

## Evaluationsverfahren der türkischen Deutschlehrkräfte im DaF-Unterricht

Rahim Şentürk , Edirne

 <https://doi.org/10.37583/diyalog.1951623>

### *Abstract (Deutsch)*

Diese Studie untersucht, welche Evaluationsverfahren türkische Deutschlehrkräfte im DaF-Unterricht anwenden und welche Verfahren sie bevorzugen. Ziel ist es, Bewertungspraktiken zu erfassen, deren Stärken und Schwächen aufzuzeigen und Empfehlungen zur Verbesserung der Sprachbewertung im Fremdsprachenunterricht abzuleiten. Grundlage bildet ein deskriptives Forschungsdesign mit einem standardisierten Fragebogen, der an 94 Deutschlehrkräfte staatlicher und privater Schulen in der Türkei verteilt wurde. Die Analyse erfolgte mit SPSS unter Anwendung deskriptiver Statistik. Die Ergebnisse geben Aufschluss über die Präferenzen der Lehrkräfte im Hinblick auf klassische (summativ) und alternative (formativ) Bewertungsmethoden und zeigen gleichzeitig Entwicklungspotenziale in Bezug auf eine kompetenzorientierte, formative Bewertung auf. Die Studie trägt zur Weiterentwicklung qualitativer Bewertungsstandards im DaF-Unterricht bei und leistet einen Beitrag zur Professionalisierung der Bewertungskompetenz von Lehrkräften.

**Schlüsselwörter:** *Deutschlehrkräfte, DaF-Unterricht, Evaluationsverfahren, deskriptive Studie, Lehrerpräferenzen.*

### *Abstract (English)*

#### ***Evaluation methods for Turkish teachers of German as a Foreign Language***

This study explores which evaluation methods are used and preferred by Turkish teachers of German as a Foreign Language. The aim is to identify current practices, highlight their strengths and weaknesses, and derive recommendations for improving language assessment in foreign language instruction. The study is based on a descriptive research design and uses a standardized questionnaire distributed to 94 teachers working at state and private schools in Turkey. The data were analyzed using SPSS with descriptive statistical methods. The results provide insights into teachers' preferences for classical (summative) and alternative (formative) assessment methods and reveal potential for the further development of competence-oriented, formative assessment practices. This study contributes to the advancement of quality standards in DaF assessment and supports the professionalization of teachers' assessment literacy.

**Keywords:** *Teachers of German, German as a second language teaching, assessment methods, descriptive study, Teacher preferences.*

## EXTENDED ABSTRACT

This study investigates the measurement and evaluation practices employed by Turkish teachers of German as a Foreign Language (GFL), focusing on the frequency, typology, and perceived pedagogical relevance of both classical and alternative assessment tools. Although the last decade has witnessed a growing discourse on formative, competence-oriented, and learner-centered evaluation in foreign language education, summative assessment practices continue to dominate instructional contexts in Turkey. Against this background, the study seeks to identify which specific techniques are most commonly implemented by teachers, how these align with current pedagogical expectations such as those outlined in the *Türkische Maarif-Modell*, and what implications arise for professionalization and innovation in assessment literacy.

The theoretical framework of the study draws on contemporary discussions in the field of language assessment, especially the contrast between summative and formative paradigms. Summative assessment—traditionally represented by standardized testing, grades, and product-oriented evaluation—serves accountability and certification purposes. In contrast, formative assessment emphasizes feedback, reflection, and learning progress, aiming to integrate assessment into the learning process itself. In addition, the study considers current developments regarding digitalization and artificial intelligence (AI) in language testing, as these technologies increasingly influence how teachers measure linguistic competence and provide feedback. Prior empirical research (e.g., Black & Wiliam 2018; Drackert & Nechaeva 2018) has underscored the transformative potential of formative approaches, while also noting persistent implementation barriers such as time constraints, large class sizes, and insufficient teacher training.

Methodologically, the research adopts a descriptive quantitative design. Data were collected through an online questionnaire distributed to 94 German teachers working in state and private schools across Turkey. The instrument, adapted from the model by Çelikkaya et al. (2010) and further refined for the DaF context, consisted of two parts: (1) demographic information (gender, teaching experience, school type, educational level) and (2) the frequency of use of nine classical and nine alternative evaluation techniques. Respondents rated each item on a four-point Likert scale (“I never use it”, “I rarely use it”, “I often use it”, “I always use it”). The data were analyzed using SPSS, employing descriptive statistics such as frequencies, percentages, means, and standard deviations to illustrate patterns of preference and practice.

The results reveal a clear predominance of classical assessment methods, such as multiple-choice tests, cloze tests, short-answer items, and oral examinations. These tools are widely valued for their objectivity, reliability, and efficiency, particularly in large classes where time and administrative resources are limited. In contrast, more elaborate alternative techniques—including portfolios, peer evaluation, and structured interviews—are used much less frequently. Nevertheless, certain formative approaches, especially self-assessment and project-based tasks, were reported as being implemented more regularly, reflecting an emerging but still cautious shift toward learner-centered evaluation.

These findings correspond with prior studies in the domains of English and Social Science education, which have likewise demonstrated the persistence of traditional testing despite curricular reforms emphasizing formative and competence-oriented assessment. For example, Drackert and Nechaeva (2018) found that while self-assessment positively affects learners’ motivation and satisfaction, its success depends on structure, clarity, and teacher guidance. Similarly, the qualitative work by Şentürk and Köksal (2025) among German teacher candidates demonstrated that formative evaluation—particularly in communicative grammar

teaching—was perceived as practical and effective in enhancing learner engagement and reflection. Such findings underline the need for systematic professional development in assessment literacy and a deeper integration of formative strategies into teacher education.

From a practical standpoint, the study recommends strengthening institutional and curricular frameworks that facilitate formative assessment, especially in overcrowded classrooms, where individualized feedback remains challenging. The inclusion of AI-supported tools—such as digital rubrics, automated feedback systems, and adaptive testing platforms—could help balance workload and ensure fairness and consistency in evaluation. These tools also align with the goals of the 2024 Turkish National Education Report and broader international agendas advocating digital transformation and equity in assessment.

However, certain limitations must be acknowledged. As a self-reported quantitative study, the data reflect teachers' perceptions rather than direct classroom observations, which restricts insight into contextual dynamics and authentic classroom behavior. Therefore, future research should employ mixed-methods approaches, combining survey data with qualitative interviews, lesson observations, and document analyses, to explore the motivations behind teachers' assessment choices and the practical challenges they encounter in implementing formative practices. Moreover, further investigation into the effectiveness of AI-assisted assessment tools could provide valuable evidence of integrating digital innovations sustainably into foreign language pedagogy.

In conclusion, this study provides empirical evidence for the persistent gap between policy aspirations and classroom realities in foreign language assessment. By mapping current trends, identifying professional needs, and highlighting areas of potential innovation, it contributes to the broader discourse on quality assurance, teacher training, and digital transformation in German as a Foreign Language education.

## Einleitung

Die Evaluation von Lernleistungen ist ein fester Bestandteil des Bildungsprozesses und eng mit dem Unterrichtsprozess verbunden. Sie ermöglichen nicht nur die Erfassung des Wissensstandes und der Kompetenzen der Lernenden, sondern auch die Ableitung gezielter Fördermaßnahmen sowie die Reflexion über die Wirksamkeit didaktischer Entscheidungen. Im Fremdsprachenunterricht, insbesondere im Bereich Deutsch als Fremdsprache (DaF), kommt der Leistungsbewertung eine zentrale Rolle zu, da sie direkte Rückschlüsse auf sprachliche Teilkompetenzen erlaubt.

In der Fachliteratur wird zwischen zwei grundlegenden Evaluationsansätzen unterschieden: der summativen und der formativen Bewertung. Die summative Evaluation dient der abschließenden Beurteilung von Lernergebnissen und wird häufig zur Notengebung oder Leistungsdokumentation eingesetzt (Harlen 2007; Tienson-Tseng 2019; Dmitrenko et al. 2021). Die formative Bewertung hingegen ist ein kontinuierlicher Prozess, der die Lernentwicklung begleitet, individuelles Feedback ermöglicht und flexible Anpassungen im Unterricht unterstützt (Şentürk 2025a). „Formative Beurteilungen begleiten den Kompetenzaufbau der Lernenden und zielen darauf ab, ihren Entwicklungsprozess gezielt zu fördern“ (Brown 2004: 6). Gemäß Grotjahn und Kleppin (2017) können Bewertungen von Leistungen entweder punktuell und produkt-/ergebnisorientiert am Ende eines Lernabschnitts stattfinden (summativ) oder kontinuierlich und prozessorientiert in den Unterricht integriert werden (formativ). Beide Ansätze sind komplementär und tragen in ihrer Kombination wesentlich zur Qualität des Lernprozesses bei (Taras 2005; Berisha et al. 2024).

Während summative Verfahren wie Multiple-Choice-Aufgaben, Lückentexte oder Zuordnungsaufgaben eine standardisierte und vergleichbare Leistungsmessung ermöglichen, werden sie auch dafür kritisiert, kognitive Tiefe und Lernprozesse nur begrenzt abbilden zu können (Çelikkaya et al. 2010). Formative Methoden wie Selbstevaluation, Peer-Feedback, projektbasierte Aufgaben oder mündliche Rückmeldungen fördern hingegen die Eigenverantwortung der Lernenden und stärken deren metakognitive Fähigkeiten (Schütze et al. 2018; Peyer et al. 2023). Besonders im kompetenzorientierten Fremdsprachenunterricht gewinnen diese alternativen Verfahren zunehmend an Bedeutung.

Im Rahmen des „türkischen Jahrhundert-Bildungsmodells“ wird die formative Bewertung ausdrücklich gefördert, da sie zur kontinuierlichen Verbesserung des Unterrichts und zur nachhaltigen Lernentwicklung beitragen soll (MEB 2024). Mess- und Bewertungsaktivitäten sollen demnach nicht nur Lernstände dokumentieren, sondern vielmehr Diagnoseinstrumente sein, um individuelle Lernprozesse gezielt zu unterstützen. Das neue Maarif-Modell betont den Einsatz kontinuierlicher, formativer Evaluationsverfahren, um den Lernprozess nachhaltig zu fördern – mit dem klaren Ziel, die Unterrichtsgestaltung konsequent auf eine tiefgreifende Lernentwicklung auszurichten.

Trotz der theoretischen Fundierung evaluativer Praktiken mangelt es bislang an empirischen Erkenntnissen darüber, welche Evaluationsverfahren im DaF-Unterricht

tatsächlich eingesetzt werden und welche Präferenzen Lehrkräfte dabei zeigen, insbesondere im türkischen Bildungskontext. Die vorliegende Studie setzt hier an, indem sie die Nutzung der Evaluationsverfahren durch türkische Deutschlehrkräfte systematisch erfasst. Ziel ist es, aktuelle Anwendungsmuster sichtbar zu machen, bestehende Bewertungspraxen einzuordnen und Impulse für eine evidenzbasierte Weiterentwicklung von Evaluationskultur im Fremdsprachenunterricht zu geben.

Vor diesem Hintergrund werden im Folgenden die beiden zentralen Evaluationsansätze — formative und summative Evaluation — theoretisch eingeordnet und hinsichtlich ihrer Funktionen im Fremdsprachenunterricht dargestellt.

## **Formative Evaluation**

Die formative Evaluation (formatives Assessment) stellt ein zentrales Element der lernbegleitenden Leistungsbeurteilung dar. Sie dient dazu, Informationen über den aktuellen Lernstand der Lernenden zu sammeln und auf dieser Grundlage den Unterricht gezielt anzupassen (Grotjahn & Kleppin 2017). „Formative Beurteilung umfasst dabei alle Aktivitäten während eines Kurses, für die Lernende Feedback erhalten, um ihr Lernen zu verbessern, unabhängig davon, ob eine Benotung erfolgt oder nicht“ (Higgins et al. 2010: 55).

Ziel der formativen Evaluation ist es, Stärken, Schwächen und Lernfortschritte sichtbar zu machen und daraus Fördermaßnahmen abzuleiten (GER 2001: 180). Zentrale Merkmale sind die Klärung von Lernzielen, die Diagnose individueller Leistungen sowie eine darauf basierende Rückmeldung und Unterstützung (Schütze et al. 2018: 701). Dadurch wird den Lernenden ermöglicht, ihren Lernprozess selbst zu überwachen und zunehmend Verantwortung für ihr Lernen zu übernehmen (vgl. Jones 2006).

Formative Evaluation erfolgt prozessbegleitend und berücksichtigt die Entwicklung von Kompetenzen im Lernverlauf (Brown 2004: 6). Sie kann sowohl formelle als auch informelle Verfahren umfassen, etwa kurze Tests, Beobachtungen, Portfolioarbeit oder Selbstbeurteilungen der Lernenden (Hughes 2003: 5). Die dabei gewonnenen Informationen dienen Lehrkräften als Grundlage zur Anpassung ihrer Unterrichtsplanung und -durchführung.

Darüber hinaus erfüllt formative Beurteilung mehrere Funktionen für Lehrkräfte: Sie unterstützt die Unterrichtsplanung, liefert Evidenzen für den Lernfortschritt der Lernenden, zeigt den Grad der Zielerreichung im Curriculum und ermöglicht zugleich eine Reflexion der eigenen Lehrpraxis (Rea-Dickins & Gardner 2000: 229). Zusammenfassend lässt sich festhalten, dass formative Evaluation zur Schaffung eines lernförderlichen, adaptiven Unterrichtsumfelds beiträgt, in dem Lernprozesse kontinuierlich begleitet und optimiert werden.

## Summative Evaluation

Die summative Evaluation wird häufig als „assessment of learning“ bezeichnet und dient der abschließenden Beurteilung dessen, was Lernende am Ende eines Kurses oder Lernabschnitts erreicht haben. Sie fasst die während eines bestimmten Zeitraums erworbenen Kenntnisse und Kompetenzen zusammen und erfolgt typischerweise am Ende eines Semesters oder einer Unterrichtseinheit (GER 2001; Brown 2004; Hughes 2003: 5). „Dabei wird die Leistung sowohl einzelner Lernender als auch von Gruppen zu einem bestimmten Zeitpunkt dokumentiert“ (Harlen 2007: 16).

Summative Beurteilungen sind in der Regel punktuell, normorientiert und kursbezogen (GER 2001). Sie erfolgen meist durch formelle Prüfungsformate wie Abschlussprüfungen oder standardisierte Tests, die den erreichten Lernstand zusammenfassend erfassen. Ziel ist es, festzustellen, „in welchem Umfang die Lernziele erreicht wurden, ohne jedoch notwendigerweise Hinweise für zukünftige Lernprozesse zu liefern“ (Brown 2004: 6).

Im Gegensatz zur prozessorientierten formativen Evaluation konzentriert sich die summative Evaluation auf das Ergebnis des Lernens. Sie dient häufig als Grundlage für Entscheidungen wie Notengebung, Leistungsdifferenzierung oder Einstufung der Lernenden in verschiedene Kompetenzniveaus. Gemäß Grotjahn und Kleppin (2017: 78) können „Leistungsbewertungen entweder summativ oder formativ erfolgen: Während summative Verfahren punktuell und ergebnisorientiert am Ende eines Lernabschnitts durchgeführt werden, sind formative Verfahren kontinuierlich, prozessorientiert und in den Unterricht integriert“. Tabelle 1 stellt die zentralen Unterschiede zwischen beiden Evaluationsformen dar und bietet einen Überblick über wesentliche Aspekte der Leistungsbewertung (Grotjahn & Kleppin 2017: 78).

<b>formative Evaluation</b>	<b>summative Evaluation</b>
Die Evaluation ist in den Unterricht integriert und findet regelmäßig statt.	Die Evaluation steht am Ende eines Lernabschnitts.
Die Evaluation ist häufig interaktiv, dialogisch und kooperativ.	Die Evaluation ist eine Momentaufnahme der Leistungen des Lernenden durch die Lehrkraft.
Die Ergebnisse der Evaluation werden von der Lehrkraft für die Optimierung der Lehrprozesse genutzt.	Die Ergebnisse der Evaluation dienen der Lehrkraft als Informationen, ob und inwieweit die Lernziele des jeweiligen Lernabschnittes von den Lernenden erreicht wurden.
Soweit es in der Unterrichtssituation sinnvoll ist, werden die Ergebnisse so an die Lernenden zurückgemeldet, dass diese selbst ihr Lernen optimieren können.	Die Ergebnisse werden an die Lernenden als ein abschließender Lernertrag zurückgemeldet. Die Ergebnisse bilden meist die Grundlage für Auslese und Einstufung der Lernenden in Leistungsniveaus.
Rückmeldungen sind in der Tendenz deskriptiv und zumeist nicht mit einer Benotung verbunden.	Rückmeldungen erfolgen durch Bewertung und Benotung.

**Tab. 1:** Unterschiede der formativen und summativen Evaluation

Zusammenfassend fungiert die summative Evaluation als abschließende Leistungsfeststellung, die den Lernerfolg dokumentiert und als Grundlage für administrative sowie pädagogische Entscheidungen dient. Sie ergänzt die formative Evaluation, welche den Lernprozess prozessbegleitend unterstützt.

## **Forschungsstand**

Die Forschung zur Leistungsbewertung im Fremdsprachenunterricht hat in den letzten Jahrzehnten zahlreiche Perspektiven hervorgebracht. Insbesondere der Gegensatz zwischen summativer und formativer Bewertung wird in der Literatur intensiv diskutiert (Tienson-Tseng 2019; Şentürk & Köksal 2025; Berisha et al. 2024; Şentürk 2025a). Während klassische Verfahren wie Multiple-Choice-Tests und mündliche Prüfungen nach wie vor weit verbreitet sind, wird die formative Methode zunehmend als wirksames Mittel zur Unterstützung des Lernprozesses betrachtet (Dmitrenko et al. 2021; Swan & Foster 2018). Formative Verfahren wie Selbstevaluation, Peer-Evaluation und kompetenzorientierte Aufgaben gelten als förderlich für selbstreguliertes Lernen und nachhaltigen Kompetenzaufbau (Peyer et al. 2023). In einer empirischen Studie mit Russischlernenden zeigen Drackert und Nechaeva (2018), dass die regelmäßige Selbsteinschätzung des Lernfortschritts zu höherer Zufriedenheit führen kann. Allerdings verdeutlicht die Studie auch, dass offene Formate der Selbstevaluation nicht von allen Lernenden als hilfreich empfunden werden, was auf die Bedeutung einer didaktisch differenzierten Gestaltung solcher Verfahren hinweist.

Aktuelle empirische Studien betonen zudem die Relevanz digitaler Tools und KI-gestützter Technologien im Bewertungsprozess. Zhai und Nehm (2023) legen nahe, dass KI-basierte Systeme Lehrkräfte insbesondere in der formativen Bewertung effektiv unterstützen und individualisiertes Feedback in Echtzeit ermöglichen können. Auch internationale Bildungsberichte (z. B. U.S. Department of Education 2023) weisen darauf hin, dass KI zunehmend als Instrument zur Qualitätssicherung und Lernprozessbegleitung eingesetzt wird. Eine systematische Übersicht von Ningsih et al. (2024) hebt hervor, dass digitale formative Tools insbesondere im Fremdsprachenunterricht das Potenzial besitzen, Motivation und Lernerfolg der Lernenden zu steigern, sofern sie didaktisch sinnvoll integriert werden. Ergänzend unterstreicht der Nationale Bildungsbericht „Bildung in Deutschland 2024“, dass KI-gestützte Technologien im Bildungswesen ein wachsendes Potenzial zur individuellen Förderung und zur Verbesserung adaptiver Lernprozesse bieten – vorausgesetzt, Lehrkräfte werden entsprechend qualifiziert (Nationaler Bildungsbericht 2024).

Im Kontext des DaF-Unterrichts in der Türkei besteht jedoch weiterhin Forschungsbedarf. Eine qualitative Studie von Şentürk und Köksal (2025) zeigt, dass die Deutschlehramtskandidaten formative Bewertungsverfahren als zentral für eine praxisnahe, lernwirksame und effektive Unterrichtsgestaltung betrachten. Kontinuierliches Feedback und kompetenzorientierte Bewertung stärken dabei das studentenzentrierte Lernen und fördern eine zeitgemäße Lehrerbildung im DaF-Kontext. Während einige Studien die Anwendung klassischer Bewertungsverfahren

bestätigen (Çelikkaya et al. 2010; Üner & Aşlıoğlu 2022), fehlt es weitgehend an systematischen Analysen zum Einsatz formativer oder KI-gestützter Methoden.

Zudem zeigen Untersuchungen zu Lehrenden, die am Anfang ihrer beruflichen Laufbahn stehen, dass die Vorbereitung, Durchführung und Bewertung von Prüfungen zur Messung der vier grundlegenden Sprachfertigkeiten als besonders zeitaufwendig und belastend wahrgenommen werden. Diese Befunde weisen auf erhebliche strukturelle Herausforderungen im Bereich der Leistungsbewertung im DaF-Unterricht hin und verdeutlichen den Bedarf an gezielter Unterstützung für Lehrkräfte (vgl. Şentürk 2025b). Die vorliegende Studie greift diese Forschungslücke auf und untersucht die Nutzung und Präferenz von Evaluationsverfahren türkischer DaF-Lehrkräfte im Kontext aktueller didaktischer und technologischer Entwicklungen.

## **Ziel und Fragestellung**

Das Ziel dieser Studie besteht darin, die von türkischen Deutschlehrkräften im Unterricht Deutsch als Fremdsprache bevorzugten Evaluationsverfahren sprachlicher Kompetenzen zu untersuchen. Durch die Analyse dieser Präferenzen sollen bestehende Bewertungspraxen hinsichtlich ihrer Stärken und Schwächen kritisch beleuchtet sowie Ansätze zur Optimierung des Fremdsprachenunterrichts entwickelt werden. Aus diesem Ziel ergibt sich folgende zentrale Forschungsfrage:

- Welche Evaluationsverfahren bevorzugen türkische Deutschlehrkräfte zwecks der Überprüfung von Sprachkompetenzen im DaF-Unterricht?

## **Methode**

Die vorliegende Untersuchung folgt einem quantitativen, deskriptiven Forschungsdesign. Ziel war es, die von türkischen Deutschlehrkräften im Fach Deutsch als Fremdsprache (DaF) bevorzugten Evaluationsverfahren systematisch zu erfassen und auf dieser Grundlage Stärken und Schwächen der gängigen Bewertungspraxen zu identifizieren. Der deskriptive Ansatz erlaubt es, Häufigkeiten, Verteilungen und zentrale Tendenzen der erhobenen Daten darzustellen, ohne kausale Zusammenhänge zu untersuchen (vgl. Dörnyei 2007; Bortz & Döring 2013).

## **Datenerhebung**

Um die erforderlichen Daten zu erheben, wurde ein Fragebogen eingesetzt. Der Fragebogen wurde auf der Grundlage der Studie von Çelikkaya et al. (2010) entwickelt und inhaltlich an den Kontext des DaF-Unterrichts angepasst. Der Fragebogen besteht sowohl aus demografischen Merkmalen der Teilnehmenden (z. B. Geschlecht, Ausbildungsgrad, Schultyp) als auch Angaben zu den verwendeten Evaluationsverfahren.

Die Antwortmöglichkeiten zu den einzelnen Verfahren basierten auf einer vierstufigen Likert-Skala mit den Kategorien: „Ich verwende nie“, „Ich verwende selten“, „Ich verwende oft“ und „Ich verwende ständig“. Diese Skalierung ermöglichte eine differenzierte Erfassung der Häufigkeit, mit der die Lehrkräfte verschiedene Evaluationsverfahren im Unterricht einsetzen.

## **Datenanalyse**

Die Datenauswertung erfolgte mit der Statistiksoftware SPSS. Im Rahmen einer deskriptiven Analyse wurden Häufigkeiten, Prozentwerte, Mittelwerte und Standardabweichungen berechnet. Diese statistischen Kennzahlen ermöglichen eine differenzierte Darstellung der Präferenzen hinsichtlich der Bewertungstechniken. Die Ergebnisse wurden grafisch (z. B. Balken- und Kreisdiagramme) und tabellarisch aufbereitet, um sowohl eine quantitative Übersicht als auch eine punktuelle Analyse der einzelnen Bewertungstechniken zu gewährleisten.

### *Validität*

- **Konstruktvalidität:** Der Fragebogen wurde auf Grundlage theoretisch fundierter Kategorien zur Leistungsbewertung im Fremdsprachenunterricht modifiziert. Die untersuchten Bewertungsmethoden orientieren sich an gängigen Konzepten der DaF-Didaktik, was die Konstruktvalidität unterstützt.
- **Inhaltsvalidität:** Die Items wurden so gestaltet, dass sie zentrale Aspekte der Bewertungspraxis im DaF-Unterricht erfassen. Expert:innen aus dem Bereich Fremdsprachendidaktik überprüften die inhaltliche Passung der Fragen zur Forschungsfrage und Zielsetzung.
- **Kriteriumsvalidität:** Die Befunde wurden mit Ergebnissen ähnlicher Studien zur Evaluation von Bewertungspraxen im Fremdsprachenunterricht verglichen. Dies stärkt die externe Vergleichbarkeit und stützt die Validierung der Ergebnisse.

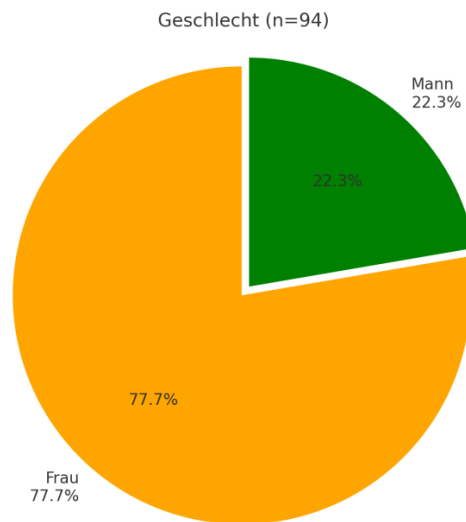
### *Reliabilität*

- **Standardisierung der Erhebung:** Die digitale Durchführung über Microsoft Forms und die Verwendung geschlossener Antwortformate auf einer vierstufigen Likert-Skala sicherten eine einheitliche Erhebungssituation und verringerten potenzielle Messfehler.
- **Vergleichbarkeit mit anderen Studien:** Die erhobenen Daten und angewandten Methoden wurden mit ähnlichen empirischen Untersuchungen im Bereich der Sprachtestdidaktik abgeglichen, um die methodische Konsistenz und Nachvollziehbarkeit zu gewährleisten.

## **Teilnehmer**

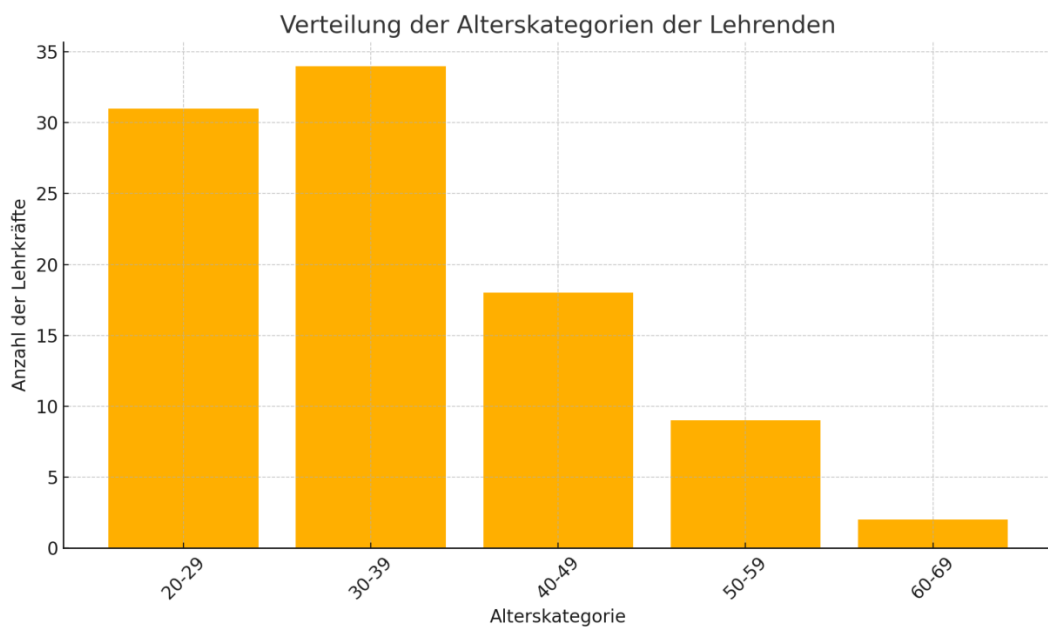
Die Teilnehmenden der Untersuchung bestehen aus 94 türkischen Deutschlehrkräften, die aktiv an öffentlichen und privaten Schulen in der Türkei arbeiten. Bei der

Stichprobenauswahl wurde eine leicht zugängliche, auf Freiwilligen basierende Methode bevorzugt. In Folgenden werden die demografischen Daten der Teilnehmenden präsentiert.



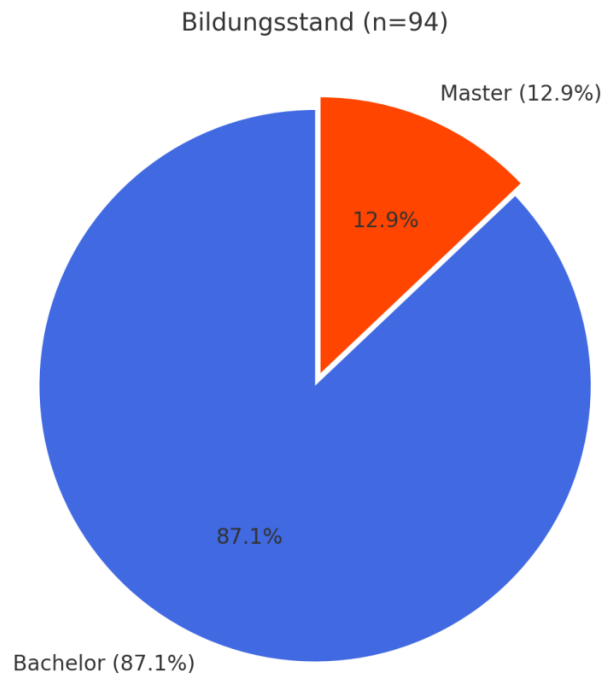
**Abb. 1:** Geschlecht der Lehrenden

Wie aus der Abbildung 1 hervorgeht, sind 77,7 % der befragten Lehrenden weiblich, während der Anteil der männlichen Lehrenden bei 22,3 % liegt.



**Abb. 2:** Alter der Lehrenden

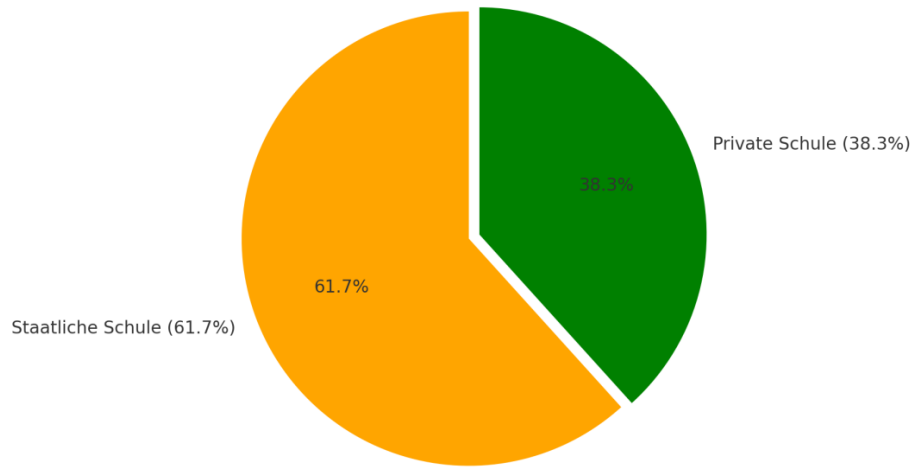
Die Abbildung 2 verdeutlicht die Verteilung der Lehrenden nach Alterskategorien. Es wird deutlich, dass die Altersgruppen 20-29 und 30-39 Jahren am stärksten vertreten sind, während die Altersgruppe 60-69 Jahren die geringste Anzahl aufweist. Aus der Grafik lässt sich ableiten, dass das Durchschnittsalter der Lehrenden bei etwa 37,4 Jahren liegt.



**Abb. 3:** Bildungsstatus der Lehrenden

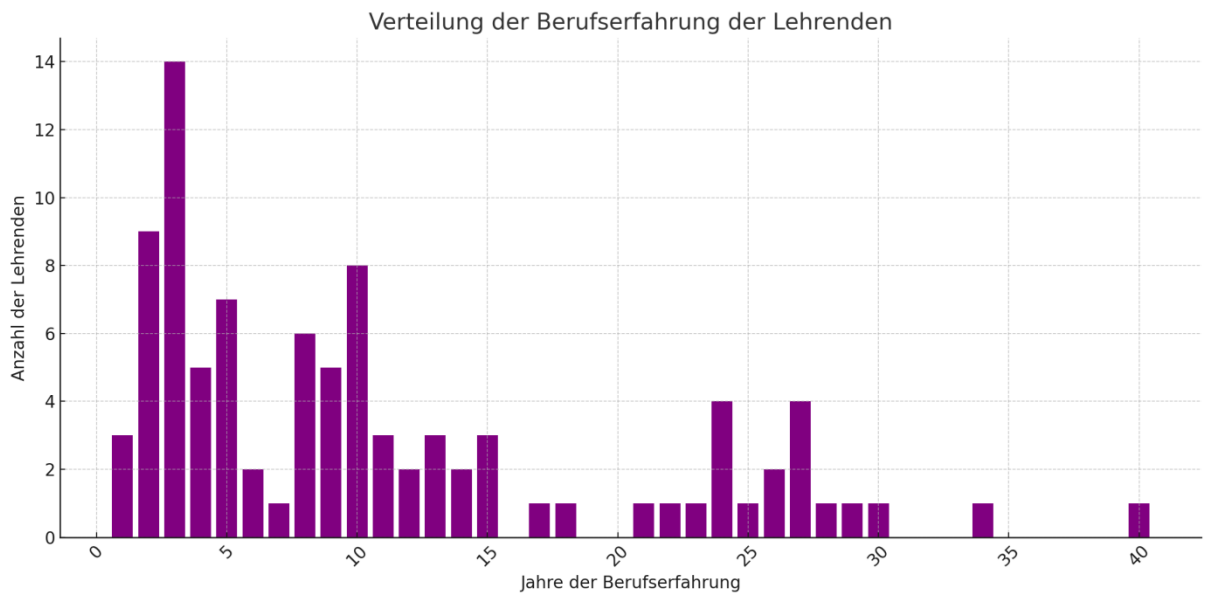
Aus der Abbildung 3 geht hervor, dass 87,1 Prozent der Lehrenden einen Bachelor-Abschluss besitzen, während 12,9 Prozent über einen Master-Abschluss verfügen.

Art der Schule, an der die Lehrkraft arbeitet (n=94)



**Abb. 4:** Schulart, an der die Lehrenden arbeiten

Die Abbildung 4 veranschaulicht die Verteilung der Lehrenden nach Schulart. Demnach arbeiten 61,7 % der Lehrkräfte an staatlichen Schulen, während 38,3 % an privaten Schulen tätig sind. Diese Aufteilung gibt Aufschluss über die Prävalenz von Deutschunterricht in verschiedenen Schulformen.



**Abb. 5:** Verteilung der Berufserfahrung

Die Abbildung 5 verdeutlicht, wie viele Lehrende jeweils über eine bestimmte Anzahl von Jahren Berufserfahrung verfügen. Die häufigsten Erfahrungswerte liegen bei 4 und 10 Jahren, was darauf hindeutet, dass es mehrere Lehrkräfte mit kürzerer

Berufserfahrung gibt. Die Lehrenden mit längerer Berufserfahrung (über 20 Jahre) sind weniger zahlreich vertreten. Diese Verteilung könnte Auswirkungen auf die Dynamik und Erfahrung in der Unterrichtspraxis haben.

## Befunde

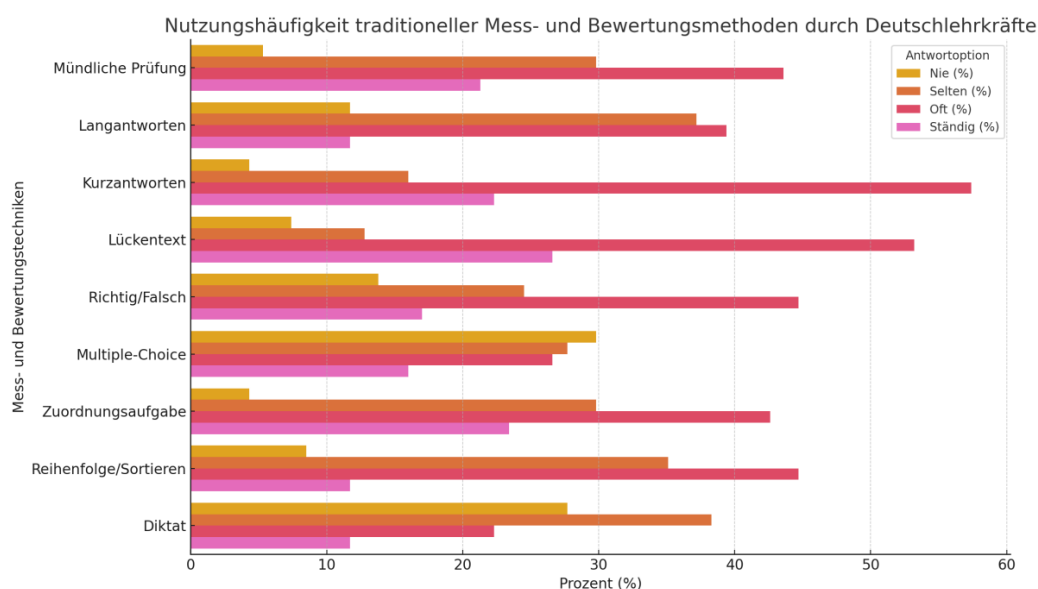
In diesem Abschnitt werden die Ergebnisse der Analyse dargestellt, wobei der Schwerpunkt auf der Nutzung von Evaluationsverfahren durch Deutschlehrkräfte liegt.

Evaluationsverfahren			Ich verwende nie	Ich verwende selten	Ich verwende oft	Ich verwende ständig	Total
1	Mündliche Prüfungen	N	5	28	41	20	94
		%	5,3	29.8	43.6	21.3	100
2	Langantworten	N	11	35	37	11	94
		%	11.7	37.2	39.4	11.7	100
3	Kurzantworten	N	4	15	54	21	94
		%	4.3	16.0	57.4	22.3	100
4	Lückentexte	N	7	12	50	25	94
		%	7.4	12.8	53.2	26.6	100
5	Richtig/Falsch Aufgaben	N	13	23	42	16	94
		%	13.8	24.5	44.7	17.0	100
6	Multiple-Choice- Aufgaben	N	28	26	25	15	94
		%	29.8	27.7	26.6	16.0	100
7	Zuordnungsaufgaben	N	4	28	40	22	94
		%	4.3	29.8	42.6	23.4	100
8	Reihenfolge-/ Sortier- Aufgaben	N	8	33	42	11	94
		%	8.5	35.1	44.7	11.7	100
9	Diktate	N	26	36	21	11	94
		%	27.7	38.3	22.3	11.7	100

**Tab. 2:** Klassische (summativ) Evaluationsverfahren

Die Tabelle 2 legt nahe, dass offene Aufgabenformate wie Langantworten und Kurzantworten, die am häufigsten genutzten Bewertungsmethoden sind, was auf ihre Effektivität zur Förderung der Ausdrucksfähigkeit der Lernenden hindeutet. Auch Lückentexte werden stark bevorzugt, da sie eine gute Balance zwischen Struktur und Ausdrucksfreiheit bieten. Dies deutet darauf hin, dass die Lehrkräfte die aktive Sprachproduktion und das kritische Denken der Lernenden fördern möchten, anstatt sich auf reine Wiedererkennung oder Reproduktion von Wissen zu beschränken. Geschlossene Aufgabenformate wie Multiple-Choice, Richtig/Falsch, Reihenfolge-Aufgaben und Diktat werden seltener verwendet. Diese Methoden scheinen für die Lehrenden weniger attraktiv zu sein, möglicherweise weil sie weniger tiefgehendes Verständnis abfragen und eher auf das Wiedererkennen von Informationen abzielen. Insgesamt nutzen Lehrende vorrangig Bewertungsmethoden, die die aktive Sprachproduktion und das kritische Denken fördern, während weniger interaktive Methoden zurückhaltender genutzt werden.

Die nachfolgende Grafik veranschaulicht die Häufigkeit der Verwendung traditioneller summativer Messinstrumente durch Deutschlehrkräfte.



**Abb. 6:** Häufigkeit der Verwendung traditioneller Evaluationsverfahren

Die Abbildung 6 zeigt deutlich, dass Kurzantworten (57,4 % „oft“ und 22,3 % „ständig“) sowie Lückentextaufgaben (53,2 % „oft“, 26,6 % „ständig“) von einem Großteil der Deutschlehrkräfte regelmäßig eingesetzt werden. Auch Zuordnungsaufgaben (42,6 % „oft“, 23,4 % „ständig“) und Reihenfolge-/Sortieraufgaben (44,7 % „oft“) sind häufig verwendete Formate. Diese hohe Nutzung weist auf eine starke Orientierung an klar strukturierten, überprüfbaren Aufgabenformaten hin, die sich leicht in standardisierte Prüfungsverfahren integrieren lassen.

Im Gegensatz dazu fällt auf, dass Multiple-Choice-Aufgaben nur von 16 % der Lehrkräfte ständig und von 26,6 % oft verwendet werden, während 29,8 % angeben, diese nie zu nutzen. Noch markanter ist der Befund bei Diktaten: Hier geben 27,7 % „nie“ und 38,3 % „selten“ als Antwort an. Dies legt nahe, dass diese Techniken im modernen Fremdsprachenunterricht deutlich weniger Relevanz besitzen – möglicherweise aufgrund ihrer begrenzten Aussagekraft über kommunikative Kompetenzen.

Zusammenfassend verdeutlicht die Grafik, dass traditionelle summative Methoden wie Kurzantworten und Lückentexte nach wie vor dominieren, während zeitintensivere oder umstrittenere Formate wie Diktate zunehmend an Bedeutung verlieren. In der folgenden Tabelle werden die alternativen Evaluationsverfahren dargestellt, die von den Lehrkräften eingesetzt werden.

Alternative Evaluationsverfahren		Ich verwende nie	Ich verwende selten	Ich verwende oft	Ich verwende ständig	Total
1 Selbstevaluation	N	12	33	37	12	94
	%	12.8	35.1	39.4	12.8	100
2 Peer-Evaluation	N	26	35	27	6	94
	%	27.7	37.2	28.7	6.4	100
3 Gruppen-Evaluation	N	21	34	32	7	94
	%	22.3	36.2	34.0	7.4	100
4 Portfolio	N	34	32	21	7	94
	%	36.2	34.0	22.3	7.4	100
5 Projektarbeit	N	8	39	28	19	94
	%	8.5	41.5	29.8	20.2	100
6 kompetenzorientierte Aufgabe	N	5	21	41	27	94
	%	5.3	22.3	43.6	28.7	100
7 Mindmap	N	25	38	22	9	94
	%	26.6	40.4	23.4	9.6	100
8 Beobachtung	N	12	17	39	26	94

		%	12.8	18.1	41.5	27.7	100
9	Interview	N	36	29	23	6	94
		%	38.3	30.9	24.5	6.4	100

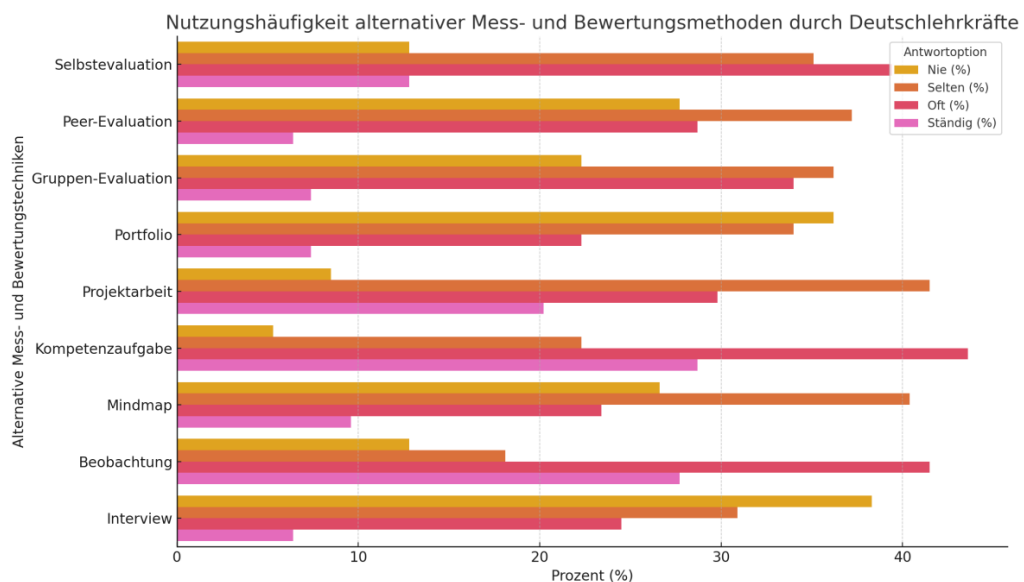
**Tab. 3:** Alternative (formativ) Evaluationsverfahren

Die Tabelle 3 verdeutlicht, in welchem Umfang alternative Evaluationsverfahren von Deutschlehrkräften im DaF-Unterricht eingesetzt werden. Es ist zu erkennen, dass die Mehrheit der Lehrkräfte Selbstevaluation, kompetenzorientierte Aufgaben und Beobachtung relativ häufig oder ständig verwendet. Diese Verfahren spiegeln eine stärker lernerzentrierte und formative Bewertungspraxis wider.

Die Selbstevaluation wird von 52,2 % (oft/ständig) der Befragten oft oder ständig genutzt, was auf die zunehmende Bedeutung der Reflexionskompetenz der Lernenden hinweist. Auch kompetenzorientierte Aufgaben werden mit 72,3 % (oft/ständig) am häufigsten eingesetzt und stellen somit das zentralste Instrument alternativer Bewertung dar. Beobachtung folgt mit 69,2 % (oft/ständig) auf Platz zwei.

Demgegenüber werden Portfolios und Interviews eher selten eingesetzt: Rund 36,2 % der Lehrkräfte verwenden Portfolio nie, und 38,3 % verzichten auf Interviews vollständig. Peer- und Gruppen-Evaluationen nehmen eine mittlere Position ein, während Projektarbeit und Mindmaps eine moderate, aber wachsende Nutzung zeigen.

Rekapitulierend verdeutlicht die Tabelle 3, dass alternative, prozess- und kompetenzorientierte Evaluationsverfahren im DaF-Unterricht zwar zunehmend Beachtung finden, jedoch noch nicht flächendeckend etabliert sind.



**Abb. 7:** Häufigkeit der Verwendung alternativer Mess- und Bewertungstechniken

Die Abbildung 7 veranschaulicht, wie häufig alternative (formative) Mess- und Evaluationsverfahren von Deutschlehrkräften eingesetzt werden. Auffällig ist, dass insbesondere kompetenzorientierte Aufgaben (43,6 % oft, 28,7 % ständig) sowie

Beobachtungen (41,5 % oft, 27,7 % ständig) vergleichsweise häufig genutzt werden. Diese Techniken scheinen im Unterricht stark verankert zu sein und deuten auf eine zunehmende Orientierung an individuellen Lernprozessen hin.

Demgegenüber werden Interviews, Portfolios (70 %) und Peer-Evaluation (54 %) von einem beträchtlichen Teil der Lehrkräfte gar nicht oder nur selten verwendet. Dies könnte auf zeitliche, methodische oder institutionelle Hürden hinweisen. Abschließend macht die Grafik deutlich, dass alternative Evaluationsverfahren zwar präsent sind, aber in ihrer Anwendung stark variieren, was auf unterschiedliche didaktische Prioritäten und Rahmenbedingungen im DaF-Unterricht schließen lässt.

Die Auswertung aus Abbildung 6 verdeutlicht, dass Deutschlehrkräfte ein ausgewogenes Verhältnis zwischen traditionellen und alternativen Evaluationsverfahren anstreben. Besonders häufig kommen strukturierte und effizient durchführbare Formate wie Kurzantworten (57,4 %), Lückentexte (53,2 %) und Zuordnungsaufgaben (42,6 %) zum Einsatz. Diese Methoden sind leicht standardisierbar und lassen sich gut in den Schulalltag integrieren. Gleichzeitig zeigen die Daten, dass auch alternative, lernprozessorientierte Ansätze zunehmend Bedeutung gewinnen. So werden etwa kompetenzorientierte Aufgaben von 43,6 % der Lehrkräfte oft und von 28,7 % ständig genutzt. Auch Beobachtungen (41,5 % oft, 27,7 % ständig) werden regelmäßig eingesetzt und deuten auf eine stärkere Berücksichtigung individueller Lernverläufe hin. Demgegenüber werden aufwendigere Verfahren wie Portfolios (36,2 % nie) oder Interviews (38,3 % nie) deutlich seltener eingesetzt. Dies lässt auf zeitliche und organisatorische Hürden schließen, die deren Anwendung im schulischen Kontext erschweren.

## **Diskussion**

Die Ergebnisse der Studie bestätigen frühere Befunde, die eine vorherrschende Nutzung klassischer Evaluationsverfahren durch Lehrkräfte aufzeigen. Wie Çelikkaya, Karakuş und Demirbaş (2010) betonen, greifen viele Lehrkräfte bevorzugt zu traditionellen Evaluationsverfahren wie Kurzantworttests und schriftlichen Prüfungen – eine Tendenz, die sich auch bei den befragten Deutschlehrkräften zeigt. Evaluationsverfahren wie Kurzantworten, Lückentexte und mündliche Prüfungen dominieren aufgrund ihrer strukturellen Klarheit und praktischen Anwendbarkeit. Im Englischunterricht ist laut Üner und Aşlıoğlu (2022) ein ähnliches Muster erkennbar: Lehrkräfte setzen überwiegend auf schriftliche und mündliche Prüfungen sowie Multiple-Choice-Aufgaben und achten dabei stark auf die Passung zum Lernstand. Dies spiegelt sich auch in dieser Studie wider.

Neben traditionellen Verfahren finden auch formative Evaluationsverfahren bei Deutschlehrkräften zunehmend Berücksichtigung. Im Gegensatz zu anderen Studien, die primär summative Formate dokumentieren, offenbart die vorliegende Untersuchung eine differenziertere Praxis: alternative Evaluationsverfahren wie Selbstevaluation, Peer-Evaluation und kompetenzorientierte Aufgaben finden durchaus Anwendung. Dies

deutet auf eine zunehmende Offenheit gegenüber formativen Verfahren hin, wenngleich deren Einsatz im Vergleich zu klassischen Verfahren noch zurückhaltend ausfällt. Diese Tendenz wird durch die qualitative Studie von Şentürk und Köksal (2025) gestützt in der die angehenden Deutschlehrkräfte formative Evaluationsverfahren – insbesondere im kommunikativen Grammatikunterricht – als praxisnah und lernförderlich einschätzen. Im Vordergrund standen insbesondere kontinuierliches Feedback, kompetenzorientierte Aufgabenstellungen und eine Bewertung, die stark auf den Lernprozess ausgerichtet ist. Gleichzeitig zeigen andere Befunde, dass die Umsetzung solcher Verfahren mit einem erhöhten Arbeitsaufwand verbunden sein kann, insbesondere für Lehrende am Beginn ihrer beruflichen Laufbahn (vgl. Şentürk 2025b). Dies könnte erklären, warum formative Verfahren trotz positiver Einstellungen in der Praxis noch vergleichsweise zurückhaltend eingesetzt werden.

Alternative Verfahren wie Portfolios oder Peer-Evaluation werden – trotz ihrer didaktischen Potenziale – seltener eingesetzt. Wie Çoruhlu, Nas und Çepni (2009) zeigen, sind hierfür vor allem Zeitmangel, fehlende Schulung und Unsicherheiten bei der Bewertung verantwortlich. Trotz der bildungspolitischen Leitlinien im Sinne des *Maarif-Modells* (MEB 2024), das formative Evaluation und kompetenzorientierte Ansätze fördert, bleibt die Unterrichtspraxis stark summativ geprägt. Viele türkische Deutschlehrkräfte greifen auf vertraute Verfahren zurück und integrieren innovative Verfahren nur punktuell. Daraus ergibt sich ein klarer Bedarf an gezielter Fortbildung, um sowohl die methodische als auch die konzeptionelle Sicherheit im Umgang mit alternativen Evaluationsverfahren zu stärken.

## **Fazit**

Die vorliegende Studie bietet einen systematischen Überblick über die Bewertungspraktiken türkischer Deutschlehrkräfte im DaF-Unterricht und leistet damit einen wichtigen Beitrag zur empirischen Erforschung von Evaluationsverfahren im schulischen Fremdsprachenkontext. Die Ergebnisse deuten darauf hin, dass Lehrkräfte bei der Wahl ihrer Evaluationsverfahren sowohl pädagogische als auch praktische Aspekte berücksichtigen. Während strukturierte, klassische Verfahren bevorzugt werden, ist zugleich ein wachsendes Interesse an formativen Methoden erkennbar, die individuelle Lernprozesse stärker in den Fokus rücken. Die Untersuchung macht deutlich, dass Bewertung nicht nur der Leistungsfeststellung dient, sondern ein zentrales Instrument zur Lernbegleitung und Kompetenzentwicklung darstellt.

## **Forschungsfrage:**

Welche Evaluationsverfahren werden zwecks der Überprüfung von Sprachkompetenzen von türkischen Deutschlehrkräften im DaF-Unterricht bevorzugt?

- Die Auswertung der Daten deutet darauf hin, dass türkische Deutschlehrkräfte vor allem klassische (summative) Evaluationsverfahren bevorzugen.

Kurzantworten (57,4 % oft, 22,3 % ständig) und Lückentexte (53,2 % oft, 26,6 % ständig) werden besonders häufig genutzt. Auch mündliche Prüfungen gehören mit über 60 % regelmäßiger Nutzung zu den bevorzugten Techniken.

- Formative Verfahren wie kompetenzorientierte Aufgaben (43,6 % oft, 28,7 % ständig) und Selbstevaluation (52,2 % oft/ständig) finden ergänzend Anwendung. Portfolios (36,2 % nie) und Interviews (38,3 % nie) hingegen werden deutlich seltener eingesetzt.

Gleichzeitig verdeutlicht die Studie einen bedeutsamen Entwicklungsbedarf im Bereich der alternativen und formativen Evaluationsverfahren, insbesondere im Hinblick auf deren systematische Integration in den schulischen Alltag. Trotz einer zunehmenden Offenheit gegenüber innovativen Verfahren bleibt deren praktische Umsetzung aus Gründen der fehlenden Schulung, Zeitressourcen oder methodischen Sicherheit begrenzt. Daraus ergibt sich ein klarer bildungspolitischer Handlungsauftrag: Um eine nachhaltige Kompetenzorientierung im Fremdsprachenunterricht zu gewährleisten, sollten Fortbildungsangebote ausgebaut, praxisnahe Leitfäden bereitgestellt und unterstützende digitale Lösungen gezielt gefördert werden.

### **Vorschläge:**

- Nach Maarif-Modell sollten Evaluationsverfahren vielfältig sein und die unterschiedlichen Fähigkeiten, besonderen Bedürfnisse und Lernprofile der Lernenden berücksichtigen. Die Lernende sollten auf eine facettenreiche Weise bewertet werden. Wie es verschiedene Ansätze gibt, Lerninhalte zu vermitteln, ist auch der Einsatz vielfältiger Evaluationsverfahren zur Erfassung der Lernleistungen der Lernenden pädagogisch sinnvoll.

- Im Sinne des Maarif-Modells, das auf kontinuierliche und kompetenzorientierte Leistungsbeurteilung abzielt, sollten formative Evaluationsverfahren systematisch in den Unterricht integriert werden.

- Da die praktische Umsetzung formativer Evaluation vor allem in großen Lerngruppen mit strukturellen Herausforderungen verbunden ist, bieten KI-gestützte Bewertungstools eine praktikable Alternative. Sie ermöglichen individualisiertes Feedback, reduzieren den administrativen Aufwand und tragen zur digitalen Transformation im Bildungssystem bei.

- In einer zukünftigen Studie könnten qualitative Methoden wie Fokusgruppeninterviews eingesetzt werden, um die Gründe für die Wahl bestimmter Evaluationsverfahren durch Lehrkräfte zu untersuchen.

- Aus den Ergebnissen der Studie lässt sich ableiten, dass Lehrkräfte im DaF-Unterricht stärker in ihrer Bewertungskompetenz (Assessment Literacy) unterstützt werden sollten.

## Limitationen

Bei der Interpretation der Ergebnisse dieser Studie sind einige Einschränkungen zu berücksichtigen. Zunächst basiert die Untersuchung auf Daten von 94 Deutschlehrkräften in der Türkei, die mittels einer Gelegenheitsstichprobe auf freiwilliger Basis gewonnen wurden. Dadurch ist die Stichprobe nicht repräsentativ für alle DaF-Lehrkräfte in der Türkei, was die Generalisierbarkeit der Ergebnisse einschränkt. Abschließend ist zu beachten, dass die Studie ausschließlich auf quantitativen Fragebogendaten basiert und keine qualitativen Datenquellen wie Unterrichtsbeobachtungen, Interviews oder Dokumentenanalysen einbezogen wurden, die ein vertiefteres Verständnis der Hintergründe der Bewertungspraxis ermöglichen hätten.

## Literaturverzeichnis

- Aydoğmuş, Ayşe / Coşkun Keskin, Sevgi** (2012): Sosyal bilgiler öğretmenlerinin süreç odaklı ölçme ve değerlendirme araçlarını kullanma durumları: İstanbul ili örneği. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 8(2), 110–123. <https://dergipark.org.tr/tr/pub/mersinefd/issue/17380/181483>
- Berisha, Fatlume / Vula, Eda / Gisewhite, Rachel / McDuffie, Hannah** (2024): The effectiveness and challenges implementing a formative assessment professional development program. *Teacher Development*, 28(1), 19-43. [https://aquila.usm.edu/fac\\_pubs/21581](https://aquila.usm.edu/fac_pubs/21581)
- Bortz, Jürgen / Döring, Nicola** (2013): *Forschungsmethoden und Evaluation: Für Human- und Sozialwissenschaftler* (4. Aufl.). Springer. DOI:10.17877/DE290R-11839
- Brown, H. Douglas** (2004): Language assessment: Principles and classroom practices. *White Plains, NY*: Pearson Education. DOI:10.1177/0265532207086784
- Çelikkaya, Tekin / Karakuş, Ufuk / Demirbaş, Çağrı** (2010): Sosyal Bilgiler Öğretmenlerinin Ölçme-Değerlendirme Araçlarını Kullanma Düzeyleri ve Karşılaştıkları Sorunlar. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi*, 11(1), 57-76. <https://dergipark.org.tr/tr/pub/kefad/issue/59507/855653>
- Çoruhlu, Tülay Şenel / Er Nas, Sibel / Çepni, Salih** (2009): Fen ve teknoloji öğretmenlerinin alternatif ölçme-değerlendirme tekniklerini kullanmada karşılaştıkları problemler: Trabzon örneği. *Van Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 6(1), 122–141. <https://dergipark.org.tr/tr/pub/yyuefd/issue/13711/165998>
- Dmitrenko, Natalia / Budas, Iuliia / Koliadych, Yulia / Poliarush, Nina** (2021): Impact of formative assessment on students' motivation in foreign language acquisition. *East European Journal of Psycholinguistics*, 8(2), 36-50. <https://doi.org/10.29038/eejpl.2021.8.2.dmi>
- Dörnyei, Zoltán** (2007): Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies. Oxford University Press. DOI:10.4000/asp.294
- Drackert, Anastasia / Nechaeva, Anastasia** (2018): *Eigene Fortschritte beim Russischlernen selbst beurteilen: Wie und wozu?* In: Y. Nishina (Hrsg.): *Beiträge zum Themenschwerpunkt „Nicht-indoeuropäische Sprachen an Hochschulsprachenzentren“* (S. 7–20). Fremdsprachen und Hochschule, 93.
- GER** (2001): *Gemeinsamer europäischer Referenzrahmen für Sprachen: Lernen, lehren, beurteilen*. Langenscheidt.

- Grotjahn, Rüdiger / Kleppin, Karin** (2017): *Prüfen, Testen, Evaluieren*. Berlin: Langenscheidt.
- Harlen, Wynne** (2007): *Assessment of learning*. London: Sage.
- Higgins, Richard / Hartley, Peter / Skelton, Alan** (2010): The conscientious consumer: Reconsidering the role of assessment feedback in student learning. *Studies in Higher Education*, 27(1), 53–64.
- Hughes, Arthur** (2003): *Testing for language teachers*. Cambridge University Press.
- Jones, Jane** (2006): Foreign languages in the primary school in England: A new pupil learning continuum. *Francophonie*, 5(31), 3–7.
- MEB** (2024): *Ölçme ve değerlendirme – Türkiye Yüzyılı Maarif-Modeli*. Abgerufen am 7. Mai 2025, von <https://tymm.meb.gov.tr/olcme-degerlendirme>
- Nationaler Bildungsbericht** (2024): *Bildung in Deutschland 2024: Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Künstlicher Intelligenz*. Bielefeld: wbv Publikation. <https://www.bildungsbericht.de/de/bildungsberichte-seit-2006/bildungsbericht-2024/pdf-dateien-2024/bildungsbericht-2024.pdf> (Letzter Zugriff: 07.06.2025).
- Ningsih, Nurul Rahayu / Rosidah, Nur Aini / Pradana, Dian Arief** (2024): *The role of formative assessment in developing English language curriculum and learning*. *International Journal of Education and Practice*, 12(1), 1–10. <https://www.researchgate.net/publication/388103671>
- Peyer, Elisabeth / Lüthi, Gabriela / Ravazzini, Nadia** (2023): Formative Beurteilung des Sprechens im DaF-Unterricht: Praktiken und Überzeugungen von Lehrpersonen der Sekundarstufe I. *Informationen Deutsch als Fremdsprache*, 50, 492 - 504. DOI:10.1515/infodaf-2023-0082
- Rea-Dickins, Pauline / Gardner, Sheena** (2000): Snares and silver bullets: Disentangling the construct of formative assessment. *Language Testing*, 17(2), 215–243.
- Reimann, Daniel** (2020): Methoden der Fremdsprachenforschung. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 25(1), 32–50. doi: 10.26083/tuprints-00021920
- Schütze, Birgit / Souvignier, Elmar / Hasselhorn, Marcus** (2018): Stichwort – Formatives Assessment. *Zeitschrift für Erziehungswissenschaft*, 21, 697 - 715. <https://doi.org/10.1007/s11618-018-0838-7>
- Swan, Malcolm / Foster, Colin** (2018): Formative assessment lessons. In: *Classroom assessment in mathematics: Perspectives from around the globe* (pp. 11-24). Cham: Springer International Publishing. [https://doi.org/10.1007/978-3-319-73748-5\\_2](https://doi.org/10.1007/978-3-319-73748-5_2)
- Şentürk, Rahim** (2025a): *Prüfen, Testen und Evaluieren im kommunikativen Grammatikunterricht der türkischen Deutschlehrerausbildung: Problemfelder, neue Ansätze und Implementierungsmöglichkeiten* [Doktora tezi, Trakya Üniversitesi Sosyal Bilimler Enstitüsü].
- Şentürk, Rahim** (2025b): *Mesleğe yeni başlayan Almanca öğretmenlerinin karşılaştıkları güçlükler ve çözüm önerileri*. In *Proceedings of the 14th International New York Academic Research Congresses on Social, Humanities, Administrative, and Educational Sciences* (New York, USA). <https://doi.org/10.30546/19023.978-9952-610-04-08.2025.5097>
- Şentürk, Rahim / Köksal, Handan** (2025): *The role of formative assessment in communicative grammar teaching: Opinions of German teacher candidates*. Vortrag präsentiert auf dem XXIII. IBANESS Congress Series on Economics, Business and Management, Plovdiv, Bulgarien.
- Taras, Maddalena** (2005): Assessment–summative and formative–some theoretical reflections. *British journal of educational studies*, 53(4), 466-478. <https://doi.org/10.1111/j.1467-8527.2005.00307.x>
- Tienson-Tseng, Heather L.** (2019): Best practices in summative assessment. In *Biochemistry Education: From Theory to Practice* (S. 219-243). American Chemical Society. 10.1021/bk-2019-1337.ch010.
- U.S. Department of Education** (2023): *Artificial intelligence and the future of teaching and learning: Insights and recommendations*. <https://www.ed.gov/sites/ed/files/documents/ai-report/ai-report.pdf>

**Üner, Berivan Tokur / Aşlıođlu, Bayram** (2022): İngilizce öğretiminde ölçme değerlendirme sürecine ilişkin öğretmen görüşleri. *EKEV Akademi Dergisi*, 89, 25-50. <https://dergipark.org.tr/en/pub/sosekev/issue/71371/1147452>

**Zhai, Xiaoming / Nehm, Ross H.** (2023): AI and formative assessment: The train has left the station. *Journal of Research in Science Teaching*, 60(6), 1390-1398. <https://doi.org/10.1002/tea.21885>