

Das Dhoch3-Modul 11 zum „Beurteilen von Sprachkompetenzen in Deutsch als Fremdsprache“

Inhalte, Leitkonzepte und Überlegungen zu Nützlichkeit und Wirksamkeit

Katharina Karges, Alice Friedland & Katrin Wisniewski

Abstract: Obwohl Beurteilungen von Sprachkompetenzen in der Praxis allgegenwärtig sind, wird die Kompetenz, Beurteilungen zu verstehen, zu gestalten und zu nutzen in der Lehrpersonenaus- und -weiterbildung kaum systematisch vermittelt. Um dem entgegenzuwirken, wurde das Dhoch3-Modul „Beurteilen von Sprachkompetenzen in DaF“ entwickelt, das dazu befähigen soll, Kernbegriffe, -konzepte und -methoden der Beurteilung zu vermitteln und für den eigenen Lehrkontext nutzbar zu machen. Wie alle Dhoch3-Module steht es vor der Herausforderung, Material für verschiedenste Standorte und curriculare Kontexte bereitzustellen, die von den Nutzenden als relevant betrachtet werden und an ihre Interessen und Wissensstände anknüpfen. Unser Beitrag führt in die Hauptinhalte von Modul 11 ein und erläutert die dafür zentralen Konzepte der *(Language) Assessment Literacy* und des *Local Language Testing*. Er diskutiert darüber hinaus mögliche Hindernisse und Gelingensbedingungen für die Wirksamkeit dieses und weiterer Module.

Schlagwörter: Dhoch3, Beurteilung, Testen, Prüfen, Lehrerbildung

Abstract: Although the assessment of language competence is ubiquitous in practice, the ability to understand, design, and make use of assessments is hardly ever taught systematically in teacher education and professional development. To address this gap, the Dhoch3 module “Assessing Language Competence in German as a Foreign Language” was developed. Its aim is to enable participants to engage with core terms, concepts, and methods of assessment and make them applicable in local teaching contexts. Like all Dhoch3 modules, it faces the challenge of providing materials that are relevant across diverse institutional settings and curricular contexts, while at the same time connecting to users’ prior knowledge and interests. Our contribution introduces the main contents of Module 11 and explains the key concepts that guided its development, namely *(Language) Assessment Literacy* and *Local Language Testing*. In addition, it discusses possible obstacles and enabling conditions for the usefulness and effectiveness of this and other modules.

Keywords: Dhoch3, Assessment, language testing, teacher training



Dieses Werk ist lizenziert unter einer Creative Commons Namensnennung 4.0 International Lizenz.

1 Einleitung

Sprachliche Kompetenzen werden überall beurteilt¹ – im Unterrichtsalltag ebenso wie in anderen Sprachlehr- und -lernkontexten. Zunehmend kommt dabei standardisierten Tests eine Schlüsselrolle zu, was sich besonders deutlich im Zusammenhang mit dem Aufenthalt von Migrant:innen in Deutschland und dem Hochschulzugang zeigt: Der Zugang zu Ausbildung, Beruf oder gesellschaftlicher Teilhabe hängt maßgeblich davon ab, dass bestimmte Sprachtests erfolgreich absolviert werden – Tests, deren Struktur und Herausforderungen Lehrende kennen sollten. Aber auch im Unterricht selbst müssen Lehrkräfte beurteilen, und dafür eigene Beurteilungsaufgaben entwickeln und auswerten können.

Gleichzeitig erfährt das Thema in der Lehrkräfteausbildung oft nur marginale Aufmerksamkeit. So haben wir es nicht nur mit einem deutlichen Mangel an qualifizierten DaF-Lehrkräften zu tun (vgl. Auswärtiges Amt 2025: 12ff.), sondern es gibt auch an Universitäten bislang kaum genügend Fachkräfte, die zukünftige Lehrende fundiert für die Nutzung und Entwicklung von Beurteilungen ausbilden könnten. Vor diesem Hintergrund liegt die Notwendigkeit von Aus- und Fortbildungen im Bereich der Beurteilung von Sprachkompetenzen auf der Hand. Dem soll das neue Modul 11 des Dhoch3-Programms wenigstens teilweise nachkommen. In diesem Artikel stellen wir das Modul und seine Besonderheiten vor und befassen uns dann vor dem Hintergrund noch fehlender Evaluierungen allgemeiner mit der Frage, welche Bedingungen erfüllt sein müssen, damit ein Dhoch3-Modul für seine Zielgruppen nützlich und wirksam sein kann.

Der Beitrag ist in zwei Teile gegliedert. In Kapitel 2 stellen wir das im März 2025 neu eingeführte Modul 11 zur Beurteilung von Sprachkompetenzen vor und erläutern zwei leitende Konzepte seiner Entwicklung – *Language Assessment Literacy* und *Local Language Testing*. In Kapitel 3 nutzen wir die Gelegenheit, etwas grundlegender über die Nützlichkeit und Wirksamkeit von Dhoch3-Modulen zu reflektieren und erste Vorschläge für ihre Erforschung zu formulieren. In Kapitel 4 diskutieren wir dann mögliche Perspektiven für die zukünftige Entwicklung und verweisen auf offene Forschungslücken.

2. Das Dhoch3-Modul 11: Beurteilen von Sprachkompetenzen in Deutsch als Fremdsprache

Kaum eine Form des Erlernens einer Fremd- oder Zweitsprache, auch des Deutschen, läuft ab, ohne dass die Kompetenzen der Lernenden (oder der Lehrenden) an gewissen Stellen beurteilt werden. Lehrkräfte beobachten ihre Lernenden während ihres Lernprozesses, greifen bei Schwierigkeiten ein und honorieren Lernfortschritte. Beim selbstgesteuerten Lernen schätzen viele Lernende ihren Lernfortschritt regelmäßig selbst ein, sei es aus eigenem Antrieb oder z. B. mithilfe eines Portfolios. In den meisten Lehr- und Lernkontexten sind zusätzlich formalisiertere Beurteilungen üblich, bei denen Lernende explizit zeigen sollen, was sie mit der Sprache tun können und/oder über sie wissen.

¹ Zur Verwendung des Begriffs „Beurteilung“ siehe Kapitel 2.

Sie erhalten dafür ein mehr oder weniger formales Feedback, eine Note oder auch ein Zertifikat. Diese breite Spannweite spiegelt sich in der Definition von Beurteilung wider, die wir in Modul 11 verwenden: Beurteilungen sind systematische, begründete Verfahren, mit deren Einsatz Informationen zum Wissen über Sprache oder die Fähigkeit, Sprache zu verwenden, zu einem bestimmten Zweck erhoben und weiterverarbeitet werden. Diese Definition basiert auf dem Language-Assessment-Begriff von Bachman & Palmer (2010: 20), den wir in Anlehnung an den Gemeinsamen Referenzrahmen für Sprachen (GER, Europarat 2001, 2020) als „Beurteilung“ übersetzen (daneben werden in der Fachliteratur auch andere Begriffe verwendet, u. a. „Leistungsmessung“, „Evaluation“ oder auch das Begriffspaar „Prüfen und Testen“). Beurteilung umfasst neben standardisierten Tests (d. h. formellen Formen der Beurteilung) wie den Goethe-Zertifikaten oder dem TestDaF also auch unterrichtsbezogene Formen des Beurteilens. Zu letzteren gehören der wöchentliche Wortschatztest ebenso wie die schuleigene Jahresabschlussprüfung, Feedback der Lehrperson zu einem vorgetragenen Rollenspiel oder die Korrektur eines schriftlichen Textes. Entscheidend ist, dass die Beurteilung einen bestimmten, zuvor festgelegten Zweck verfolgt, und dass die Ergebnisse der Beurteilung weiterverarbeitet werden, also gewisse Konsequenzen haben.

Der effektive (und effiziente) Einsatz all dieser Formen von Beurteilung verlangt von Lehrkräften ein hohes Maß an Fachwissen und -kompetenz: Aufgaben müssen gezielt ausgesucht oder entwickelt werden, ihre Bewertung muss ebenso zweckmäßig wie informativ sein und Ansprüchen wie Fairness, Praktikabilität und Validität genügen (vgl. z. B. Green 2021). Dabei sollte die Beurteilung prinzipiell dem Lernen förderlich sein, die Lernenden nicht belasten und ein möglichst repräsentatives Abbild der getesteten Fähigkeiten oder Inhalte liefern. Bei Beurteilungen, deren Konsequenzen schwerer wiegen, z. B. für den Zugang zu einer Aufenthaltserlaubnis, aber auch bei Einstufungstests oder Qualifikationsprüfungen, müssen eine Reihe weiterer Gütekriterien erfüllt sein (vgl. Bärenfänger & Geist 2021: 303; siehe auch Kane 2013). Solche Beurteilungsinstrumente – meist Tests – werden üblicherweise von Institutionen entwickelt, die darauf spezialisiert sind. Lehrpersonen müssen in diesem Kontext vor allem einschätzen können, ob und welche dieser Tests für ihre Lernenden angemessen sind, wie sie sie darauf vorbereiten, und wie die Testergebnisse zu interpretieren sind.

Dieses Verständnis von Beurteilung und von den Kompetenzen, die Lehrkräfte mitbringen sollten, um Beurteilungen angemessen einsetzen zu können, steht im Zentrum des neuen Dhoch3-Moduls 11 „Beurteilen von Sprachkompetenzen in Deutsch als Fremdsprache“. Es orientiert sich am Konzept der Language Assessment Literacy (LAL), das in der Sprachtest- und -beurteilungsforschung seit einiger Zeit intensiver diskutiert wird (z. B. Taylor 2009; Kremmel & Harding 2020; Vogt, Böhn & Tsagari 2024). Die potenziell weltweite Einsetzbarkeit des Moduls stellt außerdem Anforderungen an die Inhalte und Materialien, die im Zusammenhang mit der Strömung des Local Language Testing diskutiert werden (Yan, Dimova & Ginther 2023). Beide Konzepte werden in 0 näher erläutert, woraufhin in 0 kurz die Struktur und Inhalte des neuen Moduls vorgestellt werden. In 0 zeigen wir auf, wie wir verschiedenen Herausforderungen bei der Entwicklung des Moduls begegnet sind.

1.1 Assessment Literacy und Local Language Testing: Eine Begriffsklärung

Zur Entwicklung des Moduls 11 gehörte u. a. eine Klärung, was DaF-Lehrkräfte weltweit zum Thema „Beurteilen“ wissen und können sollten. Diese Frage ist der Kern der sogenannten *Assessment Literacy*. Ursprünglich wurde der Begriff in den USA im Zusammenhang mit der Forderung entwickelt, dass Lehrpersonen ein gutes (d. h. kritisches) Verständnis von standardisierten Tests und Testergebnissen benötigen (Stiggins 1991). Heute beschreibt er das „level of knowledge, skills, and understanding of assessment principles and practice that is [...] required by [...] test stakeholder groups, depending on their needs and context“ (Taylor 2009: 24). *Assessment Literacy* umfasst also sowohl das Wissen, welches z. B. Arbeitgeber:innen benötigen, um die Aussagekraft von (Schul-)Noten oder Testergebnissen adäquat einschätzen zu können, als auch die Fähigkeit von Lehrpersonen, in ihrem Unterricht Beurteilungen von hoher Qualität zu entwickeln und anzuwenden. Auch angemessene Rückmeldungen an Lernende und ein Bewusstsein für mögliche (positive und negative) Konsequenzen des Beurteilens gehören zu einer gut entwickelten *Assessment Literacy* von Lehrpersonen (siehe hierzu auch Vogt 2018: 6ff.; Inbar-Lourie 2017). Diese wird jedoch oft in der Lehrpersonenaus- und -weiterbildung nicht oder nur am Rande thematisiert (Vogt 2018: 10f.).

Was genau diese Kompetenzen und Wissensbereiche umfassen, wurde zwar bislang für DaF/DaZ nicht empirisch beschrieben, aus dem Bereich des Englischen (als Fremdsprache) gibt es aber allgemeine Modelle für *Assessment Literacy* sowie erste empirische Studien. So definieren z. B. Xu und Brown (2016) *Assessment Literacy* als einen andauernden Kompromiss: TALiP (Teacher Assessment Literacy in Practice) wird beschrieben als „a dynamic, complex entity combining teachers’ assessment knowledge, their conceptions of assessment, and their responses to the external contexts embedded with actual constraints and affordances in the environment“ (ebd.: 157). Lehrpersonen treffen Entscheidungen im Zusammenhang mit Beurteilungen aufgrund dieses komplexen Zusammenspiels – dementsprechend sollte die Förderung von *Assessment Literacy* auch auf diesen verschiedenen Ebenen stattfinden: Neben der Vermittlung von Wissen sollten Lehrpersonen auch ein Bewusstsein für die Bedeutung ihrer eigenen Wahrnehmung und Interpretationen sowie Kenntnisse über den eigenen Mikro- und Makrokontext entwickeln und kritisch hinterfragen (siehe auch Inbar-Lourie 2017: 7). Dem trägt das neue Dhoch3-Modul insofern Rechnung, als es viele Angebote für praktische Anwendungen und Reflexionen enthält, die durch zusätzliche Inputs ergänzt werden. Damit soll die Nützlichkeit der Inhalte für die eigene Praxis in den Vordergrund gestellt werden.

Die eigene Praxis steht auch im Zentrum des sog. *Local Language Testing*. Das Konzept unterscheidet lokale Tests von standardisierten Tests, die weltweit ohne Anpassung an lokale Bedingungen eingesetzt werden (z. B. die Goethe-Zertifikate, das Österreichische Sprachdiplom (ÖSD) oder TestDaF). „Lokal“ werden Beurteilungsinstrumente dadurch, dass sie von Stakeholdern vor Ort (mit)verantwortet werden, sei es bei ihrem Design, ihrer Anwendung und/oder ihrer Interpretation: „We define a local test as a test whose development is designed to produce scores to

help solve a local decision-making problem which emerges out of a need within the local context in which the test will be used.” (Dimova, Yan & Ginther 2022: 342). Solche Beurteilungsinstrumente zeichnen sich idealerweise dadurch aus, dass sie in Inhalt, Format und Durchführung (besser) zum Unterrichts- und Lernkontext der Testteilnehmenden passen (ebd.: 1ff.). Durch diese Nähe zum lokalen Unterrichtskontext werden auch Konzepte wie formative Beurteilung, Selbstbeurteilung und Peer Assessment wieder stärker diskutiert (ebd.: 344f.).

An dieser Stelle schließt sich der Kreis: Um gute Beurteilungsinstrumente für den eigenen Kontext zu entwickeln, Beurteilungen gezielt im Unterricht einzusetzen, aber auch um Lernende adäquat auf das Ablegen standardisierter Beurteilungsinstrumente vorzubereiten, müssen die beteiligten Personen (v. a. Lehrende und Lernende, aber auch Programmverantwortliche oder Forschende) über die nötige Language Assessment Literacy verfügen, die den Kern von Modul 11 bildet.

1.2 Überblick über Modul 11

Das neue Modul orientiert sich in Aufbau, Struktur und Gestaltung an bereits bestehenden Dhoch3-Modulen. Seine primären Nutzer:innen sind Personen, die DaF-Lehrkräfte ausbilden, also v. a. Hochschullehrende. Dhoch3-Module sind Moodle-Räume, die kopiert und deren Inhalte für die eigene Lehre angepasst werden können. Modul 11 ist in sieben Themen untergliedert, die weitgehend unabhängig voneinander genutzt werden können (siehe Tab. 1).

1	Einführung und Aktivierung von Vorwissen
2	Grundlagen des Beurteilens
3	Was wird beurteilt? Wichtige Konstrukte in DaF
4	Wie wird beurteilt? Typen von Beurteilungen
5	Beurteilungsinstrumente entwickeln (Projektarbeit)
6	Qualitätssicherung
7	Beurteilung und der Gemeinsame Europäische Referenzrahmen (GER)

Tab. 1: Themen im Modul 11

Innerhalb dieser Themen werden jeweils Materialien zur Vorbereitung der Lehre (Baustein 1), für die Lehr- und Lernpraxis (Baustein 2) und für die Vertiefung des wissenschaftlichen Interesses (Baustein 3) zur Verfügung gestellt. Im Falle von Modul 11 sind dies meist Power-Point-Präsentationen oder andere Inputformen (z. B. Texte, Videos, Tabellen), die im Unterricht nach Belieben präsentiert und/oder den Teilnehmenden zur Verfügung gestellt werden können. Außerdem enthält das Modul viele Arbeitsblätter für die Einzel-, Partner- oder Gruppenarbeit, die in einer Version für die Kursleitenden jeweils didaktisch kommentiert wurden (z. B. Hinweise auf Anpassungsmöglichkeiten, Erweiterungen oder alternative Aktivitäten). Baustein 3 enthält vor allem weiterführende Lektüre und Literaturhinweise. Hier finden sich auch einige zentrale Texte des Forschungsgebiets, die der DAAD für die Nutzung in Dhoch3 lizenziert und damit für Nutzer:innen

zugänglich gemacht hat, die sonst nur eingeschränkt Zugriff auf deutschsprachige Fachliteratur haben.

1.3 Förderung von kontextbezogener Language Assessment Literacy in Modul 11

Die Entwicklung von lokal anwendbarer Language Assessment Literacy für DaF-Lehrkräfte steht im Zentrum des neuen Dhoch3-Moduls. Es enthält Materialien, die in Aus- und Weiterbildungsangeboten (online und offline) genutzt werden können, um (angehenden) Lehrpersonen Wissen, Kompetenzen und Strategien im Bereich des Beurteilens zu vermitteln. Diese Materialien wurden so entwickelt, dass Kursleitende jederzeit die Möglichkeit haben, eigene Inhalte oder Aktivitäten zu ergänzen, um ihren Unterricht auf die Bedürfnisse der auszubildenden Lehrpersonen und den lokalen Kontext auszurichten. Zentrales Werkzeug sind dabei die Lernziele, denen sämtliche Aktivitäten und Materialien zugeordnet sind. So sollen die Lehrkräfte z. B. in die Lage versetzt werden, „in Grundzügen [zu] beschreiben, was ein Konstrukt, ein Indikator und eine Operationalisierung sind“ (Lernziel 2.2.4). Dieses Lernziel bringt die Notwendigkeit zum Ausdruck, genau zu definieren, *was* beurteilt werden soll, und sich darüber klar zu werden, ob und *wie* die Beurteilungsaufgaben und -aktivitäten dies abbilden (d. h. operationalisieren). Dem Konstrukt und seiner Operationalisierung ist dann ein ganzes Thema im Modul gewidmet (Thema 3), in dem die Teilnehmenden anhand vielfältiger Beispiele und multimodaler Inputs, z. B. Expertinnenvideos, ihr Verständnis vertiefen. Ein weiteres Lernziel bezieht sich explizit auf den Umgang mit standardisierten Beurteilungen: „[Die Kursteilnehmenden] können (internationale) standardisierte Beurteilungsinstrumente für ihre konkrete Zielgruppe benennen und dazugehörige Informationsmaterialien zur Vorbereitung ihrer Lernenden nutzen“ (Lernziel 4.5). Die Materialien zu diesem Lernziel beinhalten eine Liste einschlägiger internationaler Beurteilungsinstrumente für DaF. Aus dieser wählen die Kursteilnehmenden drei aus, die gut zu einer zuvor von ihnen beschriebenen Zielgruppe passen. Daran schließt sich eine Internetrecherche an, bei der Informationsmaterialien zu den ausgewählten standardisierten Tests gesucht und gesichtet werden sollen.

Durchgängig arbeitet das Modul zudem mit drei Anwendungskontexten, die in einer Reihe von parallelen Aufgaben beispielhaft bearbeitet werden. Grundlage ist die Annahme, dass viele Nutzer:innen des Moduls Lehrpersonen für einen der folgenden Kontexte ausbilden: 1) schulischer DaF-Unterricht außerhalb des deutschsprachigen Raums, 2) Sprachkurse für berufliche Kontexte, 3) Sprachkurse für den Zugang zu deutschsprachigen Hochschulen. In den betreffenden Aktivitäten werden Beurteilungen und Beurteilungsaufgaben behandelt, die im jeweiligen Kontext typisch sind, und es werden an diesen Beispielen Überlegungen modelliert, die die Lehrpersonen dann für ihre eigene Unterrichtsrealität übernehmen können. So wurde ein (notwendiger) Spagat versucht, denn die große Vielfalt existierender Beurteilungspraktiken lässt sich nicht annähernd exhaustiv mit konkreten Beispielen veranschaulichen.

Die Inhalte und Materialien des Moduls bedürfen also der Anpassung durch die Kursleitung bzw. einer hohen Reflexionsleistung durch die Kursteilnehmenden. Obwohl dies eine Herausforderung darstellt, gehen wir gleichzeitig davon aus, dass es eine Stärke des Moduls ist: Language Assessment Literacy kann nur entstehen, wenn die Nutzer:innen von Beurteilungsinstrumenten über ausreichend Reflexionsfähigkeit verfügen (z. B. Yan & Fan 2021: 238; Xu & Brown 2016: 157f.). Das Modul soll dazu beitragen, diese weiterzuentwickeln. Damit dies überhaupt geschehen kann, muss es genutzt werden. Welche Bedingungen dafür erfüllt sein müssen, wird im Folgenden diskutiert.

2. Zur Förderung der Nützlichkeit und Analyse der Wirksamkeit von Dhoch3-Modulen

Kapitel 3 widmet sich der grundlegenden Frage, unter welchen Bedingungen unser (und andere) Dhoch3-Modul⁶ nützlich und wirksam sein kann bzw. können und wie sich dies untersuchen ließe. Es werden konzeptionelle Vorüberlegungen ausgehend von inhaltlichen Schwerpunkten und basierend auf Rückmeldungen aus ersten Modul-Workshops entwickelt; eine eigentliche Wirkungsanalyse ließ sich jedoch noch nicht durchführen. Die Reflexionen sind deshalb als tentativ zu verstehen.

Die Frage danach, wie die Nutzung von Dhoch3-Modulen – notwendige Voraussetzung für die Entfaltung von Wirkung – angeregt werden kann, vor allem aber die noch komplexere danach, ob und wie ein Modul dann tatsächlich wirkt, wurden u. E. bislang kaum adressiert. Im Folgenden möchten wir deshalb vorsichtige Überlegungen dazu anstellen, welche Faktoren eine wirksame Nutzung von Modul 11 begünstigen oder auch hemmen könnten und über mögliche zukünftige Wirkungsanalysen reflektieren. Dabei nennen wir Aspekte, die nicht nur für Modul 11 relevant sind und teils auf strukturelle Herausforderungen und Chancen des Gesamtprogramms verweisen (vgl. z. B. Riemer & Soethe 2023 zu erforderlichen Sprachkenntnissen; Fathy & Buchholz 2023 zu Multiplikator:inneninitiativen). Die besondere Charakteristik des Kompetenzfelds „Beurteilen“ und sein Stellenwert in der Lehrkräfteausbildung im deutschsprachigen Raum lassen dabei einige Dilemmata brennglasartig besonders deutlich zutage treten.

2.1 Nützlichkeit von Dhoch3-Modulen: Plädoyer für begleitete Implementierungen

Der alltagssprachliche Terminus „Nützlichkeit“ ist nicht sauber von dem der „Wirksamkeit“ zu unterscheiden. Wir gebrauchen ihn als Hilfsbegriff für Aspekte, die in einer Evaluierungslogik einer komplexeren Analyse von Wirkung vorausgehen würden. Nützlichkeit betrifft in diesem vorläufigen Verständnis die Herstellung der Bedingungen für Wirksamkeit auf Implementierungsebene: Wird ein Dhoch3-Modul nicht genutzt, kann es erst gar keine Wirkung entfalten. Eigentlich *nützlich* ist es letztendlich aber nur dann, wenn es auch in die vorgesehene Richtung *wirkt*, also erwünschte Veränderungen anstößt. Ein wirksames Modul hat zudem größere Chancen, an weiteren Standorten

genutzt zu werden. Wirksamkeit und Nützlichkeit hängen also sehr eng miteinander zusammen und überlappen sich erheblich.

Wir betrachten Nützlichkeit als gegeben, wenn ein Modul wahrgenommen und eingesetzt wird. Die vorliegende Dhoch3-Literatur spiegelt diesbezüglich viel Engagement: In vielen Einzelstudien wurden Vorschläge zur Verwendung einzelner Module in unterschiedlichen Kontexten unterbreitet und der DAAD ergriff in den verschiedenen Programmphasen Maßnahmen zur Disseminierung der Dhoch3-Module. Auch die Anregung von Erprobungsworkshops und weitere Aktivitäten im Rahmen von Modulentwicklungen und -überarbeitungen tragen dazu bei, dass Module nicht im leeren internationalen Raum schweben.

Ob ein Angebot letztendlich genutzt wird, hat wohl auch inhaltliche Gründe. Die Chancen für eine (nachhaltige) Nutzung erhöhen sich sicherlich, wenn Modulinhalt von Zielgruppen für relevant gehalten werden, lokale Bedarfe aufgreifen, dabei aber nicht redundant sind. Zusätzlich ist zu berücksichtigen, dass das zu vermittelnde neue Wissen an schon bestehende grundlegende Wissensbestände anknüpfen können muss, damit die Module eine realistische Nutzungschance besitzen. Dass dies besser gelingen kann, wenn Module grundlegend partnerschaftlich konzipiert und auch implementiert werden, liegt auf der Hand. Ebenso spielen institutionell-strukturelle Gegebenheiten eine wichtige Rolle, z. B. verfügbare personelle Ressourcen oder Anschlussmöglichkeiten an bestehende curriculare Strukturen. Schließlich sind Lehrkräfte die zentralen Akteur:innen, die Dhoch3-Modulen Leben einhauchen können. Schart & Ohta (2018) empfehlen, Fortbildungen „von den Lehrenden her [zu] denken“ und nennen u. a. die Prinzipien Subjektivität (Berücksichtigung individueller Normen und Überzeugungen), Praxisrelevanz, Erfahrung (Möglichkeiten zum Ausprobieren) und Multiperspektivität.

Der Nützlichkeit von Modul 11 stehen einige Hemmnisse im Weg. So werden einige Dhoch3-Module von einigen Zielgruppen schon themenunabhängig als „zu schwierig“ eingeordnet (vgl. Riemer & Soethe 2023: 27), was für Modul 11 mit seinem bereits im deutschsprachigen Raum als anspruchsvoll wahrgenommenem Thema ebenso gelten dürfte. Zusätzlich werden Lehrende wenig Vorwissen mitbringen; es fehlt ihnen also nicht nur passendes Material, um Lehrveranstaltungen zum Thema „Beurteilen“ anbieten zu können. Damit ist die erste Professionalisierungsaufgabe (vgl. Abb. 1) besonders wichtig: Lehrende sind in besonderem Maße eine zunächst selbst fortzubildende Zielgruppe. Bei der Modulerstellung wurde daher versucht, niedrigschwellige und auch sprachlich nicht zu schwierige Materialien zu entwickeln, sodass es möglich ist, auch ohne jedes Vorwissen mit Modul 11 ins Thema einzusteigen und sich solides Grundwissen zu erarbeiten.

Wichtige Voraussetzungen für die Nützlichkeit des Moduls sind die Komplementarität zu bestehenden Angeboten und die Relevanz für standortspezifische Studiengänge. Nun ist weniger davon auszugehen, dass Modul 11 Redundanz schafft. Wahrscheinlich ist aber, dass nur Standorte mit einem ausreichend breiten Lehrangebot in der Lage und interessiert sein werden, Modul 11 zu nutzen. Wo DaF als eigenes Fach existiert, ist dies wesentlich wahrscheinlicher als in reinen Germanistik-Studiengängen mit DaF-Anteilen oder auch in kombinierten Studiengängen wie etwa

Wirtschaftswissenschaften mit Deutschausbildung. Strukturelle Bedingungen der Hochschulen spielen also für die Modulnützlichkeit eine entscheidende Rolle.

Eine Herausforderung, die wir für Modul 11 und strukturell ähnliche Module sehen, ist die bislang weitgehende Auskoppelung der Modulimplementierung aus den Entwicklungsprojekten: Der eigentliche Einsatz der Module wird bislang nicht systematisch und prozessorientiert begleitet und dadurch auch nicht näher erforscht. Eine entsprechende Begleitforschung würde aber eine systematische Wirkungsanalyse vorbereiten helfen und diese erheblich erleichtern (vgl. 3.2). Je spezifischer die Modulthemen sind, desto eher wird für deren Nutzung (und Nützlichkeit) eine begleitende Unterstützung nötig sein. Um standortspezifisch bedarfsgerechte, realistisch umsetzbare Implementierungen (von Modul 11) zu ermöglichen, könnten u. E. an – vereinfacht gesagt – einzelnen Standorttypen (z. B. Hochschule mit DaF-Studiengang; typischer Germanistikstudiengang mit DaF-Elementen usw.) modellhafte begleitete Implementierungen mit der Unterstützung des Modulteams umgesetzt werden. Die entsprechenden Prozesse und Dokumentationen könnten in die Moodle-Plattform eingefügt werden, würden so vielleicht für eine insgesamt verstärkte Partnerschaftlichkeit sorgen und eventuell auch die Bereitschaft strukturell vergleichbarer Standorte erhöhen, das Modul zu nutzen. Solche Modellimplementierungen wären dann zwar, wie Workshopformate, punktuell, bergen aber unseres Erachtens das Potenzial für die Übertragbarkeit in institutionell ähnliche Kontexte.

2.2 Desiderat: Wirkungsanalysen zu Dhoch3-Modulen

Die Wirkung von Dhoch3-Modulen ist bislang empirisch nicht untersucht (vgl. aber Qian, Zhao & Zheng 2022) und wäre konzeptionell und methodisch komplex. Die tatsächliche Nutzung ist Voraussetzung für die Entfaltung von Wirkungen und insofern Bestandteile einer Wirkungsanalyse. Daher müsste es zunächst darum gehen, zu definieren, welche Domäne der Wirkungslogik im Zentrum stehen soll. Abb. 1 illustriert die doppelte Professionalisierungsaufgabe der Dhoch3-Module: Sie wirken sowohl auf die Kursleitenden, die das Material in ihrer Lehre einsetzen als auch auf die zu unterrichtenden (angehenden) DaF-Lehrkräfte. Die Lehrkräfte prägen dann ihre unterschiedlichsten DaF-Praxiskontexte.

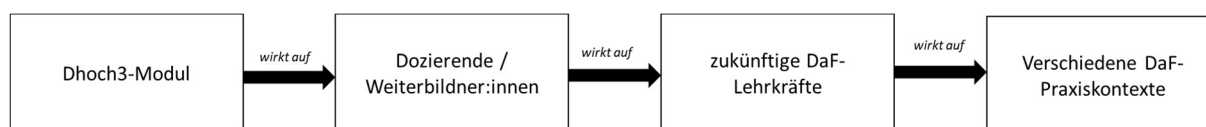


Abb. 1: Doppelt mittelbare Wirkungsentfaltung von Dhoch3-Modulen auf Lehrkräfte und Lehrpraxis

Daneben sind weitere Stakeholdergruppen und Wirkungsebenen betroffen, bspw. Veränderungen auf institutionellen Ebenen. Eine Wirkungsanalyse müsste all diese Veränderungen benennen und sauber trennen. Ferner ist plausibel, dass sich die realistisch erwartbaren Wirkungen je nach institutionellem Kontext unterscheiden. Sinnvoll könnte es etwa sein, für einige wenige „typische“ Einsatzkontexte von Dhoch3 (z. B. Germanistikstudiengang, DaF-Studiengang, Weiterbildung) gesonderte Wirkungslogiken zu formulieren.

Dann müssten Wirkungsannahmen formuliert, mithilfe relevanter, realistisch erhebbarer Indikatoren operationalisiert und mit geeigneten Methoden untersucht werden. Allgemeinere Indikatoren lassen sich aus den Programmzielen des DAAD ableiten². An vorderster Stelle steht dabei die Ausbildung angehender Lehrkräfte an Hochschulen. Dhoch3 soll offen, komplementär zu bestehenden Curricula und Strukturen und an eigene Bedarfe lokal adaptierbar sein (vgl. auch Schmälting 2018). Zudem soll die Digitalisierung vorangetrieben und die internationale Vernetzung von DaF-bezogenen Studiengängen intensiviert werden. Zentral für die Definition von Indikatoren ist zum anderen die fachliche Professionalisierungsperspektive. Sicherlich wäre die Anwendung eines Professionalisierungsmodells für eine systematische Evaluation sinnvoll (z. B. Feike et al. 2023). Was Professionalisierung im Fach DaF/DaZ bedeutet, ist jedoch umstritten (Feike et al. 2023; Legutke & Schart 2016), nicht zuletzt wegen der Heterogenität der DaF-bezogenen bzw. germanistischen Hochschullehre weltweit. Peuschel und Wittmann (2025: 178) konstatieren, dass allgemeinere Professionalisierungsansätze meist auf Kosten von Standort- und Zielgruppenspezifika gehen, eine zu starke Zuspitzung jedoch umgekehrt eine Modellbildung genauso erschweren würde wie systematische diesbezügliche Forschung. Dieser Punkt ist auch für Dhoch3-Module zentral, für die die Anschlussfähigkeit an sehr heterogene Standorte ein zentraler Nützlichkeitsfaktor ist. Jede Wirkungsanalyse zu Dhoch3 würde sich in einem Spannungsverhältnis zwischen standortspezifischen Studienkontexten, Anforderungen, Gegebenheiten und Zielen einerseits, dem Desiderat eines gemeinsamen Verständnisses von Lehrkräfteprofessionalisierung andererseits bewegen³.

Konkreter auf Modul 11 bezogen böte sich für die Analyse von Wirkungen auf Ebene der Lehrpraxis (vgl. Abb. 1, ganz rechts) der Rekurs auf die drei konkreten Anwendungskontexte an (vgl. 2.3). Um den individuellen Kompetenzzuwachs von Dozierenden und angehenden Lehrkräften zu untersuchen, sind die feingranularen Lernziele von Nutzen (ebd.). Eine ganze Reihe an quantitativen und qualitativen Methoden ist dafür denkbar, von Befragungen über Selbstaussagen bis hin zu reflektierenden Fokusgruppeninterviews. Empfehlenswert scheint ein gemischtmethodischer Ansatz, um mit der sehr großen erwartbaren Heterogenität der Kontexte umgehen zu können. Von zentraler Bedeutung im gesamten Prozess der Wirkungsanalyse wäre ein partizipativer Ansatz: Die Perspektiven der internationalen Kolleg:innen müssen eine mindestens genauso wichtige, konstitutive Rolle bei der Formulierung von Wirkungsannahmen und der Auswahl sinnvoller Indikatoren spielen wie die der Modulteams, um nicht an tatsächlichen Bedarfen und Gegebenheiten vorbeizuarbeiten.

² <https://www.daad.de/de/der-daad/querschnittsdimensionen-themen/die-deutsche-sprache-foerdern/das-projekt-dhoch3/>

³ Es lassen sich etliche weitere Punkte nennen, die ein Dhoch3-Modul mehr oder weniger nützlich und wirksam machen können. Regt ein Dhoch3-Modul Fachdiskurse (standortübergreifend) und kritischen Austausch an, in unserem Fall bspw. über die Rolle von Beurteilungen in DaF, trägt es darüber auch zur Ausprägung einer professionellen kritisch-reflexiven Haltung bei, wirkt es u. E. in eine höchst wünschenswerte Richtung. Idealerweise passiert dies nicht nur bei den Nutzer:innen des Moduls weltweit, sondern auch im deutschsprachigen Raum, zuallererst bei den Modulentwickelnden. Letztere kommen durch das Modul in Kontakt mit diversen Standortspezifika und heterogenen Lehr-Lernbedingungen für DaF weltweit und adjustieren so auch ihren eigenen Blick auf das jeweilige Modulthema. Über Publikationen sollten diese Erkenntnisse der ganzen Community bekanntgemacht werden bzw. wiederum in Curricula in deutschsprachigen Ländern einfließen.

3 Fazit

Eine Besonderheit des Themenfelds „Beurteilen“ ist seine dilemmatische Stellung zwischen erheblicher praktischer Relevanz einerseits und fachlicher Randständigkeit andererseits. So werden selbst in DaF/DaZ-Studiengängen des deutschsprachigen Raums nur in wenigen Ausnahmefällen fest curricular verankerte Module zu diesem Thema angeboten. Fragen der Beurteilung erfahren geringere Aufmerksamkeit als andere, eher als Kernbausteine empfundene Gegenstände. Außerdem ist die Zahl der DaF/DaZ-Kolleg:innen, die sich wissenschaftlich mit Fragen des Beurteilens auseinandersetzen, gering. Neben dem breiten methodischen Wissen, das nötig ist, um die Qualität von Beurteilungen einzuschätzen oder selbst Verfahren zu entwickeln, setzt der Themenbereich Wissen zu sprachlichen Kompetenzen und deren Aneignung als möglichen Beurteilungsgegenständen schon voraus; zudem findet ein großer Teil der entsprechenden Fachdiskussionen in englischer Sprache statt. Ein in der Praxis als omnipräsent wahrgenommenes Phänomen, das sehr häufig (und oft zu Recht) auch kritische Aufmerksamkeit erfährt und ein zentrales Handlungsfeld für DaF/DaZ-Lehrkräfte in ihrem Unterrichtsalltag ist, erweist sich somit schon im deutschsprachigen Raum als sowohl in der Hochschulausbildung als auch in der Forschung äußerst spärlich repräsentiert. Es ist aus dieser Perspektive also begrüßenswert, dass der DAAD nun diesem Themenbereich ein Dhoch3-Modul widmet. In unserem Artikel beschreiben wir einige Merkmale und Besonderheiten des Moduls, v. a. seine Organisation entlang von feingranularen Lernzielen und den durchgängigen Bezug auf drei Anwendungskontexte. Ferner werden zwei Hauptprinzipien näher erläutert, nämlich die *Language Assessment Literacy* und das *Local Language Testing*.

Im Anschluss geht der Beitrag auf Bedingungen für die tatsächliche Nutzung von Dhoch3-Modulen ein, die als Voraussetzung für umfassendere Wirkungen verstanden werden kann (vgl. 3.1). Angesichts einiger praktischer Hindernisse schlagen wir hier begleitete, prozessorientierte Modell-Implementierungen vor. Auch der Ablauf möglicher Wirkungsanalysen wird zumindest in groben Zügen skizziert. Wirkungsanalysen sind ein dringendes Desiderat und erfordern eine anspruchsvolle Auseinandersetzung mit Wirkungsannahmen sowie die Festlegung geeigneter Indikatoren. Dabei sollten Programmziele, Aspekte der fachlichen Professionalisierung und dezentrale internationale, standortgeleitete Perspektiven gleichermaßen eine wichtige Rolle spielen. Wirkungsanalysen sollten also mehrperspektivisch verschiedene Stakeholdergruppen und Wirkungsebenen adressieren und dabei mit unterschiedlichen, möglichst triangulierten gemischt-methodischen Designs vorgehen. Eine grundlegend gleichberechtigt-partizipative Anlage ist unabdingbar. Zwar dürfte es eine Herausforderung sein, im vorliegenden, politisch-strategisch gesteuerten, heterogenen Handlungsfeld Wirkungsanalysen durchzuführen, unseres Erachtens ist das Dhoch3-Programm aber so ausgereift und gut etabliert, dass nun solche Anstrengungen unternommen werden können (vgl. das wirkungsorientierte Monitoring des DAAD⁴).

⁴ <https://www.daad.de/de/der-daad/wer-wir-sind/daad-evaluationen/>

Zusammenfassend halten wir in unserer vorsichtigen Annäherung ein Modul für potenziell nützlich und wirksam, das an unterschiedlichen Standorten anschlussfähig ist, aus den Materialien heraus Gestalt annimmt und in der Praxis „lebt und atmet“. Idealerweise sollte es verschiedenartiges Anschlusshandeln anregen, also kein bloßer ‚Steinbruch‘ an Materialien sein, sondern organisch genutzt und transformiert werden und so auch den Charakter des Fachs insgesamt zu transformieren vermögen – wenn auch indirekt, in kleinen Schritten und langsam.

Literaturverzeichnis

- Auswärtiges Amt (Hrsg.) (2025). Deutsch als Fremdsprache weltweit 2025. Perspektiven, Herausforderungen, Strategien. Auswärtiges Amt. <https://www.auswaertiges-amt.de/resource/blob/2728424/f7ee0558008807b8e4b1ae15907d5872/deutsch-als-fremdsprache-weltweit-data.pdf>
- Bachman, Lyle F. & Palmer, Adrian S. (2010). *Language Assessment in Practice: Developing Language Assessments and Justifying Their Use in the Real World*. Oxford, New York: Oxford University Press.
- Bärenfänger, Olaf & Geist, Barbara (2021). Prüfen, Testen, Sprachstände erheben. In: Altmayer, Claus & Biebighäuser, Katrin & Haberzettl, Stefanie & Heine, Antje (Hrsg.). *Handbuch Deutsch als Fremd- und Zweitsprache: Kontexte – Themen – Methoden*. Stuttgart: Metzler, 301–318.
- Dimova, Slobodanka & Yan, Xun & Ginther, April (2022). Local tests, local contexts. *Language Testing* 39 (3), 341–354. DOI: <https://doi.org/10.1177/02655322221092392>
- Europarat (2001). *Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen*. Berlin: Langenscheidt.
- Europarat (2020). *Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen. Begleitband*. Stuttgart: Klett.
- Fathy, Hebatallah & Buchholz, Stefan (2023). Das Programm Dhoch3. Genese, aktuelle Entwicklungen und Zukunftsperspektiven. In: DAAD (Hrsg.). *Deutsch als Fremdsprache in der digitalen Welt: Zu aktuellen Entwicklungen in Lehre und Forschung*. Bonn: Deutscher Akademischer Austauschdienst e.V. (DAAD), 14–21.
- Feike, Julia & Janiková, Věra & Ōta, Tatsuya & Peuschel, Kristina & Saunders, Constanze & Schart, Michael (2023). Professionalisierung erleben, gestalten, begleiten und erforschen: Einleitung zur ersten Ausgabe von KONTEXTE. *Kontexte – Internationales Journal zur*

- Professionalisierung in Deutsch als Fremdsprache* 1 (1), 1–20. DOI: <https://doi.org/10.24403/jp.1296775>
- Green, Anthony (2021). *Exploring Language Assessment and Testing: Language in Action*. 2. Auflage. London, New York: Routledge.
- Inbar-Lourie, Ofra (2017). Language Assessment Literacy. In: Shohamy, Elana & Or, Irit G. & May, Stephen (Hrsg.). *Language Testing and Assessment*. 3. Auflage. Cham: Springer, 257–270.
- Kane, Michael T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement* 50 (1), 1–73. DOI: <https://doi.org/10.1111/jedm.12000>.
- Kremmel, Benjamin & Harding, Luke (2020). Towards a comprehensive, empirical model of language assessment literacy across stakeholder groups: Developing the Language Assessment Literacy Survey. *Language Assessment Quarterly* 17 (1), 100–120. DOI: <https://doi.org/10.1080/15434303.2019.1674855>.
- Legutke, Michael & Schar, Michael (Hrsg.) (2016). *Fremdsprachendidaktische Professionsforschung: Brennpunkt Lehrerbildung*. Tübingen: Narr Francke Attempto.
- Peuschel, Kristina & Wittmann, Gerda-Elisabeth (2025). These H »Professionalisierung von Lehrkräften«. Herausforderungen in der Professionalisierung von DaF-Lehrkräften kooperativ und kritisch-reflexiv begegnen. *Zeitschrift für Deutsch im Kontext von Mehrsprachigkeit* 41 (1–2), 166–182. <https://doi.org/10.14220/zdkm.2025.41.1.166>
- Qian, Chunchun & Zhao, Jin & Zheng, Yu (2022). Berufsbezogener DaF-Unterricht: Evaluation eines Dho3-Moduls. In: Dimova, Dimka & Müller, Jennifer & Siebold, Kathrin & Teepker, Frauke & Thaller, Florian (Hrsg.). *DaF und DaZ im Zeichen von Tradition und Innovation*. 47. Jahrestagung des Fachverbandes Deutsch als Fremd- und Zweitsprache an der Universität Marburg 2020. Göttingen: Universitätsverlag Göttingen, 327–344.
- Riemer, Claudia & Soethe, Paulo (2023). Dho3 und die internationale Zusammenarbeit in der Hochschulbildung. DaF und Germanistik: Chancen und Herausforderungen. In: DAAD (Hrsg.). *Deutsch als Fremdsprache in der digitalen Welt: Zu aktuellen Entwicklungen in Lehre und Forschung*. Bonn: Deutscher Akademischer Austauschdienst e.V. (DAAD), 22–33.
- Schar, Michael & Ohta, Tatsuya (2018). Von den Lehrenden her denken – zur Konzeption von Fortbildungen am Beispiel eines Qualifizierungsangebots für Deutschlehrkräfte in Japan. *Informationen Deutsch als Fremdsprache* 45 (5), 672–694. DOI: <https://doi.org/10.1515/infodaf-2018-0085>
- Schmaling, Benjamin (2018). Dho3: Online-Studienmodule für die Deutschlehrerausbildung an Hochschulen weltweit: kulturspezifische Modifizierbarkeit und Anwendung im Blended-

- Learning-Format. *Informationen Deutsch als Fremdsprache* 45 (5), 635–654. DOI: <https://doi.org/10.1515/infodaf-2018-0083>
- Stiggins, Richard J. (1991). Assessment Literacy. *Phi Delta Kappan* 72 (7), 534–39.
- Taylor, Lynda (2009). Developing assessment literacy. *Annual Review of Applied Linguistics* 29, 21–36. DOI: <https://doi.org/10.1017/S0267190509090035>.
- Vogt, Karin (2018). Teachers' Assessment Literacy Enhancement (TALE): Needs Analysis Report. https://taleproject.eu/pluginfile.php/2127/mod_page/content/2/TALE%20Needs%20analysis%20report.pdf
- Vogt, Karin & Bøhn, Henrik & Tsagari, Dina (2024). Language assessment literacy. *Language Teaching* 57 (3), 325–340. DOI: <https://doi.org/10.1017/S0261444824000090>
- Xu, Yueting & Brown, Gavin T. L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education* 58, 149–162. DOI: <https://doi.org/10.1016/j.tate.2016.05.010>
- Yan, Xun & Dimova, Slobodanka & Ginther, April (2023). *Local Language Testing: Practice across Contexts*. Springer Nature.
- Yan, Xun & Fan, Jason (2021). “Am I qualified to be a language tester?”: Understanding the development of language assessment literacy across three stakeholder groups. *Language Testing* 38 (2), 219–246. DOI: <https://doi.org/10.1177/0265532220929924>

Angaben zu den Personen:

Katharina Karges ist wissenschaftliche Mitarbeiterin am Herder-Institut der Universität Leipzig und assoziierte Wissenschaftlerin des Instituts für Mehrsprachigkeit der Universität Freiburg/Schweiz. Sie ist Mitautorin des Dhoch3-Moduls zum Beurteilen von Sprachkompetenzen in DaF. Zu ihren Forschungsinteressen gehören die Beurteilung (fremd)sprachlicher Kompetenzen, Mehrsprachigkeitsforschung und -didaktik sowie empirische Forschungsmethoden und Korpuslinguistik.

Kontakt: katharina.karges@uni-leipzig.de

Alice Friedland ist Mitautorin des Dhoch3-Moduls 11 zum Beurteilen von Sprachkompetenzen in DaF. Ihre Forschungsinteressen liegen in den Bereichen Washback und Assessment Literacy. Sie arbeitet als Lehrkraft für besondere Aufgaben am Sprachenzentrum und am Institut für Germanistik

und Medienkulturen der Technischen Universität Dresden und als Wissenschaftliche Mitarbeiterin am Herder-Institut der Universität Leipzig.

Kontakt: alice.friedland@uni-leipzig.de

Katrin Wisniewski ist Gerhard-Helbig-Professorin für Deutsch als Fremd- und Zweitsprache am Herder-Institut der Universität Leipzig. Gemeinsam mit Professorin Gabriele Kniffka war sie für die Projektleitung zum Dhoch3-Modul „Beurteilen von Sprachkompetenzen in DaF“ zuständig. Ihre Schwerpunkte sind das Assessment von Sprachkompetenzen, der Erwerb des Deutschen mit Fokus Grammatik, die Arbeit mit Lernerkorpora und der Zusammenhang sprachlicher Kompetenzen mit Bildungs- und Studienerfolg.

Kontakt: karin.wisniewski@uni-leipzig.de

DOI: <https://doi.org/10.24403/jp.1540575>