

Prüfen, Testen, Evaluieren

Rüdiger Grotjahn
Karin Kleppin

Klett-Langenscheidt

München

Fort- und Weiterbildungsreihe des Goethe-Instituts
Allgemeiner Herausgeber: Goethe-Institut e.V.
Herausgeber dieser Einheit: Karin Ende, Goethe-Institut e.V.
Fachlektorat: Karin Ende, Imke Mohr
Redaktion: Kathrin Stelzer, Coleen Clement
Manuskriptbetreuung: Steffi Borneleit
Illustrationen: Daniela Kohl, heutebunt
Filmproduktion: Magnus Froböse, Context Film

Diese Einheit ist Bestandteil der Fort- und Weiterbildungsreihe *Deutsch Lehren Lernen*.

1. Auflage 1 5 4 3 2 1 | 2019 18 17 16 15

© 2015 Goethe-Institut, München

Das Werk und seine Teile sind urheberrechtlich geschützt.

Jede Verwertung in anderen als den gesetzlich zugelassenen Fällen bedarf deshalb der vorherigen schriftlichen Einwilligung des Verlages.

Koordination: Sabine Wenkums

Verlagsredaktion: Esther Debus-Gregor

Covergestaltung: Bettina Lindenberg, München

Innengestaltung: Ute Weber GrafikDesign, Geretsried

Satz: kaltner verlagsmedien GmbH, Bobingen

Gesamtherstellung: Print Consult GmbH, München

ISBN 978-3-12-606970-0

ISBN 978-3-12-606970-0



Symbole



Arbeitsblatt/Arbeitsmaterialien



www-Link



Auszug aus einem Lehrwerk



Video-QR-Code: weitere Informationen auf Seite 6.



Video



Information zum verwendeten Video



Aufgabe möglichst in Partner- oder Gruppenarbeit zu bearbeiten



relevant für Präsenzphase, wenn kombinierte Fortbildungsveranstaltung besucht wird

Blau gesetzte Begriffe werden im Glossar erklärt. Sie sind dann blau, wenn sie eingeführt werden.

Weitere Informationen finden Sie auf www.goethe.de/dll.

Inhalt

| | |
|--|-----------|
| Vorwort | 5 |
| Einleitung | 7 |
| 1 Prüfen, Testen, Evaluieren: Eine erste Annäherung | 8 |
| 1.1 Ihre Perspektive als Lehrkraft | 8 |
| 1.2 Die Perspektive Ihrer Lernenden | 10 |
| 1.3 Funktionen des Prüfens, Testens und Evaluierens | 12 |
| 1.4 Begriffliche Festlegungen | 15 |
| 2 Rahmenbedingungen für das Prüfen, Testen und Evaluieren | 16 |
| 2.1 Handlungs- und Kompetenzorientierung | 17 |
| 2.2 Kompetenzniveaus und Kann-Beschreibungen | 21 |
| 2.3 Kompetenzstandards und Curricula | 29 |
| 3 Prüfen, Testen, Evaluieren: Grundlagen | 33 |
| 3.1 Grundlegende Typen der Evaluation | 33 |
| 3.2 Was wollen und können wir prüfen? | 40 |
| 3.3 Isolierte und integrierte Überprüfung von Kompetenzen | 48 |
| 3.4 Gütekriterien von Prüfungen und Tests | 50 |
| 3.5 Standardisierte Prüfungen: Beispiele | 55 |
| 3.6 Prüfungs- und Aufgabenspezifikationen | 57 |
| 3.7 Aufgabenformate in Prüfungen | 58 |
| 3.7.1 Aufgabenformate und Testkonstrukt | 59 |
| 3.7.2 Offene versus geschlossene Aufgabenformate | 61 |
| 3.7.3 Prüfungsziele und Prüfungsaufgaben | 67 |
| 3.8 Arbeitsanweisung und Situierung der Prüfungsaufgabe | 70 |
| 4 Bewertung und Rückmeldung von Prüfungsleistungen | 72 |
| 4.1 Bewertung der Prüfungsleistungen | 72 |
| 4.2 Rückmeldung der Prüfungsleistungen | 85 |
| 5 Prüfungsaufgaben für den Unterricht | 87 |
| 5.1 Hör-Sehverstehen und Leseverstehen überprüfen | 88 |
| 5.2 Schreiben überprüfen | 92 |

| | | |
|----------|---|------------|
| 5.3 | Sprechen überprüfen | 94 |
| 5.4 | Sprachmittlungskompetenz überprüfen | 97 |
| 5.5 | Sprachliches Wissen überprüfen: Lexik und Grammatik | 100 |
| 5.6 | Soziopragmatische und interkulturelle Kompetenzen überprüfen | 103 |
| 5.7 | Prüfungsaufgaben selbst erstellen | 106 |
| 5.7.1 | Prüfungsaufgabe zum Leseverstehen | 106 |
| 5.7.2 | Prüfungsaufgabe zum interaktiven Sprechen | 116 |
| 5.8 | Erfolgchancen bei Prüfungen erhöhen: Sinnvolle Test-Taking-Strategien vermitteln | 122 |
| 5.9 | Optimierung der eigenen Prüfungspraxis | 125 |
| 6 | Informelles Evaluieren im Unterricht | 127 |
| 6.1 | Evaluieren durch die Lehrkraft | 127 |
| 6.1.1 | Geplante und ungeplante Evaluation | 127 |
| 6.1.2 | Interaktionistische dynamische Evaluation | 129 |
| 6.2 | Evaluieren durch die Lernenden | 139 |
| 6.2.1 | Selbstevaluation | 140 |
| 6.2.2 | Peer-Evaluation | 145 |
| 7 | Praxiserkundungsprojekte planen | 147 |
| | Lösungsschlüssel | 149 |
| | Glossar | 161 |
| | Literaturhinweise | 167 |
| | Quellenangaben | 173 |
| | Angaben zu den Autoren | 175 |
| | Übersicht über die Videos auf der DVD | 176 |

3.2 Was wollen und können wir prüfen?


Wenn Sie dieses Teilkapitel bearbeitet haben, dann

- können Sie bei vielen vorgefertigten Prüfungen erkennen, was mit den Aufgaben überprüft werden soll,
- können Sie einschätzen, welche spezifischen Teilkompetenzen mit den entsprechenden Aufgaben tatsächlich überprüft werden können und inwieweit diese durch die Prüfung abgedeckt bzw. nicht abgedeckt sind,
- können Sie bei vielen vorliegenden Prüfungen erkennen, ob auch noch andere Merkmale die Testleistung beeinflussen,
- wissen Sie, dass Sie diese Erkenntnisse auch bei der Erstellung eigener Prüfungsaufgaben berücksichtigen sollten.

Wir laden Sie zunächst ein, drei Prüfungsaufgaben zu lösen und an sich selbst zu beobachten, wie Sie dabei vorgehen und welche Kompetenzen und Strategien Sie dabei aktivieren.


Versuchen Sie, während der Bearbeitung der folgenden Prüfungsaufgaben A–C laut zu denken und sich eventuell dabei selbst auf einem Tonträger aufzunehmen. Dies erlaubt es Ihnen, Ihren Lösungsweg später nachzuvollziehen. Hier sind die Prüfungsaufgaben A–C.

Prüfungsaufgabe A



Seite 4

Modellsatz 02 Leseverstehen
 Lesetext 1: Aufgaben 1 – 10 ca. 10 Min.



Nebenjobs und Praktika

Sie suchen für Ihre Bekannten Jobs bzw. Praktika. Schreiben Sie in das Kästchen rechts den Buchstaben des passenden Angebots. Jedes Angebot kann nur einmal gewählt werden. Es gibt nicht für jede Person ein passendes Angebot. Gibt es für eine Person kein geeignetes Stellenangebot, dann schreiben Sie den Buchstaben *I*. Das Angebot im Beispiel kann nicht mehr gewählt werden.

Sie suchen ein passendes Angebot für ...

| | | | |
|------|--|---|------|
| (01) | ... eine Bekannte, die ein Praktikum in einer Bank machen muss. | A | (01) |
| (02) | ... Ihren Bekannten, der dringend ein Praktikum bei einer Computerfirma sucht. | I | (02) |
| 1 | ... Ihre jüngere Schwester, die sich für Journalismus interessiert. | | 1 |
| 2 | ... eine Kommilitonin, die gerne Studierende aus anderen Ländern betreuen möchte. | | 2 |
| 3 | ... Ihren Freund, der Fremdsprachen studiert und bei einer Zeitung im europäischen Ausland ein Praktikum absolvieren möchte. | | 3 |
| 4 | ... eine Schülerin, die Erfahrungen mit Online-Zeitschriften sammeln möchte. | | 4 |
| 5 | ... einen befreundeten Schüler, der für die Sommerferien einen Job im Ausland sucht. | | 5 |
| 6 | ... eine Kunststudentin, die für die Semesterferien einen Job im Bereich Kulturmanagement sucht. | | 6 |
| 7 | ... Ihre deutsche Kommilitonin, die an einer Uni im Ausland arbeiten möchte. | | 7 |
| 8 | ... Ihren Freund, der Politikwissenschaften studiert und dringend einen Job neben dem Studium braucht. | | 8 |
| 9 | ... Ihren deutschen Studienkollegen, der in den Semesterferien Jugendliche in Frankreich betreuen möchte. | | 9 |
| 10 | ... eine Freundin, die gerne bei einer Kabarett- oder Theatergruppe mitmachen möchte. | | 10 |



Nebenjobs und Praktika

A

Karriere machen in der Finanzwirtschaft?

Dann sollten Sie sich um einen Praktikumsplatz bei Finanz-Dienstleistungsunternehmen oder im Bankwesen bemühen. Bewerben können sich Interessierte, die bereits eine Ausbildung im Bereich Finanzwesen gemacht haben, aber auch Studierende der Wirtschaftswissenschaften. Nähere Informationen unter: www.finanzenundwirtschaft.de

B

Organisationstalent gesucht

Die Abteilung „Internationales“ des Akademischen Auslandsamtes sucht eine Tutorin oder einen Tutor. Der Aufgabenbereich umfasst die Unterstützung ausländischer Studierender, insbesondere bei der Integration in den Universitätsalltag. Wir erwarten überdurchschnittliches Geschick in der Organisation sowie im sozialen und interkulturellen Umgang. Bewerbungen bitte an:
Frau Bäumler (Verwaltungsgebäude Raum 03).

C

Arbeiten und gleichzeitig Urlaub machen?

Das könnt ihr während der Schulferien bei der Olivenenernte in Italien, bei der Weinlese in Frankreich oder beim Erdbeerpflücken in Spanien. Neben freier Kost und Logis gibt es ein kleines Taschengeld. Die Reisekosten werden vom Austauschprogramm „Europa ohne Grenzen“ übernommen. Infos gibt's online unter: www.jobsundurlaub.de

D

Wenig Geld, aber viel Spaß

Das versprechen wir! Denn wir können zwar nicht viel zahlen, aber für kreative Studierende eine interessante Nebentätigkeit im Bereich Kleinkunst bieten. Wenn du ein Schauspieltalent bist, wenn du jonglieren oder Feuer schlucken kannst, wenn du Gitarre spielst oder sonst ein Instrument, dann solltest du Kontakt mit der Initiative „Klein, aber Kunst“ aufnehmen: E-Mail: kleinaberkunst@altavista.net

E

Hilfs-Stelle

Am Institut Frieden und Demokratie sind ab Herbst

zwei Stellen für studentische Hilfskräfte zu besetzen.

Das Institut erforscht innergesellschaftliche und internationale Konflikte und vermittelt theoretische sowie interdisziplinäre Kompetenzen.

Voraussetzungen: Studium der Politik- oder Sozialwissenschaften, Interesse an interkulturellen Fragen sowie sicherer Umgang mit dem Internet. Bewerbungen bitte an:

Herrn Klierer (Verwaltungsgebäude Raum 2/05).

F

Schnupper-Praktikum

Wie macht man eine Zeitung? Wer sich das fragt, findet sicher eine Antwort bei einem Schnupper-Praktikum „Printmedien“ für Schülerinnen und Schüler ab 16. Ein solches zweiwöchiges Praktikum kann man bei namhaften Zeitungsverlagen in Hamburg absolvieren. Die Stadt Hamburg fördert das Programm, sorgt für kostenlose Unterkunft und finanziert ein Taschengeld.

Infos unter: 0180/234517

G

Jobben in der Fremde

Die Universität unterhält internationale Partnerschaften mit mehreren Hochschulen im europäischen und außereuropäischen Ausland. Ein spezielles Austauschprogramm bietet die Möglichkeit, an einer der Partneruniversitäten in der Hochschulverwaltung ein bezahltes Praktikum zu absolvieren. Voraussetzungen sind u. a. ausreichende Sprachkenntnisse sowie ein abgeschlossenes Grundstudium (vorzugsweise der Sozial-, Kultur- oder Politikwissenschaften). Auskunft erteilt das Amt für Internationales, Raum 256 (*Frau Brauber*).

H

Virtuelles Schulmagazin sucht MitarbeiterIn

Die Tätigkeit beim Webmagazin „Skultur“ besteht darin, Texte und Fotos zu Themen rund um die Schule in Kooperation mit unserem Team für das Internet zu redigieren. Wichtig: Du solltest selbst noch zur Schule gehen sowie Interesse am Umgang mit dem Internet und an Fragen zum Webdesign mitbringen.

Interessiert? Dann schicke eine Kurzbewerbung an: info@skultur.de

(TestDaF Modellsatz 02: http://www.testdaf.de/fileadmin/Redakteur/PDF/TestDaF/Modellsatz_02/Modellsatz_02_Stand_12-07.pdf)

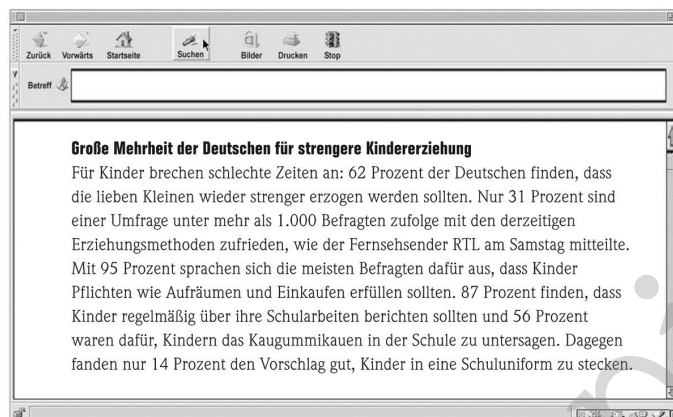
Prüfungsaufgabe B

GOETHE-ZERTIFIKAT B2

Modellsatz

Aufgabe 1 Thema 2 Dauer 65 Minuten

Im Internet lesen Sie folgende Meldung:



Schreiben Sie als Reaktion auf diese Meldung an die Online-Redaktion. Sagen Sie,

mit welchen der erwähnten Erziehungsmaßnahmen Sie persönlich (nicht) einverstanden sind.

welche Vorschläge für die Kindererziehung Sie machen möchten.

ob Kinder früher strenger erzogen wurden.

wer für die Erziehung der Kinder zuständig ist.

Hinweise:

Vergessen Sie bitte nicht Anrede und Gruß.

Die Adresse der Internetredaktion brauchen Sie nicht anzugeben.

Bei der Beurteilung wird u. a. darauf geachtet,

- ob Sie alle vier angegebenen Inhaltspunkte berücksichtigt haben,
- wie korrekt Sie schreiben,
- wie gut Sätze und Abschnitte sprachlich miteinander verknüpft sind.

Schreiben Sie mindestens 180 Wörter.

Seite 18

Goethe-Zertifikat B2 Modellsatz 200410



(Goethe-Zertifikatsprüfung B2: http://www.goethe.de/lrn/pro/gb2/B2_Modellsatz_04.pdf)

Bei der dritten Prüfungsaufgabe handelt es sich um einen sogenannten **C-Test**. Ein **C-Test** ist ein spezifischer **Lückentest**, der normalerweise aus vier bis sechs kurzen Texten mit jeweils 20–25 Lücken pro Text besteht. Beginnend mit dem zweiten Wort des zweiten Satzes fehlt bei jedem zweiten Wort die zweite Hälfte (siehe www.c-test.de). Die Bewertung der Lernenden erfolgt anhand der Zahl der korrekt rekonstruierten Wörter. Der C-Test unterscheidet sich von vielen anderen Lückenformaten dadurch, dass nach einem mechanischen Prinzip Wortteile entfernt werden und nicht etwa ganze Wörter nach z.B. inhaltlichen Kriterien. Der C-Test ist mittlerweile ein häufig eingesetztes und vor allem von Prüfungsspezialisten geschätztes Verfahren. Wenn Sie gern einmal einen C-Test am Computer bearbeiten möchten, empfehlen wir Ihnen, den Beispielttest des *onDaF* unter www.ondaf.de auszuprobieren.

Prüfungsaufgabe C

Ergänzen Sie in dem unten stehenden Text die Lücken sinngemäß wie in dem folgenden Beispiel: Ihr Unter **nicht** beginnt **mor gen**.

TEXT A

Das oft kritisierte Handwerk hat in den vergangenen Jahren offenbar viele Sympathien zurückgewonnen. Das ist das Ergebnis einer Umfrage, die der Marplan-Institut im Auftrag des Zentralverbands für das Deutsche Handwerk durchgeführt hat. Danach haben achtzig Prozent aller Bundesbürger das Handwerk für "sehr wichtig" befunden. Fünfundsiebzig Prozent ist die Wirtschaftszweig "sehr vertraut" und für fünfundsiebzig Prozent "sehr zuverlässig". Der Zentralverband teilte gestern in Bonn mit, eine ähnliche Umfrage hätte fünf Jahre zuvor ein schlechteres Ergebnis gebracht.

(<http://www.wirtschaftsdeutsch.de/ctest/ctestwirt.txt.php3>)

Aufgabe 18



Lösen Sie die Prüfungsaufgaben A–C.

a) Wie sind Sie bei der Lösung vorgegangen?

| Prüfungsaufgabe | So bin ich bei der Lösung der Prüfungsaufgabe vorgegangen: |
|-------------------------|--|
| A. TestDaF | |
| B. Goethe-Zertifikat B2 | |
| C. C-Test | |

b) Was wollten die Testerstellenden Ihrer Meinung nach mit diesen Aufgaben überprüfen?

Prüfungsaufgabe A (TestDaF) soll überprüfen:

Prüfungsaufgabe B (Goethe-Zertifikat B2) soll überprüfen:

Prüfungsaufgabe C (C-Test) soll überprüfen:

Die Frage, welche zugrunde liegenden Teilkompetenzen und Wissensbestände mit den jeweiligen Prüfungsaufgaben erfasst werden können und sollen, ist ganz zentral. Die genaue Beschreibung der Kompetenzen ist nicht nur für standardisierte Tests, sondern ebenso für Prüfungen im unterrichtlichen Kontext ausgesprochen wichtig, und zwar sowohl im Hinblick auf die Erstellung als auch den Einsatz der Prüfungen.

Performanztests

Kompetenzen kann man in der Regel nicht direkt erfassen. Man kann jedoch das Verhalten und Handeln der Lernenden beobachten und daraus schließen, welche Kompetenzen bei ihnen vorhanden sind. Das beobachtbare (sprachliche) Handeln wird häufig als **Performanz** bezeichnet. Vor allem dann, wenn das in der Prüfungssituation geforderte sprachliche Handeln in realen Situationen in ähnlicher Form vorkommt, spricht man von einem **Performanztest** oder auch von aufgabenbasiertem Evaluieren. Typische Beispiele für sprachliche Performanztests sind die Überprüfung der Schreibfähigkeit z.B. anhand des Verfassens eines Aufsatzes oder die Überprüfung der Sprechfähigkeit z.B. anhand eines Rollenspiels. Moderne handlungsorientierte Tests überprüfen die produktiven Fertigkeiten stets durch Performanztests.

Sehen Sie nun ein Beispiel für eine Aufgabe, die die Sprechfähigkeit überprüft.



Sprechen Teil 3

Teilnehmer/in A und B

Teil 3: Gemeinsam etwas planen

Sie wollen mit einem Freund am nächsten Samstag einen Ausflug machen.

Planen Sie, was Sie tun möchten. Hier sind einige Notizen:

- Wohin?
- Wie lange?
- Wie reisen?
- Was dort machen?
- Was mitnehmen?
- Wer kümmert sich um was?
- ...?

Prüfungstraining Deutsch-Test für Zuwanderer, S. 109.

Die produktiven Teilkompetenzen werden in der Regel mithilfe von offenen Aufgabenformaten geprüft. Die Prüfungsteilnehmenden bekommen für das Sprechen oder Schreiben Impulse, müssen die Aufgabe dann aber selbst ausgestalten. Im Gegensatz dazu werden die rezeptiven Teilkompetenzen in der Regel anhand von Aufgabenformaten geprüft, die die Bearbeitung stärker steuern und bei denen häufig Lösungsmöglichkeiten vorgegeben sind. Die Lernenden müssen z.B. die richtige Lösung ankreuzen, Sätze zuordnen oder Kurzantworten geben. Solche Aufgaben unterscheiden sich also deutlich von den wirklichkeitsnahen Aufgaben in Performanztests.

Allerdings sind auch im Fall der rezeptiven Teilkompetenzen performanzbasierte Überprüfungsformen durchaus denkbar. Das Leseverstehen könnte man mithilfe eines Performanztests z.B. folgendermaßen überprüfen: Die Prüfungsteilnehmenden lesen eine Montageanleitung und bauen den entsprechenden Gegenstand zusammen. Wie gut der Text verstanden wurde, könnte dann an der Qualität des Montageobjekts erkannt werden. Diese Prüfungsaufgabe entspricht zwar Handeln außerhalb der Prüfungssituation in authentischer Form. Wer allerdings schon einmal versucht hat, z.B. einen Schrank von einem Möbeldiscounter anhand der Montageanleitung zusammenzubauen, der weiß, dass das Resultat des Montageversuchs nicht nur von der Leseleistung abhängt, sondern auch von der handwerklichen Erfahrung und dem räumlichen Vorstellungsvermögen. Es werden hier also noch weitere Fähigkeiten gemessen, die aber im Hinblick auf die Lesekompetenz und die Fähigkeit zum Lesen außerhalb der Prüfungssituation nicht relevant sind.



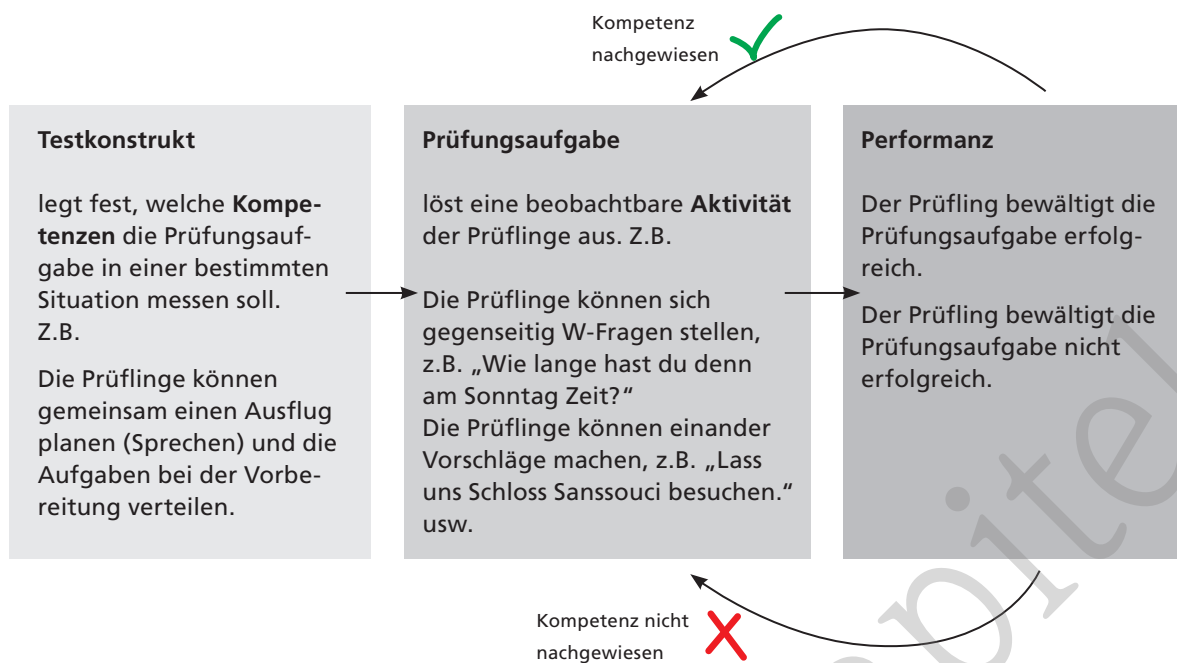
Performanztests, bei denen es in erster Linie um das gezeigte Verhalten geht, sind vor allem in bestimmten beruflichen Handlungsfeldern wie etwa Deutsch für Kellner im Hotelgewerbe sinnvoll. Charakteristisch und problematisch in Prüfungen für solche Berufsgruppen ist, dass häufig nur ein begrenztes Repertoire von sprachlichen Handlungen überprüft wird (z.B. begrüßen, Bestellungen aufnehmen, nach der Zufriedenheit fragen). Bei der Erstellung von Prüfungen sollte man darauf achten, solche Handlungssituationen möglichst genau und umfassend abzubilden. In diesem Fall zeigt dann das Prüfungsergebnis unmittelbar an, wie gut jemand in dem entsprechenden Handlungsfeld agieren kann. Ein Rückgriff auf zugrunde liegende Kompetenzen ist nicht nötig.

Testkonstrukt

Wir haben oben bereits erwähnt, dass man Kompetenzen nur indirekt erfassen kann, und zwar über die Beobachtung von Verhalten. Die Kompetenzen, die konkretem Verhalten in Handlungssituationen zugrunde liegen, bezeichnet man häufig auch als **Konstrukte**. Entsprechend werden Kompetenzen, die anhand der beobachtbaren Leistung in einer Prüfung gemessen werden sollen, unter dem Begriff **Testkonstrukt** zusammengefasst.

Prüfungsaufgaben dienen dazu, Kompetenzen sichtbar zu machen. Es handelt sich damit um sogenannte **Operationalisierungen** der Konstrukte. Operationalisierung heißt, dass die Prüflinge – gesteuert über Aufgaben – bestimmte beobachtbare Aktivitäten (Operationen) ausführen. Anhand der gezeigten Aktivitäten kann dann auf das Testkonstrukt (z.B. auf die Fähigkeit zum Schreiben alltagssprachlicher Texte) geschlossen werden. Bei einer kompetenzorientierten Prüfung wird somit deutlich zwischen dem in der Prüfung gezeigten Verhalten (z.B. Schreiben eines Textes) und den zu messenden Kompetenzen unterschieden. Daneben wird der Begriff des Testkonstrukts auch noch in einem weiten Sinne verwendet und bezieht sich auf alles, was ein Test erfassen soll.

Sehen Sie die dargestellten Zusammenhänge noch einmal in einer Grafik zusammengefasst:



Was wollen wir überprüfen?

In der Prüfungsaufgabe B, die Sie oben bearbeitet haben, ging es darum, zu einer vorgegebenen Situation eine angemessene Mail zu schreiben, indem man die passenden sprachlichen Mittel verwendet. Eine erfolgreiche Bearbeitung der Aufgabe werden wir als Zeichen (Indikator) dafür, dass die Person über die Kompetenz zum Schreiben (einer bestimmten Art) von argumentativen Texten verfügt. Deutlich komplizierter wird es, wenn es nicht um das Schreiben oder Sprechen, sondern um das Hör- oder Leseverstehen geht. Denn wie ein Text oder eine mündliche Äußerung verstanden wird, ist meistens nicht direkt am Verhalten der Lesenden oder Hörenden ablesbar, sondern muss z.B. durch Fragen zum Inhalt erschlossen werden. In Prüfungsaufgabe A waren zehn Personen mit bestimmten Bedürfnissen aufgeführt, denen jeweils ein passendes Job- bzw. Praktikumsangebot zugeordnet werden musste. Bei der Bewertung wird von der Anzahl der richtigen Zuordnungen auf das Niveau der Lesekompetenz des jeweiligen Testteilnehmenden geschlossen. Dabei setzt eine zuverlässige Einschätzung voraus, dass man genau weiß, welche Prozesse, Kenntnisse, Strategien usw. zur Lösung der Aufgabe führen (siehe auch die Begründung im Lösungsschlüssel zu Aufgabe 18).

Was überprüfen wir wirklich?

Wenn die Prüfungsaufgaben das intendierte Testkonstrukt nur zum Teil repräsentieren, dann spricht man auch von **Konstruktunterrepräsentation**. Ein extremes Beispiel ist das Testen kommunikativer Kompetenz allein über Grammatikaufgaben. Im Fall einer Konstruktunterrepräsentation ist die Messung zu ‚eng‘ in Bezug auf das angezielte Testkonstrukt. Wenn dagegen ein Teil der Unterschiede zwischen den Prüfungsergebnissen der Lernenden nichts mit dem intendierten Testkonstrukt (z.B. kommunikative Kompetenz) zu tun hat und somit auf **konstruktirrelevante Faktoren** zurückgeht, dann ist die Messung zu ‚breit‘. In den Prüfungsergebnissen finden sich dann konstruktirrelevante Unterschiede zwischen den Prüfungsteilnehmenden. Oder anders ausgedrückt: Die Ergebnisse weisen damit eine sogenannte **konstruktirrelevante Varianz** auf.

Ein Fall von konstruktirrelevanter Varianz liegt ebenfalls vor, wenn sich Prüfende in einem Prüfungsgespräch so verhalten, dass sie die Prüflinge verängstigen, oder auch dann, wenn Prüfungsteilnehmende unter Prüfungsangst leiden. Lernende berichten z.B. immer wieder, dass gerade Hörverstehensaufgaben für sie mit Ängsten und Stress verbunden sind. Manchmal haben sie Angst, dass sie etwas akustisch nicht mitbekommen. Manchmal glauben sie, dass sie alles verstehen müssen, dies aber gar nicht schaffen können. Manchmal fühlen sie sich auch überfordert, weil sie nicht hinreichend an authentische Hörsituationen gewöhnt sind. Wer mit solchen Ängsten zu kämpfen hat, kann in einer Evaluationssituation nicht zeigen, was er eigentlich leisten kann.

Wenn Sie wissen, dass bei Ihren Lernenden solche Ängste bestehen, dann können Sie versuchen, den Stress durch eine geeignete Vorbereitung auf die Prüfung abzubauen. Sie können die Lernenden darauf hinweisen, dass sie in einem Hörverstehenstest in der Regel nicht jedes Detail verstehen müssen. Sie können sie gezielt auf das Hören von authentischen Situationen vorbereiten und dabei ähnliche Fragen beantworten bzw. Aufgaben lösen lassen, wie sie in Hörverstehenstests vorkommen. Natürlich sollten die Lernenden dabei zusätzlich ihr eigenes Tun reflektieren.

Außerdem sollten Sie darauf achten, dass während der Durchführung eines Hörverstehenstests möglichst keine Störgeräusche wie z.B. Lärm vor dem Unterrichtsraum auftreten oder dass einige Prüflinge sehr nah, andere Prüflinge hingegen sehr weit von der Hörquelle entfernt sitzen. Auch hier spiegeln die Leistungsunterschiede in den Prüfungsergebnissen dann nicht nur das zu messende Konstrukt, nämlich Hörverstehenskompetenz wider, sondern auch den konstruktirrelevanten Störfaktor Nebengeräusche oder den Abstand zur Hörquelle.

Nehmen wir eine weitere Situation an: Eine Lehrerin gibt Hinweise zum Inhalt einer Schreibprüfung. Insbesondere weniger kompetente Prüflinge bereiten sich dann manchmal so vor, dass sie einen Text mithilfe anderer vorformulieren und diesen Text auswendig lernen. Passt der vorbereitete Text zur Aufgabenstellung in der Prüfung, erhalten die weniger kompetenten Prüflinge vermutlich sogar eine hohe Punktzahl. Die Aufgabe testet dann allerdings nicht allein die Schreibkompetenz, sondern in erster Linie die konstruktirrelevante Fähigkeit, einen auswendig gelernten Text zu reproduzieren.

Eine andere wichtige Quelle konstruktirrelevanter Varianz ist das Verhalten der Bewertenden im Fall offener Aufgabenformate wie z.B. bei einem mündlichen Interview oder beim Schreiben eines Briefes. Wenn einige Bewertende z.B. systematisch strenger oder milder als andere urteilen, dann führt dies auch zu konstruktirrelevanten Unterschieden in den Prüfungsergebnissen.

Wie bereits in den vorangehenden Beispielen angedeutet, handelt es sich bei der konstruktirrelevanten Varianz um einen systematischen Messfehler, den wir bei der Entwicklung, dem Einsatz und der Bewertung von Testaufgaben aus Gründen der Fairness möglichst weitgehend ausschließen bzw. kontrollieren sollten. Daneben gibt es natürlich stets noch unsystematische Messfehler aufgrund nicht oder nur schwer kontrollierbarer Zufallseinflüsse. Dazu zählen z.B. schwankende Aufmerksamkeit und Motivation der Testteilnehmerinnen und Testteilnehmer oder plötzliche Erkrankungen.

Wichtig ist, dass wir uns immer darüber Gedanken machen, was wir überhaupt überprüfen wollen und können – vor allem dann, wenn es sich um sogenannte „schwer messbare“ Kompetenzen handelt wie im Fall bestimmter Aspekte interkultureller (kommunikativer) Kompetenz. Während man soziokulturelles Wissen über Höflichkeitskonventionen relativ leicht überprüfen kann, sind z.B. Einstellungen gegenüber der Zielkultur schwieriger messbar (siehe Teilkapitel 5.6).

Abschließend möchten wir Sie nun bitten, eine Prüfungsaufgabe im Hinblick auf ein bestimmtes Testkonstrukt zu bewerten.

Aufgabe 19

Wird das Testkonstrukt „sich schriftlich für einen Praktikumsplatz bewerben“ durch die folgende Prüfungsaufgabe angemessen erfasst? Begründen Sie Ihre Meinung.

Eine Kollegin möchte in ihrem Unterricht überprüfen, ob ihre Lernenden sich in Deutschland schriftlich für einen Praktikumsplatz bewerben können. Sie berichtet Ihnen von ihrer Prüfungsvorbereitung: „Ich habe in einem deutschen Bewerbungsbrief, den ich im Internet gefunden habe, eine Reihe von schwierigen Wörtern entfernt. Die Schülerinnen und Schüler sollten die fehlenden Wörter ergänzen.“

Ich bin der Meinung, dass _____