

Praktikables und nachhaltiges Testen fremdsprachlicher Leistung

Nancy Kontomitrou¹

This paper firstly analyses the quality criterion of “practicality” in developing language tests. In this context, Bachman and Palmer's (1996) classification of important resources for language testing is presented and examples of resources for each category are then discussed. In addition, the relationship between the quality criterion of “practicality” and the quality criterion of “economy” is examined. Finally, I analyse the factor of sustainability as an important aspect which should also play an important role in language testing. In this context, I propose that sustainability should be considered as a quality criterion of language tests. Subsequently, opportunities are presented to design sustainable and practicable formal and informal tests in order to save resources and protect the environment. Finally, conclusions are drawn from these proposals.

In diesem Beitrag wird zunächst das Qualitätskriterium der „Praktikabilität“ bei der Entwicklung von Sprachtests analysiert. In diesem Zusammenhang wird die Klassifizierung wichtiger Ressourcen für Sprachtests von Bachman und Palmer (1996) präsentiert und Beispiele von Ressourcen für jede Kategorie werden diskutiert. Darüber hinaus wird das Verhältnis zwischen dem Qualitätskriterium der „Praktikabilität“ und dem Qualitätskriterium der „Ökonomie“ untersucht. Abschließend analysiere ich den Faktor der Nachhaltigkeit als einen wichtigen Aspekt, der auch beim Testen fremdsprachlicher Leistung eine wichtige Rolle spielen sollte. In diesem Zusammenhang schlage ich vor, dass Nachhaltigkeit als Qualitätskriterium von Tests betrachtet werden sollte. Anschließend werden Möglichkeiten aufgezeigt, nachhaltige und praktikable formelle und informelle Tests zu entwerfen, um Ressourcen zu schonen und die Umwelt zu schützen. Schließlich werden aus diesen Vorschlägen Schlussfolgerungen gezogen.

1 Einleitung

Ein Sprachtest soll über bestimmte Qualitätsmerkmale verfügen, damit sein Ziel erreicht werden kann und damit man sich auf dessen Ergebnisse wissenschaftlich stützen kann. Die Erfüllung von Testgütekriterien bis zu einem hohen Grad kann auf eine hochqualitative Messung der fremdsprachlichen Leistung der Testkandidaten bzw. Lerner und auf das Vorhandensein einer Qualität bei der Beurteilung des Tests und seiner Ergebnisse hinweisen.

¹ Nancy Kontomitrou ist Dozentin für Deutsch als Fremdsprache im Fachbereich für Deutsche Sprache und Literatur der Nationalen und Kapodistrias-Universität Athen. Zu ihren wissenschaftlichen Interessen gehören die Fremdsprachendidaktik, der Fremdsprachenerwerb, die Analyse und Erstellung didaktischen Materials, die Evaluation im Fremdsprachenunterricht und die Testentwicklung und -beurteilung.

E-Mailadresse: nancykon@gs.uoa.gr

Die Berücksichtigung der „Praktikabilität“ als Gütekriterium von Tests ist ein Bereich, der mehrere Dimensionen beinhaltet, und diesem Gütekriterium sollte parallel zu den anderen Testgütekriterien ein hoher Stellenwert beigemessen werden. Dieses Gütekriterium ist sowohl bei der Erstellung von Tests als auch bei der Testanwendung und der Testauswertung zu berücksichtigen. In Zusammenhang mit diesem Kriterium steht auch die „Ökonomie“. Schließlich ist die Nachhaltigkeit, die heutzutage einen immer höheren Stellenwert in vielen Bereichen auf der ganzen Welt einnimmt, ein wichtiger Aspekt, der auch beim Testen fremdsprachlicher Leistung zu beachten ist.

2 Das Testgütekriterium der „Praktikabilität“

Nach Bachman / Palmer (1996, 36) bezieht sich das Qualitätskriterium der „Praktikabilität“ auf „das Verhältnis der zur Verfügung stehenden Ressourcen für die Testentwicklung, -konstruktion und -anwendung zu den dafür benötigten Ressourcen“². „Immer wenn die für eine Prüfung notwendigen Ressourcen die vorhandenen Ressourcen überschreiten, ist die Praktikabilität einer Prüfung unzureichend. Die Praktikabilität kann sich auf die Herstellung, die Durchführung und die Auswertung von Prüfungen beziehen.“ (Grotjahn / Kleppin 2015, 54). Testentwickler sollen sich also bei der Festlegung der Vorgaben von Tests Gedanken darüber machen, auf welche Art und Weise der Test entwickelt, durchgeführt und ausgewertet werden sollte, damit die Ressourcen, die für die Konstruktion, Durchführung und Auswertung des Tests erforderlich sind, auch mit den Ressourcen, über die man verfügt, übereinstimmen.

2.1 Klassifikation von Ressourcen nach Bachman/Palmer (1996)

Nach Bachman / Palmer (1996, 36/37) ist es von großer Wichtigkeit zu spezifizieren, was mit dem Begriff „Ressourcen“ gemeint ist, um das Gütekriterium der Praktikabilität verstehen zu können. Um die Ressourcen zu klassifizieren, sind drei „allgemeine Kategorien“ (ebd. 36) zu berücksichtigen:

- a) Die „zur Verfügung stehenden menschlichen Leistungspotenziale“³ („human resources“)
- b) Die „materiellen Ressourcen“ („material resources“)
- c) Die „Zeit“ („time“).

² Bachman / Palmer (1996, 36) definieren Praktikabilität folgendes: „We can define Practicality as the relationship between the resources that will be required in the design, development, and use of the test and the resources that will be available for these activities”.

³ „Human Resources“ auf Duden online. URL: https://www.duden.de/rechtschreibung/Human_Resources (Abrufdatum 24.6.22)

Bei der ersten Kategorie handelt es sich um die menschlichen Leistungspotenziale bzw. die Arbeitskräfte, die bei der Erstellung, Durchführung und Auswertung von Tests benötigt werden. Zu den Human Resources gehören die Testautoren⁴, die Bewerter⁵, die Testadministratoren⁶ und das für die Unterstützung von Tests benötigte Büro- und technische Personal⁷.

Die zweite Kategorie betrifft Aspekte, die mit der Ausstattung und den Materialien zu tun haben. In Bezug auf die Ausstattung ist die räumliche Ausstattung⁸, wie zum Beispiel Räume für die Entwicklung, die Durchführung von Tests und Räume für die Verwaltung (s. Bachman / Palmer 1996, 37), ein wichtiger Faktor. Zudem ist mit dem Aspekt der Ausstattung auch technisches Material, wie zum Beispiel Computer, Textverarbeitungsprogramme, Videorecorder (s. ebd.), Drucker, Kopiergeräte, und automatische Geräte zur Notengebung, gemeint. Schließlich gehören zu dieser Kategorie auch Materialien, wie zum Beispiel Papier und Bilder (s. ebd.).

Die dritte Kategorie hat mit dem zeitlichen Faktor zu tun und betrifft die Zeit für die Entwicklung des Tests⁹, nämlich die Zeit von dem Beginn des Prozesses der Testerstellung bis zur Bekanntmachung der Noten (s. ebd.), und die Zeit für bestimmte Aktivitäten¹⁰ (z.B. Konstruktion, Schreiben, Verwaltung, Notengebung, Analyse)¹¹.

2.2 Praktikabilität bei der Erstellung, Anwendung und Auswertung von Tests

Die im vorigen Unterkapitel präsentierten Kategorien von Ressourcen sind, wie schon erwähnt wurde, sowohl bei der Erstellung als auch bei der Testanwendung und der Testauswertung von großer Bedeutung.

Wenn man versucht zu kategorisieren, welche spezifischen Ressourcen bei der Testerstellung, welche bei der Testanwendung und welche bei der Testauswertung benötigt werden, könnten anhand von Beispielen folgende Bereiche genannt werden:

Human Resources

- Testerstellung: Testadministratoren, wissenschaftliche Verantwortliche / Testdirektoren, Testentwickler bzw. Testautoren, Testpersonen für die

⁴ „test writers“ (Bachman / Palmer (1996, 36)

⁵ „scorers or raters“ (ebd.)

⁶ „test administrators“ (ebd.)

⁷ „clerical and technical support personnel“ (ebd. 37)

⁸ „space“ (ebd.)

⁹ „Development time“ (ebd.)

¹⁰ „Time for specific tasks“ (ebd.)

¹¹ „designing, writing, administering, scoring, analyzing“ (ebd.)

Pilotphase des Tests, Korrektoren / Korrektorinnen, zuständige Personen für das Layout des Tests, zuständige Personen für die technische Bearbeitung des Sounds für die Aktivität „Hörverstehen“, zuständige Personen für die Aufnahme von bestimmten Testaufgaben, zuständige Personen für die Schulung der Prüfer mündlichen und schriftlichen Ausdrucks

- Testanwendung: Testadministratoren, Prüfer / Prüferinnen mündlichen Ausdrucks, Beobachter des Prüfungsverfahrens mündlicher Prüfungen, technisches Personal für die Kontrolle des Sounds, Personal für das Kopieren / Verteilen der Kandidatenblätter / Prüferhefte, Personal für die Prüfungsaufsicht bei schriftlichen Tests, Personal für die Begleitung der Testkandidaten bei mündlichen Prüfungen zu den Prüfungsräumen, Personal für die Sammlung und Aufbewahrung der Antwortbögen, Personal für das Senden der Antwortbögen der Testkandidaten an die zuständige Zentralstelle
- Testauswertung: Testadministratoren, Personal für die Bewerter-schulung und die Vorbereitung der Materialien für die Bewerber schriftlichen Ausdrucks, Koordinatoren der Bewerber schriftlichen Ausdrucks, technisches Personal für die automatische Berechnung der Gesamtnoten, Personal für die Bekanntmachung der Testergebnisse, Personal für die Vorbereitung der Zertifikate, Mitarbeiter für die statistische Analyse der Ergebnisse und die Beurteilung der Prüfung nach den Testgütekriterien

Materielle Ressourcen

- Testerstellung:
Räume: Räume für die Testadministration, für Testdesigner, für die Schulung von Testautoren, technisch ausgestatteter Raum / Studio für die Vorbereitung der Hördateien für das Testen des Hörverstehens
Ausstattung: Computer und Internetanschluss, technische Geräte für die Vorbereitung der Hördateien für das Testen des Hörverstehens, Programme für das Layout
Materialien: Templates der verschiedenen Prüfungsphasen, Computer, USB-Sticks, Antwortbögen / Prüfungshefte, Kugelschreiber / Bleistifte
- Testanwendung:
Räume: für die Testadministration, Räume für die Durchführung von schriftlichen Prüfungen, spezifisch ausgestattete Räume im Falle des Testens der mündlichen Rezeption, spezifisch ausgestattete Räume im Falle des Testens des Hör-Seh-Verstehens (z.B. mit Videogeräten oder Computer / Laptops / Tablets / Kopfhörer), Räume für die Durchführung von mündlichen Prüfungen

Ausstattung: Computer / Laptops / Tablets und Internetanschluss in den Räumen für die Testadministration, Computer / Laptops / Tablets und Kopfhörer im Raum der Prüfung im Falle eines computerbasierten Tests, Videogeräte und Kopfhörer im Raum der Prüfung im Falle des Testens des Hör-Seh-Verstehens, Kopiergeräte, VBI¹²

Materialien: Antwortbögen / Prüfungshefte, Kugelschreiber / Bleistifte

- Testauswertung

Räume: für die Testadministration, für die Bewerter-schulung, für das Bewertungsverfahren, für die Aufbewahrung der Antwortbögen der Testkandidaten, für technische Geräte, die bei der Berechnung der Endnoten nötig sind.

Ausstattung: Geräte für die automatische Berechnung der Noten, Computer und Internetanschluss, Regale für die Aufbewahrung der schriftlichen Leistungen und der Antwortbögen der Testkandidaten

Materialien: Antwortbögen / Prüfungshefte der Testkandidaten, Kugelschreiber (meist schwarze oder blaue), Bewerterausweise mit der Bewerbernummer, Aufkleber für die Verbergung der Noten des 1. Bewerter im Falle von 2 Bewertern, Blätter Papier für die Notizen der Bewerter

Zeit

- Testerstellung: Zeit für die organisatorischen Aspekte der Prüfung, für die Bestimmung des Ziels der Prüfung, für die Entwicklung der Vorgaben, für die Entwicklung der Testaufgaben / Items, für die Entwicklung der Skalen zur Bewertung mündlichen und schriftlichen Ausdrucks, für die Pilotphase der Prüfung, für mögliche Änderungen nach der Pilotphase, für das Testdesign, für die Aufnahme der Hördateien
- Testanwendung: Zeit für organisatorische Aspekte kurz vor, während und kurz nach der Durchführung der Prüfung, für das Kopieren und Verteilen der Prüfungsaufgaben / der Kandidatenblätter, für die Bearbeitung der Prüfungsaufgaben jeder Prüfungsphase, zwischen den verschiedenen Prüfungsphasen, das Sammeln der Antwortbögen / Prüfungshefte
- Testauswertung: Zeit für organisatorische Aspekte, für die Bewerter-schulung, für die Bewertung schriftlichen Ausdruck und von Ergänzungsaufgaben, für die Berechnung der Noten, für die Bekanntmachung der Testergebnisse, für die Analyse der Testergebnisse, für die Berechnung der Schwierigkeit der Aufgaben, für die Beurteilung der Prüfung nach den Testgütekriterien

¹² Diese technische Ausstattung wird beim Staatlichen Sprachzertifikat (KPG) verwendet: https://rcel2.enl.uoa.gr/kpg/images/2nd_english.png

3 „Praktikabilität“ und „Ökonomie“

Für Bachman / Palmer (1996, 35/36) handelt es sich bei der Praktikabilität um ein Kriterium, das eine wichtige Rolle für die Erfüllung des übergeordneten Kriteriums der „Nützlichkeit“ spielt. Die erwähnten Autoren „berücksichtigen auch den finanziellen Faktor als wesentliches Merkmal der Praktikabilität eines Tests (Wiedenmayer 2006, 30). In diesem Zusammenhang ist es wichtig die Kosten einzuschätzen, so dass auch überlegt wird, ob die Summe der verfügbaren Ressourcen in der spezifischen Testsituation ausreichen (Bachman / Palmer 1996, 37), da „in vielen Testsituationen die Summe der verfügbaren Ressourcen im Wesentlichen eine Funktion des Budgets sind“¹³ (ebd.)

Andere Wissenschaftler erwähnen bei den Gütekriterien nicht das Kriterium der „Praktikabilität“, sondern das Kriterium der „Ökonomie“, das als ein Nebengütekriterium von Tests betrachtet wird.

„Ein Test ist dann ökonomisch, wenn er

- eine kurze Durchführungszeit beansprucht,
- wenig Material verbraucht,
- einfach zu handhaben,
- als Gruppentest durchführbar,
- schnell und bequem auszuwerten ist“ (Lienert / Raatz 1998, 12).

Obwohl beide Begriffe sich auf Ressourcen beziehen, können sie nicht als Synonym verwendet werden. Beim Kriterium der Praktikabilität ist zu berücksichtigen, ob die Ressourcen, die bei der Testerstellung, Testanwendung und Testauswertung geplant und benötigt werden, auch verfügbar sind. Obwohl bei der Erfüllung dieses Kriteriums auch ökonomische Faktoren eine wichtige Rolle spielen, ist nicht auszuschließen, dass bei einer hohen Verfügbarkeit von Ressourcen, was nicht immer ökonomisch ist, eine Praktikabilität vorhanden ist. Eine schnelle und bequeme Auswertung bei einem Test kann zum Beispiel zur Erfüllung des Kriteriums der Ökonomie beitragen, wenn zum Beispiel ein Bewerter mit einer geeigneten und bequemen Bewertungsskala die schriftlichen Leistungen der Kandidaten bewertet. Wenn aber die Human Resources es erlauben, dass mehr Bewerter die schriftlichen Leistungen der Kandidaten bewerten, ist der Verlauf des Bewertungsverfahrens nicht so schnell und nicht so ökonomisch, aber das Kriterium der Praktikabilität kann wegen des Vorhandenseins der Human Resources in diesem Fall erfüllt werden.

Aus diesen Überlegungen lässt sich feststellen, dass es auf die Situation ankommen könnte, ob die Erfüllung des Kriteriums der Praktikabilität auch die

¹³ „...in many testing situations the total amount of resources available is essentially a function of budget“ (ebd.).

Erfüllung des Kriteriums der Ökonomie bedeutet. Natürlich wird in vielen Fällen festgestellt, dass praktikable Tests auch ökonomische Tests sind.

In Bezug auf das Kriterium der Praktikabilität ist es auch wichtig zu berücksichtigen, dass dieses Kriterium „für eine bestimmte Testsituation“¹⁴ (Bachman / Palmer 1996, 37) eingeschätzt werden kann, denn „verschiedene Typen von Ressourcen und die damit verbundenen Kosten in einem unterschiedlichen Grad in den verschiedenen Testentwicklungsphasen und bei der Testanwendung benötigt werden“¹⁵ (ebd.).

4 „Praktikabilität“ und „Nachhaltigkeit“

In Zusammenhang mit dem Kriterium der „Praktikabilität“ sollte auch ein ebenfalls wichtiger Aspekt betont werden, nämlich die „Nachhaltigkeit“, die auch beim Testen fremdsprachlicher Leistung eine wichtige Rolle spielen sollte und einen Beitrag zum Umweltschutz leisten kann. „Wir sollten mit den begrenzt zur Verfügung stehenden Ressourcen sorgsam umgehen und nicht auf Kosten der Menschen in anderen Regionen der Erde und auf Kosten zukünftiger Generationen leben. Nachhaltigkeit betrifft unsere Umwelt, alle Bereiche unseres Lebens und Wirtschaftens. Nachhaltiges Handeln ist also eine Aufgabe der ganzen Gesellschaft – national und international. Wir müssen unsere Erde für alle und auf Dauer bewohnbar erhalten.“ (<https://www.bmu.de/themen/nachhaltigkeit-digitalisierung/nachhaltigkeit>). Es gibt viele Möglichkeiten den Aspekt der Nachhaltigkeit auch bei den Aspekten, die das Testen fremdsprachlicher Leistung betreffen, zu berücksichtigen.

Auf der Grundlage der oben präsentierten, die Nachhaltigkeit betreffenden Aspekte könnte die Frage gestellt werden, ob die „Nachhaltigkeit“ als ein weiteres Gütekriterium von Tests angesehen werden könnte. Dieses Kriterium könnte wie folgt definiert werden: Ein Test ist nachhaltig, wenn in seiner Testentwicklung, Testanwendung und Testauswertung der Aspekt der Umwelt berücksichtigt wird und wenn dabei ein sorgsamer Umgang mit den auf der Erde begrenzt verfügbaren Ressourcen berücksichtigt wird.

4.1 Möglichkeiten der Berücksichtigung der Praktikabilität und der Nachhaltigkeit in Sprachtests

Es gibt viele Möglichkeiten praktikable Tests zu entwickeln und die Nachhaltigkeit auch in dem Bereich des Testens von Fremdsprachen zu berücksichtigen. Die Berücksichtigung der Nachhaltigkeit kann aber nicht immer das Vorhandensein

¹⁴ „...Practicality can only be determined for a specific testing situation” (ebd.)

¹⁵ „Different types of resources, and their associated costs, will be required in varying degrees at different stages of test development and use” (ebd.)

der Praktikabilität gewährleisten. Je nach Situation können nachhaltig entwickelte Tests mehr oder weniger praktikabel sein, denn es liegt an den Ressourcen an, die bei jeder Testsituation zur Verfügung stehen.

Eine Möglichkeit praktikables und nachhaltiges Testen zu gewährleisten sind gestaffelte Prüfungen. Ein Beispiel einer gestaffelten Prüfung ist die Prüfung des Staatlichen Sprachzertifikats (KPG), die vom griechischen Ministerium für Bildung und Religionsangelegenheiten organisiert und verwaltet wird. In dieser Prüfung wird dafür gesorgt, dass Kosten für die Teilnahme an der Prüfung und Zeit (Vgl. Τσοπάνογλου, in Zusammenarbeit mit Δεδρίνοβ 2012, 8) gespart werden. „Die Testkandidaten werden bei der Prüfung auf dem Niveau A sowohl auf dem A1 als auch auf dem A2, so auch auf den anderen Niveaus B und C, getestet. Je nach ihrer Leistung und je nach erreichter Punktzahl (s. auch https://rcel2.enl.uoa.gr/kpg/files/KPG_Prodiagrafes.pdf) können die Testkandidaten entsprechend ein Zertifikat auf dem Niveau A1 oder A2, auf dem Niveau B1 oder B2 oder auf dem Niveau C1 oder C2 erhalten“ (Kontomitrou 2021, 24). Die Prüfung des Staatlichen Sprachzertifikats ist praktikabel und „Faktoren, die bei der Beurteilung der Prüfung als praktikabel eine wichtige Rolle spielen, sind die praktikable Bewertung und Interpretation der Ergebnisse durch das gemischte Bewertungssystem, die praktikable Testdurchführung (z.B. im mündlichen Teil, für den ein Leitfaden für Prüfer den Prüferinnen und Prüfern zur Verfügung gestellt wird), die Möglichkeit von online mündlichen Prüfungen in Orten, in denen kein Prüfer / keine Prüferin zur Verfügung steht, durch die gestaffelte Prüfung von Niveaus und durch die geschlossenen Items und den Ergänzungsaufgaben, die meist mit einer als korrekt geltenden Antwort zu ergänzen sind“ (ebd. 27). Diese Faktoren können auch für eine Nachhaltigkeit beim Testen sorgen.

Eine weitere Möglichkeit nachhaltige Tests zu entwickeln ist der Einsatz von digitalen Tests. Solche Tests sind „ressourcenschonend und nachhaltig“ (<https://www.goethe.de/ins/gr/de/sta/ath/prf/ddd.html>) und werden als Möglichkeit parallel zu der Möglichkeit von papierbasierten Tests schon bei vielen Prüfungen eingesetzt, wie zum Beispiel bei Prüfungen des Goethe-Instituts (<https://www.goethe.de/ins/gr/de/sta/ath/prf/ddd.html>) und des TestDaF-Instituts (<https://www.testdaf.de/de/>). Auch für das Staatliche Sprachzertifikat (KPG) wird ein Projekt die digitale Form der KPG-Prüfung (e-KPG) ermöglichen. Dieses Projekt wird die Vorbereitung und Durchführung des e-KPG Tests für die Sprachen „Englisch“, „Französisch“ und „Deutsch“ (<https://www.espa.gr/el/Pages/ProclamationsFS.aspx?item=5219>) zum Gegenstand haben und wird außer den weiteren Vorteilen, auch dazu beitragen, dass die KPG-Prüfung auch zu einem noch höheren Grad nachhaltig ist. Die Nachhaltigkeit, die durch die digitale Form der e-KPG Prüfung berücksichtigt wird, wird auch durch den vorgesehenen adaptiven

Charakter (ebd.) parallel zu dem schon vorhandenen gestaffelten Charakter der Prüfung, der auch bei der digitalen Form behalten wird, verstärkt.

Computer-adaptive Tests sind also noch eine Möglichkeit die Nachhaltigkeit und auch die Praktikabilität zu berücksichtigen. „Der Schwierigkeitsgrad der dem Prüfling gestellten Aufgaben richtet sich dabei nach dessen bisherigem Abschneiden. Der adaptive Test benötigt weniger Aufgaben zur Bewertung eines Kandidaten und damit weniger Zeit“ (Gutzat / Pauen / Voss 2001, 13). Da computer-adaptive Tests unter anderem auch darauf abzielen „das Testverfahren zu verkürzen“ (Tschirner 2001, 72), sind sie praktikabel und auch nachhaltig. Diese Tests sind auch nachhaltig, da keine Kandidatenblätter verwendet werden und so weniger Papier verbraucht wird.

Schließlich bieten einige Institutionen die Möglichkeit von zu Hause aus eine Sprachprüfung abzulegen. Ein Beispiel ist die Version „TOEFL iBT Home Edition“ (<https://www.ets.org/toefl/test-takers/ibt/why/options>). Eine solche Version von Tests ist außer praktikabel auch nachhaltig, da Kosten und Ressourcen gespart werden können.

4.2 Vorschläge für die Entwicklung und den Einsatz von praktikablen und nachhaltigen Tests im Fremdsprachenunterricht

Praktikabilität und Nachhaltigkeit sind zwei Aspekte, die sowohl bei formellen Tests, also Tests „die häufig präzisierend als standardisiert bezeichnet werden“ (Grotjahn 2001, 33) und „das Ergebnis langwieriger Bemühungen von Testspezialisten sind“ (ebd.), als auch bei informellen Tests, bei denen es sich „meist um ein deutlich weniger aufwändiges Erzeugnis von Unterrichtspraktikern“ (ebd.) sind und „sich zudem in der Regel auf eine konkrete Unterrichtssituation“ (ebd.) beziehen, einen wichtigen Stellenwert einnehmen sollten. Da es bei Tests, die das Erreichen von bestimmten Lernzielen nach einer oder mehreren Unterrichtseinheiten zum Gegenstand haben, die Entwicklung von digitalen Tests, die Entwicklung von gestaffelten Prüfungen oder die Entwicklung von computer-adaptiven Tests eventuell schwer zu erstellen, nicht praktikabel und in vielen Fällen nicht geeignet wären, sollte nach mehreren Möglichkeiten nachgedacht werden, die ein praktikables und nachhaltiges Testen erlauben könnten.

Ein Vorschlag wäre zum Beispiel, das Drucken und Kopieren von Kandidatenblättern zu reduzieren. Das kann zum Beispiel mit Hilfe von Gruppenprüfungen erreicht werden. Jede Gruppe könnte die Testaufgaben auf einem Computer oder Tablet bekommen, wenn solche elektronischen Geräte verfügbar sind, und jeder Testkandidat / jede Testkandidatin könnte die Antworten auf einem recycelten Blatt Papier schreiben. Wenn elektronische Geräte nicht verfügbar sind,

könnte jeder Testkandidat / jede Testkandidatin der Gruppe mit einer unterschiedlichen Seite des Tests beginnen.

Eine weitere Möglichkeit sparsam mit den verfügbaren Ressourcen umzugehen, wäre eine Datenbank mit informellen Tests oder Testaufgaben zu erstellen und auf diese Weise Zeit, materielle Ressourcen und Human Ressourcen zu reduzieren, indem derselbe Test mehrmals in verschiedenen Gruppen, die denselben Unterrichtsstoff mit denselben Lernzielen bearbeitet haben, angewendet wird.

Schließlich besteht die Möglichkeit festzustellen, ob eine Verkürzung eines Tests möglich ist, indem man zum Beispiel feststellt, ohne welche Testaufgaben weitere wichtige Gütekriterien von Tests nicht verringert werden. Wenn festgestellt wird, dass die Reliabilität, die Validität und die Objektivität hoch bleiben, auch wenn bei einigen Tests die Testaufgaben reduziert werden, dann wäre die Verkürzung eines Tests eine Möglichkeit auch mit den Ressourcen sparsam umzugehen und nachhaltige und praktikable Tests im DaF-Unterricht zu erstellen, anzuwenden und auszuwerten.

5 Schlussfolgerung

Aus den in diesem Beitrag präsentierten und analysierten Möglichkeiten und Vorschlägen, praktikable und nachhaltige Tests zu erstellen, lässt sich schlussfolgern, dass das Kriterium der Praktikabilität eine wichtige Rolle bei der Testentwicklung, -anwendung und -auswertung spielen sollte. In diesem Zusammenhang sollte außerdem nach neuen Möglichkeiten nachgedacht werden, die die Erfüllung dieses Kriteriums ermöglichen. Beim Prozess des Testens fremdsprachlicher Leistung sollte auch die Nachhaltigkeit berücksichtigt werden und als weiteres Gütekriterium von Tests betrachtet werden. Für die Erfüllung dieses Kriteriums sollte auch untersucht werden, welche weiteren Möglichkeiten es heutzutage und in Zukunft geben könnte, nachhaltiges Testen fremdsprachlicher Leistung zu gewährleisten.

Eine Weiterforschung im Bereich der Berücksichtigung der Praktikabilität und der Nachhaltigkeit bei digitalen und papierbasierten Sprachtests ist notwendig und es sollte auch in der didaktischen Praxis das Feedback von Lehrenden, die diese Möglichkeiten einsetzen, erforscht und analysiert werden. Wichtig dabei ist auch den Washback-Effekt von praktikablen und nachhaltigen Tests zu erforschen und festzustellen, welche weiteren Bereiche der Bildung von solchen Tests profitieren könnten. Der Zusammenhang von Theorie und Praxis ist für das Erreichen bestmöglicher Forschungsergebnisse im Bereich der Entwicklung von praktikablen und nachhaltigen Sprachtests von erheblicher Bedeutung.

Literaturverzeichnis

- Bachman, Lyle F. / Palmer, Adrian S.: Language Testing in Practice: Designing and Developing Useful Language Tests [1996]. Oxford: Oxford University Press
- Gutzat, Bärbel / Pauen, Peter / Voss, Josef: Computer- und Internetinsatz bei TestDaF. In P. Tischer (Hrsg.): Neue Medien - Neues Lehren - Neues Lernen: Fremdsprachenunterricht im Informationszeitalter (S. 267–284). Bochum: AKS-Verlag. (2001): <<https://www2.testdaf.de/fileadmin/Redakteur/PDF/Forschung-Publikationen/aks.pdf>>
- Grotjahn, Rüdiger: Leistungsmessung und Leistungsbeurteilung [2001]. Band A. Πάτρα: Ελληνικό Ανοικτό Πανεπιστήμιο
- Grotjahn, Rüdiger / Kleppin, Karin: Prüfen, Testen, Evaluieren. Deutsch Lehren Lernen [2015]. Band 7. München: Klett-Langenscheidt
- Kontomitrou, Nancy: Staatliches Sprachzertifikat. Testentwicklung und Qualitätskriterien. In: Aktuell. Band 52, ΠΕΚΑΓΕΠΕ (2021): 24-29. <<https://www.pdv.org.gr/img/Aktuell52.final.pdf>>
- Lienert, Gustav. A. / Raatz, Ulrich: Testaufbau und Testanalyse [1998]. Weinhheim: Beltz Psychologie Verlags Union.
- Tschirner, Erwin: Leistungsmessung und Leistungsbeurteilung [2001]. Band B. Πάτρα: Ελληνικό Ανοικτό Πανεπιστήμιο
- Τσοπάνογλου, Αντώνης, σε συνεργασία με την Βασιλική Δενδρινού: Επιλογές του ΚΠΓ για την αποτίμηση της γλωσσομάθειας (2012). RCeL. <http://rcel.enl.uoa.gr/periodical/articles/Article7_Tsopanoglou_Dendrinou_periodical.pdf>
- Wiedenmayer, Dafni: DaF-Testen. Testentwicklung und Testbeurteilung [2006]. Athen: DaF extra Verlag

Quellen aus dem Internet

- <https://www.bmuv.de/themen/nachhaltigkeit-digitalisierung/nachhaltigkeit>
- <https://www.espa.gr/el/Pages/ProclamationsFS.aspx?item=5219>
- <https://www.goethe.de/ins/gr/de/sta/ath/prf/ddd.html>
- <https://www.testdaf.de/de/>
- https://rcel2.enl.uoa.gr/kpg/files/KPG_Prodiagrafes.pdf
- <https://www.ets.org/toefl/test-takers/ibt/why/options>
- https://rcel2.enl.uoa.gr/kpg/images/2nd_english.png
- https://www.duden.de/rechtschreibung/Human_Resources