

Digitales Prüfen prozess- und stakeholderorientiert denken mit der LUCA Office Simulation



VIOLA DEUTSCHER

Prof. Dr., Professorin an der
Universität Mannheim
viola.deutscher@
uni-mannheim.de



ANDREAS RAUSCH

Prof. Dr., Professor an der
Universität Mannheim
rausch@uni-mannheim.de



JÜRGEN SEIFRIED

Prof. Dr., Professor an der
Universität Mannheim
juergen.seifried@
uni-mannheim.de

Innovationen in der beruflichen Bildung, die im Rahmen der Digitalisierung in Schule und Betrieb erfolgen, drohen letztendlich zu scheitern, wenn sie nicht auch im Prüfungssystem abgebildet werden. In diesem Sinne befasst sich der Beitrag mit der Konzeption eines prozess- und stakeholderorientierten Prüfungsmodells für die berufliche Bildung zur Organisation und Durchführung von Prüfungen sowohl für Zwischen- und Abschlussprüfungen (Prüfungen auf zentraler Ebene) als auch für betriebliche und schulische Lernerfolgskontrollen (dezentrale Ebene). Die Potenziale digitaler kaufmännischer Zwischen- und Abschlussprüfungen mit der LUCA Office Simulation werden anhand eines Umsetzungsbeispiels skizziert. Im Ausblick wird reflektiert, welche Aspekte die Potenziale des digitalen Prüfens in der beruflichen Bildung aktuell noch einschränken.

Digitale Prüfungen in der beruflichen Bildung

Im Zuge der Digitalisierung und Automatisierung von Tätigkeiten steigen die Kompetenzanforderungen an Fachkräfte (vgl. FREY/OSBOURNE 2017). Aus Sicht der Berufsbildung gewinnen somit die Förderung komplexerer beruflicher Handlungskompetenzen und die Arbeit mit digitalen Tools an Bedeutung (vgl. RAUSCH u. a. 2021). Für das berufliche Lehren und Lernen ist dabei zentral, dass dieser Wandel nicht nur an den verschiedenen Lernorten, sondern auch in Prüfungen konstruktiv aufgegriffen wird (vgl. SEVERING/WEIß 2011; PELLEGRINO 2012). Die Digitalisierung von Prüfungen leistet daher einen wichtigen Beitrag zur Qualitätssicherung in der beruflichen Bildung. Für die Akzeptanz von digitalen Prüfungen sind aus unserer Sicht zwei Aspekte von zentraler Bedeutung: Zum einen sollten Prüfungsumgebungen umfassende Lösungen für verschiedene Prüfungsphasen unter Einbeziehung aller an der Prüfung Beteiligten bereitstellen. Zum anderen sollten digitale Prüfungsumgebungen intuitiv zu bedienen sein (Benutzerfreundlichkeit). Idealerweise sind solche Prüfungsumgebungen für Zwischen- und Abschlussprüfungen daher gleichzeitig auch als Lern- und Prüfungsumgebungen für Lernende und Lehrende an beruflichen Schulen und in Ausbildungsbetrieben verfügbar, sodass alle Beteiligten bereits im Vorfeld der Prüfungen Erfahrungen mit entsprechenden Prozessen und Prüfungsumgebungen sammeln können.

Konzeptioneller Rahmen für digitale Prüfungen

Prüfen in der beruflichen Bildung

Bei der Gestaltung beruflicher Prüfungen im Sinne einer summativen Feststellung der beruflichen Handlungsfähigkeit sind grundsätzlich zwei Ebenen zu differenzieren: die *zentrale Ebene* der Zwischen- und Abschlussprüfungen sowie die *dezentrale Ebene* des Prüfungsgeschehens in Ausbildungsbetrieben und an beruflichen Schulen (z. B. Klassenarbeiten).

Auf der zentralen Ebene sind verschiedene Sozialpartner und Institutionen involviert, deren Zusammenarbeit in Prüfungsumgebungen zu koordinieren ist. So sind neben den zu prüfenden Personen sowie Prüferinnen und Prüfern insbesondere auch die zuständigen Stellen (i. d. R. die Kammern) als Beteiligte zu berücksichtigen. Darüber hinaus kommt bei zentralen Zwischen- und Abschlussprüfungen insbesondere der Arbeit in den Prüfungsausschüssen eine hohe Bedeutung zu, in denen sowohl Arbeitgeber- als auch Arbeitnehmerseite sowie mindestens eine Lehrkraft vertreten sind (vgl. GAJEK u. a. 2021). Mit Blick auf dezentrale Prüfungen in Schulen und Ausbildungsbetrieben sind Lehrkräfte und Ausbilder/-innen neben den Lernenden die wesentlichen Nutzergruppen.

Wenn man nun von digitalem Prüfen, *E-Assessments* oder *Computer-Based Assessments* spricht, wird hiermit oft eine Prüfung am Bildschirm eines digitalen Endgeräts assozii-

iert. Solch enge Definitionen finden sich sowohl in der Forschungsliteratur als auch in einschlägigen Praxis- und Lernmaterialien zur Prüfungsthematik (vgl. KLEIJ u. a. 2012). Diese Auffassung ist zwar nicht falsch, führt aber zu einem verkürzten Verständnis von digitalem Prüfen. Umfassendere Definitionen beziehen dagegen neben der digitalen Prüfungsdurchführung auch die digitale Prüfungserstellung, die Automatisierung der Prüfungsauswertung sowie die Bereitstellung von digitalem Feedback zur Leistung mit ein und stellen damit auf eine Digitalisierung des gesamten Prüfungsprozesses ab (vgl. CRISP 2011).

Gestaltung des (digitalen) Prüfungsprozesses

Die Gestaltung von digitalen Prüfungsprozessen, die sich im Wesentlichen nicht von jener der analogen Variante unterscheidet, umfasst folgende Schritte:

1. Im Zuge der Prüfungserstellung müssen Aufgaben erstellt, ggf. angepasst und schließlich bezogen auf die Repräsentativität und Reihenfolge der Aufgaben innerhalb der Prüfung didaktisch sinnvoll zusammengestellt werden.
2. Für die Prüfungsdurchführung muss ein räumlicher und zeitlicher Rahmen festgelegt werden, in dem die Prüfung stattfinden soll. Während der Prüfungsdurchführung muss i. d. R. eine Prüfungsaufsicht erfolgen, um die Durchführungsobjektivität der Prüfung zu gewährleisten.
3. Zur Prüfungsauswertung müssen auf Basis von Bewertungsrastern eine Auswertung der erbrachten Prüfungsleistungen und schließlich auch eine Beurteilung dieser Daten in Form eines Abgleichs mit dem pädagogischen Erwartungshorizont erfolgen.
4. Abschließend muss ein Prüfungsfeedback an die Prüferinnen und die geprüften Personen erfolgen.

Während des gesamten Prozesses sind an verschiedenen Stellen die Wahrung von Gütekriterien (Objektivität, Reliabilität, Validität) und eine faire und transparente Prüfungsabwicklung unter Beachtung des Datenschutzes sowie bei Zwischen- und Abschlussprüfungen die Wahrung der in der Ausbildungsordnung festgelegten Prüfungsanforderungen zu gewährleisten. Im Hinblick auf die Validität der Prüfungen sind sechs Qualitätskriterien – u. a. in Anlehnung an Arbeiten zur Authentizität beruflicher Assessments (vgl. GULIKERS/BASTIAENS/KIRSCHNER 2004) – zu beachten (vgl. DEUTSCHER/WINTHER 2022):

1. Zunächst muss die Prüfungsumgebung technisch den berufstypischen Arbeitsplatz möglichst realitätsnah abbilden. Erfolgt die Aufgabenbearbeitung am Arbeitsplatz z. B. vorwiegend computerbasiert (z. B. bei kaufmännischen Aktivitäten), so sind Paper-Pencil-Formate in Prüfungen nur bedingt valide und es müssen die zur

Arbeitsbewältigung üblicherweise eingesetzten (digitalen) Arbeitswerkzeuge abgebildet werden. Findet die Arbeitstätigkeit dagegen überwiegend im nicht-digitalen Modus statt (z. B. im Bereich der Pflegeberufe), sind digitale Prüfungsmodi hinsichtlich ihrer Validität durchaus auch kritisch zu hinterfragen.

2. Zudem müssen die Prüfungsinhalte relevant und auch repräsentativ in Bezug auf das berufliche Curriculum sein und bei Zwischen- und Abschlussprüfungen den Prüfungsanforderungen entsprechen.
3. Weiterhin müssen für die Prüfungsinhalte Aufgaben gestaltet werden, die sich hinsichtlich ihrer Aufgabenschwierigkeit an den realen Gegebenheiten am Arbeitsplatz orientieren.
4. Auch muss die Prüfungsumgebung den organisationalen Kontext, in dem die Arbeitstätigkeiten stattfinden, möglichst realitätsnah abbilden, z. B. mithilfe eines Modellunternehmens.
5. Dem Verständnis von beruflicher Handlungskompetenz folgend sollten sich die Prüfungsaufgaben zudem auf vollständige Arbeits- und Geschäftsprozesse beziehen. Entsprechend sollte auch die Sequenzierung von Aufgaben in beruflichen Prüfungen der »natürlichen« Abfolge beruflicher Handlungen bzw. Tätigkeiten entsprechen.
6. Schließlich zeichnen sich berufliche Tätigkeiten oftmals durch ein hohes Maß an sozialer Interaktion aus, sodass die Ab- oder Nachbildung sozialer Interaktionen empfehlenswert ist. Dies kann durch die soziale Situierung der zu prüfenden Person in die Organisation sowie durch die Zuweisung einer Rolle in dieser Organisation erfolgen. Über diese soziale Einbettung können dann auch soziale Kompetenzaspekte (z. B. wertschätzende schriftliche Kommunikation in E-Mails oder Posts) im Zuge des Prüfungsgeschehens miterfasst werden.

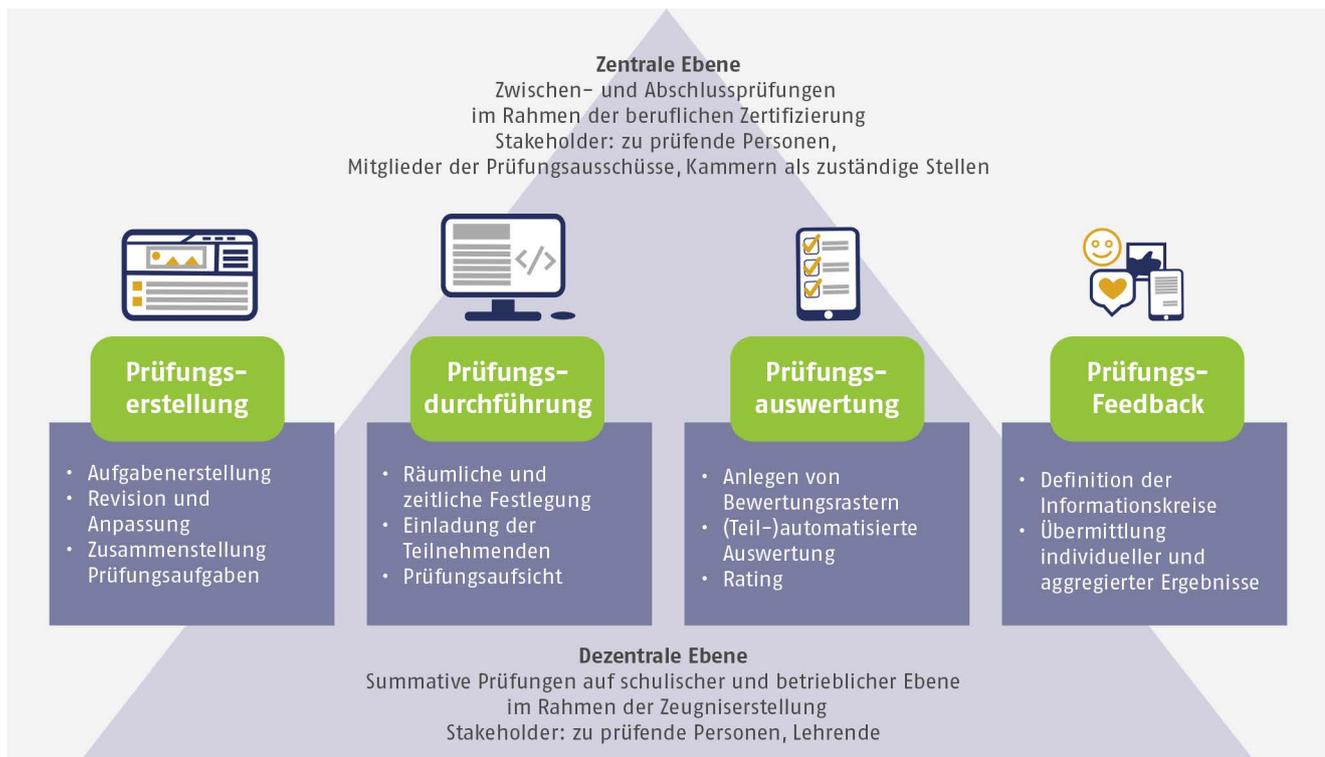
Abbildung 1 gibt einen zusammenfassenden Überblick über die Gestaltungsparameter (digitaler) Prüfungsprozesse in der beruflichen Bildung. Digitale Prüfungen werden gemäß den obigen Überlegungen zusammenfassend definiert als *technologiebasierte Stützung (durch Hard- und Softwarelösungen) des gesamten Prüfungsprozesses (Prüfungserstellung, -durchführung, -auswertung und -Feedback) auf zentraler oder dezentraler Ebene unter Einbeziehung aller relevanten Stakeholder*.

Gestaltung digitaler Zwischen- und Abschlussprüfungen mit der LUCA Office Simulation

Die skizzierten Schritte und Qualitätsmerkmale von Prüfungsprozessen sind gleichermaßen sowohl für digitale als auch für analoge Prüfungsformate relevant. Was digitales Prüfen nun von analogem Prüfen unterscheidet, sind verschiedene technische Möglichkeiten bei der Bewältigung

Abbildung 1

Darstellung beruflicher Prüfungsebenen und -prozesse



der oben dargestellten vier Prozessschritte. Neue Möglichkeiten für digitale berufliche Prüfungen mit besonderem Fokus auf die kaufmännischen Zwischen- und Abschlussprüfungen werden nachfolgend am Beispiel einer webbasierten Unterrichts- und Assessmentplattform, der LUCA Office Simulation (vgl. Infokasten), für die oben skizzierten Prozessschritte skizziert.

Digitale Prüfungserstellung

Bei der digitalen Prüfungserstellung besteht die Möglichkeit, Aufgabenentwürfe dauerhaft in ihrer digitalen Darbietungsform zu sichern (z. B. über Cloud-Technologie). Über geschützte Aufgabenpools lässt sich auf bestehende Aufgaben im System zugreifen, welche mühelos duplizierbar

und modifizierbar sind, sodass sich der Aufwand für die Entwicklung von Prüfungsaufgaben an dieser Stelle deutlich reduzieren kann. Ein weiterer Vorteil bei der Prüfungserstellung bei webbasierten Systemen besteht in der Möglichkeit der kollaborativen Aufgabenentwicklung. Dazu können weitere an der Prüfung Mitwirkende über das System eingeladen sowie zur Prüfergruppe hinzugefügt werden und dann in Echtzeit kollaborativ an den Aufgabenentwürfen arbeiten. Im Aufgabenerstellungssystem (*LUCA Editor*; vgl. Abb. 2, S. 24) kann neben einem Modellunternehmen auch ein hierzu passendes System zur Ressourcenplanung (sog. ERP-System) angelegt werden. Zudem besteht die Möglichkeit, die Aufgabenstellung sozial eingebettet als E-Mail mit Handlungsaufforderungen durch Kolleginnen und Kollegen oder Vorgesetzte anzulegen. Dabei wird nicht – wie auch in analogen Prüfungen durchaus üblich – ein Bild einer E-Mail dargestellt, sondern es muss direkt im Mailprogramm gearbeitet werden (korrektes Versenden etc.). Neben komplexen Aufgaben mit überwiegend offenen Antwortoptionen lassen sich aber auch geschlossene Aufgabenformate in die Prüfungsaufgaben einbauen. Im *LUCA Manager* werden die in Form von Arbeitsszenarien angelegten Prüfungsaufgaben dann – der Geschäftsprozesslogik folgend – zu »Prüfungsheften« zusammengestellt. Ein Beispiel für einen geschäftsprozessorientierten Testheftentwurf findet sich bei KLOTZ (2015).

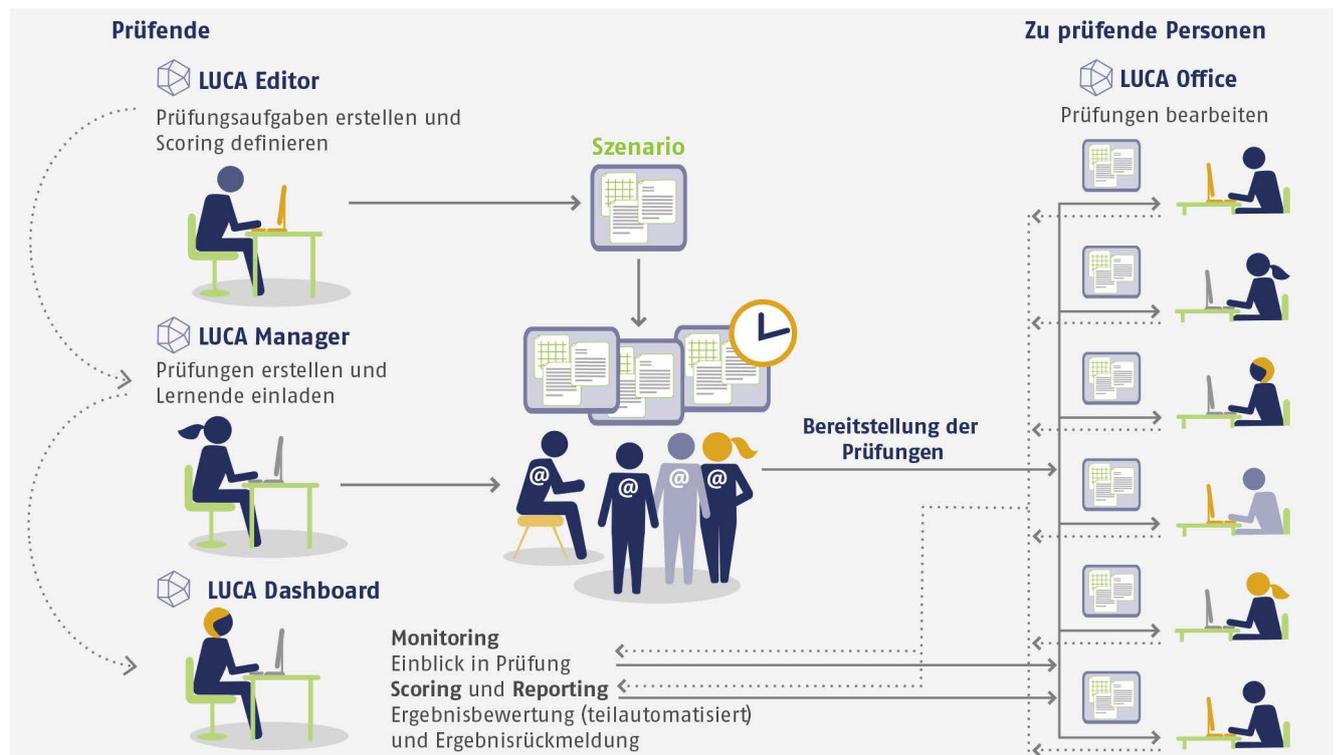
Die Lern- und Assessmentplattform LUCA

LUCA wurde im Rahmen der Forschungs- und Transferinitiative ASCOT+ des Bundesministeriums für Bildung und Forschung (BMBF) im Zeitraum von 2019 bis 2022 an der Universität Mannheim entwickelt. Sie ist als browser-basierter Online-Dienst für Lehr- und Assessmentzwecke kostenfrei verfügbar. Der Name LUCA spielt auf Luca Pacioli (1445–1514), einen Pionier der kaufmännischen Didaktik, an.

Weitere Informationen: <https://luca-office.de/>

Abbildung 2

Architektur des LUCA-Prüfungssystems



Digitale Prüfungsdurchführung

Die Prüfungsdurchführung erfolgt im *LUCA Office*, das typische, im Funktionsumfang allerdings reduzierte Bürowerkzeuge bereitstellt. Die zu prüfenden Personen werden über einen Link im *LUCA Manager* zur Prüfung eingeladen, navigieren frei durch die Umgebung und nutzen verschiedene digitale Werkzeuge, die zur Lösung des Arbeitsauftrags benötigt werden (z. B. E-Mail-Client, ERP-System, Tabellenkalkulations-Software). Gleichzeitig können die Prüfenden den Aufgabenfortschritt in der Gruppenansicht und den individuellen Stand der Bearbeitung einzelner zu prüfender Personen über das *LUCA Monitoring* verfolgen. Eine Prüfungsaufgabe beginnt mit einem Arbeitsauftrag (ggf. inklusive angehängter Dateien) und endet mit der Antwort-Mail auf die Auftrags-Mail oder durch Ablauf der vorgesehenen Bearbeitungszeit. Die Prüfung insgesamt endet, wenn alle einzelnen Prüfungsaufgaben einer Prüfung abgeschlossen wurden oder die Prüfungszeit abgelaufen ist.

Digitale Prüfungsauswertung

Bei digitalen Prüfungen können bestimmte Auswertungen voll automatisiert erfolgen. Hierzu gleicht ein Algorithmus die Lösungen mit den zuvor von den Prüfern eingetragenen korrekten Lösungen bzw. Lösungsräumen (bei mehreren korrekten Lösungen) ab. Automatisierungsoptionen

bestehen bei komplexeren Plattformen – wie LUCA – nicht nur für geschlossene Antworten (z. B. Multiple-Choice-Aufgaben), sondern auch für Logdaten (z. B. in Bezug auf die Eintragung von Werten oder Bezeichnungen in einer Tabellenkalkulation). Die Bewertung von Logdaten als Werte des Verhaltens einer Person in einer Prüfungsumgebung ist gerade für die Erfassung von Handlungskompetenz im Sinne des tatsächlich gezeigten Verhaltens in Testsituationen eine hochgradig authentische Bewertungsform. Darüber hinaus können unterstützend digitale Elemente für die manuelle Auswertung angeboten werden. Typische Leistungsdimensionen für schriftliche Ausarbeitungen sind z. B. inhaltliche Argumentation, Strukturierung von Texten, Angemessenheit der Ausdrucksweise sowie Rechtschreibung und Grammatik. Dazu müssen diese Leistungsdimensionen in digitalen Bewertungsrastern genauer definiert sowie detaillierte Vorgaben für die Bewertung der einzelnen Leistungskriterien entwickelt werden (vgl. JONSSON/SVINGBY 2007). Eine weitere Besonderheit bei digitalen Auswertungssystemen liegt darin, dass Prüfende nach Prüfungsdurchführung die automatisierten Korrekturen direkt abrufen und prüfen sowie offene Aufgaben – allein oder kollaborativ in Echtzeit – bewerten können. Dies ist gerade bei Zwischen- und Abschlussprüfungen aufgrund der hohen Relevanz der Testergebnisse von Vorteil.

Digitales Prüfungsfeedback

Das Prüfungsfeedback ist für die verschiedenen Anspruchsgruppen unterschiedlich auszugestalten. Kammern beispielsweise erhalten die Prüfungsergebnisse aggregiert in Form einer Übersicht über Durchschnittswerte für die (vorab durch die Prüfenden definierten) Kompetenzdimensionen und kennungsbasiert auf Individualebene der Prüfungsteilnehmenden in Bezug auf die erreichten Punkte (ggf. differenziert nach verschiedenen Kompetenzdimensionen). Die zu prüfenden Personen hingegen erhalten ihre individuellen Prüfungsergebnisse über ihren LUCA-Zugang.

Abbildung 2 gibt einen zusammenfassenden Überblick über den digitalen Prüfungsprozess mit der LUCA Office Simulation.

Grenzen von digitalen Prüfungen

Neben den aufgezeigten Potenzialen digitaler Prüfungen sind auch deren Grenzen in den Blick zu nehmen. Zunächst ist zu konstatieren, dass die Entwicklung authentischer Prüfungsformate – wie auch die Entwicklung authentischer Unterrichtsformate – aufwändig ist. Der Arbeitsschwerpunkt verschiebt sich dabei im Prüfungsprozess stärker von der Prüfungsauswertung (am Ende des Prüfungsprozesses) in Richtung der Aufgabenerstellung (zu Beginn des Prozesses), da bereits hier multidimensionale Bewertungsraster anzulegen sind, die dann im Zuge der Prüfungsauswertungsphase automatisierte Auswertungen ermöglichen. Die Erstellung dieser Bewertungsraster ist gerade für geschäftsprozessorientierte und problemhaltige Prüfungsaufgaben nicht trivial, da solche Aufgaben zur Steigerung der Authentizität

idealerweise offene Antwortformate und die Berücksichtigung verschiedener Lösungswege und Folgefehler vorsehen (vgl. z. B. KLOTZ 2015). Einschränkend ist in Bezug auf die automatisierte Auswertung von Logdaten in Zwischen- und Abschlussprüfungen zudem anzumerken, dass derzeit nach § 42 Absatz 4 Satz 2 des BBiG nur Antwort-Wahl-Aufgaben explizit für automatisierte Auswertungen vorgesehen sind. Um das Potenzial von Logdaten für Zwischen- und Abschlussprüfungen nutzen zu können, wäre hier eine Erweiterung um die neuen technischen Möglichkeiten perspektivisch zu diskutieren. Erwähnenswert ist zudem, dass aktuell die Möglichkeiten im Rahmen der automatisierten Datenauswertung oft noch hinter den Erwartungen zurückbleiben (ungeachtet der aktuellen Entwicklungen im Bereich der KI). Komplexere Verfahren des automatisierten Bewertens nutzen zwar Machine-Learning-Algorithmen (solche sind auch in LUCA über R-Scripts implementierbar). Diese Verfahren benötigen allerdings eine große Menge an gelabelten Trainingsdaten, d. h. vollständig und korrekt ausgewertete Fälle. Zudem können die hieraus resultierenden Algorithmen nur für gleiche oder sehr ähnliche Aufgaben verwendet werden. Dies macht dieses Verfahren für gleichartige Prüfungen mit vielen Teilnehmenden auf zentraler Ebene interessant. Für die kommenden Jahre ist diesbezüglich mit weiteren Entwicklungen zu rechnen. ◀



Abbildungen von S. 23 und S. 24 zum Download
 Abb. 1: www.bwp-zeitschrift.de/g12069
 Abb. 2: www.bwp-zeitschrift.de/g12072

LITERATUR

- CRISP, G.: *Teacher's handbook on e-Assessment: A handbook to support teachers in using e-assessment to improve and evidence student learning and outcomes*. San Francisco 2011
- DEUTSCHER, V.; WINTHER, E.: *Assessment of vocational competences – Definitions, issues and quality criteria*. In: HARTEIS, C.; GIJBELS, D.; KYINDT, E. (Hrsg.): *Research approaches on workplace learning. Insights from a growing field*. Cham 2022, S. 305–320
- FREY, C. B.; OSBOURNE, M. A.: *The future of employment: How susceptible are jobs to computerisation?* In: *Technological Forecasting and Social Change* 48 (2017) 114, S. 254–280
- GAJEK, S.; HEWLEET, C.; KAISER, F.; KAHL-ANDERSEN, A.; KEUP-GOTTSCHALCK, M.; LABUSCH-SCHÖNWANDT, G.: *Handbuch für Prüfende in der beruflichen Bildung*. Baden-Baden 2021
- GULIKERS, J. T. M.; BASTIAENS, T. J.; KIRSCHNER, P. A.: *A Five-Dimensional Framework for Authentic Assessment*. In: *Educational Technology Research and Development* 52 (2004) 3, S. 67–86
- JONSSON, A.; SVINGBY, G.: *The use of scoring rubrics: Reliability, validity and educational consequences*. In: *Educational Research Review* 2 (2007) 2, S. 130–144

KLEIJ, F. VAN DER; EGGEN, T. H.; TIMMERS, C. F.; VELDKAMP, B. P.: *Effects of feedback in a computer-based assessment for learning*. In: *Computers & Education* 58 (2012) 1, S. 263–272

KLOTZ, V.: *Diagnostik beruflicher Kompetenzentwicklung: eine wirtschaftsdidaktische Modellierung für die kaufmännische Domäne*. Wiesbaden 2015

PELLEGRINO, J.W.: *The design of an assessment system focused on student achievement. A learning sciences perspective on issues of competence, growth and measurement*. In: BERNHOLT, S.; NEUMANN, K.; NENTWIG, P. (Hrsg.): *Making it tangible – Learning outcomes in science education*. Münster 2012, S. 79–107

RAUSCH, A.; DEUTSCHER, V.; SEIFRIED, J.; BRANDT, S.; WINTHER, E.: *Die web-basierte Bürosimulation LUCA – Funktionen, Einsatzmöglichkeiten und Forschungsausblick*. In: *Zeitschrift für Berufs- und Wirtschaftspädagogik* 117 (2021), S. 372–394

SEVERING, E.; WEIß, R. (Hrsg.): *Prüfungen und Zertifizierung in der beruflichen Bildung – Anforderungen – Instrumente – Forschungsbedarf*. Bielefeld 2011