

MODUL 3

ITEMTYPEN

EINIGE ASPEKTE DER ITEMERSTELLUNG

1. TEXTE

- a) Authentizität
 - i) situationsbezogene Authentizität
 - ii) interaktionsbezogene Authentizität
- b) Schwierigkeit der Texte
 - i) Sprachliche Struktur des Textes
 - ii) Kontextuelle Einbettung des Textes
 - iii) Inhalt des Textes
 - iv) Art der Interaktion
- c) Schwierigkeit in Hörtexten
 - i) Interaktion der Sprecher
 - ii) Zeitliche Bezüge und Kontext
 - iii) Sprache

2. ITEMTYPEN

- a) Einige allgemeine Regeln
- b) Multiple-Choice- und andere Auswahlitems
 - i) Einzelne und textbezogene Multiple-Choice-Items
 - ii) Richtig/Falsch-Items
 - iii) Lückentest mit Multiple-Choice-Items
 - iv) Lückentest mit Schüttelkasten
 - v) Textabschnitte zuordnen
 - vi) Zuordnungsaufgaben
 - vii) Mehrfachzuordnung
 - viii) Fehlerkorrektur: Ein Wort zu viel
- c) Items mit subjektivem Beantwortungsspielraum
 - i) Kurzantwortaufgaben
 - ii) Satzergänzung
 - iii) Lückentest mit offenen Lücken
 - iv) Umformungen
 - v) Wortbildung
 - vi) Lückentest zur Wortbildung

- vii) Aus Stichwörtern einen Text erstellen
- viii) Korrektur lesen/Fehler korrigieren
- ix) Informationstransfer

3. NICHT-ITEMBASIERTE AUFGABENTYPEN

- a) Schreiben: Aufgaben zum erweiterten Schreiben
 - i) Aufgaben zum Schreiben mit detailliertem Input
 - ii) Aufgaben zum Schreiben mit geringem Input
 - iii) Beispiele für Aufgaben zum Schreiben auf unterschiedlichen Stufen
- b) Sprechen: Aufgaben zum Sprechen
 - i) Vortrag
 - ii) Vorgabe von Bildern
 - iii) Schriftliche Vorgaben
 - iv) Aufgaben mit einer Informationslücke

4. ANWEISUNGEN

5. LÖSUNGSSCHLÜSSEL, BEWERTUNGSVORGABEN, BEWERTUNGSANLEITUNG

6. ÜBUNGEN

ITEMTYPEN

EINIGE ASPEKTE DER ITEMERSTELLUNG

Ein Item zu schreiben heißt u. a., einen Stimulus zu geben, der eine geforderte Antwort hervorruft. Die Form des Stimulus hängt dabei vom gewählten Itemtyp ab. Die Auswahl an Itemtypen kann durch die Art des Inputs (d. h. der Vorgaben) eingeschränkt werden, wird aber primär durch die erwartete Antwort bestimmt. Oder anders ausgedrückt: Der gewählte Itemtyp sollte bei der Bearbeitung die sprachlichen Fähigkeiten hervorrufen, die getestet werden sollen.

Bevor ein Testautor Items schreiben kann, muss er in den Techniken der Itemerstellung geschult werden und deren Umsetzung verstehen, um ein angemessenes Verhältnis zwischen Stimulus und Reaktion herstellen zu können. Erreicht er dies nicht, entstehen Probleme. So ist es z. B. möglich, Items zu einem Text zu schreiben, die man richtig beantworten kann, ohne den Text verstanden zu haben. Entweder versteht man das den Items zugrunde liegende Muster oder man findet Hinweise im Text; manchmal können auch Formulierungen in den Items direkt einer Textpassage zugeordnet werden. Dadurch kann das Item zwar „korrekt“ beantwortet werden, jedoch beweist dies nicht notwendigerweise, dass etwas Sinnvolles getestet wurde. Auch kann sich ein Stimulus sehr gut für ein bestimmtes Item eignen, ohne dass dieses Item dem Prüfungsziel entspricht.

Man sollte nicht davon ausgehen, dass sich die Schwierigkeit eines Items leicht aus dem Verhältnis von Text und Antwortmöglichkeiten bestimmen lässt. Stimulus und Antwort haben ihre eigenen sprachlichen Merkmale, aber die Aufgabe als Verbindung zwischen beiden kann zusätzlich zu ihren sprachlichen Anforderungen noch kognitive Komplexität mit hineinbringen. Das Weltwissen spielt ebenfalls eine Rolle. Wie sich Modelle der Sprachkompetenz in der Konzeptionsphase eines Tests niederschlagen, wurde in Modul 1 erläutert. Man muss alle Merkmale der kommunikativen Kompetenz berücksichtigen, die für das Prüfungsziel relevant sind.

Bevor der Testautor beginnt, Items zu schreiben, muss er sich über Folgendes im Klaren sein:

- warum gerade dieser Itemtyp für den Test ausgewählt wurde
- auf welche Kompetenzbereiche der Kandidaten die Items zielen sollen

Terminologie

Ein Test besteht aus einer Reihe von **Aufgaben**. Die geschlosseneren Aufgabentypen (die beim Lese- und Hörverstehen, bei der Überprüfung von Grammatik und Wortschatz und bei Schreibaufgaben auf Satzebene verwendet werden), bestehen aus einer Kombination von **Anweisung**, **Input** (wie z. B. einem **Text** als Stimulus) und der **Antwort** des Kandidaten, die durch die unterschiedlichen **Items** (entweder produktiv oder rezeptiv) hervorgerufen wird. Die Items werden mithilfe eines **Lösungsschlüssels** oder einer **Bewertungsvorgabe** bewertet. Man muss unterscheiden zwischen itembasierten Aufgaben und den Aufgaben, die in Tests zum freien Schreiben bzw. zum Sprechen verwendet werden und die sich aus **Anweisung**, **Input** und **Antwort** zusammensetzen; hier wird die Antwort mithilfe einer **Bewertungsanleitung** oder mithilfe **definierter Kriterien** bewertet, im Gegensatz zu einer Bewertung mithilfe eines Lösungsschlüssels bzw. Bewertungsvorgabe.

Im Folgenden wird näher auf einige Aspekte eingegangen, die die Texte, die Itemtypen, die nicht-itembasierten Aufgabentypen, die Anweisungen, die Bewertungsvorgaben und den Lösungsschlüssel betreffen.

1. Texte

Wenn man Texte für eine Aufgabe sucht, ist es sehr wichtig, sie im Hinblick auf die entsprechende Testpopulation auszuwählen. Der sprachliche Schwierigkeitsgrad muss angemessen und das Thema für die angenommene Altersgruppe und die Voraussetzungen der Kandidaten geeignet sein. Themen, die möglicherweise als beleidigend empfunden werden oder gar Bestürzung auslösen könnten, sollten vermieden werden. Diese Aspekte werden in Modul 2 in den Ausführungen zum Prozess der Testerstellung behandelt.

Zwei Aspekte, die zur Textauswahl gehören, sollen hier diskutiert werden: die Authentizität und die Schwierigkeitsfaktoren eines Textes.

a) Authentizität

Bei der Auswahl von Texten für den Unterricht und für Tests wird die Frage der Authentizität häufig diskutiert. Kommt es den Bedürfnissen der Kandidaten eher entgegen, wenn man in einer Prüfung Texte im Original aus Zeitungen oder Zeitschriften aufnimmt (z. B. beim Leseverstehen), oder sollte man Texte anbieten, die vom Testautor für die Prüfung geschrieben wurden?

Zeitungen und Zeitschriften erscheinen angemessener, weil sie einem realitätsnahen Sprachgebrauch entsprechen; sie wurden für Muttersprachler geschrieben und nicht zu Prüfungszwecken. Es kann ein Ziel des Lernalers sein, mit Texten so umgehen zu können, wie Muttersprachler dies tun. Wenn dies so ist, sollte er auch anhand authentischer Textvorlagen geprüft werden. Ein Text, der nur geschrieben wurde, um ein bestimmtes sprachliches Phänomen zu testen, hat möglicherweise keine Ähnlichkeit mit der Sprache, die Muttersprachler außerhalb einer Testsituation verwenden.

Gibt es eine breitere Definition von Authentizität?

Manche vertreten die Meinung, Authentizität sei eine Folge der Interaktion zwischen Leser und Text und nicht nur eine Qualität des Textes an sich. Selbst wenn man die vielen Zeitungen und Zeitschriften nur flüchtig liest, stellt man eine große Bandbreite an Sprachgebrauch fest. Dies zeigt, dass nicht alle Texte für alle Leser authentisch sind. Wer der Leser ist, welches Leseziel er hat, welche Schreibabsicht der Autor verfolgt sowie das Ausmaß an kultureller Übereinstimmung zwischen Leser und Text sind Faktoren, die bestimmen, ob der Text für den Leser authentisch ist. Gibt es wenig Übereinstimmung zwischen der im Text vorausgesetzten Sachkenntnis und kulturellem Wissen und der Sachkenntnis und dem kulturellen Wissen des Lesers, wird das Lesen zu einer wenig authentischen Aktivität (denken Sie nur an einen älteren Opernliebhaber, der versucht, ein Musikmagazin für Jugendliche zu lesen).

Muttersprachler wählen bis zu einem gewissen Grad die Texte aus, die für sie authentisch sind, und meiden die anderen. Wie kann man nun angemessene Texte aus authentischen Quellen wie Zeitungen und Zeitschriften auswählen und sicher sein, dass sie für Lerner authentisch sind, die vielleicht nie im Land der Zielsprache waren und bei denen nicht das soziale und kulturelle Wissen eines Muttersprachlers vorausgesetzt werden kann, für den die Texte geschrieben wurden?

Es reicht mit Sicherheit nicht aus, Artikel oder Anzeigen aus Zeitungen auszuschneiden und anzunehmen, dass sie sich für den Unterricht und für Tests eignen, nur weil sie aus einer authentischen Quelle stammen. Wenn dem Lerner das Wissen fehlt, das bei den Lesern vorausgesetzt wird, für die der Text geschrieben wurde, dann wird er gezwungen, Wort für Wort zu lesen, um den Text zu verstehen, wodurch das Lesen alles andere als authentisch wird.

Es muss jedoch eine Verbindung zwischen der Testaufgabe einerseits und dem Gebrauch der Sprache in Realsituationen und außerhalb von Tests andererseits geben; in diesen Situationen möchte sich der Lerner sprachlich behaupten können, Situationen, die wir im Test abbilden wollen. Außerdem stellt sich die Frage der Augenscheinvalidität, d. h., inwieweit der Test für die Verbraucher so aussieht, als prüfe er für sie relevante kommunikative Aktivitäten.

In den späten 70er Jahren erfuhr die Definition von Authentizität eine wichtige Erweiterung. Widdowson (1978) und später Bachman (1990) sehen zwei Ebenen der Authentizität: die situationsbezogene und die interaktionsbezogene Authentizität.

i) situationsbezogene Authentizität

Die situationsbezogene Authentizität kann definiert werden als das Ausmaß, in dem die Merkmale des Aufgabenformats die Merkmale der Realsituation widerspiegeln, in der die Sprache verwendet wird.

Bei der Entwicklung einer Aufgabe mit situationsbezogener Authentizität muss man zunächst die wichtigen Merkmale identifizieren, die die Aufgabe in der Sprachverwendungssituation bestimmen; dabei geht man von Rahmenvorgaben zu den Merkmalen des Aufgabenformats aus. Dann können Testaufgaben entwickelt werden, die diese wichtigen Merkmale aufweisen.

ii) interaktionsbezogene Authentizität

Bachman definiert die interaktionsbezogene Authentizität als Interaktion zwischen Testaufgabe und Testteilnehmer:

„If our objective is to design language test tasks that correspond to non-test language use, then test tasks must incorporate the goal-directed, purposive nature of language as communication, which means that they must involve the test taker in functions other than simply demonstrating his knowledge of the language.“
(Bachman, 1990)

Diese Sicht von Authentizität bedingt, dass für Testentwickler und Testautoren Folgendes wichtig ist:

- Texte, Situationen, Kontexte und Aufgaben anzubieten, die Realsituationen widerspiegeln, ohne sie exakt abzubilden
- Situationen und Aufgaben zu verwenden, die den zukünftigen Testteilnehmern voraussichtlich bekannt und relevant für ihr Sprachniveau sind
- das kommunikative Ziel der Aufgabe und den intendierten Adressaten durch einen Kontext zu verdeutlichen
- die Kriterien für eine erfolgreiche Aufgabenerfüllung zu verdeutlichen

b) Schwierigkeit der Texte

Testentwickler und Testautoren müssen sich ständig mit der Frage beschäftigen, was die Schwierigkeit eines Textes ausmacht. Sowohl bei schriftlichen als auch bei gesprochenen Texten müssen sie bedenken, dass eine Reihe von Faktoren bestimmen, wie schwer ein Text für den Leser bzw. Hörer zu verstehen ist (innerhalb und außerhalb einer Testsituation).

i) Sprachliche Struktur des Textes

Ein Text, der sich aus kurzen einfachen Sätzen zusammensetzt und kein Passiv verwendet, wird als leichter angesehen als ein Text, der aus langen verschachtelten Sätzen mit häufigen Passivkonstruktionen besteht. Dem Kandidaten relativ vertraute Strukturen und Vokabeln sind leichter zu verstehen als weniger vertraute.

ii) Kontextuelle Einbettung des Textes

Es erleichtert das Verarbeiten schriftlicher und gesprochener Texte, wenn sich der Text direkt an den Leser bzw. Hörer wendet, statt ihn Mäuschen spielen und eine Interaktion zwischen anderen Personen beobachten zu lassen. Auch die visuelle Unterstützung durch ein Video (in einem Test zum Hörverstehen) oder durch Bilder und Diagramme erleichtern das Textverstehen; außerdem ist es leichter, einen Text ohne zeitliche Begrenzung zu bearbeiten. Die Aufgabe kann auch dadurch erleichtert werden, dass der Kontext als zwingenden Grund der Bearbeitung eine „Informationslücke“ aufweist, die der Kandidat füllen muss.

iii) Inhalt des Textes

Handelt es sich um eine Erzählung, dann sind wenige klar unterscheidbare Personen am leichtesten zu verstehen. So wird z. B. eine Geschichte über zwei Männer und zwei Frauen, die unterschiedlich alt sind, die unterschiedliche Namen haben und die sich in ihrer Persönlichkeit klar voneinander unterscheiden, als leichter empfunden als eine Geschichte mit einer ganzen Reihe nur diffus gezeichneter Personen.

Je geradliniger die Handlung ist, also in klarer Abfolge der Ereignisse und ohne Rückblenden, desto leichter ist sie zu verstehen. Wenn es eine klare Verbindung zwischen den Ereignissen gibt, wie Ursache und Wirkung, wird der Text als leichter empfunden als ein Text, in dem die Ereignisse scheinbar nicht miteinander verbunden sind. Ein Hörer oder Leser, der eine neue Geschichte mühelos in sein Hintergrundwissen integrieren kann, findet den Text zudem leichter als jemand, dem diese Wissensstrukturen fehlen.

iv) Art der Interaktion

Ein sehr formeller Text, der eine distanzierte Beziehung ausdrückt, oder ein sehr informeller Text, der eine eher persönliche Beziehung ausdrückt, können beide größere Schwierigkeiten verursachen als Texte, die in einem relativ neutralen oder gemäßigt informellen Stil verfasst sind.

c) Die Schwierigkeit in Hörtexten

Bei schriftlichen Texten besteht die Möglichkeit, die Verbindungen zwischen Textteilen zu sehen, im Text vor- und zurückzugehen oder den Text als Ganzes zu betrachten; bei gehörten Texten ist dies jedoch nicht möglich. Wer Items zum Hörverstehen erstellt, sollte beim Schreiben der Items oder bei der Auswahl der Hörtexte neben dem Aspekt der sprachlichen Schwierigkeit (Komplexität der Strukturen und des Wortschatzes) auch die unten genannten Faktoren bedenken. Sie alle beeinflussen den Grad der Textverarbeitung, der über das reine Verstehen hinaus erforderlich ist, und haben somit auch eine Auswirkung auf die Schwierigkeit des Textes.

i) Interaktion der Sprecher

- Der Monolog ist eine Diskursform, der man am leichtesten folgen kann, besonders wenn der Sprecher sich direkt an den Hörer zu wenden scheint.
- Ein Gespräch mit zwei unterschiedlichen Stimmen (ein Mann und eine Frau oder ein Erwachsener und ein Kind) ist die zweitleichteste Form.
- Ein Gespräch zwischen zwei Personen gleichen Geschlechts und Alters, oder ein Gespräch zwischen mehr als zwei Sprechern, ist schwerer zu verstehen.
- Ein Gespräch zwischen Sprechern mit klar definierten Rollen, wie z. B. Elternteil und Kind, ist leichter zu verstehen.
- Ein Gespräch zwischen Sprechern mit gleichen Rollen, z. B. zwischen Kollegen gleichen Geschlechts und ähnlichem Status, die über eine Situation am Arbeitsplatz sprechen, ist im Allgemeinen schwerer zu verstehen.

ii) Zeitliche Bezüge und Kontext

- Ein Text mit verschiedenen Schauplätzen, wechselnden Zeitbezügen und vielen unterschiedlichen Ereignissen ist schwerer zu verstehen als ein Text mit nur wenigen Ereignissen, die zudem alle in der gleichen Umgebung und zur gleichen Zeit stattfinden.
- Ein Text, der von Anfang an in einen klaren Kontext eingebettet ist, ist leichter zu verstehen.

iii) Sprache

- Ein kurzer Text mit hoher Informationsdichte, auf den sich eine verhältnismäßig große Anzahl von Items bezieht, ist für Kandidaten schwer zu verarbeiten, selbst wenn das sprachliche Niveau des Textes angemessen erscheint. Redundanzen in Form von Erklärungen, Umformulierungen und Wiederholungen können dazu beitragen, die Schwierigkeit eines Textes zu verringern.
- Mit hohem Tempo gesprochene informelle Sprache, Umgangssprache und die Verwendung zusammengezogener Formen, das scheinbare Fehlen einer klaren Textstruktur und häufiger Sprecherwechsel stellen höhere Anforderungen an das Hörverstehen als eine stärker formelle Sprache mit üblicherweise langsamerem Sprechtempo, weniger häufigen Sprecherwechseln und mehr Gemeinsamkeiten mit der Schriftsprache.
- Ein Sprecher mit einem natürlich langsamen Sprechtempo und einer ausdrucksvollen Stimme ist leichter zu verstehen als ein Sprecher, der schnell oder monoton spricht. Es hilft auch, wenn das Sprechtempo des Sprechers gleich bleibt.

2. Itemtypen

Welcher Itemtyp ist besonders angemessen, um eine bestimmte Fertigkeit in einem bestimmten Test zu überprüfen? Dies ist eine wichtige Frage, die normalerweise in der Konzeptionsphase beantwortet wird.

Die große Anzahl unterschiedlicher Itemtypen, die in Sprachtests verwendet werden, lässt sich auf verschiedene Weise kategorisieren:

- Einige werden als objektive Items bezeichnet, weil kein menschliches Urteil erforderlich ist, um sie zu bewerten, während andere eine (kurze) schriftliche Antwort verlangen und subjektiv bewertet werden.
- Einige überprüfen die rezeptiven Fertigkeiten, andere die produktiven.
- Einige beziehen sich auf einen Text, während andere (dekontextualisierte) Einzelaufgaben darstellen.

Welches ist das wichtigste Kriterium, um den Wert eines Itemtyps zu beurteilen?

Obgleich einige Itemtypen häufiger verwendet werden als andere, kann man nicht sagen, dass sie deshalb die besten sind. Das wichtigste Kriterium, um den Wert eines Itemtyps zu beurteilen, ist seine jeweilige Angemessenheit als Aufgabe in einem Sprachtest, der für eine bestimmte Situation gedacht ist und der ein definiertes Ziel hat. Der Itemtyp, der das gewünschte sprachliche Verhalten am direktesten misst, ist im Allgemeinen auch der am besten geeignete.

a) Einige allgemeine Regeln

Es gibt einige allgemeine Regeln, die man befolgen sollte, wenn man Items schreibt, unabhängig davon, um welchen Itemtyp es sich handelt:

- Die Items sollten sich immer auf wichtige Informationen beziehen.
- Die üblichen grammatischen Regeln sollten eingehalten werden.
- Wenn ein neuer Itemtyp verwendet wird, sollte ein Beispiel vorgegeben werden (außer das Lösungsverfahren ist so klar, dass dies nicht nötig ist).
- Wenn Items sich auf einen Text beziehen, muss die korrekte Antwort es erforderlich machen, den Text zu lesen und zu verstehen – es sollte nicht möglich sein, das Item allein durch Weltwissen oder Hintergrundwissen korrekt zu lösen.
- Items, die sich auf einen Text beziehen, können vor oder nach dem Text stehen. Diejenigen, die vor dem Text stehen, sollten das Verstehen der Gesamtaussage des Textes prüfen, während die nachstehenden ein detailliertes Lesen erfordern sollten oder überprüfen, ob Schlussfolgerungen gezogen werden können.

Die folgende Klassifizierung unterteilt Items in zwei weit gefasste Kategorien:

Auswahlitems

Diese Items verlangen vom Kandidaten, dass er unter mehreren Optionen die richtige Antwort auswählt, wie z. B. bei Multiple-Choice-Items mit drei oder vier Optionen, bei Richtig/Falsch oder bei verschiedenen Zuordnungsaufgaben.

Items mit subjektivem Beantwortungsspielraum

(engl. *Candidate-supplied response item types*)

Hierzu gehören Items, bei denen die Kandidaten die Antwort selbst formulieren, z. B. bei einer Kurzantwort oder einem Lückentest mit offenen Lücken.

Im Allgemeinen werden Tests, die aus Auswahlitems bestehen, als objektiver hinsichtlich der Bewertung angesehen als solche, in denen der Kandidat die Antworten selbst produzieren muss.

Es ist wichtig, noch einmal zu betonen, dass kein Itemtyp für sich genommen nützlicher als ein anderer ist. Die Auswahl des angemessenen Itemtyps hängt von den spezifischen Zielen und den Schwerpunkten des Testentwicklers ab. In den folgenden Beschreibungen der Itemtypen und den Anmerkungen, die ihnen folgen, werden Hinweise zu den Fertigkeiten gegeben, die üblicherweise mit einem bestimmten Itemtyp assoziiert werden.

Das Testen von Sprechen und Schreiben

Beim Testen der Fertigkeiten Sprechen und Schreiben können bestimmte Teilfertigkeiten, wie z. B. Grammatik, Lexik, Orthografie, Aussprache usw., gemessen werden. Items, die Teilfertigkeiten des Schreibens überprüfen, finden sich in Tests oder Subtests mit Bezeichnungen wie „Grammatik und Sprachgebrauch“ oder „Sprachbausteine“. Wenn die Fertigkeiten Sprechen oder Schreiben auf dieser Ebene von isolierten Elementen getestet werden, geschieht dies meist mithilfe itembasierter Subtests.

Die Sprech- und Schreibfertigkeiten, die bei einer Interaktion ins Spiel kommen, erfordern das Strukturieren von Gedanken und Argumenten sowie den Aufbau zusammenhängender Äußerungen; diese Fertigkeiten können im Allgemeinen nicht mit itembasierten Tests geprüft werden, sondern erfordern andere Aufgaben.

Die folgenden Abschnitte enthalten Überlegungen zu einer Reihe von Item- und Aufgabentypen. Sie sind nicht umfassend, möchten aber die gebräuchlichsten Item- und Aufgabentypen vorstellen, besonders diejenigen, die von den ALTE-Mitgliedern in ihren Fremdsprachenprüfungen eingesetzt werden.

b) Multiple-Choice- und andere Auswahlitems

Wenn in einem Test Auswahlitems verwendet werden, wird der Testentwickler vermutlich die folgenden Eigenschaften für vorteilhaft halten. Auswahlitems sind:

- weltweit fast allen Kandidaten bekannt
- unabhängig von der Schreibfertigkeit
- schnell und leicht zu bewerten und geeignet für die Korrektur mit einer Lösungsschablone oder einem Scanner
- objektiv zu bewerten
- zeitökonomisch, sodass der Kandidat innerhalb kurzer Zeit eine Menge Items lösen kann, die eine Reihe von Prüfungszielen erfassen; dies trägt zur Reliabilität des Tests bei

Andererseits muss darauf hingewiesen werden, dass Auswahlitems auch **kritisiert** werden, und zwar aus folgenden Gründen:

- Sie überprüfen das Erkennen und nicht die Produktion.

- Sie sind eingeschränkt in dem, was sie überprüfen können.
- Sie erlauben dem Kandidaten nicht, eine große Bandbreite von Fertigkeiten zu demonstrieren.
- Sie sind häufig abhängig von der Lesefertigkeit.
- Sie sind rateanfällig; selbst bei einem Item mit drei Distraktoren gibt es eine Chance von 1 zu 4, durch Raten das Item richtig zu lösen; bei weniger Distraktoren erhöht sich diese Chance entsprechend.
- Gute Items zu erstellen, ist schwer und zeitraubend.
- Sie haben möglicherweise einen schlechten Einfluss auf den Unterricht, wenn sich dieser zu sehr darauf konzentriert, diesen Itemtyp zu üben.

Die Entscheidung, ob ein bestimmter Itemtyp eingesetzt wird oder nicht, hängt davon ab, was getestet werden soll und warum. Kein Itemtyp ist absolut gesehen richtig oder falsch. Unter dem Begriff Multiple-Choice- oder Auswahlitem lassen sich eine ganze Reihe von Techniken subsumieren. Ihnen allen ist gemein, dass der Kandidat zwischen vorgegebenen Optionen wählen muss; er hat kein einziges Wort selbst zu produzieren. Im Folgenden werden die gebräuchlichsten Auswahlitems vorgestellt.

i) Einzelne und textbezogene Multiple-Choice-Items

Hier ein Beispiel für ein einzelnes Multiple-Choice-Item:

The singer ended the concert her most popular song.

A by B with C in D as

Der Satz mit einer Lücke ist der *Stamm*, der im Beispiel oben gefolgt wird von vier *Optionen*. B ist die *Lösung* (engl. *key*) und A, C und D sind die *Distraktoren*, die von den schwächeren Kandidaten gewählt werden könnten. Man kann drei, vier oder fünf Optionen vorgeben.

Hier ein Beispiel für ein textbezogenes Item:

Then he saw a violin in a shop. It was of such high quality that even top professional players are rarely able to afford one like it. 'I'd never felt money was important until then,' Colin explained. 'Even with the money I'd won, I wasn't sure I could afford to buy the violin, so I started to leave the shop. Then I thought I'd just try it, and I fell in love with the beautiful sound it made. I knew it was perfect both for live concerts and for recordings.'

When Colin first found the violin, what did he think?

- A He might not have enough money to buy it.
- B He should not spend all of his money on it.
- C He was not a good enough player to own it.
- D He could not leave the shop without it.

Textbezogene Items werden häufig als Frage formuliert, gefolgt von drei, vier oder fünf Optionen, unter ihnen auch die Lösung.

Multiple-Choice-Items werden häufig verwendet, um das Lese- und Hörverstehen zu prüfen.

Hier ein Beispiel zum Hörverstehen. Der Text handelt von einer jungen Amerikanerin, die in London ein Kaffeehaus eröffnet hat.

<i>Interviewer:</i>	Now, Ally, you run this company with your husband, Scott, so tell me how did it all start?
<i>Ally:</i>	Well, I've known Scott since I was fifteen and after we'd both finished college in the States, he came to England because he had got a job in a bank. We weren't married then but I decided to follow him over here. I had a degree in Media Studies, so I got a job in magazine publishing very quickly.

Ally first decided to go to England because she
A would have the chance to study there
B was offered a job in London
C wanted to be with her boyfriend

Im Beispiel oben wird der Stamm des Multiple-Choice-Items als unvollständiger Satz formuliert, der durch die Optionen ergänzt wird.

<i>Ally:</i>	The coffee thing started when on my first morning here I told Scott I'd walk him to work and we'd stop for a latte, that's a milky coffee, by the way. Scott looked at me blankly, and I just assumed he'd been working away too hard and hadn't discovered where you could get great latte coffees in London. I couldn't believe they weren't available! Years passed. I brought the subject up everywhere we went and people's eyes would light up.
--------------	---

What surprised Ally about London?
A a certain type of coffee was not on sale
B people were not interested in the quality of coffee
C the coffee bars were not conveniently located

Im Beispiel oben wird der Stamm als Frage formuliert.

Eine der Entscheidungen, die man fällen muss, wenn man textbezogene Items schreibt, ist, ob man sie als **Frage** oder als **unvollständigen Satz** formuliert. In machen Tests wird nur der eine oder der andere Typ verwendet, während andere Testentwickler eine Mischung aus Frage und unvollständigem Satz bevorzugen.

Welche Regeln gelten für das Schreiben von einzelnen und textbezogenen Multiple-Choice-Items?

- Ein Item sollte ein relevantes Element testen.
- Items sollen nicht voneinander abhängig sein, d. h., die Antwort auf ein Item sollte nicht die Antwort auf ein weiteres Item beeinflussen.
- Es sollte nur eine Lösung geben, die klar und eindeutig sein muss.
- Die Distraktoren sollten, obwohl sie falsch sind, trotzdem so plausibel sein, dass schwächere Kandidaten sie als Lösungsmöglichkeit ansehen.
- Die Optionen sollten ein kohärentes Ganzes bilden; es sollte nicht drei ähnlich aussehende Optionen geben und eine, die sich deutlich von den anderen unterscheidet.
- Wenn die Optionen den Stamm ergänzen, sollte jede Ergänzung einen grammatisch korrekten Satz ergeben. Bei einzeln stehenden Multiple-Choice-Items sollten keine Distraktoren mit inkorrekt Grammatik verwendet werden.
- Jede Option sollte möglichst gleich lang sein.
- Um die Leselast zu reduzieren, sollte jede Information, die in den Optionen wiederholt wird, aus diesen herausgenommen und in den Stamm aufgenommen werden.
- Optionen, die sich fast von selbst ausschließen, indem sie Wörter wie „immer“ oder „nie“ verwenden, sollten vermieden werden.
- Alle Optionen sollten eine ähnliche grammatische Struktur und eine ähnliche Komplexität haben.
- Negationen sollten soweit wie möglich vermieden werden, aber wenn der Stamm eine Verneinung enthält, dann sollte diese fett gedruckt werden, und alle Optionen sollten ohne Negationen formuliert werden.
- Lexikalische Hinweise, die den Kandidaten zur richtigen Antwort führen (Wort-für-Wort-Erkennen zwischen Text und Item) sollten vermieden werden. Im folgenden Beispiel ist es das Wort „fruit“ im Stamm und in einer der Optionen, das einen Hinweis auf die Lösung gibt:

Which of the following is mentioned in the recipe for Rich Fruit Cake?

- A almonds
- B milk
- C dried fruit
- D apples

- Die Position der Lösung sollte zufällig sein und jeder Buchstabe (A, B, C oder A, B, C, D) sollte gleich häufig vorkommen.

ii) Richtig/Falsch-Items

Bei einem Richtig/Falsch-Item muss der Kandidat beurteilen, ob eine Aussage im Hinblick auf den Text richtig oder falsch ist; dieser Itemtyp wird häufig in Tests zum Lese- oder Hörverstehen eingesetzt.

Beispiel:

Die Items gehören zu einem Hörtext, in dem sich zwei Personen über Fernsehen unterhalten.

	YES	NO
Tony and Rachel both dislike watching cartoons.		
Tony and Rachel both prefer watching television alone.		
Rachel thinks her mother can afford to buy her a television.		
Tony has kept his promise about watching television at night.		
Rachel wants to be able to choose when she watches television.		

Der Nachteil dieses Aufgabentyps liegt darin, dass er nur eine Wahl zwischen zwei Alternativen bietet. Damit kann ein Kandidat, der rät, immer noch eine relativ hohe Punktzahl erreichen. Das einfache Lösungsverhalten, das von den Kandidaten gefordert wird, rechtfertigt einen Einsatz nur auf einem niedrigen Niveau des Leseverstehens; für Tests zum Hörverstehen hingegen ist dieses Item geeignet. Bei Tests zum Leseverstehen kann man die Ratewahrscheinlichkeit dadurch reduzieren, dass man eine dritte Option einführt. Damit ergibt sich eine Wahl zwischen „Richtig/Falsch/Nicht im Text“ bzw. „Ja/Nein/Dazu wird nichts gesagt“, aber es könnte angemessener sein, Multiple-Choice-Items zu verwenden.

Der Typ „Ja/Nein/Dazu wird nichts gesagt“ eignet sich nicht für Tests zum Hörverstehen, da er zu großen Unsicherheiten führt. Es ist sehr schwer zu entscheiden, ob etwas gesagt wurde oder nicht, wenn man im Text nicht nachlesen kann, denn dies ist ja im Allgemeinen beim Hörverstehen nicht möglich.

iii) Lückentest mit Multiple-Choice-Items

In einem Lückentest werden Wörter getilgt und dadurch Lücken im Text geschaffen, die der Kandidat ergänzen muss, normalerweise durch ein oder zwei Wörter. Bei diesem Typ gibt es unterschiedliche Varianten:

- Die Lücken können mechanisch gesetzt werden, indem z. B. jedes sechste oder siebte Wort getilgt wird (dies ist ein Cloze-Test).
- Bestimmte Wörter können in unregelmäßigen Abständen im Text getilgt werden.

Im folgenden Lückentest werden Multiple-Choice-Items zu den Lücken vorgegeben. Es gibt aber auch Lückentests, bei denen die Lücken „offen“ sind und der Kandidat die richtige Ergänzung selbst finden und schreiben muss. Diese Lückentests mit offenen Lücken werden bei Items, zu denen Kandidaten Antworten produzieren, behandelt.

Lückentests mit Multiple-Choice-Items werden meist verwendet, um Leseverstehen, Grammatik oder Wortschatz zu testen.

Beispiel:

THE LANGUAGES OF THE WORLD							
Thousands of languages are spoken in the world today. Populations that(1)..... similar cultures and live only a short distance(2)..... may still speak languages that are quite distinct and not(3)..... understood by neighbouring populations.							
1. A	share	B	belong	C	keep	D	own
2. A	far	B	apart	C	divided	D	separated
3. A	closely	B	freely	C	smoothly	D	readily

Die Lücken im Text wurden dort gesetzt, wo der Testautor ein bestimmtes Element testen wollte, im Gegensatz zu dem Verfahren, bei dem die Lücken im Text mechanisch gesetzt werden. Man unterscheidet deshalb zwischen Lückentest (bewusste Tilgung) und Cloze-Test.

Weil durch die vorgegebenen Multiple-Choice-Items die Antwortmöglichkeiten begrenzt werden, wird dieser Itemtyp oft in Tests zum Lesen oder zur Grammatik bzw. zum Sprachgebrauch in einem Subtest eingesetzt, in dem der Wortschatz geprüft werden soll; in einem Lückentest mit offenen Lücken gäbe es zu viele Lösungsmöglichkeiten für jede Lücke. Sowohl ein Lückentest mit Multiple-Choice-Items als auch ein Lückentest mit offenen Lücken können eingesetzt werden, um grammatische Strukturen zu testen. Der Nachteil bei der bewussten Tilgung ist, dass dieser Itemtyp nur sehr eingeschränkt Kenntnisse testen kann, und dies auch nur auf der Satzebene.

Welche Grundregeln gibt es für das Schreiben von Lückentests?

- Wie bei anderen Multiple-Choice-Items auch darf es nur eine richtige Antwort geben und die Optionen müssen ein kohärentes Ganzes bilden. Ein Weg, um zu guten Distraktoren zu kommen, besteht darin, einigen Lernern den Test als Lückentext mit offenen Lücken zu geben und einige der falschen Antworten anschließend als Distraktoren zu verwenden.
- Die erste Lücke sollte nicht zu nah an den Textanfang gesetzt werden. Die nachfolgenden Lücken sollten nicht so nah beieinanderliegen, dass der Kandidat nicht mehr erkennen kann, welche Struktur getestet werden soll. Als Faustregel gilt, dass zwischen sieben und zwölf Wörter zwischen den Lücken stehen sollten.
- Man sollte möglichst vermeiden, das erste Wort in einem Satz oder Verneinungen zu tilgen. Es ist auch nicht ratsam, Wörter zu tilgen (meistens sind dies Adjektive oder Adverbien), ohne die der Satz auch komplett wäre.
- Den Kandidaten wird in der Anweisung gesagt, dass sie die Lücken mit einem Wort füllen sollen. Daher sollten zusammengezogene Formen, Wörter mit Bindestrich und jede andere Form, die den Kandidaten verwirren könnte, nicht getilgt werden.

iv) Lückentest mit Schüttelkasten

Ein ähnlicher Test wie der oben beschriebene ist der Lückentest mit einem Schüttelkasten, der die Wörter enthält, die in die Lücken passen, und einige zusätzliche Wörter, die nicht passen. Dieser Test eignet sich zum Überprüfen des Leseverstehens auf der Grundstufe.

Beispiel:

Choose a suitable word from the list given above the passage for each of the gaps, and write your answer in the space on the answer sheet.

sun	late	paper	on	went
second	work	met	a	see
but	had	enjoyed		

.....

Carla's Weekend

On Saturday morning Carla ...(1)..... down to Bournemouth to(2).... some friends. It was(3).... beautiful day so they stayed(4).... the beach all afternoon. Carla got up ...(5)....on Sunday morning. She read the(6).... and then remembered that she(7).... arranged to meet her sister for lunch. They(8).... a film in the evening.

v) Textabschnitte zuordnen

Anbei ein Beispiel für eine weitere Variante des Lückentests. Sie besteht aus einem Text, in dem sechs Abschnitte fehlen. Dem Kandidaten werden sieben Abschnitte vorgegeben, von denen er sechs in die passende Textstelle einfügen muss. Dies ist ein Test zum Leseverstehen auf relativ hohem Niveau, der das Globalverstehen eines längeren Textes sowie das Erkennen seiner Kohärenzmerkmale prüft.

Beispiel:

Ten Days Under the Sea

It was seven on a July morning, and I felt as if I had been awake for hours. I was standing with my research team in a motorboat speeding towards Conch Reef off Key Largo, Florida. We were about to start our first mission in the Aquarius underwater habitat, a six-person research station situated 6.5 kilometres off Key Largo and 15 metres below the surface. When we reached our destination on the reef, we would descend into the clear, dark blue water and stay there for rather a long time.

1. My team and I had gone through a lot to get to where we were. There had been a year of planning, four days of intensive training and, in my case, a lifetime of ambition to work underwater as a marine biologist.
2. My team-mates, pensive and quiet, seemed to be ruminating on much the same theme as we arrived at the barge, moored, and exchanged our dry shirts and sandals for damp wet-suits and ungainly fins. After years of use, the scuba gear I donned had the comfort of well-used tools, except for one crucial omission.
3. The Aquarius habitat would be our only refuge and the surface a dangerous place, where we could die in minutes. Within 24 hours of submerging, our bodies would become saturated with nitrogen gas. In this state, a rapid return to the surface would induce a severe and possibly crippling or even fatal case of decompression sickness, better known as 'the bends'.
4. Yet as I plunged into the water, I was freed from my concerns and from the weighty terrestrial world. Finally, I was able to focus my attention on the immediate goals of my research and the excitement and challenges of living underwater. My colleagues hovered below me as I adjusted my mask and then sank below the surface of the water.
5. As I continued my descent, the reef beneath me took shape, becoming a rolling landscape of underwater hills and valleys cloaked in a forest of sponges, stony corals and soft corals waving in the current. As on previous occasions, I was impressed by the odd similarity between this reefscape and a frosty winter scene in my native England.
6. My surveys suggested that many of the corals that had managed to survive by attaching themselves to solid surfaces did so in cracks and on the undersides of ledges and pieces of rubble. I began to wonder whether this kind of settlement provided a springboard for growth onto the open reef. During our first mission in Aquarius, my colleagues and I would begin trying to find out.

A

My familiar red face mask no longer had a snorkel attached to the strap. The most basic component of my regular equipment was conspicuous in its absence, reminding me that, where I was going, the surface would no longer provide a safe haven from trouble.

B

The lengthy periods in the water inevitably made us feel a greater affinity for the resident marine life than for our fellow humans visiting regularly from the surface with food and supplies. These feelings were accentuated at night, when we made short forays into an inky darkness teeming with marine creatures.

C

I was reminded of how I had first had an idea about this project some five years before, when I was struck by the devastation wreaked by Hurricane Hugo on Caribbean reefs, particularly those off St John in the US Virgin Islands. It had taken nearly six years for those reefs to recover.

D

Still, I couldn't help thinking about the things I would miss while living underneath the sea: sunshine, fresh air, open spaces, even the squadrons of pelicans that soared silently over the boat.

E

Although I had long been aware of this fact, I realised there was no turning back as I sat on the diving platform at the stern of the boat, straining to prevent myself from being pushed into the water by the heavy set of oxygen tanks on my back.

F

For the next ten days, in fact, we would be 'aquanauts', living every marine researcher's fantasy: we would spend as many as six hours a day working in the water and then retire to a warm, dry, comfortable shelter for meals, discussions, relaxation and sleep.

G

I had started hundreds of dives in similar fashion, but this one was different. Instead of surfacing after a brief visit, my colleagues and I would be down as deep as 30 metres for nearly three hours, completing three times the tricky manoeuvre of exchanging empty scuba tanks for full ones at depth.

Lösung: 1 F 2 D 3 A 4 E 5 G 6 C

Eine solche Aufgabe sollte man entweder auf einer Seite präsentieren oder auf zwei gegenüberliegenden Seiten, damit der Kandidat zwischen dem Text und den Optionen, aus denen er die Antworten auswählen muss, mühelos vor und zurückgehen kann. Dass es sechs Lücken aber sieben Optionen gibt, verhindert eine automatische Zuordnung der letzten Option, wenn alle vorhergehenden richtig gewählt wurden.

vi) Zuordnungsaufgaben

Bei den Zuordnungsaufgaben gibt es verschiedene Varianten. Ihnen gemein ist, dass Elemente aus zwei separaten Listen miteinander in Verbindung gebracht werden müssen. In ihrer einfachsten Variante wird sie auf der Grundstufe verwendet, um grammatische Strukturen zu überprüfen, wobei die erste Hälfte des Satzes in der linken Spalte vorgegeben wird und der Kandidat die richtige Ergänzung aus der rechten Spalte auswählen muss.

Eine etwas komplexere Variante der Zuordnungsaufgabe besteht darin, zwei Zusammenstellungen von kürzeren Texten vorzugeben, wobei die eine Spalte oft Personen beschreibt, die einen bestimmten Bedarf haben, z. B. eine bestimmte Art von Urlaub zu machen, eine Unterkunft zu finden oder ein bestimmtes Buch zu lesen; die andere Spalte zeigt Urlaubsangebote, Bücher etc., von denen aber jeweils nur eins genau den Anforderungen der jeweiligen Person entspricht.

Im folgenden Beispiel werden in der linken Spalte fünf kurze Beschreibungen von Personen vorgegeben, von denen jeder ein Buch kaufen möchte. Rechts davon finden sich acht Buchbesprechungen. Die Aufgabe ist, für jede Person das passende Buch zu finden.

Mit dieser Aufgabe wird das Leseverstehen getestet.

Beispiel:

Ali enjoys reading crime stories which are carefully written so that they hold his interest right to the end. He enjoys trying to guess who the criminal really is while he's reading.		A <i>London Alive</i> This author of many famous novels has now turned to writing short stories with great success. The stories tell of Londoners' daily lives and happen in eighteen different places – for example, one story takes place at a table in a café, another in the back of a taxi and another in a hospital.	B <i>Burnham's Great Days</i> Joseph Burnham is one of Britain's best-loved painters these days, but I was interested to read that during his lifetime it was not always so. Art historian Peter Harvey looks at how Burnham's work attracted interest at first but then became less popular.
Monica is a history teacher in London. She enjoys reading about the history of people in other parts of the world and how events changed their lives.		C <i>The Missing Photograph</i> Another story about the well-known policeman, Inspector Manning. It is written in the same simple but successful way as the other Manning stories – I found it a bit disappointing as I guessed who the criminal was halfway through!	D <i>Gone West</i> A serious look at one of the least-known regions of the United States. The author describes the empty villages which thousands left when they were persuaded by the railway companies to go West in search of new lives. The author manages to provide many interesting details about their history.
Silvia likes reading true stories which people have written about themselves. She's particularly interested in people who have had unusual or difficult lives.		E <i>The Letter</i> The murder of a television star appears to be the work of thieves who are quickly caught. But they escape from prison and a young lawyer says she knows who the real criminals are. Written with intelligence, this story is so fast-moving that it demands the reader's full attention.	F <i>Let me tell you ...</i> The twenty stories in this collection describe the lives of different people who were born in London in 1825. Each story tells the life history of a different person. Although they are not true, they gave me a real feeling for what life used to be like for the ordinary person.

Daniel is a computer salesman who spends a lot of time travelling abroad on planes. He enjoys detective stories which he can read easily as he gets interrupted a lot.		G <i>The Last Journey</i> John Reynolds' final trip to the African Congo two years ago unfortunately ended in his death. For the first time since then, we hear about where he went and what happened to him from journalist Tim Holden, who has followed Reynolds' route.	H <i>Free at Last!</i> Matthew Hunt, who spent half his life in jail for a crime he did not do, has written the moving story of his lengthy fight to be set free. Now out of prison, he has taken the advice of a judge to describe his experiences in a book.
Takumi doesn't have much free time so he reads short stories which he can finish quickly. He likes reading stories about ordinary people and the things that happen to them in today's world.			

Diese Art von Zuordnungsaufgabe kann auch schon auf einem relativ niedrigen Niveau eingesetzt werden. In der Anweisung wird dem Kandidaten gesagt, dass jede Option nur einmal gewählt werden kann. Dieser Hinweis ist ein wesentlicher Teil der Anweisung. Wenn es wie im Beispiel oben mehr Optionen als Personen gibt, wird eine automatische Zuordnung der letzten Option verhindert. Bei dieser Art von Aufgabe ist das Layout besonders wichtig; die ganze Aufgabe sollte auf einer Seite oder auf zwei gegenüberliegenden Seiten präsentiert werden, damit der Kandidat alle Texte leicht überfliegen kann.

vii) Mehrfachzuordnung

Bei einer Mehrfachzuordnung werden in der einen Spalte Fragen oder Satzanfänge, die sich auf einen schriftlichen Text beziehen, vorgegeben. In der anderen Spalte stehen die Optionen (Wörter oder Sätze), die aber diesmal mehrfach gewählt werden können.

Beispiel:

Diese Items folgen einem Text über Gase, die die Erde aufheizen.

What are the sources of the following gases?	A industry
carbon dioxide	B insects
1..... 2..... 3.....	C decay
CFCs 4	D motor vehicles
methane 5 6	E generating electricity from
nitrous oxide 7 8	fossil fuels
ozone 9	F reaction with sunlight
	G household products

Der Unterschied zwischen dieser Zuordnungsaufgabe und dem vorangegangenen Beispiel besteht darin, dass hier nicht angegeben wird, wie oft man die Optionen wählen kann. Da die Optionen im Verlauf der Aufgabenbearbeitung nicht weniger werden, wird die Aufgabe nicht zunehmend leichter. Sie ist eine ökonomische Art von Zuordnung und wird dazu eingesetzt, das Leseverstehen bis zu einem hohen Niveau zu überprüfen.

viii) Fehlerkorrektur: Ein Wort zu viel

Bei diesem Aufgabentyp gibt es in den meisten Zeilen des Textes ein Wort zu viel, das nicht in den Text passt. Die Kandidaten müssen dieses Wort finden und in die rechte Spalte neben die entsprechende Zeile schreiben; wenn es kein zusätzliches, falsches Wort in der Zeile gibt, muss in der Spalte rechts ein Häkchen gemacht werden.

Beispiel:

The Ski Shop	
Three years ago I spent six months working in a ski shop.	1. ✓
I had always been enjoyed skiing, and so I thought it would	2. been
be a good opportunity to earn a little bit money and to	3. bit
practise on my favourite sport. I learnt a lot while I was	4.
working there, even though it was hard work. I can now	5.
tell of someone's skiing ability just by watching them carry a	6.
pair of skis. Most people are usually agree that good skiers	7.
pick their skis up with lots confidence and carry them over	8.
their shoulder pointing forwards.	

Dieser Aufgabentyp verlangt vom Kandidaten, dass er sein Wissen über die Strukturen der Sprache einsetzt. Deshalb eignet er sich besonders zur Überprüfung des grammatischen Wissens. In bestimmten Fällen könnte dieser Itemtyp einer authentischen Situation entsprechen und deshalb sehr ansprechend sein.

Trotzdem hat dieser Itemtyp auch Nachteile. Es ist schwer, Items zu finden, die plausible Fehler beinhalten oder die nicht (wäre dies eine Realsituation) auf mehr als eine Art und Weise korrigiert werden könnten. Das Beispiel oben liegt auf der Stufe B2; der vorgegebene Text muss dem entsprechen, was ein Kandidat auf dieser Stufe selbst schreiben könnte. Die Fehler sollten für diese Stufe typisch sein, z. B. im Englischen „must to/must“, „said him/told him“, „married with/married to“, „bored/boring“, „the news is/are“ usw.

In der Kategorie der Auswahlitems gibt es noch viele weitere Varianten, z. B. aus einer Liste von Optionen die korrekten Überschriften für Texte suchen, Sätze durch Auswahl aus einer Liste von Optionen ergänzen, Bilder inhaltlich Hörtexten zuordnen oder aus verschiedenen Optionen die richtige Beschriftung für ein Diagramm finden. Die wesentlichen Merkmale dieser Varianten stimmen mit denen der oben gezeigten Beispiele überein.

c) Items mit subjektivem Beantwortungsspielraum (engl. *Candidate-supplied response item types*)

Items, bei denen der Kandidat die Antwort selbst produzieren muss, können unterschiedliche Formate haben. Das reicht von Satzergänzungen oder Kurzantworten auf eine Frage, bei der der Kandidat nur ein Wort, eine Zahl oder einen kurzen Satz schreiben muss, bis hin zu Items, die eine längere und komplexere Produktion erfordern. Im Vergleich zu den Auswahlitems sind diese Items dadurch gekennzeichnet, dass sie:

- einfacher zu entwerfen sind
- eine größere Bandbreite an Inhalten testen
- die Ratewahrscheinlichkeit minimieren
- einen kreativen Gebrauch der Sprache ermöglichen
- sowohl Fertigkeiten als auch Teilfertigkeiten messen
- eine bessere Rückwirkung auf den Unterricht haben
- bei der Bewertung einen ähnlichen Grad an Objektivität erlauben wie Auswahlitems

Aber diese Itemtypen haben auch ihre Grenzen. Es ist schwer, die Items so zu schreiben, dass die erwartete Antwort klar ist, selbst bei den Items, bei denen die Produktion stark gesteuert wird. Es gibt oft akzeptable Alternativen anstelle einer einzigen richtigen Antwort. Das erschwert die Korrektur, macht sie zeitraubend und setzt oft eine Bewertung durch Beurteiler statt einer Auswertung durch Schablone oder Scanner voraus.

Gemeinsam ist diesen Itemtypen, dass die Kandidaten selbst bei den Items, bei denen die Antwort sehr gesteuert ist, Sprache produzieren, statt aus einer Anzahl vorgegebener Optionen die richtige Antwort auszuwählen. Obwohl sie in Tests zum Lese- und Hörverstehen oder in Tests zu grammatischen Strukturen eingesetzt werden, enthalten sie immer **ein Element des Schreibens**. Manchmal erfolgt die Produktion in der Sprache des Kandidaten, aber häufiger muss die Antwort in der Zielsprache erfolgen. Dadurch entsteht ein Problem hinsichtlich der **Rechtschreibung**. Deshalb muss der Lösungsschlüssel festlegen, wie damit umzugehen ist. Wenn orthografische Fehler sanktioniert werden, können Kandidaten mit Defiziten in der Rechtschreibung nicht ihre wirklichen sprachlichen Fähigkeiten zeigen. Bewertet man die Fehler in der Rechtschreibung nicht, dann entsteht das Problem zu entscheiden, ab welchem Punkt ein falsch geschriebenes Wort das gemeinte Wort vielleicht nicht mehr zweifelsfrei erkennen lässt. Ein anderes Problem ist die sprachliche **Korrektheit** und zu entscheiden, ab welchem Punkt Fehler nicht mehr akzeptabel sind. Probleme dieser Art führen zu einer subjektiveren Bewertung.

i) Kurzantwortaufgaben

Dieser Itemtyp besteht aus einer Frage oder Stichwörtern, auf die mit einem Wort oder einem kurzen Satz reagiert werden muss. Der genaue Umfang der geforderten Antwort sollte angegeben werden. Der Itemtyp bezieht sich auf einen Text und wird in Tests zum Lese- und Hörverstehen eingesetzt.

Beispiel:

Der Text zum Hörverstehen ist ein Gespräch, in dem sich ein Mann nach einer Zugverbindung erkundigt.

TRAIN	
To:	Newcastle
Day of journey:	
Train leaves at:	
Return ticket costs:	
Food on train:	Drinks and
Address of Travel Agency	22 Street

Im Test besteht die Lösung der Items entweder aus einem Wort, einer Nummer oder einem kurzen Satz, der nicht länger als drei oder vier Wörter ist. Wie umfangreich die Antworten sein sollen, muss den Kandidaten mitgeteilt werden. Die Items müssen so formuliert werden, dass man sie nicht alleine mit Weltwissen oder gesundem Menschenverstand lösen kann.

Im Beispiel oben wird die Orthografie nur beim letzten Item explizit getestet. Der Name (Mallet Street) wird in der Aufnahme buchstabiert (wie es im wirklichen Leben vorkommen könnte) und der Punkt wird nur vergeben, wenn der Kandidat das Wort im Antwortbogen richtig geschrieben hat.

ii) Satzergänzung

Bei diesem Itemtyp wird ein Teil des Satzes vorgegeben. Der Kandidat muss dem Text die notwendige Information entnehmen, um den Satz zu vervollständigen. Dieser Itemtyp wird in Tests zum Lese- und Hörverstehen eingesetzt.

Beispiel:

Der Text ist das gleiche Gespräch wie im Beispiel oben (ein Mann erkundigt sich nach einer Zugverbindung). Das Verstehen könnte auch mit folgenden Items getestet werden:

1. The man wants to travel on
2. His train leaves at
3. A return ticket costs
4. On the train the man can buy drinks and
5. The address of the Travel Agency is 22 Street.

Im Allgemeinen ist es besser, die Lücken ans Ende der Aussagen zu setzen, damit dem Kandidaten genug Kontext gegeben wird, um die Items zu lösen.

Wenn man solche Items schreibt, muss man sicherstellen, dass es nur eine richtige Antwort gibt oder zumindest, dass die alternativen Lösungsmöglichkeiten begrenzt sind und alle im Lösungsschlüssel angegeben werden können. Davon hängt eine erfolgreiche Bewertung ab.

iii) Lückentest mit offenen Lücken

Die verschiedenen Varianten von Lückentests wurden im Kapitel zu Multiple-Choice- und anderen Auswahlitems beschrieben. In einem Lückentest mit offenen Lücken werden die Lücken vom Testautor bestimmt, der damit bestimmte Strukturen testen möchte. Die Aufgabe des Kandidaten ist, jede Lücke im Text durch ein passendes Wort zu ergänzen.

Beispiel:

A NEW CRUISE SHIP One (1) the biggest passenger ships in history, the Island Princess, carries people on cruises around the Caribbean. More than double (2) weight of the Titanic (the large passenger ship which sank in 1912), it was (3) large to be built in (4) piece. Instead, forty-eight sections (5) total were made in different places. The ship was then put together (6) these sections at a shipbuilding yard in Italy.
--

Der Lückentest mit offenen Lücken eignet sich zum Testen von grammatischen Strukturen, z. B. können Präpositionen und Teile eines Verbs getilgt werden. Häufig gibt es nur eine richtige Antwort. Wortschatz ist leichter mithilfe eines Lückentests mit Multiple-Choice-Items zu überprüfen, weil es bei offenen Lücken oft zu viele mögliche Lösungen gibt und ein solcher Test deswegen bei der Auswertung nicht sehr praktikabel ist. Die Lücken sollten zwischen jedem siebten und zehnten Wort gesetzt werden.

iv) Umformungen

Bei diesem Itemtyp wird dem Kandidaten ein Satz vorgegeben, dem der Satzanfang eines zweiten Satzes folgt und der mit anderen grammatischen Strukturen das Gleiche aussagen soll. So kann z. B. der erste Satz im Aktiv formuliert sein und der zweite muss dann im Passiv geschrieben werden. Die Aufgabe des Kandidaten ist, den zweiten Satz richtig zu ergänzen.

Dieser Itemtyp wird eingesetzt, um die grammatischen Strukturen oder das Schreiben auf Satzebene zu überprüfen.

Beispiel:

1. Maria missed the ferry because her car broke down. If..... 2. Please do not smoke in this area of the restaurant. Customers are requested.....
--

Bei dieser Art von Items ist es Voraussetzung für ein gutes Funktionieren der Items, dass es nur eine richtige Antwort bzw. eine sehr eingeschränkte Anzahl an Alternativen gibt. Es ist auch wichtig, vorher zu überlegen, was man testen möchte und für jedes Item die richtige Antwort und alle weiteren möglichen Lösungen zu bedenken.

Zum Beispiel gibt es beim ersten Item drei Elemente, die man testen möchte, für die es jeweils einen Punkt gibt. Der Lösungsschlüssel sieht folgendermaßen aus:

	Punkte
Maria's / her car had not broken down	1
Maria / she would not have missed / failed to catch the ferry	1 + 1

ODER

Maria's / her car had not broken down	1
Maria / she would / could	1
have caught / got / been able to catch the ferry	1

Der Lösungsschlüssel muss alle akzeptablen Antworten enthalten. Deshalb werden für die zweite Satzhälfte die verschiedenen Lösungsmöglichkeiten aufgeführt.

Eine Möglichkeit, den Fokus solcher Items etwas einzugrenzen, besteht darin, die Lücke in die Mitte des Satzes zu setzen, damit die Lösung von den Strukturen auf beiden Seiten eingegrenzt wird. Und wenn man vorgibt, wie viele Wörter die Lösung enthalten soll, schränkt das zusätzlich die Lösungsmöglichkeiten ein (im folgenden Beispiel sind es drei Wörter).

Beispiel:

She moved to New York in order to find singing work.

She moved to New Yorkwanted to find singing work.

Because of... / so ... / since ...

*... she**

* Abweichung vom Original

v) Wortbildung

Bei diesem Itemtyp wird ein Wort im Satz getilgt und ein verwandtes Wort wird dem Kandidaten als Lösungshilfe vorgegeben. Wenn z. B. der Kontext das Nomen *Sänger* verlangt, wird das Verb *singen* vorgegeben. Der Kandidat muss das im Kontext richtige Nomen schreiben. Dieser Itemtyp wird in Tests zu den Sprachstrukturen bzw. zum Schreiben verwendet, deren Fokus auf der Überprüfung von **lexikalischem Wissen** liegt.

Beispiel:

1. There were over fifty in the orchestra.

(MUSIC)

2. My parents me to learn to drive.

(COURAGE)

Bei dieser Art von Items ist es wichtig, dass der Satz einen kurzen und eindeutigen Kontext für die Lösung vorgibt. Im Allgemeinen sieht es natürlicher aus, in den Sätzen statt „er“ oder „sie“ richtige Namen zu verwenden.

Wenn Items zur Wortbildung in einen längeren Text eingebettet werden, steht dem Kandidaten mehr Kontext zur Verfügung als bei den Einzelsätzen oben. Im Beispiel unten erlaubt es der Kontext, auch die Verneinung zu testen (in Item 2):

In Holland, people were so desperate to own tulip bulbs that their (1) became quite extraordinary. It was not (2) for people to sell all their (3) in order to buy a single tulip bulb. The situation became so serious that laws were passed with the (4) of controlling this trade in tulips.

BEHAVE

COMMON

POSSESS

INTEND

vi) Lückentest zur Wortbildung

Neben den gezeigten Beispielen können Items zur Wortbildung auch als Variante des Lückentests mit offenen Lücken geschrieben werden. Die Aufgabe besteht in diesem Fall aus einem längeren Text, bei dem in jeder Zeile ein Wort getilgt wurde und eine andere grammatische Form dieses Wortes vorgegeben wird. Der Kandidat muss die Stelle des fehlenden Wortes finden und seine korrekte Form schreiben.

Dieser Itemtyp kann in Tests zu den **Sprachstrukturen** bzw. zum **Schreiben auf Satzebene** verwendet werden.

vii) Aus Stichwörtern einen Text erstellen

Bei dieser Aufgabe werden die lexikalischen Bestandteile eines jeden Satzes in reduzierter Form vorgegeben, die stichwortartigen Notizen ähneln. Die Aufgabe des Kandidaten ist es, aus diesen Stichwörtern einen grammatisch korrekten Satz zu formulieren, indem er die notwendigen Umformungen und Ergänzungen (Präpositionen, Artikel, Hilfsverben usw.) vornimmt. Ein Schrägstrich markiert die Stelle im Satz, an der eine Ergänzung oder Umformung notwendig ist.

Dieser Aufgabentyp wird gelegentlich in Tests zu den **Sprachstrukturen** bzw. zum **Schreiben** verwendet.

Beispiel:

Mit den folgenden Stichwörtern soll ein Antwortschreiben auf eine Einladung verfasst werden.

Dear Mr Harris

- a) I be very pleased / meet you / teachers' conference / London last year.
- b) It be kind / you / invite me / come and see you while I be / England / this summer.
- c) I hope / pay a visit / your school / 26th and 27th June if / not be inconvenient.
- d) Please / not rearrange / programme / me.
- e) I be very happy / fit in / whatever you / do at that time.
- f) I like / stay overnight / 26th June and hope / arrange accommodation / me.
- g) I telephone you once / reach London / confirm / exact time / arrival / school.
- h) I look forward / meet / again.

Yours sincerely

Was bisher über Umformungsaufgaben gesagt wurde, trifft auch auf diese Aufgabe zu. Der Testautor muss sich darüber im Klaren sein, was er testen möchte, und als Teil der Aufgabe einen Lösungsschlüssel erstellen. Der Text oben enthält vierzig Elemente, die getestet werden und jedes wird mit einem Punkt bewertet. Der Lösungsschlüssel für den ersten Satz sieht folgendermaßen aus:

I was very pleased (1 mark) to meet you (1 mark) during/(while we were) at (1 mark) the (1 mark) teachers' conference (which was held) in London last year (1 mark).

Ein Nachteil dieser Aufgabe liegt darin, dass das Erstellen eines Lösungsschlüssels kompliziert ist und die korrekte Bewertung nur schwer zu bewerkstelligen.

viii) Korrektur lesen/Fehler korrigieren

Dieser Aufgabentyp besteht aus einem Text, in dem jede nummerierte Zeile ein falsches Wort enthält. Der Kandidat muss erst das falsche Wort finden und dann die korrekte Form ans Ende der Zeile schreiben. Bei einer einfacheren Variante dieses Aufgabentyps wird das falsche Wort hervorgehoben und der Kandidat muss es nur richtig schreiben.

Dies ist eine stark gesteuerte Form des Aufgabentyps „Items mit subjektivem Beantwortungsspielraum“, den man am häufigsten in Tests zu **Sprachstrukturen** findet.

Im folgenden Beispiel gibt es entweder keinen Fehler in der Zeile oder es gibt einen Fehler bei der Rechtschreibung oder der Zeichensetzung.

What colour can do for you	
Today, colour is a dazzling background to our lives in a way that our	1 ✓
ancesters can only have dreamed about. We take colour pictures of	2 ancestors
our holidays, watch colour TV go shopping in supermarkets which	3 TV, go
vibrate with colour and we have colour printers attached to our home	4
computers. We worry about the right colours for decorating the house	5
and we have favourites and pet hates where clothes are conserved.	6

Bei dieser Aufgabe gibt es pro Item nur eine Lösung. Es ist für den Testautor wichtig zu wissen, welche relevanten Elemente überprüft werden sollen.

ix) Informationstransfer

Bei diesem Aufgabentyp müssen Informationen, die in einer bestimmten Form gegeben werden, in einer anderen Form weitergegeben werden. So kann man z. B. in einem Text Informationen über die Eigenschaften eines Geräts finden und diese verwenden, um ein Diagramm zu diesem Gerät zu beschriften; oder es sollen Zahlen aus einem Text in eine Tabelle oder Grafik übertragen werden.

Dieser Itemtyp überprüft **Sprachstrukturen** und Teilfertigkeiten des **Schreibens**.

Beispiel:

Der Input besteht aus einer E-Mail, die ein Vertreter an seinen Vorgesetzten geschrieben hat. Die Aufgabe besteht darin, diese E-Mail als Bericht für die Zentrale umzuschreiben. Der „Transfer“, um den es hier geht, betrifft das Register; die E-Mail ist informell, der Bericht soll formell sein.

E-Mail

John

Here's my impression of this year's sales conference. We were all very happy about where it was held – no problems getting to the centre. When we got there, the manager greeted us warmly and everything connected with the dinner that evening was pretty good. Unfortunately, we had problems the next day – it was incredibly hot and stuffy in the conference room so we felt very uncomfortable. Then they told us at lunch that there were lots of things on the menu we couldn't have but they didn't tell us why.

Bericht

The (1) itself was found to be very satisfactory and (2) to the centre was very straightforward. In addition, on (3) we were given a very warm (4) by the manager, and the dinner arrangements were perfectly acceptable. However, the (5) in the conference room was extremely poor, causing great discomfort to all the delegates. As for lunch, many of the dishes were (6) and no (7) was given.

Wie bei den anderen Itemtypen gibt es auch bei Items mit subjektivem Beantwortungsspielraum zusätzliche Varianten, z. B. längere Antworten auf Fragen oder Aufgaben, bei denen Informationen in eine Karte, ein Diagramm oder eine Grafik eingetragen werden müssen. Die aufgeführten Beispiele sollten einen Eindruck von der Vielfalt möglicher Itemtypen geben, bei denen der Kandidat zur Lösung des Items Sprache produzieren muss.

3. Nicht-itembasierte Aufgabentypen

a) Schreiben: Aufgaben zum erweiterten Schreiben

Das erweiterte Schreiben kann auf verschiedene Weise getestet werden. Die Aufgaben unterscheiden sich im Hinblick darauf, wie sehr sie die Antwort des Kandidaten steuern. Auch die Menge an Input kann von Aufgabe zu Aufgabe variieren; sie kann von der Vorgabe einer Überschrift reichen, gefolgt von Anweisungen, wie viele Wörter zu schreiben sind, bis hin zu einem Input von einigen hundert Wörtern (z. B. Auszüge aus Briefen, Artikeln, Anzeigen).

Je höher die Steuerung, desto größer ist die Menge an Input und desto größer ist die Leselast für den Kandidaten.

Jede der oben erwähnten Arten von Input hat ihre Vor- und Nachteile.

i) Aufgaben zum Schreiben mit detailliertem Input

Ein detaillierter Input hat zur Folge, dass die Antworten der Kandidaten einheitlicher sind, wodurch eine schnellere und zuverlässigere Bewertung möglich wird. Es bedeutet aber auch, dass sowohl Lesen als auch Schreiben getestet werden, was vom Prüfungsziel ablenken und die Frage der Validität der Aufgabe als einer Schreibaufgabe aufwerfen könnte.

Die Aufgabe hat jedoch den Vorteil, dass die Situationseinbettung authentischer gestaltet werden kann. Man kann den Kandidaten Auszüge aus Briefen oder andere Materialien als Input vorlegen und eine Antwort verlangen, die sich an einen bestimmten Leser wendet, und das Register verwendet, das durch den Input vorgegeben wird.

Anbei ein Beispiel für die Stufe B1:

This is part of a letter you receive from your English penfriend.

Help! It's my brother's 14th birthday next
month and I can't think of a present to give
him. What do teenage boys like getting as
presents in your country?

Now write a letter answering your penfriend's question.

ii) Aufgaben zum Schreiben mit geringem Input

Eine Schreibaufgabe, die nur aus einem Aufsatzthema besteht, führt zu sehr unterschiedlichen Antworten, bei denen eine faire Bewertung schwieriger wird. Im Vergleich zu einer Schreibaufgabe mit detailliertem Input prüft sie nur in geringem Maße das Verstehen des Inputs. Eine solche Aufgabenstellung bringt jedoch ein anderes Problem mit sich, denn gute Aufsätze schreiben zu können, kann auch vom Bildungshintergrund und der Kultur der Kandidaten abhängen. Das aber ist etwas, was der Testentwickler vermutlich nicht testen will.

Man kann dieses Problem durch Aufgaben entschärfen, die einen Input aufweisen, der über das Thema des Aufsatzes hinausgeht. Dabei werden Angaben zur erwarteten Textsorte, zum Schreibanlass und zum potenziellen Leser gemacht. Die schriftliche Produktion ist somit gesteuert, als wenn nur ein Thema vorgegeben wird. Den Beurteilern kann eine Bewertungsanleitung an die Hand gegeben werden, die den erwarteten Inhalt skizziert und die Kriterien aufführt, mit denen die sprachliche Produktion bewertet werden soll. Die Kandidaten können eigene Ideen ausdrücken und das Thema bis zu einem gewissen Grad ihren eigenen Interessen anpassen.

Nachstehend ein Beispiel für die Stufe B2:

A British TV company is thinking of making a film about life in your area and has asked you to give them some information. Write a **report** describing the advantages of living in your area and saying how the area might change in the future.

Der Beurteiler bewertet die schriftliche Produktion mithilfe einer für diese Aufgabe entwickelten Bewertungsanleitung:

Inhalt	Der Bericht sollte die Vorteile des Wohngebiets beschreiben sowie mögliche zukünftige Veränderungen.
Spektrum	Spektrum der sprachlichen Mittel zur Beschreibung, Meinungsäußerung und Erklärung. Lexik für die Themenbereiche Stadt/Land, Wohnen usw.
Aufbau und Kohäsion	Ein klar strukturierter Text mit Einleitung und Schluss. Untergliederungen sind von Vorteil.
Register und Format	Durchgehend korrektes Register (neutral/formal). Die formellen Aspekte eines Berichts sind nicht wichtig.
Potentieller Leser	Wird über die Wohngegend des Schreibers informiert.

Ein weiterer wichtiger Aspekt ist die Anzahl der angebotenen Themen. Wenn nur ein Thema zur Bearbeitung angegeben wird, ist dieses vermutlich nicht für alle Kandidaten von gleichem Interesse und von gleicher Relevanz. Wenn es hingegen eine Auswahl an Themen gibt, ist es schwierig, die jeweiligen schriftlichen Produktionen zu vergleichen, besonders wenn sie sich in der Textsorte unterscheiden.

Eine Möglichkeit wäre, die Aufgaben zum Schreiben zu unterteilen, wobei im ersten Teil das Thema für alle obligatorisch wäre, und im zweiten Teil gewisse Wahlmöglichkeiten angeboten werden könnten. Das Thema, das alle Kandidaten bearbeiten müssen, hätte einen Input, der keinen der Kandidaten benachteiligt, während Aufgaben im zweiten Teil auf besondere Interessen oder auf verschiedene Altersgruppen eingehen könnten. Die obligatorische Aufgabe erlaubt eine zuverlässigere Bewertung, während die zweite Aufgabe den Kandidaten die Möglichkeit gibt, ihr Bestes zu geben, indem sie über ein Thema schreiben, das sie wirklich beschäftigt und interessiert.

iii) Beispiele für Aufgaben zum Schreiben auf unterschiedlichen Stufen

a) Das folgende Beispiel ist einem Test für die Grundstufe entnommen. Es gibt einen vergleichsweise umfangreichen Lese-Input. Die Tatsache, dass die beiden Fertigkeiten nicht getrennt werden können, wird bewusst hingenommen, und der Test wird deshalb als Test zum Lesen und Schreiben bezeichnet. Die erwarteten Antworten sind vergleichbar und einfach zu bewerten.

- Read this note from Jane.
- Then address the envelope to Jane's friend at the travel agency.
- The envelope is on the back of your answer sheet.

You asked me about the Happy Tours Travel Agency. It's in the centre of Belford. You'll see it at the end of Valley Road; the number is 104. My friend there is John Spratt. The post code is BL2 9XR and the telephone number is 0871 4266.

b) Die folgende Aufgabe ist auf dem gleichen Niveau wie die vorherige Aufgabe, erlaubt bei der Lösung jedoch mehr Freiheit und Fantasie.

You must see your friend, David, before tomorrow evening.

Write a note to David.

Say:

- Why you want to see him.
- Where and when to meet you.

Write 20-30 words.

c) Das folgende Beispiel ist einem Test auf einer etwas höheren Stufe entnommen; es handelt sich aber immer noch um die Grundstufe. Durch die Kombination von Anweisung und Input wird die Produktion stark gesteuert, allerdings ist die Antwort nicht vollkommen vorhersehbar.

You have just had a very difficult week with a lot of problems at home and at college. Using the information in the diary, write a letter of about 100 words to a friend, telling him/her about your bad week.

Monday	Washing machine broke down – water everywhere.
Tuesday	No trains until 9 o'clock. Late for college.
Wednesday	O.K.!!
Thursday	Cat disappeared.
Friday	English test - really difficult.
Saturday	Broke one of grandmother's valuable plates - worth £500.

d) Das folgende Beispiel ist einem Test auf der Mittelstufe entnommen und zeigt eine Aufgabe mit detailliertem Input, der die sprachliche Produktion stark steuert.

Radford is a small industrial town without many facilities for its growing population.

You are a member of the local council, which is meeting to decide how to use an empty site near the centre of the town.

The four most popular ideas for the development are given below.

Using the information given, complete the paragraphs on the next page, giving reasons for your choices.

(Illustrations of a street market, a park, an arts centre and a sports hall follow)

Die Kandidaten müssen bei ihren Ausführungen drei Abschnitte vervollständigen. Der Anfang eines jeden Abschnitts wird vorgegeben:

- a) I am most in favour of
.....
- b) My second choice would be
.....
- c) However, I would not.....
.....

Wenn der Testentwickler solche sehr gesteuerten Aufgaben zum Schreiben einsetzen möchte, muss der Testautor eine Musterantwort mitliefern. Er muss genau aufführen, welche inhaltlichen und sprachlichen Aspekte positiv bewertet werden sollen, damit eine leicht handhabbare Bewertungsanleitung erstellt werden kann.

e) Das nächste Beispiel entstammt dem gleichen Test wie das Beispiel d) oben, zeigt jedoch eine Aufgabe des entgegengesetzten Typs, nämlich ein Aufsatzthema, das den Kandidaten viel Spielraum gibt, ihre eigene Meinung auszudrücken und ihre Kreativität einzusetzen. Diese beiden Beispiele zeigen deutlich die unterschiedlichen Anforderungen, die an die Beurteiler bei der Bewertung von Aufgaben zum schriftlichen Ausdruck gestellt werden.

You recently took part in a class discussion about choosing an interesting job. Your teacher has now asked you to write a composition, answering the following question and giving reasons for your choice.

Would you rather be a politician, a teacher or a musician?

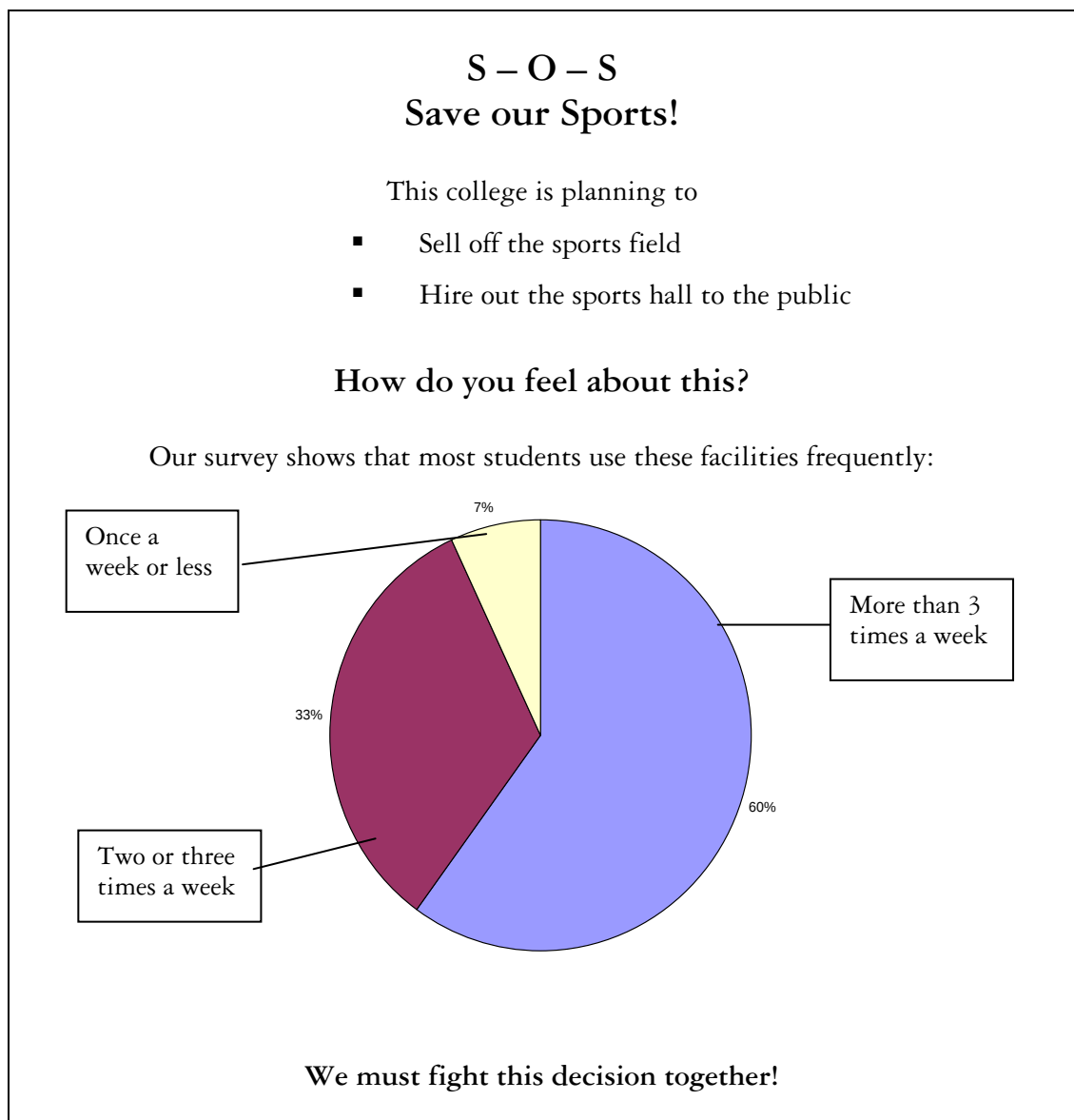
f) Das folgende Beispiel ist eine Aufgabe, bei der die Kandidaten einen Input von ca. 400 Wörtern verarbeiten und die darin enthaltenen Informationen angemessen umsetzen müssen, um die Aufgabe (einen Artikel und einen Brief zu schreiben) lösen zu können. Das Beispiel ist der Prüfung *Certificate of Advanced English* auf der Stufe C1 entnommen. Die Kandidaten müssen das Input-Material sorgfältig lesen und die wichtigen Informationen herausfiltern. Input-Material kann aus unterschiedlichen Kombinationen von Texten und Notizen bestehen, die manchmal auch von Bildern oder Diagrammen begleitet sein können. Die Aufgabe unterteilt sich oft in mehrere Teile. Die Aufgabentypen in den verschiedenen Teilen sind unterschiedlich. Es kann sich um das Schreiben eines formellen Briefes, eines Berichts, eines Artikels handeln oder um jede Kombination dieser Textsorten.

Part 1 (CAE)

1 You are a member of the student welfare committee at your college. The committee has recently received a memo from the college Principal, announcing major cuts to the existing sports facilities. The committee is opposed to these reductions and has prepared the poster shown below.

You have offered to write an article for inclusion in the next college newsletter in order to ensure that students have as much information as possible, and to get their support. You have also been asked to reply briefly to the Principal's memo, informing him of the results of the survey and making him aware of the committee's opposition to the cuts.

Read the poster below and the Principal's memo on the next page, to which you have added your comments. Then, **using the information carefully**, write the article and letter as instructed on the next page.



To: The Student Welfare Committee

From: E G Baton, Principal

REVIEW OF SPORTS FACILITIES

*We
didn't!*

As you know, the college has been forced to consider various cost-cutting measures recently.

In connection with this, my colleagues and I have been discussing the college sports facilities. In our view, these are very much under-used. We have estimated that a small number of students use the sports hall, and the sports field seems to be used only twice a month for football matches.

*Not true –
NB our
survey**What about
practice?!*

Due to this apparent lack of student interest, I have decided on some immediate changes. I intend to make the sports hall available for hire to the public during the day. I would also like the committee to introduce a student membership charge – students will be required to pay for these facilities.

All day?

Finally, the sports field has been put up for sale. At least one buyer is interested in the site for building development.

*Unfair**Never!!*

I trust the committee will support my decision.

E. G. Baton

Principal.

Schreiben Sie:

- a) einen **Artikel** für die Studentenzeitschrift Ihrer Hochschule (ca. 150 Wörter)
- b) einen angemessenen **Brief** an den Dekan (ca. 100 Wörter). Die Anschrift müssen Sie nicht angeben.

Drücken Sie sich so weit wie möglich mit Ihren eigenen Worten aus.

g) Das letzte Beispiel ist eine Aufgabe, die der gleichen Prüfung entnommen ist wie das Beispiel f). Diese Aufgabe gibt mehr Input vor als nur ein Thema, über das geschrieben werden soll, und erlaubt eine gewisse Freiheit bei der Ausführung, nennt jedoch deutlich die Rahmenbedingungen, an die sich die Kandidaten halten müssen. Es ist in diesem Fall nicht weiter problematisch, dass das Thema nicht unbedingt für alle Kandidaten attraktiv und relevant ist, da diese Aufgabe eine von vier Aufgaben ist, aus denen der Kandidat wählen kann.

You are a member of the students' committee at a college where there are many students from all over the world. The college has a lot of sports facilities which students can use in their spare time. There is also a wide range of recreational activities which students can participate in. The committee is concerned that many students are not taking advantage of all there is on offer.

You have been asked to write a leaflet which:

informs the students about the facilities and activities that are available
points out the benefits of taking up these opportunities
encourages students to use the facilities and join in the activities

Write the text for the leaflet.

b) Sprechen: Aufgaben zum Sprechen

Bei Tests zum Sprechen wird häufig über das Prüfer-/Beurteilertraining, die Rolle der Prüfer und die Zuverlässigkeit ihrer Bewertungen diskutiert, aber selten über die Aufgaben selbst. Bei den Testverfahren und den Aufgaben gibt es jedoch eine Reihe von Merkmalen, die bei der Entwicklung eines Tests zum Sprechen in Betracht gezogen werden müssen, unabhängig von den Prüfern.

Diese Merkmale betreffen Folgendes:

- ob der Test als Gespräch mit einem Prüfer durchgeführt wird oder in einem Sprachlabor mit Aufgaben von einem Tonträger
- ob Rollenspiele eingesetzt werden oder ob die Kandidaten aufgefordert werden, über sich und ihre persönlichen Ansichten und Meinungen zu sprechen
- ob die Vorgaben mündlich, schriftlich, in Form von Bildern oder als Kombination von allem gegeben werden
- wie viele Kandidaten gleichzeitig geprüft werden und ob sie aufgefordert werden, untereinander zu kommunizieren oder nur mit dem Prüfer zu sprechen

Wie bei den Aufgaben zum Schreiben steuert der Input die mündliche Produktion auf unterschiedliche Weise. Auch hier gilt: Je komplexer der Input, desto gesteueter die Antwort und umso leichter die Bewertung.

In seiner einfachsten und am wenigsten gesteuerten Form besteht ein Test zum Sprechen aus Fragen des Prüfers an den Kandidaten, die sich an einer Themenliste orientieren. Der Test könnte mit einer „Aufwärmphase“ beginnen, in der der Prüfer einige allgemeine Fragen stellt. Insgesamt kann man jedoch nicht sagen, dass sich ein solches Prüfungsgespräch aus Aufgaben zusammensetzt.

Etwas mehr Steuerung besitzt ein Test zum Sprechen, wenn der Prüfer eine Aufgabe stellt, indem er mündliche Anweisungen und einen Input (in Form von Bildern, einem Text usw.) vorgibt und dem Kandidaten eine gewisse Zeit lässt, seine Antwort zu formulieren. Die Sprache des Prüfers und auch die des Kandidaten können erheblich dadurch gesteuert werden, dass die Prüfer sich an vorgegebene Fragen halten müssen (an einen sogenannten *interlocutor frame*), der sicherstellt, dass das Prüfungsgespräch in seiner Grundform für alle Kandidaten gleich abläuft, selbst wenn sich einzelne Aufgaben unterscheiden können. Dieser Punkt wird in Modul 4 ausführlicher diskutiert.

Die Wahl des Aufgabentyps wird bis zu einem gewissen Grad davon beeinflusst, ob die Kandidaten einzeln, paarweise oder in größeren Gruppen geprüft werden. In Paar- oder Gruppenprüfungen sind Aufgaben möglich, die eine Diskussion untereinander oder die Zusammenarbeit bei der Lösung eines Problems erfordern. Aufgaben, die den Kandidaten hintereinander gestellt werden, können thematisch verbunden sein. Weiter unten werden Beispiele gegeben, die die Bandbreite der Aufgabentypen in einem Prüfungsgespräch verdeutlichen sollen.

i) Vortrag

Die Kandidaten werden gebeten, eine bestimmte Zeit über ein Thema zu sprechen; der Vortrag kann entweder einige Zeit vorher vorbereitet werden oder das Thema wird kurz vor der Prüfung bekannt gegeben. Dieses Verfahren erlaubt dem Kandidaten, über Themen zu sprechen, die ihn betreffen und interessieren. Der Nachteil ist, dass der Kandidat den Vortrag vorher auswendig lernen kann, da die Themen zum Teil in verschiedenen Prüfungen eingesetzt werden und sich dadurch wiederholen können.

Wenn in einem Prüfungsgespräch mehr als ein Kandidat geprüft wird, kann eine Aufgabe darin bestehen, dass sich die Kandidaten gegenseitig zu ihrem Vortrag Fragen stellen und beantworten. Damit wird sichergestellt, dass jeder Kandidat dem anderen aufmerksam zuhört. Die Aufgabe überprüft ihre interaktive Kommunikationsfähigkeit.

Das folgende Beispiel zeigt die Anweisungen, die den Kandidaten bei der Aufgabe „Vortrag“ gegeben werden.

In this part of the test you are each going to talk on your own for about two minutes. You need to listen while your partner is speaking because you'll be asked to comment afterwards.

I'm going to give you a card with a question written on it and I'd like you to tell us what you think. There are also some ideas on the card for you to use if you like.

A

How are young people's lives changing?

- learning methods
- opportunities
- expectations

B

What role can old people play in society today?

- the family
- the workplace
- public life

Im folgenden Beispiel, einem Test für Wirtschaftsenglisch auf fortgeschrittenem Niveau, benötigt der Kandidat ein breites Spektrum an Fachlexik und an Fachwissen, um Vorträge zu fachlichen Themen halten zu können:

- A
Customer Relations: the importance of making customers feel valued.
- B
Strategic Planning: how to decide whether to purchase or rent company premises.
- C
Finance: how to decide whether to float a company on the stock market.

ii) Vorgabe von Bildern

Kandidaten sehen sich ein oder mehrere Fotos an. Je nach ihrem sprachlichen Niveau kann die Aufgabe darin bestehen zu beschreiben, was sie auf dem Foto sehen, was dort passiert, bei mehreren Fotos ein gemeinsames Thema vorzuschlagen oder ihre Meinung zum Inhalt des Fotos zu äußern. Dann kann eine Diskussion über das Thema des Fotos folgen.

Beispiel:

In Teil 2 der Prüfung zum *First Certificate in English* vom Juni 2004 bekommen die Kandidaten zwei Fotos: eins von einem Bergsteiger, der auf der Spitze eines Berges zeltet, und ein Foto von Personen, die eine Wildwassertour machen. Die Aufgabe sieht folgendermaßen aus:

Interlocutor: Now, (*Candidate A*), here are your two photographs. **They show people enjoying the natural world.** Please let (*Candidate B*) have a look at them.

Hand over picture sheet to (Candidate A).

I'd like you to compare and contrast these photographs, and **say why you think the people have chosen to go to these places.**

Remember, (*Candidate A*), you have only about a minute for this, so don't worry if I interrupt you. All right?

1 minute

Candidate A:.....

Interlocutor: Thank you. *Retrieve photographs.*

(*Candidate B*), **would you like to spend time in either of these places?**

Approx. 20
seconds

Candidate B:.....

Interlocutor: Thank you.

iii) Schriftliche Vorgaben

Diskussionen können durch eine kurze schriftliche Vorgabe stimuliert werden, wie z. B.:

Today's news events are tomorrow's history

Die Kandidaten sollen zu dieser Aussage Stellung nehmen und einige Ereignisse von nationaler und internationaler Bedeutung schildern, die in letzter Zeit oder während der Lebenszeit der Kandidaten stattgefunden haben, und von denen sie glauben, dass sie die Geschichte beeinflusst haben oder beeinflussen werden. Mit einem so minimalen Input wie diesem ist die Aufgabe sehr anspruchsvoll.

In einer Paarprüfung können Aufgaben gestellt werden, bei denen die Kandidaten zusammen arbeiten müssen. Den Kandidaten können schriftliche Anweisungen oder Bilder vorgegeben werden, die die Grundlage für die gemeinsam zu lösende Aufgabe bilden. Eine solche Aufgabe prüft die interaktive Kommunikationsfähigkeit (hören, sich in das Gespräch einklinken, das Gespräch eröffnen, reagieren).

Anbei ein Beispiel aus einem Test auf der Stufe B2.

I'd like you to imagine that a busy international hotel is looking for staff for the holiday season. Here are some of the jobs available.

Candidates are given picture representing 7 different jobs.

Talk to each other about how difficult it would be to do these jobs without training. Then say which job you would find the most difficult.

iv) Aufgaben mit einer Informationslücke

Bei Tests zum Sprechen, besonders bei Paarprüfungen, wird häufig versucht, ein echtes Kommunikationsbedürfnis zu stimulieren. Dies geschieht dadurch, dass man etwas gemeinsam herausfinden oder gemeinsam ein Problem lösen muss, oft auf der Grundlage von Bildern. Anbei einige Varianten für solche Aufgaben:

- Jeder Kandidat hat ein Bild, das der andere nicht sehen kann. Obwohl beide sehr ähnliche Bilder erhalten, sind sie doch unterschiedlich; ein Kandidat muss dem anderen sein Bild beschreiben und die Unterschiede zu seinem Bild feststellen.
- Jeder Kandidat erhält einen Satz gleicher Fotos, die aber jeweils in verschiedener Reihenfolge angeordnet sind. Ein Kandidat beschreibt eines seiner Fotos und der andere muss herausfinden, welche Nummer dieses Foto in seinem Satz hat.

Solche Aufgaben eignen sich am besten für Tests auf der Grundstufe, da die zu produzierende Sprache einfach und bis zu einem gewissen Grad repetitiv ist.

4. Anweisungen

Die Definition von Anweisung, die in der Einführung von Modul 1 gegeben wurde, lautete „Instruktionen, die den Kandidaten gegeben werden, um ihre Antwort auf eine bestimmte Testaufgabe zu steuern“. Die Anweisungen sollten auch angeben, wie und wo die Antwort festgehalten werden soll.

Die Anweisung, die der Prüfer einem Kandidaten gibt, sollte die Aufgabe so klar und deutlich wie möglich darstellen. Es sollte keine Verwirrung entstehen und keine Klarstellung notwendig werden, da beides beim Kandidaten Angstgefühle erzeugen und seine Leistung beeinträchtigen kann, was sich auf die Reliabilität und die Validität der Leistungsmessung auswirkt.

In Tests zum Hörverstehen sollten die Anweisungen auf dem Aufgabenblatt stehen und auch vom Band zu hören sein.

In Tests zum Sprechen (d. h. bei Prüfungsgesprächen) werden die Anweisungen mündlich vom Prüfer gegeben. Hier kann der Kandidat Gelegenheit zum Nachfragen bekommen, was als Teil der Interaktion zu rechtfertigen ist.

Im Anschluss ein Beispiel für eine Anweisung bei einer Aufgabe zum Schreiben, das dem *First Certificate in English* entnommen ist:

A British TV company is thinking of making a film about life in your area and has asked you to give them some information. Write a report describing the advantages of living in your area **and** saying how the area might change in the future.

Write your **report**.

Die Fragen, die man sich bei der Formulierung einer Anweisung stellen muss, sind:

- **Wie klar ist die Anweisung?** (Ist es möglich, die Aufgabe falsch zu verstehen?)
- **Ist sie leicht zu verstehen?** (Ist die Sprache auf dem richtigen Niveau? Dies ist bei Tests auf niedrigem Niveau besonders wichtig.)
- **Wie angemessen ist sie?** (Gibt sie alle nötigen Informationen?)
- **Wie relevant ist sie?** (Gibt sie nur die nötigen Informationen?)
- **Sind die Anweisungen einheitlich?** (Die Formulierungen der Anweisungen sollten im ganzen Test standardisiert sein, damit der Kandidat soweit wie möglich auf bekannte Muster von Anweisungen trifft. Die Anweisungen sollten auch über verschiedene Versionen eines Tests hinweg einheitlich sein.)

Das Schreiben von Anweisungen ist ein wichtiger Bestandteil der Aufgabenerstellung. Testautoren sollten sich deshalb eine Checkliste anlegen, die folgendermaßen aussehen könnte:

- Stimmt die Anweisung mit den Richtlinien für den Test überein?
- Wenn die Anweisung dem Kandidaten neu ist, wurde dann mindestens ein klares Beispiel gegeben?
- Ist die Anweisung sprachlich korrekt und dem Niveau des Tests angemessen? In einem Test zum Leseverstehen bedeutet dies, dass die Sprache der Anweisung im Niveau *niedriger* sein muss als das Niveau, das man testen möchte.
- Ist die Sprache einfach und klar?
- Gibt es überflüssige Formulierungen?
- Gibt es doppelte Verneinungen?
- Gibt es zweideutige Formulierungen und sind Missverständnisse möglich? Es ist für Testautoren hilfreich, einen Kollegen zu bitten, das Item zu lösen, da man als Autor oft zu „nah“ am Text ist, um alle potenziellen Probleme zu erkennen.
- Gibt die Anweisung alle notwendigen Informationen und nennt sie die Einschränkungen?

Wichtige Einzelheiten, die voraussichtlich in einer Anweisung genannt werden, sind:

- Wo genau befindet sich der Input, auf den sich die Anweisung bezieht (z. B. Seitenzahl)?
- Wie viele Wörter werden in der Antwort verlangt?
- Kann eine Option einmal oder mehrmals gewählt werden?
- Können die Optionen in beliebiger Reihenfolge geschrieben werden oder nicht?
- Wie viele Wörter sollen bei einer Aufgabe zum Schreiben ungefähr geschrieben werden?
- Sind die Angaben zu den Wahlmöglichkeiten zwischen verschiedenen Aufgaben klar?
- Wie oft wird der Hörtext gehört?
- Inwieweit kann das Inputmaterial in der Antwort verwendet werden?
- Gibt es klare Hinweise zu den Kriterien, die eine erfolgreiche Ausführung der Aufgabe definieren?

Diese detaillierten Informationen sollten in Handreichungen für Testautoren aufgeführt werden, damit diese sich bei der Testerstellung daran halten können und sie zukünftig immer einheitlich umsetzen können. Der Testautor kann sich Schablonen für die Anweisungen erstellen, die ihm das Schreiben von Items erleichtern.

5. Lösungsschlüssel, Bewertungsvorgaben, Bewertungsanleitung

Jede Testaufgabe muss begleitet sein von:

- einer entsprechenden Anweisung
- einem korrekten Lösungsschlüssel (d. h. den richtigen Antworten bei objektiven Items)
- Bewertungsvorgaben für die produktiven Teilfertigkeiten mit allen richtigen Antworten (Kurzantwortaufgaben und gelenkte Aufgaben zum Schreiben)
- einer Bewertungsanleitung für die subjektiv zu bewertenden Tests, die die Bewertungskriterien und die Merkmale der Aufgabenerfüllung definiert

Eine Checkliste für Testautoren könnte folgende Fragen enthalten:

- Wurde eine angemessene Musterantwort geschrieben?
- Wenn es mehr als eine richtige Antwort gibt, wurden dann alle möglichen Lösungen in den Bewertungsvorgaben aufgeführt?
- Ist der Lösungsschlüssel bzw. sind die Bewertungsvorgaben klar und einfach zu handhaben?
- Gibt es klare Hinweise zu den Punktwerten für jede richtige Antwort oder für eine teilweise richtige Antwort?
- Ist die Anzahl möglicher Lösungen ausreichend begrenzt? Wenn viele richtige Antworten möglich sind, führt das zu Problemen bei der Auswertung, da der Test ggf. von Personen ausgewertet wird, die keine Sprachexperten sind.
- Wurden alle Einschränkungen genannt? (Z. B.: Die Kandidaten müssen aus einer Liste mit fünf Aufgaben zwei auswählen – kein Punkt wird vergeben, wenn mehr als zwei Aufgaben ausgewählt wurden.)

Es gibt verschiedene Methoden zur Bewertung des Sprechens und des Schreibens, die in Modul 4 vorgestellt werden. Im Allgemeinen geschieht die Bewertung anhand einer Bewertungsanleitung, die z. B. die zu bewertenden sprachlichen Leistungen in Bereiche unterteilt wie Aussprache, Intonation, Strukturen und jeden Bereich auf einer Skala von 1-5 bewertet. Um dem Beurteiler zu helfen, wird eine kurze Beschreibung einer typischen Leistung auf jeder Stufe beigelegt. Der Gesamtpunktwert für die Aufgabe zum Sprechen oder zum Schreiben ergibt sich dadurch, dass für jeden Bereich ein bestimmter Punktwert (aus 5) vergeben wird und diese Punktwerte aus allen Bereichen (z. B. insgesamt sechs) addiert werden. Dies ergibt dann einen bestimmten Gesamtpunktwert aus einem Maximum von 30.

Eine etwas andere Methode der Bewertung wird bei Aufgaben angewendet, wie sie die Beispiele f) und g) zur Fertigkeit Schreiben oben zeigen. Da diese Aufgaben in unterschiedlichem Ausmaß durch den Input gesteuert werden, ist es notwendig, **aufgabenspezifische Bewertungsanleitungen** zu erstellen. Unter den Überschriften **Inhalt, Aufbau und Kohäsion, Spektrum, Register, potenzieller Leser und Korrektheit der Sprache** werden kurze Beschreibungen der erwarteten Leistung gegeben, die einen Punktwert von 5 erhalten würden. Dies wird ergänzt durch allgemeine Beschreibungen der Leistungen, die auf jeder Stufe erwartet werden. Aufgabenspezifische Bewertungsanleitungen werden von den Testautoren erstellt und in den Erprobungen kontrolliert, indem man sie zur Bewertung vieler Kandidatenbeispiele einsetzt.

Der Testautor sollte immer eine **Musterantwort** schreiben, unabhängig davon, ob dies verlangt wird oder nicht. Es ist wichtig zu kontrollieren, ob das Thema mit der vorgegebenen Anzahl an Wörtern behandelt werden kann und ob dies auf dem sprachlichen Niveau möglich ist, das von den Kandidaten erwartet wird. Bei Aufgaben dieser Art können zwar mögliche Fehler in der Erprobung festgestellt werden, aber es sollte alles unternommen werden, um Fehler schon vorher auszuschließen.

6. ÜBUNGEN

1. In jedem der folgenden Items gibt es einen offensichtlichen Fehler.

Sagen Sie, was mit jedem Item nicht stimmt und entscheiden Sie, ob man es so umformulieren kann, dass daraus ein akzeptables Item wird oder ob Sie es verwerfen würden.

- a) *In the story, the merchant was unhappy because it*
- A *rained.*
 - B *was windy.*
 - C *was dark*
 - D *was windy and rainy and he had forgotten his overcoat.*
- b) A *tea* B *tee* C *party* D *dinner*
- c) A *ran* B *fast* C *runs* D *is running*
- d) *What is the physical state of sodium at room temperature?*
- A *fluid*
 - B *gas*
 - C *liquid*
 - D *solid*
- e) *What happened in the period 1980 - 1985?*
- A *The birth rate rose.*
 - B *The birth rate fell.*
 - C *The birth rate remained unchanged.*
 - D *The death rate fell.*

-
- f) *They said they*
- A *had gone.*
 - B *have went.*
 - C *had go.*
 - D *had went.*
- g) *When is it not appropriate not to be absent from class?*
- A *When you are sick.*
 - B *When you are young.*
 - C *While class is in session.*
 - D *Whenever the teacher is angry.*
- h) *We learn from this passage that Napoleon was*
- A *British.*
 - B *French.*
 - C *Polish.*
 - D *German.*
- i) *The boy took the newspaper*
- A *because he wanted to read it.*
 - B *because he wanted to wrap a gift in it.*
 - C *because he wanted to dispose of it.*
 - D *because he wanted to remove an article.*

j) *At the end of the lesson the teacher..... up the children's books.*

A *picked*

B *collected*

C *took*

D *pulled*

k) *Martha arrived late for her interview because*

A *she overslept that morning.*

B *her train arrived late.*

C *of the train drivers' strike.*

D *she missed her train.*

l) *Why did Mr Young go to the bank?*

A *to get a new cheque book*

B *to buy a sandwich*

C *to ask for some advice*

D *to pay in a cheque*

m) *Which magazine should you buy if your hobby is gardening?*

A *Beautiful Homes*

B *At Home With The Stars*

C *Home and Family*

D *Homes and Gardens*

2.

Read the four test items and then do the task which follows them.

1.

My village is called Compton. There are only three shops in my village. There is a bakery, a butcher's and a newsagent's. The bakery is open every day from Monday to Saturday. It opens at eight o'clock every morning and closes at four o'clock. The butcher's opens the same hours on the same days, but it is closed all day Wednesday. The newsagent's is open from seven o'clock every morning Monday to Saturday. It closes nearly every day at five o'clock. On Wednesdays it closes at one o'clock.

Which shops are open in Compton on a Saturday morning?

Answer: All three; the bakery, butcher's and newsagent's.

2.

By now I've got the whole place clearly charted.

Our garden, first: where I did not invent

Blinding theologies of flowers and fruits,

And wasn't spoken to by an old hat.

And here we have that splendid family

I never ran to when I got depressed,

The boys all biceps and the girls all chest,

Their comic Ford, their farm where I could be

'Really myself'. I'll show you come, to that,

The bracken where I never trembling sat,

What, in your own words, is Philip Larkin, the poet writing about?

3.

Friends, Romans, countrymen, lend me your ears;

I come to bury Caesar, not to praise him.

The evil that men do lives after them,

The good is oft interred with their bones;

Who spoke these words?

4.

A man is returning home from a day at the market. He has bought a chicken, a fox and a bag of grain. Don't ask me why he bought a fox! He comes to a river. It is too deep to wade across and too rough to swim. There is a small boat moored at the bank, but it is too small for the man and all his possessions. He can only take one thing across at a time. He can row backwards and forwards as many times as he likes, but if he leaves the fox and the chicken alone, the fox will eat the chicken. If he leaves the chicken and grain, then the chicken will eat the grain. How can he get everything across the river?

Answer:

- i) He takes the chicken across and leaves it.
- ii) He comes back and takes the fox across and leaves it.
- iii) He brings the chicken back with him.
- iv) He takes the grain across and leaves it with the fox.
- v) He comes back for the chicken.

Sehen Sie sich die Texte und die Items noch einmal an und überlegen Sie, was jedes Item testen soll und wie schwierig es ist.

Unten finden Sie eine Liste der Punkte, die in den Beispielen veranschaulicht wurden. Können Sie dieser Liste (einen) weitere(n) Punkt(e) hinzufügen?

- sprachliche Komplexität des Textes oder des Stimulus
- sprachliche Komplexität der geforderten Antwort
- kognitive Komplexität der Aufgabe
- erforderliches Weltwissen anstelle von Sprachkenntnissen
- sprachliche Merkmale des Texttyps