

Kompetenzorientierte Prüfungserstellung

Eine digitale Workbench zur Konstruktion von Prüfungsaufgaben in der kaufmännischen Berufsbildung



ESTHER WINTHER
Prof. Dr., Professorin an der
Universität Duisburg Essen
esther.winther@uni-due.de



JOANA REIMER
Wiss. Mitarbeiterin an der
Universität Duisburg Essen
joana.reimer@uni-due.de



ANNE-CATHRIN VONARX
Wiss. Mitarbeiterin an der
Universität Duisburg-Essen
anne-cathrin.vonarx@
uni-duisburg-essen.de

Der Beitrag zeigt Möglichkeiten einer digitalen Workbench zum Erstellen kaufmännischer Prüfungsaufgaben auf. Das Kernelement der Workbench ist eine umfangreiche Datenbank, die empirisch geprüfte Aufgaben und Gesamtprüfungen enthält und als Grundlage für die Konstruktion neuer kompetenzorientierter Aufgaben dient. Im Fokus des Beitrags stehen dabei zwei zentrale Herausforderungen: Prüfungsaufgaben authentisch zu erstellen und den Schwierigkeitsgrad der Aufgaben zu bestimmen. Hierzu wird erläutert, wie die Workbench dazu beiträgt, diese Herausforderungen erfolgreich handhaben zu können.

Herausforderungen bei der Erstellung standardisierter Prüfungsaufgaben

Berufliche Abschlussprüfungen spielen eine wichtige Rolle: Sie erfassen die berufliche Handlungsfähigkeit und eröffnen für die erfolgreich Teilnehmenden den Zugang in eine qualifizierte Beschäftigung. Gleichzeitig bieten Abschlusszeugnisse ein verlässliches, vor allem aber ein von allen Beteiligten anerkanntes Testat zur beruflichen Leistungsfähigkeit der jeweiligen Inhaber/-innen.

In der Wissenschaft ist mittlerweile gut belegt, dass externe, also lernortunabhängig erstellte, zentrale Prüfungen eine qualitätssichernde bzw. -steigernde Funktion auf das vorauslaufende (Aus-)Bildungssystem und die Leistungen der Prüflinge haben (vgl. WEIB 2011; WÖRMANN 2018). Hinterfragt werden jedoch aus Forschung und Praxis mitunter folgende Punkte:

- Bauen sich die Prüfungsaufgaben über für einen konkreten Beruf authentische Situationen auf, die beruflich adäquates Handeln in Geschäftsprozessen erfassbar machen (Authentizität als Designkriterium)?
- Bilden die Prüfungsaufgaben die – entsprechend den schulischen Rahmenlehrplänen und den betrieblichen Ausbildungsordnungen – zu erwartenden Leistungsfähigkeiten der Auszubildenden in einem konkreten Beruf angemessen ab, sodass die Gesamtprüfung eine zufallsfreie Prognose beruflicher Handlungsfähigkeit

ermöglicht (schwierigkeitsgenerierende Merkmale als Designkriterium)?

- Sind die gestellten Prüfungsaufgaben logisch und verständlich (Plausibilität) und gewähren eine zuverlässige, valide und faire Bewertung (psychometrische Standards der Auswertungsheuristiken)? (vgl. ausführlich WINTHER 2011).

Auch die Rechtssicherheit von Prüfungen ließe sich durch eine forschungsbegleitete Anpassung der Prüfungserstellung und der Prüfungsauswertung verbessern (wie beispielsweise im DFG-Schwerpunktprogramm 1293 belegt; vgl. KLOTZ/WINTHER 2016; WINTHER 2011). Hier setzt die Workbench an.

Die digitale Workbench

Die digitale Workbench ist ein Online-Tool für das Prüferehrenamt, das dabei hilft, Prüfungsaufgaben für kaufmännische Ausbildungsberufe kompetenzorientiert und modellkonform zu konstruieren und schlüssig in eine Gesamtprüfung zu überführen. Hierzu werden modelltheoretische und psychometrische Kompetenzannahmen in Geschäftsprozesse und zugehörige Aufgaben übersetzt. Es geht also darum, wie Modelle und psychologische Messungen genutzt werden können, um in Geschäftsprozessen die Leistung (Fähigkeiten, Eigenschaften oder Kenntnisse) ei-

Abbildung 1

Aufbau der Workbench und ausgewählte Funktionalitäten



ner Person zu bewerten. Die Workbench bildet den gesamten Workflow zwischen der zentralen Aufgabenstelle für kaufmännische Abschluss- und Zwischenprüfungen (AkA) und den Aufgabenentwickelnden ab.

Die Workbench hat den Anspruch, ein sich permanent weiterentwickelndes Tool zu sein, das integrierte Informations- und Weiterbildungsangebote, verbesserte Kommunikationskanäle sowie größtmögliche Transparenz im Workflow zwischen allen verantwortlichen Prüfungsbeauftragten¹ anbietet und auf die Bedarfe ehrenamtlicher Prüfender anpasst. Hierfür verfügt die Workbench über drei Funktionsbereiche: »Arbeitsumgebung«, »Archiv & Ressourcen« sowie »Weiterbildung & Support« (vgl. Abb. 1).

Mit einer Workbench arbeiten

Die Workbench wurde auf Basis von User Cases erstellt. Hierbei wurde deutlich, dass die ehrenamtlichen Prüfungserstellenden insbesondere bei der Konstruktion authentischer Prüfungsaufgaben sowie bei der Ex-ante-Bewertung der Schwierigkeitsgrade von Aufgaben (digitale) Unterstützung suchen. Insbesondere im Hinblick auf diese

Herausforderungen bietet die Workbench automatisierte Algorithmen an, die die Gesamtqualität der Prüfung erhöhen.

Authentizität als Designkriterium für die Erstellung von Prüfungsaufgaben

Um berufliche Fachkompetenzen unter einer größtmöglichen Nähe zur beruflichen Ausbildungsrealität erfassen zu können, stellt Authentizität ein wichtiges Kriterium dar (vgl. WINTHER 2010). Der Begriff Authentizität bleibt in der beruflichen Bildung unspezifisch und wird meistens schlagwortartig verwendet (vgl. ACHTENHAGEN/WEBER 2003). Authentizität existiert nicht per se (vgl. WINTHER 2010), ist subjektiv (vgl. HUANG 2002) und muss modelliert werden: In der beruflichen Bildung wird Authentizität als Zielkategorie für die Gestaltung von Lern- und Prüfungsprozessen formuliert, um didaktische Konstruktionen maßgeblich über aus der Berufswelt modellierte Situationen zu definieren (vgl. GULIKERS/BASTIAENS/KIRSCHNER 2004; DEUTSCHER/WINTHER 2018). Vor diesem Hintergrund ist ein grundlegendes Verständnis der Bedeutung von Authentizität wichtig und darüber, wie sie sich in Prüfungsaufgaben realisieren lässt.

In verschiedenen Projekten im Kontext der beruflichen Kompetenzmessung und der kompetenzorientierten Aufgaben- und Prüfungskonstruktion (u. a. ASCOT und ASCOT+)

¹ Die Prüfungsausschüsse sind drittelparitätlich besetzt; ihnen gehören Beauftragte der Arbeitnehmer und Arbeitgeber in gleicher Anzahl sowie mindestens eine Lehrkraft einer berufsbildenden Schule an.

konnte gezeigt werden, dass eine wiederholte Aktivierung des Authentizitätskonzepts für den Erstellungsprozess notwendig ist, da nicht von einer umfassend verankerten Wissensbasis ausgegangen werden kann. Dies bestätigte sich auch im Rahmen eines Wissenschafts-Praxis-Diskurses im September 2022, der qualitativ evaluiert wurde: An diesem Diskurs haben Mitglieder aus zwei Fachausschüssen der Aufgabenstelle für kaufmännische Abschluss- und Zwischenprüfungen (AkA) für die Ausbildungsberufe Industriekaufmann/-frau (Gruppe A; $n_A = 12$) sowie Kaufmann/-frau für Spedition und Logistikdienstleistung (Gruppe B; $n_B = 9$) teilgenommen. Die Befunde zeigen, dass sowohl bei den Spontanassoziationen zu dem Begriff Authentizität als auch bei der Bewertung von Aussagen zu Authentizität fehlerhafte Interpretationen bestehen. Mit dem Begriff Authentizität assoziieren beide Gruppen (Gruppe A: 7/13 sowie Gruppe B: 4/11) Begriffe wie »Realität«, »Wirklichkeit«, »Wahrheit«; die Aussage: »Authentizität ist für mich die Realität, die sich meinen Auszubildenden an ihren Arbeitsplätzen zeigt« erhält die höchste Zustimmung. Gerade die Gleichsetzung von Realität und Authentizität erschwert jedoch den Zugang zu den Konstruktionslogiken sowohl in Instruktions- als auch Assessmentprozessen.

In der Workbench wird dieser Qualifizierungsbedarf zweifach adressiert: Erstens wird Authentizität als inhaltliches Thema in Weiterbildungs- und Fortbildungsangeboten (sowohl in einem Webinar als auch in Microlearnings) behandelt. Hierbei werden den ehrenamtlichen Prüfenden wichtige Prinzipien und Techniken vermittelt, um Aufgaben möglichst authentisch zu formulieren. Dies umfasst beispielsweise die Anwendung von aktuellen Fachkenntnissen und die Berücksichtigung von realen Anwendungssituationen. Zweitens werden über Blaupausen von authentischen Materialien konkrete Vorlagen bereitgestellt, die den Kon-

struktionsprozess direkt unterstützen. Diese Blaupausen dienen als Orientierungshilfen und bieten eine Struktur für die Erstellung von Aufgaben. Sie enthalten beispielhafte Formulierungen, Aufgabentypen und mögliche Lösungswege, die den Prüfenden als Ausgangspunkt dienen können. Auf diese Weise werden sie dabei unterstützt, praxisnahe und authentische Prüfungsaufgaben zu erstellen.

Durch die Kombination beider Unterstützungsangebote ermöglicht die Workbench den ehrenamtlichen Prüfenden eine effektive Gestaltung von Prüfungsaufgaben, die den Anforderungen an Authentizität gerecht werden und den Lernenden eine realitätsnahe Prüfungserfahrung bieten.

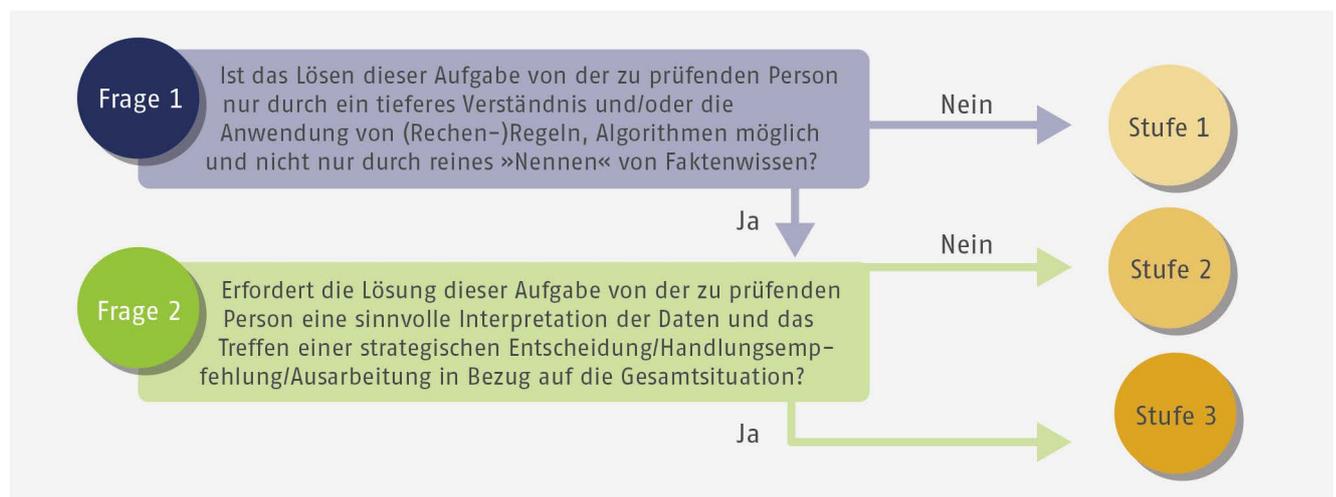
Schwierigkeitsprognostik als Designkriterium für die Erstellung von Prüfungsaufgaben

Um sowohl leistungsstarke als auch leistungsschwache Auszubildende angemessen bewerten zu können, müssen die angebotenen Prüfungsaufgaben in ihrer Schwierigkeit variieren. Hierzu sind die Prüfungsaufgaben nach den Prinzipien der funktionalen Modellierung, der inhaltlichen Spezifität und der Art der kognitiven Beanspruchung zu entwickeln und entsprechend dem angenommenen Antwortverhalten der zu Prüfenden zu codieren (vgl. WINTHER 2010).

- Die funktionale Modellierung bewertet Prüfungsaufgaben dahingehend, wie anspruchsvoll es ist, die zur Lösung notwendigen Schritte aus der zugrundeliegenden Anforderungssituation zu extrahieren. Die Aufgabenkonstruktion wird hier auf *Kontextebene* beurteilt.
- Die inhaltliche Spezifität bewertet Prüfungsaufgaben hinsichtlich des zu ihrer Lösung benötigten (Fach-)Wissens. Die Aufgabenkonstruktion wird hier auf *Inhaltsebene* beurteilt.

Abbildung 2

Entscheidungsbaum zur Einschätzung der Aufgabenschwierigkeit (Beispiel)



- Die Art der kognitiven Beanspruchung bewertet Prüfungsaufgaben mit Blick auf die zu ihrer Lösung erforderlichen Leistungsfähigkeiten der Auszubildenden. Die Aufgabenkonstruktion wird hier auf kognitiver *Prozessebene* beurteilt.

Die systematische Konstruktion von Prüfungsaufgaben entlang der o. g. Merkmale stellt sicher, dass

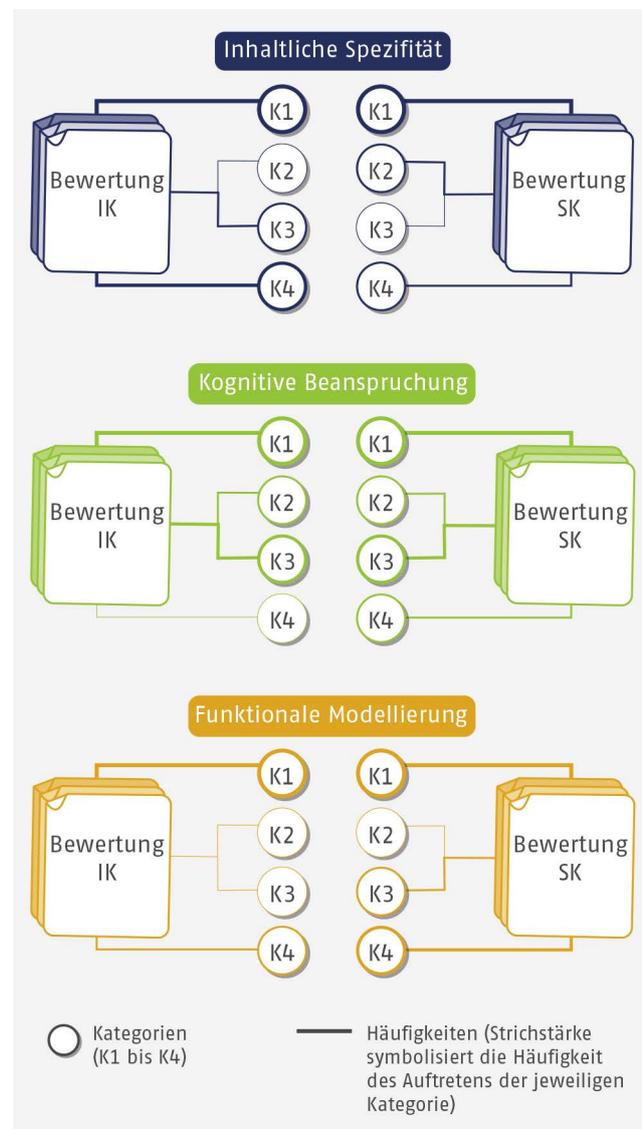
- (1) Annahmen über kognitive Theorien in die Prüfungsaufgaben einfließen und
- (2) Prüfungsaufgaben formuliert werden, die gut zwischen den zu prüfenden Personen trennen können, da sie unterschiedliche Fähigkeitsstufen ansprechen.

Über die Workbench sind Microlearnings abrufbar, die in die Aufgabenkonstruktion mittels schwierigkeitsgenerierender Merkmale einführen. Zur Konstruktion der Lerneinheiten wurden Vor-Ort-Trainings durchgeführt, um die Bedarfe der ehrenamtlichen Prüfungserstellenden in Lerninhalte zu übersetzen und um das didaktische Konzept zu prüfen. Ausgangspunkt der Trainings ist eine standardisierte Metrik, die eine Ex-ante-Bewertung der Aufgabenschwierigkeiten entlang einfacher Entscheidungsbäume ermöglicht (Abb. 2 stellt beispielhaft den Entscheidungsbaum für das Merkmal »Art der kognitiven Beanspruchung« dar).

In den Vor-Ort-Trainings, die im September 2022 stattfanden, wurden ehrenamtliche Prüfungserstellende der Ausbildungsberufe Industriekaufmann/-frau (Gruppe A; $n_A = 12$) sowie Kaufmann/-frau für Spedition und Logistikdienstleistung (Gruppe B; $n_B = 9$) darin fortgebildet, Prüfungsaufgaben anhand schwierigkeitsgenerierender Merkmale zu bewerten. Hierzu wurden ihnen die Entscheidungsbäume in Kombination mit Prüfungsaufgaben präsentiert. In der Transferphase, die sich an das Training anschloss, waren die Prüfungserstellenden aufgefordert, verschiedene Prüfungsaufgaben hinsichtlich ihrer schwierigkeitsgenerierenden Merkmale (richtig) zu bewerten. Der Bewertungsprozess wurde beobachtet und die Diskussion sprachlich aufgezeichnet, um Aussagen darüber treffen zu können, wie gut die Entscheidungsbaum-Metrik verstanden wird und ob sie sich als prognostisches Instrument zur A-priori-Schwierigkeitsbewertung eignet. Aus dem Diskussionsverlauf lassen sich verschiedene Kategorien herausfiltern, anhand derer sich aufklären lässt, wie die ehrenamtlichen Prüfenden ihren Bewertungsprozess strukturiert haben:

- auf den Entscheidungsbaum zurückgegriffen (K1)
- Prüfungsaufgaben im Vergleich betrachtet (K2)
- bekannte Operatoren und eigene Routinen herangezogen (K3)
- Eigenschaften des jeweiligen schwierigkeitsgenerierenden Merkmals unabhängig vom Entscheidungsbaum diskutiert (K4)

Abbildung 3
Prozess der Bewertung von Prüfungsaufgaben anhand ausgewählter Kriterien



In Abbildung 3 sind die Häufigkeiten, mit denen die jeweiligen Kategorien im Bewertungsprozess auftraten und zu einer richtigen Bewertung geführt haben, auf Basis von Strichstärken visualisiert.

Mit Blick auf die einzelnen schwierigkeitsgenerierenden Merkmale konnten folgende Befunde gesichert werden:

- Zur Bewertung der *inhaltlichen Spezifität* haben alle Aufgabenerstellenden bei jeder zu bewertenden Prüfungsaufgabe den Entscheidungsbaum (K1) als einzige oder zentrale Orientierung verwendet. Die zweithäufigste Entscheidungsgrundlage stellt die Vertrautheit mit dem Fachwissen dar (K4); im Bewertungsprozess wird folglich auf eine charakteristische Eigenschaft des schwierigkeitsgenerierenden Merkmals zurückgegriffen.
- Zur Bewertung der *kognitiven Beanspruchung* wurde der Entscheidungsbaum (K1) als erste und grundlegende

Anlaufstelle verwendet. Bei diesem schwierigkeitsgenerierenden Merkmal greifen die Prüfungsaufgabenersteller/-innen zudem auf ihnen bekannte Konzepte und Routinen der Aufgabenerstellung zurück (K3).

- Die *funktionale Modellierung* wird unter expliziter Nutzung des Entscheidungsbaums (K1) bewertet. Bei der Bewertung der funktionalen Modellierung ist auffallend, dass zusätzlich die wahrgenommene Komplexität der zu bewertenden Prüfungsaufgaben (K4) diskutiert wurde.

Die Befunde zeigen, dass eine weitgehend standardisierte Metrik die Prüfungserstellenden bei der Bewertung von Aufgabenschwierigkeiten unterstützen kann und sich über die Metrik sowie über charakteristische Eigenschaften einzelner schwierigkeitsgenerierender Merkmale zwischen den Prüfungserstellenden sehr schnell ein Konsens erzeugen lässt, der die geschätzte Aufgabenschwierigkeit sehr nah an die empirisch ermittelte Aufgabenschwierigkeit heranführt. Die Studie zeigt, welche Bedeutung ein einheitlicher Referenzrahmen haben kann: Er schafft Orientierung und ist eine zentrale Grundlage für die reliable Prüfungserstellung.

Eine digitale Workbench als handlungsleitendes Instrument für die Erstellung kompetenzorientierter Prüfungsaufgaben

Die Workbench bildet einen ersten Orientierungsrahmen zur standardisierten Aufgabenerstellung, unterstützt hierdurch eine Vereinheitlichung bezüglich der Anforderungen und trägt dazu bei, dass die Auszubildenden unter vergleichbaren Bedingungen geprüft werden.

Perspektivisch kann die Workbench auf Basis der vorgestellten Überlegungen und Erkenntnisse immer weiterentwickelt werden und zur Erstellung von sowohl Prüfungsaufgaben als auch Gesamprüfungen bezüglich der Standardisierung, Qualität sowie Effizienz beitragen.

Die Workbench setzt erste Impulse zur Digitalisierung des Prüfungswesens. Hierdurch wird eine effektivere Konstruktion von Prüfungsaufgaben ermöglicht. Aufgaben können schnell aktualisiert und den aktuellen Bedürfnissen angepasst werden. Dies trägt dazu bei, dass Prüfungen relevant und auf dem neuesten Stand bleiben. Ebenso ergeben sich künftig neue Möglichkeiten zur Prüfungsdurchführung wie beispielsweise Online-Prüfungen.

Darüber hinaus wird durch die Digitalisierung eine effizientere Kollaboration gefördert. Hier kann die Workbench als Plattform dienen, um den Austausch und die Validierung von Prüfungen zu fördern. Prüfungsaufgaben können in digitaler Form zwischen verschiedenen Institutionen und auch Berufsfeldern ausgetauscht werden. Dies führt zu einer breiteren Perspektive und einer verbesserten Vergleichbarkeit von Kompetenzen.

Durch diese perspektivischen Ansätze kann die digitale Workbench sowohl einen bedeutenden Beitrag zum Erstellungsprozess leisten als auch innovative Ansätze zur Lernförderlichkeit aufzeigen und einen einfacheren und effizienteren Austausch ermöglichen.

Zusammenfassend kann die Nutzung der digitalen Workbench zur Erstellung von sowohl Prüfungsaufgaben als auch Gesamprüfungen dazu beitragen, einheitliche, authentische und angemessen anspruchsvolle Prüfungen zu entwickeln. Durch die Effizienz, Anpassungsfähigkeit und Lernförderlichkeit der digitalen Workbench wird das Prüfungswesen professionalisiert und werden die Potenziale der Digitalisierung genutzt, um die Qualität und Relevanz der Prüfungen immer weiter zu verbessern. ◀



Abbildungen von S. 17 und S. 18 zum Download
Abb. 1: www.bwp-zeitschrift.de/g12063
Abb. 2: www.bwp-zeitschrift.de/g12066

LITERATUR

ACHTENHAGEN, F.; WEBER, S.: »Authentizität« in der Gestaltung beruflicher Lernumgebungen. In: BREDOW, A.; DOBISCHAT, R.; ROTTMANN, J. (Hrsg.): Berufs- und Wirtschaftspädagogik von A bis Z. Baltmannsweiler 2003, S. 185–199

DEUTSCHER, V.; WINTHER, E.: A Conceptual Framework for Authentic Competence Assessment in VET: A Logic Design Model. In: MCGRATH, S.; MULDER, M.; PAPIER, J.; SUART, R. (Hrsg.): Handbook of Vocational Education and Training. Cham 2018

GULIKERS, J. T.; BASTIAENS, T. J.; KIRSCHNER, P. A.: A five-dimensional framework for authentic assessment. In: Educ Technol Res Dev 52 (2004) 3, S. 67–86

HUANG, H. M.: Towards constructivism for adult learners in online learning environments. In: British Journal of Educational Technology (2002) 33, S. 27–37

KLOTZ, V.K.; WINTHER, E.: Zur Entwicklung domänenverbundener und domänenspezifischer Kompetenz im Ausbildungsverlauf. In: Zeitschrift für Erziehungswissenschaft 19 (2016) 4, S. 765–782

WEIß, R.: Prüfungen in der beruflichen Bildung – ein vernachlässigter Forschungsgegenstand. In: SEVERING, E.; WEIß, R. (Hrsg.): Prüfungen und Zertifizierung in der beruflichen Bildung – Anforderungen – Instrumente – Forschungsbedarf. Bielefeld 2011, S. 37–52

WINTHER, E.: Kompetenzmessung in der beruflichen Bildung. Bielefeld 2010

WINTHER, E.: Kompetenzorientierte Assessments in der beruflichen Bildung – Am Beispiel der Ausbildung von Industriekaufleuten. In: Zeitschrift für Berufs- und Wirtschaftspädagogik 107 (2011) 1, S. 33–54

WÖBMAN, L.: Central Exit Exams Improve Student Outcomes (IZA World of Labor 419). München 2018