

TOBIAS KÄRNER / FRANZISKA RÖHRIG / JULIA WARWAS

Situative Anforderungscharakteristika lernprozessbegleitender Diagnostik im Unterricht

Systematisches Literaturreview und Fragebogenentwicklung

Situational demands of formative assessment of student learning during class Systematic literature review and questionnaire development

KURZFASSUNG: Ziel des Beitrags ist die literaturgestützte Systematisierung sowie empirische Erfassung von Anforderungscharakteristika lernprozessbegleitender Diagnosen im laufenden Unterricht. Basierend auf einem systematischen Literaturreview sowie auf der Verarbeitung von Fragebogendaten von insgesamt 150 Lehrpersonen werden die folgenden Dimensionen identifiziert, inhaltlich beschrieben, empirisch erfasst und im Ausblick hinsichtlich ihrer forschungs- sowie unterrichtspraktischen Verwendungsmöglichkeiten diskutiert: (1) Vielfalt an Informationen, (2) Informationsdefizite, (3) unsichere Informationsgrundlage, (4) Gefahr situativer Wahrnehmungs- und Urteilsverzerrungen, (5) Zeit- und Handlungsdruck, (6) hohe Entscheidungsverantwortung, (7) fehlende Regeln und Handlungsschemata.

Schlagworte: Anforderungsanalyse, Lehrpersonen, Diagnostik unterrichtlichen Erlebens und (Lern-)Verhaltens von Schüler/innen

ABSTRACT: The paper aims at a literature-based systematization and empirical measurement of demands that characterize formative assessment of student learning during class. Drawing on a systematic literature review and statistical analysis of survey data from 150 teachers, we identify, describe and empirically measure the following dimensions before discussing the resulting framework's scientific as well as practical utility: (1) variety of information, (2) information deficits, (3) uncertain informational basis, (4) risk of situational cognitive biases, (5) time pressure and pressure to act, (6) high decision-making responsibility, (7) lack of action rules and heuristics.

Keywords: demand analysis, teachers, diagnosis of students' experiences and (learning) behaviour during class



1. Ausgangslage und Zielstellung des Beitrags

In der beruflichen Bildung sowie in der Lehrer/innenbildung im Allgemeinen ist eine zunehmende unterrichtspraktische sowie wissenschaftliche Beschäftigung mit interindividuellen Unterschieden zwischen den Lernenden festzustellen (HEINRICHS/REINKE 2019; TERHART 2014). Dies erfolgt vor dem Hintergrund von Professionalisierungsdiskursen, denen zufolge eine adressatengerechte Unterrichtsgestaltung (z. B. BRÜHWILER/BLATCHFORD 2011), binnendifferenzierende Lernangebote (z. B. HUGENER/KRAMMER 2010) und individuelle, konstruktive Lernunterstützung zentrale Ziel- und Bestimmungsgrößen erfolgreichen Unterrichtens darstellen (z. B. HERMKES/MACH/MINNAMEIER 2018). Alle diese Maßnahmen setzen jedoch akkurate diagnostische Urteile über unterschiedsbestimmende Faktoren in einer Klasse voraus, um zu inhaltlich begründeten und eben auch adressatengerechten didaktischen Entscheidungen zu gelangen. Die Vielfalt der in der Literatur als entscheidungsrelevant thematisierten Merkmale von Lernenden ist groß. Hierzu zählen insbesondere fachbezogenes Wissen (z. B. TANGGAARD/NIELSEN/JØRGENSEN 2015), domänenspezifische (Problemlöse-)Kompetenzen (z. B. RAUSCH/KÖGLER/SEIFRIED 2019; SEMBILL/RAUSCH/KÖGLER 2013), kognitive Grundfähigkeiten bzw. Intelligenz und metakognitive Merkmale (z. B. LEHMANN/SEEBER 2013; PETSCH/NORWIG/NICKOLAUS 2012) sowie psychosoziale und emotional-motivationale Charakteristika (z. B. REINKE/HEINRICHS 2019; SEMBILL et al. 2002).

Die lehrpersonenseitige Fähigkeit, entsprechende diagnostische Urteile möglichst zutreffend fällen zu können, wird als diagnostische Kompetenz bezeichnet (SCHRADER 2010, 2013). Zur Quantifizierung der Akkuratheit, mit der entsprechende Einschätzungen getroffen werden, haben sich in der Literatur die drei nachfolgenden statistischen Kennwerte etabliert (SCHRADER/HELMKE 1987). Bezogen auf Schüler/innenmerkmale informiert die Niveauelemente darüber, ob das Lehrendenurteil die klassendurchschnittliche Ausprägungshöhe des interessierenden Merkmals exakt erfasst, überschätzt oder unterschätzt. In der Streuungskomponente kommt zum Ausdruck, ob die Lehrperson die Bandbreite bestehender Unterschiede zwischen den Merkmalsträger/innen korrekt einstuft, überzeichnet oder nivelliert. Mithilfe der Rangordnungskomponente kann ermittelt werden, wie gut es der Lehrperson gelingt, die Merkmalsträger/innen in eine den realen Verhältnissen entsprechende Rangreihe zu bringen, also bspw. treffsicher die besten und schlechtesten Schüler/innen der Klasse zu identifizieren. Hinsichtlich der diagnostischen Akkuratheitsmerkmale bestehen jedoch personenbezogene Unterschiede, welche u. a. auf die jeweils vorliegende lehrpersonenseitige Expertise zurückzuführen sind. Experten gelingt es hierbei schneller und passgenauer, sich ein Bild eines/r Schülers/in und der jeweiligen pädagogischen Situation zu machen, auch dann, wenn die verfügbaren Informationen mehrdeutig sind oder sogar zueinander in Widerspruch stehen, was als „holistic recognition of patterns“ (BERLINER 2004, S. 207) bezeichnet wird und auf ein umfangreiches und tiefgreifendes episodisches Gedächtnis zurückzuführen ist (BERLINER 1988, 2001). Im Vergleich zeigen sich bei Novizen oft Schwierigkeiten, verfügbare (mehrdeutige) Informationen zu einem schlüssigen

Gesamtbild zu integrieren (BERLINER 2001; BÖHMER/HÖRSTERMANN/GRÄSEL/KROLAK-SCHWERDT/GLOCK 2015).

Abgesehen von der personenspezifischen Expertise bzgl. der Einschätzung von Schüler/innenmerkmalen zeichnen bestehende Befunde zur Urteilsgenauigkeit von Lehrpersonen bzgl. der o. g. *interindividuellen* Unterschiede, abhängig vom Beurteilungsinhalt, ein eher ernüchterndes Bild: Fallen die Einschätzungen von fachbezogenen Leistungen und allgemeinen kognitiven Fähigkeiten der Schüler/innen noch moderat akkurat aus (z. B. HOGE/COLADARCI 1989; MACHTS et al. 2016; SÜDKAMP/KAISER/MÖLLER 2012), so zeigen sich für die Beurteilung emotional-motivationaler und psychosozialer Lernendenmerkmale meist nur geringe Einschätzungsgenauigkeiten (z. B. HOSENFELD/HELMKE/SCHRADER 2002; PRAETORIUS/KOCH/SCHUNPFLUG/ZEINZ/DRESEL 2017; STANG/URHAHNE 2016). Die zugehörigen Studien haben allerdings relativ zeitstabile Schüler/innenmerkmale wie etwa die generelle Leistungsängstlichkeit oder das schulische Fähigkeitsselbstkonzept zum Gegenstand. Für eine adressatengerechte Gestaltung unterrichtlicher Interaktionen und Lerngelegenheiten ist hingegen (auch) die Diagnose situationsgebundener und damit *intraindividuell* schwankender *Erlebensqualitäten und Verhaltensweisen* der Schüler/innen während des Lernprozesses von hoher Relevanz (WARWAS/KÄRNER/GOLYSZNY 2015). Zugleich ist davon auszugehen, dass die Komplexität, Vielschichtigkeit, Mehrdimensionalität und Perspektivenabhängigkeit unterrichtlicher Situationen (BECK 1996; HEID 2001) die akkurate Einschätzung lernbegleitender Erlebensqualitäten und Verhaltensweisen von Lernenden erheblich erschweren bzw. sogar beeinträchtigen kann (zsf. WARWAS/KÄRNER/GOLYSZNY 2015). Da nach unserem Wissen derzeit noch keine systematische Anforderungsanalyse für diese wichtige pädagogische Aufgabe vorliegt, setzt sich der vorliegende Beitrag das *Ziel*, Anforderungscharakteristika lernprozessbegleitender Diagnosen im laufenden Unterricht zum einen auf Basis einer literaturgestützten Systematisierung auszdifferenzieren und zum anderen mithilfe strukturprüfender Verfahren an Befragungsdaten zu untersuchen, ob auch Lehrpersonen in ihrer Wahrnehmung lernprozessdiagnostischer Anforderungen zwischen den literaturgestützt ermittelten Charakteristika unterscheiden. Orientiert am *Vorgehen* der arbeitspsychologischen Anforderungsanalyse (SCHULER 2014) werden daher in einem ersten Schritt die situationsspezifischen Anforderungen diagnostischer Leistungen während des Unterrichtsgeschehens literaturbasiert identifiziert und beschrieben. Die so gewonnenen Beschreibungen dienen als inhaltliche Grundlage für die Konstruktion von Skalen, deren dimensionale Struktur in einem zweiten Schritt faktorenanalytisch überprüft wird.

2. Inhaltliche Bestimmung von situativen Anforderungscharakteristika lernprozessbegleitender Diagnostik

Bevor situative Anforderungscharakteristika lernprozessbegleitender Diagnostik im Unterrichtsprozess inhaltlich näher bestimmt werden, soll zunächst erläutert werden, was im Folgenden unter Diagnostik im pädagogischen Sinne, unter Prozessbezug und

This material is under copyright. Any use outside of the narrow boundaries of copyright law is illegal and may be prosecuted.

This applies in particular to copies, translations, microfilming as well as storage and processing in electronic systems.

© Franz Steiner Verlag, Stuttgart 2020

unter einer diagnostischen Situation verstanden wird. Nach VAN OPHUYSEN und LINTORF (2013, S. 60) umfasst die *pädagogische Diagnostik* „das systematische Sammeln von Informationen mittels vornehmlich sozialwissenschaftlicher, wenngleich unstandardisierter Instrumente sowie deren Analyse und Integration, um professionelles pädagogisches Handeln vorzubereiten und zu unterstützen.“ KARST, KLUG und UFER (2017, S. 102) definieren eine *diagnostische Situation* im Allgemeinen als „Verhältnisse oder Umstände innerhalb oder außerhalb des Unterrichts, in denen pädagogische Diagnostik durch die Lehrperson stattfindet.“ Daneben sind diagnostische Situationen den Autor/innen zufolge (1) pädagogischen Handlungen vor- bzw. untergeordnet und haben zum Ziel, (2) Informationen für zukünftige pädagogische Entscheidungen zu gewinnen. Die Autoren/innen betonen weiterhin, dass die aus der diagnostischen Entscheidung resultierenden pädagogischen bzw. didaktischen Handlungen nicht mehr Gegenstand einer diagnostischen Situation sind. Ferner sind sie auf (3) einzelne Schüler/innen bezogen und umfassen (4) die Beurteilung lernrelevanter Merkmale. Anders als im Rahmen der Statusdiagnostik üblich, bei welcher in erster Linie zeit- und situationsinvariante Merkmale (z. B. Intelligenz) erfasst und bewertet werden, hat die *prozessbezogene Diagnostik* die Einschätzung von zeit- und bedingungsabhängigen Erlebensqualitäten und Verhaltensweisen von Schüler/innen (z. B. aktuelle Unter- bzw. Überforderung, situatives Interesse) während des laufenden Unterrichts zum Ziel (AUFSCHNAITER et al. 2015; PAWLIK 1976). Da eine diagnostische Situation in dem hier verstandenen Sinne nur in geringem Maße planbar ist und während des Unterrichts stattfindet, wird sie vorrangig dem sog. „On-the-Fly Assessment“ innerhalb der formativen Diagnostik zugeordnet (vgl. GLOGGER-FREY/HERPPICH 2017; SHAVELSON et al. 2008).

Die aus der lernprozessbegleitenden Diagnostik *resultierenden Anforderungen* für die Lehrperson sind vorwiegend situationsspezifisch und konzentrieren sich auf *Bedingungen* des diagnostischen Auftrags und dessen *Ausführung*. Um vorwiegend Anforderungen der diagnostischen Tätigkeiten zu erfassen, wird unsere Anforderungsanalyse daher an der Ebene des situationsgebundenen Verhaltens ausgerichtet (vgl. SCHULER 2014; SONNTAG/FRIELING/STEGMAIER 2012). Dementsprechend sollen nicht konkrete (Teil-)Aufgaben oder Schritte in einer diagnostischen Situation (z. B. das Aufstellen diagnostischer Hypothesen) oder die dafür jeweils notwendigen Eigenschaften, Fertigkeiten und Fähigkeiten der Lehrkraft erfasst werden. Vielmehr konzentriert sich die Anforderungsanalyse auf die an Lehrpersonen gestellten *situationsabhängigen, verhaltensrelevanten Anforderungen*, welche mit der Durchführung diagnostischer Tätigkeiten während des laufenden Unterrichts verbunden sind (hierzu VAN OPHUYSEN/LINTORF 2013).

2.1 Systematisches Literaturreview zur inhaltlichen Bestimmung der Anforderungsdimensionen

Durch eine systematische Literaturrecherche wurden einschlägige Publikationen identifiziert, welche Hinweise auf potenzielle diagnosebezogene Anforderungen an Lehrkräfte liefern (RÖHRIG 2019). Als Richtlinie diente dabei das für systematische Berichte

und Meta-Analysen entwickelte *Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Statement* (LIBERATI et al. 2009), wobei individuelle Anpassungen im Detaillierungsgrad der Vorgaben des PRISMA Statements gemäß Zielsetzung der qualitativen Analyse vorgenommen wurden. Als Informationsquellen der Recherche dienten nationale und internationale Wissenschaftsdatenbanken, Google Scholar sowie lokale Datenbanken. In Folge des explorativen Charakters der Entwicklungsphase wurden einschlägige *Literaturdatenbanken* gewählt (Tabelle 1).

Tab. 1. Informationsquellen

| | |
|------------------------------|---|
| Elektronische Datenbanksuche | ERIC, JSTOR, Fachportal Pädagogik, PsycARTICLES, PsycINFO, Teacher Reference Center |
| Lokale Datenbanken | KonSearch, Lokaler Katalog der Universitätsbibliothek Konstanz |
| Manuelle Suche | Google Scholar |

Trotz methodischer Nachteile wurde als Suchstrategie eine manuelle Handsuche gewählt und auf die Verwendung eines starren Suchstrings verzichtet. Es können zwar *Schlüsselbegriffe* wie beispielsweise „Anforderung“, „Lehrkräfte“ oder „Diagnostik“ (bzw. deren jeweilige englischsprachige Begriffe) abgeleitet werden, weshalb eine Definition von Suchkombinationen mittels Boolescher Operatoren grundsätzlich denkbar gewesen wäre. Dies erschien uns jedoch angesichts der teils ambivalenten Begriffsverständnisse innerhalb der Literatur nicht zielführend und hätte zu einer kaum bewältigbaren Menge an initialen Treffern geführt. Hinsichtlich der Quellenauswahl wurden Titel, Abstracts und Volltexte gesichtet und unter Zuhilfenahme der folgenden *Auswahlkriterien* selektiert:

- Der Zeitpunkt der Veröffentlichung ist unbegrenzt, da ein möglichst repräsentativer und zeitlich offener Überblick über mögliche diagnosebezogene Anforderungen an Lehrpersonen geschaffen werden soll.
- Die Veröffentlichung muss ferner in deutscher oder englischer Sprache verfügbar sein.
- Daneben ist es notwendig, dass sich die Veröffentlichung inhaltlich mit Anforderungen in diagnostischen Situationen beschäftigt.
- Neben explizit diagnosebezogenen Anforderungen an Lehrkräfte werden auch Informationen aus verwandten Berufen, wie der Diagnostik im medizinischen und psychologischen Bereich sowie im Managementbereich, eingeschlossen, da die generelle Idee der Diagnostik (gezielt Informationen erfassen, um daraus Schlussfolgerungen, Entscheidungen und Handlungen abzuleiten) auch in den genannten Bereichen gilt.
- Die Veröffentlichungen können darüber hinaus schulfachspezifische wie auch schulfachübergreifende Anforderungen in diagnostischen Situationen einbeziehen.

Insgesamt konnten über die systematische Suche 45 Quellen identifiziert werden, welche sich entsprechend der in Tabelle 2 dargestellten deskriptiven Merkmale charak-

terisieren lassen (eine Übersicht findet sich in Anhang 1 – Auflistung der gesichteten Quellen).

Tab. 2. Deskriptive Merkmale der gesichteten Quellen

| Kriterium | | Anzahl Quellen | Prozent |
|----------------------------|--|-----------------------|----------------|
| Stichprobenumfang | K | 45 | 100.0 |
| Veröffentlichungs- jahr | vor 2010 | 7 | 15.6 |
| | ab 2010 | 38 | 84.4 |
| Publikationssprache | Deutsch | 39 | 86.7 |
| | Englisch | 6 | 13.3 |
| Land | Deutschland | 40 | 88.9 |
| | Europa (ohne Deutschland, Schweiz) | 2 | 4.4 |
| | Schweiz | 1 | 2.2 |
| | USA | 2 | 4.4 |
| Beitragsart | Beitrag in Sammelband | 24 | 53.3 |
| | Dissertation | 2 | 4.4 |
| | Journal/Fachzeitschrift | 12 | 26.7 |
| | Monografie | 6 | 13.3 |
| | Zeitungsartikel | 1 | 2.2 |
| Untersuchungs- kontext | Betriebswirtschaftlicher Bereich | 2 | 4.4 |
| | Fächerübergreifender Schulbereich | 28 | 62.2 |
| | Fachspezifischer Schulbereich (Biologie) | 2 | 4.4 |
| | Fachspezifischer Schulbereich (Mathematik) | 2 | 4.4 |
| | Medizinischer Bereich | 2 | 4.4 |
| | Psychologischer Bereich | 9 | 20.0 |

Im Zuge der Quellenanalyse konnten durch die Autor/innen des Beitrags die folgenden sieben Anforderungsdimensionen inhaltlich gefasst und kategorisiert werden: (1) Vielfalt an grundsätzlich verfügbaren Informationen bei gleichzeitig vorliegenden (2) Informationsdefiziten, eine (3) unsichere Informationsgrundlage, die mögliche (4) Gefahr situativer Wahrnehmungs- und Urteilsverzerrungen aufseiten der Lehrperson, (5) Zeit- und Handlungsdruck, eine (6) hohe Entscheidungsverantwortung sowie (7) fehlende Regeln und Handlungsschemata. Die identifizierten Dimensionen werden nachfolgend näher beschrieben.



2.2 Vielfalt an Informationen

Im Zusammenhang mit diagnostischen Situationen in pädagogischen Handlungsvollzügen gehen zahlreiche Autor/innen von einem „Information overload“ aus, was eine seitens der Lehrperson wahrgenommene Informationsüberflutung beschreibt. Auf die Lehrkraft strömen aus vielen unterschiedlichen Quellen sensorische Eindrücke ein (HASCHER 2011; KLUG 2017), welche sich auf verbale, paraverbale und nonverbale Beobachtungen und Einschätzungen beziehen (FÖRSTER/BÖHMER 2017; LEICHNER 1979). Die Lehrkraft muss dabei permanent bewerten, welche der verfügbaren Informationen für eine pädagogische Entscheidung relevant bzw. irrelevant oder gar irreführend sind (BEHRMANN/GLOGGER-FREY 2017; KLUG 2017). Anstelle einer isolierten Betrachtung von Schüler/innenmerkmalen wird dabei von Lehrpersonen gefordert, ebenso Kontextinformationen ihres Auftretens (z. B. bzgl. Lerninhalt, Unterrichtsverlauf, Mediennutzung etc.) zu berücksichtigen (HESSE/LATZKO 2017). Urteile müssen in der Folge auf Basis von sehr vielen verfügbaren Informationen getroffen werden (HASCHER 2011). Aufgrund grundsätzlich beschränkter kognitiver Verarbeitungskapazitäten ist es hierbei erforderlich, dass Lehrkräfte eine Differenzierung relevanter und irrelevanter Informationen vornehmen (BÖHMER/ENGLICH/BÖHMER 2017; HETMANEK/VAN GOG 2017) und so der Forderung nachkommen, ausschließlich valide Signale bzw. Indikatoren der interessierenden Merkmale zu verwerten (FÖRSTER/BÖHMER 2017; HESSE/LATZKO 2017). Das durch hochgradige Multidimensionalität geprägte Informationsgefüge im Unterrichtsprozess (DOYLE 1977) wird auch dadurch bedingt, dass Lehrkräfte mit ihren Einschätzungen selten nur ein einziges Ziel verfolgen. Vielmehr müssen parallel multiple Ziele Berücksichtigung finden, die über die (Un-)Wichtigkeit verschiedener Erlebens- und Verhaltensindikatoren mitentscheiden (HETMANEK/VAN GOG 2017). Hinzu kommt, dass die Diagnostik während des Unterrichts quasi „nebenbei“ stattfindet (POTT 2019; PRAETORIUS/HETMANEK/HERPPICH/UFER 2017) und somit diagnostische Urteile häufig „on-the-fly“ getroffen werden. Zugehörige Tätigkeiten wie z. B. Informationssammlung und -integration müssen oftmals parallel ausgeführt werden (ALTMANN/NÜCKLES 2017; SCHRADER 2017). Die Lehrperson kann sich dabei meist nicht auf einzelne Schüler/innen im Detail konzentrieren, sondern hat gleichzeitig unterschiedliche Beurteilungsgegenstände und situative Gegebenheiten im Blick zu behalten (WARWAS/KÄRNER/GOLYSZNY 2015).

2.3 Informationsdefizite

Lernprozessbezogene Einschätzungen von Lehrpersonen werden in der Literatur weiterhin mit dem erschwerenden Aspekt der unvollständigen Informationslage in Verbindung gebracht, da davon ausgegangen wird, dass Lehrpersonen valide und relevante Merkmale des Erlebens und Verhaltens in Unterrichtssituationen in einem holistischen Sinne nicht *vollständig* beobachten bzw. kognitiv verarbeiten können (HESSE/LATZKO 2017). Diagnostische Situationen im laufenden Unterricht weisen einen geringen Grad

an Transparenz auf und erschweren damit die Urteilsfindung, da Lehrkräfte vorwiegend hypothetische Konstrukte bewerten (HESSE/LATZKO 2017), welche nicht direkt beobachtbar sind (FÖRSTER/BÖHMER 2017; HASCHER 2011; POTT 2019). HESSE und LATZKO (2017) gehen davon aus, dass Lehrkräfte im Unterricht ihre Einschätzungen nur valide durch den Einsatz merkmalspezifischer Erhebungsinstrumente generieren können. Fehlen Lehrpersonen jedoch aussagekräftige lernprozessbezogene Informationen für eine situationsgebundene diagnostische Entscheidung, so ist es i. d. R., bspw. aufgrund retrospektiver Verzerrungen, nicht möglich, diese nachträglich zu gewinnen (BEHRMANN/GLOGGER-FREY 2017). Auch sollen in aktuelle diagnostische Entscheidungen Informationen über früheres Lernverhalten die Schüler/innen einbezogen werden, die unter Umständen – beispielsweise aufgrund fehlender Dokumentation – der Lehrperson nicht vorliegen (BEHRMANN/VAN OPHUYSEN 2017). Ohne entsprechende Informationen (z. B. Aufzeichnungen im Klassenbuch) sind getroffene Entscheidungen möglicherweise zu einseitig akzentuiert (BALOGH/MILLER/BALL 2015; BEHRMANN/VAN OPHUYSEN 2017), was eine oft notwendige multiperspektivische Betrachtung behindert bzw. unmöglich macht.

2.4 Unsichere Informationsgrundlage

Wahrgenommene Informationen können weiterhin mehrdeutig sein und/oder im Widerspruch zueinander stehen (FERNÁNDEZ-BALLESTEROS et al. 2001; WESTHOFF/HORNKE/WESTMEYER 2003). Eine Lehrperson ist hierbei mit möglicherweise irreführenden Informationen konfrontiert, die eine Behinderung ihrer diagnostischen Tätigkeiten darstellen können (BEHRMANN/GLOGGER-FREY 2017; BÖHMER/ENGLICH/BÖHMER 2017). Einer aus dem wissenschaftlichen Kontext heraus geforderten „Informationsvalidität“ kann in der Unterrichtspraxis meist nicht konsequent nachgekommen werden, was in deutlichem Kontrast zur geforderten „Objektivität“ lehrpersonenseitiger Einschätzungen und Bewertungen steht (FÖRSTER/BÖHMER 2017). Valide schüler/innenseitige Verhaltensindikatoren, aus denen beispielsweise aktuelle Erlebensqualitäten oder Fähigkeitsniveaus abgeleitet werden könnten, sind i. d. R. nicht verfügbar. Überdies können Schüler/innen wissentlich oder unwissentlich falsche Informationen liefern. Insgesamt ist es also für Lehrpersonen aufgrund einer unsicheren Informationsgrundlage schwierig einzuschätzen, ob die subjektiv repräsentierte Umwelt den objektiven Gegebenheiten entspricht (STIEGLITZ 2008).

2.5 Gefahr situativer Wahrnehmungs- und Urteilsverzerrungen

Es ist davon auszugehen, dass Lehrkräfte beim Treffen diagnostischer Entscheidungen durch verzerrte Informationsaufnahme, -verarbeitung und -bewertung behindert werden (BEHRMANN/VAN OPHUYSEN 2017; FERNÁNDEZ-BALLESTEROS et al. 2001; STIEGLITZ 2008; WESTHOFF/HORNKE/WESTMEYER 2003). Neben Erwartungseff-

fekten bzw. Vorurteilen (ORTNER 2010; SEMBILL 1987; SÜDKAMP/KAISER/MÖLLER 2017) und Sympathien bzw. Antipathien (LORENZ 2011) zählen systematische Beurteilungstendenzen (zsf. HESSE/LATZKO 2017) zu den beurteilerabhängigen Hindernissen einer situationsangemessenen Urteilsbildung. HESSE und LATZKO (2017) merken hierzu an, dass „Lehrerdiagnosen im Unterricht [...] häufig auch durch eine pädagogisch günstige Voreingenommenheit getragen“ werden. LEICHTNER (1979) weist in diesem Zusammenhang auf den Umstand hin, dass hypothesenbestätigende Informationen eher Berücksichtigung finden als hypothesenablehnende Informationen, obwohl die Bedeutsamkeit der Informationen in gleichem Maße ausgeprägt sei. Daneben wird davon ausgegangen, dass motivationale Orientierungen und individuelle Selbstregulationsfähigkeiten diagnostische Urteile beeinflussen können (BEHRMANN/VAN OPHUYSEN 2017; KARING/SEIDEL 2017a). Beispielsweise wird an die Lehrkraft die Anforderung gestellt, ihre eigenen Emotionen situationsangemessen zu regulieren, was Menschen im Allgemeinen sowie abhängig von den jeweiligen situativen Gegebenheiten unterschiedlich gut gelingt (vgl. SCHLADITZ/DRÜGE 2017). Des Weiteren können Lehrkräfte bei diagnostischen Aufgaben mit Interessenskonflikten konfrontiert sein, da ihr eigenes Normen- und Wertesystem nicht zwangsläufig mit demjenigen der Lernenden oder den rechtlichen Vorgaben (z. B. landeseinheitliche Richtlinien zur Bewertung von Schülern/innen) übereinstimmt (FERNÁNDEZ-BALLESTEROS et al. 2001; WESTHOFF, HORNKE/WESTMEYER 2003). Neben den genannten beurteilerabhängigen Hindernissen situationsangemessener Einschätzungen können aber auch Gegebenheiten der Diagnosesituation selbst die Urteilsfindung systematisch beeinflussen. Derartige situationsbedingte Verzerrungen sind insb. aus der Forschung zur Leistungsbewertung in mündlichen Prüfungssituationen bekannt, wo sie als Primingeffekte, reihenfolgeinduzierte Kontrasteffekte und Halo-Effekte nachgewiesen werden konnten (im Überblick INGENKAMP/LISSMANN 2008).

2.6 Zeit- und Handlungsdruck

Im Allgemeinen ist die Unterrichtszeit (i. d. R. 45-Min.-Taktung) beschränkt und kann von der Lehrperson nur sehr begrenzt beeinflusst werden (KARST/KLUG/UFER 2017). Lehrpersonen haben in der Folge nicht beliebig viel Zeit diagnostische Einschätzungen vorzunehmen; es fehlen ihnen somit für eine ausführliche Informationssammlung und -bewertung häufig zeitliche Kapazitäten (FÖRSTER/KARST 2017; GLOGGER-FREY/HERPPICH 2017; HETMANEK/VAN GOG 2017; SCHRADER 2017). In Folge des Zeitdrucks stehen mehrdimensionale und differenzierte Urteile aufgrund des hohen Aufwands bei der Urteilsbildung meist in Konflikt mit weiteren unterrichtsbezogenen Tätigkeiten (HESSE/LATZKO 2017). Ein bewusstes und explizites Urteilen ist folglich oft nicht umsetzbar (ebd.; HERPPICH et al. 2017). Vielmehr müssen diagnostische Tätigkeiten unmittelbar vor einer zu treffenden pädagogischen bzw. didaktischen Entscheidung spontan und sehr schnell durchgeführt werden (BEHRMANN/VAN OPHUYSEN 2017; KAISER et al. 2017). Dabei kann sich die Lehrkraft i. d. R. nicht unmittelbar aus dem Un-

terrichtverlauf zurückziehen, um über die diagnostische Tätigkeit nachzudenken bzw. eine getroffene Entscheidung eingängig zu reflektieren (HESSE/LATZKO 2017).

2.7 Hohe Entscheidungsverantwortung

Diagnostische Tätigkeiten sind weiterhin mit Anforderungen in Bezug auf die erlebte Entscheidungsverantwortung verbunden. Je nach wahrgenommener Schwere der mit einem diagnostischen Urteil verbundenen Konsequenzen messen Lehrpersonen der Wichtigkeit einer exakten Einschätzung interessierender Schüler/innenmerkmale variierende Bedeutung bei und betreiben bei der situativen Informationsauswahl und -integration unterschiedlich hohe(n) Aufwand und Sorgfalt (BEHRMANN/VAN OPHUYSEN 2017; BÖHMER/ENGLISH/BÖHMER 2017; FERNÁNDEZ-BALLESTEROS et al. 2001; HESSE/LATZKO 2017; HETMANEK/VAN GOG 2017; KARST/KLUG/UFER 2017; SCHRADER 2017; WESTHOFF/HORNKE/WESTMEYER 2003). Eng damit verbunden ist der Umstand, dass diagnostische Urteile je nach Situation unterschiedlich verbindlich sind bzw. sein müssen (KARST/KLUG/UFER 2017) und bei Fehlentscheidungen bzw. falschen Urteilen „Kosten“ im Sinne von negativen Folgen für den/die betreffende/n Schüler/in entstehen können (BEHRMANN/VAN OPHUYSEN 2017). Mit zunehmender Schwere der Konsequenzen wächst daher idealerweise der Anspruch, möglichst akkurate, transparente und faire Urteile zu bilden (ebd.) und konsistente und inhaltlich begründete Entscheidungen zu treffen (FÖRSTER/BÖHMER 2017). Daneben müssen Lehrkräfte mit ihrer Positionsmacht als Beurteiler verantwortungsvoll umgehen (HOEFERT 1995), was beispielsweise die vertrauliche Behandlung von Informationen über Schüler/innen einschließt (BEAUDUCEL/LEUE 2014). Zugleich können sie mit ethischen Problemen ihrer Entscheidung konfrontiert sein (ZIEGLER/BÜHNER 2012) und müssen sich der Tragweite und Bedeutung diagnostischer Aufgaben bewusst sein, was unter Umständen eine situationsangemessene Distanzierung erforderlich macht (SCHLADITZ/DRÜGE 2017).

2.8 Fehlende Regeln und Handlungsschemata

Regeln bzw. Maßstäbe im Kontext *summativer* diagnostischer Tätigkeiten (z. B. Notengebung, Leistungsfeststellung) sind für Lehrpersonen grundsätzlich verfügbar, welche im Sinne von Handlungsempfehlungen hilfreich, aber gleichzeitig auch einschränkend wirken können. Diese umfassen institutionelle Richtlinien und rechtliche Rahmenbedingungen (z. B. Schulgesetze) (BEAUDUCEL/LEUE 2014; KARST/KLUG/UFER 2017) wie auch professionelle bzw. berufsethische Leitlinien (FERNÁNDEZ-BALLESTEROS et al. 2001; WESTHOFF/HORNKE/WESTMEYER 2003). Jedoch fehlen Lehrpersonen im laufenden Unterrichtsprozess oft verbindliche Regeln, Orientierungsmaße und Handlungsschemata zum Fällen diagnostischer Urteile (HERPPICH et al. 2017). Beispielsweise sind Zeitpunkte zur Durchführung formativer diagnostischer Einschätzungen nicht

verbindlich festgelegt. Auch das diagnostische Ziel ist meist nicht vorgegeben bzw. definiert (BEHRMANN/GLOGGER-FREY 2017). Vielmehr müssen sich Lehrkräfte oft selbst für einen angemessenen Beurteilungsmaßstab entscheiden, also inwiefern sie sich bei der Einschätzung interessierender Schüler/innenmerkmale an formalisierten Kriterien wie Lernzielen, an sozialen Vergleichsgrößen oder an individuellen Entwicklungsverläufen orientieren (HESSE/LATZKO 2017). Verbindliche Bewertungskriterien liegen ebenfalls i. d. R. nicht vor und es fehlen meist Interpretations- oder Bewertungsempfehlungen bzw. -standards (HERPPICH et al. 2017). Die Schwierigkeit beginnt in diesem Zusammenhang bereits bei der Informationsaufnahme aufgrund mangelhafter bzw. unvollständiger Operationalisierung der Beurteilungskriterien (HOEFERT 1995). Lehrpersonen treffen die Informationsauswahl und erbringen Einschätzungen im laufenden Unterricht folglich schwerpunktmäßig nach ihrem eigenen Ermessen (HERPPICH et al. 2017).

3. Empirische Erfassung situativer Anforderungscharakteristika einer lernprozessbegleitenden Diagnostik

Die literaturgestützt ermittelten situativen Anforderungscharakteristika lernprozessbegleitender Diagnostik (siehe Abschnitt 2) wurden in einem zweiten Schritt operationalisiert. In diesem Abschnitt werden daher das Vorgehen bei der Datenerhebung und die Stichprobe, der Umgang mit fehlenden Werten sowie das Vorgehen bei der Itementwicklung und der Prüfung der Faktorenstruktur beschrieben.

3.1 Datenerhebung und Stichprobe

Als Grundlage der empirischen Überprüfung wurde eine anonyme und onlinebasierte Datenerhebung durchgeführt. Da es sich bei Lehrpersonen erfahrungsgemäß um eine eher schwer zu erreichende Zielgruppe handelt, haben wir uns bei der Stichprobengewinnung für das Schneeballverfahren entschieden (vgl. DÖRING/BORTZ 2016). Die Einladung zur Befragung wurde über universitäre Mailverteiler an eingeschriebene und ehemalige Studierende des beruflichen sowie allgemeinbildenden Lehramts versandt sowie in einschlägigen Internetforen für Lehrpersonen veröffentlicht. Die angeschriebenen Personen wurden darüber hinaus auch gebeten, die Anfrage zur Studienteilnahme an ihnen bekannte Lehrkräfte weiterzuleiten. Die Items wurden randomisiert abgefragt, um einerseits abbruchbedingte, systematisch fehlende Werte aufgrund des Reihenfolgeplatzes zu vermeiden und um andererseits Reihenfolgeeffekte in den gegebenen Antworten möglichst auszuschließen.

Insgesamt riefen 455 Personen den bereitgestellten Link zum Fragebogen auf, 200 Personen fingen an, den Fragebogen zu bearbeiten, wobei 50 Personen die Bearbeitung bereits vor den relevanten Items abbrachen bzw. diese nicht ernsthaft bearbeiteten, womit letztendlich 150 verwertbare Fragebögen übrig blieben.

This material is under copyright. Any use outside of the narrow boundaries of copyright law is illegal and may be prosecuted.

This applies in particular to copies, translations, microfilming as well as storage and processing in electronic systems.

© Franz Steiner Verlag, Stuttgart 2020

Die finale Stichprobe besteht damit aus 150 praktizierenden Lehrkräften (103 weiblich, 46 männlich, 1 divers) mit einem Durchschnittsalter von 34,23 Jahren ($SD = 10,57$; Min. = 19; Max. = 71) und einer durchschnittlichen Anzahl an Dienstjahren (exklusive dem Referendariat) von 6,97 ($SD = 9,52$, Min. = 0, Max. = 42). Der jüngste Teilnehmer gab an, als Lehrkraft für Deutsch als Fremdsprache tätig zu sein, womit auch das niedrige Alter erklärbar ist. Somit wurde der Teilnehmer nicht aus der Stichprobe herausgenommen, da die Voraussetzungen der Zielgruppe „praktizierende Lehrkraft“ als gegeben anzusehen sind. Zum Erhebungszeitpunkt befanden sich 57 der Befragten im Vorbereitungsdienst.

Die überwiegende Mehrheit der Befragten war zum Erhebungszeitpunkt in Deutschland als Lehrkraft tätig ($n = 144$, 96 %; zusätzlich eine Person aus Österreich und fünf Personen aus der Schweiz). Bei den in Deutschland unterrichtenden Lehrpersonen sind die folgenden Bundesländer vertreten: Baden-Württemberg ($n = 81$, 56,3 %), Bayern ($n = 33$, 22,9 %), Nordrhein-Westfalen ($n = 13$, 9 %), Hessen ($n = 11$, 7,6 %), Niedersachsen ($n = 2$, 1,4 %), Rheinland-Pfalz ($n = 1$, 0,7 %), Saarland ($n = 1$, 0,7 %), Berlin ($n = 1$, 0,7 %) und Schleswig-Holstein ($n = 1$, 0,7 %).

Bezogen auf die Gesamtstichprobe ist der überwiegende Teil der Befragten an Grundschulen ($n = 47$), an Gymnasien ($n = 39$) oder an beruflichen Schulen (z. B. Berufsschule, Berufsfachschule, Fachoberschule) ($n = 33$) tätig. Daneben befinden sich in der Stichprobe Lehrkräfte aus Schularten mit mehreren Bildungsgängen (z. B. Gemeinschaftsschule) ($n = 15$), dem Bildungsgang Hauptschule (z. B. Hauptschule, Mittelschule, Werkrealschule) ($n = 11$), Förderschulen ($n = 9$), Realschulen ($n = 9$) und sonstigen Schulen (z. B. Bildungsträger, Schule für Gesundheits- und Krankenpflege) ($n = 3$).

In der Gesamtstichprobe sind folgende Unterrichtsfächer bzw. Fachgebiete vertreten: Mathematik, Informatik, Naturwissenschaften und Technik (MINT) ($n = 88$), Deutsch ($n = 76$), Gesellschaftswissenschaften (z. B. Erdkunde, Geschichte, Politik) ($n = 60$), Kunst, Musik und Textilgestaltung ($n = 53$), Fremdsprachen ($n = 40$), Sport ($n = 41$), Religion, Philosophie und Ethik ($n = 32$), Wirtschaftswissenschaften (z. B. Büromanagement, Gesamtwirtschaft, Übungsfirma) ($n = 30$), Deutsch als Fremd- bzw. Zweitsprache ($n = 4$) und Sonstiges (z. B. Hauswirtschaft) ($n = 3$).¹

3.2 Umgang mit fehlenden Werten

Aufgrund des umfangreichen Fragebogens sind vereinzelte Datenlücken zu verzeichnen. Für die relevanten Items ergibt sich ein mittlerer Anteil fehlender Werte von 17,1 % (Minimum: $n = 20$, 13,3 %; Maximum: $n = 36$; 24 %). Die Elimination einzelner Items mit fehlenden Werten wurde als nicht zielführend eingeschätzt, da nach einer näheren Prüfung nicht davon ausgegangen werden kann, dass die fehlenden Werte systematisch begründbar sind. Die annähernd gleichmäßige Verteilung fehlender Werte je Item ist

1 Die höhere Gesamtanzahl bei den Schularten und den unterrichteten Fächern kommt zustande, da manche Personen an mehreren Schultypen tätig sind bzw. mehrere Fächer unterrichten.

auch mit der randomisierten Abfrage der Items zu erklären. Die Abwägung eines fall- bzw. listenweisen Fallausschlusses wurde verworfen, da sich angesichts der Verteilung fehlender Werte über alle Items hinweg eine sehr reduzierte Anzahl an Beobachtungen (Analyse mit listenweisem Fallausschluss: $n = 85$) ergeben hätte. Ebenfalls nachteilig erschien die Ersetzung der fehlenden Daten durch den Mittelwert, da mit diesem Verfahren einhergehende Verzerrungen (z. B. starke Verzerrungen der Schätzer der Kovarianzen) zu nicht mehr positiv definiten Korrelationsmatrizen führen können (vgl. RAESSLER/RUBIN/SCHENKER 2008; RUBIN 1987). Um bei den statistischen Analysen angesichts der sowieso bereits verhältnismäßig kleinen Stichprobe keine zusätzliche Reduktion hinnehmen zu müssen, wurden die fehlenden Werte unter Verwendung der korrespondierenden SPSS-Standardroutine *multiple imputation* imputiert. Es wurden für die interessierenden Items jeweils 5 Werte für jeden fehlenden Wert geschätzt, welche entsprechend gemittelt wurden. Die weiterführenden Datenanalysen basieren demnach auf dem imputierten Datensatz. Nennenswerte Abweichungen in den deskriptiven Werten zwischen dem imputierten und dem nicht-imputierten Datensatz sind nicht zu verzeichnen (siehe Tabelle 4 unten).

3.3 Itementwicklung und Prüfung der Faktorenstruktur

Der Prozess der Itemformulierung verlief iterativ und durchlief verschiedene Stufen. Zunächst wurden insgesamt 575 Rohitems formuliert. Die Ausformulierung der Rohitems erfolgte durch die Autor/innen auf Grundlage der in Abschnitt 2 beschriebenen Anforderungscharakteristika und der in Anhang 1 gelisteten Referenzquellen. Anschließend wurden aufgrund von inhaltlichen Redundanzen, ungeeigneten bzw. unverständlichen Itemformulierungen sowie aus erhebungsökonomischen Gründen insgesamt 424 Items entfernt. Aus dem ursprünglichen Itempool verblieben somit 151 Items, welche im Zuge zweier leitfadengestützter Experteninterviews durch praktizierende Lehrpersonen (eine weibliche Grundschullehrkraft und ein männlicher Gymnasiallehrer mit jeweils mehrjähriger Berufserfahrung) nochmals auf inhaltliche Verständlichkeit und sprachliche Korrektheit überprüft und ggf. überarbeitet wurden.

Hinsichtlich des Abfrageformats wurden zunächst unterschiedliche Varianten anhand bestehender Fragebögen aus dem Bereich der psychologischen Arbeits- bzw. Anforderungsanalyse (z. B. Instrument zur stressbezogenen Arbeitsanalyse; SEMMER/ZAPF/DUNCKEL 1998) sowie in der Literatur vorgeschlagene Skalen (z. B. BÜHNER 2011) gesichtet und gegeneinander abgewogen. Die Wahl fiel aus testtheoretischen und -ökonomischen Gründen auf ein geschlossenes Antwortformat mit einer fünfstufigen Ratingskala (1 = „trifft gar nicht zu“ bis 5 = „trifft völlig zu“).

Die Prüfung der Faktorenstruktur auf Grundlage der erhobenen Daten erfolgte zunächst mittels explorativer Faktorenanalyse (Hauptachsenanalyse mit Varimax Rotation) unter Verwendung der Software SPSS 25® (IBM®, Chicago, USA). Hierbei orientierten wir uns an dem bei ROST und SCHERMER (1989) angegebenen Vorgehen: Eigenwert > 1 , Scree-Test und Ellenbogenkriterium, pro Komponente mehr als drei Markiertvariab-

len, Faktorladungen der einzelnen Items > 0.4 sowie keine starken Mehrfachladungen. Unter Anwendung der genannten Kriterien verblieben insgesamt 42 finale Items.

Die anschließende konfirmatorische Prüfung der Faktorenstruktur erfolgte mit der Software Mplus4 (MUTHÉN/MUTHÉN 1998–2010) unter Rückbezug auf die gängigen Gütekriterien. Die Gütekriterien des Gesamtmodells geben Hinweise auf eine sehr gute Modellpassung und auch die Fit-Indizes der einzelnen Modelle sind als gut bis sehr gut zu bewerten (Abbildung 1, Tabelle 3). Vergleicht man das 1-Faktormodell (Loglikelihood = -7339.372, AIC = 14930.744, BIC= 15310.084) mit dem 7-Faktormodell (Loglikelihood = -7070.328, AIC = 14434.655, BIC= 14877.219), so schneidet letztgenanntes Modell sowohl hinsichtlich der Fit-Indizes als auch hinsichtlich der gängigen Informationskriterien besser ab.

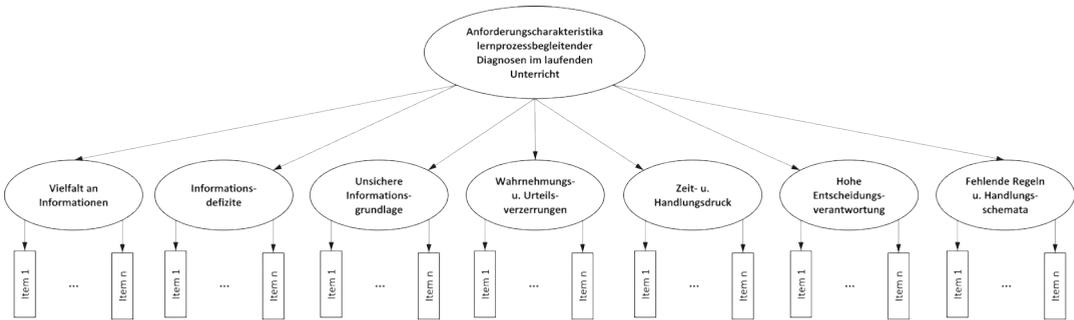


Abb. 1. Messmodell „Anforderungscharakteristika lernprozessbegleitender Diagnosen im laufenden Unterricht“

Tab. 3. Modelltests und Fit-Indizes

| Modell | χ^2 | d. f. | p | $\chi^2/d. f.$ | RMSEA | CFI | TLI | SRMR |
|---|----------|-------|---------|----------------|---------|-------|-------|-------|
| 1-Faktormodell | 1380.247 | 819 | ≤ 0.001 | 1.685 | 0.068 | 0.606 | 0.586 | 0.101 |
| 7-Faktormodell | 842.159 | 798 | 0.135 | 1.055 | 0.019 | 0.969 | 0.967 | 0.065 |
| Vielfalt an Informationen | 4.871 | 9 | 0.845 | 0.541 | ≤ 0.001 | 1.000 | 1.000 | 0.023 |
| Informationsdefizite | 2.328 | 2 | 0.312 | 1.164 | 0.033 | 0.997 | 0.991 | 0.022 |
| Unsichere Informationsgrundlage | 9.529 | 9 | 0.390 | 1.059 | 0.020 | 0.996 | 0.993 | 0.035 |
| Gefahr situativer Wahrnehmungs- und Urteilsverzerrungen | 5.917 | 5 | 0.314 | 1.183 | 0.035 | 0.996 | 0.992 | 0.024 |
| Zeit- und Handlungsdruck | 22.903 | 14 | 0.062 | 1.636 | 0.065 | 0.951 | 0.926 | 0.045 |
| Hohe Entscheidungsverantwortung | 22.494 | 20 | 0.314 | 1.125 | 0.029 | 0.981 | 0.973 | 0.045 |
| Fehlende Bewertungs- und Handlungsregeln | 5.292 | 9 | 0.808 | 0.588 | ≤ 0.001 | 1.000 | 1.000 | 0.023 |

Hinweise: n = 150; χ^2 = Chi-Quadrat-Wert; d. f. = Anzahl der Freiheitsgrade, CFI = Comparative Fit Index; TLI = Tucker-Lewis-Index; SRMR = Standardized Root Mean Square Residual; RMSEA = Root-Mean-Square-Error of Approximation

4. Deskriptive und korrelative Befunde

4.1 Deskriptive Ergebnisse

Die deskriptiven Item- und Skalenkennwerte sind in Tabelle 4 aufgeführt; die jeweiligen Itemformulierungen finden sich im Anhang 2 (Fragebogen zur lehrpersonenseitigen Einschätzung diagnosebezogener Anforderungen, LEDA).

Im Allgemeinen sind die Trennschärfen mit Werten zwischen 0.334 und 0.667 als angemessen ($r_{it} > 0.3$) zu bewerten. Die Skalenreliabilitäten erweisen sich mit Cronbachs Alpha Werten von 0.703 bis 0.809 als akzeptabel bis gut. Ebenfalls umfassen die Skalen mit 4 bis 8 Items pro Skala eine ausreichende Anzahl (vgl. WEIBER/MÜHLHAUS 2010). Die Itemschwierigkeiten bewegen sich im Bereich von 0.56 bis 0.87 und sind damit noch als angemessen zu bewerten, wobei vereinzelte Items in Folge eines hohen Mittelwertes Itemschwierigkeiten über 0.8 annehmen, was auf leichte Deckeneffekte insb. bei den Items des Faktors „Vielfalt an Informationen“ hinweist. Abbildung 2 veranschaulicht die Mittelwerte und Standardabweichungen der Skalen grafisch.

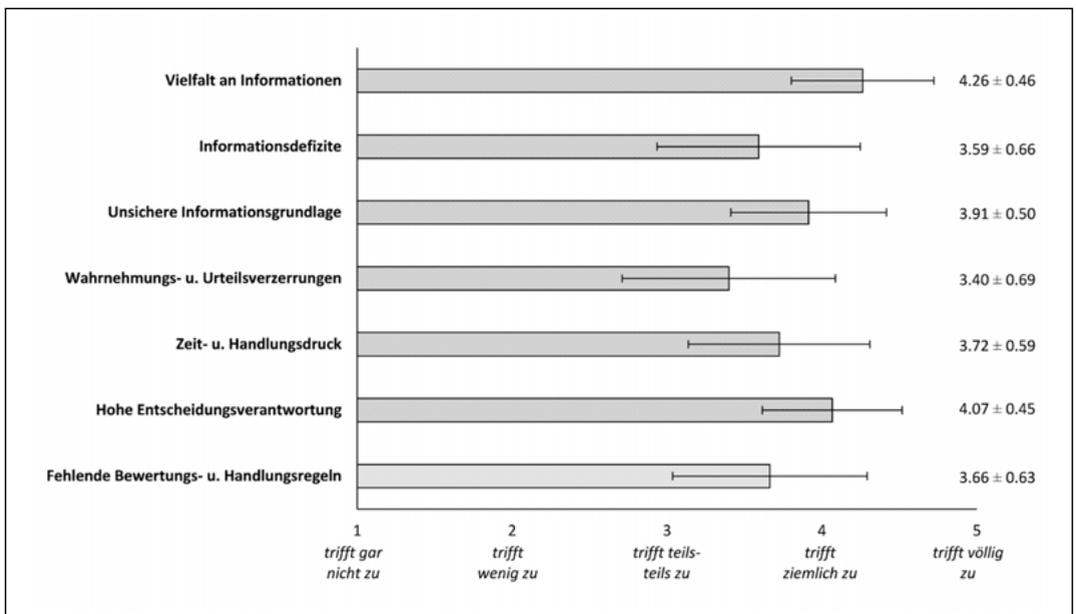


Abb. 2. Deskriptive Ergebnisse für die Skalenmittelwerte

Insgesamt zeigen sich mittlere bis hohe Mittelwertausprägungen im Bereich von $M = 3.40$ und $M = 4.26$ (wobei 3 = „trifft teils-teils zu“ und 4 = „trifft ziemlich zu“). Die Standardabweichungen nehmen Werte zwischen 0.45 und 0.69 an.

Tab. 4. Deskriptive Item- und Skalennennwerte

| Item | n | M | SD | n | M | SD | n | M | SD | Min. | Max. | P _i | λ _{ij} | CFAf.l. | r _{it} | α | m |
|---|-------------------|------|-------|-------------|------|-------|---|---|------|------|------|----------------|-----------------|---------|-----------------|-------|---|
| | (nicht-imputiert) | | | (imputiert) | | | | | | | | | | | | | |
| <i>Vielfalt an Informationen (Inf_VI)</i> | | | | | | | | | | | | | | | | | |
| Inf_VI_1 | 143 | 4.32 | 0.526 | 150 | 4.26 | 0.460 | | | | | | | 0.604 | 0.431 | 0.496 | 0.743 | 6 |
| Inf_VI_2 | 122 | 4.26 | 0.758 | 150 | 4.25 | 0.716 | 2 | 5 | 0.81 | | | | 0.619 | 0.389 | 0.496 | | |
| Inf_VI_3 | 126 | 4.37 | 0.666 | 150 | 4.33 | 0.631 | 2 | 5 | 0.83 | | | | 0.533 | 0.365 | 0.438 | | |
| Inf_VI_4 | 122 | 4.22 | 0.733 | 150 | 4.21 | 0.688 | 2 | 5 | 0.80 | | | | 0.566 | 0.429 | 0.486 | | |
| Inf_VI_5 | 126 | 4.23 | 0.812 | 150 | 4.18 | 0.760 | 1 | 5 | 0.80 | | | | 0.631 | 0.451 | 0.551 | | |
| Inf_VI_6 | 129 | 4.17 | 0.751 | 150 | 4.13 | 0.717 | 2 | 5 | 0.78 | | | | 0.475 | 0.309 | 0.419 | | |
| Inf_VI_6 | 128 | 4.51 | 0.676 | 150 | 4.48 | 0.653 | 2 | 5 | 0.87 | | | | 0.475 | 0.309 | 0.419 | | |
| <i>Informationsdefizite (Inf_DEF)</i> | | | | | | | | | | | | | | | | | |
| Inf_DEF_1 | 137 | 3.56 | 0.748 | 150 | 3.59 | 0.656 | | | | | | | | | | 0.711 | 4 |
| Inf_DEF_2 | 120 | 3.73 | 0.950 | 150 | 3.72 | 0.891 | 1 | 5 | 0.68 | | | | 0.652 | 0.579 | 0.510 | | |
| Inf_DEF_3 | 122 | 3.67 | 0.991 | 150 | 3.67 | 0.924 | 1 | 5 | 0.67 | | | | 0.731 | 0.673 | 0.603 | | |
| Inf_DEF_4 | 125 | 3.26 | 0.952 | 150 | 3.27 | 0.887 | 1 | 5 | 0.57 | | | | 0.570 | 0.504 | 0.475 | | |
| Inf_DEF_4 | 123 | 3.72 | 0.943 | 150 | 3.71 | 0.885 | 1 | 5 | 0.68 | | | | 0.539 | 0.475 | 0.404 | | |
| <i>Unsichere Informationsgrundlage (Inf_UN)</i> | | | | | | | | | | | | | | | | | |
| Inf_UN_1 | 145 | 3.91 | 0.581 | 150 | 3.91 | 0.503 | | | | | | | | | | 0.706 | 6 |
| Inf_UN_2 | 126 | 3.94 | 0.892 | 150 | 3.93 | 0.833 | 1 | 5 | 0.73 | | | | 0.589 | 0.489 | 0.480 | | |
| Inf_UN_3 | 126 | 3.70 | 0.923 | 150 | 3.70 | 0.865 | 1 | 5 | 0.68 | | | | 0.506 | 0.436 | 0.390 | | |
| Inf_UN_4 | 127 | 3.94 | 0.800 | 150 | 3.95 | 0.745 | 2 | 5 | 0.74 | | | | 0.589 | 0.437 | 0.484 | | |
| Inf_UN_5 | 121 | 3.83 | 0.799 | 150 | 3.84 | 0.733 | 1 | 5 | 0.71 | | | | 0.565 | 0.413 | 0.458 | | |
| Inf_UN_6 | 125 | 4.03 | 0.851 | 150 | 4.03 | 0.789 | 2 | 5 | 0.76 | | | | 0.525 | 0.413 | 0.425 | | |
| Inf_UN_6 | 128 | 4.06 | 0.811 | 150 | 4.03 | 0.768 | 2 | 5 | 0.76 | | | | 0.449 | 0.344 | 0.392 | | |
| <i>Wahrnehmungs- und Urteilsverzerrungen (WZ)</i> | | | | | | | | | | | | | | | | | |
| WZ_1 | 143 | 3.39 | 0.750 | 150 | 3.40 | 0.689 | | | | | | | | | | 0.809 | 5 |
| WZ_2 | 130 | 3.50 | 0.966 | 150 | 3.48 | 0.925 | 1 | 5 | 0.62 | | | | 0.731 | 0.674 | 0.665 | | |
| WZ_3 | 125 | 3.36 | 0.979 | 150 | 3.34 | 0.926 | 1 | 5 | 0.59 | | | | 0.751 | 0.692 | 0.667 | | |
| WZ_4 | 124 | 3.00 | 0.961 | 150 | 3.57 | 0.901 | 1 | 5 | 0.64 | | | | 0.597 | 0.535 | 0.502 | | |
| WZ_5 | 114 | 3.38 | 0.972 | 150 | 3.37 | 0.886 | 1 | 5 | 0.59 | | | | 0.653 | 0.577 | 0.557 | | |
| WZ_5 | 122 | 3.20 | 1.012 | 150 | 3.23 | 0.937 | 1 | 5 | 0.56 | | | | 0.673 | 0.628 | 0.590 | | |

This material is under copyright. Any use outside of the narrow boundaries of copyright law is illegal and may be prosecuted.

This applies in particular to copies, translations, microfilming as well as storage and processing in electronic systems.

© Franz Steiner Verlag, Stuttgart 2020

| Item | n | M | SD | n | M | SD | n | M | SD | Min. | Max. | P _i | λ _{ij} | CFAf.l. | r _{it} | α | m |
|--|-------------------|------|-------|-------------|------|-------|---|---|------|-------|-------|----------------|-----------------|---------|-----------------|---|---|
| | (nicht-imputiert) | | | (imputiert) | | | | | | | | | | | | | |
| <i>Zeit- und Handlungsdruck (ZeHa)</i> | | | | | | | | | | | | | | | | | |
| ZeHa_1 | 147 | 3,72 | 0,689 | 150 | 3,72 | 0,586 | 1 | 5 | 0,71 | 0,693 | 0,652 | 0,553 | 0,750 | 7 | | | |
| ZeHa_2 | 125 | 3,86 | 1,011 | 150 | 3,83 | 0,944 | 1 | 5 | 0,70 | 0,540 | 0,459 | 0,451 | | | | | |
| ZeHa_3 | 128 | 3,80 | 0,905 | 150 | 3,80 | 0,851 | 2 | 5 | 0,66 | 0,464 | 0,468 | 0,453 | | | | | |
| ZeHa_4 | 124 | 3,65 | 1,076 | 150 | 3,64 | 1,012 | 1 | 5 | 0,62 | 0,525 | 0,480 | 0,476 | | | | | |
| ZeHa_5 | 121 | 3,46 | 0,992 | 150 | 3,49 | 0,918 | 1 | 5 | 0,66 | 0,557 | 0,533 | 0,454 | | | | | |
| ZeHa_6 | 122 | 3,63 | 1,038 | 150 | 3,62 | 0,960 | 1 | 5 | 0,76 | 0,429 | 0,396 | 0,334 | | | | | |
| ZeHa_7 | 127 | 4,04 | 0,987 | 150 | 4,04 | 0,926 | 1 | 5 | 0,66 | 0,627 | 0,543 | 0,551 | | | | | |
| ZeHa_7 | 128 | 3,65 | 0,919 | 150 | 3,64 | 0,869 | 1 | 5 | 0,66 | 0,627 | 0,543 | 0,551 | | | | | |
| <i>Hole Entscheidungsverantwortung (HVA)</i> | | | | | | | | | | | | | | | | | |
| HVA_1 | 147 | 4,06 | 0,527 | 150 | 4,07 | 0,452 | | | | | | | 0,703 | 8 | | | |
| HVA_2 | 124 | 4,19 | 0,820 | 150 | 4,18 | 0,760 | 2 | 5 | 0,80 | 0,546 | 0,414 | 0,441 | | | | | |
| HVA_3 | 128 | 4,22 | 0,752 | 150 | 4,20 | 0,714 | 2 | 5 | 0,80 | 0,417 | 0,297 | 0,335 | | | | | |
| HVA_4 | 125 | 4,43 | 0,699 | 150 | 4,39 | 0,663 | 2 | 5 | 0,85 | 0,421 | 0,279 | 0,361 | | | | | |
| HVA_5 | 124 | 3,80 | 0,954 | 150 | 3,79 | 0,894 | 1 | 5 | 0,70 | 0,536 | 0,478 | 0,452 | | | | | |
| HVA_6 | 120 | 4,06 | 0,873 | 150 | 4,06 | 0,805 | 1 | 5 | 0,77 | 0,477 | 0,382 | 0,398 | | | | | |
| HVA_7 | 128 | 3,58 | 1,001 | 150 | 3,58 | 0,943 | 1 | 5 | 0,65 | 0,444 | 0,418 | 0,342 | | | | | |
| HVA_8 | 125 | 4,38 | 0,668 | 150 | 4,33 | 0,642 | 2 | 5 | 0,83 | 0,519 | 0,332 | 0,425 | | | | | |
| HVA_8 | 123 | 4,02 | 0,945 | 150 | 4,01 | 0,871 | 2 | 5 | 0,75 | 0,496 | 0,431 | 0,418 | | | | | |
| <i>Fehlende Regeln und Handlungsschemata (FHR)</i> | | | | | | | | | | | | | | | | | |
| FHR_1 | 145 | 3,69 | 0,758 | 150 | 3,66 | 0,628 | | | | | | | 0,762 | 6 | | | |
| FHR_2 | 122 | 3,52 | 0,998 | 150 | 3,53 | 0,925 | 1 | 5 | 0,63 | 0,697 | 0,643 | 0,600 | | | | | |
| FHR_3 | 130 | 3,93 | 0,982 | 150 | 3,95 | 0,918 | 1 | 5 | 0,74 | 0,635 | 0,581 | 0,523 | | | | | |
| FHR_4 | 117 | 3,58 | 0,958 | 150 | 3,61 | 0,874 | 1 | 5 | 0,65 | 0,535 | 0,466 | 0,456 | | | | | |
| FHR_5 | 123 | 3,61 | 1,021 | 150 | 3,59 | 0,950 | 1 | 5 | 0,65 | 0,536 | 0,508 | 0,463 | | | | | |
| FHR_6 | 123 | 3,67 | 1,069 | 150 | 3,64 | 0,992 | 1 | 5 | 0,66 | 0,636 | 0,629 | 0,535 | | | | | |
| FHR_6 | 122 | 3,68 | 0,990 | 150 | 3,68 | 0,915 | 1 | 5 | 0,67 | 0,503 | 0,458 | 0,446 | | | | | |

Hinweise: fünfstufige Likert-Skala: 1 = „trifft gar nicht zu“, 2 = „trifft wenig zu“, 3 = „trifft teils-teils zu“, 4 = „trifft ziemlich zu“, 5 = „trifft völlig zu“; P_i = Itemschwierigkeit; λ_{ij} = vollstandardisierte Faktorladungen aus der explorativen Faktorenanalyse; CFA f.l. = Faktorladungen der konfirmatorischen Faktorenanalyse; r_{it} = Itemtrennschärfe; α = Cronbachs Alpha; m = Anzahl an Items

4.2 Korrelative Ergebnisse

Die latenten Skalenkorrelationen zeigen, dass die erfassten Anforderungscharakteristika bis auf zwei Ausnahmen durchwegs gering bis moderat positiv miteinander in Zusammenhang stehen (Tabelle 5). Sehr klar differenzieren die Befragten bspw. zwischen den Bedingungen eines hohen Zeit- und Handlungsdrucks im Unterricht einerseits und den Schwierigkeiten, die aus Informationsvielfalt (0.176) oder Entscheidungsverantwortung (0.101) resultieren, während die Anforderungsgehalte unsicherer und defizitärer Informationslagen nachvollziehbarerweise als ähnlich eingestuft werden (0.803).

Tab. 5. Latente Skalenkorrelationen

| Faktoren | (1) | (2) | (3) | (4) | (5) | (6) |
|--|----------|----------|----------|----------|----------|--------|
| (1) Vielfalt an Informationen | | | | | | |
| (2) Informationsdefizite | 0.544*** | | | | | |
| (3) Unsichere Informationsgrundlage | 0.683*** | 0.803*** | | | | |
| (4) Wahrnehmungs- u. Urteilsverzerrungen | 0.421*** | 0.742*** | 0.673*** | | | |
| (5) Zeit- u. Handlungsdruck | 0.176 | 0.626*** | 0.575*** | 0.570*** | | |
| (6) Hohe Entscheidungsverantwortung | 0.746*** | 0.374*** | 0.625*** | 0.275** | 0.101 | |
| (7) Fehlende Regeln u. Handlungsschemata | -0.045 | 0.382*** | 0.232* | 0.242* | 0.621*** | -0.138 |

*Hinweise: *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$*

5. Schlussteil

5.1 Zusammenfassung der zentralen Ergebnisse

Der Beitrag hatte die literaturgestützte Systematisierung von Anforderungscharakteristika lernprozessbegleitender Diagnosen im laufenden Unterricht sowie deren strukturelle Validierung in empirischen Befragungsdaten von Lehrpersonen zum Ziel. Am Ansatz der arbeitspsychologischen Anforderungsanalyse orientiert (vgl. SCHULER 2014), wurden in einem ersten Schritt an Lehrpersonen gestellte situationsgebundene Anforderungen beim Fällen lernprozessdiagnostischer Urteile inhaltlich bestimmt (Abschnitt 2). Die ermittelten Dimensionen wurden in einem zweiten Schritt als eigenständige, jedoch nicht voneinander unabhängige Anforderungscharakteristika faktorenanalytisch bestätigt (Abschnitt 3). Sämtliche erfassten Anforderungsdimensionen werden gemessen an der zugrundeliegenden Ratingskala (1–5) seitens der Befragten überdurchschnittlich hoch eingeschätzt. Die vergleichsweise größten Anforderungen stellen aus ihrer Sicht die Vielfalt an zu verarbeitenden Informationen, die unsichere Informationsgrundlage sowie die hohe Entscheidungsverantwortung dar (Abschnitt 4).

5.2 Limitationen der Studie

Limitationen unserer Studie zeigen sich zum einen in Bezug auf die durchgeführte systematische Literatursichtung. Zwar orientierten wir uns bzgl. des methodischen Vorgehens am PRISMA Statement (LIBERATI et al. 2009) inklusive der Verwendung einschlägiger wissenschaftlicher Fachdatenbanken (z. B. ERIC, Fachportal Pädagogik) und definierten klare Auswahlkriterien für die aufzunehmenden Quellen, jedoch finden sich in der Auswahl mehr Quellen von deutschsprachigen als von internationalen Forschenden: 40 von 45 Quellen beziehen sich auf Ansätze aus dem nationalen Kontext, lediglich fünf Quellen beinhalten Informationen zu diagnosebezogenen Anforderungen, die über diesen Kontext hinausweisen (Tabelle 2). Eine mögliche Erklärung hierfür liegt darin, dass der Diagnostik-Begriff im internationalen (englischsprachigen) Forschungsraum möglicherweise anders konnotiert ist als dies im deutschsprachigen Raum der Fall ist und somit entsprechende Veröffentlichungen im Suchprozess nicht sichtbar wurden. Hinsichtlich der Aktualität der eingeschlossenen Quellen ist anzumerken, dass 38 von 45 Quellen ab dem Jahr 2010 veröffentlicht wurden, weshalb insgesamt von aktuellen Quellen auszugehen ist. Dennoch kann angenommen werden, dass die identifizierten und erfassten Anforderungscharakteristika für eine lernprozessbegleitende Diagnostik tendenziell als „zeitlos“ einzustufen sind. Auch aufgrund der disziplinübergreifenden Quellensuche ist es plausibel, dass die identifizierten Charakteristika als allgemeingültig einzustufen sind. Da es jedoch ergänzende domänenspezifische Besonderheiten bei der Diagnostik lernrelevanter Merkmale zu beachten gibt, bspw. im Rahmen der Erfassung von Sprachfähigkeiten in beruflichen Übergangsklassen, müssten diese im jeweiligen Anwendungsfall berücksichtigt werden.

Zum zweiten beziehen sich Limitationen auf die empirische Erfassung der betrachteten Anforderungsdimensionen. Bezogen auf die formulierten Items ist anzumerken, dass manche Items eher problematisierend formuliert wurden und so möglicherweise zu Demand-Effekten bei den Teilnehmenden führen könnten (z. B. ... sind nicht alle notwendigen Informationen für eine diagnostische Entscheidung verfügbar; ... fehlt der Lehrperson Zeit, um umfassende Informationen zu sammeln; siehe Anhang 2). Durch die betreffenden Items wird somit möglicherweise suggeriert, dass der Lehrberuf einen eher herausfordernden Beruf darstellt und sich das berufliche Handeln überwiegend unter schwierigen, einschränkenden Bedingungen vollzieht. In weiterführenden Studien wäre daher zu eruieren, die verwendeten Items alternativ durchweg positiv auszuformulieren oder noch stärker alternierende (positive wie auch negative) Items als bisher zu verwenden.

Wenngleich wir zum Zwecke der strukturellen Validierung lehrpersonenseitig wahrgenommener lernprozessdiagnostischer Anforderungen auf eine praktikable Stichprobe zurückgreifen konnten, weist diese einige (mögliche) Einschränkungen auf. Einerseits konnten wir nur auf eine verhältnismäßig kleine Stichprobe zurückgreifen. Andererseits ist nicht auszuschließen, dass unter den Teilnehmer/innen eine gewisse Selbstselektion stattfand, indem schwerpunktmäßig nur diejenigen Lehrpersonen an der Befragung teilgenommen haben, welche entweder besonders an der Thematik in-

teressiert sind oder welche diagnostische Situationen im Unterrichtsprozess besonders intensiv als Anforderung wahrnehmen. Weiterhin ist unsere Stichprobe nicht repräsentativ für Lehrkräfte in Deutschland, da der Großteil der Teilnehmer/innen zum Befragungszeitraum in Baden-Württemberg tätig war, was möglicherweise dem zugrundeliegenden Schneeballverfahren geschuldet ist. Daher haben wir auch darauf verzichtet, bundesland- sowie schulartspezifische Analysen durchzuführen. Auch zeichnet sich unsere Stichprobe mit einem Mittel von 34.23 Jahren durch ein relativ niedriges Durchschnittsalter sowie im Mittel mit 6.97 Dienstjahren durch eine verhältnismäßig geringe Berufserfahrung als Lehrperson aus, was die eingeschätzten Anforderungscharakteristika lernprozessbegleitender Diagnostik möglicherweise beeinflussen könnte. Weiterhin haben wir die statistische Modellprüfung an derselben Stichprobe vorgenommen, an der bereits der Schritt der finalen Itemauswahl vorgenommen wurde. D. h., die Auswahl der finalen 42 Items aus dem ursprünglichen Itempool von 151 könnte ggf. durch stichprobenspezifische Eigenheiten beeinflusst worden sein. In einem nachfolgenden Schritt müsste daher das Messmodell an einer neuen Stichprobe entsprechend überprüft werden.

Darüber hinaus ist die gewählte Methode, Anforderungscharakteristika lernprozessbegleitender Diagnostik zunächst mittels systematischem Literaturreview und anschließend mittels Fragebogenitems zu erfassen, abwägend zu betrachten. Geht man davon aus, dass empirische Studien zur lehrpersonenseitigen Wahrnehmung sich immer nur auf die bereits existierende, gelebte Praxis beziehen können und zwar aus Sicht der Betroffenen selbst, wären entsprechende Folgestudien erforderlich, welche beispielsweise mit beobachtungsbasierten Ansätzen arbeiten und damit einen differenzierteren und umfassenderen Blick auf die unterrichtliche „Diagnosepraxis“ ermöglichen.

5.3 Ausblick

Mit der unternommenen Systematisierung und empirischen Erfassung von situationsabhängigen Anforderungen, welche mit der Durchführung lernprozessdiagnostischer Tätigkeiten während des Unterrichts verbunden sind, verfolgten wir zunächst einen analytisch-beschreibenden Ansatz. Daher wären in Folgeuntersuchungen personenbezogene Variablen (z. B. fachliche und überfachliche Lehrendenkompetenzen, Maße der kognitiven Verarbeitungskapazität) sowie kontextbezogene Variablen (z. B. Klassenkompositionsmerkmale) in den Blick zu nehmen, welche ihrerseits Effekte auf die wahrgenommene Höhe der einzelnen Anforderungsdimensionen ausüben könnten. Forschungsseitig wären weitere Einsatzgebiete denkbar. So könnten bspw. in Experimentalstudien zu Urteilsprozessen von Lehrkräften in „simulierten Klassenräumen“ (SÜDKAMP/MÖLLER/POHLMANN 2013) die identifizierten Anforderungscharakteristiken als Situationsmerkmale explizit modelliert und unter der *ceteris-paribus*-Regel systematisch variiert werden, um deren mögliche differentielle oder auch additive Effekte auf die Urteilsakkuratheit zu prüfen. Dies eröffnet zum einen eine Validierungsmöglichkeit der hier vorgeschlagenen Klassifikation. Zum anderen ließen sich Aussagen darüber gewinnen, ob

alle Dimensionen gleichermaßen oder einzelne besonders markant oder aber insb. die Kumulation situativer Anforderungen die Diagnosegüte beeinflussen. Auch im Rahmen von Feld- bzw. Fragebogenstudien wären die Anforderungsdimensionen als relevante situationsbedingte Moderatoren der diagnostischen Urteilsgenauigkeit zu berücksichtigen. Angesichts der teils sehr hohen Streuungen, die empirisch erfasste Gütemaße diagnostischer Urteile in bislang vorliegenden Studien aufweisen, dürften situationsabhängig variierende Schwierigkeiten der Urteilsfindung zusätzlich zu den bisher berücksichtigten Moderatoren (darunter Testkenntnis und Berufserfahrung; z. B. SÜDKAMP/KAISER/MÖLLER 2012) zur Varianzaufklärung beitragen.

Da diagnostische Fähigkeiten gemeinhin als wesentliche Komponente von Lehrprofessionalität gelten, ihre Messung sich jedoch weitgehend auf die rechnerische Bestimmung abweichungsbasierter Maße der Urteilsakkuratheit beschränkt (zsf. SÜDKAMP/PRAETORIUS 2017), wäre zu überlegen, inwiefern der individuelle Umgang mit situativen Anforderungscharakteristika in ein erweitertes Konstrukt diagnostischer Kompetenz einfließen könnte. Dies würde einem Kompetenzverständnis, welches im Allgemeinen Fähigkeiten zur verantwortungsbewussten Bewältigung beruflicher Anforderungen umfasst, gerecht werden. Entsprechende Anknüpfungspunkte wären auch für die Ausbildung von Lehrer/innen (erste und zweite Phase) zu sehen, in welcher ein professioneller Umgang mit situativen Anforderungscharakteristika lernprozessbegleitender Diagnostik thematisiert werden sollte. Hierbei wäre es beispielsweise denkbar, Videovignetten einzusetzen, welche Verweise auf die beschriebenen Anforderungscharakteristika beinhalten. Diese könnten Reflexionsprozesse über typische situative Auftretensmuster, drohende Konsequenzen für Lehrkräfteurteile sowie mögliche Gegenmaßnahmen anstoßen (vgl. SEIFRIED/WUTTKE 2017).

Weiterhin wäre zu eruieren, inwieweit neue Technologien i. S. digitaler Diagnose- und Entscheidungsunterstützungssysteme Lehrpersonen bei der pädagogischen Prozessdiagnostik im konkreten Unterrichtsgeschehen situationsadäquat unterstützen können. Wie unsere berichteten Analysen verdeutlichen, sehen Lehrpersonen die prozessbegleitende Diagnostik als anforderungsreich an, da meist nicht alle relevanten Informationen wahrgenommen und verarbeitet werden können. Die Komplexität, Vielschichtigkeit, Mehrdimensionalität und Perspektivenabhängigkeit unterrichtlicher Situationen überfordern viele Lehrkräfte aufgrund natürlicher Kapazitätsgrenzen des menschlichen Informationsverarbeitungssystems zwangsläufig. Dies öffnet wiederum Urteilsverzerrungen Tür und Tor, da Heuristiken zum Einsatz kommen, um die Komplexität vorherrschender multidimensionaler Informationsgefüge zu reduzieren, was nicht selten mit Fehleinschätzungen einhergeht. Praktikable Möglichkeiten, Lehrpersonen von überfordernden mentalen Operationen der Informationssammlung, -verdichtung und -integration zu entlasten und sie bei der Prozessdiagnostik im konkreten Unterrichtsgeschehen zu unterstützen (ohne ihnen dabei nachgelagerte Entscheidungen über pädagogische oder didaktische Maßnahmen abzunehmen), werden daher in digitalen Technologien (z. B. Systeme zur Durchführung formativer Assessments) gesehen, auf welche zumindest Teile der damit verbundenen Aufgaben ausgelagert werden können (z. B. zsf. HASSELHORN/SCHNEIDER/TRAUTWEIN 2014). Eine aktuelle ka-

tegoriengeleitete Sichtung generischer sowie anwendungsspezifischer Funktionen von digitalen Systemen weist jedoch darauf hin, dass gegenwärtig insbesondere solche Systeme am Markt unterrepräsentiert erscheinen, welche Lehrpersonen durch Möglichkeiten der gezielten Analyse und Aufbereitung situationsbezogener Informationen beim Treffen begründeter pädagogisch-didaktischer Entscheidungen unterstützen (KÄRNER et al. 2019). Bis dato stellen auch Evaluationsstudien des Einsatzes entsprechender Unterstützungssysteme eher eine Seltenheit dar (z. B. FÖRSTER/SOUVIGNIER 2015). Hierbei wäre anzunehmen, dass durch entsprechende Softwareunterstützung bestimmte Anforderungen (insb. Vielfalt an Informationen, Informationsdefizite, unsichere Informationsgrundlage) reduziert werden. Ein konkretes Beispiel einer solchen Software stellt das *Teachers' Diagnostic Support System* (TDSS) dar, welches Lehrpersonen durch Möglichkeiten der gezielten Analyse und Aufbereitung von inhalts- und schüler/innenbezogenen Informationen beim Treffen begründeter pädagogisch-diagnostischer Entscheidungen unterstützen soll. Das als Client-Server-Modell implementierte TDSS erlaubt hierbei die Erfassung schüler/innen- sowie lernkontextbezogener Informationen und stellt der Lehrperson verschiedene Auswertungsmöglichkeiten und -darstellungen in Echtzeit zur Verfügung, darunter etwa individuelle Entwicklungsverläufe einzelner Schüler/innen oder unterrichtliche Erlebensqualitäten in Abhängigkeit von aktuell vorherrschenden Sozialformen (für eine detailliertere Beschreibung siehe KÄRNER/WARWAS/SCHUMANN 2020).

Insbesondere zeigen sich hier Potenziale für die Unterstützung des Lehrens und Lernens an beruflichen Schulen. Angesichts einer Implementierung von formativen Assessments im berufsbildenden Schulwesen zum Zwecke einer datenbasierten Unterrichtsdiagnostik, -durchführung und -entwicklung erscheint es sinnvoll und zielführend, zukünftige Entwicklungen standardisierter, domänenspezifischer Aufgaben- und Testformate voranzubringen, welche beispielsweise für spezifische Ausbildungsberufe und dort verortete fachspezifische Lernfelder nutzbar sind. Orientiert an den jeweiligen berufsspezifischen Lehrplänen wären somit perspektivisch Aufgaben- und Testformate sowie didaktisch-methodische Konzepte für die Realisierung von formativen Assessments zu entwickeln, welche als fester Bestandteil einer unterrichtsbegleitenden Lernverlaufdiagnostik mit dem Ziel einer bestmöglichen Lernförderung Einzug in den berufsschulischen Unterrichtsalltag finden könnten. Da sich entsprechende Aufgaben- und Testformate letztendlich an zu entwickelnden Kenntnissen und Fertigkeiten der Schüler/innen und somit an angestrebten Lern- und Bildungszielen orientieren (ACHTENHAGEN 2012), sollte insbesondere bei der Testentwicklung nicht zu kurz gegriffen werden. Ein mehrdimensionales Kompetenzverständnis, welches neben fachwissensbezogenen auch affektiv-motivationale sowie soziale Aspekte beinhaltet (z. B. RAUSCH/KÖGLER/SEIFRIED 2019; SEMBILL/RAUSCH/KÖGLER 2013), sowie ein partizipatives Bildungsverständnis, in welchem die Lernenden dahingehend ernst genommen werden, „an der Bestimmung der Zwecke, Inhalte und Realisierungsformen ihres Engagements wirklich und einflussreich mitzuwirken“ (HEID 1992, S. 106), erscheinen als wünschenswerte Zielgrößen entsprechender pädagogisch-technologischer Entwicklungen (zsf. KÄRNER et al., in Vorbereitung).

This material is under copyright. Any use outside of the narrow boundaries of copyright law is illegal and may be prosecuted.

This applies in particular to copies, translations, microfilming as well as storage and processing in electronic systems.

© Franz Steiner Verlag, Stuttgart 2020

Literatur

- ACHTENHAGEN, F. (2012). The curriculum-instruction-assessment triad. *Empirical research in vocational education and training*, 4(1), 5–25.
- ALTMANN, A. F. / NÜCKLES, M. (2017). Empirische Studie zu Qualitätsindikatoren für den diagnostischen Prozess. In A. SÜDKAMP / A.-K. PRAETORIUS (Hrsg.), *Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (S. 142–149). Münster: Waxmann.
- AUFSCHNAITER, C. V. / CAPPELL, J. / DÜBBELDE, G. / ENNEMOSER, M. / MAYER, J. / STIENSMEIER-PELSTER, J., et al. (2015). Diagnostische Kompetenz. Theoretische Überlegungen zu einem zentralen Konstrukt der Lehrerbildung. *Zeitschrift für Pädagogik*, 61(5), 738–758.
- BALOGH, E. / MILLER, B. T. / BALL, J. (2015). Improving diagnosis in health care. *Military Medicine*, 181(3), 183–185.
- BEAUDUCEL, A. / LEUE, A. (2014). *Psychologische Diagnostik*. Göttingen: Hogrefe.
- BECK, K. (1996). Die ‚Situation‘ als Bezugspunkt didaktischer Argumentation – Ein Beitrag zur Begriffspräzisierung. In W. SEYD / R. WITT (Hrsg.), *Situation, Handlung, Persönlichkeit. Kategorien wirtschaftspädagogischen Denkens* (S. 87–98). Hamburg: Feldhaus.
- BEHRMANN, L. / GLOGGER-FREY, I. (2017). Produkt- und Prozessindikatoren diagnostischer Kompetenz. In A. SÜDKAMP / A.-K. PRAETORIUS (Hrsg.), *Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (S. 134–142). Münster: Waxmann.
- BEHRMANN, L. / KAISER, J. (2017). Das Modell pädagogischer Diagnostik nach Ingenkamp und Lissmann. In A. SÜDKAMP / A.-K. PRAETORIUS (Hrsg.), *Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (S. 59–63). Münster: Waxmann.
- BEHRMANN, L. / VAN OPHUYSEN, S. (2017). Das Vier-Komponenten-Modell der Diagnosequalität. In A. SÜDKAMP / A.-K. PRAETORIUS (Hrsg.), *Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (S. 38–42). Münster: Waxmann.
- BERLINER, D. C. (1988). Implications of studies on expertise in pedagogy for teacher education and evaluation. In J. PFLEIDERER (Ed.), *New directions for teacher assessment* (pp. 39–68). Princeton: Educational Testing Service.
- BERLINER, D. C. (2001). Learning about and learning from expert teachers. *International Journal of Educational Research*, 35, 463–482.
- BERLINER, D. C. (2004). Describing the behavior and documenting the accomplishments of expert teachers. *Bulletin of Science, Technology & Society*, 24(3), 200–212.
- BÖHMER, M. / ENGLICH, B. / BÖHMER, I. (2017). Schülerbeurteilungen aus der Perspektive dualer Prozessmodelle der sozialen Urteilsbildung. In A. SÜDKAMP / A.-K. PRAETORIUS (Hrsg.), *Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (S. 50–54). Münster: Waxmann.
- BÖHMER, I. / HÖRSTERMANN, T. / GRÄSEL, C. / KROLAK-SCHWERDT, S. / GLOCK, S. (2015). Eine Analyse der Informationssuche bei der Erstellung der Übergangsempfehlung: Welcher Urteilsregel folgen Lehrkräfte? *Journal for Educational Research Online*, 7(2), 59–81.
- BRÜHWILER, C. (2017). Diagnostische und didaktische Kompetenz als Kern adaptiver Lehrkompetenz. In A. SÜDKAMP / A.-K. PRAETORIUS (Hrsg.), *Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (S. 123–134). Münster: Waxmann.
- BRÜHWILER, C. / BLATCHFORD, P. (2011). Effects of class size and adaptive teaching competency on classroom processes and academic outcome. *Learning and Instruction*, 21(1), 95–108.
- BÜHNER, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson.
- DÖRING, N. / BORTZ, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Berlin: Springer
- DOYLE, W. (1977). Learning the classroom environment: An ecological analysis. *Journal of Teacher Education*, 28(6), 51–55.

- FERNÁNDEZ-BALLESTEROS, R. / BRUYN, E. E. J. DE / GODOY, A. / HORNKE, L. F. / LAAK, T. J., VIZCARRO, C., et al. (2001). Guidelines for the Assessment Process (GAP): A proposal for discussion. *European Journal of Psychological Assessment*, 17(3), 187–200.
- FÖRSTER, N. / BÖHMER, I. (2017). Das Linsenmodell – Grundlagen und exemplarische Anwendungen in der pädagogisch-psychologischen Diagnostik. In A. SÜDKAMP / A.-K. PRAETORIUS (Hrsg.), *Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (S. 46–50). Münster: Waxmann.
- FÖRSTER, N. / KARST, K. (2017). Modelle diagnostischer Kompetenz: Gemeinsamkeiten und Unterschiede. In A. SÜDKAMP / A.-K. PRAETORIUS (Hrsg.), *Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (S. 63–65). Münster: Waxmann.
- FÖRSTER, N. / SOUVIGNIER, E. (2015). Effects of providing teachers with information about their students' reading progress. *School Psychology Review*, 44(1), 60–75.
- GLOGGER-FREY, I. / HERPPICH, S. (2017). Formative Diagnostik als Teilaspekt diagnostischer Kompetenz. In A. SÜDKAMP / A.-K. PRAETORIUS (Hrsg.), *Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (S. 42–46). Münster: Waxmann.
- HASCHER, T. (2011). Diagnostizieren in der Schule. In: A. BARTZ / C. KLOFT / J. FABIAN / S. HUBER / H. ROSENBUSCH / H. SASSENSCHIEDT (Hrsg.), *PraxisWissen SchulLeitung. Basiswissen und Arbeitshilfen zu den zentralen Handlungsfeldern von Schulleitung* (S. 1–8). Kronach: Carl Link.
- HASSELHORN, M. / SCHNEIDER, W. / TRAUTWEIN, U. (Hrsg.) (2014). *Lernverlaufsdiagnostik*. Göttingen: Hogrefe.
- HEID, H. (1992). Was „leistet“ das Leistungsprinzip? *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 88(2), 91–108.
- HEID, H. (2001). Situation als Konstrukt. Zur Kritik objektivistischer Situationsdefinitionen. *Schweizerische Zeitschrift für Bildungswissenschaften*, 23(3), 513–528.
- HEINRICHS, K. / REINKE, H. (Hrsg.) (2019). *Heterogenität in der beruflichen Bildung im Spannungsfeld zwischen Erziehung, Förderung und Fachausbildung*. Bielefeld: wbv.
- HELMKE, A. (08. Januar 2009). Die pädagogische Diagnostik führt ein Schattendasein. *Frankfurter Allgemeine Zeitung* (6), S. 8–9.
- HELMKE, A. / LENSKE, G. (2013). Unterrichtsdiagnostik als Voraussetzung für Unterrichtsentwicklung. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 31(2), 214–233.
- HERMKES, R. / MACH, H. / MINNAMEIER, G. (2018). Scaffolding von Problemlöseprozessen im Buchführungsunterricht. In E. WITTMANN / D. FROMMBERGER / B. ZIEGLER (Hrsg.), *Jahrbuch der berufs- und wirtschaftspädagogischen Forschung 2018* (S. 67–80). Budrich: Opladen.
- HERPPICH, S. / PRAETORIUS, A.-K. / HETMANEK, A. / GLOGGER-FREY, I. / UFER, S. / LEUTNER, D., et al. (2017). Ein Arbeitsmodell für die empirische Erforschung der diagnostischen Kompetenz von Lehrkräften. In A. SÜDKAMP / A.-K. PRAETORIUS (Hrsg.), *Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (S. 75–93). Münster: Waxmann.
- HESSE, I. / LATZKO, B. (2017). *Diagnostik für Lehrkräfte*. Opladen: Barbara Budrich.
- HETMANEK, A. / VAN GOG, T. (2017). Förderung von diagnostischer Kompetenz: Potenziale von Ansätzen aus der medizinischen Ausbildung. In A. SÜDKAMP / A.-K. PRAETORIUS (Hrsg.), *Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (S. 209–216). Münster: Waxmann.
- HOEFERT, H.-W. (1995). Rollen des Diagnostikers. In W. SARGES (Hrsg.), *Management-Diagnostik* (S. 78–87). Göttingen: Hogrefe.
- HOGUE, R. D. / COLADARCI, T. (1989). Teacher-Based Judgments of Academic Achievement: A Review of Literature. *Review of Educational Research*, 59(3), 297–313.
- HOSENFELD, I. / HELMKE, A. / SCHRADER, F.-W. (2002). Diagnostische Kompetenz: Unterrichts- und lernrelevante Schülermerkmale und deren Einschätzung durch Lehrkräfte in der Unter-

- richtsstudie SALVE. In M. PRENZEL / J. DOLL (Eds.), *Bildungsqualität von Schule: Schulische und außerschulische Bedingungen mathematischer, naturwissenschaftlicher und überfachlicher Kompetenzen* (S. 65–82). *Zeitschrift für Pädagogik*, Beiheft, 45. Weinheim: Beltz.
- HUGENER, I. / KRAMMER, K. (2010). Differenzierende Massnahmen zur Individualisierung des Unterrichts. In K. REUSSER / C. PAULI / M. WADIS (Hrsg.), *Unterrichtsgestaltung und Unterrichtsqualität. Ergebnisse einer internationalen und schweizerischen Videostudie zum Mathematikunterricht* (S. 91–106). Münster: Waxmann.
- INGENKAMP, K. / LISSMANN, U. (2008). *Lehrbuch der Pädagogischen Diagnostik*. Weinheim: Beltz.
- JANSEN, C. / MEYER, M. (2016). *Diagnostizieren im Dialog. Ein Leitfaden für den individualisierten Unterricht*. Weinheim: Beltz.
- KAISER, J. / PRAETORIUS, A.-K. / SÜDKAMP, A. / UFER, S. (2017). Die enge Verwobenheit von diagnostischem und pädagogischem Handeln als Herausforderung bei der Erfassung diagnostischer Kompetenz. In A. SÜDKAMP / A.-K. PRAETORIUS (Hrsg.), *Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (S. 114–123). Münster: Waxmann.
- KARING, C. / SEIDEL, T. (2017a). Ausblick zur Förderung diagnostischer Kompetenz. In A. SÜDKAMP / A.-K. PRAETORIUS (Hrsg.), *Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (S. 240–246). Münster: Waxmann.
- KARING, C. / SEIDEL, T. (2017b). Förderung diagnostischer Kompetenz. In A. SÜDKAMP / A.-K. PRAETORIUS (Hrsg.), *Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (S. 201–202). Münster: Waxmann.
- KÄRNER, T. (2017). A mixed-methods study of physiological reactivity to domain-specific problem solving: methodological perspectives for process-accompanying research in VET. *Empirical Research in Vocational Education and Training*, 9(10), 1–30.
- KÄRNER, T. / FENZL, H. / WARWAS, J. / SCHUMANN, S. (2019). Digitale Systeme zur Unterstützung von Lehrpersonen – Eine kategoriengeleitete Sichtung generischer und anwendungsspezifischer Systemfunktionen. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 115(1), 39–65.
- KÄRNER, T. / KELLER, T. / SCHNEIDER, A. / ALBANER, D. / SCHUMANN, S. (in Vorbereitung). Technologiegestützte formative Diagnostik: Sichtung und Kategorisierung verfügbarer Systeme zur Unterstützung der datengestützten Entscheidungsfindung von Lehrpersonen.
- KÄRNER, T. / WARWAS, J. / SCHUMANN, S. (2020). A Learning Analytics Approach to Address Heterogeneity in the Classroom: The Teachers' Diagnostic Support System. *Technology, Knowledge and Learning*. Online First Publication: <https://doi.org/10.1007/s10758-020-09448-4>
- KARST, K. (2017). Diagnostische Kompetenz und unterrichtliche Situationen. In A. SÜDKAMP / A.-K. PRAETORIUS (Hrsg.), *Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (S. 25–29). Münster: Waxmann.
- KARST, K. / HARTIG, J. / KAISER, J. / LIPOWSKY, F. (2017). Mehrebenenmodelle als Werkzeuge zur Analyse diagnostischer Kompetenz von Lehrkräften – ein lineares Mischmodell (LMM) und seine Anwendung in R. In A. SÜDKAMP / A.-K. PRAETORIUS (Hrsg.), *Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (S. 153–174). Münster: Waxmann.
- KARST, K. / KLUG, J. / UFER, S. (2017). Strukturierung diagnostischer Situationen im inner- und außerunterrichtlichen Handeln von Lehrkräften. In A. SÜDKAMP / A.-K. PRAETORIUS (Hrsg.), *Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (S. 102–114). Münster: Waxmann.
- KLUG, J. (2017). Ein Prozessmodell zur Diagnostik und Förderung von selbstreguliertem Lernen. In A. SÜDKAMP / A.-K. PRAETORIUS (Hrsg.), *Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (S. 54–59). Münster: Waxmann.
- KOLVENBACH, H. (2018). Diagnose-oder Befunderhebungsfehler? Ein medizintheoretisch fundiertes Abgrenzungsmodell. *Medizinrecht*, 37(2), 950–958.
- KRELL, M. (2017). Evaluating an instrument to measure mental load and mental effort considering different sources of validity evidence. *Cogent Education*, 4(1), 1–10.

- LEHMANN, R. H. / SEEBER, S. (Hrsg.) (2013). *ULME III. Untersuchung von Leistungen, Motivation und Einstellungen der Schülerinnen und Schüler in den Abschlussklassen der Berufsschulen*. Münster: Waxmann.
- LEICHNER, R. (1979). *Psychologische Diagnostik*. Weinheim: Beltz.
- LIBERATI, A. / ALTMAN, D. G. / TETZLAFF, J. / MULROW, C. / GÖTZSCHE, P. C. / IOANNIDIS, J. PA., et al. (2009). The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. *PLoS Medicine*, 6(7), 1–28.
- LORENZ, C. (2011). *Diagnostische Kompetenz von Grundschullehrkräften: strukturelle Aspekte und Bedingungen*. Dissertation. Universität Bamberg.
- MACHTS, N. / KAISER, J. / SCHMIDT, F. T. / MÖLLER, J. (2016). Accuracy of teachers' judgments of students' cognitive abilities: A meta-analysis. *Educational Research Review*, 19, 85–103.
- MUTHÉN, L. K. / MUTHÉN, B. O. (1998–2010). *Mplus User's Guide*. Sixth Edition. Muthén & Muthén, Los Angeles, CA. <https://www.statmodel.com/download/usersguide/Mplus%20Users%20Guide%20v6.pdf> [06.02.2019].
- ORTNER, T. (2010). Aktuelle Herausforderungen an die Psychologische Diagnostik. In K. D. KUBINGER / T. M. ORTNER (Hrsg.), *Psychologische Diagnostik in Fallbeispielen* (S. 18–27). Göttingen: Hogrefe.
- PAWLIK, K. (1976). Modell- und Praxisdimensionen psychologischer Diagnostik. In K. PAWLIK (Hrsg.), *Diagnose der Diagnostik: Beiträge zur Diskussion der psychologischen Diagnostik in der Verhaltensmodifikation* (S. 13–43). Stuttgart: Klett.
- PETSCH, C. / NORWIG, K. / NICKOLAUS, R. (2012). Individuelle Förderung in der beruflichen Grundbildung: Das berufsbezogene Strategietraining BEST. *Die berufsbildende Schule*, 64(11/12), 317–323.
- POTT, A. (2019). *Diagnostische Deutungen im Lernbereich Mathematik: Diagnostische Kompetenzen von Lehramtsstudierenden für sonderpädagogische Förderung und den Primarbereich*. Dissertation. Technische Universität Dortmund.
- PRAETORIUS, A.-K. / HETMANEK, A. / HERPPICH, S. / UFER, S. (2017). Herausforderungen bei der empirischen Erforschung diagnostischer Kompetenz. In A. SÜDKAMP / A.-K. PRAETORIUS (Hrsg.), *Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (S. 95–102). Münster: Waxmann.
- PRAETORIUS, A.-K. / KOCH, T. / SCHEUNPFLUG, A. / ZEINZ, H. / DRESEL, M. (2017). Identifying determinants of teachers' school-related motivations using a Bayesian cross-classified multi-level model. *Learning and Instruction* 52, 148–160.
- RAESSLER, S. / RUBIN, D. B. / SCHENKER, N. (2008). Incomplete Data: Diagnosis, Imputation, and Estimation. In E. D. DE LEEUW / J. J. HOX / D. A. DILLMAN (Eds.), *International Handbook of Survey Methodology* (pp. 370–386). New York: Lawrence Erlbaum.
- RAUSCH, A. / KÖGLER, K. / SEIFRIED, J. (2019). Validation of Embedded Experience Sampling (EES) for Measuring Non-cognitive Facets of Problem-Solving Competence in Scenario-Based Assessments. *Frontiers in Psychology*, 10, 1–16.
- REINKE, H. / HEINRICHS, K. (2019). Lernende mit sozial-emotional schwierigen Ausgangslagen an beruflichen Schulen – Herausforderungen für Lehrkräfte an der Schnittstelle von pädagogischer Diagnostik, Förderung und Beratung. In K. HEINRICHS / H. REINKE (Hrsg.), *Heterogenität in der beruflichen Bildung im Spannungsfeld zwischen Erziehung, Förderung und Fachausbildung* (S. 63–77). Bielefeld: wbv.
- RÖHRIG, F. (2019). *Entwicklung eines Fragebogens zur lehrpersonenseitigen Einschätzung diagnosebezogener Anforderungen*. Unveröffentlichte Masterarbeit, Fachbereich Wirtschaftswissenschaften, Universität Konstanz.
- ROST, D. H. / SCHERMER, F. J. (1989). Diagnostik des Leistungsangsterlebens. *Diagnostica*, 35, 287–314.

- RUBIN, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. Hoboken, NJ: Wiley-IEEE.
- SCHLADITZ, S. / DRÜGE, M. (2017). Psychotherapeutische Ausbildung. *Psychotherapeut*, 62(1), 47–53.
- SCHMIDT-ATZERT, L. / AMELANG, M. (2012). *Psychologische Diagnostik*. Heidelberg: Springer.
- SCHRADER, F.-W. (2010). Diagnostische Kompetenz von Eltern und Lehrern. In D. H. ROST (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 102–108). Weinheim/Basel: Beltz.
- SCHRADER, F.-W. (2013). Diagnostische Kompetenz von Lehrpersonen. *Beiträge zur Lehrerbildung*, 31(2), 154–165.
- SCHRADER, F.-W. (2017). Diagnostische Kompetenz von Lehrkräften – Anmerkungen zur Weiterentwicklung des Konstrukts. In A. SÜDKAMP / A.-K. PRAETORIUS (Hrsg.), *Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (S. 247–255). Münster: Waxmann.
- SCHRADER, F.-W. / HELMKE, A. (1987). Diagnostische Kompetenz von Lehrern: Komponenten und Wirkungen. *Empirische Pädagogik*, 1(1), 27–52.
- SCHULER, H. (2014). Arbeits- und Anforderungsanalyse. In H. SCHULER / U. P. KANNING (Hrsg.), *Lehrbuch der Personalpsychologie* (S. 62–97). Göttingen: Hogrefe.
- SEIFRIED J. / WUTTKE E. (2017). Der Einsatz von Videovignetten in der wirtschaftspädagogischen Forschung: Messung und Förderung von fachwissenschaftlichen und fachdidaktischen Kompetenzen angehender Lehrpersonen. In C. GRÄSEL / K. TREMPER (Hrsg.), *Entwicklung von Professionalität pädagogischen Personals* (S. 303–322). Wiesbaden: Springer VS.
- SEMBILL, D. (1987). Wirtschaftslehreunterricht: Einige Forschungsergebnisse zum Zusammenhang von Verbalurteilen und Handeln am Beispiel der Unterrichtseinheit „Kaufvertrag“. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 83(3), 213–232.
- SEMBILL, D. / RAUSCH, A. / KÖGLER, K. (2013). Non-Cognitive Facets of Competence. Theoretical Foundations and Implications for Measurement. In K. BECK / O. ZLATKIN-TROITSCHANSKAIA (Eds.), *From Diagnostics to Learning Success* (pp. 199–211). Rotterdam: Sense.
- SEMBILL, D. / WOLF, K. / WUTTKE, E. / SCHUMACHER, L. (2002). Self-Organized Learning in Vocational Education – Foundation, Implementation, and Evaluation. In K. BECK (Ed.), *Teaching-Learning Processes in Vocational Education. Foundations of Modern Training Programms* (pp. 267–295). Frankfurt a. M.: Lang.
- SEMMER, N. / ZAPF, D. / DUNCKEL, H. (1998). *ISTA. Instrument zur stressbezogenen Arbeitsanalyse*. Version 6.0. Bern: Eigenverlag.
- SHAVELSON, R. J. / YOUNG, D. B. / AYALA, C. C. / BRANDON, P. R. / FURTAK, E. M. / RUIZ-PRIMO, M. A., et al. (2008). On the Impact of Curriculum-Embedded Formative Assessment on Learning: A Collaboration between Curriculum and Assessment Developers. *Applied Measurement in Education*, 21(4), 295–314.
- SONNTAG, K. / FRIELING, E. / STEGMAIER, R. (2012). *Lehrbuch Arbeitspsychologie*. Bern: Huber.
- STANG, J. / URHAHNE, D. (2016). Wie gut schätzen Lehrkräfte Leistung, Konzentration, Arbeits- und Sozialverhalten ihrer Schülerinnen und Schüler ein? Ein Beitrag zur diagnostischen Kompetenz von Lehrkräften. *Psychologie in Erziehung und Unterricht*, 63(3), 204–219.
- STIEGLITZ, R.-D. (2008). *Diagnostik und Klassifikation in der Psychiatrie*. Stuttgart: Kohlhammer.
- SÜDKAMP, A. / KAISER, J. / MÖLLER, J. (2012). Accuracy of Teachers' Judgments of Students' Academic Achievement: A Meta-Analysis. *Journal of Educational Psychology*, 104(3), 743–762.
- SÜDKAMP, A. / KAISER, J. / MÖLLER, J. (2017). Ein heuristisches Modell der Akkuratheit diagnostischer Urteile von Lehrkräften. In A. SÜDKAMP / A.-K. PRAETORIUS (Hrsg.), *Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (S. 33–38). Münster: Waxmann.
- SÜDKAMP, A. / MÖLLER, J. / POHLMANN, B. (2013). Der Simulierte Klassenraum Ein Instrument zur Untersuchung. In E. M. LANKES (Hrsg.), *Pädagogische Professionalität als Gegenstand empirischer Forschung* (S. 87–97). Waxmann: Münster.

- SÜDKAMP, A. / PRAETORIUS, A.-K. (Hrsg.) (2017). *Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen*. Münster: Waxmann.
- TANGGAARD, L. / NIELSEN, K. / JØRGENSEN, C. H. (2015). Students' experiences of ability-based streaming in vocational education. *Education + Training*, 57(7), 723–737.
- TERHART, E. (2014). Umgang mit Heterogenität: Anforderungen an Professionalisierungsprozesse. *Lehren und Lernen*, 8(9), 7–12.
- VAN OPHUYSEN, S. / LINTORF, K. (2013). Pädagogische Diagnostik im Schulalltag. In S.-I. BEUTEL / W. BOS / R. PORSCH (Hrsg.), *Lernen in Vielfalt: Chance und Herausforderung für Schul- und Unterrichtsentwicklung* (S. 55–76). Münster: Waxmann.
- WARWAS, J. / KÄRNER, T. / GOLYSZNY, K. (2015). Diagnostische Sensibilität von Lehrpersonen im Berufsschulunterricht: Explorative Prozessanalysen mittels Continuous-State-Sampling. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 111(3), 437–454.
- WEIBER, R. / MÜHLHAUS, D. (2010). *Strukturgleichungsmodellierung: Eine anwendungsorientierte Einführung in die Kausalanalyse mit Hilfe von AMOS, SmartPLS und SPSS*. Heidelberg: Springer.
- WESTHOFF, K. / HORNKE, L. F. / WESTMEYER, H. (2003). Richtlinien für den diagnostischen Prozess. *reportPsychologie*, 28(9), 504–516.
- ZIEGLER, M. / BÜHNER, M. (2012). *Grundlagen der psychologischen Diagnostik*. Wiesbaden: Springer.



Anhang 1 – Auflistung der gesichteten Quellen

| Nr. | Autor(en) | Jahr | Sprache | Land | Beitragsart | Untersuchungsbereich |
|-----|-------------------------------------|------|----------|-------------|-----------------------|--|
| 1 | ALTMANN/NÜCKLES (2017) | 2017 | Deutsch | Deutschland | Beitrag in Sammelband | Fachspezifischer Schulbereich (Biologie) |
| 2 | BALOGH/MILLER/BALL (2015) | 2015 | Englisch | USA | Journal | Medizinischer Bereich |
| 3 | BEAUDUCEL/LEUE (2014) | 2014 | Deutsch | Deutschland | Monographie | Psychologischer Bereich |
| 4 | BEHRMANN/GLOGGER-FREY (2017) | 2017 | Deutsch | Deutschland | Beitrag in Sammelband | Fächerübergreifender Schulbereich |
| 5 | BEHRMANN/KAISER (2017) | 2017 | Deutsch | Deutschland | Beitrag in Sammelband | Fächerübergreifender Schulbereich |
| 6 | BEHRMANN/VAN OPHUYSEN (2017) | 2017 | Deutsch | Deutschland | Beitrag in Sammelband | Fächerübergreifender Schulbereich |
| 7 | BÖHMER/ENGLICH/BÖHMER (2017) | 2017 | Deutsch | Deutschland | Beitrag in Sammelband | Fächerübergreifender Schulbereich |
| 8 | BRÜHWILER (2017) | 2017 | Deutsch | Schweiz | Beitrag in Sammelband | Fächerübergreifender Schulbereich |
| 9 | DOYLE (1977) | 1977 | Englisch | USA | Journal | Fächerübergreifender Schulbereich |
| 10 | FERNÁNDEZ-BALLESTEROS et al. (2001) | 2001 | Englisch | Europa | Journal | Psychologischer Bereich |
| 11 | FÖRSTER/BÖHMER (2017) | 2017 | Deutsch | Deutschland | Beitrag in Sammelband | Fächerübergreifender Schulbereich |
| 12 | FÖRSTER/KARST (2017) | 2017 | Deutsch | Deutschland | Beitrag in Sammelband | Fächerübergreifender Schulbereich |
| 13 | GLOGGER-FREY/HERPPICH (2017) | 2017 | Deutsch | Deutschland | Beitrag in Sammelband | Fächerübergreifender Schulbereich |
| 14 | HASCHER (2011) | 2011 | Deutsch | Deutschland | Journal | Fächerübergreifender Schulbereich |
| 15 | HELMKE/LENSKE (2013) | 2013 | Deutsch | Deutschland | Journal | Fächerübergreifender Schulbereich |
| 16 | HELMKE (08. JANUAR 2009) | 2009 | Deutsch | Deutschland | Zeitungsartikel | Fächerübergreifender Schulbereich |
| 17 | HERPPICH et al. (2017) | 2017 | Deutsch | Deutschland | Beitrag in Sammelband | Fächerübergreifender Schulbereich |
| 18 | HESSE/LATZKO (2017) | 2017 | Deutsch | Deutschland | Beitrag in Sammelband | Fächerübergreifender Schulbereich |

| Nr. | Autor(en) | Jahr | Sprache | Land | Beitragsart | Untersuchungsbereich |
|-----|-------------------------------|------|----------|-------------|-----------------------|--|
| 19 | HETMANEK/VAN GOG (2017) | 2017 | Deutsch | Deutschland | Beitrag in Sammelband | Fächerübergreifender Schulbereich |
| 20 | HOEFERT (1995) | 1995 | Deutsch | Deutschland | Beitrag in Sammelband | Betriebswirtschaftlicher Bereich |
| 21 | JANSEN/MEYER (2016) | 2016 | Deutsch | Deutschland | Monographie | Fächerübergreifender Schulbereich |
| 22 | KAISER et al. (2017) | 2017 | Deutsch | Deutschland | Beitrag in Sammelband | Fächerübergreifender Schulbereich |
| 23 | KARING/SEIDEL (2017a) | 2017 | Deutsch | Deutschland | Beitrag in Sammelband | Fächerübergreifender Schulbereich |
| 24 | KARING/SEIDEL (2017b) | 2017 | Deutsch | Deutschland | Beitrag in Sammelband | Fächerübergreifender Schulbereich |
| 25 | KÄRNER (2017) | 2017 | Englisch | Deutschland | Journal | Betriebswirtschaftlicher Bereich |
| 26 | KÄRNER/WARWAS/SCHUMANN (2020) | 2020 | Englisch | Deutschland | Journal | Fächerübergreifender Schulbereich |
| 27 | KARST (2017) | 2017 | Deutsch | Deutschland | Beitrag in Sammelband | Fächerübergreifender Schulbereich |
| 28 | KARST et al. (2017) | 2017 | Deutsch | Deutschland | Beitrag in Sammelband | Fachspezifischer Schulbereich (Mathematik) |
| 29 | KARST/KLUG/UFER (2017) | 2017 | Deutsch | Deutschland | Beitrag in Sammelband | Fächerübergreifender Schulbereich |
| 30 | KLUG (2017) | 2017 | Deutsch | Deutschland | Beitrag in Sammelband | Fächerübergreifender Schulbereich |
| 31 | KOLVENBACH (2018) | 2018 | Deutsch | Deutschland | Journal | Medizinischer Bereich |
| 32 | KRELL (2017) | 2017 | Englisch | Deutschland | Journal | Fachspezifischer Schulbereich (Biologie) |
| 33 | LEICHNER (1979) | 1979 | Deutsch | Deutschland | Monographie | Psychologischer Bereich |
| 34 | LORENZ (2011) | 2011 | Deutsch | Deutschland | Dissertation | Fächerübergreifender Schulbereich |
| 35 | ORTNER (2010) | 2010 | Deutsch | Deutschland | Beitrag in Sammelband | Psychologischer Bereich |
| 36 | POTT (2019) | 2019 | Deutsch | Deutschland | Dissertation | Fachspezifischer Schulbereich (Mathematik) |
| 37 | PRAETORIUS/KOCH et al. (2017) | 2017 | Deutsch | Deutschland | Beitrag in Sammelband | Fächerübergreifender Schulbereich |

This material is under copyright. Any use outside of the narrow boundaries of copyright law is illegal and may be prosecuted.

This applies in particular to copies, translations, microfilming as well as storage and processing in electronic systems.

© Franz Steiner Verlag, Stuttgart 2020

| Nr. | Autor(en) | Jahr | Sprache | Land | Beitragsart | Untersuchungsbereich |
|-----|----------------------------------|------|---------|-------------|-----------------------|-----------------------------------|
| 38 | SÜDKAMP/KAISER/MÖLLER (2017) | 2017 | Deutsch | Deutschland | Beitrag in Sammelband | Fächerübergreifender Schulbereich |
| 39 | STIEGLITZ (2008) | 2008 | Deutsch | Deutschland | Monographie | Psychologischer Bereich |
| 40 | SCHLADITZ/DRÜGE (2017) | 2017 | Deutsch | Deutschland | Journal | Psychologischer Bereich |
| 41 | SCHMIDT-ATZERT/AMELANG (2012) | 2012 | Deutsch | Deutschland | Monographie | Psychologischer Bereich |
| 42 | SCHRADER (2017) | 2017 | Deutsch | Deutschland | Beitrag in Sammelband | Fächerübergreifender Schulbereich |
| 43 | WARWAS/KÄRNER/GOLYSZNY (2015) | 2015 | Deutsch | Deutschland | Journal | Fächerübergreifender Schulbereich |
| 44 | WESTHOFF/HORNKE/WESTMEYER (2003) | 2003 | Deutsch | Europa | Journal | Psychologischer Bereich |
| 45 | ZIEGLER/BÜHNER (2012) | 2012 | Deutsch | Deutschland | Monographie | Psychologischer Bereich |

This material is under copyright. Any use outside of the narrow boundaries of copyright law is illegal and may be prosecuted.

This applies in particular to copies, translations, microfilming as well as storage and processing in electronic systems.

© Franz Steiner Verlag, Stuttgart 2020

Anhang 2 – Fragebogen zur lehrpersonenseitigen Einschätzung diagnosebezogener Anforderungen (LEDA)

Bei der Einschätzung von Schülerbefindlichkeiten und -verhaltensweisen während des laufenden Unterrichts ...

Vielfalt an Informationen (Inf_VI)

- Inf_VI_1 ... muss die Lehrperson eine Vielzahl an unterschiedlichen Informationen in eine diagnostische Entscheidung einbeziehen, da Schülerinnen und Schüler sich je nach Unterrichtssituation anders verhalten (z. B. Klassengespräch versus Gruppenarbeit).
- Inf_VI_2 ... muss die Lehrperson eine diagnostische Entscheidung auf Basis von vielen verschiedenen Informationen treffen.
- Inf_VI_3 ... fließen vielfältige Informationen ein (z. B. Schülerpersönlichkeit, aktueller Unterrichtsinhalt).
- Inf_VI_4 ... müssen unterschiedliche Aspekte berücksichtigt werden (z. B. Schülerin/ Schüler, Thema, Verhalten unter den Schülern).
- Inf_VI_5 ... muss die Lehrkraft je nach verfolgten Lernzielen unterschiedliche Dinge besonders im Blick haben.
- Inf_VI_6 ... muss die Lehrkraft gleichzeitig mehrere Dinge im Blick haben (z. B. Klassenklima, Verhalten einzelner Schüler, Lärm).

Informationsdefizite (Inf_DEF)

- Inf_DEF_1 ... sind **nicht** alle notwendigen Informationen für eine diagnostische Entscheidung verfügbar.
- Inf_DEF_2 ... stehen der Lehrkraft **nicht** alle notwendigen Informationen über die Schülerin/den Schüler zur Verfügung.
- Inf_DEF_3 ... sind Vorinformationen über die Schülerinnen oder Schüler **nicht** vollständig verfügbar. (Beispiel: Informationen fehlen, weil die Lehrkraft diese nicht dokumentiert oder vergessen hat.)
- Inf_DEF_4 ... kommt die Lehrkraft **nicht** immer an alle notwendigen Informationen, weil Schülerrinnen und Schüler je nach Situation nicht bereit sind, Informationen über sich preis zu geben.

Unsichere Informationsgrundlage (Inf_UN)

- Inf_UN_1 ... muss die Lehrkraft eine diagnostische Entscheidung auf Grundlage von mehrdeutigen Informationen treffen.
- Inf_UN_2 ... muss die Lehrkraft mit irreführenden Informationen umgehen.
- Inf_UN_3 ... kann die Lehrkraft die vorhandenen Informationen unterschiedlich deuten.
- Inf_UN_4 ... können unterschiedliche Informationen zueinander in Widerspruch stehen.
- Inf_UN_5 ... muss sich die Lehrperson bewusst sein, dass eine diagnostische Entscheidung grundsätzlich falsch sein kann.
- Inf_UN_6 ... muss die Lehrperson irreführende Informationen erkennen (z. B. Fehlangaben durch Schülerinnen oder Schüler).

Bei der Einschätzung von Schülerbefindlichkeiten und -verhaltensweisen während des laufenden Unterrichts ...

Gefahr situativer Wahrnehmungs- und Urteilsverzerrungen (WZ)

| | |
|------|--|
| WZ_1 | ... ist die Lehrkraft möglicherweise voreingenommen, sodass Urteile nicht objektiv sein können. |
| WZ_2 | ... besteht die Gefahr, dass die Lehrkraft eigene Erwartungen selbst bestätigt, obwohl die Situation sich faktisch anders darstellt. |
| WZ_3 | ... wird die Lehrperson durch persönliche Einstellungen beeinflusst. |
| WZ_4 | ... treten systematische Beurteilungstendenzen auf (z. B. besonders milde oder strenge Urteile). |
| WZ_5 | ... ist es aufgrund von möglichen Vorurteilen schwierig, objektiv zu entscheiden. |

Zeit- und Handlungsdruck (ZeHa)

| | |
|--------|---|
| ZeHa_1 | ... fehlt der Lehrperson Zeit, um umfassende Informationen zu sammeln. |
| ZeHa_2 | ... muss die Lehrkraft sehr schnell eine diagnostische Entscheidung treffen. |
| ZeHa_3 | ... fehlen Rückzugsmöglichkeiten, um eigene Urteile zu überdenken. |
| ZeHa_4 | ... sind umfassende Einschätzungen nicht mit der Komplexität des Unterrichtsprozesses zu vereinen. |
| ZeHa_5 | ... ist aufgrund von Zeitdruck ein umfassendes Einschätzen der Schülerinnen und Schüler nicht möglich . |
| ZeHa_6 | ... muss die Lehrkraft die verfügbare Unterrichtszeit (z. B. 45 Minuten) beachten und kann daher nicht beliebig viele Informationen erheben. |
| ZeHa_7 | ... muss die Lehrperson eine diagnostische Entscheidung beiläufig treffen (z. B. aufgrund mangelnder Zeit, Unterrichtskomplexität). |

Hohe Entscheidungsverantwortung (HVA)

| | |
|-------|---|
| HVA_1 | ... muss die Lehrperson sowohl negative als auch positive Folgen einer diagnostischen Entscheidung für die Schülerin/den Schüler berücksichtigen. |
| HVA_2 | ... muss sich die Lehrperson bewusst sein, dass falsche Urteile negative Folgen haben können (z. B. Arbeitsverweigerung einer Schülerin). |
| HVA_3 | ... werden faire diagnostische Entscheidungen von der Lehrperson verlangt. |
| HVA_4 | ... werden genaue diagnostische Entscheidungen von der Lehrperson verlangt. |
| HVA_5 | ... werden nachvollziehbare diagnostische Entscheidungen von der Lehrperson verlangt. |
| HVA_6 | ... ist die Lehrkraft mit möglichen ethischen Problematiken konfrontiert. (Beispiel: Fehleinschätzungen der Lehrperson können eine schlechte Förderung eines Schülers zur Folge haben.) |
| HVA_7 | ... muss sich die Lehrkraft der Bedeutung einer objektiven Einschätzung bewusst sein. |
| HVA_8 | ... muss die Lehrkraft eine anspruchsvolle diagnostische Entscheidung treffen. |

Bei der Einschätzung von Schülerbefindlichkeiten und -verhaltensweisen während des laufenden Unterrichts ...

Fehlende Bewertungs- und Handlungsregeln (FHR)

- FHR_1 ... gibt es keine klar definierten Regeln, welche Informationen für eine diagnostische Entscheidung relevant sind (z. B. ob eine individuelle Hilfestellung nötig ist).
- FHR_2 ... gibt es keine einheitlichen Verfahren, sondern jede Lehrperson handelt nach ihrem eigenen Ermessen.
- FHR_3 ... liegen der Lehrkraft keine verbindlichen Bewertungskriterien vor (z. B. um Überforderung festzustellen).
- FHR_4 ... liegen keine Interpretationsstandards zur Bewertung diagnostischer Informationen vor. (Beispiel: Wann ist Mitarbeit einer Schülerin/eines Schülers noch als ausreichend zu bewerten?).
- FHR_5 ... gibt es keine eindeutigen Regeln für die Lehrperson, wie sie eine diagnostische Entscheidung zu treffen hat.
- FHR_6 ... gibt es keine festgelegten Regeln, unter welchem Maßstab die Schülerinnen und Schüler einzuschätzen sind (z. B. kriterienorientiert versus individuumsorientiert).
-

TOBIAS KÄRNER

Professur für Wirtschaftspädagogik, insbesondere Lehr- und Lernprozesse (560A),
 Institut für Bildung, Arbeit und Gesellschaft, Universität Hohenheim, Fruwirthstraße 47,
 D-70593 Stuttgart, E-Mail: tobias.kaerner@uni-hohenheim.de

FRANZISKA RÖHRIG

Seminar für Ausbildung und Fortbildung der Lehrkräfte Freiburg (Berufliche Schulen)
 Rieselfeldallee 1, D-79111 Freiburg im Breisgau

JULIA WARWAS

Professur für Wirtschaftspädagogik, insbesondere Theorie und Didaktik beruflicher
 Bildung (560B), Institut für Bildung, Arbeit und Gesellschaft, Universität Hohenheim,
 Fruwirthstraße 47, D-70593 Stuttgart, E-Mail: julia.warwas@uni-hohenheim.de



This material is under copyright. Any use outside of the narrow boundaries
 of copyright law is illegal and may be prosecuted.

This applies in particular to copies, translations, microfilming
 as well as storage and processing in electronic systems.

© Franz Steiner Verlag, Stuttgart 2020