German

TestDaF
Test Deutsch als Fremdsprache

INDEX

Introduction to the TestDaF-Institut in Hagen/Germany	2
TestDaF Reading Comprehension	6
TestDaF Listening Compehension	11

Introduction to the TestDaF-Institut in Hagen/Germany

The TestDaF-Institut was founded in January 2001 to develop and administer the TestDaF (**T**est **D**eutsch als **F**remdsprache), a measure of students' proficiency in German as a foreign language. The main objective was to facilitate foreign students' access to German institutions of Higher Education. Among the founding members of the Board are the German Academic Exchange Service (DAAD), the Goethe-Institute and the German Rectors' Conference. The TestDaF-Institut is affiliated to the University of Hagen and the Ruhr-University Bochum.

The test format of the TestDaF was developed by a group of experts during a three year project period from 1998 to 2000 commissioned by the DAAD. This development project was financed by the Federal Ministry of Education and Research and the Federal Foreign Office. The DAAD fixed the terms of reference for the test development in 1998, but is no longer involved in the development of test tasks or items.

The TestDaF is an international examination that can be taken in the candidate's home country. It is currently offered in more than 340 licensed test centres in 77 countries including Germany. Test centres must meet strict criteria for accessibility, security, and technical capacity. Their exam officers are trained on administration security and logistics. Since 2001 the number of test administrations per year has increased considerably: two administrations in 2001, three in 2002, four worldwide and two extra administrations in China in 2003 and 2004. In 2005 five administrations will take place worldwide and two again in China. The candidature has grown from 1,190 participants in 2001 to 8,981 participants in 2004.

See: www.testdaf.de

Test Development

The TestDaF-manual for item writers provides the necessary information about the test's construct and its content. Item writers get a clear definition and detailed description of the tasks and items for each subtest.
There are six main stages in the development process of all four TestDAF subtests: commissioning, editing/vetting, paper construction, pretesting, pretest review and question paper production.

Stages of test development:

1 Commissioning

Trained item writers (external consultants) are commissioned to write the test tasks and items for the question papers. All item writers are qualified in Applied Linguistics and teaching German as a foreign language. Item writers get training from the TestDaF-Institut, concerning the theoretical basis of the test, the test format and item writing. Furthermore, a manual for item writing containing a detailed description of the tasks and four subtests is provided for their work.

2 Editing/Vetting

During the process of item writing the consultants stay in contact with the test developers responsible for the specific subtest concerned (reading comprehension, listening comprehension, oral production and written production) at the TestDaF-Institut. Item writers have to check the quality of their material according to the definitions and criteria of the TestDaF-manual for item writers and make the necessary modifications to reach an acceptable standard for pretesting. The test developers view the material submitted, to assess its suitability for use as test tasks and items, and to reject unsuitable, or problematic material. At the end of the editing phase an internal group of experts selects the most appropriate tasks for pretesting.

3 Paper Construction

The selected tasks are combined into a complete question paper for each subtest and then into a complete test of four subtests for a TestDaF exam. The internal group of experts will make sure that no overlap of topics exists across the four subtests. Also, marking keys are added to the tasks.

4 Pretesting

Pretesting provides the statistical information about the items necessary to ensure the quality of all TestDaF test material. First the test material, i.e. the complete question papers of four subtests, is piloted on a small group of native and non-native speakers in German in order to get information about performance aspects of the test material. After this first check of the test material a meeting is held at the TestDaF-Institut between psychometricians and test developers to discuss the performance of the material. If necessary, test tasks and items are revised before the second step is started: a pretesting procedure organised in Germany and in other countries with a larger representative group of potential test-takers. Each reading and listening comprehension question paper is pretested with 150-200 candidates and each written and oral production paper with 100-150 candidates. The cut-off scores to reach the level TDN 3, 4 or 5 are fixed during the pretesting procedure. The difficulty of the items in the reading and listening comprehension subtest is evaluated by linking them to an additional calibrated anchor-test, in this case a C-test. The scorebands corresponding to each level are fixed on the basis of this analysis.

5 Pretest Review

After pretesting, a meeting is held at the TestDaF-Institut between psychometricians and test developers to discuss the performance of the material. The meeting reviews the results of the pretests and evaluates the measurement characteristics of tasks and items. Furthermore a group of experts, composed of experienced raters and members of the TestDaF-Institute, meets after pretesting to evaluate the validity of new test tasks, especially those for the two productive skills. If any profound adjustment is necessary at that stage, tasks and items will have to be revised and then again pretested.

6 Question Paper Construction and Production

After the pretest review, the question papers of each subtest displaying satisfactory statistical features enter the item bank. The internal group of experts is responsible for building up a new combination of pretested question papers for each subtest for use in composing the question papers of a complete exam. The group of experts makes sure that a range of topics is maintained in the different versions of the exam.

Statistical Analyses

All the TestDaF test materials are validated by means of careful psychometric and statistical analysis before use in live tests. Feedback is gathered systematically by the institute's data analysts and psychometricians during pretesting and again after the live examinations. More specifically, relevant features of the TestDaF analytic procedures include the following:

- Each item or task is subjected to an extensive piloting and trialling process before being considered for inclusion in a TestDaF exam.
- Rasch theory provides the primary basis for constructing as reliable, valid, and fair measures of each language skill as possible.
- Test equation procedures based on a Rasch approach are used to ensure equivalence from one TestDaF exam to the next.
- With writing and speaking, an examinee's performance is graded by trained raters using a detailed list of criteria; raters are monitored on a regular basis.
- Many-facet Rasch measurement is routinely applied to detect and measure rater effects in the writing and speaking sections; reported scores are adjusted for differences in rater severity/leniency.

For more detail see:

Eckes, T. (2003). Qualitätssicherung beim TestDaF: Konzepte, Methoden, Ergebnisse. Fremdsprachen und Hochschule, 69, 43–68.

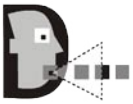
Eckes, T. (2004). Beurteilerübereinstimmung und Beurteilerstrenge: Eine Multifacetten-Rasch-Analyse von Leistungsbeurteilungen im „Test Deutsch als Fremdsprache" (TestDaF). Diagnostica, 50, 65–77.

Eckes, T., Ellis, M., Kalnberzina, V., Pi_orn, K., Springer, C., Szollás, K., & Tsagari, C. (in press). Progress and problems in reforming public language examinations in Europe: Cameos from the Baltic States, Greece, Hungary, Poland, Slovenia, France, and Germany.

Background to the Examination

Overall aim: Assessment of language proficiency within an academic context at an intermediate to advanced level. Candidates have to show that they can understand and produce written and spoken texts that are relevant for everyday life and for studies at institutions of Higher Education. They have to manage administrative affairs and to organise their studies at a German university. As a proficiency test the TestDaF evaluates the candidate's language performance separately in four skills at three levels: TestDaF-Niveaustufe 3 (TDN 3), 4 (TDN 4) and 5 (TDN 5). These levels are related to the levels of the Common European Framework B2 and C1. During the project period (1998-2000) the TestDaF levels were linked to the framework of the Association of Language Testers in Europe (ALTE) by using anchor-items calibrated to the ALTE levels.

Structure of the TestDaF

Subtests	Structure	Content
 Reading	3 parts, 30 items 60 min	• comprehension of global meaning as well as of specific information • complex information processing, e.g. decoding of implicit meaning or underlying concepts
 Listening	3 parts, 25 items 40 min	
 Writing	1 item 60 min	• describing graphs and tables • commenting on controversial questions, making reference to different positions and presenting a personal point of view
 Speaking	10 items 30 min	• making requests • asking for, giving or obtaining information • verbalising information given in a graph, diagram or table • presenting and disputing controversial issues
Total time:	3 h 10 min + break (1 h)	

TestDaF Reading Comprehension

The subtest of Reading Comprehension consists of three parts. Part one shows several short texts of brochures, programs or university calendars offering information for orientation (350 words on the whole, matching task). The candidates' ability to read several texts for specific information is assessed. In part two a 450-550 word journalistic text about a socio-political or scientific topic is used, in conjunction with a multiple choice task. The candidates should be able to understand the main ideas and detailed information. Part three provides an academic text with complex structures and content about a scientific problem or development (500-600 words, selected response task: yes/no/not in the text). The text is written for non-experts in the subject. Candidates are assessed on their ability to recognise point of views from the text, to recognise the gist of a paragraph and to understand implicit information.

Test to be analysed	TestDaF, Reading Comprehension, part 3
Target Language	German
Task	Part 3 – questions 21-30
Rubrics	Target Language
Items	Target Language
Time to do total subtest	60 minutes
Target Level	CEFR C1
Age-group sector	Adults or university students
Domain:	Educational domain
Communicative activities	
Reading for information and argument	Understanding a university-level text about a scientific problem or development.

Reading Items

TestDaF Test Deutsch als Fremdsprache

Leseverstehen

Lesetext 3: Aufgaben 21 – 30

ca. 20 min

Lesen Sie den Text und lösen Sie die Aufgaben.

Die Sammelleidenschaft beim Menschen

Evolutionsbiologen beschäftigen sich seit geraumer Zeit mit der Sammelleidenschaft des Menschen. Der Mensch sammelt die unterschiedlichsten Gegenstände, unabhängig von ihrem materiellen Wert. Selbst Schlüsselanhänger, Gesteinsproben oder Murmeln* können den Sammeldrang wecken. Keineswegs stellt das Sammeln lediglich das Anhäufen nutzloser Gegenstände zum Zeitvertreib dar, wie gemeinhin angenommen wird. Vielmehr kommt dem Sammeln in vielerlei Hinsicht eine wichtige Rolle zu. Wie kommt also diese Sammelleidenschaft zustande, und welche Funktion hat sie?

Die Sammelleidenschaft kann am stärksten in der Kindheit konstatiert werden, in einem Alter also, in dem Nützlichkeitserwägungen noch keine Rolle spielen können. Kinder beginnen in der Regel schon früh, oft unvermittelt und intensiv mit dem Sammeln. Das Sammeln bestimmter Gegenstände muss also schon in der Kindheit etwas außerordentlich Wichtiges sein. Es ließ sich beobachten, dass Kinder beim Sammeln sogar bereit sind, ihre guten sozialen Beziehungen zu gefährden. Das unterscheidet das Sammeln vom Spielen: Ein Spiel wird zwar sehr intensiv betrieben, wird aber auch rasch wieder beendet, während das Sammeln oft über Monate und Jahre, nicht selten ein ganzes Leben lang, anhält.

Offensichtlich handelt es sich hier um ein Charakteristikum des Menschen. Beobachtungen in der Tierwelt, insbesondere bei Menschenaffen, haben gezeigt, dass bei Tieren kein Sammelverhalten festzustellen ist. Doch warum sollte ausgerechnet die scheinbar so nutzlose Eigenart des Sammelns in besonderer Weise menschentypisch sein? Und wie konnte dieses Verhalten überhaupt entstehen, wenn es doch so offensichtlich ohne Nutzwert ist? Für die Evolutionsbiologen sind solche Fragen eine große Herausforderung.

In einer bestimmten Hinsicht befindet sich das Sammeln in bester Übereinstimmung mit der lebendigen Natur: Beide charakterisiert die Vielfalt. Sammlungen bilden ausnahmslos Vielfalt ab. Ohne Vielfalt keine Sammlung. Die bloße Anhäufung von Gleichartigem reizt einen echten Sammler nicht. In der Vielfältigkeit könnte nun der Schlüssel zum Ursprung des Sammelns zu suchen sein: Sammeln ist das Organisieren von Vielfalt, eine Fertigkeit, der für die Menschen entscheidende Bedeutung zukommt und die vielleicht deshalb schon von Kindheit an trainiert wird.

Das menschliche Gehirn entwickelt und strukturiert sich über das Sammeln von Daten, wie wir es im Computerzeitalter nennen würden. Wie bei einem Computer bleiben Programme nutzlos, wenn ihnen nicht bestimmte Daten zur Verfügung gestellt werden, um jene Verknüpfungen zu ermöglichen, die schließlich zu den gewünschten Ergebnissen führen. Auch unser Gehirn sammelt zunächst unsystematisch Daten. Wie ein Schwamm saugt es Unmengen davon auf, weit mehr, als uns bewusst wird. Die Hauptaufgabe des Gehirns besteht darin, diesen Datenstrom der Sinne zu kanalisieren, zu regulieren und zunehmend zu ordnen. Dabei wird das Brauchbare, Vernünftige und Wichtige vom Datenüberfluss befreit. So übt sich schon das junge Gehirn in Systematik und lernt, auszuwählen und zuzuteilen. Genau dies wird beim Sammeln an konkreten Objekten praktiziert und trainiert. Das Gehirn ordnet, stellt Unterschiede fest und Übereinstimmungen her, bildet Hierarchien und Klassen. Die Sammlung ist kein Chaos, sondern Ordnung, die wächst und gedeiht. Sie trainiert die Speicherkapazitäten des Gehirns, schafft spezielle Kenntnisse und Erinnerungen. Ganz zutreffend sprechen wir von „Erinnerungsstücken", wenn wir etwas mitgenommen und angesammelt haben.

Die Sammelleidenschaft entwickelte sich in einem Millionen von Jahren währenden Prozess der Menschwerdung. Aller Wahrscheinlichkeit nach stellt sie eine überlebensnotwendige Anpassung dar, die mit der Entwicklung der sozialen und geistigen Fähigkeiten einherging. Der Mensch wurde zum Datensammler, weil sein Gehirn größer und leistungsfähiger wurde. Darin steckt die Grundlage für sein hochdifferenziertes Sozialverhalten und für all die gesellschaftlichen und wissenschaftlichen Leistungen. Ohne die Fähigkeit und Bereitschaft, Daten zu sammeln, aufzubereiten, sie zu bewahren, um sie weiterzugeben, wären alle Kenntnisse und Errungenschaften der Menschheit allenfalls zufällig entstanden und rasch wieder verloren gegangen.

*Murmel = kleine bunte Glaskugel

Nach: J.H. Reichholf: „Die Sammelleidenschaft beim Menschen." In: Psychologie Heute, 04/ 2001

Lesetext 3: Aufgaben 21 – 30

Markieren Sie die richtige Antwort.

		Ja	Nein	Text sagt dazu nichts	
(01)	Viele Sammler interessieren sich besonders für moderne Kunst.			**X**	(01)
(02)	Menschen sammeln Gegenstände, weil sie sich langweilen.		**X**		(02)
21	Kinder denken nicht an die Verwertbarkeit ihrer Sammelobjekte.				21
22	Kinder fühlen sich sozial akzeptiert, wenn sie wichtige Gegenstände ansammeln.				22
23	Kinder riskieren beim Sammeln Streit untereinander.				23
24	Die Evolutionsbiologen fanden nur bei manchen Affenarten einen Drang zum Sammeln.				24
25	Evolutionsbiologen interessiert in erster Linie der Nutzwert menschlichen Verhaltens.				25
26	Sammlungen zeichnen sich durch ein breites Spektrum aus.				26
27	Ein Charakteristikum des menschlichen Gehirns ist es, Wichtiges von Unwichtigem unterscheiden zu können.				27
28	Die Fähigkeit, gesammelte Gegenstände zu kategorisieren, nimmt mit zunehmendem Alter ab.				28
29	Sammeln ist eine Übung für das Gedächtnis.				29
30	Die Fähigkeit, große Mengen von Informationen zu verarbeiten führte zur Entwicklung spezifisch menschlicher Kompetenzen.				30

Content analysis

Reading component

The subtest of Reading Comprehension consists of three parts. Part one shows several short texts of brochures, programs or university calendars offering information for orientation (350 words on the whole, matching task). The candidates' ability to read several texts for specific information is assessed. In part two a 450-550 word journalistic text about a socio-political or scientific topic is used, in conjunction with a multiple choice task. The candidates should be able to understand the main ideas and detailed information. Part three provides an academic text with complex structures and content about a scientific problem or development (500-600 words, selected response task: yes/no/not in the text). The text is written for non-experts in the subject. Candidates are assessed on their ability to recognise point of views from the text, to recognise the gist of a paragraph and to understand implicit information.

Text Characteristics:

(Analysis of example tasks and items has been carried out in conjunction with the Dutch CEF Grid, the CEFR, the ALTE can-do statements and the DIALANG performance descriptors.)

Test to be analysed:	TestDaF (Paper)
Task:	Part 3 – question 21-30, Selected sample questions: 22, 23, 25
Skill:	Reading
Rubric in L1/Target Language:	TL
Target Language:	German
Item in L1/Target Language:	TL
Time to do total task:	20 m (total subtest 60 m)
1. Text Source:	authentic text, journal article
2. Authenticity:	adapted for non-experts
3. Discourse type:	expository text
4. Discourse subtype:	explications example: broader accounts of abstract phenomena, e.g. lectures, talks
5. Domain:	educational
6. Topic:	scientific
7. Nature of content:	mostly concrete
8. Text length:	576 words
9. Vocabulary:	rather extended
10. Grammar:	frequent compound sentences
Comprehensible by learner at CEF level	C 1

Item Characteristics:

Item 22

Item type	Selected response, three options (Key option = not in the text)
Operations involved in answering:	Evaluating explicit information in the text concerning the main idea.
	Intensive reading and detailed comprehension.
Item level estimated	C1

Item 23

Item type	Selected response, three options (Key option = yes)
Operations involved in answering:	Evaluating implicit information in text concerning the main idea.
	Intensive reading and detailed comprehension.
Item level estimated	C1

Item 25

Item type	Selected response, three options (Key option = no)
Operations involved in answering:	Inferring from information implicit in the text. Drawing a conclusion considering several sentences.
	Intensive reading and detailed comprehension.
Item level estimated	C1

Statistical analysis

Statistical Report (Whole Task)

mean facility (p)	**.58**
mean discrimination (Pb)	**.30**
mean difficulty (Rasch)	**3.95 Logits**

Statistical Report (Individual Items)

	item 22	item 23	item 25
Facility (p)	.40	.35	.34
Discrimination (i.d.)	.48	.46	.21
Difficulty (Rasch)	1.38 SE = .18	1.64 SE = .18	1.71 SE = .18
Sample size	168		

TestDaF Listening Compehension

The subtest of Listening Comprehension consists of three parts. In part one the candidate listens to a dialogue in a situation of everyday life at the university (short answer questions). The candidates' ability to understand and select information when listening to another speaker is assessed. In part two a radio interview with 3 or 4 participants about topics concerning higher education in general is played (true/false task). The candidates should be able to understand the main ideas and detailed information. Part three provides a short lecture or interview with an expert or university professor about a scientific problem or development (short answer questions). The text is adapted for non-experts in the subject. Candidates are assessed on their ability to recognise point of views from the text, to recognise the gist of a paragraph and to understand implicit information.

Test to be analysed	TestDaF, Listening Comprehension, part 3
Target Language	German
Task	Part 3 – questions 19-25
Rubrics	Target Language
Items	Target Language
Time to do total subtest	40 minutes
Target Level	CEFR C1
Age-group sector	Adults or university students
Domain:	Educational domain
Communicative activities	
Listening for information and argument	Following and understanding lectures with scientific content. Understanding meanings, point of views.
Listening as a member of a life audience	Taking notes while listening to a lecture.

Listening Items

TestDaF Test Deutsch als Fremdsprache

Hörverstehen

Hörtext 3: Aufgaben 19 – 25

Sie hören einen kurzen Vortrag von Professor Kaesler über Jugendliche und ihre Moral.
Sie hören diesen Vortrag **zweimal.**

Lesen Sie jetzt die Aufgaben 19 – 25.

Hören Sie nun den Text ein erstes Mal.
Beantworten Sie beim Hören die Fragen 19 – 25 in Stichworten.

Über Jugendliche und ihre Moral

(0)	Was beschreibt Prof. Kaesler in seinem Beispiel?	(0)	*den Verkauf eines Mopeds*
19	Was ist das Ergebnis der Umfrage?	19	
20	Wovon ist die Soziologin Nunner-Winkler überzeugt?	20	
21	Wie wird das Verhalten moralisch engagierter Jugendlicher im Berufsleben bewertet?	21	
22	Zu welchem Forschungsergebnis kamen Soziologen an der Universität Dortmund?	22	
23	Welche Sinnfrage stellen sich Jugendliche?	23	
24	Welchen Maßstab setzt Prof. Kaesler für die Qualität einer guten Gesellschaft?	24	
25	Was möchte die Soziologie wieder erreichen?	25	

Content analysis

Listening component

The subtest of Listening Comprehension consists of three parts. In part one the candidate listens to a dialogue in a situation of everyday life at the university (short answer questions). The candidates' ability to understand and select information when listening to another speaker is assessed. In part two a radio interview with 3 or 4 participants about topics concerning higher education in general is played (true/false task). The candidates should be able to understand the main ideas and detailed information. Part three provides a short lecture or interview with an expert or university professor about a scientific problem or development (short answer questions). The text is adapted for non-experts in the subject. Candidates are assessed on their ability to recognise point of views from the text, to recognise the gist of a paragraph and to understand implicit information.

Text Characteristics:

(Analysis of example tasks and items has been carried out in conjunction with the Dutch CEF Grid, the CEFR, the ALTE can-do statements and the DIALANG performance descriptors.)

Test to be analysed:	TestDaF (Paper)
Task:	Part 3 – questions 19-25 Selected sample questions: 19, 21, 24
Skill:	Listening
Rubric in L1/Target Language:	TL
Target Language:	German
Item in L1/Target Language:	TL
Time to do total task:	10-12 minutes
1. Text Source:	authentic interview or lecture
2. Authenticity:	adapted for non-experts, audio recorded
3. Discourse type:	expository text
4. Discourse subtype:	explications example: broader accounts of abstract phenomena, e.g. lectures, talks
5. Domain:	educational
6. Topic:	scientific
7. Nature of content:	Concrete introduction but then fairly extensive abstract content
8. Text length:	300 seconds
9. Vocabulary:	rather extended
10. Grammar:	frequent compound sentences
11. Text speed:	normal
12. Number of participants in the text:	One
13. Accent/standard	standard pronunciation
14. Clarity of articulation	clearly articulated
15. How often played	played twice
Comprehensible by learner at CEF level	C 1

Item Characteristics:

Item 19

Item type	Constructed response, short answer questions
Operations involved in answering:	Evaluating explicit information in the text concerning the main idea.
Item level estimated	C1

Item 21

Item type	Constructed response, short answer questions
Operations involved in answering:	Evaluating explicit information in the text.
Item level estimated	C1

Item 24

Item type	Constructed response, short answer questions
Operations involved in answering:	Inferring explicit information in the text connecting parts of the text.
Item level estimated	C1

Statistical analysis

Statistical Report (Whole Task)

mean facility (p)	**.41**
mean discrimination (Pb)	**.44**
mean difficulty (Rasch)	**0.93**

Statistical Report (Individual Items)

	item 19	item 21	item 24
Facility (p)	.32	.24	.33
Discrimination (i.d.)	.50	.32	.35
Difficulty (Rasch)	1.41 SE = .17	1.94 SE = .18	1.39 SE = .17
Sample size	198		