

Profil 8:

Netzwerke – Strukturen von Wissen, Akteuren und Prozessen in der beruflichen Bildung

**Digitale Festschrift
für BÄRBEL FÜRSTENAU**



Eveline WUTTKE¹, Susan SEEBER², Hanna MEINERS² & Lütfiye TURHAN¹

(¹ Universität Frankfurt, ² Universität Göttingen)

Entwicklung und Evaluation eines Trainingskonzeptes für die Gestaltung problemhaltiger technologiebasierter Lern- und Prüfungsaufgaben

Online unter:

https://www.bwpat.de/profil8_fuerstenaу/wuttke_et_al_profil8.pdf

in

bwp@ Profil 8 | September 2023

Netzwerke – Strukturen von Wissen, Akteuren und Prozessen in der beruflichen Bildung

Teil 2: Design und Evaluation konstruktivistischer Lehr-Lern-Arrangements

Hrsg. v. **Mandy Hommel, Carmela Aprea & Karin Heinrichs**

www.bwpat.de | ISSN 1618-8543 | bwp@ 2001–2023



www.bwpat.de



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

Berufs- und Wirtschaftspädagogik - online

Entwicklung und Evaluation eines Trainingskonzeptes für die Gestaltung problemhaltiger technologiebasierter Lern- und Prüfungsaufgaben

Abstract

In kaufmännischen Curricula spiegeln sich schon seit geraumer Zeit Ausbildungsziele wider, die auf die Herausbildung und Förderung komplexer (Problemlöse-)Kompetenzen gerichtet sind (z. B. im Rahmenlehrplan für Industriekaufleute (Kultusministerkonferenz 2002)). Auch mit Blick auf die digitalisierungsbezogenen Veränderungen der Kompetenzanforderungen an kaufmännische Arbeitsplätze, die von einer Abnahme an Routineaufgaben und einer Zunahme von problemorientierten Aufgaben ausgehen (vgl. ausgewählte Beiträge in Schumann/Seeber/Abele 2022), sind diese Kompetenzen stärker zu adressieren. Neu geordnete Curricula des kaufmännischen Bereichs, z. B. des Ausbildungsberufs der Bankkaufleute, haben diese Entwicklungen aufgegriffen und zielen daher nochmals verstärkt auf Problemlösekompetenz ab.

Eine konsequente Umsetzung dieses anspruchsvollen curricularen Ziels kann jedoch nur dann systematisch erreicht werden, wenn curriculare Vorgaben, Unterricht in der Berufsschule und Lernen im Betrieb sowie Zwischen- und Abschlussprüfungen den von Pellegrino (2010) geforderten Zusammenhang aufweisen (vgl. dazu auch Biggs (1996), constructive alignment). Insbesondere die Prüfungspraxis steht in der Kritik. So wiesen Weber und Achtenhagen (2014) schon vor einiger Zeit auf die fehlende, jedoch notwendige Verknüpfung von Curriculum, Instruktion und Assessment hin. Auch eine aktuelle Analyse von Aufgaben aus kaufmännischen Zwischen- und Abschlussprüfungen zeigt, dass nach wie vor überwiegend Wissen abgefragt wird und die berufliche Problemlösekompetenz eine untergeordnete Rolle spielt (Wuttke et al. 2022). Für das berufliche Bildungspersonal stellt sich also sowohl für die Instruktion als auch für die Lernprozessdiagnostik und die Zwischen- und Abschlussprüfungen die Frage, wie problemhaltige Lern- und Prüfungsaufgaben zu konzipieren sind, um die angestrebte Kompetenz gezielt zu fördern und zu erfassen.

Um Auszubildende und Lehrende im kaufmännischen Bereich zur Entwicklung problemhaltiger (Prüfungs-)Aufgaben zu befähigen, wurde ein Trainingskonzept entwickelt. Durch das Training wurden die genannten Zielgruppen geschult, Aufgaben nach ihrem Problemgehalt zu beurteilen und auszuwählen sowie problemhaltige Aufgaben unter Nutzung digitaler Medien und einer digitalen Lehr-Lern-Plattform selbst zu konzipieren. Das Training wurde mit N=152 (angehenden) Lehrkräften und Auszubildenden zwischen November 2020 und Oktober 2022 durchgeführt. Die Ergebnisse der Evaluation zeigen, dass die Erwartungen der Teilnehmenden an das Training bezüglich des Erstellens von problemhaltigen Prüfungsaufgaben erfüllt wurden und dass sie mit dem Training zufrieden sind. Die Teilnehmenden weisen zudem einen signifikanten Wissenszuwachs bei den Inhalten der Trainingsbausteine auf. Die von den Trainingsteilnehmenden entwickelten Aufgaben wurden an einer Stichprobe von Auszubildenden (N=1.416) erprobt (Gütekriterien, Schwierigkeitsgrad).

Development and evaluation of a training concept for enhancing skills in the design of problem-oriented technology based learning and examination tasks

As the commercial workplace becomes increasingly digitized, routine tasks are being replaced with more complex tasks requiring problem-solving skills. In response to the growing demand for a workforce equipped with such skills, the learning objectives of current curricula for many post-secondary vocational education programs in Germany now reflect a growing trend to promote the development of problem-solving competences. However, recent research shows that theoretical in-class instruction, practical in-company training, and assessments often are misaligned and assessments of learning remain predominantly knowledge-based. To be able to promote and track the development of problem-solving skills in a commercial context, both in-school and in-company vocational instructors need to know how to create valid and reliable assessment items. We developed a technology-based intervention for vocational instructors in the commercial sector to learn how to design problem-based assessment items using digital media and a digital learning platform. First, we taught the N=152 (prospective) teachers and instructors how to assess pre-made tasks according to their problem-related content. Next, we taught them how to design problem-based tasks themselves. The tasks created by the participants were then administered to a sample of trainees in vocational education (N=1.416) whose performance was analyzed to determine the appropriateness and effectiveness of the items. The vocational instructors demonstrated a significant increase in knowledge of the content covered upon completion of the modules.

The tasks developed by the training participants were tested on a sample of trainees (N=1.416) (quality criteria, level of difficulty).

***Schlüsselwörter:** Professionalisierung von Lehr- und Prüfungspersonal, Problemlösekompetenz, Technologiebasierung, Training*

***Keywords:** Professionalization of teaching and examination staff, problem-solving skills, use of technology, training*

1 Notwendigkeit eines Trainings zur Erstellung problemhaltiger technologiebasierter Aufgaben: Einleitung

Wenn es um die Gestaltung von komplexen Lehr-Lern-Arrangements und problembasiertem Lernen in verschiedenen Kontexten wie Schule und Hochschule geht, ist Bärbel Fürstenau seit vielen Jahren Expertin (z. B. Fürstenau 1994; Fürstenau 1999; Schlicht/Hommel/Fürstenau 2022). Mit Blick auf die im Abstract angesprochene Triade von Curriculum, Instruktion und Assessment hat sie sich insbesondere der Instruktion (natürlich unter Berücksichtigung der curricularen Gegebenheiten) gewidmet. Unser Projekt nimmt dagegen das Assessment in den Blick und wir knüpfen dabei gerne an die jahrelange Arbeit an, die sie – in verschiedenen Arbeitskontexten – zum Thema Problemorientierung geleistet hat.

Wenn der Blick auf die drei Teile der Triade gerichtet wird, zeigt sich, dass in der wirtschaftspädagogischen Praxis, der Ausbildung von Kaufleuten, die Problemorientierung in sehr unterschiedlichem Maße umgesetzt ist. So spiegeln sich in kaufmännischen Curricula schon seit geraumer Zeit Ausbildungsziele wider, die auf die Herausbildung und Förderung komplexer (Problemlöse-)Kompetenzen gerichtet sind (z. B. im Rahmenlehrplan für Industriekaufleute (Kultusministerkonferenz 2002)). Aktuell verstärkt sich diese Entwicklung noch, was mit Blick auf die durch die Digitalisierung sich verändernden Kompetenzanforderungen an kaufmännischen Arbeitsplätzen gut nachvollziehbar ist. Wenn Routineaufgaben abnehmen und eine Zunahme von problemorientierten Anforderungen an kaufmännischen Arbeitsplätzen zu beobachten ist (vgl. Geiser et al. 2021; ausgewählte Beiträge in Schumann/Seeber/Abele 2022), ist Problemlösekompetenz stärker zu adressieren. Neu geordnete Curricula des kaufmännischen Bereichs, z. B. der Ausbildungsberuf der Bankkaufleute, zielen daher nochmals stärker auf Problemlösekompetenz ab.

Eine konsequente Umsetzung dieses anspruchsvollen curricularen Zieles wird jedoch nur dann systematisch erreicht werden, wenn neben den curricularen Vorgaben auch der Unterricht in der Berufsschule und das Lernen im Betrieb sowie die Zwischen- und Abschlussprüfungen den von Pellegrino (2010) geforderten Zusammenhang aufweisen (vgl. dazu auch Biggs (1996), *constructive alignment*). Insbesondere die Prüfungspraxis steht dabei schon seit längerem in der Kritik und es wird bemängelt, dass es an der notwendigen Verknüpfung von Curriculum, Instruktion und Assessment fehlt (Weber/Achtenhagen 2014). Der Schluss liegt nahe, dass die mangelnde Passung der Prüfungen zu den curricularen Vorgaben und (teilweise) der instruktionalen Umsetzung auf fehlende Kompetenzen des Prüfungspersonals bei der Entwicklung und Umsetzung problemhaltiger Aufgaben zurückzuführen ist. Diese Annahme war Ausgangspunkt für die Entwicklung eines entsprechenden Trainings. Um die Annahme der unzureichenden Aufgabenqualität zu prüfen, wurde im Vorfeld eine Analyse schriftlicher Abschlussprüfungen der Ausbildungsberufe Industriekaufmann*frau (IK) und Kaufmann*frau für Büromanagement (KBM) durchgeführt. Die beiden Ausbildungsberufe wurden gewählt, da (1) im Falle der IK ein Beruf erfasst wird, in dem sich zentrale kaufmännische Kerntätigkeiten widerspiegeln und der in vielfältiger Weise curriculare Überschneidungen zu anderen kaufmännischen Berufen aufweist, d. h. als kaufmännischer Kernberuf gilt (vgl. Seifried et al. 2016), (2) KBM ein Ausbildungsberuf ist, der in vielfältigen Branchen (Handwerk, Industrie, Handel und Öffentlicher Dienst) und Unternehmenstypen (Kleinst- bis Großunternehmen) für unterschiedliche kaufmännische Handlungsfelder ausgebildet wird und (3) beide Berufe zu den quantitativ bedeutsamen kaufmännischen Berufen zählen (vgl. Datenbank DAZUBI; Bundesinstitut für Berufsbildung 2019). Darüber hinaus ist die Anforderung, (komplexe) berufliche Probleme zu lösen, Gegenstand der Curricula beider Ausbildungsberufe.

Analysiert wurden 1.454 Aufgaben aus schriftlichen Zwischen- und Abschlussprüfungen der Jahrgänge 2015/16 bis 2019 der beiden Ausbildungsberufe. Die Auswertung hinsichtlich des Problemgehalts der Aufgaben sowie Nähe zu berufsbezogenen Problemlöseprozessen erfolgte mittels einer qualitativen Inhaltsanalyse. Die Analyse bestätigt die bisherige Kritik an der fehlenden Problemhaltigkeit von Prüfungsaufgaben und machen erneut deutlich, dass bisherige

Prüfungsaufgaben für beide Berufe kaufmännische Problemsituationen nicht hinreichend abbilden und damit Problemlösekompetenz nicht hinreichend messen (Wuttke et al. 2022). Ein Ansatzpunkt zur Erklärung der vor allem in Prüfungsaufgaben unzureichend adressierten Problemhaltigkeit besteht darin, dass die Entwicklung komplexer problemhaltiger Prüfungsaufgaben und ihre technologiebasierte Umsetzung zur authentischen Abbildung beruflicher Handlungssituationen in der Aus- und Fortbildung von Lehr- und Prüfungspersonal bisher kaum thematisiert wurde. So zeigt eine Suche in Veranstaltungsdatenbanken zur Fortbildung für die Zielgruppe, dass kaum Veranstaltungen existieren, die explizit darin schulen, problemhaltige schriftliche Prüfungsaufgaben zu erstellen. Fortbildungsangebote der Kammern hingegen stellen vor allem auf die Handlungsorientierung ab, wobei Fragen der Situierung und des vollständigen Handlungszyklus eine zentrale Rolle einnehmen (vgl. ausgewählte Beiträge in BWP 3 2014; Lorig et al. 2014).

Diese Erkenntnis bildete den Ausgangspunkt für die Entwicklung und Umsetzung des Trainings zur Erstellung problemhaltiger und technologiebasierter Aufgaben für (angehende) Lehrende, Auszubildende sowie Prüfende in kaufmännischen Berufen. Eine computerbasierte Umsetzung wurde gewählt, weil sich damit berufliche Anforderungssituationen authentischer modellieren lassen und der Computer als digitales Endgerät zudem ein zentrales Werkzeug für kaufmännische Auszubildende darstellt.

Das Training wurde so konzipiert, dass (1) insbesondere Merkmale problemhaltiger Aufgaben (siehe Abschnitt 2.2) thematisiert wurden. Um dabei eine ansprechende Gestaltung und effektive Nutzung digitaler Medien sicherzustellen, wurde das Trainingskonzept zudem um eine (2) mediendidaktische Schulung erweitert. Daneben wurden auch (3) diagnostische Aspekte, insbesondere diagnostische Gütekriterien, thematisiert. Als Tool zur technologiebasierten Erstellung der Aufgaben wurde auf eine (4) digitale Lehr-Lern-Plattform zurückgegriffen. Eine Schulung zu deren Nutzung war weiterer Bestandteil des Trainings. In den folgenden Abschnitten skizzieren wir die Grundlagen der Trainingsentwicklung und die Ausgestaltung der Trainingsbausteine (Kapitel 2), beschreiben die methodische Umsetzung der Evaluation (Kapitel 3) und deren Ergebnisse (Kapitel 4). Abschließend diskutieren wir die Befunde und geben einen Ausblick auf Implikationen und weitere Forschungsarbeiten (Kapitel 5).

2 Grundlagen der Trainingsentwicklung

2.1 Berücksichtigung von Kriterien erfolgreicher Trainings

Von der Planung, über die Konzeption bis hin zur Durchführung des Trainings wurden einschlägige Kriterien erfolgreicher Trainings berücksichtigt:

- (1) Im *Trainingsangebot* wurde die Relevanz der Thematik hervorgehoben, was sich positiv auf die Fortbildungsmotivation auswirken sollte (Gräsel/Fussangel/Schellenbach-Zell 2008; Lipowsky 2011, 402; Lipowsky/Rzejak 2021, 59ff.). Grundsätzlich kann jedoch ohnehin von einer eher hohen Motivation ausgegangen werden, weil die Gestaltung von Aufgaben einen Kern der Lehr- und Prüfungspraxis bildet. Dazu kommt, dass

durch die Pandemie die Relevanz von Lehr-Lern-Plattformen, digitalen Lernumgebungen und technologiebasierten Aufgaben noch deutlicher hervorgehoben wurde.

- (2) Bei der *Gestaltung der Trainingsmodule* wurde der Forschungsstand zu effektiven Trainings berücksichtigt und Input- sowie Anwendungsphasen im Wechsel angeboten (Cademartori et al. 2017, Krille 2019). Zudem wurde immer wieder zur Reflexion angeregt (z. B. zur bisherigen Prüfungspraxis; Cademartori et al. 2017, 32; Lipowsky 2009, 355).
- (3) Da bisherige Befunde mit Blick auf den *zeitlichen Umfang* von Trainings bislang keinen eindeutigen Schluss zulassen (Cademartori et al. 2017, 31f.; Lipowsky 2009, 350f.), aber zu erwarten ist, dass sehr kurze Maßnahmen kaum langfristig wirksam werden (Lipowsky 2009, 350f.; Lipowsky/Rzejak 2021, 56f.; Seifried/Wuttke 2017, 117f.), wurde das Training über einen mehrtägigen Zeitraum angelegt. Es beinhaltet eine zweitägige Inputphase, eine tutoriell begleitete Selbstlernphase und eine Konsolidierungsphase. In dieser wurde insbesondere die Reflexion der Selbstlernphase durch Expert*innen- und Peer-Feedbacks gefördert, was sich positiv auf den Trainingserfolg auswirken sollte (Cademartori et al. 2017, 42; Lipowsky 2010, 63; Lipowsky/Rzejak 2021, 53f.).

Zur Unterstützung des Trainings und für die Bereitstellung aller Materialien in der Nachtrainingsphase wurde die Plattform ILIAS eingesetzt, die auch Bestandteil des Trainings ist.

2.2 Grundlagen der Entwicklung der Trainingsbausteine

Bei der Gestaltung der Trainingsbausteine (siehe Abb. 1) lag der Schwerpunkt auf der Aufgabenerstellung und dabei auf Merkmalen problemhaltiger Aufgaben (Trainingsbaustein 1). Um dabei auch eine ansprechende Gestaltung und effektive Nutzung digitaler Medien sicherzustellen, wurde zudem eine mediendidaktische Schulung angeboten (Trainingsbaustein 2). Als Tool zur technologiebasierten Erstellung der Aufgaben wurde auf eine digitale Lehr-Lern-Plattform (ILIAS bzw. Moodle) zurückgegriffen, deren Nutzung in Trainingsbaustein 3 thematisiert wurde. Zusätzlich wurden diagnostische Aspekte, insbesondere diagnostische Gütekriterien, thematisiert (Trainingsbaustein 4).

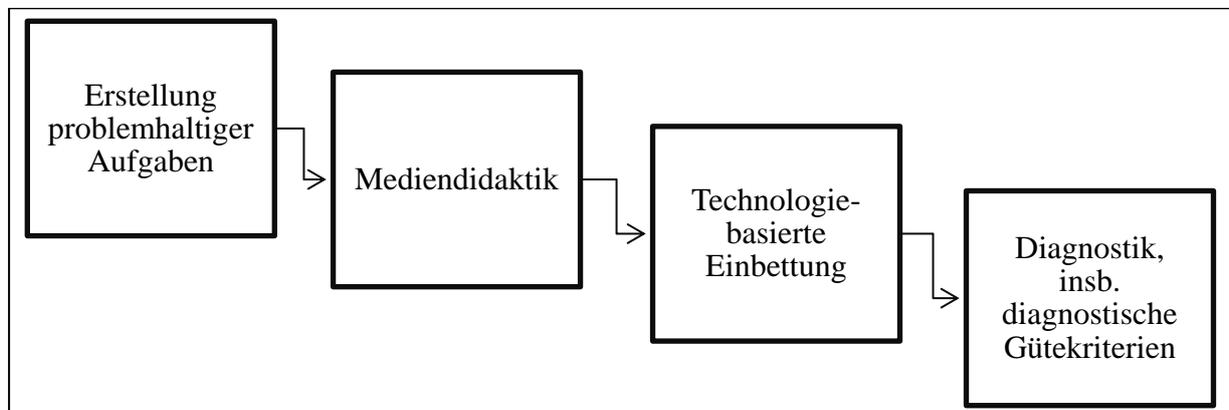


Abbildung 1: Aufbau des Trainings

Trainingsbaustein 1: Merkmale problemhaltiger Aufgaben

Bevor auf Merkmale problemhaltiger Aufgaben Bezug genommen werden kann, muss geklärt werden, was Probleme und problemhaltige Aufgaben sind und wie sie sich von anderen Aufgabentypen unterscheiden. Bei der Entwicklung des Trainingsbausteines wurde dazu auf die Problemlöseforschung zurückgegriffen, die als zentrales Merkmal eines Problems die Barriere bezeichnet, die sich zwischen dem (unterwünschten) Anfangszustand und dem erwünschten Endzustand befindet. In der gegebenen Situation fehlen dann ad hoc Strategien und Handlungsschemata, die den Übergang vom unerwünschten Anfangszustand zum erwünschten Zielzustand ermöglichen (Dörner 1987, 10). Darin liegt auch der Unterschied zwischen (Routine-)Aufgaben und Problemen: Während zu Ersteren Lösungswege und -mittel bekannt sind und „nur“ abgearbeitet werden müssen, ist dies bei Problemen und problemhaltigen Aufgaben nicht der Fall (siehe hierzu z. B. Dörner 1987; Vollmeyer/Funke 1999, 214). Hier müssen Lösungswege erst generiert werden.

Nicht alle Probleme sind in gleichem Maße komplex, sondern das Maß der Komplexität bestimmt sich durch die Ausprägung verschiedener Dimensionen (Dörner 1987; Jonassen/Hung 2012):

- (1) *Klarheit und Eindeutigkeit* (des Ausgangszustandes und des anzustrebenden Zielzustands, der verfügbaren Informationen und der Offenheit potentieller Lösungen);
- (2) *(Eigen-)dynamik* einer Situation (Veränderbarkeit auch ohne eigenes Zutun des Problemlösenden);
- (3) *Vernetztheit* (der Variablen, verbunden mit erwünschten sowie unerwünschten Neben- und Folgewirkungen einer Problemlösung);
- (4) *Interessenkonflikte* (Notwendigkeit der Berücksichtigung von möglicherweise wenig kompatiblen Interessen verschiedener Stakeholder);
- (5) *Handlungskontrolle und Reflexion*, d.h. die Notwendigkeit zur Begründung von Entscheidungen in Problemsituationen.

Je nach Ausprägung der einzelnen Merkmale variiert das Maß der Komplexität und Problemhaltigkeit. Je mehr Merkmale vollständig erfüllt sind, desto problemhaltiger ist eine Aufgabe.

Im Verlauf des Problemlöseprozesses müssen dann bestimmte Schritte durchlaufen werden, in denen verschiedene Kompetenzfacetten zum Tragen kommen (Seifried et al. 2016; Bransford/Stein 1993; Dörner 1987; Funke 2006, 2011; Sembill 1992; Wuttke et al. 2015; Wuttke/Wolf 2007; siehe hierzu auch Wuttke et al. 2022):

- (1) Zu Beginn des Problemlöseprozesses hat die Analyse und Bewertung der Ausgangslage sowie die Formulierung von Zielen zu erfolgen. Hier sind auch die für die Problemlösung notwendigen Informationen zu suchen und zu verarbeiten.

- (2) Danach ist eine Lösung zu entwickeln, die dann durchgeführt und kontrolliert (Handlungskontrolle) werden muss.
- (3) Abschließend ist zu reflektieren, ob mit der entwickelten Lösung das Problem tatsächlich gelöst wurde.

Diese Grundlagen bilden den Ausgangspunkt des ersten Trainingsbausteins und auf dieser Grundlage werden Merkmale von (komplexen) Problemen (Bransford/Stein 1993; Dörner 1987; Funke 2006, 2011; Sembill 1992; Wuttke et al. 2015; Wuttke/Wolf, 2007; siehe hierzu auch Wuttke et al. 2022) näher ausgeführt und ein Raster zu den Merkmalen problemhaltiger Aufgaben etabliert. Das Raster kann im weiteren Verlauf des Trainings immer wieder bei der Reflexion und Evaluation vorhandener Aufgaben und der Entwicklung neuer, problemhaltiger Aufgaben herangezogen werden, um deren Problemhaltigkeit zu bestimmen bzw. diese herzustellen.

Die Erstellung entsprechender Aufgaben durch Teilnehmende erfolgt in einer Selbstlernphase. Zu den Aufgaben wird im Anschluss ein Feedback bezogen auf den Problemgehalt und andere Qualitätsmerkmale (z. B. Situierung, Medieneinbindung) gegeben, wenn nötig finden Anpassungen statt.

Trainingsbaustein 2: Mediendidaktische Grundlagen der technologiebasierten Umsetzung

Bei Prüfungsaufgaben wird häufig deren fehlende Authentizität kritisiert (Deutscher/Winther 2019; Ressel/Kiesecker 2019; Seifried et al. 2015, 4; Wuttke et al. 2022). Authentische Aufgaben orientieren sich an der beruflichen Realität und bilden berufstypische Anforderungssituationen ab. Technologiebasierte Prüfungen bieten vielfältigere Gestaltungsmöglichkeiten für eine authentische Situierung und Umsetzung von Aufgaben als papierbasierte Formate. Beispielsweise können durch Videovignetten oder den Einsatz virtueller Agenten reale Situationen authentisch simuliert oder es kann mit berufstypischer Software gearbeitet werden. Allerdings kommen die Vorteile technologiebasierter Aufgaben nur dann zum Tragen, wenn mediale Gestaltungsprinzipien berücksichtigt werden. Zu den wichtigsten Gestaltungsprinzipien gehören laut Clark und Mayer (2016) das (1) Modalitätsprinzip (Low/Sweller 2005; Moreno/Mayer 1999), das (2) räumliche Kontiguitätsprinzip (Ayres/Sweller 2005; Moreno/Mayer 1999) und das (3) Redundanzprinzip (Mayer/Johnson 2008).

- (1) Das Modalitätsprinzip setzt auf das Ansprechen verschiedener Sinneskanäle. Insbesondere wenn es um die Erläuterung von Abbildungen geht, wird ein ergänzender Einsatz von Sprechtext zur schriftlichen Erläuterung empfohlen. Hierdurch können verschiedene Sinneskanäle für die Informationsaufnahme parallel genutzt werden. Die visuelle Aufmerksamkeit wechselt dadurch nicht zwischen der Abbildung und dem Text (Low/Sweller 2005). Dies soll eine bessere Nutzung der kognitiven Ressourcen ermöglichen.
- (2) Das räumliche Kontiguitätsprinzip besagt, dass Beschriftungen möglichst nah an den korrespondierenden Abbildungselementen platziert werden sollen, um aufwendige visuelle Suchprozesse zu reduzieren (Ayres/Sweller 2005; Moreno/Mayer 1999). So ist insbesondere bei einzelnen und kleinen Bildschirmen der darstellbare Inhalt zumeist

stark begrenzt. Während bei papierbasierten Aufgaben, die sich z. B. auch über mehrere Seiten erstrecken können, die Seiten nebeneinandergelegt werden können, müssen Lernende und Prüflinge bei dem gewählten Bildschirmausschnitt eine Eingrenzung der darzustellenden Inhalte vornehmen.

- (3) Das Redundanzprinzip beschreibt die zeitgleiche Darbietung von identischem auditivem und visuellem Text, die vermieden werden soll (Mayer/Johnson 2008). Dieses Prinzip adressiert insbesondere einen möglichen Einsatz von Audio- und Videovignetten. So muss bei der Gestaltung der eingesetzten Medienelemente durch die Aufgabenerstellenden bereits im Vorfeld der intendierte Informationsgehalt einzelner Elemente beachtet und diese aufeinander abgestimmt werden.

Die konsequente Umsetzung dieser Gestaltungsprinzipien verringert die Gefahr, dass die begrenzte Kapazität des Wahrnehmungs- und Informationsverarbeitungssystems der Lernenden und Prüflinge überlastet wird (Clark/Mayer 2016). Der Trainingsbaustein thematisiert diese Gestaltungsgrundlagen und ergänzt sie um:

- (4) allgemeine Rahmenbedingungen digitaler Aufgaben (z. B. Erzeugen eines situativen Rahmens für das Aufgabenszenario bei der Erstellung digitaler Aufgaben, Nutzung von Text-, Bild und Multimediaelementen)
- (5) rechtliche Aspekte der Nutzung von Medienelementen, Einblick in Arten von Nutzungslizenzen;
- (6) Fragengestaltung technologiebasierter Aufgaben, d. h. Auswahl geeigneter Fragetypen zum Messen der Problemlösekompetenz sowie die Formulierung der digital zu präsentierenden Aufgabenstellung bzw. der möglichen Antwortalternativen (geschlossene, halboffene und offene Fragetypen)

Trainingsbaustein 3: Einführung in die digitale Lehr-Lern-Plattform

Die technologiebasierte Umsetzung von Aufgaben erlaubt eine breite Auswahl unterschiedlicher Fragetypen. Damit kann sichergestellt werden, dass berufliche Tätigkeiten besser simuliert werden, als bei den häufig in Prüfungssituationen genutzten geschlossenen Fragetypen (z. B. Single- oder Multiple-Choice Fragen). Elektronische Lern- und Prüfungssysteme erlauben zusätzlich den Einsatz von Intermediate Constraint-Fragen (Scalise/Gifford 2006). Diese ordnen sich zwischen den geschlossenen und offenen Fragetypen ein und ermöglichen die Interaktion zwischen den Lernenden bzw. Prüflingen und dem System. Beispiele hierfür sind Anordnungs- oder Zuordnungsfragen, bei denen die Position von Elementen (z. B. Eigenschaften von Bewerber*innen in der Personalauswahl) verändert werden muss. Weiterhin können auch Kompetenzen in der Nutzung von Anwendungssoftware realistisch an digitalen Endgeräten überprüft werden.

Ein weiterer Trainingsbaustein ist deshalb die Einführung in und die Arbeit mit den Lehr-Lern-Plattformen ILIAS oder Moodle, um die praktische Umsetzung der mediendidaktischen Gestaltung zu ermöglichen. Beide Plattformen sind frei nutzbar (open-source) und können über ein

beliebiges internetfähiges Endgerät aufgerufen werden. Den Trainingsteilnehmenden werden umfangreiche Auswahlmöglichkeiten hinsichtlich der einsetzbaren Fragetypen und medialen Gestaltungselemente bereitgestellt. Somit kann für Aufgaben ein individueller, situativer Rahmen geschaffen werden, der die Authentizität fördert. Das Ziel der technischen Schulung ist es, die Aufgabenerstellenden zum Implementieren technologiebasierter Aufgaben zu befähigen.

Im Training werden die Anforderungen aus den drei Trainingsbausteinen, nämlich die problemhaltige und die technologiebasierte Umsetzung von Aufgaben dadurch verbunden, dass die problemhaltigen Aufgaben direkt auf einer digitalen Lehr-Lern-Plattform unter Berücksichtigung mediendidaktischer Aspekte erstellt und dort dann auch von Auszubildenden bearbeitet werden können.

Bei der mediendidaktischen Gestaltung und technologiebasierten Umsetzung geht es somit um die authentischere Situierung des zu bearbeitenden Problems und um das präzisere Messen von in der Praxis erforderlichen Kompetenzen als dies in papierbasierten Prüfungen der Fall ist (Brink/Lautenbach 2011; Joint Information Systems Committee 2007). Ein weiterer Vorteil ist die praktische Erleichterung der Durchführung von Tests. Einerseits können bestehende Aufgaben und Medien in Datenbanken (z. B. Aufgaben- und Medienpools) gespeichert und einfach wiederverwendet bzw. modifiziert werden. Andererseits ermöglicht die teilweise automatisierte Auswertung die Reduktion des Korrekturaufwands sowie – und das ist sicher das wichtigere Argument – die Erhöhung der Objektivität und Reliabilität der Prüfungen. Die Ergebnisse sind je nach Gegebenheit vergleichbar mit der Bewertung durch menschliche Korrektor*innen (Larkey 1998). Durch technologiebasiertes Prüfen kann damit insgesamt die Validität, Reliabilität und Objektivität in der Kompetenzmessung gesteigert (Kirschner et al. 2017; Palm 2008) sowie die Testeffizienz verbessert werden (Hewlett/Kahl-Andresen 2014).

Trainingsbaustein 4: Diagnostische Gütekriterien problemhaltiger Aufgaben

Prüfungen und Zertifikate haben insbesondere eine selektive Funktion (Brunner/Stanat/Pant 2014, 485f.; Severing 2011, 18). Sie dienen als Entscheidungsgrundlage im Rahmen der Auswahl von Personen für eine bestimmte Arbeits- und Lernumwelt, etwa bei der Auswahl von Ausbildungsabsolvent*innen für eine im Unternehmen zu besetzende Stelle oder im Rahmen der Auswahl einer geeigneten Lernumgebung für eine bestimmte Person, etwa bei der Auswahl einer konkreten Berufsausbildung aus einem breiten Spektrum möglicher Ausbildungsberufe (Brunner et al. 2014, 485). Abschlussprüfungen entscheiden zudem über die weitere berufliche Laufbahn und die auf den Zertifikaten ausgewiesenen Ergebnisse können weitreichende Folgen für den (weiteren) beruflichen Werdegang von Ausbildungsabsolvent*innen haben. Es ist folglich unerlässlich, dass Abschlussprüfungen sichere Rückschlüsse auf die interessierenden Merkmale (hier: Kompetenzen) einer Person erlauben. Die von den Trainingsteilnehmenden zu entwickelnden Aufgaben müssen daher, ebenso wie wissenschaftliche Tests im Allgemeinen, als Instrumente pädagogischer Diagnostik den Ansprüchen der wissenschaftlichen Gütekriterien genügen (Schrader 2013, 155). Daher werden diese ebenfalls im Training in den Blick genommen. Im Vordergrund stehen die (Haupt-)Gütekriterien der Objektivität, Reliabilität und Validität (vgl. hierzu Moosbrugger/Kelava 2020, 17), wobei im Training der Fokus vor allem auf Fragen der Auswertungsobjektivität und der Inhaltsvalidität liegt.

3 Methodische Umsetzung: Evaluationsdesign und Durchführung des Trainings

3.1 Evaluationsdesign

Der Erfolg des Trainings wurde mittels eines Pre-/Posttestdesigns überprüft. Die Evaluation erfolgte in Anlehnung an die Evaluationsstufen von Kirkpatrick und Kirkpatrick (2006, 21ff.), nämlich „Reaktion“, „Lernen“, „Verhalten“ und „Resultate“. Nachfolgend werden die Befunde der ersten drei Ebenen als Erfolgskriterium herangezogen und berichtet. Auf der Verhaltensebene, der Konzeption von problemorientierten Aufgaben, wurden die Aufgaben auf ihre Problemhaltigkeit geprüft und zu einem späteren Zeitpunkt (nach Abschluss aller Trainings) bei einer Auszubildendenstichprobe eingesetzt, um Informationen über deren diagnostischen Qualität zu erhalten. Die Evaluation auf der vierten Ebene entspräche einer Veränderung der diagnostischen Praxis, insbesondere der Prüfungspraxis im Rahmen der Abschlussprüfungen. Eine solche Weiterentwicklung der Prüfungspraxis stellt einen längerfristigen Prozess dar und erfordert zudem entsprechende Unterstützungsstrukturen, teils Veränderungen rechtlicher Rahmenbedingungen zur Prüfungsdurchführung.

3.2 Durchführung, Stichprobe und Instrumente

Die Trainings fanden im Zeitraum von November 2020 bis Oktober 2022 in mehreren Durchläufen mit insgesamt 152 Teilnehmer*innen statt. Pandemiebedingt wurde ein Großteil der Trainings virtuell durchgeführt, einige wenige konnten in Präsenz angeboten werden. Die Stichprobe ist folgendermaßen zusammengesetzt (siehe Tab. 1):

Tabelle 1: Stichprobe der Trainingsteilnehmer*innen

	N	Alter M (SD)	Geschlecht	Berufserfahrung M (SD)	Fachsemester/ Ausbildungsjahr	Berufsausbildung (BA) absolviert
Lehrende, Ausbildende, Prüfer*innen	63	48,2 (8,6)	m = 20 w = 43	16,5 (10,2)	---	50 (davon 40 eine kaufmännische BA)
Angehende Lehrkräfte (LiV an kaufm. Schulen und Masterstudierende WiPäd)	89	27,4 (3,9)	m = 39 W = 50	---	LiV: 1. AJ: 4 2. AJ: 1 Studierende: 1./2. Semester: 56 3. Semester: 21 4. Semester: 4	47 (davon 44 eine kaufmännische BA)

Die beiden ersten Ebenen des Trainingserfolgs (Reaktion und Lernen) wurden in einem Pre-/Posttestdesign evaluiert. Zu beiden Messzeitpunkten (vor Beginn und nach Abschluss des Trainings) wurden folgende Instrumente eingesetzt:

- (Erfüllte) Teilnehmer*innenerwartungen (in Anlehnung an Krille (2017) entwickelt auf Basis des VIE-Ansatzes¹ nach Vroom (1964), siehe dazu auch Noe und Schmitt (1986), Tharenou (2001) und Zaniboni et al. (2011)). Alle fünf Skalen (mit insgesamt 41 Items) weisen eine sehr gute Reliabilität auf (Pretest/Posttest: Erstellung von Aufgaben, $\alpha=0.903/0.933$; Unterstützung von Lehren und Lernen durch Digitalisierung, $\alpha=0.886/0.901$; E-Learning und Mediendidaktik, $\alpha=0.895/0.906$; Gütekriterien, $\alpha=0.925/0.894$; Erfahrungsaustausch und Netzwerken, $\alpha=0.908/0.896$)
- Im Posttest wurde zudem erfasst, wie die Teilnehmer*innen den Nutzen des Trainings einschätzen (in Anlehnung an Grohmann/Kauffeld (2013) nach Greiwe (2020, 124ff.)). Hierzu wurden die Skalen „wahrgenommene Qualität der Instruktion“ ($\alpha=0.815$), „wahrgenommene Nützlichkeit der digitalen Plattform“ ($\alpha=0.887$), „wahrgenommene Benutzerfreundlichkeit der digitalen Plattform“ ($\alpha=0.907$), „Nutzungsintention der digitalen Plattform“ (0.902), „Zufriedenheit mit dem Training“ ($\alpha=0.867$), Nutzen des Trainings ($\alpha=0.836$) sowie „eingeschätzter Kompetenzerwerb“ ($\alpha=0.863$) eingesetzt.
- Mit Blick auf den Wissenszuwachs wurden ein Wissenstest zu Merkmalen problemhaltiger Aufgaben sowie ein Wissenstest zu diagnostischen Gütekriterien (beide Eigenentwicklung) und ein Instrument zur Selbsteinschätzung der Medienkompetenz (Niegemann/Korbach, in Vorbereitung) eingesetzt. Der Wissenstest zu Merkmalen problemhaltiger Aufgaben umfasst fünf Items ($\alpha=0.520/0.706$). Der Test zu den diagnostischen Gütekriterien umfasst acht Items, die Reliabilität ist zufriedenstellend ($\alpha=0.717/0.699$). Die Selbsteinschätzung der Medienkompetenz umfasst die drei Skalen Mediennutzung, Medienkunde (Datenschutz und -sicherheit) und Medienkritik. Die Reliabilitäten im Pre- und Posttest sind zufriedenstellend bis akzeptabel (Mediennutzung $\alpha=0.611/0.642$; Medienkunde $\alpha=0.727/0.656$; Medienkritik $\alpha=0.775/0.797$).

Für die dritte Ebene (Verhalten) wurden die von den Teilnehmenden entwickelten Aufgaben für die Ausbildungsberufe IK und KBM als Erfolgskriterium herangezogen.

4 Ergebnisse

Ein Blick auf die Befunde zeigt, dass die Erwartungen der Teilnehmer*innen weitgehend erfüllt wurden, sie dem Training einen Nutzen zusprechen und das Training auch durchaus als erfolgreich bewerten.

Ebene 1: Reaktion auf das Training

Die Teilnehmenden zeigen sich insgesamt zufrieden bis sehr zufrieden mit dem Training² ($M=3.00$; $SD=0.65$) und sehen darin einen mittelmäßigen bis starken Nutzen ($M=3.00$; $SD=0.77$). Besonders gute Rückmeldung erhält die Qualität der Instruktion ($M=3.40$;

¹ Der Valenz-Instrumentalitäts-Erwartungs-Ansatz besagt, dass Individuen Leistungsbereitschaft in einer spezifischen Situation zeigen, wenn ein das Erreichen eines gesetzten Ziels in Aussicht steht (Vroom 1964).

² Teilnehmende bewerteten ihre Zufriedenheit, den Trainingsnutzen, die Instruktionsqualität sowie den eingeschätzten Lernerfolg auf einer vierstufigen Likert-Skala (1 = „trifft gar nicht zu“ bis 4 = „trifft voll zu“).

SD=0.51). Mit einem Median von $M=3.20$ ($SD=0.61$) wird der Lernerfolg ebenfalls als eher hoch eingeschätzt.

Ebene 2: Wissenszuwachs

Die Ergebnisse der Wissenstests³ bestätigen grundsätzlich die subjektive Einschätzung zum Lernerfolg. Teilnehmende weisen im Vergleich des Pre- und Posttests einen deutlichen Lernzuwachs⁴ zu den Merkmalen problemhaltiger Aufgaben auf ($M_{Pre|AM}=3.26$; $SD_{Pre|AM}=1.27$; $M_{Post|AM}=3.77$; $SD_{Post|AM}=1.45$; $t_{Post-Pre|AM}(119)=-3.47$; $p_{AM}<0.05$). Wir gehen davon aus, dass die intensiven Anwendungsphasen und die tutorielle Begleitung während des Trainings einen wesentlichen Beitrag hierzu geleistet haben. Mit Blick auf das Wissen über diagnostische Gütekriterien ließ sich kein signifikanter Zuwachs feststellen ($M_{Pre|GK}=4.64$; $SD_{Pre|GK}=2.29$; $M_{Post|GK}=4.69$; $SD_{Post|GK}=2.18$; $t_{Post-Pre|GK}(142)=-0.29$; $p_{GK}>0.05$). Dies kann darauf zurückgeführt werden, dass das Thema insgesamt einen relativ knappen Anteil im Training einnahm und weniger Lerngelegenheiten und Lernzeit zur Anwendung zur Verfügung standen.

Die Ergebnisse des Fragebogens zur Medienkompetenz unterstützen die Notwendigkeit eines Bausteins zur Mediendidaktik. Bei einer detaillierteren Betrachtung nach Zielgruppen fällt auf, dass ein allgemeiner Professionalisierungs- bzw. Weiterbildungsbedarf besteht. So weisen insbesondere angehende und erfahrene Lehrkräften Schwächen in der Kompetenzfacette der Medienkritik auf, während für das betriebliche Bildungspersonal eine stärkere Sensibilität nachgewiesen werden konnte. Hierbei kann von einem Einfluss der betrieblichen Arbeitsumgebung ausgegangen werden. Gegenüber der digitalen Mediennutzung zeigen hingegen die erfahrenen Lehrenden und das betriebliches Bildungspersonal eine höhere Einschätzung der eigenen Kompetenz als die angehenden Lehrenden. Ein erster Erklärungsansatz könnte hierbei die aktive Nutzung digitaler Medien während der Coronapandemie sein. Eine detaillierte Ergebnisdarstellung ist in Meiners et al. (2022) zu finden. Der Pre-Posttest-Vergleich zur Medienkompetenz zeigt, dass Teilnehmende mit Blick auf die Medienkunde⁵ nach dem Training einen signifikant besseren Wert erzielen ($z=5.148$; $p<0.05$). Die im Training genutzte digitale Lehr-Lern-Plattform⁶ wird als nützlich bis sehr nützlich ($M=3.00$; $SD=0.591$) und als mittelmäßig bis stark benutzerfreundlich eingeschätzt ($M=3.00$; $SD=0.565$). Teilnehmende zeigen eine Tendenz dazu, die digitale Plattform nutzen zu wollen ($M=3.00$; $SD=0.740$). Die Werte unterstreichen die Relevanz der im Training eingesetzten Bausteine (2) zur Mediendidaktik und (3) zur Technikschiulung.

Ebene 3: Verhalten

³ Im Wissenstest zu den Merkmalen problemhaltiger Aufgaben konnten Teilnehmende bis zu 5 Punkte erreichen, im Wissenstest zu diagnostischen Gütekriterien bis zu 8 Punkte.

⁴ Da der Pre- und Posttest aus mehreren Fragebogen- und Testteilen bestand (vgl. Tabelle 3), kam es vor, dass Teilnehmende manche Teile übersprungen haben. Dies führte dazu, dass die Anzahl der Teilnehmenden in den einzelnen Fragebogen-/Testteilen variiert. So haben an den Wissenstests nicht alle Teilnehmenden teilgenommen, weshalb hierzu die N niedriger sind als zur ersten Ebene der Evaluation.

⁵ Zur Medienkunde wurde eine Likert-Skala von 1 = „stimme gar nicht zu“ bis 6 = „stimme voll zu“ eingesetzt.

⁶ Teilnehmende bewerteten die wahrgenommene Nützlichkeit und Benutzerfreundlichkeit der digitalen Plattform sowie ihre Nutzungsintention der digitalen Plattform auf einer vierstufigen Likert-Skala (1 = „trifft gar nicht zu“ bis 4 = „trifft voll zu“).

Die Ergebnisse der Erprobungsstudien bei kaufmännischen Auszubildenden bestätigen den Trainingserfolg. So zeigen durchgeführte Think-Aloud-Studien bei 26 kaufmännischen Auszubildenden, dass zum Lösen der aus den Trainings gewonnenen Aufgaben zumeist die Schritte des Problemlösens durchlaufen werden müssen. Eine Auswertung erfolgte in Form einer qualitativen Inhaltsanalyse der Think-Aloud-Protokolle, in der anhand der Problemlöseschritte analysiert wurde. Kritik an den Aufgaben selbst wurde kaum formuliert. Auch die Ergebnisse der quantitativen Erprobungsstudie mit knapp 1.300 Auszubildenden sind zufriedenstellend und zeigen, dass die Aufgaben sich eignen, um Problemlösekompetenz zu erfassen.

Ebenso zeigen die Ergebnisse der quantitativen Studien zur Funktionalität und diagnostischen Qualität der entwickelten problemorientierten Aufgaben gute Ergebnisse. Es wurden insgesamt 37 Aufgaben zu kaufmännischen Problemsituationen in den Bereichen Personalwirtschaft, Marketing und Absatz, Finanzierung und Investition sowie Beschaffung und Bevorratung entwickelt. Davon entfielen 26 Aufgaben auf IK und 11 Aufgaben auf KBM, die im Testheftdesign eingesetzt wurden. Die Aufgaben wurden von insgesamt 1.416 Auszubildenden getestet (IK=655, davon gültig N=632; KBM=761, davon gültig N=741). Insgesamt lagen damit Testdaten von 1.373 Teilnehmenden aus den beiden Berufen aus dem 2. und 3. Ausbildungsjahr vor. Aufgrund des Testheftdesigns schwanken die Antworthäufigkeiten zwischen 16 und 166 bei den 27 IK-Aufgaben und 103 bis 326 Antworten für die 11 KBK-Aufgaben. Anzumerken ist, dass ein Testheft aus drei bis fünf Aufgaben bestand. Auszubildenden bearbeiteten jeweils ein Testheft zufällig. Eine IK-Aufgabe musste aufgrund ihres hohen Anforderungsniveaus ausgeschlossen werden, da keine Person dieses Item korrekt löste.

Die analysierten Aufgabenpools mit 25 bzw. 11 Aufgaben für die beiden Ausbildungsberufe sind geeignet, Problemlösekompetenz zu erfassen. Ebenfalls weisen die Aufgaben unterschiedliche Schwierigkeitsgrade auf, sodass eine trennscharfe Prognostik von Kompetenzen der Auszubildenden möglich ist. Die Aufgaben wurden auf der Grundlage eines eindimensionalen Rasch-Modells als standardisiertes IRT-Modell skaliert. Mit Hilfe der marginalen Maximum-Likelihood-Schätzung wurden sowohl Personenfähigkeitsparameter (gewichtete Maximum-Likelihood-Schätzer (WLE)) als auch zugehörige Reliabilitäten (EAP/PV) bestimmt. Die Personenparameter streuen im Bereich zwischen -3,919 bzw. -3,938 und 2,564 bzw. 3,324 in Bezug auf die IK- bzw. KBM-Auszubildenden. Zusätzlich wurden Itemschwierigkeitsparameter inklusive der zugehörigen Fit-Werte bestimmt. Die problemhaltigen Aufgaben weisen in der Schwierigkeit eine gute Streuung auf. Die Fit-Werte der Itemparameterschätzungen sind sehr zufriedenstellend: Die gewichteten Abweichungsquadrate (weighted MNSQ) liegen alle in einem akzeptablen Intervall zwischen 0,8 und 1,2. Die Trennschärfe aller Items liegt bei mindestens 0,2. Zusammenfassend bleibt festzuhalten, dass die Aufgabenpools für beide Gruppen kognitiv unterschiedlich anspruchsvoll gestaltet sind, sodass eine trennscharfe Prognostik von Kompetenzen der Auszubildenden im Bereich des Problemlösens möglich ist.

Sowohl die kognitive Validierung der im Training entwickelten Aufgaben (einschließlich der für das Training genutzten „Musteraufgaben“) als auch die quantitative Erprobung verweisen auf einen guten Trainingserfolg auf der dritten Evaluationsebene. Allerdings – und dies ist als Einschränkung anzumerken – wissen wir nicht, wie die Aufgaben ohne ein solches Training entwickelt worden wären. Hier kann allenfalls geltend gemacht werden, dass die meisten

Aufgaben nach der Diskussion mit den Trainingsteilnehmer*innen anhand von Kriterien zur Klassifikation problemhaltiger Aufgaben nochmals eine Überarbeitung erfahren haben, um dem entsprechenden Anspruch gerecht zu werden.

5 Diskussion und Ausblick

Insgesamt zeigen die Trainingsergebnisse, dass Interventionen zur Professionalisierung des Bildungs- und Prüfungspersonals im Bereich der Konstruktion diagnostisch abgesicherter problemorientierter Aufgaben so anzulegen sind, dass verschiedene Perspektiven miteinander verschränkt werden. So erwies es sich für die Aufgabenkonstruktion als sehr hilfreich, nicht nur die theoriegeleitete Konstruktion problemhaltiger kaufmännischer Aufgaben als Kriterium heranzuziehen, sondern zugleich Merkmale der Authentizität, der Handlungsorientierung und der digitalen Umsetzung mit Hilfe von Medien und deren Implementation in ein Lernmanagementsystem bei der Konzeption zu berücksichtigen. Die Verbindung mit mediendidaktischen und technologischen Inhalten erwies sich als Vorteil, da die Teilnehmenden ohnehin vor der Herausforderung standen, eine digitale Umsetzung von Lernangeboten zu entwickeln. Ob dies jedoch zuvorderst der Zeit während der Corona-Pandemie geschuldet war, in der ein hoher Weiterbildungsbedarf im Umgang mit digitalen Medien zutage trat, oder ob dies auch mit eher allgemeinen Digitalisierungsprozessen in Arbeits- und Lernumwelten zusammenhängt, ist nicht eindeutig zu bestimmen.

Hinsichtlich der Förderung der einzelnen Kompetenzfacetten, die im Training adressiert wurden, offenbarten sich jedoch auch deutlich Unterschiede. Während sich das Wissen über Kriterien zur Strukturierung des Problemgehalts von Aufgaben positiv entwickelte und auch Kompetenzen zur Beurteilung und Gestaltung problemhaltiger Aufgaben (ermittelt über einen entsprechenden Test und über die während des Trainings entwickelten Aufgaben) einen Zuwachs zu verzeichnen hatten, fiel das Ergebnis im Bereich der Diagnostik kritischer aus. Dies könnte zum einen damit zusammenhängen, dass unter den Trainingsteilnehmenden überwiegend (angehendes) berufsschulisches Bildungspersonal war, das (noch) keine Erfahrungen in der Prüfungspraxis hatte und das Thema möglicherweise als weniger relevant erachtete. Zudem handelt es sich um ein eher komplexes Thema, das aufgrund der zeitlichen Limitationen des Trainings und der Schwerpunktsetzung auf die Aufgabenkonstruktion und technologiebasierte Umsetzung sehr stark komprimiert werden musste. Insofern ist das Trainingsergebnis zu diesem Baustein eher darauf zurückzuführen, dass ein zentrales Kriterium erfolgreicher Trainings, Übung und Anwendung (Krille 2019; Cademartori et al. 2017; Lipowsky 2009), nicht hinreichend umgesetzt werden konnte. Hier wird auch das Erfordernis einer Weiterentwicklung des Trainings gesehen, zum einen mit Blick auf das didaktisch-methodische Arrangement, zum anderen aber auch mit Blick auf die Zielgruppen, bei denen sich in den genutzten diagnostischen Verfahren während des Ausbildungsprozesses deutliche Unterschiede zwischen berufsschulischem und betrieblichem Bildungspersonal zeigen.

Ein sicherlich größeres Desiderat der Trainingsstudie ist die langfristige Wirkung des Trainings. Wir wissen nicht, ob das Training tatsächlich dazu führt, dass Lehrende und Auszubildende künftig bei der Konstruktion von Lern- und Diagnostikaufgaben, die sich ja in ihrem

Anforderungsniveau entsprechen sollten (Curriculum-Instruktions-Assessment-Triade, vgl. Pellegrino 2012, Weber/Achtenhagen 2014), verstärkt eine bewusste Modellierung des Problemlöseraums über die Situierung von Anforderungen und die Formulierung von Aufgaben bzw. Arbeitsaufträgen vornehmen. Hierfür wäre eine Langzeit- bzw. weiterführende Studie notwendig, in der die Teilnehmenden des Trainings im Hinblick auf die Entwicklung und Nutzung technologiebasierter problemorientierter Aufgabenstellung für Lern- und Diagnostikzwecke beobachtet und/oder befragt werden. Daher können zum Transfer des Trainings und zur Weiterentwicklung der geförderten Kompetenzen im Handlungsfeld keine Aussagen getroffen werden. Auch können etwaige Wirkungen als Multiplikator*innen nicht abgeschätzt werden.

Literatur

Ayres, P./Sweller, J. (2005): The split-attention principle in multimedia learning. In: Mayer, R.: The Cambridge handbook of multimedia learning. New York, 206-226.

Biggs, J. (1996): Enhancing teaching through constructive alignment. Higher Education, (32), 347-364. Online: <https://link.springer.com/content/pdf/10.1007/BF00138871.pdf?pdf=button> (23.02.2023).

Bransford, J./Stein, B. S. (1993): The ideal problem solver. A guide for improving thinking, learning, and creativity. 2. Aufl. New York. Online: https://www.tntech.edu/cat/pdf/useful_links/idealproblemsolver.pdf (23.02.2023).

Brink, R./Lautenbach, G. (2011): Electronic assessment in higher education. In: Educational Studies, 37(5), 503-512.

Brunner, M./Stanat, P./Pant, H. A. (2014): Diagnostik und Evaluation. Pädagogische Psychologie. Weinheim/Basel.

Bundesinstitut für Berufsbildung (2014): Prüfungen im dualen System. In: BWP, Berufsbildung in Wissenschaft und Praxis, 3/2014. Online: <https://www.bwp-zeitschrift.de/dienst/publikationen/de/7273> (23.02.2023).

Bundesinstitut für Berufsbildung (2023): Liste der am stärksten besetzten Ausbildungsberufe. Online: <https://www.bibb.de/dienst/dazubi/de/1867.php>.

Cademartori, I./Seifried, J./Wuttke, E./Krille, C./Salge, B. (2017): Developing a training programme to promote professional error competence in accounting. In: Wuttke, E./Seifried, J. (Hrsg.): Professional Error Competence of Preservice Teachers: Evaluation and Support, 29-46. https://doi.org/10.1007/978-3-319-52649-2_3.

Clark, R. C./Mayer, R. E. (2016): E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning. New Jersey.

Deutscher, V./Winther, E. (2019): Zentrale Abschlussprüfungen in der Berufsausbildung - ein Plädoyer. Berufsbildung in Wissenschaft und Praxis, 48 (6), 11-15. Online: https://www.uni-du.de/imperia/md/content/biwi/einrichtungen/vet/deutscher_winther_2019_zentrale_abschlusspruefungen.pdf (14.03.2020).

Frey, C.B./Osborne, M.A. (2017): The future of employment: How susceptible are jobs to computerisation? In: Technological Forecasting and Social Change, Volume 114, 2017, 254-280.

Dörner, D. (1987): Problemlösen als Informationsverarbeitung. 3. Aufl. Kohlhammer Standards Psychologie: Studententext: Teilgebiet Denkpsychologie. Stuttgart.

Funke, J. (2006): Komplexes Problemlösen. In: Funke, J. (Hrsg.): Denken und Problemlösen. Bd. 8. Göttingen u. a., 375-446.

Funke, J. (2011): Problemlösen. In: Betsch, T./Funke, J. & Plessner, H. (Hrsg.): Denken – Urteilen, Entscheiden, Problemlösen. Allgemeine Psychologie für Bachelor. Berlin, 137-199.

Fürstenau, B. (1994): Komplexes Problemlösen im betriebswirtschaftlichen Unterricht. Wiesbaden.

Fürstenau, B. (1999): Förderung von Problemlösefähigkeit im planspielgestützten Unterricht. In: Unterrichtswissenschaft, 27(2), 135-158.

Geiser, P./Seeber, S./Busse, J./Schumann, M./Lange, A./Weber, S./Achtenhagen, F./Hiller, F./Zarnow, S./Hackenberg, T. (2021): Kompetenzfacetten in digitalisierten kaufmännischen Arbeitsplatzsituationen – Eine vergleichende Perspektive von Auszubildenden und Lehrkräften. In: Zeitschrift für Berufs- und Wirtschaftspädagogik, 117, 2021/4, 630-657. <https://doi.org/10.25162/zbw-2021-0024>.

Gräsel, C./Fussangel, K./Schellenbach-Zell, J. (2008): Transfer einer Unterrichtsinnovation. Das Beispiel Chemie im Kontext. In: Lankes, E.-M. (Hrsg.): Pädagogische Professionalität als Gegenstand empirischer Forschung. Münster, 207-218.

Greiwe, C. (2020): Kompetenzen im Nachhaltigkeitsmanagement: Eine Interventionsstudie mit angehenden kaufmännischen Lehrkräften. Volume 61. Bielefeld.

Grohmann, A./Kauffeld, S. (2013): Evaluating training programs: development and correlates of the Questionnaire for Professional Training Evaluation. In: International Journal of Training and Development, 17 (2), 135-155.

Hewlett, C./Kahl-Andresen, A. (2014): Prüfungsökonomie statt Prüfungsqualität. In: Berufsbildung in Wissenschaft und Praxis, 3, 7-9.

Jonassen, D. H./Hung, W. (2012): Problem Solving. In: Seel, N. M. (Hrsg.): Encyclopedia of the sciences of learning. New York, 2680-2683.

Joint Information Systems Committee (2007): Effective Practice with e-Assessment: An overview of technologies, policies and practice in further and higher education. Online: <https://www.webarchive.org.uk/wayback/archive/20140613220103/http://www.jisc.ac.uk/media/documents/themes/elearning/effpraceassess.pdf> (23.02.2023).

Kirkpatrick, D. L./ Kirkpatrick, J. D. (2006): Evaluating training programs. The four levels. 3. Aufl. San Francisco.

Kirschner, P. A./Park, B./Malone, S./Jarodzka, H. (2017): Toward a cognitive theory of multimedia assessment (CTMMA). In: Spector, M./Lockee, B./Childress, M. (Hrsg.): Learning,

design, and technology: An international compendium of theory, research, practice, and policy, 1-23. https://doi.org/10.1007/978-3-319-17727-4_53-1.

Krille, C. (2017): Prospective Teachers' Training Motivation and Its Influence on Training Success. In: Wuttke, E./Seifried, J. (Hrsg.): Professional Error Competence of Preservice Teachers. Evaluation and Support. Cham, 99-114.

Krille, C. (2019): Fortbildungsmotivation von (wirtschaftspädagogischen) Lehrkräften. Inaugural-Dissertation, Johann Wolfgang Goethe-Universität Frankfurt a.M.

Kultusministerkonferenz (2002): Rahmenlehrplan für den Ausbildungsberuf Industriekaufmann/Industriekauffrau (Beschluss der Kultusministerkonferenz vom 14.06.2002). Online: <https://www.kmk.org/fileadmin/pdf/Bildung/BeruflicheBildung/rlp/industriekfm.pdf> (23.02.2023).

Larkey, L. S. (1998): Automatic essay grading using text categorization techniques. In: Croft, W. B./Moffat, A. (Hrsg.): Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval. New York, 90-95.

Lipowsky, F. (2009): Unterrichtsentwicklung durch Fort- und Weiterbildungsmaßnahmen für Lehrpersonen. In: Beiträge zur Lehrerinnen- und Lehrerbildung, 27 (3), 346-360. Online: [heBildung/rlp/industriekfm.pdf](https://www.kmk.org/fileadmin/pdf/Bildung/BeruflicheBildung/rlp/industriekfm.pdf) (23.02.2023).

Lipowsky, F. (2010): Lernen im Beruf. Empirische Befunde zur Wirksamkeit von Lehrerfortbildung. In: Müller, F. H./Eichenberger, A./Lüders, M./Mayr, J. (Hrsg.): Lehrerinnen und Lehrer lernen. Konzepte und Befunde zur Lehrerfortbildung. Münster, 51-70.

Lipowsky, F. (2011): Theoretische Perspektiven und empirische Befunde zur Wirksamkeit von Lehrerfort- und -weiterbildung. In: Terhart, E./Bennewitz, H./Rothland, M. (Hrsg.): Handbuch der Forschung zum Lehrerberuf. Münster, 398-417.

Lipowsky, F./Rzejak, D. (2021): Fortbildungen für Lehrpersonen wirksam gestalten. Ein praxisorientierter und forschungsgestützter Leitfaden. Gütersloh.

Lorig, B./Bretschneider, M./Gutschow, K./Mpangara, M./Weber-Höller, R. (2014): Kompetenzbasierte Prüfungen im dualen System – Bestandsaufnahme und Gestaltungsperspektiven. Bonn. Online: https://www.bibb.de/dienst/dapro/daprodocs/pdf/eb_42333.pdf (23.02.2023).

Low, R./Sweller, J. (2005): The modality principle in multimedia learning. In: Mayer, R. (Hrsg.): The Cambridge handbook of multimedia learning. New York, 227-246.

Mayer, R. E./Johnson, C. I. (2008): Revising the redundancy principle in multimedia learning. In: Journal of Educational Psychology, 100(2), 380.

Meiners, H./Hartmann, P./Niegemann, H./Seeber, S./Wuttke, E./Schumann, M. (2022): Digitale Medienkompetenz als Voraussetzung für die Erstellung von Prüfungsaufgaben. In: Schumann, S./Seeber, S./Abele, S. (Hrsg.): Digitale Transformation in der Berufsbildung. Konzepte, Befunde und Herausforderungen, Wirtschaft – Beruf – Ethik, Bd. 41. Bielefeld, 123-144.

Moosbrugger, H./Kelava, A. (2020): Qualitätsanforderungen an Tests und Fragebogen ("Gütekriterien"). In: Moosbrugger H./Kelava, A. (Hrsg.): Testtheorie und Fragebogenkonstruktion. 3. Aufl. Berlin, 13-38.

Moreno, R./Mayer, R. E. (1999): Cognitive principles of multimedia learning: The role of modality and contiguity. In: Journal of educational psychology, 91(2), 358-368. Online: <https://psycnet.apa.org/doi/10.1037/0022-0663.91.2.358> (30.05.2023).

Noe, R. A./Schmitt, N. (1986): The influence of trainee attitudes on training effectiveness: Test of a model. In: Personnel Psychology, 39 (3), 497-523. Online: <https://psycnet.apa.org/doi/10.1111/j.1744-6570.1986.tb00950.x> (23.02.2023).

Palm, T. (2008): Performance assessment and authentic assessment: A conceptual analysis of the literature. In: Practical Assessment, Research, and Evaluation, 13(1), 4.

Pellegrino, J. W. (2010): The Design of an Assessment System for the Race to the Top: A Learning Sciences Perspective on Issues of Growth and Measurement. Educational Testing Service. Online: <https://pdfs.semanticscholar.org/53b7/6668fec653df7db1261304bd43a4ce64e42d.pdf> (23.02.2023)

Ressel, T./Kiesecker, D. (2019): Die duale Kompetenzprüfung. Konzept zur Weiterentwicklung der Abschlussprüfung zu einem Kompetenznachweis für die Lernorte Schule und Betrieb. IG Metall. Online: https://wap.igmetall.de/docs_Duale_Kompetenzpruefung_Langfassung_c0a0d2a45a0e13c11d72ca159d7ba82d771dcff6.pdf (23.02.2023).

Schlicht, J./Hommel, M./Fürstenau, B. (2022): Digital gestütztes, problembasiertes forschendes Lernen von (künftigen) Lehr- und Führungskräften. In: Bildung und Beruf, Juli und August 2022, 259-264.

Schrader, F. W. (2013): Diagnostische Kompetenz von Lehrpersonen. In: Beiträge zur Lehrerbildung, 31(2), 154-165.

Schumann, S./Seeber, S./Abele, S. (2022): Digitalisierung und digitale Medien in der Berufsbildung: Konzepte und empirische Befunde. Bielefeld.

Seifried, J./Rausch, A./Kögler, K./Brandt, S./Eigenmann, R./Schley, T. (2016): Problemlösekompetenz angehender Industriekaufleute – Konzeption des Messinstruments und ausgewählte empirische Befunde (DomPL-IK). In: Beck, K./Landenberger, M./Oser, F. (Hrsg.): Technologiebasierte Kompetenzmessung in der beruflichen Bildung. Ergebnisse aus der BMBF-Förderinitiative ASCOT, Wirtschaft – Beruf – Ethik, Bd. 32, 1. Aufl. Bielefeld, 119-138.

Seifried, J./Wuttke, E. (2017): Teacher Training as a Contribution to Teachers' professional Development: Conclusions from a Research Programme to Foster Professional Error Competence in Accounting. In: Wuttke, E./Seifried, J. (Hrsg.): Professional error competence of pre-service teachers. Evaluation and support (SpringerBriefs in education). Cham, 115-120.

Seifried, J./Wuttke, E./Türling, J./Krille, C./Paul, O. (2015): Teachers' strategies for handling student errors – the contribution of teacher training programs. In: Gartmeier, M./Gruber, H./Ha-

scher, T./Heid, H. (Hrsg.): Fehler. Ihre Funktionen im Kontext individueller und gesellschaftlicher Entwicklung. Errors; their functions in context of individual and societal development. Münster, 177-188.

Severing, E. (2011): Prüfungen und Zertifikate in der beruflichen Bildung: eine Einführung. In: Severing, E./Weiß, R. (Hrsg.): Prüfungen und Zertifizierungen in der beruflichen Bildung. Anforderungen – Instrumente – Forschungsbedarf. Bielefeld, 15-36.

Tharenou, P. (2001): The relationship of training motivation to participation in training and development. In: Journal of Occupational and Organizational Psychology, 74 (5), 599-621.

Vroom, V. H. (1964): Work and motivation. Oxford.

Weber, S./Achtenhagen, F. (2014): Einige unmaßgebliche Gedanken zur aktuellen Diskussion um eine gelingende Curriculumentwicklung. In: bwp@ Berufs- und Wirtschaftspädagogik – online, Profil 3: Digitale Festschrift für Tade Tramm zum 60. Geburtstag. Online: http://www.bwpat.de/profil3/weber_achtenhagen_profil3.pdf (30.05.2023).

Wuttke, E./Seeber, S./Geiser, C./Turhan, L. (2022): Zur Problemhaltigkeit von Aufgaben in kaufmännischen Abschluss- und Zwischenprüfungen – Ergebnisse aus Aufgabenanalysen. In: Zeitschrift für Berufs- und Wirtschaftspädagogik, 118 (1), 25-52.

Wuttke, E./Seifried, J./Brandt, S./Rausch, A./Sembill, D./Martens, T. (2015): Modellierung und Messung domänenspezifischer Problemlösekompetenz bei angehenden Industriekaufleuten. Entwicklung eines Testinstruments und erste Befunde zu kognitiven Kompetenzfacetten. In: Zeitschrift für Berufs- und Wirtschaftspädagogik, 111(2), 189-207.

Wuttke, E./Wolf, K. D. (2007): Entwicklung eines Instrumentes zur Erfassung von Problemlösefähigkeit – Ergebnisse einer Pilotstudie. In: Europäische Zeitschrift für Berufsbildung, 2(41), 99-118.

Zaniboni, S./Fraccaroli, F./Truxillo, D. M./Bertolino, M./Bauer, T. N. (2011): Training valence, instrumentality, and expectancy scale (T-VIES-it): Factor structure and nomological network in an Italian sample. In: Journal of Workplace Learning, 23 (2), 133-151.

Zitieren dieses Beitrags

Wuttke, E.; Seeber, S., Meiners, H. & Turhan, L. (2023): Entwicklung und Evaluation eines Trainingskonzeptes für die Gestaltung problemhaltiger technologiebasierter Lern- und Prüfungsaufgaben. In: *bwp@ Profil 8: Netzwerke – Strukturen von Wissen, Akteuren und Prozessen in der beruflichen Bildung. Digitale Festschrift für Bärbel Fürstenau zum 60. Geburtstag*, hrsg. v. Hommel, M./Aprea, C./Heinrichs, K., 1-20. Online: https://www.bwpat.de/profil8_fuerstenau/wuttke_etal_profil8.pdf (14.09.2023).

Die Autorinnen



Prof. Dr. EVELINE WUTTKE

Goethe Universität Frankfurt, Professur für Wirtschaftspädagogik, insbes. empirische Lehr-Lern-Forschung

Theodor-W.-Adorno-Platz 4, 60629 Frankfurt

wuttke@em.uni-frankfurt.de

<https://www.wiwi.uni-frankfurt.de/abteilungen/wipaed/professoren/wuttke/team.html>



Prof. Dr. SUSAN SEEBER

George-August-Universität Göttingen, Professur für Wirtschaftspädagogik und Personalentwicklung

Platz der Göttinger Sieben 5, 37073 Göttingen

susan.seeber@wiwi.uni-goettingen.de

<https://www.uni-goettingen.de/de/31829.html>



HANNA MEINERS

George-August-Universität Göttingen, Professur für Wirtschaftspädagogik und Personalentwicklung

Platz der Göttinger Sieben 5, 37073 Göttingen

hanna.meiners@uni-goettingen.de

<https://www.uni-goettingen.de/de/hanna+meiners+m.ed./640518.html>



LÜTFIYE TURHAN

ehem. Goethe Universität Frankfurt, Professur für Wirtschaftspädagogik, insbes. empirische Lehr-Lern-Forschung

Theodor-W.-Adorno-Platz 4, 60629 Frankfurt

luetfiye.turhan@gmail.com

<https://www.wiwi.uni-frankfurt.de/abteilungen/wipaed/professoren/wuttke/team.html>