opladen: Leske + Budrich.
- ACHTENHAGEN, F. (2007). Wirtschaftspädagogische Forschung zur beruflichen Kompetenzentwicklung. In J. VAN BUER & C. WAGNER (Hrsg.), *Qualität von Schule. Ein kritisches Handbuch* (S. 481–494). Frankfurt am Main: Lang.
- ACHTENHAGEN, F. & WINTHER, E. (2010). Kompetenzdiagnostik in der beruflichen Aus- und Weiterbildung. In J. WARWAS & D. SEMBILL (Hrsg.), *Schule zwischen Effizienzkriterien und Sinnfragen* (S. 208–228). Baltmannsweiler: Schneider Hohengehren.
- BAETHGE, M., ACHTENHAGEN, F., ARENDS, L., BABIC, E. & BAETHGE-KINSKY, V. (2006). *Berufsbildungs-PISA. Machbarkeitsstudie*. Stuttgart: Steiner.
- BECK, K. (2005). Ergebnisse und Desiderate zur Lehr-Lern-Forschung in der kaufmännischen Berufsausbildung. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 101(4), 533–556.
- BECK, K. (2009). Strategien empirischer Forschung zur Professionalität von Lehrpersonen – Ein kritischer Blick aus methodologischer Sicht. In O. ZLATKIN-TROITSCHANSKAIA, K. BECK, D. SEMBILL, R. NICKOLAUS & R. MULDER (Hrsg.), *Lehrerprofessionalität. Bedingungen, Genese, Wirkungen und Messung* (S. 237–247). Weinheim: Beltz.
- BECKETT, D. & HAGER, P. J. (2002). *Life, work and learning. Practice in postmodernity*. London: Routledge.
- BETSCH, T., FUNKE, J. & PLESSNER, H. (2011). *Denken – Urteilen, Entscheiden, Problemlösen*. Heidelberg: Springer.

- BRANSFORD, J. & STEIN, B. S. (1993). *The ideal problem solver. A guide for improving thinking, learning, and creativity* (2nd ed.). New York: Freeman.
- CHOMSKY, N. (1965). *Aspects of the Theory of Syntax*. Cambridge: The MIT Press.
- DÖRNER, D. (1985). Verhalten und Handeln. In D. DÖRNER & H. SELG (Hrsg.), *Psychologie – Eine Einführung in ihre Grundlagen und Anwendungsfehler* (S. 73–86). Stuttgart: Kohlhammer.
- DÖRNER, D. (1987). *Problemlösen als Informationsverarbeitung* (3. Aufl.). Stuttgart: Kohlhammer.
- DÖRNER, D. (2000). *Logik des Mißlingens – Strategisches Denken in komplexen Situationen* (13. Aufl.). Reinbek bei Hamburg: Rowohlt.
- DÖRNER, D. (2008). *Bauplan für eine Seele* (2. Aufl.). Reinbek bei Hamburg: Rowohlt.
- DÖRNER, D. & KAMINSKI, G. (1988). Handeln – Problemlösen – Entscheiden. In K. IMMELMANN, K. R. SCHERER, C. VOGEL & P. SCHMOOCK (Hrsg.), *Psychobiologie. Grundlagen des Verhaltens* (S. 375–415). Stuttgart: Gustav Fischer.
- DÖRNER, D., REITHER, F. & STÄUDEL, T. (1983). Emotion und problemlösendes Denken. In H. MANDL & G. L. HUBER (Hrsg.), *Emotion und Kognition* (S. 61–84). München: Urban und Schwarzenberg.
- DUNCKER, K. (1945). *On Problem-Solving*. Washington: APA.
- ERICSSON, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. In K. A. ERICSSON, N. CHARNES, P. J. FELTOVICH & R. R. HOFFMANN (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 683–703). Cambridge: University Press.
- EYSENCK, M. W. (1994). Intelligence. In M. W. EYSENCK (Ed.), *The Blackwell Dictionary of cognitive psychology* (pp. 192–193). Oxford: Blackwell.
- FULLER, A. & UNWIN, L. (2003). Fostering Workplace Learning: looking through the lens of apprenticeship. *European Educational Research Journal*, 2(1), 41–55.
- FUNKE, J. (2003). *Problemlösendes Denken*. Stuttgart: Kohlhammer.
- FUNKE, J. & FRENSCH, P. (1995). Complex problem solving research in North America and Europe: An integrative view. *Foreign Psychology*, 5, 42–47.
- GLASER, R. (1994). Expertise. In M. W. EYSENCK (Ed.), *The Blackwell Dictionary of cognitive psychology* (pp. 139–142). Oxford: Blackwell.
- GREIFF, S. (2012). *Individualdiagnostik komplexer Problemlösefähigkeit*. Münster: Waxmann.
- GREIFF, S. & FUNKE, J. (2010). Systematische Erforschung komplexer Problemlösefähigkeit anhand minimal komplexer Systeme. In E. KLIEME (Hrsg.), *Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes* (S. 216–227). Beiheft 56 der Zeitschrift für Pädagogik. Weinheim: Beltz.
- GRUBER, H., HARTEIS, C. & REHRL, M. (2006). Professional Learning: Erfahrung als Grundlage von Handlungskompetenz. *Bildung und Erziehung*, 59(2), 193–203.
- GSCHWENDTNER, T., GEISSEL, B. & NICKOLAUS, R. (2010). Modellierung beruflicher Fachkompetenz in der gewerblich-technischen Grundbildung. In E. KLIEME (Hrsg.), *Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes* (S. 258–269). Beiheft 56 der Zeitschrift für Pädagogik. Weinheim: Beltz.
- HAHN, D. (1987). Controlling: Stand und Entwicklungstendenzen unter besonderer Berücksichtigung des CIM-Konzeptes. In A.-W. Scheer (Hrsg.), *Rechnungswesen und EDV: 8. Saarbrücker Arbeitstagung 1987* (S. 3–39). Heidelberg: Physica.
- HAHN, D. & HUNGENBERG, H. (2001). *PuK: Planung und Kontrolle, Planungs- und Kontrollsysteme, Planungs- und Kontrollrechnung* (6. Auflage). Wiesbaden: Gabler.
- HERL, H. E., O'NEIL, JR., H. F., CHUNG, G. K. W. K., BIANCHI, C., WANG, S.-L., MAYER, R., LEE, C. Y., CHOI, A., SUEN, T. & TU, A. (1999). *Final report for validation of problem-solving measures. CSE Technical Report 501*. CRESST, University of California.
- JONASSEN, D. H. (2010). *Research Issues in Problem Solving. The 11th International Conference on Education Research. New Educational Paradigm for Learning and Instruction*.

- September 29 – October 1, 2010. <https://aectorg.yourwebhosting.com/publications/whitepapers/2010/JonassenI CER.pdf> (aufgerufen am 03.12.2014).
- JONASSEN, D. H. & HUNG, W. (2012). Problem solving. In N. M. SEEL (Ed.) *Encyclopedia of the sciences of learning* (pp. 2680–2683). New York: Springer.
- KIEFER, T., ROBITZSCH, A. & WU, M. (2013). *TAM: test analysis modules*. Retrieved from <http://CRAN.R-project.org/package=TAM>.
- KLIEME, E. (2004). Was sind Kompetenzen und wie lassen sie sich messen? *Pädagogik*, 56(6), 10–13.
- KMK (2002). *Rahmenlehrplan für den Ausbildungsberuf Industriekaufmann/Industriekauffrau vom 14.06.2002*.
- KUHL, J. (2001). *Motivation und Persönlichkeit*. Interaktionen Psychischer Systeme. Göttingen: Hogrefe.
- LEHMANN, R. H. & SEEGER, S. (Hrsg.). (2007). *ULME III. Untersuchung von Leistungen, Motivation und Einstellungen der Schülerinnen und Schüler in den Abschlussklassen der Berufsschulen*. Hamburg: Behörde für Bildung und Sport.
- LIPSHITZ, R. & BAR-ILAN, O. (1996). How problems are solved: Reconsidering the phase theorem. *Organizational Behavior and Human Decision Processes*, 65(1), 48–60.
- MARTENS, T. (2012). Was ist aus dem Integrierten Handlungsmodell geworden? In W. KEMPF & R. LANGEHEINE (Hrsg.), *Item-Response-Modelle in der sozialwissenschaftlichen Forschung* (S. 210–229). Berlin: Regener.
- MASTERS, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47, 149–174.
- MAYER, R. E. (2011). Problem solving and reasoning. In V. GROVER AUKRUST (Ed.), *Learning and cognition in education* (pp. 112–117). Oxford: Elsevier.
- MERKERT, J. (2013). *Typische Arbeitstätigkeiten angehender Industriekaufleute – Eine empirische Analyse mit Blick auf das Controlling*. Unveröff. Masterarbeit, Universität Mannheim.
- NERDINGER, F. W., BLICKLE, G. & SCHAPER, N. (2014). *Arbeits- und Organisationspsychologie* (3., vollst. überarb. Aufl.). Berlin und Heidelberg: Springer.
- NEUBERGER, O. (2006). *Mikropolitik und Moral in Organisationen. Herausforderung der Ordnung*. Stuttgart: Lucius & Lucius.
- NEWELL, A. & SIMON, H. A. (1972). *Human problem solving*. Englewood Cliffs, New Jersey: Prentice-Hall.
- NICKOLAUS, R., ROSENDAHL, J., GSCHWENDTNER, T., GEISSEL, B. & STRAKA, G. A. (2010). Erklärungsmodelle zur Kompetenz- und Motivationsentwicklung bei Bankkaufleuten, Kfz-Mechatronikern und Elektronikern. In J. SEIFRIED, E. WUTTKE, R. NICKOLAUS & P. E. SLOANE (Hrsg.), *Lehr-Lern-Forschung in der kaufmännischen Berufsbildung. Ergebnisse und Gestaltungsaufgaben* (S. 73–87). Beiheft 23 der Zeitschrift für Berufs- und Wirtschaftspädagogik. Stuttgart: Steiner.
- NOKES, T. J., SCHUNN, C. D. & CHI, M. T. H. (2011). Problem solving and human expertise. In V. GROVER AUKRUST (Ed.), *Learning and cognition in education* (pp. 104–111). Oxford: Elsevier.
- OECD (2013). PISA 2015 – Draft collaborative problem solving framework. URL: <http://www.oecd.org/pisa/pisaproducts/Draft%20PISA%202015%20Collaborative%20Problem%20Solving%20Framework%20.pdf> (aufgerufen am 03.12.2014)
- ROSA, H. (2005). *Beschleunigung. Die Veränderung der Zeitstrukturen in der Moderne*. Berlin: Suhrkamp.
- RAUSCH, A. (2011). *Erleben und Lernen am Arbeitsplatz in der betrieblichen Ausbildung*. Wiesbaden: VS-Verlag.
- RAUSCH, A., SCHLEY, T. & WARWAS, J. (in Vorb.). Problem solving in everyday office work—A diary study on differences between apprentices and skilled employees.
- RAUSCH, A., SEIFRIED, J. & KÄRNER, T. (in Vorb.). Embedded Experience Sampling (EES) as an integrated approach of measuring non-cognitive facets in the computer-based assessment of problem-solving competence in the business domain.

- RAUSCH, A. & WUTTKE, E. (in Vorb.). Foundation of a multi-faceted model of domain-specific problem-solving competence and evaluation by different stakeholders in the business domain.
- REICHMANN, T. (2001). *Controlling mit Kennzahlen und Managementberichten: Grundlagen einer systemgestützten Controlling-Konzeption* (6. Aufl.). München: Vahlen.
- RUIZ-PRIMO, M. A. (2009). *Towards a Framework for Assessing 21st Century Science Skills*. Aufgerufen am 28.04.2011. Verfügbar unter http://www7.nationalacademies.org/BOSE/Ruiz_Primo_21st_Century_Paper.pdf (aufgerufen am 03.12.2014).
- SCHMITT, N., CORTINA, J. M., INGERICK, M. J., & WIECHMANN, D. (2003). Personnel selection and employee performance. In W. C. BORMAN, D. R. ILGEN, & R. J. KLIMOSKI (Eds.), *Handbook of Psychology Industrial and organizational psychology* (Vol. 12, pp. 77–105). Hoboken, NJ: Wiley.
- SCHMIDT, F. L. & HUNTER, J. E. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology*, 86, 162–173.
- SCHOENFELD, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics. In D. GROUWS (Ed.): *Handbook of research on mathematics teaching and learning* (p. 334–370). New York: Macmillan.
- SCHOENFELD, A. H. (2013). Reflections of problem solving in theory and practice. *The Mathematics Enthusiast*, 10(1&2), 9–34.
- SCHUMACHER, L. (2002). *Emotionale Befindlichkeit und Motive in Lerngruppen*. Hamburg: Kovač.
- SCHWAB, M. (2012). *Controlling-Aufgaben und Controlling-Leitbild in der Ausbildung zum Industriekaufmann*. Unveröff. Masterarbeit, Universität Mannheim.
- SEEBER, S. (2008). Ansätze zur Modellierung beruflicher Fachkompetenz in kaufmännischen Ausbildungsberufen. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 104(1), 74–97.
- SEMBILL, D. (1984). *Modellgeleitete Interaktionsanalysen im Rahmen einer forschungsorientierten Lehrerbildung – am Beispiel von Untersuchungen zum „Kaufvertrag“*. Berichte des Seminars für Wirtschaftspädagogik der Georg-August-Universität Göttingen, Band 7, Göttingen.
- SEMBILL, D. (1992). *Problemlösefähigkeit, Handlungskompetenz und Emotionale Befindlichkeit – Zielgrößen forschenden Lernens*. Göttingen: Hogrefe.
- SEMBILL, D. (1995). Der Wille zum Nicht-Müssen – Gestaltungskraft im Spannungsverhältnis von Innovation und Organisation. In G. P. BUNK & R. LASSAHN (Hrsg.), *Festschrift für Artur Fischer zum 75. Geburtstag* (S. 125–146). Steinbach bei Gießen: Ehgart & Albohn.
- SEMBILL, D. (2008). Führung und Zeit – gesellschaftliche, institutionelle und unterrichtliche Perspektiven. In J. Warwas & D. Sembill (Hrsg.), *Zeit-gemäße Führung – zeitgemäßer Unterricht* (S. 81–98). Baltmannsweiler: Schneider Hohengehren.
- SEMBILL, D. (2010). Emotionen – Auslöser, Begleiter und Ziele individuellen und sozialen Handelns. In R. NICKOLAUS, G. PÄTZOLD, H. REINISCH & T. TRAMM (Hrsg.), *Handbuch der Berufs- und Wirtschaftspädagogik* (S. 80–84). Bad Heilbrunn: Klinkhardt.
- SEMBILL, D., RAUSCH, A. & KÖGLER, K. (2013). Non-Cognitive Facets of Competence: Theoretical Foundations and Implications for Measurement. In K. BECK/O. ZLATKIN-TROITSCHANSKAIA (Eds.): *From Diagnostics to Learning Success. Proceedings in Vocational Education and Training* (pp. 199–211). Rotterdam: Sense Publishers B.V.
- SLOANE, P. E. (2008). Vermessene Bildung – Überlegungen zur Entwicklung des Deutschen Qualifikationsrahmens. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 104(4), 481–502.
- SMITH, C. A., ORGAN, D. W. & NEAR, J. P. (1983). Organizational citizenship behavior: Its nature and antecedents. *Journal of Applied Psychology*, 68(4), 653–663.
- STENSTRÖM, M.-L. & TYNJÄLÄ, P. (Eds.). (2009). *Towards Integration of Work and Learning. Strategies for Connectivity and Transformation*. Berlin: Springer.
- STOCK-HOMBURG, R. (2013). Zukunft der Arbeitswelt 2030 als Herausforderung des Personalmanagements. In R. Stock-Homburg (Hrsg.), *Handbuch Strategisches Personalmanagement* (2. überarb. und erw. Aufl.) (S. 603–629). Wiesbaden: Springer Gabler. Wiesbaden.

- VAN GOG, T. (2012). Expertise. In N. M. SEEL (Ed.), *Encyclopedia of the sciences of learning* (pp. 1238–1240). New York, NY: Springer.
- VERSCHAFFEL, L., VAN DOOREN, W. & DE SMEDT, B. (2012). Mathematical learning. In N. M. SEEL (Ed.), *Encyclopedia of the sciences of learning* (pp. 2121–2123). New York: Springer.
- WAINER, H., BRADLOW, E. T. & WANG, X. (2007). *Testlet response theory and its applications*. New York: Cambridge University Press.
- WANG, W. C., & WILSON, M. (2005). Exploring local item dependence using a random-effects facet model. *Applied Psychological Measurement*, 29, 296–318.
- WEINERT, F. E. (2001). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In F. E. WEINERT (Hrsg.), *Leistungsmessung in Schulen* (S. 17–32). Weinheim: Beltz.
- WEISS, R. H. (2006). *CFT 20-R. Grundintelligenztest Skala 2 – Revision*. Göttingen: Hogrefe.
- WINTHER, E. (2010). *Kompetenzmessung in der beruflichen Bildung*. Bielefeld: Bertelsmann.
- WOLF, K. D. (2003). *Gestaltung und Einsatz einer internetbasierten Lernumgebung zur Unterstützung Selbstorganisierten Lernens*. Hamburg: Kovač.
- WUTTKE, E. (1999). *Motivation und Lernstrategien in einer selbstorganisationsoffenen Lernumgebung. Eine empirische Untersuchung bei Industriekaufleuten*. Frankfurt am Main: Lang.
- WUTTKE, E. & WOLF, K. D. (2007). Entwicklung eines Instruments zur Erfassung von Problemlösefähigkeit – Ergebnisse einer Pilotstudie. *Europäische Zeitschrift für Berufsbildung*, 41(2), 99–118.
- ZLATKIN-TROITSCHANSKAIA, O., BECK, K., SEMBILL, D., NICKOLAUS, R. & MULDER, R. (Hrsg.) (2009). *Lehrprofessionalität. Bedingungen, Genese, Wirkungen und ihre Messung*. Weinheim: Beltz.

Anschrift der Autorinnen und Autoren:

Prof. Dr. Eveline Wuttke, Professur für Wirtschaftspädagogik, insbesondere empirische Lehr-Lern-Forschung, Goethe Universität Frankfurt, Theodor-W.-Adorno-Platz 4, 60629 Frankfurt, wuttke@em.uni-frankfurt.de

Prof. Dr. Jürgen Seifried, Lehrstuhl für Wirtschaftspädagogik II, Universität Mannheim, L 4,1, 68131 Mannheim, seifried@bwl.uni-mannheim.de

Steffen Brandt, Wissenschaftszentrum Kiel, Fraunhoferstraße 13, 24118 Kiel, info@artofreduction.com

Dr. Andreas Rausch, Otto-Friedrich-Universität Bamberg, Lehrstuhl für Wirtschaftspädagogik, Kärntenstraße 7, 96052 Bamberg, andreas.rausch@uni-bamberg.de

Prof. Dr. Detlef Sembill, Otto-Friedrich-Universität Bamberg, Lehrstuhl für Wirtschaftspädagogik, Kärntenstraße 7, 96052 Bamberg, detlef.sembill@uni-bamberg.de

Dr. Thomas Martens, Deutsches Institut für Internationale Pädagogische Forschung (DIPF), Technologiebasiertes Assessment, Solmsstraße 73, 60486 Frankfurt, martens@dipf.de

Prof. Dr. Karsten Wolf, Universität Bremen, Fachbereich 12 (Erziehungswissenschaften), Postfach 330440, 28359 Bremen, wolf@uni-bremen.de