



Das Schülerlabor als Ort authentischer Wissenschaftsvermittlung? Entwicklung und Validierung eines Fragebogens zur Erfassung der Authentizitätswahrnehmung der Wissenschaftsvermittlung im Schülerlabor

L. Finger¹ · V. van den Bogaert¹ · J. Fleischer¹ · J. Raimann² · K. Sommer³ · J. Wirth¹

Eingegangen: 17. Juni 2021 / Angenommen: 16. Februar 2022 / Online publiziert: 1. April 2022
© Der/die Autor(en) 2022

Zusammenfassung

Im vorliegenden Beitrag wird ein Fragebogen zur mehrdimensionalen Erfassung der Wahrnehmung von Authentizität in der Wissenschaftsvermittlung (FEWAW) vorgestellt. Die Entwicklung des Messinstruments mit den Authentizitätsdimensionen Vermittler, Ort, Vorgehen und Innovation erfolgte auf der Grundlage theoretischer Modelle zur Wahrnehmung von Authentizität der Wissenschaftsvermittlung. In einer Validierungsstudie mit $N=337$ Schülerinnen und Schülern im Alfried Krupp-Schülerlabor der Ruhr-Universität Bochum konnte die faktorielle Struktur mittels konfirmatorischer Faktorenanalysen bestätigt werden. Die internen Konsistenzen der vier Dimensionen sind zufriedenstellend (Cronbachs $\alpha \geq 0,81$). Außerdem konnte die Messinvarianz des Instruments für das Geschlecht sowie die Klassenstufe bestätigt werden. Die Zusammenhänge der Authentizitätsdimensionen mit Außenkriterien, wie z. B. epistemologischen Überzeugungen, liefern weitere Hinweise auf die Validität des Messinstruments. Insgesamt stellt der FEWAW ein geeignetes Instrument für die mehrdimensionale Erfassung der Authentizitätswahrnehmung in der Wissenschaftsvermittlung dar.

Schlüsselwörter Wissenschaftsvermittlung · Schülerlabor · Authentizität · Fragebogen · Validierung

Out-of-school Labs as Places for Authentic Science Education? Construction and Validation of a Questionnaire for Measuring the Perception of Authenticity of Science Education in Out-of-school Labs

Abstract

This paper presents a questionnaire for multidimensional assessment of the perception of authenticity in science education (FEWAW). The development of the questionnaire with the authenticity dimensions instructor, location, method and innovation was based on theoretical models of the perception of authenticity in science education. In a validation study with $N=337$ students in the out-of-school-lab at the Ruhr-University Bochum, the factor structure was confirmed by means of confirmatory factor analyses. The internal consistencies of all dimensions were satisfactory (Cronbach's $\alpha \geq 0.81$). Moreover, results proof measurement invariance for gender and class level. The correlations of the authenticity dimensions with external criteria such as epistemological beliefs provide further evidence for the validity of the subscales. Overall, the FEWAW is a suitable instrument for the multidimensional assessment of perceptions of authenticity in science education.

Keywords Science education · Out-of-school lab · Authenticity · Questionnaire · Validation study

✉ L. Finger
lena.finger@rub.de

¹ Lehrstuhl für Lehr-Lernforschung, Ruhr-Universität Bochum, Bochum, Deutschland

² Lehrstuhl für Sozialpsychologie, FernUniversität Hagen, Hagen, Deutschland

³ Didaktik der Chemie, Ruhr-Universität Bochum, Bochum, Deutschland

Einleitung

Die Förderung einer wissenschaftlichen Grundbildung der breiten Öffentlichkeit, als eine zentrale Bedingung für individuelle Entscheidungs-, Handlungs- und Diskursfähigkeit in modernen Gesellschaften, wird seit langem als ein zentrales Bildungsziel betrachtet (Fischer 1998; OECD 2013; Haupt und Hempelmann 2015). Die Vermittlung wissenschaftlicher Grundbildung in der Schule wird jedoch vereinzelt als unzureichend kritisiert (Gräber und Nentwig 2002; Euler und Weßnigk 2011), da u. a. wissenschaftspropädeutische Aufgaben, die Schülerinnen und Schüler im Unterrichtsalltag bearbeiten, selten den Kernmerkmalen echter wissenschaftlicher Arbeit entsprechen (Chinn und Malhotra 2002; Capps und Crawford 2013). Über solche „simple inquiry tasks“ (Chinn und Malhotra 2002) werden den Schülerinnen und Schülern im Unterricht nicht nur ein vereinfachtes Bild wissenschaftlicher Praktiken, sondern auch wenig elaborierte epistemologische Überzeugungen vermittelt (Bromme und Kienhues 2014). Um ein angemessenes Verständnis wissenschaftlicher Denk- und Arbeitsweisen aufzubauen und der Entstehung lernhinderlicher Überzeugungen entgegenzuwirken, ist vielmehr eine authentische Wissenschaftsvermittlung erforderlich (Bauer et al. 2017).

In den letzten zwei Jahrzehnten sind unterschiedliche Konzepte entstanden, die eine authentische Wissenschaftsvermittlung möglich machen sollen. Beispielsweise bieten Schülerlabore speziell für Schülerinnen und Schüler zugeschnittene Angebote im Sinne einer guten Ergänzung zum herkömmlichen Unterricht. Das zentrale Anliegen der Schülerlabore ist es, „Wissenschaft durch erfahrungsbasierte Zugänge erlebbar zu machen“ (Euler et al. 2015, S. 761) und wissenschaftliche Arbeitsweisen im Rahmen aktiver Auseinandersetzung mit dem Lerngegenstand zu vermitteln. Dabei wird die authentische Gestaltung der Lernumgebung nach konstruktivistischen Prinzipien als wesentlicher Vorteil gegenüber traditionellem Unterricht mit seinem tendenziell transmissiven Verständnis von Lernen gewertet (Euler und Weßnigk 2011; Euler et al. 2015). Scharfenberg und Bogner (2015) schreiben der Authentizität der Lernumgebung Schülerlabor sogar eine direkte Wirkung auf die Entwicklung des aktuellen Interesses (vgl. Krapp 2001) von Schülerinnen und Schülern zu.

In der Schülerlaborforschung besteht Einigkeit über die Bedeutung und die Notwendigkeit von authentischen Lernumgebungen für gelingende Wissenschaftsvermittlung und Interessenförderung. Diese programmatische Wirkung objektiv vorliegender Authentizität kann jedoch nur erfolgen, wenn die Lernumgebung Schülerlabor von den Schülerinnen und Schülern auch subjektiv als authentisch wissenschaftlich wahrgenommen wird (Sommer et al. 2018). Für empirische Untersuchungen von Zusammenhängen der Authentizität von Lernumgebungen mit Personenmerkmalen,

wie Interesse oder Fachwissen, reicht es demzufolge nicht aus, die Authentizität von Lernumgebungen objektiv anzunehmen. Vielmehr bedarf es der Möglichkeit, die subjektive Wahrnehmung der Authentizität zu erfassen und mit Personenmerkmalen in Relation zu setzen (Gulikers et al. 2006; Betz 2018).

Derzeit liegt jedoch kein Instrument vor, das eine differenzierte Betrachtung von Authentizitätswahrnehmung in der Wissenschaftsvermittlung und damit die systematische Untersuchung der Zusammenhänge von subjektiv authentischen Lernumgebungen und Personenmerkmalen erlaubt. Aus diesem Grund gibt es bislang auch nur wenige empirische Studien, die die subjektive Wahrnehmung von Authentizität berücksichtigen. Da die Bedeutung von wahrgenommener Authentizität für die Gestaltung erfolgreicher Lehr-Lernkontexte jedoch einer weiteren empirischen Überprüfung bedarf (Sommer et al. 2018), leistet die theoriegeleitete Entwicklung und Validierung eines Fragebogens zur Messung von Authentizitätswahrnehmung dazu einen wichtigen Beitrag.

Theoretischer Hintergrund

Authentizität als Merkmal von Lernumgebungen

Die Definitionen von Authentizität als Merkmal einer Lernumgebung sind vielfältig. Scharfenberg und Bogner (2015) kritisieren gar das Fehlen einer allgemein anerkannten Definition dieses zentralen Merkmals konstruktivistischer Lernumgebungen.

In diesem Beitrag wird in Anlehnung an Honebein et al. (1993) unter Authentizität die Annäherung einer Lernsituation an reale Anwendungssituationen verstanden. Dabei kann Authentizität als Kontinuum betrachtet werden, das sich zwischen zwei Endpunkten bewegt. Dies ist zum einen die echte Anwendungssituation bzw. die Realität und zum anderen eine stark didaktisierte Lernumgebung bzw. eine stark vereinfachte Darstellung der Realität (Betz et al. 2016; s. a. Chinn und Malhotra 2002; Sommer et al. 2018).

Anker-Hansen und Andréa (2019) nehmen auf Grundlage eines systematischen Literaturreviews eine klare Differenzierung verschiedener Anwendungssituationen vor. Sie unterscheiden zwischen Anwendungssituationen, die für Schülerinnen und Schüler alltäglich und nützlich sind einerseits und Anwendungssituationen im professionellen Kontext andererseits. Mit Blick auf den professionellen bzw. beruflich-wissenschaftlichen Kontext gilt eine Lernumgebung als authentisch, wenn sie den Alltag und die Arbeitsweisen einer Wissenschaftlerin oder eines Wissenschaftlers realistisch abbildet und eine spezifische Komplexität aufweist, die sich u. a. aus einer hohen Informationsdichte in spezifischen Wissensbereichen und den

vielfältigen Tätigkeitsbereichen im beruflichen Alltag ergibt (Lee und Songer 2003; van Vorst et al. 2015). Dem vorliegenden Beitrag liegt dieses letztere Verständnis von Authentizität im professionellen Kontext zugrunde und ist von der Anpassung einer Lernumgebung oder eines Lerninhalts an die alltägliche Lebenswelt der Schülerinnen und Schüler abzugrenzen.

Authentizität ist ein relationales Konzept (Honebein et al. 1993). Das bedeutet, dass die Authentizität eines Gegenstands bzw. einer Tätigkeit nur in Bezug auf dessen bzw. deren Ähnlichkeit zu etwas Anderem beurteilt werden kann. Dabei beeinflussen individuelle Repräsentationen des Vergleichsgegenstands die subjektive Wahrnehmung von Authentizität. In authentischen Lernsituationen (in unserem Sinne) werden beruflich-wissenschaftliche Kontexte inszeniert, um Schülerinnen und Schülern einen Einblick in diese Bereiche zu geben und darüber hinaus den Wissenstransfer zu fördern (Mandl et al. 1995). Die subjektive Wahrnehmung der Authentizität der Lernumgebung wird nun dadurch beeinflusst, welche individuellen Repräsentationen dieses beruflich-wissenschaftlichen Kontexts bei den einzelnen Schülerinnen und Schülern vorliegen, d. h. wie hoch der Grad der Vertrautheit mit diesem spezifischen beruflich-wissenschaftlichen Kontext ist (Habig et al. 2018).

Theoretische Modellierung von Authentizität der Wissenschaftsvermittlung

Bislang gibt es zahlreiche Ansätze, die Authentizität von Lernumgebungen inhaltlich genauer zu differenzieren. Trotz verschiedener Konzeptualisierungen besteht Konsens über die Mehrdimensionalität von Authentizität. Die Lernumgebung wird stets in verschiedene Bereiche klas-

sifiziert, die mehr oder weniger authentisch sein können. So unterscheidet Scharfenberg (2005) im Kontext der Wissenschaftsvermittlung Lernorte, Vermittler, Materialien, Methoden sowie Probleme als verschiedene Dimensionen von Authentizität. Herrington und Herrington (2006) unterscheiden neun Charakteristika authentischer Lernumgebungen, die stärker die Aktivitäten der Schülerinnen und Schüler während des Lernens beschreiben, wie etwa Reflexion und gemeinsame Wissenskonstruktion. Das Modell der Authentizität der Wissenschaftsvermittlung (Betz et al. 2016) betrachtet Authentizität ebenfalls als mehrdimensionales Konstrukt. Es geht jedoch über die bloße Klassifikation von Merkmalen der Lernumgebung hinaus und bindet die Authentizität der Lernumgebung in ein Wirkungsmodell ein. Die subjektive Wahrnehmung von Authentizität steht im Zentrum dieses Wirkungsmodells, das im Folgenden vorgestellt werden soll.

Das mehrdimensionale Modell der Authentizität der Wissenschaftsvermittlung (Betz et al. 2016) beschreibt die Entwicklung eines subjektiven Authentizitätsempfindens auf der Grundlage von Authentisierungsprozessen, die durch die Interaktion von individuellen Personenmerkmalen und Merkmalen der Lernumgebung entstehen. Authentisierung bedeutet, dass die Schülerinnen und Schüler eine objektiv (mehr oder weniger) wissenschaftlich authentisch gestaltete Lernumgebung immer vor dem Hintergrund individueller Überzeugungen, Interessen, Fähigkeiten und Wissen wahrnehmen. Die daraus resultierende Wahrnehmung von Authentizität kann Effekte auf Interesse, Motivation, Überzeugungen sowie Wissen und Kompetenzen der Schülerinnen und Schüler haben (vgl. Abb. 1).

Im Modell der Authentizität der Wissenschaftsvermittlung werden sieben Merkmale der Lernumgebung unter-

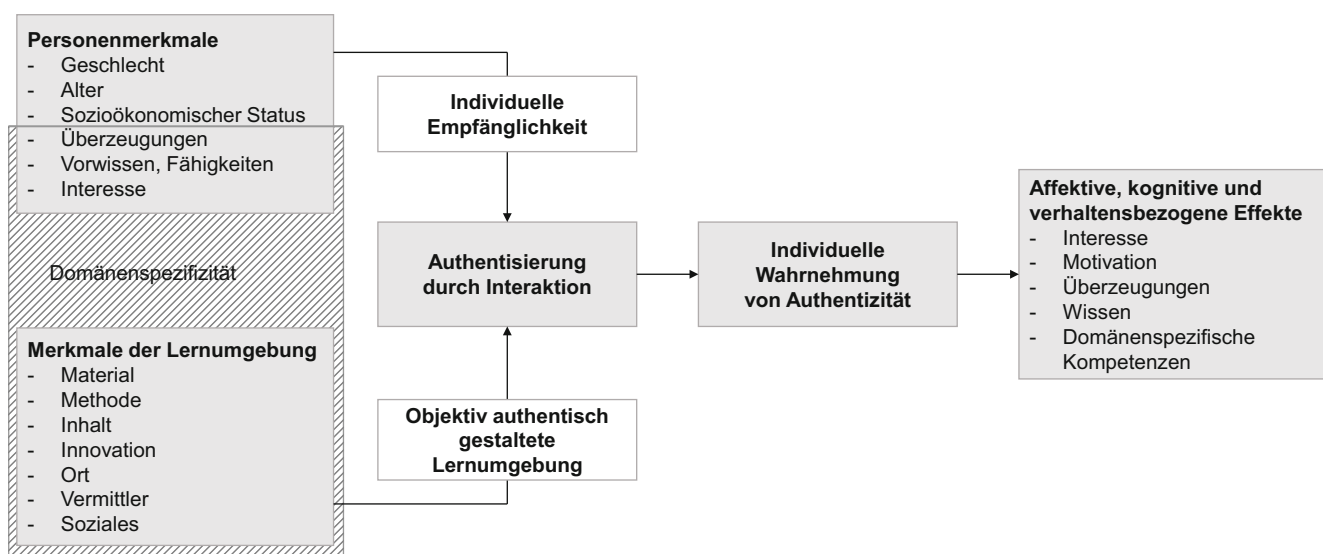


Abb. 1 Modell der Authentizität der Wissenschaftsvermittlung (Betz et al. 2016) mit Ergänzungen (schraffierter Kasten) durch Nachtigall et al. (2018a)

schieden, die einen Spielraum zur Gestaltung von objektiv authentischen Lernumgebungen bieten. Zu diesen Merkmalen zählen das Material, die Methoden, der Inhalt, der Innovationsgehalt, der Lernort, der Vermittler sowie soziale Interaktionen. Vereinzelt Studien deuten darauf hin, dass bspw. authentische Materialien einen positiven Effekt auf das Interesse (Peacock 1997) und epistemologische Überzeugungen (Mierwald et al. 2018) von Schülerinnen und Schülern haben können. Auch die Variation des Lernorts (Schülerlabor vs. Schule) zeigt bei Schülerinnen und Schülern eine unterschiedlich stark ausgeprägte Wahrnehmung von Authentizität zugunsten des Schülerlabors, die die Entwicklung von situationalem Interesse beeinflusst (Betz 2018; Schüttler et al. 2021). Um die positiven Wirkungen authentischer Lernumgebungen wie eine Förderung von domänenspezifischen Kompetenzen oder die Förderung von für den Lernerfolg relevanten Variablen wie Motivation und Interesse (Mandl et al. 1995; Betz et al. 2016) zu entfalten, reicht eine objektiv authentisch gestaltete Lernumgebung jedoch nicht aus. Die Authentizität der Lernumgebung muss von den Schülerinnen und Schülern als solche wahrgenommen werden (vgl. Relationalität der Authentizität). Personenmerkmale, wie das Geschlecht, das Alter oder epistemologische Überzeugungen, sind die Grundlage für eine individuelle Empfänglichkeit gegenüber der Lernumgebung. So haben Schülerinnen und Schüler womöglich eine stereotype Vorstellung von Wissenschaft (vgl. Hagenkötter et al. 2021), die die Wahrnehmung von Authentizität beeinflusst. Zentral bleibt die Frage, ob und in welchem Ausmaß die Schülerinnen und Schüler die Situation bzw. die Lernumgebung als authentisch, d. h. als echte Wissenschaft, wahrnehmen. Durch die Fokussierung auf die subjektive Wahrnehmung von Authentizität kann die Wirkung der didaktischen Überlegungen überprüft werden. Werden Fehlkonzepte bei den Schülerinnen und Schülern deutlich, ist dies bspw. ein Zeichen für Aufklärungsbedarf über tatsächliche Wissenschaftspraktiken. Aus den Annahmen des Wirkungsmodells folgt insgesamt, dass die subjektive Wahrnehmung von Authentizität und die auf sie wirkenden Faktoren zukünftig stärker in den Blick genommen werden müssen. Zudem sind die objektive Authentizität einer Lernumgebung genauso wie ihre subjektive Wahrnehmung

– in Übereinstimmung mit bisherigen Konzeptualisierungen – als mehrdimensionale Konstrukte zu betrachten.

Im Rahmen einer empirischen Überprüfung des Modells haben Nachtigall et al. (2018a) eine Adaptation vorgenommen (Abb. 1, schraffierter Kasten links). Nicht nur die individuellen Personenmerkmale, wie z. B. Vorwissen und Erfahrungen der Schülerinnen und Schüler, sind domänenspezifisch, sondern auch die Merkmale der Lernumgebung. In den vielfältigen Fachgebieten können die verschiedenen Gestaltungsmerkmale der Lernumgebung einen unterschiedlich starken Einfluss auf die Wahrnehmung von Authentizität haben. Durch die spezifische Gestaltung einer Laborumgebung in naturwissenschaftlichen Projekten könnte bspw. der Ort den größten Einfluss auf die Wahrnehmung der Authentizität ausüben. Es ist bisher jedoch nicht geklärt, ob in den verschiedenen Domänen bestimmte Gestaltungsmerkmale eine stärkere Rolle in der Wahrnehmung von Authentizität einnehmen als andere. Durch die Entwicklung eines Instruments, das die Wahrnehmung von Authentizität mehrdimensional erhebt, kann in zukünftiger Forschung unter der Berücksichtigung möglicher Interaktionen von Gestaltungs- und Personenmerkmalen u. a. der Frage nach der Einflussgröße der einzelnen Merkmale in unterschiedlichen Domänen nachgegangen werden.

Messung von Authentizitätswahrnehmung in Schülerlaboren

In Studien, die die programmatische Wirkung von Schülerlaboren auf Personenmerkmale der Schülerinnen und Schüler, wie etwa Interesse, Motivation und Fachwissen, systematisch untersuchen (Brandt 2005; Scharfenberg 2005; Guderian 2007; Brandt et al. 2008; Itzek-Greulich et al. 2015, 2017), wird die Authentizität der Schülerlabore zunehmend als wichtige unabhängige Variable betrachtet, die die Wirkung des Schülerlabors auf Personenmerkmale der Schülerinnen und Schüler maßgeblich beeinflussen kann (Engeln 2004; Pawek 2009; Damerau 2012; Betz 2018; Mierwald et al. 2018; Schüttler et al. 2021). In diesen Studien wird eine Messung der subjektiven Authentizitätswahrnehmung durchgeführt, da sich aus der Relationalität des Authentizitätskonstrukts (Honebein et al. 1993) die Notwendigkeit zur

Tab. 1 Übersicht bisheriger Messinstrumente zur Wahrnehmung von Authentizität der Wissenschaftsvermittlung

Quelle	Begriffsverständnis Authentizität	Stichprobe (N)	Itemanzahl (Cronbachs α)
Engeln (2004)	Qualität des Bezugs zur Wissenschaft	334	3 (0,66)
Glowinski (2007)	Einblick in die reale Forschungswelt	458	4 (0,66)
Pawek (2009)	Möglichst große Annäherung an die reale Forschungspraxis	730	4 (0,80)
Damerau (2012)	Möglichst große Annäherung an die reale Forschungswelt	631	3 (0,81)
Betz (2018)	Realistischer Einblick in die Wissenschaft	153	5/9 (je 0,74)
Nachtigall et al. (2018b)	Lernaktivität im Schülerlabor	121	8 (0,82)
Schüttler et al. (2021)	Verwendete Laborgeräte	148	4 (0,85)

Messung subjektiv wahrgenommener Authentizität ableiten lässt.

Tab. 1 fasst zentrale Merkmale bisheriger Instrumente zur Messung von Authentizitätswahrnehmung zusammen. Eine erste Operationalisierung stammt von Engeln (2004), die Authentizität als eindimensionales Konstrukt versteht. In den Untersuchungen von Glowinski (2007), Pawek (2009) und Damerau (2012) wird Authentizität ebenfalls als eindimensionales Konstrukt erfasst. Charakteristisch für die Untersuchungen von Nachtigall et al. (2018b), Betz (2018) sowie Schüttler et al. (2021) ist, dass Authentizität der Wissenschaftsvermittlung genauer differenziert wird und sich der Umfang der Messinstrumente z. T. deutlich erhöht. Nachtigall et al. (2018b) sowie Schüttler et al. (2021) fokussieren aber lediglich einzelne Facetten von authentischen Lernumgebungen. Betz (2018) entwickelt die Items des Messinstruments zwar gemäß des Modells der Authentizität der Wissenschaftsvermittlung (Betz et al. 2016), fasst die Items der vielfältigen Dimensionen jedoch zu nicht mehr als zwei Subskalen zusammen.

Erste Untersuchungen zur Rolle der Authentizität im Schülerlabor belegen – trotz unterschiedlicher und meist wenig ausdifferenzierter Operationalisierung des Authentizitätsbegriffs – kurzfristige Effekte subjektiv authentischer Lernumgebungen auf das Interesse, das Fähigkeitsselbstkonzept und epistemologische Überzeugungen von Schülerinnen und Schülern (Engeln 2004; Glowinski 2007; Pawek 2009; Damerau 2012; Betz 2018; Mierwald et al. 2018). Da bislang kein Instrument vorliegt, das eine mehrdimensionale und hinreichend umfassende Betrachtung von subjektiver Authentizitätswahrnehmung erlaubt, sind systematische Untersuchungen des Zusammenhangs von authentischen Lernumgebungen und Personenmerkmalen in der Wissenschaftsvermittlung kaum zu finden (Schüttler et al. 2021). Um diese Lücke zu schließen, ist es das Ziel der vorliegenden Arbeit, einen mehrdimensionalen *Fragebogen zur Erfassung der Wahrnehmung von Authentizität in der Wissenschaftsvermittlung* (FEWAW) zu entwickeln und seine psychometrische Qualität zu evaluieren. Dadurch sollen weitere empirische Überprüfungen der Wirkungszusammenhänge von objektiver Authentizität der Wissenschaftsvermittlung, subjektiver Authentizitätswahrnehmung und Personenmerkmalen, wie Interesse oder Fachwissen, in Lehr-Lernkontexten möglich werden (Sommer et al. 2018).

Entwicklung eines Testinstruments zur Messung subjektiver Authentizitätswahrnehmung

Betz et al. (2016) haben mit dem Modell der Authentizität der Wissenschaftsvermittlung bereits einen wichti-

gen Beitrag zur konzeptionellen Ausdifferenzierung des Begriffs der Authentizität geleistet und die Grundlage für eine angemessene Operationalisierung des Konstrukts geschaffen. Dieses Modell stellt die theoretische Grundlage für die Entwicklung des *Fragebogens zur Erfassung der Wahrnehmung von Authentizität in der Wissenschaftsvermittlung* (FEWAW) dar.

Differenzierung der Subdimensionen

Das Konstrukt Authentizitätswahrnehmung in der Wissenschaftsvermittlung wird über fünf Dimensionen operationalisiert, denen im Kontext des Schülerlabors eine besonders zentrale Rolle zukommt. Unabhängig von den unterschiedlichen Kategorien, in die Schülerlabore eingeordnet werden können (z. B. klassisches Schülerlabor oder Schülerlabor mit Berufsorientierung), heben bereits Haupt und Kollegen (2013) die Bedeutung des Orts bei diesen außerschulischen Lernangeboten hervor. Außerdem werden Material, Methode sowie Vermittler zu den allgemeinen Schülerlabor-Kriterien gezählt. Auch Mierwald (2020) macht auf die zentrale Rolle von fachspezifischen Materialien im Schülerlabor aufmerksam. Scharfenberg (2005) betrachtet Authentizität im Kontext der Wissenschaftsvermittlung ebenfalls als mehrdimensionales Konstrukt und differenziert authentische Lernorte, Vermittler, Materialien und Methoden sowie authentische Probleme. Authentizität wird erzielt durch die Bearbeitung realer Probleme, mit denen Wissenschaftlerinnen und Wissenschaftler konfrontiert sind (Edelson et al. 1999), unter Einbezug angemessener wissenschaftlicher Methoden (Lee und Songer 2003).

Vor diesem Hintergrund werden im FEWAW die folgenden Dimensionen differenziert: *Ort, Vermittler, Material, Methode* und *Innovation* (im Sinne von wissenschaftlich authentischen Problemen). Die Itemformulierungen ergeben sich über alle Dimensionen hinweg aus dem Verständnis von Authentizität als Kontinuum mit den Endpunkten stark didaktisiert versus real. Die Items bilden einen dieser Endpunkte des Kontinuums ab (besonders starke Annäherung der Lernumgebung an die beruflich-wissenschaftliche Realität), zu denen die Schülerinnen und Schüler Zustimmung oder Ablehnung äußern können. Die Einschätzung dieser objektiven Merkmale der Lernumgebung durch die Schülerinnen und Schüler gibt Aufschluss über die subjektive Wahrnehmung der Authentizität. Die Items sind insofern offen formuliert, als dass sie stereotype Vorstellungen von Schülerinnen und Schülern über Wissenschaft und wissenschaftliches Personal (Hagenkötter et al. 2021) nicht direkt adressieren. Gleichwohl kann die subjektive Wahrnehmung der Situation, die erfasst wird, von Stereotypen geprägt sein (vgl. Fehlkonzepte von Wissenschaft). Die Itemformulierungen basieren auf den folgenden Definitionen:

Ort Der Ort bildet den das Lernsetting räumlich umgebenden Rahmen. Es wird abgebildet, wie authentisch die besuchten Gebäude, Räume und Orte durch die Schülerinnen und Schüler wahrgenommen werden. Dazu zählen auch Orte außerhalb eines Schülerlabors, wie ein Botanischer Garten, eine Ausgrabungsstätte oder ein Forschungsschiff. Als authentisch gilt ein Ort dann, wenn er hauptsächlich für wissenschaftliche Zwecke hergerichtet und ausgestattet wurde und wenn dort Forschung betrieben wird oder zumindest werden kann (Sommer et al. 2020). Instrumente und Materialien (wie z.B. nicht genutzte Bücher in einer Bibliothek oder nicht genutzte Messinstrumente) tragen zur wissenschaftlichen Ausstattung eines Raumes bei.

Vermittler Vermittler sind Personen, die die Rolle der Lehrenden übernehmen. Diese Dimension bildet ab, wie authentisch diese Personen durch die Schülerinnen und Schüler wahrgenommen werden. Als authentisch gelten Vermittler dann, wenn sie in ihrem beruflichen Alltag als Wissenschaftlerin oder Wissenschaftler agieren und durch ihre Person einen direkten Zugang zur wissenschaftlichen Forschung der jeweiligen Disziplin herstellen. Euler und Weßnigk (2011) weisen in diesem Zusammenhang auf die Bedeutung von Rollenmodellen hin.

Material Diese Dimension bildet ab, wie authentisch das eingesetzte Material durch die Schülerinnen und Schüler wahrgenommen wird. Als authentisch gilt ein Material dann, wenn es nicht (oder nur sehr wenig) didaktisch aufbereitet ist (Sommer et al. 2018). Unter Materialien werden alle Dinge verstanden, mit denen Schülerinnen und Schüler arbeiten, mit Ausnahme der Instruktionsmaterialien und Anleitungen. Dazu gehören zum einen alle Untersuchungsgegenstände, wie z.B. Substanzen, Bodenproben, Originaltexte oder Zeitzeugeninterviews. Zum anderen fallen darunter auch Geräte und Instrumente, mit denen etwas untersucht wird, wie z.B. Mikroskope, Indikatorlösungen oder Papier und Bleistift.

Methode Frage- bzw. Problemstellungen können mit spezifischen wissenschaftlichen Methoden bearbeitet werden. Entsprechend bildet diese Dimension ab, wie wissenschaftlich die Schülerinnen und Schüler ihr Arbeiten und Vorgehen wahrnehmen. Wenngleich innerhalb wissenschaftlicher Disziplinen ein Methodenpluralismus besteht, wird im Rahmen von Schülerlaborbesuchen in den meisten Fällen auf nur eine dieser wissenschaftlichen Methoden fokussiert. Als authentisch gilt eine Methode dann, wenn sie einem wissenschaftlichen Vorgehen der jeweiligen Disziplin entspricht. Dazu gehört auch die Frage, wie groß das Ausmaß der Teilnahme am systematischen Prozess der Erkenntnisgewinnung ist.

Innovation Die Innovation beschreibt die Aktualität der Forschungsfrage bzw. der Problemstellung, die bearbeitet wird. Als authentisch gilt eine Fragestellung dann, wenn sie in real bedeutsame Themenfelder aktueller Forschung eingebettet ist und Schülerinnen und Schüler durch ihre Forschungsergebnisse einen eigenen Beitrag zu deren Lösung leisten können. Durch die Beantwortung einer innovativen, neuartigen Forschungsfrage entsteht ein tatsächlicher Erkenntnisgewinn für die jeweilige Disziplin und nicht nur für die Schülerin oder den Schüler selbst (Sommer et al. 2020). Je zentraler im Projekt die Erforschung von etwas Neuem ist und bestehende wissenschaftliche Erkenntnisse nicht nur rezipiert bzw. repliziert werden, desto innovativer ist der dargebotene Inhalt.

Pilotierung und Überarbeitung

Um die Verständlichkeit der Items und deren Passung zu den theoretisch hergeleiteten Dimensionen empirisch zu prüfen, wurde eine erste Version des FEAW mit 17 Items und fünfstufiger Likert-Skala (1 = *völlig falsch* bis 5 = *völlig richtig*) entwickelt und pilotiert. Im Rahmen einer Expertenbefragung sind neun Expertinnen und Experten¹ aufgefordert worden, die einzelnen Items den übergeordneten Dimensionen zuzuordnen. Diese Expertenbefragung erzielte eine hohe Übereinstimmung² und kann – über die theoriegeleitete Konstruktion des Fragebogens und seiner Dimensionen hinaus – als Hinweis auf Inhaltsvalidität gewertet werden (Jenßen et al. 2015). Im Rahmen der darauffolgenden Pilotierung des FEAW an einer Stichprobe von 288 Schülerinnen und Schülern im naturwissenschaftlichen Schülerlabor (Alter $M=15,39$, $SD=2,69$; Anteil weiblich 50,4 %) konnte nach Ausschluss einzelner Items eine mehrfaktorielle Struktur mittels konfirmatorischer Faktorenanalyse bestätigt werden ($\chi^2=48,32$, $df=43$, $CFI=0,987$, $TLI=0,979$, $RMSEA=0,034$, $SRMR=0,028$). Die Subskala *Innovation* wies mit $\alpha=0,29$ keine hinreichende interne Konsistenz auf, was eine Überarbeitung der Items erforderte. Die internen Konsistenzen der übrigen Subskalen waren zufriedenstellend (Cronbachs $\alpha \geq 0,73$). Um parallel erste Hinweise auf die Konstruktvalidität der Subskala *Ort* zu ermitteln, wurde eine quasi-experimentelle Feldstudie durchgeführt, die nur den Lernort variierte (Schule/Schülerlabor). Wie erwartet, war die Authentizitätswahrnehmung in Bezug auf das Gestaltungsmerkmal *Ort* im Schülerlabor ($M=3,94$, $SD=1,00$) signifikant höher als in der Schule ($M=2,73$, $SD=1,25$), $F(1, 352)=101,78$, $p<0,001$, $\eta^2=0,224$. Der große Effekt, den die Lernum-

¹ Fachdidaktikerinnen und Fachdidaktiker der Ruhr-Universität aus unterschiedlichen Disziplinen mit Erfahrung in der Durchführung von Schülerlaborprojekten.

² Übereinstimmung bei allen Items jeweils mind. 94 %.

Tab. 2 Darstellung der externen Skalen zur Erhebung der konvergenten Außenkriterien

Quelle	Konstrukt und Beispielimem	<i>N</i>	<i>M</i>	<i>SD</i>	α
Damerau (2012)	Allgemeine Authentizität: „Ich habe heute einen Einblick in den Berufsalltag von Wissenschaftlern bekommen.“	253	3,65	1,05	0,90
Nachtigall et al. (2018b) ^a	Authentizität von Material und Aufgaben: „Ich denke, die Aufgaben stellen einen wichtigen Teil der Arbeit eines Wissenschaftlers dar.“	203	3,51	0,59	0,68
Huber et al. (2009) ^a	Expertise der Projektleitung: „Die Projektleitung ist ein Experte bzgl. des Projektthemas.“	202	4,62	0,52	0,88
Kremer (2010)	Überzeugung zur Herkunft des Wissens: „Anfänger können noch keine wissenschaftlichen Methoden anwenden.“	317	2,35	0,81	0,80
Moschner und Gruber (2017) ^a	Überzeugung zur Reflexivität von Wissen: „Wissen entwickelt sich weiter, wenn man sich kritisch damit auseinandersetzt.“	264	3,92	0,68	0,80

Skalen jeweils von 1 = *völlig falsch* bis 5 = *völlig richtig*

^aEinsatz nur im naturwissenschaftlichen Schülerlabor

gebung im naturwissenschaftlichen Schülerlabor auf die Authentizitätswahrnehmung bzgl. des Orts hat, ist ein erster Validitätshinweis für diese Subskala.

Methoden

Stichprobe

Zwischen März und November 2020 wurden in der Hauptstudie 19 Klassen unmittelbar nach ihrem Besuch im Alfred Krupp-Schülerlabor an der Ruhr-Universität Bochum befragt. An der Befragung nahmen Projekte aus acht verschiedenen Fächern teil (Anteil naturwissenschaftlich 82,5 %). Insgesamt haben an der freiwilligen Befragung 337 Schülerinnen und Schüler (Alter $M=16,61$, $SD=1,63$; Anteil weiblich 52,2 %) von vier unterschiedlichen Schulformen teilgenommen.

Instrumente

Erfassung von Authentizität in der Wissenschaftsvermittlung

Die Items des FEAW (detaillierte Darstellung in Tab. 3 im Ergebnisteil) können auf einer fünfstufigen Skala von 1 = *völlig falsch* bis 5 = *völlig richtig* bewertet werden. Für den FEAW lautet die Instruktion: „Die folgenden Aussagen beziehen sich auf das Projekt, an dem du teilgenommen hast. Bitte bewerte, ob die folgenden Aussagen bezogen auf das Projekt deiner Meinung nach richtig oder falsch sind. Kreuze bitte genau eines der fünf Kästchen an.“

Erfassung von Außenkriterien zur Validierung

Neben dem Ziel, die angenommene Faktorenstruktur des FEAW zu prüfen, sollen weitere Hinweise für die Konstruktvalidität des Messinstruments untersucht werden. Dazu werden mit der Statistiksoftware SPSS 26 theoretisch angenommene Zusammenhänge der Subskalen des FEAW mit verschiedenen Außenkriterien empirisch geprüft (Tab. 2).

Zunächst erwarten wir, dass die Skala zur Erfassung von allgemeiner Authentizität (Damerau 2012) mit allen Subskalen des FEAW eine positive Korrelation zeigt. Des Weiteren wird die Skala von Nachtigall et al. (2018b) eingesetzt, deren Verständnis von Authentizität sich darauf bezieht, „inwiefern die Aufgaben als passend für die Wissenschaftspraxis empfunden“ (S. 311) werden. Im Sinne von konvergenter bzw. diskriminanter Validität sollte diese Skala besonders hoch mit den Subskalen *Methoden* und *Material* und weniger hoch mit den verbleibenden Subskalen des FEAW korrelieren.

Weiterhin erwarten wir einen positiven Zusammenhang zwischen der Subskala *Vermittler*, die die Wahrnehmung der Projektleitung als echte Wissenschaftlerin bzw. als echter Wissenschaftler erfragt, und dem Ausmaß wahrgenommener Expertise der Projektleitung, gemessen mit einem adaptierten Fragebogen von Huber et al. (2009), da beide Skalen sinnverwandte Inhalte abfragen.

Zusätzlich werden zwei Skalen zur Erfassung epistemologischer Überzeugungen eingesetzt. Epistemologische Überzeugungen sind individuelle Annahmen über die Natur des Wissens sowie den Prozess des Wissenserwerbs (Hofer und Pintrich 1997; Mason 2010; Kremer und Kapitzka 2020) und stellen eine individuelle Disposition dar, die die Empfänglichkeit der Lernenden für die gestaltete Lernumgebung beeinflusst (Schommer-Aikins und Hutter 2002; Ma-

son 2010; Betz et al. 2016). Im Hinblick auf die ökonomische Umsetzbarkeit der gesamten Befragung wurde aus beiden Bereichen jeweils eine Skala eingesetzt. Die Skala Reflexivität des Wissens (Moschner und Gruber 2017) ist den Überzeugungen zur Natur des Wissens zuzuordnen und fragt individuelle Annahmen über die Veränderbarkeit des Wissens ab. Elaborierte Überzeugungen werden auf dieser Skala durch hohe Werte repräsentiert. Die Skala Herkunft des Wissens (Kremer 2010) ist den Überzeugungen zum Prozess des Wissenserwerbs zuzuordnen, wobei eine elaborierte Überzeugung durch niedrige Werte angezeigt wird. Da vor dem Hintergrund dieser individuellen Vorstellungen von Wissenschaft und wissenschaftlichen Prozessen die Authentizität der Schülerlabore subjektiv bewertet wird (vgl. Relationalität der Authentizität), nehmen wir die im Folgenden beschriebenen Zusammenhänge zwischen epistemologischen Überzeugungen und der Wahrnehmung von Authentizität an. Wir erwarten, dass die Subskala *Innovation*, die das Ausmaß der eigenen Beteiligung bei der Klärung einer wissenschaftlichen Forschungsfrage beschreibt, positiv mit elaborierten epistemologischen Überzeugungen zusammenhängt (positiver Zusammenhang mit der Skala Reflexivität des Wissens, negativer Zusammenhang mit der Skala Herkunft des Wissens). So schätzen Schülerinnen und Schüler den Innovationsgrad des Schülerlabors bspw. geringer ein, je fester sie der Überzeugung sind, dass sie selbst keine wissenschaftlichen Methoden anwenden oder neues Wissen generieren können.

Analysen zur Überprüfung des Messmodells

Zur Prüfung der faktoriellen Struktur des FEAW werden konfirmatorische Faktorenanalysen in Mplus (Version 8; Muthén und Muthén 1998–2017) durchgeführt. Es werden zwei Modelle unterschiedlicher Komplexität spezifiziert und hinsichtlich ihrer Passung auf die Daten verglichen. Als erstes Modell wird ein einfaktorielles Modell getestet (Generalfaktormodell), da Authentizität in bisherigen Studien häufig als eindimensionales Konstrukt operationalisiert wurde. Um die Annahme zu prüfen, dass Authentizität vielmehr ein mehrdimensionales Konstrukt ist, wird außerdem das theoretisch angenommene fünffaktorielle Modell getestet, das auf dem Modell von Betz et al. (2016) basiert. Für die Beurteilung der Modellpassung werden verschiedene Maße herangezogen. Der χ^2 -Test sollte nicht als einziges Kriterium betrachtet werden, da er aufgrund seiner Sensitivität bei großen Stichproben nicht zuverlässig ist. Als Faustregel für eine gute Modellpassung gilt eher das Verhältnis von $\chi^2/df \leq 2$ (Schermelleh-Engel und Moosbrugger 2003). Es werden außerdem der Root Mean Square Error of Approximation (RMSEA), das Standardized Root Mean Square Residual (SRMR), der Tucker-Lewis Index (TLI) und der Comparative Fit Index (CFI) herangezogen. Für

RMSEA und SRMR gelten Werte unter 0,05 als Zeichen für eine gute Modellpassung und Werte zwischen 0,05 und 0,08 zeigen eine akzeptable Modellpassung an. Für TLI und CFI gelten Werte über 0,95 bzw. 0,97 als Zeichen für eine gute Modellpassung (Hu und Bentler 1999; Schermelleh-Engel und Moosbrugger 2003; Kline 2011). Des Weiteren wird Cronbachs Alpha als Maß der Reliabilität der einzelnen Subskalen bestimmt. Als Richtwert für eine hinreichend reliable Skala gilt ein Wert von $\alpha \geq 0,80$ (Gäde et al. 2020).

Der FEAW soll entsprechend der breiten Zielgruppe von Schülerlaboren (vgl. Haupt 2015) über verschiedene Klassen hinweg genutzt werden. Deshalb wurden sowohl das Geschlecht (männlich/weiblich) als auch die Klassenstufe (Mittelstufe/Oberstufe) für die Messinvarianzprüfung ausgewählt. Die Einteilung der Variable Klassenstufe ist wie folgt: Klasse 7 bis 10 bilden die Mittelstufe und Klasse 11 bis 13 bilden die Oberstufe. Zur Überprüfung der Messinvarianz wird der *step-up*-Ansatz gewählt (Christ und Schlüter 2012). Das bedeutet, die Messinvarianz wird jeweils anhand zunehmend restriktiver Modelle mittels χ^2 -Differenztest in Mplus geprüft. Die getesteten Modelle umfassen die konfigurale Invarianz (gleiche Faktorenstruktur), die metrische Messinvarianz (invariante Faktorenstruktur und Faktorladungen) sowie die skalare Messinvarianz (invariante Faktorenstruktur, Faktorladungen und Itemintercepts), mit deren Vorliegen die Bedingung für die Berechnung von latenten Mittelwertunterschieden zwischen Gruppen erfüllt ist (Kleinke et al. 2017).

Die Betrachtung aller in die Analysen einfließenden Items des FEAW und die Relation von leeren Datenfeldern zu allen Datenfeldern zeigt, dass fehlende Werte nur zu einem geringen Anteil vorliegen (<2%). Die fehlenden Werte (Item-Nonresponse) in Bezug auf den FEAW zeigen ein allgemeines Ausfallmuster und werden als ignorierbare Nichtantworten (missing at random) behandelt (Schafer und Graham 2002). Durch die Tatsache, dass ganze Klassenverbände an der Validierungsstudie teilgenommen haben, liegen kaum Unit-Nonresponses vor. Jedoch sind unterschiedliche Fragebogensets eingesetzt worden, was eine Reduktion der vorliegenden Daten bzgl. der Außenkriterien per Design bedingt. So wurden die Authentizität von Material und Aufgaben (Nachtigall et al. 2018b), die Expertise der Projektleitung (Huber et al. 2009) und Überzeugungen zur Reflexivität von Wissen (Moschner und Gruber 2017) nur in naturwissenschaftlichen Schülerlaborprojekten erhoben (vgl. Anmerkung zu Tab. 2). Da die Daten insgesamt eine Abweichung von der Normalverteilung zeigen, werden die Analysen in Mplus mit dem MLR-Schätzer durchgeführt, der bei Verletzungen der Normalverteilungsvoraussetzung empfohlen wird (Christ und Schlüter 2012).

Tab. 3 Itemformulierungen mit Itemmittelwerten, Standardabweichungen und Faktorladungen (λ) basierend auf den Ergebnissen der explorativen Faktorenanalyse

	<i>M</i>	<i>SD</i>	λ_{F1}	λ_{F2}	λ_{F3}	λ_{F4}
<i>Ort</i>						
An dem Ort, an dem ich war, wird echte Forschung betrieben	4,02	1,07	0,800	–	–	–
Ich habe an einem Ort gearbeitet, an dem geforscht wird	3,82	1,17	0,716	–	0,500	–
Ich war in einer Umgebung, in der geforscht wird	3,96	1,05	0,792	–	0,502	–
<i>Vermittler</i>						
Mir stand ein echter Forscher zur Seite	3,35	1,35	–	0,874	–	–
Mir hat ein Wissenschaftler persönlich Forschung nähergebracht	3,29	1,30	–	0,767	–	–
Die Person, die das Projekt mit mir durchgeführt hat, ist selbst Wissenschaftler	3,32	1,34	–	0,773	–	–
<i>Vorgehen</i>						
Mir wurden Materialien aus der Wissenschaft bereitgestellt	4,36	0,99	–	–	0,801	–
Ich konnte Dinge nutzen, die auch in der Forschung genutzt werden	3,91	1,16	0,605	–	0,838	–
Ich habe wissenschaftliche Methoden angewendet	4,31	0,98	–	–	0,847	–
Ich bin selbst so vorgegangen, wie man in der Wissenschaft vorgeht	3,78	1,11	0,538	–	0,717	–
<i>Innovation</i>						
Ich habe versucht, etwas für die Wissenschaft Neues herauszufinden	2,07	1,27	–	–	–	0,755
Ich habe der Wissenschaft bei der Beantwortung einer wichtigen Frage geholfen	2,13	1,25	–	–	–	0,842
Ich habe einen Beitrag zur Beantwortung einer aktuellen Forschungsfrage geleistet	2,25	1,32	–	–	–	0,732

Tab. 4 Fitstatistiken der konfirmatorischen Faktoranalysen des FEAW (N= 337)

Faktoren	χ^2	<i>df</i>	χ^2/df	CFI	TLI	RMSEA	SRMR
1	707,48	65	10,88	0,564	0,477	0,171	0,139
4	86,57	59	1,47	0,981	0,975	0,037	0,042

Ergebnisse

Überprüfung des Messmodells

Die theoretisch angenommene Faktorenstruktur wurde mittels einer konfirmatorischen Faktorenanalyse geprüft. Bei der Überprüfung dieses fünffaktoriellen Modells ergab sich das Problem einer Missspezifikation (Korrelationen außerhalb des zulässigen Bereichs). Um das Problem zu identifizieren, ist außerplanmäßig eine explorative Faktorenanalyse durchgeführt worden. Das Ergebnis zeigt, dass nur vier der angenommenen fünf Faktoren extrahiert werden, bei der die Subskalen *Material* und *Methode* einen gemeinsamen Faktor bildeten (Tab. 3). Für diese vierfaktorielle Lösung ergab sich die Annäherung an eine Einfachstruktur, in der jedes Item substantiell auf dem zugehörigen Faktor lädt ($\lambda \geq 0,7$),

während die Nebenladungen nahezu alle gering sind³. Da eingesetzte Materialien und Methoden sich in der Praxis durchaus gegenseitig bedingen können, ist eine Zusammenlegung beider Dimensionen auch inhaltlich gut begründbar. Aus diesem Grund wurde für die weiteren Berechnungen ein vierfaktorielles Modell spezifiziert, bei dem die beiden Subskalen *Material* und *Methode* zu einer Skala *Vorgehen* zusammengefasst sind. In Tab. 3 sind die Formulierungen der Items sowie die Itemstatistiken der finalen Fassung des FEAW dargestellt.

Für das vierfaktorielle Modell ergab sich die beste Modellpassung (siehe Tab. 4). Der χ^2 -Differenztest, für den eine Korrektur mit dem MLR-scaling correcting factor durchgeführt wurde (Satorra und Bentler 2010), bestätigte die signifikant bessere Passung des vierfaktoriellen Modells

³ In Anlehnung an Backhaus et al. (2018) gelten Nebenladungen $\geq 0,50$ als hoch und sollten zur Interpretation der Faktoren herangezogen werden. Nebenladungen $< 0,50$ werden in Tab. 3 nicht dargestellt.

Tab. 5 Statistiken zur Überprüfung der Messinvarianz in Bezug auf das Geschlecht und Klassenstufe sowie Modellvergleiche

Modell	χ^2	<i>df</i>	CFI	RMSEA	SRMR	$\Delta\chi^2$	Δdf	<i>p</i>
<i>Messinvarianz Geschlecht</i>								
1 (konfigurale MI)	182,52	118	0,958	0,058	0,056	–	–	–
2 (metrische MI)	193,82	127	0,957	0,056	0,067	11,64	9	0,235
3 (skalare MI)	205,50	136	0,955	0,056	0,071	11,46	9	0,246
<i>Messinvarianz Klassenstufe</i>								
1 (konfigurale MI)	159,43	118	0,973	0,046	0,051	–	–	–
2 (metrische MI)	172,97	127	0,970	0,046	0,061	13,50	9	0,141
3 (skalare MI)	180,21	136	0,971	0,044	0,064	7,09	9	0,628

gegenüber eines Generalmodells ($\Delta\chi^2=446,845$; $\Delta df=6$; $p<0,001$).

Die Ergebnisse zur Überprüfung der Messinvarianz in Bezug auf das Geschlecht (52,2% weiblich) und die Klassenstufe (22,9% Mittelstufe) sind in Tab. 5 dargestellt.

Insgesamt wird in Bezug auf das Geschlecht die Annahme skalarer Messinvarianz gestützt, da sich für das restriktivere Modell 2 keine signifikant schlechtere Passung ergibt als für das Modell 1 ($\Delta\chi^2=11,64$; $\Delta df=9$; $p=0,235$) und für Modell 3 keine schlechtere Passung als für Modell 2 ($\Delta\chi^2=11,46$; $\Delta df=9$; $p=0,246$). Auch in Bezug auf die Klassenstufe stützen die Modellvergleiche von Modell 1 mit dem restriktiveren Modell 2 ($\Delta\chi^2=13,50$; $\Delta df=9$; $p=0,141$) sowie von Modell 2 und 3 ($\Delta\chi^2=7,09$; $\Delta df=9$; $p=0,628$) die Annahme von skalarer Invarianz. Die Überprüfung von Mittelwertsunterschieden in Bezug auf das Geschlecht zeigt, dass sich die Wahrnehmung von Authentizität zwischen Schülerinnen und Schülern lediglich in Bezug auf den Ort signifikant unterscheidet (*Vorgehen*: $t(327)=-1,10$, $p=0,272$; *Ort*: $t(328)=-2,38$, $p=0,018$; *Vermittler*: $t(327)=0,06$, $p=0,949$; *Innovation*: $t(327)=1,36$, $p=0,175$).

Überprüfung der Validität

Im Folgenden werden die latenten Korrelationen der Subskalen untereinander (Tab. 6) sowie die manifesten Korrelationen des FEWAW mit verschiedenen Außenkriterien dargestellt, die Aufschluss über die Konstruktvalidität des entwickelten Messinstruments geben können. Die Effektstärke der manifesten Korrelationen bewerten wir in Anlehnung an Döring und Bortz (2016). Es zeigt sich eine positive Kor-

relation zwischen *Ort* und *Vermittler* ($r=0,48$), zwischen *Ort* und *Vorgehen* ($r=0,42$) sowie *Vorgehen* und *Vermittler* ($r=0,43$). Die Subskala *Innovation* zeigt nur bedingt einen Zusammenhang mit den anderen Subskalen, eine positive Korrelation ergibt sich zwischen *Innovation* und *Vermittler* ($r=0,21$).

Bei der Analyse der Zusammenhänge zwischen unterschiedlichen Außenkriterien und den Subskalen des FEWAW zeigen sich überwiegend positive Zusammenhänge (Tab. 7).

Es bestehen erwartungskonforme, positive Zusammenhänge zwischen den einzelnen Subskalen des FEWAW und der wahrgenommenen *allgemeinen Authentizität* des Schülerlabors. Dabei sind die Zusammenhänge mit *Vermittler* ($r=0,53$, $p<0,001$) und *Vorgehen* ($r=0,67$, $p<0,001$) hoch, mit *Ort* zeigt sich eine mittlere ($r=0,44$, $p<0,001$) und mit *Innovation* eine schwache Korrelation ($r=0,18$, $p=0,005$). Die Skala *Authentizität der Aufgaben* zeigt erwartungskonform einen besonders starken Zusammenhang mit der Subskala *Vorgehen* ($r=0,54$, $p<0,001$). Die Korrelationen mit den übrigen Subskalen *Ort* ($r=0,37$, $p<0,001$), *Vermittler* ($r=0,42$, $p<0,001$) und *Innovation* ($r=0,25$, $p<0,001$) fallen hingegen geringer aus. Die Skala *Expertise der Projektleitung* zeigt mit einer mittleren positiven Korrelation ($r=0,33$, $p<0,001$) erwartungsgemäß den stärksten Zusammenhang mit der Subskala *Vermittler*. Es zeigen sich schwache positive Zusammenhänge mit *Ort* ($r=0,27$, $p<0,001$) und *Vorgehen* ($r=0,21$, $p=0,003$). Die epistemologischen Überzeugungen zur *Herkunft des Wissens* zeigen erwartungswidrig einen schwachen positiven Zusammenhang mit *Innovation* ($r=0,26$, $p<0,001$) und einen schwachen negativen Zusammenhang mit *Ort* ($r=-0,13$, $p=0,024$). Die

Tab. 6 Reliabilitäten (interne Konsistenzen), Skalenmittelwerte und Standardabweichungen sowie latente Korrelationen der Subskalen

	α	<i>M</i>	<i>SD</i>	1	2	3	4
1 Ort	0,82	3,93	0,93	–	–	–	–
2 Vermittler	0,84	3,32	1,12	0,48	–	–	–
3 Vorgehen	0,87	4,09	0,91	0,42	0,43	–	–
4 Innovation	0,81	2,16	1,10	-0,03	0,21	-0,08	–

Berechnungen für $307 \leq N \leq 312$, Skala jeweils von 1 = *völlig falsch* bis 5 = *völlig richtig*
 Statistisch bedeutsame Korrelationen ($p<0,05$) sind fettgedruckt

Tab. 7 Manifeste Korrelationen der Authentizitätsdimensionen mit Außenkriterien

Außenkriterium	Authentizitätsdimension	Ort	Vermittler	Vorgehen	Innovation
Allgemeine Authentizität ^a		0,44	0,53	0,67	0,18
Authentizität der Aufgaben ^b		0,37	0,42	0,54	0,25
Expertise der Projektleitung ^b		0,27	0,33	0,21	–0,11
Herkunft des Wissens ^c		–0,13	0,08	–0,05	0,26
Reflexivität des Wissens ^a		0,36	0,23	0,44	–0,20

Berechnungen für ^a252 ≤ N ≤ 264, ^b201 ≤ N ≤ 203, ^c316 ≤ N ≤ 317
 Statistisch bedeutsame Korrelationen ($p < 0,05$) sind fettgedruckt

Skala zur *Reflexivität des Wissens* steht in einem schwachen negativen Zusammenhang zur *Innovation* ($r = -0,20$, $p = 0,002$), während mittlere positive Korrelationen mit *Ort* ($r = 0,36$, $p < 0,001$) und *Vorgehen* ($r = 0,44$, $p < 0,001$) sowie ein schwacher Zusammenhang mit *Vermittler* ($r = 0,23$, $p < 0,001$) bestehen.

Diskussion

Das Ziel der vorliegenden Studie war die Entwicklung eines Fragebogens, der die Wahrnehmung von Authentizität in der Wissenschaftsvermittlung mehrdimensional erfasst, um dem Gegensatz zwischen der vielfach betonten Bedeutsamkeit der Authentizität in Lehr-Lernkontexten (Euler et al. 2015) und ihrer tendenziell wenig ausdifferenzierten Erfassung zu begegnen. Der FEWAW wurde theoriebasiert entwickelt und beruht auf dem Modell der Authentizität der Wissenschaftsvermittlung (Betz et al. 2016), welches verschiedene Merkmale der Lernumgebung unterscheidet, die einen Spielraum zur Gestaltung von objektiv authentischen Lernumgebungen bieten. Gemäß der Differenzierung von Anker-Hansen und Andrée (2019) wurde der Begriff Authentizität in dieser Studie mit Bezug auf den professionellen, beruflich-wissenschaftlichen Kontext verstanden. Zur Überprüfung der psychometrischen Qualität des Messinstruments gehörte nicht nur die Prüfung der angenommenen Faktorenstruktur mittels konfirmatorischer Faktorenanalyse, sondern auch die Überprüfung von Messinvarianz sowie die Prüfung von Zusammenhängen mit unterschiedlichen Außenkriterien als Hinweis auf die Konstruktvalidität des Instruments.

Messmodell

Die Ergebnisse der vorliegenden Validierungsstudie schließen an theoretische Überlegungen zur Wahrnehmung von Authentizität an. Die Analysen zur Überprüfung des Messmodells zeigen eine signifikant bessere Passung des mehrdimensionalen Messmodells. Damit kann die Grundannahme des theoretischen Modells (Betz et al. 2016), das Authentizität als mehrdimensionales Konstrukt versteht, empi-

risch bestätigt werden. Der FEWAW, der die Wahrnehmung der Authentizität in der Wissenschaftsvermittlung misst, beinhaltet die vier Subdimensionen *Ort*, *Vermittler*, *Vorgehen* und *Innovation*, die mit Reliabilitäten von $0,81 \leq \alpha \leq 0,87$ gute Werte aufweisen (Gäde et al. 2020). Auch die messfehlerfreien Korrelationen der Subdimensionen untereinander ($r \leq 0,48$) zeigen, dass die einzelnen Dimensionen empirisch gut voneinander trennbar sind. Die Subskala *Vorgehen* entstand aus einer Zusammenlegung der Skalen *Material* und *Methode*. Das zunächst angenommene fünffaktorielle Modell, bei dem die Subskalen *Material* und *Methode* getrennt voneinander bestehen, musste im Rahmen der vorliegenden Untersuchung verworfen werden. Damit können auch die theoretischen Annahmen mit Blick auf die Differenzierung von *Material* und *Methode* (Betz et al. 2016) empirisch nicht bestätigt werden. Es ist möglich, dass sich die eingesetzten Materialien und angewandten Methoden in einem Schülerlaborprojekt zu einem gewissen Grad gegenseitig bedingen und die beiden Subdimensionen von Schülerinnen und Schülern deshalb nicht genau differenziert werden können. Bereits im Rahmen der Pilotierung zeigte sich, dass beide Subdimensionen hoch korrelieren und die Zusammenlegung beider Subskalen zu einer deutlichen Verbesserung im Modellfit führt. Um das Ergebnis aus der Pilotierung abzusichern und den theoretischen Annahmen des Modells der Authentizität der Wissenschaftsvermittlung (Betz et al. 2016) zu folgen, sind im Rahmen der vorliegenden Untersuchung jedoch nochmals beide Skalen als eigenständige Dimensionen modelliert und geprüft worden. Zwischen den Subskalen *Ort* und *Vorgehen* zeigen sich in der Faktorenmatrix einer explorativen Faktorenanalyse jedoch hohe Nebenladungen. Dieses systematische Muster ergibt sich möglicherweise aus den Itemformulierungen, da Prozesse wissenschaftlichen Arbeitens und die Vorgehensweise „Forschen“ als beschreibendes Merkmal des Ortes genutzt werden („Ich war an einem Ort, an dem geforscht wird“). Zudem fragt die Subskala *Vorgehen* u. a. die wahrgenommene Authentizität der Materialien ab, die durchaus zur wissenschaftlichen Ausstattung eines Raumes bzw. Ortes beitragen können.

Neben der Überprüfung des Messmodells konnte in Bezug auf das Geschlecht und die Klassenstufe das Vorliegen

skalärer Messinvarianz bestätigt werden, was eine wichtige Grundvoraussetzung für den Vergleich latenter Mittelwerte zwischen verschiedenen Gruppen ist (Christ und Schlüter 2012). Das Ergebnis, dass sich die Wahrnehmung von Authentizität zwischen Schülerinnen und Schülern lediglich in Bezug auf den Ort signifikant unterscheidet, steht größtenteils im Einklang mit bisherigen empirischen Befunden, die keinen Geschlechtsunterschied belegen (Engeln 2004; Damerau 2012). Die Abweichungen in Bezug auf die Wahrnehmung des Ortes entstehen möglicherweise dadurch, dass bisherige Messinstrumente die Authentizität nicht differenziert erfasst haben.

Konstruktvalidität

Erste Hinweise auf konvergente Validität lieferten die positiven Zusammenhänge der Subskalen des FEAW mit dem Fragebogen von Damerau (2012) zur eindimensionalen Erfassung allgemeiner Authentizität und dem Fragebogen von Nachtigall et al. (2018b) zur Erfassung der Authentizität von Materialien und Aufgaben. Die Formulierung der Items zur Erfassung der allgemeinen Authentizität kann dazu beigetragen haben, dass mit der Subskala *Innovation* nur ein schwacher Zusammenhang besteht. Um einen Einblick in die Wissenschaft erhalten zu können, ist es eventuell nicht notwendig, selbst aktiv zu sein und neues Wissen zu generieren. Der Ort oder dargebotene Materialien könnten für einen Einblick in den Arbeitsalltag einer Wissenschaftlerin bzw. eines Wissenschaftlers zentraler sein. Bisher ist die Frage nach der Gewichtung der einzelnen Merkmale authentischer Lernumgebungen ungeklärt (vgl. Nachtigall et al. 2018b). Es stehen weitere Untersuchungen aus, ob sich die Merkmale authentischer Lernumgebungen additiv summieren oder ob es ggf. eine zentrale Dimension gibt, die die Wahrnehmung der übrigen Dimensionen überlagert, sodass alle Dimensionen unabhängig von der tatsächlichen Ausgestaltung als authentisch wahrgenommen werden. Der positive Zusammenhang zwischen der Skala Expertise der Projektleitung (Huber et al. 2009) und der Subskala *Vermittler* zeigt sich erwartungskonform und kann ebenfalls als Hinweis auf die Konstruktvalidität des FEAW gewertet werden.

Zudem sind Zusammenhänge des FEAW mit epistemologischen Überzeugungen betrachtet worden. In der vorliegenden Untersuchung stehen die Überzeugungen zur Herkunft des Wissens (Kremer 2010) erwartungswidrig in einem positiven Zusammenhang mit der *Innovation*. Das bedeutet, dass Schülerinnen und Schüler mit tendenziell elaborierten Überzeugungen, die der Meinung sind, dass wissenschaftliches Wissen nicht nur durch Wissenschaftlerinnen und Wissenschaftler, sondern auch selbst erzeugt werden kann (Bromme und Kienhues 2014), den Innovationsgrad der Schülerlaborprojekte geringer einschätzen. Der

schwache negative Zusammenhang zwischen Überzeugungen zur Herkunft des Wissens und der Subskala *Ort* bedeutet, dass Schülerinnen und Schüler, die selbst der Meinung sind, keine wissenschaftlichen Methoden anwenden zu können, das Schülerlabor auch nicht als einen Ort wahrnehmen, an dem echte Wissenschaft betrieben wird. Die Überzeugungen zur Reflexivität des Wissens (Moschner und Gruber 2017) stehen erwartungswidrig in einem negativen Zusammenhang zur *Innovation*, während positive Korrelationen mit den übrigen Subskalen bestehen. Auch hier schätzen Schülerinnen und Schüler mit tendenziell elaborierten Überzeugungen, die der Meinung sind, dass sich wissenschaftliches Wissen weiterentwickelt (Bromme und Kienhues 2014), die Innovation der Schülerlaborprojekte geringer ein. Auch wenn sich die Zusammenhänge zwischen der Skala *Innovation* und den epistemologischen Überzeugungen nicht in der erwarteten Richtung gezeigt haben, so bleibt das angenommene gegenläufige Muster der Korrelationen bestehen. Es zeigt sich insgesamt, dass die *Innovation* geringer bewertet wird, je elaborierter die Überzeugungen der Schülerinnen und Schüler sind. Unter der Annahme, dass das im Rahmen von Schülerlaborprojekten generierte Wissen meist nur für die Schülerinnen und Schüler selbst, nicht aber für die wissenschaftliche Gemeinschaft neu ist, erscheinen diese Korrelationsmuster im Nachhinein plausibel. Diese Annahme lässt sich bei retrospektiver Betrachtung der untersuchten Schülerlaborprojekte in dem Maße stützen, als dass die gewonnenen Daten bspw. nicht in wissenschaftliche Publikationen münden.

Limitationen

Eine Limitation der vorliegenden Studie ist die hierarchische Struktur der angefallenen Daten, die durch die Erhebungen im Klassenverband begründet ist. Alle teilnehmenden Schülerinnen und Schüler (Individualebene) stammen aus 19 Klassen (Kontexteinheiten). Die Anzahl der Beobachtungen pro Kontexteinheit war jedoch zu gering, um die Mehrebenenstruktur der Daten in den Analysen zu berücksichtigen (Kleinke et al. 2017). Für zukünftige Untersuchungen muss es daher das Ziel sein, eine ausreichend große Stichprobe zu generieren, um die hierarchische Struktur der Daten angemessen berücksichtigen zu können. Außerdem sollte im Rahmen folgender Untersuchungen der Zusammenhang zwischen den Subskalen des FEAW und weiteren geeigneten konvergenten und diskriminanten Außenkriterien untersucht werden, um die Konstruktvalidität des Fragebogens weiter abzusichern. Ein geeignetes konvergentes Außenkriterium ist bspw. das Interesse am Projektthema, zu dem bereits empirische Befunde über einen positiven Zusammenhang mit der Authentizitätswahrnehmung vorliegen (Engeln 2004; Pawek 2009).

Durch das breite Fächerspektrum des Alfred Krupp-Schülerlabors und den Einsatz des FEAW im naturwissenschaftlichen sowie im geisteswissenschaftlichen Bereich gibt es zunächst keine Beschränkung der Ergebnisse auf einzelne Fächer oder Projekte. Auch wenn die Verteilung der Schülerinnen und Schüler auf beide Bereiche sehr ungleichmäßig ist (Anteil naturwissenschaftlich 82,5%), entspricht die vorliegende Verteilung in etwa der in deutschen Schülerlaboren vorherrschenden Präferenz für ein naturwissenschaftliches Angebot (LernortLabor 2015). In einer nachfolgenden Untersuchung wäre es dennoch empfehlenswert, stärker geisteswissenschaftliche Projekte in den Blick zu nehmen, da sie eine ebenso zentrale Rolle in der Wissensvermittlung spielen, es bisher aber kaum empirische Untersuchungen zu deren Wahrnehmung und Wirkung gibt.

Trotz einiger Limitationen der vorliegenden Untersuchung konnten die theoretischen Grundannahmen des Modells der Authentizität der Wissensvermittlung (Betz et al. 2016) zur Mehrdimensionalität der Authentizität bestätigt werden und der in dieser Studie validierte Fragebogen zur Erfassung der Wahrnehmung von Authentizität in der Wissensvermittlung (FEAW) stellt ein geeignetes Instrument dar, um den Wirkungszusammenhang von Authentizität in Lehr-Lernkontexten weiter zu untersuchen.

Danksagung Die vorliegende Studie wurde innerhalb des Promotionskollegs „Metakognitives Monitoring in authentischen Lehr-/Lernkontexten im Schülerlabor“ (MeMo-akS) durch die Professional School of Education der Ruhr-Universität Bochum gefördert. Wir danken dem Alfred Krupp-Schülerlabor für die stetige Unterstützung, allen an dieser Studie beteiligten Schulklassen sowie den Gutachter*innen für die wertvollen Hinweise zur Überarbeitung des Manuskripts.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

Literatur

- Anker-Hansen, J., & Andrée, M. (2019). The pursuit of authenticity in science education. *Nordic Studies in Science Education*, 15(1), 498–510. <https://doi.org/10.5617/nordina.4723>.
- Backhaus, K., Erichson, B., Plinke, W., & Weiber, R. (2018). *Multivariate Analysemethoden*. Springer.
- Bauer, J., Berthold, K., Hefter, M. H., Prenzel, M., & Renkl, A. (2017). Wie können Lehrkräfte und ihre Schülerinnen und Schüler lernen, fragile Evidenz zu verstehen und zu nutzen? *Psychologische Rundschau*, 68(3), 188–192. <https://doi.org/10.1026/0033-3042/a000363>.
- Betz, A. (2018). Der Einfluss der Lernumgebung auf die (wahrgenommene) Authentizität der linguistischen Wissensvermittlung und das Situationale Interesse von Lernenden. *Unterrichtswissenschaft*, 46, 261–278. <https://doi.org/10.1007/s42010-018-0021-0>.
- Betz, A., Flake, S., Mierwald, M., & Vanderbeke, M. (2016). Modelling authenticity in teaching and learning contexts: a contribution to theory development and empirical investigation of the construct. In C. K. Looi, J. Polman, U. Cress & P. Reimann (Hrsg.), *Transforming learning, empowering earners: The International Conference of the Learning Sciences (ICLS) 2016* (Bd. 2, S. 815–818). International Society of the Learning Sciences. <https://doi.org/10.22318/icls2016.110>.
- Brandt, A. (2005). *Förderung von Motivation und Interesse durch außerschulische Experimentierlabors*. Dissertation. Universität Bielefeld.
- Brandt, A., Möller, J., & Kohse-Höinghaus, K. (2008). Was bewirken außerschulische Experimentierlabors? *Zeitschrift für Pädagogische Psychologie*, 22(1), 5–12. <https://doi.org/10.1024/1010-0652.22.1.5>.
- Bromme, R., & Kienhues, D. (2014). Wissenschaftsverständnis und Wissenschaftskommunikation. In T. Seidel & A. Krapp (Hrsg.), *Pädagogische Psychologie* (S. 55–81). Beltz.
- Capps, D. K., & Crawford, B. A. (2013). Inquiry-based instruction and teaching about nature of science: are they happening? *Journal of Science Teacher Education*, 24, 497–526. <https://doi.org/10.1007/s10972-012-9314-z>.
- Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic inquiry in schools. A theoretical framework for evaluating inquiry tasks. *Science Education*, 86(2), 175–218. <https://doi.org/10.1002/sce.10001>.
- Christ, O., & Schlüter, E. (2012). *Strukturgleichungsmodelle mit Mplus. Eine praktische Einführung*. Oldenbourg. <https://doi.org/10.1524/9783486714807>.
- Damerau, K. (2012). *Molekulare und Zell-Biologie im Schülerlabor. Fachliche Optimierung und Evaluation der Wirksamkeit im BeLL Bio (Bergisches Lehr-Lern-Labor Biologie)*. Dissertation. Bergische Universität Wuppertal.
- Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Springer.
- Edelson, D. C., Gordin, D. N., & Pea, R. D. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum design. *The Journal of the Learning Sciences*, 8(3, 4), 391–450.
- Engeln, K. (2004). *Schülerlabors: Authentische, aktivierende Lernumgebungen als Möglichkeit, Interesse an Naturwissenschaften und Technik zu wecken*. Logos.
- Euler, M., & Weßnig, S. (2011). Schülerlabore und die Förderung kreativer Potenziale. *Plus Lucis*, 1–2, 32–38.
- Euler, M., Schüttler, T., & Hausamann, D. (2015). Schülerlabore: Lernen durch Forschen und Entwickeln. In E. Kircher, R. Girwitz & P. Häubler (Hrsg.), *Physikdidaktik* (S. 759–782). Springer. https://doi.org/10.1007/978-3-642-41745-0_26.
- Fischer, H. E. (1998). Scientific Literacy und Physiklernen. *Zeitschrift für Didaktik der Naturwissenschaften*, 4(2), 41–52.

- Gäde, J. C., Schermelleh-Engel, K., & Werner, C. S. (2020). Klassische Methoden der Reliabilitätsschätzung. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 305–334). Springer. https://doi.org/10.1007/978-3-662-61532-4_14.
- Glowinski, I. (2007). *Schülerlabore im Themenbereich Molekularbiologie als Interesse fördernde Lernumgebungen*. Dissertation. Christian-Albrechts-Universität Kiel.
- Gräber, W., & Nentwig, P. (2002). Scientific Literacy – Naturwissenschaftliche Grundbildung in der Diskussion. In W. Gräber, P. Nentwig, T. R. Koballa & E. H. Evans (Hrsg.), *Scientific Literacy* (S. 7–20). Leske + Budrich.
- Gudian, P. (2007). *Wirksamkeitsanalyse außerschulischer Lernorte. Der Einfluss mehrmaliger Besuche eines Schülerlabors auf die Entwicklung des Interesses an Physik*. Dissertation. Humboldt-Universität Berlin.
- Gulikers, J. T. M., Bastiaens, Th. J., & Kirschner, P. A. (2006). Authentic assessment, student and teacher perceptions: the practical value of the five-dimensional framework. *Journal of Vocational Education & Training*, 58, 337–357. <https://doi.org/10.1080/13636820600955443>.
- Habig, S., van Vorst, H., & Sumfleth, E. (2018). Merkmale kontextualisierter Lernaufgaben und ihre Wirkung auf das situationale Interesse und die Lernleistung von Schülerinnen und Schülern. *Zeitschrift für Didaktik der Naturwissenschaften*, 24, 99–114. <https://doi.org/10.1007/s40573-018-0077-8>.
- Hagenkötter, R., Nachtigall, V., Rolka, K., & Rummel, N. (2021). „Meistens sind Forscher älter, meist tragen die eine Brille“ – Schülervorstellungen über Wissenschaftler*innen. *Unterrichtswissenschaft*. <https://doi.org/10.1007/s42010-021-00110-1>.
- Haupt, O. J. (2015). In Zahlen und Fakten. Der Stand der Bewegung! In Lern-ortLabor (Hrsg.), *Schülerlabor-Atlas 2015* (S. 34–53). Klett.
- Haupt, O. J., & Hempelmann, R. (2015). Schülerlabore in Art und Form. Eine Typensache! In Lern-ortLabor (Hrsg.), *Schülerlabor-Atlas 2015* (S. 14–21). Klett.
- Haupt, O. J., Domjahn, J., Martin, U., Skiebe-Corrette, P., Vorst, S., Zehren, W., & Hempelmann, R. (2013). Schülerlabor – Begriffsschärfung und Kategorisierung. *Der mathematische und naturwissenschaftliche Unterricht*, 66(6), 324–330.
- Herrington, A., & Herrington, J. (2006). What is an authentic learning environment? In A. Herrington & J. Herrington (Hrsg.), *Authentic learning environments in higher education*. INFOSCI.
- Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67(1), 88–140. <https://doi.org/10.3102/00346543067001088>.
- Honebein, P. C., Duffy, T. M., & Fishman, B. J. (1993). Constructivism and the design of learning environments: context and authentic activities for learning. In T. M. Duffy, J. Lowyck, D. H. Jonassen & T. M. Welsh (Hrsg.), *Designing environments for constructive learning* (S. 87–108). Springer.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>.
- Huber, F., Vollhardt, K., & Meyer, F. (2009). Helden der Werbung? Eine Untersuchung der Relevanz von Werbefiguren für das Konsumentenverhalten. *Marketing ZFP*, 31(3), 183–196. <https://doi.org/10.15358/0344-1369-2009-3-183>.
- Itzek-Greulich, H., Flunger, B., Vollmer, C., Nagengast, B., Rehm, M., & Trautwein, U. (2015). Effects of a science center outreach lab on school students' achievement—are student lab visits needed when they teach what students can learn at school? *Learning and Instruction*, 38, 43–52. <https://doi.org/10.1016/j.learninstruc.2015.03.00>.
- Itzek-Greulich, H., Flunger, B., Vollmer, C., Nagengast, B., Rehm, M., & Trautwein, U. (2017). Effectiveness of lab-work learning environments in and out of school: a cluster randomized study. *Contemporary Educational Psychology*, 48, 98–115. <https://doi.org/10.1016/j.cedpsych.2016.09.005>.
- Jenßen, L., Duneckacke, S., & Blömeke, S. (2015). Qualitätssicherung in der Kompetenzforschung. Empfehlungen für den Nachweis von Validität in Testentwicklung und Veröffentlichungspraxis. *Zeitschrift für Pädagogik*, 61, 11–31. Beiheft.
- Kleinke, K., Schlüter, E., & Christ, O. (2017). *Strukturgleichungsmodelle mit Mplus*. De Gruyter. <https://doi.org/10.1515/9783486989458>.
- Kline, R. B. (2011). *Principles and practice of structural equation modeling*. Guilford: .
- Krapp, A. (2001). Interesse. In D. Rost (Hrsg.), *Handwörterbuch pädagogische Psychologie* (S. 286–293). Beltz.
- Kremer, K. H. (2010). *Die Natur der Naturwissenschaften verstehen – Untersuchungen zur Struktur und Entwicklung von Kompetenzen in der Sekundarstufe I*. Dissertation. Universität Kassel.
- Kremer, K., & Kapitza, M. (2020). Untersuchung von epistemischen Überzeugungen im Schülerlabor – Möglichkeiten und Herausforderungen. In K. Sommer, J. Wirth & M. Vanderbeke (Hrsg.), *Handbuch Forschen im Schülerlabor* (S. 79–90). Waxmann.
- Lee, H.-S., & Songer, N. B. (2003). Making authentic science accessible to students. *International Journal of Science Education*, 25(8), 923–948. <https://doi.org/10.1080/09500690305023>.
- LernortLabor (Hrsg.). (2015). *Schülerlabor-Atlas 2015*. Klett.
- Mandl, H., Gruber, H., & Renkl, A. (1995). Situiertes Lernen in multimedialen Lernumgebungen. In L. J. Issing & P. Klimsa (Hrsg.), *Informationen und Lernen mit Multimedia* (S. 167–178). Beltz, PVU.
- Mason, L. (2010). Beliefs about knowledge and revision of knowledge: on the importance of epistemic beliefs for intentional conceptual change in elementary and middle school students. In L. D. Bendixen & F. C. Feucht (Hrsg.), *Personal epistemology in the classroom* (S. 258–291). Cambridge: .
- Mierwald, M. (2020). *Historisches Argumentieren und epistemologische Überzeugungen. Eine Interventionsstudie zur Wirkung von Lernmaterialien im Schülerlabor*. Springer VS.
- Mierwald, M., Lehmann, T., & Brauch, N. (2018). Zur Veränderung epistemologischer Überzeugungen im Schülerlabor: Authentizität von Lernmaterial als Chance der Entwicklung einer wissenschaftlich angemessenen Überzeugungshaltung im Fach Geschichte? *Unterrichtswissenschaft*, 46, 279–297. <https://doi.org/10.1007/s42010-018-0019-7>.
- Moschner, B., & Gruber, H. (2017). Erfassung epistemischer Überzeugungen mit dem FEE. In A. Bernholt, H. Gruber & B. Moschner (Hrsg.), *Wissen und Lernen* (S. 17–38). Waxmann.
- Muthén, L. K., & Muthén, B. O. (2017). *Mplus user's guide* (8. Aufl.). Muthén & Muthén.
- Nachtigall, V., Keuschnig, A., Behrendt, L., & Brune, L. (2018a). Authentic learning and teaching in an out-of-school lab—First steps towards empirical investigation of a theoretical model. In J. Kay & R. Luckin (Hrsg.), *Rethinking learning in the digital age: making the learning sciences count, 13th International Conference of the Learning Sciences (ICLS) 2018* (Bd. 2, S. 1061–1064). International Society of the Learning Sciences. <https://doi.org/10.22318/csl2018.1061>.
- Nachtigall, V., Rummel, N., & Serova, K. (2018b). Authentisch ist nicht gleich authentisch – Wie Schülerinnen und Schüler die Authentizität von Lernaktivitäten im Schülerlabor einschätzen. *Unterrichtswissenschaft*, 46, 299–319. <https://doi.org/10.1007/s42010-018-0020-1>.
- OECD (2013). *PISA 2015 draft science framework*. OECD.
- Pawek, C. (2009). *Schülerlabore als interessesfördernde außerschulische Lernumgebungen für Schülerinnen und Schüler aus der Mittel- und Oberstufe*. Dissertation. Christian-Albrechts-Universität Kiel.
- Peacock, M. (1997). The effect of authentic materials on the motivation of EFL learners. *ELT Journal*, 51(2), 144–156.

- Satorra, A., & Bentler, P.M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika*, 75(2), 243–248. <https://doi.org/10.1007/s11336-009-9135-y>.
- Schafer, J.L., & Graham, J.W. (2002). Missing data: our view of the state of the art. *Psychological Methods*, 7(2), 147–177. <https://doi.org/10.1037/1082-989X.7.2.147>.
- Scharfenberg, F.-J. (2005). *Experimenteller Biologieunterricht zu Aspekten der Gentechnik im Lernort Labor: Empirische Untersuchung zu Akzeptanz, Wissenserwerb und Interesse*. Dissertation. Universität Bayreuth.
- Scharfenberg, F.-J., & Bogner, F.X. (2015). Empirische Analyse. Leistung zählt! In LernortLabor (Hrsg.), *Schülerlabor-Atlas 2015* (S. 24–31). Klett.
- Schermelleh-Engel, K., & Moosbrugger, H. (2003). Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8(2), 23–74.
- Schommer-Aikins, M., & Hutter, R. (2002). Epistemological beliefs and thinking about everyday controversial issues. *The Journal of Psychology*, 136(1), 5–20.
- Schüttler, T., Watzka, B., Girwidz, R., & Ertl, B. (2021). Die Wirkung der Authentizität von Lernort und Laborgeräten auf das situationale Interesse und die Relevanzwahrnehmung beim Besuch eines naturwissenschaftlichen Schülerlabors. *Zeitschrift für Didaktik der Naturwissenschaften*. <https://doi.org/10.1007/s40573-021-00128-z>.
- Sommer, K., Wirth, J., & Rummel, N. (2018). Authentizität der Wissenschaftsvermittlung im Schülerlabor – Einführung in den Thementeil. *Unterrichtswissenschaft*, 46, 253–260. <https://doi.org/10.1007/s42010-018-0022-z>.
- Sommer, K., Firstein, A., & Rothstein, B. (2020). Authentizität (der Wissenschaftsvermittlung) im Schülerlabor. In K. Sommer, J. Wirth & M. Vanderbeke (Hrsg.), *Handbuch Forschen im Schülerlabor* (S. 21–30). Waxmann.
- Van Vorst, H., Dorschu, A., Fechner, S., Kauertz, A., Krabbe, H., & Sumfleth, E. (2015). Charakterisierung und Strukturierung von Kontexten im naturwissenschaftlichen Unterricht – Vorschlag einer theoretischen Modellierung. *Zeitschrift für Didaktik der Naturwissenschaften*, 21(1), 29–39. <https://doi.org/10.1007/s40573-014-0021-5>.