

2020–22

Jahresbericht *Annual Report*



Inhaltsverzeichnis

Contents

| | |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Vorwort <i>Preface</i> | 6 |
| 1 Editorial <i>Editorial</i> | 8 |
| 2 Angebote der Gesellschaft für Akademische Studienvorbereitung und Testentwicklung e.V. <i>Product range of the Society for Academic Study Preparation and Test Development e.V.</i> | 11 |
| 2.1 TestDaF | 11 |
| 2.1.1 Konzeption und Einführung des digitalen TestDaF <i>Concept and introduction of the digital TestDaF</i> | 11 |
| 2.1.2 Digitaler TestDaF: Konzept zur Prüfungsvorbereitung <i>Digital TestDaF: Concept for exam preparation</i> | 15 |
| 2.1.3 Konzeption und Durchführung des papierbasierten TestDaF <i>Concept and implementation of the paper-based TestDaF</i> | 16 |
| 2.1.4 Teilnehmendenentwicklung <i>Development of test taker numbers</i> | 16 |
| 2.1.5 Qualitätsmanagement: Zertifizierung von Beurteilenden <i>Quality control: certification of assessors</i> | 17 |
| 2.2 TestAS | 18 |
| 2.2.1 Konzept, Aufbau und Einführung des digitalen TestAS <i>Concept, structure and introduction of the digital TestAS</i> | 19 |
| 2.2.2 Konzept und Aufbau des papierbasierten TestAS <i>Concept and structure of the paper-based TestAS</i> | 20 |
| 2.2.3 Teilnehmendenentwicklung des papierbasierten TestAS <i>Development of test taker numbers and test dates of the paper-based TestAS</i> | 21 |
| 2.2.4 TestAS in Hochschulzugangsverfahren <i>TestAS in university admission procedures</i> | 22 |
| 2.2.5 Digitaler Mastertest (dMAT) <i>Digital master test (dMAT)</i> | 22 |
| 2.3 onSET | 24 |
| 2.3.1 Konzept und Aufbau <i>Concept and structure</i> | 24 |
| 2.3.2 Teilnehmendenentwicklung <i>Development of test taker numbers</i> | 25 |
| 2.3.3 onSET-Erweiterungen <i>onSET extensions</i> | 25 |
| 2.4 Deutsch-Uni Online | 27 |
| 2.4.1 Konzept und Aufbau <i>Concept and structure</i> | 27 |

© g.a.s.t. 2024

Herausgeber

Publisher

Dr. Jörn Weingärtner

Gesellschaft für Akademische Studienvorbereitung und Testentwicklung e.V. (g.a.s.t.)

c/o TestDaF-Institut, Universitätsstr. 134, 44799 Bochum

Fotos

Photos

Adobe Stock, teilweise generiert mit KI

Adobe Stock, some generated with KI

Beiträge

Articles

Mitarbeiterinnen und Mitarbeiter von g.a.s.t.

Staff members of g.a.s.t.

| | | |
|------------|------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| 2.4.2 | Teilnehmendenentwicklung <i>Development of user numbers</i> | 29 |
| 2.4.3 | DUO vor Ort – weltweit <i>DUO vor Ort – weltweit</i> | 30 |
| 3 | Forschung und Entwicklung <i>Research and development</i> | 33 |
| 3.1 | Bildungs- und Digitalisierungsprojekte <i>Education and Digitalisation projects</i> | 33 |
| 3.1.1 | Digitaler Campus <i>Digitaler Campus</i> | 33 |
| 3.1.2 | Bildungsraum Digital (BIRD) <i>Digital education space (BIRD)</i> | 34 |
| 3.1.3 | Profis D – Fach.Deutsch.Digital für Studium und Beruf: ein Verbundprojekt <i>Profis D – Fach.Deutsch.Digital für Studium und Beruf: a joint project</i> | 34 |
| 3.2 | TestDaF-Forschung <i>TestDaF research</i> | 37 |
| 3.2.1 | Online-Standard-Setting beim digitalen TestDaF <i>Online standard setting for the digital TestDaF</i> | 37 |
| 3.2.2 | Konstrukt integrierter Schreibaufgaben <i>Construct of integrated writing tasks</i> | 39 |
| 3.2.3 | Teilnehmendenbefragung <i>Participant survey</i> | 40 |
| 3.2.4 | Evaluation der Vorbereitungsunterlagen für den digitalen TestDaF <i>Evaluation of the preparation materials for the digital TestDaF</i> | 41 |
| 3.2.5 | Simultane Messung von Strenge- und Zentraltendenzen <i>Simultaneous measurement of severity and central tendencies</i> | 42 |
| 3.2.6 | Statistische Indikatoren der Zentraltendenz <i>Statistical indicators of central tendency</i> | 43 |
| 3.2.7 | Analyse von differenziellen Beurteilungsfunktionen <i>Analysis of differential judgement functions</i> | 43 |
| 3.3 | TestAS-Forschung und -Entwicklung <i>TestAS research and development</i> | 45 |
| 3.3.1 | Vorhersage von Studienerfolg internationaler Studierender (LIMES) <i>Predicting the academic success of international students (LIMES)</i> | 45 |
| 3.4 | onSET-Forschung und -Entwicklung <i>onSET research and development</i> | 48 |
| 3.4.1 | Mobiler onSET <i>Mobile onSET</i> | 48 |
| 3.4.2 | Faktor Zeit beim C-Test <i>Time factor in the C-test</i> | 49 |
| 3.5 | DUO-Lernplattform <i>DUO learning platform</i> | 51 |
| 3.6 | Forschungskooperationen <i>Research cooperation</i> | 54 |
| 3.6.1 | Herder-Institut (Universität Leipzig) <i>Herder Institute (University of Leipzig)</i> | 54 |
| 3.6.2 | Sprachtechnologie-Labor (Universität Duisburg-Essen) <i>Language Technology Laboratory (University of Duisburg-Essen)</i> | 55 |

| | | |
|----------|------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| 4 | g.a.s.t.-Akademie <i>g.a.s.t. Akademie</i> | 57 |
| 4.1 | Seminar für Prüfungsbeauftragte <i>Seminar for examination officers</i> | 57 |
| 4.2 | Workshops und Webinare zum Testen, Evaluieren und E-Learning <i>Workshops and webinars for Testing, evaluation and e-learning</i> | 57 |
| 4.3 | Seminare nach Maß im In- und Ausland <i>Customised workshops in Germany and abroad</i> | 58 |
| 5 | Aufträge <i>Third-party orders</i> | 61 |
| 5.1 | Institut National des Langues (INL) <i>Institut National des Langues (INL)</i> | 61 |
| 5.2 | Italienische Bildungsdirektion der Autonomen Provinz Bozen <i>Italian Education Directorate of the Autonomous Province of Bolzano</i> | 62 |
| 5.3 | Zentralstelle für das Auslandsschulwesen (ZfA) <i>Zentralstelle für das Auslandsschulwesen (ZfA)</i> | 62 |
| 5.4 | Bundesagentur für Arbeit <i>Federal Employment Agency</i> | 63 |
| 6 | Weitere Aktivitäten von g.a.s.t. e. V. <i>Further activities of g.a.s.t. e. V.</i> | 65 |
| 6.1 | Gremiensitzungen <i>Committee meetings</i> | 65 |
| 6.2 | Mitgliedschaften <i>Memberships</i> | 65 |
| 6.3 | Beirats- und Gremientätigkeit <i>Advisory and committee work</i> | 65 |
| 6.4 | Gutachtertätigkeit <i>Reviewing activities</i> | 66 |
| 6.5 | Bildungsmessen und Fachtagungen <i>Education fairs and conferences</i> | 66 |
| 6.6 | g.a.s.t. auf der IDT in Wien, Motto: *mit.sprache.teil.haben <i>g.a.s.t. at the IDT in Vienna, slogan: *mit.sprache.teil.haben</i> | 67 |
| 7 | Wirtschaftliche Entwicklung <i>Financial review</i> | 69 |
| 8 | Vorträge und Publikationen <i>Lectures and publications</i> | 71 |
| 8.1 | Vorträge und Präsentationen (Auswahl) <i>Lectures and presentations (selection)</i> | 71 |
| 8.2 | Publikationen <i>Publications</i> | 73 |



Dr. Michael Harms
Vorsitzender des Vorstands
der Gesellschaft für
Akademische Studienvorbereitung
und Testentwicklung (g.a.s.t.) e.V.

Stellvertretender Generalsekretär
des Deutschen Akademischen
Austauschdienstes (DAAD)

Dr Michael Harms
*Chairman of the Board of the Society
for Academic Study Preparation and
Test Development (g.a.s.t)*

*Deputy Secretary General of the German
Academic Exchange Service (DAAD)*

Vorwort

Wandel und Entwicklung, zumeist verbunden mit stetem Wachstum, sind seit der Gründung des Vereins prägend für die Arbeit von g.a.s.t. Lange Zeit konnten diese aktiv in eigener Initiative gestaltet werden. Wie in vielen anderen Bereichen markierte hier der Jahresbeginn 2020 eine deutliche Zäsur. Der Ausbruch der Pandemie brachte den internationalen Austausch zu einem weitgehenden Stillstand. Direkt davon betroffen war auch die Arbeit von g.a.s.t. Statt aktiver Gestaltung der Zukunft, beispielhaft durch Digitalisierung, dem Leitbegriff des letzten Jahresberichts, musste nun in einen permanenten Krisenreaktionsmodus gewechselt werden. Dies galt umso mehr, als im vergangenen Jahr zusätzlich zu den anhaltenden Folgen der Pandemie Russlands Angriff auf die Ukraine hinzukam, und damit zwei Länder betraf, die traditionell zu den teilnahmestärksten Herkunftsländern zählten. Die Teilnehmendenzahlen in allen Testprodukten und der DUO sind im Jahr 2020 deutlich gefallen und haben sich seitdem nur auf niedrigem Niveau konsolidiert. Die Spielräume von g.a.s.t. sind so sukzessive kleiner geworden. Umso wichtiger wurden in den letzten drei Jahren Projekte wie etwa Programme für Geflüchtete aus der Ukraine sowie der Bildungsraum Digital (BIRD) und der Digitale Campus, in denen g.a.s.t. sowohl inhaltlich wie auch im Bereich technischer Innovation einen bedeutenden Beitrag leistet. Es zeichnet g.a.s.t. aus, dass selbst in diesen Zeiten die Forschungstätigkeit nicht gelitten hat. Die enge Verknüpfung von Theorie und Praxis ist und bleibt ein fester Baustein der DNA des Hauses.

Es bedurfte großer Anstrengungen und eines hohen Maßes an Mut, Reißfestigkeit und Kreativität sowie nicht zuletzt der Unterstützung wichtiger Partner und Förderer, um diesen Herausforderungen erfolgreich zu begegnen. Besonderer Dank sei hier dem Bundesministerium für Bildung und Forschung und dem Auswärtigen Amt ausgesprochen. Die Krisen in der Welt sind leider keineswegs beendet oder bewältigt. Dennoch gibt es zum Ende des Jahres 2022 für g.a.s.t. erste Anzeichen der Hoffnung auf eine Trendwende zum Besseren. Es zeugt von hoher Resilienz des gesamten Hauses und großem unternehmerischem Geist, dass g.a.s.t. die Herausforderungen der Zeit angenommen hat und aus eigener Kraft Wege durch die Krise und aus der Krise fand.

Nicht nur weltpolitisch wurde eine Zeitenwende eingeläutet. Auch für g.a.s.t. als Organisation war der Zeitraum dieses Berichts von großen Veränderungen geprägt. In kurzer Folge sind mit Dr. Hans-Joachim Althaus, dem Gründungsgeschäftsführer von g.a.s.t., sowie Dr. Gabriele Kecker und Dr. Thomas Eckes, den beiden langjährigen Stellvertretenden Institutsleiter*innen, die drei Persönlichkeiten, die g.a.s.t. aufgebaut, entwickelt und in den ersten mehr als 20 Jahren entscheidend geprägt haben, in den Ruhestand getreten. Kein Dank ist groß genug, um die Leistung und das Verdienst um den Verein, das TestDaF-Institut, die Etablierung von standardisierten Testverfahren und nicht zuletzt um den internationalen Austausch und den Studienstandort Deutschland gebührend zu fassen. Wir freuen uns, dass alle drei g.a.s.t. weiterhin verbunden bleiben. Auch der Vorsitz des Vorstands wurde im Jahr 2021 übergeben. Ich danke meinem Vorgänger Christian Müller im Namen des Vorstands für sein Engagement und seinen Einsatz für g.a.s.t.

Ich wünsche Ihnen viel Freude und Gewinn bei der Lektüre,

Preface

Change and development, usually combined with steady growth, have characterised the work of g.a.s.t. since the organisation was founded. For a long time, g.a.s.t. was privileged to shape its future and development on its own initiative. As in many other areas, the beginning of 2020 marked a significant turning point. The outbreak of the pandemic brought international exchange virtually to a standstill. Unsurprisingly, the work of g.a.s.t. was no exception. Instead of actively shaping the future, for example through digitalisation, the major theme of the last annual report, we now had to switch to a permanent crisis response mode. This was all the more true last year, when the effects of Russia's attack on Ukraine drastically increased the challenges already posed by the pandemic. The war in Ukraine affects two countries that have traditionally been among the countries of origin with the highest number of participants. The number of participants in all test products and the DUO fell significantly in 2020 and has only consolidated at a low level since then. As a result, g.a.s.t.'s room for manoeuvre was markedly reduced. As a result, projects such as programmes for refugees from Ukraine, the Digital Education Space (BIRD) and the Digital Campus, in which g.a.s.t. makes a significant contribution both in terms of content and technical innovation, have grown in importance during the last three years. Despite these challenges, it is remarkable that even in these times, research activities have not suffered. The close link between theory and practice is and remains an integral part of the organisation's DNA.

It took great effort and a high degree of courage, resilience and creativity, not to mention the support of important partners and sponsors, to successfully meet these challenges. Special thanks are due to the Federal Ministry of Education and Research and the Federal Foreign Office. Unfortunately, the crises in the world are by no means over or overcome. Nevertheless, at the end of 2022, g.a.s.t. is seeing the first signs of hope for a turnaround. The fact that g.a.s.t. has accepted the challenges of the times and found its own way through and out of the crisis is testament to the resilience of the entire organisation and its entrepreneurial spirit.

The past years were not only a watershed in global politics. For g.a.s.t. as an organisation, the period covered by this report was also marked by major changes. Dr. Hans-Joachim Althaus, the founding Managing Director of g.a.s.t., as well as Dr. Gabriele Kecker and Dr. Thomas Eckes, the two long-standing Deputy Directors of the Institute, retired in quick succession – three personalities who built up, developed and shaped g.a.s.t. in its first 20-plus years. No words are sufficient to express gratitude and respect for their achievements and merits for the society, the TestDaF Institute, the establishment of standardised test procedures and, last but not least, for international exchange and Germany as a place to study. We are delighted that all three will remain associated with g.a.s.t. The chairmanship of the Executive Board was also handed over in 2021. On behalf of the Executive Board, I would like to thank my predecessor Christian Müller for his commitment and dedication to g.a.s.t.

I hope you enjoy and benefit from reading this report,



Dr. Jörn Weingärtner
Leiter des TestDaF-Instituts

Geschäftsführer der Gesellschaft für
Akademische Studienvorbereitung
und Testentwicklung e.V.

Dr. Jörn Weingärtner
Director of the TestDaF-Institut

Chief Executive of the Society for
Academic Study Preparation and
Test Development (g.a.s.t.)

1 Editorial

Liebe Leserinnen und Leser,

Krise – kein anderes Wort kennzeichnet und prägt den Zeitraum dieses Berichts markanter als dieser. Seit Beginn des Jahres 2020 ist die Welt im Dauerkrisenmodus. Auch an g.a.s.t. gehen diese Herausforderungen nicht vorbei. Nach Jahren stetigen Wachstums in den Teilnehmendenzahlen, vor allem des TestDaF, brach die Nachfrage nach Testverfahren quasi über Nacht ab 2020 um bis zu 50% ein. Es ist dabei noch nicht absehbar, ob es sich um eine vorübergehende Zäsur handelt und es ein dem „Davor“ ähnliches „Danach“ geben oder ob „Davor“ und „Seitdem“ das Begriffspaar der Zukunft sein wird. Insofern ist dieser Jahresbericht vor allem ein Krisenbericht, was sich auch daran ablesen lässt, dass wir dieses Mal gleich drei Jahre in einem Rückblick zusammenfassen.

Wie in den Vorjahren richtet sich dieser Bericht an einen breiten Adressat*innenkreis mit möglicherweise ebenso breiten wie unterschiedlichen Informationsinteressen. Wir legen Rechenschaft über die Arbeit in den letzten drei Jahren ab, möchten aber auch über neue Entwicklungen informieren. Dazu gehören die Einführung des digitalen TestAS, neue Forschungsprojekte, etwa zur Rolle der Zeit in Sprachtestverfahren, und technische Innovationen, zum Beispiel für den Digitalen Campus und die Nationale Bildungsplattform.

Da die Nachfrage im Kerngeschäft des Testens und digitalen Lehrens und Lernens deutlich zurückgegangen ist, wurden Projekte im Berichtszeitraum wichtiger, um den Erhalt des Vereins und die Gestaltungsmöglichkeiten zu sichern. Besonderer Dank für die Einrichtung und Finanzierung von Projekten und das Vertrauen in g.a.s.t., die Projekte entweder selbst zu realisieren oder zu ihnen beizutragen, kommt hier dem Auswärtigen Amt, dem BMBF und dem DAAD zu. Gleichzeitig erwies sich gerade in der Krise die Wichtigkeit grenzübergreifender Kooperation und internationalen Austauschs in ganz praktischer Weise, also die Werte, die, bezogen auf den akademischen Bereich, zur Gründung von g.a.s.t. führten. Während die ganze Welt von der Pandemie und ihren Folgen betroffen war, brachte der Angriff Russlands auf die Ukraine dort ganz anderes Leid und Herausforderungen mit sich. Der Umgang der ukrainischen Bevölkerung und der Testzentren in der Ukraine lehrte Demut in der Krise und rückte die Perspektive auf die eigenen Herausforderungen zurecht. Es erfüllt mich mit größtem Respekt, dass Testzentren in der Ukraine selbst unter Kriegsbedingungen Testverfahren durchgeführt haben, zum Teil in Schutzkellern. Auch dies ist leider derzeit Teil der neuen Realität.

Da ich erst zum Ende des Berichtszeitraums Teil von g.a.s.t. wurde, möchte ich die Gelegenheit nutzen, meinen persönlichen Dank und Respekt auszusprechen. An allererster Stelle möchte ich mich bei meinem Vorgänger, Dr. Hans-Joachim Althaus, für den Übergang und die großzügige Übergabe bedanken. In den ersten mehr als 20 Jahren von g.a.s.t. wurden hohe Maßstäbe gesetzt, denen ich mich uneingeschränkt verpflichtet fühle. Ebenso möchte ich mich bei allen bedanken, die g.a.s.t. in einer schwierigen Zeit begleitet und unterstützt haben: bei allen Testzentren, die allen Anfechtungen zum Trotz Tests in hoher Prozessqualität durchgeführt haben, und allen Förderern und Kooperationspartnern für Treue und manchmal auch Geduld mit einer g.a.s.t. im Krisenmodus in den letzten drei Jahren.

Ich wünsche eine anregende Lektüre,

1 Editorial

Dear readers,

Crisis – no other word characterises and shapes the period covered by this report more conclusively than this one. The world has been in permanent crisis mode since the beginning of 2020. And g.a.s.t. is not immune to these challenges. After years of steady growth in number of participants, especially for the TestDaF, demand for testing procedures plummeted by up to 50% virtually overnight in 2020. It is not yet clear whether this is a temporary break and whether there will be an “after” similar to the “before” or whether “before” and “ever since” will be the terms defining the future. In this respect, this annual report is above all a crisis report, which can also be seen from the fact that this time we are summarising three years in one review.

As in previous years, this report is addressing a broad audience with potentially equally broad and diverse demands of information. We give an account of our work over the past three years, but also wish to inform you about new developments. These include the introduction of the digital TestAS, new research projects, for example on the role of time in language testing procedures, and technical innovations, for instance for the Digital Campus and the National Education Platform.

As demand in the core business of testing and digital teaching and learning has significantly declined, projects became more important in the reporting period in order to ensure the organisation’s continued existence and its ability to shape the future. Special thanks are due to the Federal Foreign Office, the Federal Ministry of Education and Research (BMBF) and the German Academic Exchange Service (DAAD) for setting up and funding projects and entrusting g.a.s.t. with their realisation or a contribution towards their realisation. At the same time, the crisis has demonstrated the importance of cross-border cooperation and international exchange, i.e. the values that led to the founding of g.a.s.t. While the whole world was affected by the pandemic and its consequences, Russia’s attack on Ukraine brought with it completely different hardship and challenges. The way the Ukrainian population and the test centres in Ukraine dealt with the crisis taught us humility and put our own challenges into perspective. It fills me with the greatest respect that test centres in Ukraine have carried out test procedures even under wartime conditions, sometimes in underground shelters. Unfortunately, this too is currently part of the new reality.

As I became part of g.a.s.t. only at the end of the reporting period, I would like to take this opportunity to express my personal thanks and respect. First and foremost, I would like to thank my predecessor, Dr. Hans-Joachim Althaus, for the transition and the generous handover. In the first 20-plus years of g.a.s.t., high standards were set to which I feel fully committed. I would also like to thank all those who have accompanied and supported g.a.s.t. during a difficult time: all the test centres that have carried out tests of highest process quality despite all the challenges, and all the sponsors and partners for their loyalty and sometimes patience with a g.a.s.t. in crisis mode over the last three years.

I wish you an inspiring read,



2 Angebote der Gesellschaft für Akademische Studienvorbereitung und Testentwicklung e.V.

2 Product range of the Society for Academic Study Preparation and Test Development e.V.

2.1 TestDaF

2.1.1 Konzeption und Einführung des digitalen TestDaF

Der digitale TestDaF wurde erstmals im Oktober 2020 in lizenzierten Testzentren weltweit durchgeführt. Die Prüfung wird internetgestützt an Laptops oder PCs durchgeführt und erfordert eine geeignete technische Ausstattung und Testumgebung (Internetzugang, Datenbandbreite, Upload- und Download-Geschwindigkeit).

Der digitale TestDaF enthält vier Prüfungsteile, die den Teilkompetenzen Lesen, Hören, Schreiben und Sprechen zugeordnet sind. Er besteht aus insgesamt 23 eigens entwickelten Aufgabentypen. Die Abfolge der Prüfungsteile und die Bearbeitungszeit für die einzelnen Aufgaben sind standardisiert und verbindlich festgelegt.

2.1 TestDaF

2.1.1 Concept and introduction of the digital TestDaF

The digital TestDaF was conducted for the first time in October 2020 in licensed test centres worldwide. The test is conducted online on laptops or PCs and requires suitable technical equipment and test environment (internet access, data bandwidth, upload and download speed).

The digital TestDaF contains four parts corresponding with the sub-skills of reading, listening, writing and speaking. It consists of a total of 23 specially developed task types. The sequence of the test sections and the time required to complete the individual tasks are standardised and binding.

| | | | |
|---------------------------------------------------------------------------------------|------------------------------------|-------------------------------------------------------|----------------------------------------------|
|  | Lesen <i>Reading</i> | 7 Aufgaben 34 Items <i>7 tasks 34 items</i> | ca. 55 Min. <i>approx. 55 min.</i> |
|  | Hören <i>Listening</i> | 7 Aufgaben 30 Items <i>7 tasks 30 items</i> | ca. 40 Min. <i>approx. 40 min.</i> |
|  | Schreiben <i>Writing</i> | 2 Aufgaben <i>2 tasks</i> | ca. 60 Min. <i>approx. 60 min.</i> |
|  | Sprechen <i>Speaking</i> | 7 Aufgaben <i>7 tasks</i> | ca. 35 Min. <i>approx. 35 min.</i> |



Bevor neue Aufgaben eingesetzt werden, müssen sie zur Sicherung der Fairness und Vergleichbarkeit erprobt werden. Die Teilnehmenden bearbeiten daher neben den Echaufgaben eine variable Anzahl an sogenannten Einstreuaufgaben. Diese werden unter Realbedingungen im Test eingesetzt, fließen aber nicht in die Ermittlung der Testergebnisse der Prüfungsteilnehmenden ein. Teilnehmende bearbeiten entsprechend ggf. zwei Aufgaben desselben Aufgabentyps direkt nacheinander. Welche Aufgaben erprobt und welche für die Ermittlung der Testergebnisse verwendet werden, ist den Teilnehmenden nicht bekannt. Das bedeutet, dass die Teilnehmenden alle angezeigten Aufgaben bearbeiten müssen. Durch die Einstreuaufgaben erhöht sich die Gesamtdauer der Prüfung um maximal 30 Minuten.

Die Themen der Aufgaben stammen wie im papierbasierten TestDaF aus den Geistes- und Gesellschaftswissenschaften, Naturwissenschaften, Ingenieurwissenschaften und Technik sowie zusätzlich aus den Wirtschaftswissenschaften, der Medizin und den Humanwissenschaften. Sie sind für Nicht-Spezialist*innen verständlich aufbereitet. Es ist also kein fachliches Vorwissen zur Bewältigung der Aufgaben notwendig.

Die Bearbeitung der Aufgaben im Prüfungsteil Schreiben sowie die Beantwortung der Kurzantworten im Hören erfordern die Eingabe von Text am Computer mittels Tastatur. Der Prüfungsteil Sprechen wird in einem semi-direkten Format durchgeführt, dies bedeutet, dass die Prüfungsteilnehmenden sich in einer simulierten Gesprächssituation am Computer befinden. Dabei hören die Teilnehmenden die Aufgabenstellung über Kopfhörer, lesen sie gleichzeitig am Bildschirm mit und antworten dann einem fiktiven Gesprächspartner bzw. einer Gesprächspartnerin. Dadurch kann sichergestellt werden, dass alle Teilnehmenden diesen Prüfungsteil unter den gleichen Bedingungen ablegen und beispielsweise Vorbereitungs- und Sprechzeiten, Vorgaben und Hilfestellungen für alle Teilnehmenden identisch sind. Auch der mündliche Teil des Tests orientiert sich an den praktischen

Before new tasks are used, they must be trialled to ensure fairness and comparability. In addition to the real tasks, the participants therefore work on a variable number of so-called scatter tasks. These are used in the test under regular conditions, but are not included in the calculation of the participants' test results. Participants may work on two tasks of the same task type directly one after the other. Participants do not know which tasks are tested and which are used to determine the test results. This means that the participants must complete all the tasks displayed. The interspersed tasks increase the total duration of the test by a maximum of 30 minutes.

As in the paper-based TestDaF, the topics of the tasks include aspects of humanities and social sciences, natural sciences, engineering and technology as well as economics, and medicine. They are thus designed that no prior specialised knowledge is required to understand and complete the tasks.

Completing the tasks in the Writing and the Listening sections requires entering text on the computer using a keyboard. The speaking section of the exam is conducted in a semi-direct format: test takers find themselves in a simulated conversation situation on the computer. Participants listen to the task via headphones, read it on the screen at the same time and then answer a fictitious dialogue partner. This ensures that all participants take this part of the test under the same conditions and that, for example, preparation and speaking times, instructions and assistance are strictly identical for all participants. The oral part of the test is also based on the practical study requirements. Traditional oral performance requirements are presentations. These require longer monologue text production, which can be recorded well using a semi-direct format. The oral utterances are recorded and assessed by trained assessors after the test, as are the written performances and the answers to the semi-open listening tasks.

Studienanforderungen. Klassische mündliche Leistungsanforderungen sind Referate und Präsentationen. Diese erfordern längere monologische Textproduktion, die mit einem semi-direkten Format gut erfasst werden kann. Die mündlichen Äußerungen werden aufgezeichnet und nach der Prüfung ebenso wie die schriftlichen Leistungen und die Antworten zu den halboffenen Aufgaben im Hören von geschulten Beurteilenden bewertet.

Das Konstrukt des digitalen TestDaF ist auf einen handlungsorientierten Ansatz kommunikativer Kompetenz ausgerichtet, wie er beispielsweise im *Gemeinsamen europäischen Referenzrahmen für Sprachen (GER)* sowie dem Begleitband, dem *Companion Volume* (Europarat 2001; 2020) beschrieben wird. Gleichzeitig werden im digitalen TestDaF in den vier Prüfungsteilen Lesen, Hören, Schreiben und Sprechen relevante, studienfachübergreifende kommunikative Aufgaben aus dem Hochschulkontext abgebildet, die im Rahmen eines Hochschulstudiums von Studienbewerbenden bewältigt werden müssen:

- Führen von Gesprächen zur Studienorganisation und zu Alltagsfragen
- Führen fachbezogener Gespräche mit Kommiliton*innen und Lehrkräften unter Bezugnahme auf wissenschaftliche Literatur
- Halten von Kurzpräsentationen
- Fachliche Interaktion in Seminaren
- Folgen von Diskussionen, Folgen und argumentativ gestütztes Beziehen von Position
- Aktive Rezeption von Vorlesungen/Vorträge, Lesen von Vorlagen oder Präsentationen, Festhalten in Notizen
- Verarbeitung von Lektüre, Verfassen von annotativen Texten
- Textproduktion in E-Mails, Forumsbeiträgen, Protokollen/Hausarbeiten

The construct of the digital TestDaF is based on a practice-oriented approach to communicative competence, as described, for example, in the Common European Framework of Reference for Languages (CEFR) and the Companion Volume (Council of Europe 2001; 2020). At the same time, the four sections of the digital TestDaF – Reading, Listening, Writing and Speaking – cover relevant, interdisciplinary communicative tasks from the university context that must be mastered by applicants as part of a university degree programme:

- *Conducting discussions on study organisation and everyday issues*
- *Conducting subject-related discussions with fellow students and teachers with reference to scientific literature*
- *Giving short presentations*
- *Professional interaction in seminars*
- *Following discussions and taking a position based on arguments*
- *Active reception of lectures/presentations, reading templates or presentations, taking notes*
- *Processing readings, writing annotative texts*
- *Text production in e-mails, forum posts, minutes/housework*

Die Auswahl basiert auf den Ergebnissen einer Bedarfsanalyse, die g.a.s.t. im Jahr 2011 mit Studierenden und Hochschullehrkräften verschiedener Fachrichtungen und Hochschulen in Deutschland durchgeführt hat (Arras 2012, Marks 2015).

In ihrem Studienalltag arbeiten Studierende häufig teilkompetenzübergreifend. Sie verbinden beispielsweise Lesen und Schreiben themenbezogen in einem Arbeitsprozess miteinander. Dies wurde bei der Entwicklung des digitalen TestDaF berücksichtigt. Daher werden die vier Teilkompetenzen Lesen, Hören, Schreiben und Sprechen in den vier Prüfungsteilen des digitalen TestDaF nicht nur isoliert erfasst – d.h. möglichst ohne Beteiligung anderer Teilkompetenzen –, sondern entsprechend den zuvor gegebenen Beispielen zusätzlich auch teilkompetenzübergreifend in sogenannten integrierten Testaufgaben.

Von den in den Ergebnissen der Bedarfsanalyse genannten Textsorten musste für das standardisierte Prüfungsformat des TestDaF eine Auswahl getroffen werden. Das Verfassen einer Hausarbeit wäre aus naheliegenden Gründen zu umfangreich. Daher können in der Prüfung nur bestimmte Schreibhandlungen oder Textteile verlangt werden, die beim Schreiben einer Hausarbeit relevant sind. Die Textsorte E-Mail ist aufgrund des vielfach formelhaften Einsatzes von Redemitteln als alltags-spezifische Textsorte entweder zu leicht für das angestrebte GER-Niveau B2/C1 oder mit studienfachbezogenen Inhalten zu spezifisch und nicht in einem fächerübergreifenden Test abbildbar. Daher werden diese Textsorten im digitalen TestDaF nicht verwendet.

Der Einführung des neuen digitalen Formats ging eine mehrjährige Entwicklungsphase voraus, in der einzelne Aufgabentypen an g.a.s.t.-Testzentren in Deutschland und weltweit sowie die Software zur Durchführung der Prüfung erprobt wurden. Nach Festlegung des Testformats folgten Pilotierungen von Testaufgaben in weltweiten Field Tests für die Itembank. Die Auswertung der Qualität dieser Aufgaben durch psychometrische Analysen und begleitende qualitative Forschung war darauf ausgerichtet, die zu messende Sprachkompetenz der Teilnehmenden valide, konsistent und zuverlässig zu erfassen und die Qualität der Testaufgaben sowie die damit verbundenen sprachlichen Anforderungen über verschiedene Testsätze oder Testversionen hinweg zu garantieren und konstant zu halten (vgl. Kecker & Eckes, 2022; Kecker, Zimmermann & Eckes, 2022)¹.

Darüber hinaus legt g.a.s.t. großen Wert darauf, im Sinne der Fairness sowohl Teilnehmende als auch Lehrkräfte, die diese Gruppe auf die Prüfungen vorbereiten, sowie andere Stakeholder durch Materialien hinreichend über die Prüfungen zu informieren. Entsprechend haben Teilnehmende im Vorfeld der Prüfung die Möglichkeit, sich durch eine Demo-Version auf

The selection is based on the results of a needs analysis conducted by g.a.s.t. in 2011 with students and university lecturers from various disciplines and universities in Germany (Arras 2012, Marks 2015).

In their day-to-day studies, students simultaneously apply various skills. For example, they combine reading and writing in a subject-related work process. This was taken into account when developing the digital TestDaF. For this reason, the four sub-competences of reading, listening, writing and speaking are not only assessed in isolation in the four parts of the digital TestDaF – i.e. without involving other sub-competences if possible – but also across sub-competences in so-called integrated test tasks in accordance with the examples given above.

A selection of the text types mentioned in the results of the needs analysis had to be made for the standardised TestDaF examination format. Writing a full term paper would be too extensive for obvious reasons. Therefore, only certain writing activities or parts of texts that are relevant for writing a term paper can be expected in the examination. Due to the often formulaic use of phrases as an everyday text type, the e-mail text type is either too easy for the target CEFR level B2/C1 or too specific with subject-related content and cannot be mapped in an interdisciplinary test. Therefore, these text types are not used in the digital TestDaF.

The introduction of the new digital format was preceded by a development phase of several years, during which the software for conducting the test and individual item types were tested at g.a.s.t. test centres world-wide. Once the test format had been set out, test items were piloted in worldwide field tests for the item bank. The evaluation of the quality of these tasks through psychometric analyses and accompanying qualitative research was aimed at assessing the language competence of the participants to be measured in a valid, consistent and reliable manner and guaranteeing and maintaining the quality of the test tasks and the associated linguistic requirements across different test sets or test versions (cf. Kecker & Eckes, 2022; Kecker, Zimmermann & Eckes, 2022)¹.

In addition, g.a.s.t. attaches great importance to providing sufficient information about the examinations to both participants and instructors who prepare this group for the examinations, as well as other stakeholders, in the interests of fairness. Accordingly, participants have the opportunity to prepare for the exam using a demo version in the run-up to the exam. There are also video tutorials on the website in which participants can find out everything they need to know about the requirements of the individual task types. For test centres, extensive material on the exam concept and format as well as the

die Prüfung vorzubereiten. Ebenfalls gibt es auf der Webseite Video-Tutorials, in denen Teilnehmende alles Wissenswerte zu den Anforderungen der einzelnen Aufgabentypen erfahren. Für Testzentren stehen umfangreiche Materialien zum Prüfungskonzept und -format sowie den einzelnen Prüfungsteilen im Testzentren-Portal zur Verfügung. Ebenfalls finden sich dort Materialien zur Prüfungsvorbereitung (s. nächstes Kapitel).

2.1.2 Digitaler TestDaF: Konzept zur Prüfungsvorbereitung

Um eine angemessene und an der Entwicklung von zentralen kommunikativen Kompetenzen ausgerichtete Prüfungsvorbereitung zu ermöglichen, wurde ein detailliertes Prüfungsvorbereitungskonzept für den digitalen TestDaF ausgearbeitet. Ziel des Konzepts ist es, Lehrende an lizenzierten TestDaF-Testzentren dabei zu unterstützen, in ihrer Institution Vorbereitungskurse für TestDaF-Teilnehmende zu entwickeln. Die Zielsetzung des Vorbereitungskonzepts wurde mit der – in Kooperation mit erfahrenen Lehrkräften (u.a. der DaF-Abteilung der Ruhr-Universität Bochum und der interDaF, Leipzig) – Entwicklung umfangreicher Lehr- und Lernmaterialien praktisch umgesetzt. Diese Materialien bestehen aus sechs Lerneinheiten, die jeweils in fünf bis zehn Unterrichtseinheiten unterteilt sind. Die Unterrichtseinheiten sind hochschulrelevanten Themenbereichen und Kompetenzen gewidmet und enthalten Übungen und Lernaktivitäten, inklusive Bild- und Audiomaterialien, für die Verwendung im Online- oder Präsenzunterricht. Ergänzt wird das Unterrichtsmaterial durch einen kompletten Modellsatz des digitalen TestDaF, den Lernende über die Lernplattform DUO (Deutsch-Uni Online) aufrufen und bearbeiten können. Die Bearbeitung kann in den Unterricht integriert werden, Lösungsschlüssel und Lösungshinweise ermöglichen es Lernenden jedoch zudem, den Modellsatz eigenständig, ohne Anweisung von Lehrenden durchzuarbeiten.

Ergänzt werden die Materialien durch Handreichungen in Form von didaktischen Hinweisen für Lehrende, in denen die Zielsetzung und die methodische Umsetzung der einzelnen Übungsschritte im Detail erläutert werden. Die Handreichungen für Lehrkräfte enthalten in jeder Unterrichtseinheit didaktische Angaben zur Arbeit mit den Lernmaterialien. Sie beschreiben die Lernziele der einzelnen Aufgabenschritte, geben Hinweise dazu, was bei der Durchführung zu beachten ist und bieten Tipps und Anregungen zu weiteren Umsetzungsmöglichkeiten einzelner Aufgabenschritte.

Die Lern- und Unterrichtseinheiten sind insgesamt so gestaltet, dass sie je nach institutioneller Voraussetzungen, Bedürfnisse und Anforderungen unabhängig voneinander im Unterricht eingesetzt werden können. Somit erfüllt diese Materialsammlung die Funktion eines „Werkzeugkastens“, mit deren Hilfe Lehrkräfte ihre Vorbereitungskurse maßgeschneidert für unterschiedliche Gruppen von Lernenden planen und durchführen können.

individual parts of the exam are available on the test centre portal. Materials for exam preparation can also be found there (next chapter).

2.1.2 Digital TestDaF: Concept for exam preparation

A detailed exam preparation concept for the digital TestDaF has been developed to enable appropriate exam preparation geared towards the development of key communicative competences. The aim of the concept is to support teachers at licensed TestDaF test centres in developing preparation courses for TestDaF participants at their institution. The aim of the preparation concept was realised in practice with the development of comprehensive teaching and learning material in cooperation with highly experienced teachers (including the DaF department of the Ruhr University Bochum and interDaF, Leipzig). This material consists of six learning units, each divided into five to ten teaching units. The teaching units are dedicated to subject areas and competences relevant to higher education and contain exercises and learning activities, including visual and audio materials, for use in online or face-to-face lessons. The teaching material is supplemented by a complete model set of the digital TestDaF, which learners can access and work on via the DUO (Deutsch-Uni Online) learning platform. The work can be integrated into lessons, but solution keys and instructions also enable learners to work through the model set independently, without input from teachers.

The material is supplemented by handouts in the form of didactic notes for teachers, in which the objectives and methodological implementation of the individual exercise steps are explained in detail. The handouts for teachers contain didactic information on working with the learning material in each teaching unit. They describe the learning objectives of the individual exercise steps, provide information on what needs to be taken into account when carrying them out and offer tips and suggestions on further ways of implementing individual exercise steps.

The learning and teaching units are designed in such a way that they can be used independently of each other in the classroom, depending on the institutional conditions, needs and requirements. This collection of materials therefore fulfils the function of a “toolbox” that teachers can use to plan and implement their preparation courses tailored to different groups of learners.

¹ Kecker, G. & Eckes, T. (2022). Der digitale TestDaF: Aufbruch in neue Dimensionen des Sprachtestens. *Informationen Deutsch als Fremdsprache*, 49(4), 289-324.
Kecker, G., Zimmermann, S. & Eckes, T. (2022). Der Weg zum digitalen TestDaF: Konzeption, Entwicklung und Validierung. In P. Gretschnig & N. Wulff (Hrsg.), *Deutsch als Zweit- und Fremdsprache in Schule und Beruf* (S. 393-410). Paderborn: Brill Schöningh.

2.1.3 Konzeption und Durchführung des papierbasierten TestDaF

Die fortlaufende Validierung des papierbasierten TestDaF durch die Forschungsarbeit bei g.a.s.t. sowie das Qualitätssiegel der *Association of Language Testers in Europe* (ALTE) belegen die weiterhin hohe Qualität der Sprachprüfung. Teilnehmende und zulassende Stellen an Hochschulen können so darauf vertrauen, dass der TestDaF ein hochwertiger Sprachnachweis für die Zulassung zu einem Studium an Hochschulen in Deutschland ist.

Die Anforderungen der Prüfungsaufgaben im TestDaF sind auf die Sprachhandlungen zugeschnitten, die Studienbewerbende zu Beginn eines Studiums bewältigen müssen. Der TestDaF prüft die Sprachkompetenz auf einem fortgeschrittenen Niveau. Die TestDaF-Niveaus TDN 3, TDN 4 und TDN 5 decken die Niveaus B2 bis C1 des Gemeinsamen europäischen Referenzrahmens (GER) ab.

Die Teilkompetenzen im papierbasierten TestDaF werden weitgehend getrennt geprüft, die Ergebnisse in den Prüfungsteilen Leseverstehen (LV), Hörverstehen (HV), Schriftlicher Ausdruck (SA) und Mündlicher Ausdruck (MA) getrennt voneinander auf dem Zertifikat ausgewiesen. Dennoch ist in den Prüfungsteilen HV, SA und MA auch das Leseverstehen relevant, da Teilnehmende die Anleitungen und Aufgabenstellung verstehen müssen. In den Prüfungsteilen HV1 und HV3 müssen außerdem Kurzantworten formuliert werden. Hier spielt also auch Schreibfertigkeit eine Rolle. Im digitalen TestDaF wurden diese Ansätze weiterentwickelt, die Testaufgaben sind kompetenzübergreifend angelegt. In den Prüfungsteilen Hören, Schreiben und Sprechen des digitalen TestDaF gibt es daher zusätzlich zu den isolierten Testaufgaben auch solche, die zwei Teilkompetenzen einbinden (vgl. Kapitel 2.1.1).

2.1.4 Teilnehmendenentwicklung

In den Jahren 2020 bis 2022 legten deutlich weniger Teilnehmende den TestDaF ab als noch in den Jahren vor der Pandemie: Die Anzahl der TestDaF-Absolvent*innen fiel im Jahr 2020 auf 25.005, nachdem sie im Jahr zuvor noch knapp 47.500 betragen hatte. Nach einem leichten Anstieg der Teilnehmendenzahlen im Jahr 2021 lag die Zahl der TestDaF-Teilnehmenden im Jahr 2022 bei rund 23.000. Beeinflusst wurden diese Zahlen neben der Pandemie auch durch die geopolitische Lage in Ländern mit traditionell hoher Nachfrage wie der Russischen Föderation und der Ukraine.

Seit Einführung des TestDaF im Jahr 2001 bis Ende 2022 wurden 521.110 Einzelprüfungen mit Deutschlernenden aus 189 Ländern durchgeführt. Die am stärksten vertretene Nation ist China mit 180.928 Teilnehmenden (34,7%). Auf Platz zwei liegt wie bisher die Russische Föderation mit insgesamt 30.795 Teilnehmenden (5,9%). Die Türkei erreicht mit 21.334 Teilnehmenden Platz drei.

2.1.3 Concept and implementation of the Paper-based TestDaF

The ongoing validation of the paper-based TestDaF by the research work at g.a.s.t. and the quality seal of the Association of Language Testers in Europe (ALTE) confirm the high quality of the language test for admission to a degree programme at universities in Germany.

The requirements of the TestDaF examination tasks are tailored to the language skills that applicants have to master at the beginning of their studies. The TestDaF tests language competence at an advanced level. TestDaF levels TDN 3, TDN 4 and TDN 5 cover levels B2 to C1 of the Common European Framework of Reference for Languages (CEFR).

The sub-competences in the paper-based TestDaF are largely tested separately and the results in the reading comprehension (LV), listening comprehension (HV), written expression (SA) and oral expression (MA) sections are shown separately on the certificate. Nevertheless, reading comprehension is also relevant in the HV, SA and MA sections of the examination, as participants must understand the instructions and tasks. In the HV1 and HV3 sections of the exam, short answers must also be formulated. Hence, writing skills also play a role here. In the digital TestDaF, these approaches have been further developed and the test tasks are designed to cover all competences. In the Listening, Writing and Speaking sections of the digital TestDaF, in addition to the isolated test tasks, there are tasks that integrate two sub-competences (see chapter 2.1.1).

2.1.4 Development of test taker numbers

In the years 2020 to 2022, significantly fewer participants took the TestDaF than in the years before the pandemic: the number of TestDaF graduates fell from just under 47,500 in 2019 to 25,005 in 2020. After a slight increase in the number of participants in 2021, the number of TestDaF participants in 2022 stabilised at around 23,000. In addition to the pandemic, participation was affected by the geopolitical situation in countries with traditionally high demand, such as the Russian Federation and Ukraine.

Since its introduction in 2001 until the end of 2022, 521,110 individual TestDaF examinations were conducted with German learners from 189 countries. China ranks top with 180,928 participants (34.7%) with the Russian Federation in second place with a total of 30,795 participants (5.9%). Turkey ranked third with 21,334 participants.

2.1.5 Qualitätsmanagement: Zertifizierung von Beurteilenden

Personen, die im Auftrag von g.a.s.t. Prüfungsleistungen aus den Prüfungsteilen Hören, Schreiben und Sprechen des papierbasierten oder des digitalen TestDaF abnehmen, müssen ein mehrschrittiges Zertifizierungsverfahren durchlaufen.

Der erste Schritt ist die Teilnahme an einer ein- bzw. zweitägigen Erstschtung. Hierzu werden interessierte DaF-Lehrpersonen, die die Voraussetzungen für die Tätigkeit als Beurteilende erfüllen², eingeladen. In dieser Schulung machen Referent*innen von g.a.s.t. die Teilnehmenden mit den Aufgabenformaten, den Beurteilungsverfahren und den Maßnahmen zur Qualitätssicherung vertraut. Schriftliche und mündliche Teilnehmendenleistungen bzw. die Antworten im Hören werden beurteilt und die Beurteilungen diskutiert. Im Anschluss an die Schulung müssen die Teilnehmenden nachweisen, dass sie das Gelernte praktisch anwenden können, indem sie ausgewählte mündliche und schriftliche Teilnehmendentexte bzw. Kurzantworten in Probebeurteilungen bewerten.

Die korrekte Einstufung dieser Texte und dazugehörigen Begründungen werden anschließend von g.a.s.t.-Mitarbeiter*innen geprüft. Alle Bewerber*innen erhalten eine ausführliche Rückmeldung zu ihren Beurteilungen. Entsprechen die Einstufungen, Begründungen und Formalia den Vorgaben, beginnt eine Probephase, in der die Beurteilenden Leistungen von Prüfungsteilnehmenden bewerten und dabei testmethodisch begleitet werden. Dazu werden Strenge bzw. Milde der Beurteilenden und die Konsistenz ihrer Beurteilungen beobachtet. Liegen die erhobenen Werte innerhalb eines vorgesehenen Rahmens, erhalten die Beurteilenden ein Zertifikat, das zwei Jahre lang gültig ist.

Anschließend nehmen die Beurteilenden alle zwei Jahre an einem Training zur weiteren Zertifizierung teil. Hier werden die Kompetenzen vertieft und das eigene Beurteilungsverhalten durch Diskussionen in der Gruppe reflektiert. Auch an dieses Training schließt sich eine Bewertungsphase mit Probebeurteilungen an.

Bis Ende 2022 wurden insgesamt 120 Beurteilende für den digitalen und 280 Beurteilende für den papierbasierten TestDaF zertifiziert.

2.1.5 Quality control: certification of assessors

Assessment staff in the listening, writing and speaking sections of the paper-based or digital TestDaF on behalf of g.a.s.t. must undergo a multi-step certification procedure.

The first step is to take part in a one or two-day initial training course. In this course, specially trained g.a.s.t. staff familiarise the participants with the task formats², assessment procedures and quality assurance measures. Written and oral performance and listening answers are assessed and the assessments are discussed. Following the training, participants must demonstrate that they can apply what they have learnt in practice by assessing selected oral and written participant texts or short answers in mock assessments.

Subsequently, the correct grading of these texts and the corresponding reasoning for the assessment are checked by g.a.s.t. employees. All applicants receive detailed feedback on their assessments. If the grading, justifications and formalities meet the requirements, a trial phase begins in which the assessors evaluate the performance of examination participants and are accompanied by test methodology. The rigour or leniency of the assessors and the consistency of their assessments are observed. If the assessed values are within the specified range, the assessors receive a certificate that is valid for two years.

Assessors undergo further certification training every two years. The skills are deepened and their own assessment behaviour is reflected on through discussions in the group. This training is also followed by an assessment phase with trial assessments.

By the end of 2022, a total of 120 assessors had been certified for the digital TestDaF and 280 assessors for the paper-based TestDaF.



2.2 TestAS

Der Test für Ausländische Studierende/Test for Academic Studies (TestAS) ist ein Studieneignungstest für internationale Bewerber*innen um einen Studienplatz in Deutschland. Der TestAS ist kein Wissenstest, sondern misst allgemeine und studienfeldbezogene kognitive Fähigkeiten und Anwendungskompetenzen.

Studieninteressierte erhalten mit dem TestAS-Zertifikat eine realistische Einschätzung ihrer Fähigkeiten. Hochschulen können mit dem Testergebnis ihre Zulassungsentscheidung auf ein verlässliches Kriterium stützen, das für alle Bewerber*innen gleich ist, unabhängig von Unterschieden im Schul- und Notensystem der Herkunftsländer. Der TestAS liefert somit eine gute Prognose für den Studienerfolg.

Der TestAS wurde von der ITB Consulting GmbH, Bonn, entwickelt. Die weltweite Organisation des TestAS liegt bei g.a.s.t. in Bochum. Bis Ende 2015 wurden Entwicklung und Einsatz des TestAS durch den DAAD aus Mitteln des Bundesministeriums für Bildung und Forschung (BMBF) gefördert.

2.2 TestAS

The Test für Ausländische Studierende /Test for Academic Studies (TestAS) is a study aptitude test for international applicants for a study place in Germany. The TestAS is not a knowledge test, but measures general and study-related cognitive skills and transfer competences.

Prospective students receive a realistic assessment of their abilities with the TestAS certificate. With the test result, universities can base their admission decision on a reliable criterion that is the same for all applicants, regardless of differences in the school and grading system of the countries of origin. The TestAS therefore provides a good prognosis for academic success.

TestAS was developed by ITB Consulting GmbH, Bonn. The overall organisation of TestAS is carried out by g.a.s.t. in Bochum. Until the end of 2015, the development and use of TestAS was funded by the Federal Ministry of Education and Research (BMBF) through the channels of the DAAD.

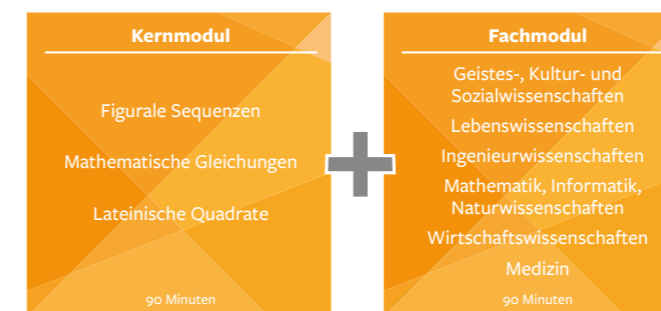
Seit Juni 2022 wird parallel zum papierbasierten TestAS auch der digitale TestAS angeboten. Der digitale TestAS wurde in Kooperation mit der Abteilung für Differentielle Psychologie und Psychologische Diagnostik am Institut für Psychologie und Pädagogik der Universität Ulm und dem Fachgebiet Psychologische Diagnostik am Institut für Psychologie der Universität Kassel entwickelt. Ziel der Kooperation war neben der Digitalisierung des TestAS auch die Revision des Testformats. Die Fähigkeiten, die in beiden Formaten geprüft werden, sind identisch. Der digitale TestAS stellt daher die gleichen Anforderungen an die Teilnehmenden wie der papierbasierte TestAS. Beide Tests werden in deutscher und englischer Sprache angeboten.

2.2.1 Konzept, Aufbau und Einführung des digitalen TestAS

Für den digitalen TestAS wurden auf Grundlage neuester wissenschaftlicher Erkenntnisse eigene zeitgemäße Aufgabentypen entwickelt, erprobt und für die digitale Darstellung und Bearbeitung optimiert. Dadurch wird sichergestellt, dass auch der digitale TestAS höchsten Anforderungen an Präzision und Zuverlässigkeit genügt.

Durch die Entwicklung in enger fachlicher Zusammenarbeit mit deutschen Hochschulen spiegelt der Test die Anforderungen für die Studienfelder Geistes-, Kultur- und Sozialwissenschaften, Lebenswissenschaften, Mathematik, Informatik und Naturwissenschaften, Ingenieurwissenschaften, Wirtschaftswissenschaften sowie Medizin sehr gut wider und ist somit ein valides Instrument zur Einschätzung der Studieneignung von Bewerbenden.

Der digitale TestAS besteht aus zwei Prüfungsteilen: das **Kernmodul** testet die allgemeine Studieneignung, die **studienfeldbezogenen Fachmodule** prüfen die studienfeldspezifische Eignung. Die Prüfungsdauer ist bei Erhalt der Messgenauigkeit im Vergleich zum papierbasierten TestAS um 80 Minuten reduziert.



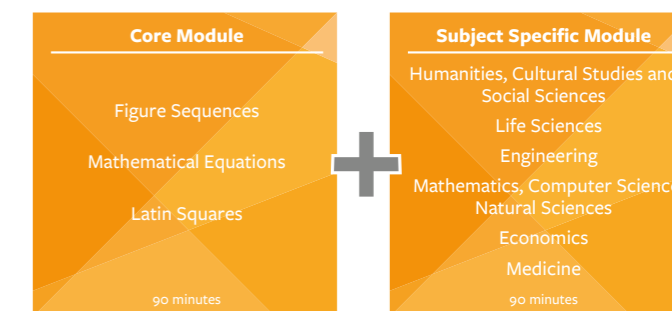
Since June 2022, the digital TestAS has been offered alongside the paper-based TestAS. The digital TestAS was developed in cooperation with the Department of Differential Psychology and Psychological Diagnostics at the Institute of Psychology and Education at the University of Ulm and the Department of Psychological Diagnostics at the Institute of Psychology at the University of Kassel. The aim of the cooperation was not only to digitise the TestAS but to revise the test format. The skills tested in both formats are identical. The digital TestAS therefore places the same demands on the participants as the paper-based TestAS. Both tests are offered in German and English.

2.2.1 Concept, structure and introduction of the digital TestAS

Based on the latest academic research, we have developed and tested our own contemporary item types for the digital TestAS and optimised them for digital display and processing. This ensures that the digital TestAS also fulfils the highest standards of precision and reliability.

Due to close cooperation with German universities, the test reflects the requirements for the fields of humanities, cultural and social sciences, life sciences, mathematics, computer science and natural sciences, engineering, economics and medicine very well and is therefore a valid instrument for assessing the suitability of applicants.

*The digital TestAS consists of two parts: the **core module** tests general aptitude for study, while the **subject-specific modules** test aptitude for a specific field of study. The duration of the test is reduced by 80 minutes compared to the paper-based TestAS while maintaining the accuracy of measurement.*



Das Kernmodul zur Messung der allgemeinen Studieneignung besteht aus drei Subtests. Gemessen werden allgemeine kognitive Fähigkeiten (wie Arbeitsgedächtnis, Logik, analytisches Denken), die für ein grundständiges Studium in Deutschland relevant sind. Die Aufgabentypen eignen sich optimal für die digitale Darstellung. Die Aufgaben des Kerntests des digitalen TestAS sind über die Instruktionen hinaus vollständig sprachfrei. Das Kernmodul ermöglicht den Vergleich der Teilnehmenden unabhängig vom absolvierten Fachmodul.

Die Fachmodule wurden als Reaktion auf den Bedarf der Hochschulen im Vergleich zum papierbasierten TestAS um die beiden Fachrichtungen Medizin und Lebenswissenschaften ergänzt. Zur Definition und Gewichtung der Prüfungsinhalte wurden Taxonomien und Rating-Studien erstellt. So wurde sichergestellt, dass die Module repräsentativ für die verschiedenen Studienfelder sind. Die Prüfungsaufgaben bestehen aus einer Kombination von fachtypischer Problemstellung (Input/Stamm) und mehreren Single-Choice-Fragen. In den Aufgaben werden grundlegende fachliche Kenntnisse, Anwendungskompetenz und die Fähigkeit zum Transfer getestet. Dabei ist auch der Umgang mit fachtypischen Darstellungsformen wie Texten, Tabellen, Abbildungen und Formeln entscheidend. Die Fragestellungen sind hierbei nicht reproduktiv, sondern problemorientiert.

Der digitale TestAS wurde 2022 erstmals am 22. Juni mit 293 Teilnehmenden an 28 Testzentren und erneut am 22. November durchgeführt.

2.2.2 Konzept und Aufbau des papierbasierten TestAS

Das Konzept für den TestAS wurde in den Jahren 2005 bis 2007 unter Mitwirkung einschlägiger Expertenkommissionen (Anforderungsworkshops) erstellt und seither kontinuierlich weiterentwickelt. Der TestAS besteht aus dem sogenannten **Kerntest**, der die allgemeine Studieneignung misst, sowie aus einem von **vier Fachmodulen** (s. Abb. auf der nächsten Seite). Während der Kerntest für alle Teilnehmenden einheitlich und verbindlich ist, wählen die Teilnehmenden bei der Anmeldung zum Test eines der folgenden vier Fachmodule:

- Geistes-, Kultur- und Gesellschaftswissenschaften (GKG)
- Ingenieurwissenschaften (ING)
- Mathematik, Informatik und Naturwissenschaften (MIN)
- Wirtschaftswissenschaften (WIWI)

Die Fachmodule messen Fähigkeiten, die nötig sind, um typische Anforderungen in einem bestimmten Studienfeld zu erfüllen. Solche Fähigkeiten sind z.B. „Problemlösen“, „Strukturen analysieren“ oder „schlussfolgerndes Denken“.

Der Kerntest und die vier Fachmodule setzen sich jeweils aus verschiedenen Aufgabengruppen zusammen. Die Aufgabengruppen der Fachmodule sind auf die Anforderungen in den

The core module for measuring general study aptitude consists of three subtests. It measures general cognitive abilities (such as working memory, logic, analytical thinking) that are relevant for an undergraduate degree programme in Germany. The task types are ideally suited for digital presentation. The tasks of the core test of the digital TestAS are completely language-free beyond the instructions. The core module allows participants to be compared regardless of the subject module they have completed.

In response to the universities' demands, the subject modules were expanded to include the two subject areas of medicine and life sciences in comparison to the paper-based TestAS. Taxonomies and rating studies were created to define and weight the examination content. This ensured that the modules are representative of the various fields of study. The examination tasks consist of a combination of subject-specific problems (input/stem) and several single-choice questions. The tasks test basic subject-specific knowledge, application skills and the ability to transfer knowledge. The use of typical subject-specific forms of presentation such as texts, tables, figures and formulae is crucial. The questions are not reproductive, but problem-oriented.

The digital TestAS was conducted for the first time in 2022 on 22 June with 293 participants at 28 test centres and again on 22 November.

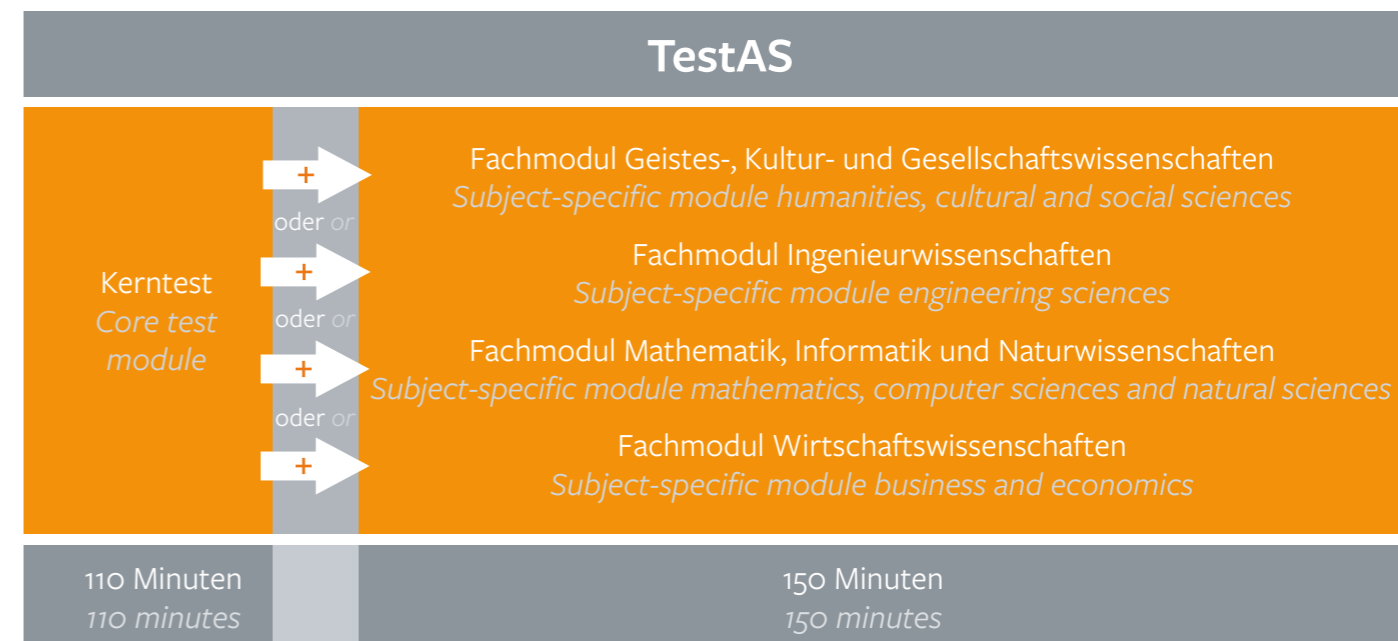
2.2.2 Concept and structure of the paper-based TestAS

*The concept of the TestAS was developed between 2005 and 2007 with the involvement of relevant expert commissions (requirements workshops) and has been continuously developed since then. The TestAS consists of the so-called **core test**, which measures general aptitude for study, and one of **four specialised modules** (s. Fig. next page). While the core test is standardised and binding for all participants, participants choose one of the following four subject modules when registering for the test:*

- *Humanities, cultural and social sciences (GKG)*
- *Engineering Sciences (ING)*
- *Mathematics, computer science and natural sciences (MIN)*
- *Economics (WIWI)*

The subject modules measure skills that are necessary to fulfil typical requirements in a specific field of study. Such skills are, for example, “problem solving”, “analysing structures” or “deductive thinking”.

The core test and the four specialised modules are each made up of different groups of tasks. The task groups of the specialised modules are tailored to the requirements of the four spe-



vier Fachgebieten zugeschnitten. Für die Testerstellung stehen derzeit 3.530 einzelne Aufgaben (Items) zur Verfügung. Hieraus lässt sich eine Vielzahl unterschiedlicher Testsätze zusammensetzen, die in ihrem Schwierigkeitsniveau kalibriert sind. Bis Ende 2019 wurden jährlich rund 300 Items neu entwickelt und unter Echtbedingungen erprobt. Als Einstreuaufgaben laufen die neuen Items für die Teilnehmenden nicht erkennbar in den einzelnen Testläufen mit; die Antworten auf diese Items gehen nicht in die Ermittlung der Testergebnisse ein.

Ende 2022 boten rund 400 Testzentren weltweit den TestAS an.

2.2.3 Teilnehmendenentwicklung des papierbasierten TestAS

2020 legten 7.878 TN den TestAS ab. Trotz weltweiter pandemiebedingter Einschränkungen blieb die Zahl damit im Vergleich zu 2019 stabil. Im Jahr 2021 musste pandemiebedingt ein Rückgang von 22% verzeichnet werden. 2022 stiegen die Teilnehmendenzahlen und bewegten sich mit 7.239 TN fast auf dem Niveau von 2020.

Die Prüfung stellt für die Teilnehmenden ein aussagekräftiges Zertifikat für die Zulassung an einer deutschen Hochschule dar. Gerade im Fachbereich Medizin wurden neue Nutzer-Hochschulen gewonnen.

Der TestAS ist fester Bestandteil des Akademischen Prüfstellensverfahrens (APS) der deutschen Botschaft in der Volksrepublik China (seit 2009) und Vietnam (seit 2012). Neu hinzugekommen ist im September 2022 der Einsatz an der APS in Indien, wo der TestAS aktuell noch optional im Rahmen des Auswahlverfahrens

cialist areas. There are currently 3,530 individual tasks (items) available for test creation. These can be used to compile a large number of different test sets that are calibrated in terms of their level of difficulty. By the end of 2019, around 300 new items had been developed each year and tested under real conditions. As interspersed tasks, the new items are not recognisable to the participants in the individual test runs; the answers to these items are not included in the calculation of the test results.

At the end of 2022, around 400 test centres worldwide were offering TestAS.

2.2.3 Development of test taker numbers of the paper-based TestAS

In 2020, 7,878 participants took the TestAS. Despite global pandemic-related restrictions, the number remained stable. In 2021, a decline of 22% was recorded due to the pandemic. In 2022, the number of participants increased and was almost on a par with 2020 at 7,239 participants.

The test provides participants with a meaningful certificate for admission to a German university. New user universities have been gained, particularly in the field of medicine.

The TestAS is also used in the Academic Evaluation Centre Procedure (APS) of the German Embassy in the People's Republic of China (since 2009) and Vietnam (since 2012). A new addition in September 2022 is the use of the TestAS at the APS in India, where it is currently used optionally as part of the selection process. TestAS thus makes a significant contribution to quality assurance for Chinese, Vietnamese and Indian students at German

eingesetzt wird. Der TestAS trägt damit wesentlich zur Qualitätssicherung des Zugangs chinesischer, vietnamesischer und indischer Studierender an deutschen Hochschulen bei. An der Vietnamese German University hat sich der TestAS seit 2011 zudem als Zugangsprüfung für die hochschuleigenen Bachelor-Programme bewährt.

Insgesamt haben seit der ersten Prüfung im Jahr 2007 85.625 Teilnehmende den papierbasierten TestAS abgelegt.

2.2.4 TestAS in Hochschulzugangsverfahren

Derzeit nutzen rund 50 deutsche Hochschulen den TestAS im Rahmen von Zulassungsverfahren. Das TestAS-Ergebnis kann dabei verpflichtender Bestandteil des Zulassungsverfahrens sein oder fakultativ zur Verbesserung der Chancen auf einen Studienplatz verwendet werden. Die Möglichkeiten zur Nutzung des TestAS sind vielfältig; Hochschulen können den TestAS entsprechend ihrer Anforderungen individuell einsetzen. Hierfür stehen fünf beispielhafte Nutzungsmodelle zur Verfügung, die unter www.testas.de/nutzungsmodelle vorgestellt werden.

Um Hochschulverwaltungen wie auch Vertretungen der Fakultäten ausführlich über den TestAS zu informieren, finden zum einen regelmäßig TestAS-Webinare und -Workshops statt, zum anderen besuchen Mitarbeitende von g.a.s.t. einzelne Hochschulen und beraten individuell vor Ort.

2.2.5 Digitaler Mastertest (dMAT)

Hochschulen fordern verstärkt Zugangsprüfungen für Master-Studierende, die ähnlich wie der TestAS standardisiert sind, die fachliche Eignung der Studienbewerbenden messen und digital und weltweit abgelegt werden können. Die Nachfrage stammt derzeit insbesondere aus Fächern und Studiengängen mit hochkompetitiven Auswahlverfahren. Die Hochschulen fordern dabei eine stärkere fachliche Differenzierung als beim TestAS für grundständige Studiengänge und einen höheren Anteil an Vorwissen.

Daher hat g.a.s.t. 2020, wiederum in Zusammenarbeit mit den Universitäten Ulm und Kassel, begonnen, einen Zugangstest für Master-Studierende zu entwickeln. Der grundlegende Aufbau des digitalen Mastertests (dMAT) ist identisch mit dem Aufbau des TestAS (s. Abb. auf der nächsten Seite): Im Kernmodul werden allgemeine kognitive Fähigkeiten gemessen, im Fachmodul wird die fachspezifische Eignung geprüft. Anders der digitale TestAS misst der dMAT im Bachelor erworbenes Wissen und Anwendungskompetenz. Zudem sind die Fachbereiche enger gefasst und dadurch spezifischer auf Masterstudiengänge angepasst. Die Fachmodule werden in besonders enger Abstimmung mit Hochschulen entwickelt.

universities. At the Vietnamese German University, the TestAS has also proven its worth as an entrance exam for the university's own Bachelor's programmes since 2011.

A total of 85,625 participants have taken the paper-based TestAS since the first examination in 2007.

2.2.4 TestAS in university admission procedures

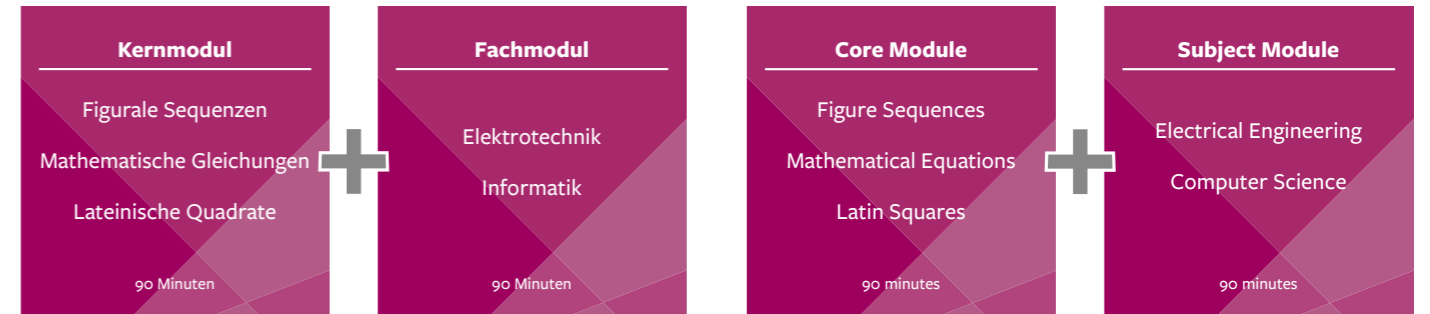
Around 50 German universities currently use the TestAS as part of the admissions process. The TestAS result can be a mandatory part of the admission procedure or can be used optionally to improve the chances of gaining entrance. There are many ways to use the TestAS; universities can use the TestAS according to their individual requirements. Five exemplary models using TestAS are presented at www.testas.de/nutzungsmodelle.

In order to provide university administrations and faculty representatives with detailed information about TestAS, TestAS webinars and workshops are held regularly, and g.a.s.t. staff visit universities and provide individual advice on site.

2.2.5 Digital master test (dMAT)

Universities are increasingly demanding standardised entrance examinations for Master's students to measure the subject-related aptitude of applicants offered digitally and worldwide. Demand is currently increasing in subjects and degree programmes with highly competitive selection procedures in particular. Universities are expecting greater subject differentiation than in the TestAS for undergraduate degree programmes and a higher proportion of prior knowledge.

Hence, g.a.s.t., again in collaboration with the universities of Ulm and Kassel, has started to develop an entrance test for Master's students in 2020. The basic structure of the digital Master's test (dMAT) is identical to the structure of the TestAS (s. Fig. next page): General cognitive abilities are measured in the core module, while subject-specific aptitude is tested in the subject module. Unlike the digital TestAS, though, the dMAT measures knowledge and application skills acquired in a Bachelor's programme. In addition, the subject areas are more narrowly defined and therefore more specifically adapted to Master's degree programmes. The specialised modules are developed in particularly close cooperation with universities.



In Kooperation mit dem Karlsruher Institut für Technologie (KIT) hat g.a.s.t. das Fachmodul Elektrotechnik entwickelt. Das Fachmodul Informatik wurde zusammen mit der Universität Paderborn entwickelt. Für beide Fachmodule wurde in einem ersten Schritt eine Taxonomie erstellt, die die Inhalte der Fachmodule zu definieren halfen. Anhand dieser Taxonomien wurde über Rating-Studien analysiert, welche Gewichtung dieser Inhalte am besten geeignet ist, die Anforderungen des Studiengangs abzubilden. Nach gemeinsamen Workshops haben Professor*innen sowie wissenschaftliche Mitarbeitende der Elektrotechnik des KIT und der Informatik der Universität Paderborn Aufgaben entworfen, die in Zusammenarbeit aus sprachlicher und testmethodischer Perspektive überarbeitet wurden. Hierbei lag ein Fokus auf der Herausforderung, die inhaltlich komplexen Fachmodulaufgaben dennoch sprachlich so einfach wie möglich zu gestalten.

In co-operation with the Karlsruhe Institute of Technology (KIT), g.a.s.t. has developed the electrical engineering module. The computer science module was developed together with the University of Paderborn. In a first step, a taxonomy was created for both specialised modules which helped to define the contents of the specialised modules. Based on these taxonomies, rating studies were used to analyse which weighting of this content is best suited to reflect the requirements of the degree programme. Following joint workshops, professors and junior researchers from the electrical engineering department at KIT and the computer science department at Paderborn University designed tasks that were reviewed in collaboration from a linguistic and testing methodology perspective. The focus here was on the challenge of making the complex content of the specialised module tasks as simple as possible in terms of language.





2.3 onSET

2.3.1 Konzept und Aufbau

Der onSET ist ein Online-Testsystem zur Einstufung von Fremdsprachenkenntnissen. Das von g.a.s.t. entwickelte System erlaubt die Messung der allgemeinen Sprachkompetenz; eine Einstufung nach einzelnen Sprachfertigkeiten erfolgt nicht.

Alle im onSET angebotenen Tests sind Lückentests, deren Aufbau dem sogenannten C-Test-Prinzip folgt: In kurzen, authentischen Texten werden durch systematische Tilgung von Wortteilen jeweils 20 Lücken erzeugt. Teilnehmende haben die Aufgabe, in jedem Text die Lücken korrekt zu füllen. Jeder Test bildet Aufgaben aus den Niveaustufen von A2 bis C1 ab. Die Anzahl der korrekt gefüllten Lücken gibt Auskunft über das erreichte GER-Niveau.

Der onSET verfügt über eine komplett internetgestützte Testanwendung. Alle Komponenten des Tests und der Testdurchführung sind online verfügbar: Registrierung der Teilnehmenden,

2.3 onSET

2.3.1 Concept and structure

onSET is an online test system for the self-assessment of foreign language skills. The system developed by g.a.s.t. measures general language competence; it does not assess individual language skills.

onSET tests are structured according to the so-called C-test principle: In short, authentic texts, 20 gaps are created by systematically deleting words or parts of words. Participants are requested to fill in the gaps correctly in each text. Each test contains tasks from levels A2 to C1. The number and level of difficulty of correctly filled gaps provides information about the CEFR level achieved.

The onSET is a completely web-based test application. All components of the test and test administration are available online: Registration of participants, set-up and administration of test

Einrichtung und Verwaltung von Testterminen, Buchung von Testterminen, Testbearbeitung und Ergebnismittlung und Übermittlung.

2.3.2 Teilnehmendenentwicklung

Im Jahr 2020 haben 10.929 onSET-Prüfungen mit der deutschen Sprachversion (onSET-Deutsch) oder mit der englischen Version (onSET-English) stattgefunden; dies entspricht einem Rückgang gegenüber dem Vorjahr (mit insgesamt 25.967 Prüfungen) um 58%. Dieser starke Rückgang ist auf die Covid-19-Pandemie zurückzuführen. Im ersten Pandemiejahr wurden 3.905 Prüfungen in der sogenannten fixierten Version (vgl. Kapitel 2.3.3) durchgeführt, die auch in den Folgejahren auf Grund der pandemischen Lage an Bedeutung gewonnen hat. 2021 wurden 14.021, davon 5.851 Online-Prüfungen in der fixierten Version, abgelegt; 2022 waren es 13.312, davon 5.643 fixierte Online-Prüfungen. Seit Einführung des onSET im Oktober 2006 haben insgesamt 266.677 Online-Prüfungen zur Spracheinstufung stattgefunden.

2.3.3 onSET-Erweiterungen

Der onSET wurde im Oktober 2006 weltweit eingeführt, damals noch unter der Bezeichnung „onDaF – Online-Einstufungstest Deutsch als Fremdsprache“. Anfang 2015 wurde eine Version für Englisch als Fremdsprache freigeschaltet. Die englische Sprachversion onSET-English besteht wie die deutsche Sprachversion onSET-Deutsch aus acht Lückentexten, die den Teilnehmenden im LOFT-Verfahren (LOFT = „linear-on-the-fly-testing“) zur Bearbeitung an lizenzierten Testzentren unter Beaufsichtigung dargeboten werden. Bei diesem Verfahren erhält jede(r) Teilnehmende eine andere Zusammenstellung von Texten, die automatisiert aus einer kalibrierten Itembank entnommen werden; dabei ist sichergestellt, dass die zu bearbeitenden Texte für alle Teilnehmenden im Mittel gleich schwer sind. Versionen des onSET, die diesem Modus der Testdarbietung folgen, werden als „Standardversionen“ bezeichnet.

Seit seiner Einführung haben einzelne Testzentren den Wunsch geäußert, Teilnehmenden die Möglichkeit einzuräumen, den onSET zu Hause ohne Aufsicht abzulegen. Das Interesse an einer solchen Form der Testdurchführung hat sich mit Beginn der Corona-Pandemie im Jahr 2020 noch deutlich verstärkt. Um diesem Wunsch bzw. Interesse zu entsprechen, hat g.a.s.t. schon früh damit begonnen, so genannte „fixierte Versionen“ des onSET zur Verfügung zu stellen (dies war insbesondere bei den kostenlosen onSET-Angeboten für Geflüchtete in den Jahren 2016 bis 2019 der Fall). Bei einer fixierten onSET-Version bearbeiten alle Teilnehmenden dieselben Texte in derselben Reihenfolge, ohne dabei beaufsichtigt zu werden. Die für diesen Zweck verwendeten Texte werden nach den gleichen Kriterien, die auch für die Standardversion gelten, aus der onSET-Itembank entnommen und manuell zusammengestellt.

dates, booking of test dates, test processing, result generation and transmission.

2.3.2 Development of test taker numbers

In 2020, 10,929 onSET examinations took place with the German language version (onSET-Deutsch) or with the English version (onSET-English); this marks a decrease of 58% compared to the previous year (with a total of 25,967 examinations). This sharp decline is due to the Covid-19 pandemic. In the first year of the pandemic, 3,905 examinations were carried out in the so-called fixed version (see section 2.3.3), i.e. the unsupervised version taken at home. This version gained in importance in subsequent years due to the pandemic situation. In 2021, 14,021 examinations were taken, of which 5,851 were online examinations in the fixed version; in 2022, 13,312 examinations were taken, of which 5,643 were fixed online examinations. Since the introduction of onSET in October 2006, a total of 266,677 online language placement exams have taken place.

2.3.3 onSET extensions

onSET was launched worldwide in October 2006, at that time under the name “onDaF – Online-Einstufungstest Deutsch als Fremdsprache”. A version for English as a foreign language was released at the beginning of 2015. The English language version onSET-English, like the German language version onSET-Deutsch, consists of eight texts that are presented to participants in the LOFT procedure (LOFT = “linear-on-the-fly-testing”) for completion at licensed test centres under supervision. In this procedure, each participant receives a different set of texts, which are automatically taken from a bank of calibrated items; this ensures that the texts to be worked on are on average equally difficult for all participants. Versions of the onSET following this mode of test presentation are referred to as “standard versions”.

Since its introduction, individual test centres have expressed the desire to give participants the opportunity to take the onSET at home without supervision. Interest in this type of test has increased significantly since the start of the coronavirus pandemic in 2020 (see above 2.3.2.). In order to meet this desire and interest, g.a.s.t. started early on to provide so-called “fixed versions” of the onSET (this was particularly the case with the free onSET product ranges for refugees in the years 2016 to 2019). In a fixed onSET version, all participants work on the same texts in the same order without being supervised. The texts used for this purpose are taken from the onSET item bank and compiled manually according to the same criteria that apply to the standard version.

In Kooperation mit dem Arbeitskreis der Sprachenzentren an Hochschulen (AKS) sind weitere Sprachversionen in Vorbereitung: Zusätzlich zu den deutschen und englischen Sprachversionen wurde im Herbst 2020 damit begonnen, Aufgaben (Lückentexte) für französische, italienische und spanische onSET-Versionen zu entwickeln. Auch für diese drei neuen Sprachen ist vorgesehen, Standardversionen und fixierte Versionen anzubieten.

Zusätzlich wird das onSET-System künftig die Möglichkeit bieten, auch Kurzversionen mit sechs Texten oder Versionen mit nur vier Texten (fokussierte Versionen) bearbeiten zu lassen. Diese decken nicht mehr die gesamten Niveaustufen von A2 bis C1 ab, sondern beschränken sich je nach Anwendungszweck auf einen eingegrenzteren Bereich des GER. Kurzversionen können z.B. die Bereiche A2 bis B2 oder B1 bis C1 erfassen. Fokussierte Versionen erlauben die gezielte Differenzierung zwischen den benachbarten Niveaus A2 und B1 oder B1 und B2 etc. Dies dient einem effizienteren Einsatz bei genauerer Vorkenntnis der Sprachniveaus der Teilnehmenden. Die für die Abgrenzung der GER-Niveaus erforderlichen Cut-Scores werden jeweils auf der Grundlage der Prototypgruppenmethode des Standard-Settings (Eckes 2010, 2017)³ ermittelt.

Zusammenfassend lassen sich die verschiedenen, nach Abschluss der Erweiterungen verfügbaren Versionen des onSET nach vier Dimensionen ordnen: (1) Sprache, (2) Modus, (3) Länge und (4) GER-Bereich. Die folgende Abbildung veranschaulicht die onSET-Systematik. Testzentren werden frei wählen können, welche Version für ihre Zwecke geeignet ist. Sie können aber auch festlegen, dass an ihrer Einrichtung Teilnehmende grundsätzlich nur eine einzige Kombination aus den vier Dimensionen ablegen können, also z.B. nur den onSET-Deutsch als lange Standardversion für die Niveaus A2 bis C1.

In cooperation with the Arbeitskreis der Sprachenzentren an Universitäten (AKS), further language versions are in preparation: In addition to the German and English language versions, work began in autumn 2020 on developing tasks for French, Italian and Spanish onSET versions. There are also plans to offer standard and fixed versions for these three new languages.

In future, the onSET system will also offer the option of editing short versions with six texts or versions with just four texts (focussed versions). These will no longer cover the entire levels from A2 to C1, but will be limited to a narrower range of the CEFR, depending on the intended use. Short versions can, for example, cover the ranges A2 to B2 or B1 to C1. Focused versions allow targeted differentiation between the neighbouring levels A2 and B1 or B1 and B2, etc. This allows for more efficient use with more precise prior knowledge of the participants' language levels. The cut scores required for the differentiation of CEFR levels are determined on the basis of the prototype group method of the standard setting (Eckes 2010, 2017)³.

To summarise, the various versions of onSET available after completion of the extensions can be classified according to four dimensions: (1) language, (2) mode, (3) length and (4) CEFR range. The following diagram illustrates the onSET system. Test centres will be free to choose the version most suitable for their purposes. However, they can also specify that participants at their centre can only take a single combination of the four dimensions, e.g. only the onSET German as the long standard version for levels A2 to C1.



onSET-Versionen onSET-Versions

| Sprache Language | onSET-Deutsch | onSET-English | onSET-Français | onSET-Español | onSET-Italiano | onSET-... |
|---------------------------|-----------------------------------------------------------------|---------------|-----------------------------------------------------------|--------------------------------------------------------------------------------------|--------------------------------------------------------------------|-----------|
| Modus Mode | Standard (LOFT/überwacht) Standard version (LOFT/supervised) | | | Fixiert (ein festes Set/nicht überwacht) Fixed version (fixed set/not supervised) | | |
| Länge Duration | Lang (8 Texte/40 Min.) Long version (8 texts/40 min.) | | Kurz (6 Texte/30 Min.) Short version (6 texts/30 min.) | | Fokussiert (4 Texte/20 Min.) Focussed version (4 texts/20 min.) | |
| GER-Bereich CEFR Range | A2–C1 A1–B2 | | A1–B1 A2–B2 B1–C1 | | A1/A2 A2/B1 B1/B2 B2/C1 | |

³ Eckes, T. (2010). Der Online-Einstufungstest Deutsch als Fremdsprache (onDaF): Theoretische Grundlagen, Konstruktion und Validierung. In R. Grotjahn (Hrsg.), C-Test: Beiträge aus der aktuellen Forschung/The C-test: Contributions from current research (S. 125–192). Frankfurt am Main: Peter Lang.
Eckes, T. (2017). Setting cut scores on an EFL placement test using the prototype group method: A receiver operating characteristic (ROC) analysis. Language Testing, 34(3), 383–411.

2.4 Deutsch-Uni Online

2.4.1 Konzept und Aufbau

Die Deutsch-Uni Online (DUO) bietet Online-Deutschkurse an, die sprachlich, kulturell und propädeutisch auf einen Studien- oder Forschungsaufenthalt an einer deutschsprachigen Hochschule vorbereiten oder studienbegleitend besucht werden können. Im Netzwerk der Deutsch-Französischen Hochschule (DFH) fördert DUO außerdem den Aufbau der notwendigen Kenntnisse und Fertigkeiten in der französischen Wissenschafts- und Fachsprache.

Bei der Entwicklung der DUO-Kurse werden kognitions- und erwerbslinguistischer Forschungsergebnisse und Erkenntnisse einer wissenschaftlich fundierten Medien-, Fach- und Fremdsprachendidaktik berücksichtigt. Seit zwanzig Jahren entstehen so handlungs- und kompetenzorientierte Sprachlernangebote, die die Bedürfnisse von Studierenden in den Mittelpunkt stellen.

2.4 Deutsch-Uni Online

2.4.1 Concept and structure

Deutsch-Uni Online (DUO) offers online German courses to prepare students linguistically, culturally and propaedeutically for a study or research stay at a German-speaking university. Within the network of the Franco-German University (DFH), DUO also promotes the development of the necessary knowledge and skills in French academic and technical language.

The DUO courses are developed on the basis of recent research on cognitive and acquisitional linguistics as well as on specialised and foreign language didactics. For twenty years, this has resulted in practice- and competence-oriented language learning programmes that focus on the needs of students.

Das modular aufgebaute Kursangebot für Deutsch als Fremdsprache ermöglicht strukturiertes Sprachenlernen mit einer durchgehenden Progression von A1 bis C1 des GER. Das Angebot für Französisch als Fremdsprache konzentriert sich auf den Bereich der selbstständigen Sprachverwendung. Lernende können zwischen Kursformaten mit mehr oder weniger intensiver Betreuung durch eine DaF-Lehrkraft wählen oder sich für einen Sprachkurs ganz ohne Tutorierung entscheiden.

Die Betreuung durch DUO-Tutor*innen ist ein Kernelement der Didaktik in DUO-Kursen. Der persönliche Kontakt über verschiedene Feedback- und Nachrichtenkanäle der DUO-Lernumgebung, per E-Mail oder Voice-Chat ermöglicht gezielte ausführliche Rückmeldung zu schriftlichen und mündlichen Sprachleistungen sowie individuelle Lernberatung. Die DUO-Lernumgebung stellt eine Vielzahl an Funktionen und Werkzeugen für die Interaktion und Kooperation zwischen den Lernenden zur Verfügung. Die DUO-Tutor*innen moderieren die synchrone und asynchrone Kommunikation und leiten kooperative Aufgaben in Forum, Chat oder virtuellem Klassenzimmer an.

The modular course programme for German as a foreign language enables structured language learning with a continuous progression from A1 to C1 of the CEFR. The French as a foreign language programme focuses on the area of independent language use. Learners can choose between course formats with more or less intensive support from a German as a foreign language teacher or opt for a language course without any tutoring at all.

The support provided by DUO tutors is a core element of the didactics in DUO courses. Personal contact via various feedback and messaging channels in the DUO learning environment, by email or voice chat enables targeted, detailed feedback on written and oral language performance as well as individualised learning advice. The DUO learning environment provides a variety of functions and tools for interaction and cooperation between learners. The DUO tutors moderate synchronous and asynchronous communication and guide co-operative tasks in the forum, chat or virtual classroom.

Die Flexibilität des Sprachlernangebots ermöglicht den vielseitigen Einsatz der DUO in institutionellen Kontexten. Hochschulen im In- und Ausland, Mittlerorganisationen oder Sprachkursanbieter nutzen DUO kurstragend, kursergänzend oder kursbegleitend, rein online oder hybrid, mit tutorieller Betreuung durch DUO oder durch eigene von DUO für diese Aufgaben geschulte Lehrkräfte. DUO berät und betreut ihre Partner*innen in allen Phasen der E-Learning-Implementierung und bei der Entwicklung maßgeschneiderter Lehr- und Lernpläne. In verschiedenen internationalen Projekten (s. Kap. 2.4.3) unterstützt DUO den Aufbau von Strukturen für den Deutschunterricht im Ausland.

Die Mitarbeitenden der DUO tragen regelmäßig zum wissenschaftlichen Austausch auf nationalen und internationalen Tagungen bei und forschen zum computerunterstützten Fremdsprachenunterricht, zur tutoriellen Betreuung in Online-Kursen sowie zum Mehrwert digitaler Lernangebote im Kontext der Internationalisierung der Hochschulen.

The flexibility of the language learning programme allows the use of DUO in a wide variety of institutional contexts. Universities in Germany and abroad, intermediary organisations or language course providers use DUO to support courses, supplement courses or accompany courses, purely online or hybrid, with tutorial support from DUO or from their own teachers trained by DUO for these tasks. DUO advises and supports its partners in all phases of e-learning implementation and in the development of customised teaching and learning plans. DUO supports the development of structures for teaching German abroad in various international projects (see section 2.4.3). DUO staff regularly contribute to academic exchange at national and international conferences and conduct research into computer-assisted foreign language teaching, tutorial support in online courses and the added value of digital learning opportunities in the context of the internationalisation of universities.



2.4.2 Teilnehmendenentwicklung

Nach einem Höchststand mit knapp 5.800 Lernenden im Jahr 2019 verzeichnete die DUO in den Corona-Jahren von 2020 bis 2022 einen starken Rückgang der Teilnehmendenzahlen.

2020 schlug hier zunächst vor allem die vorzeitige Beendigung des Programms „Idiomas sem Fronteiras (IsF) – Alemão“ (Sprachen ohne Grenzen – Deutsch) in Brasilien zu Buche. Die Teilnehmendenzahl halbierte sich annähernd von gut 1.500 Lernenden 2019 auf knapp 800 Lernende 2020. Der Bedarf in anderen größeren Programmen wie zum Beispiel dem Projekt „Lernen – Lehren – Helfen“ des Instituts für Deutsch als Fremdsprache der Ludwig-Maximilians-Universität für Geflüchtete in Bayern sowie in der Kooperation mit dem Freshman Institute in Aachen zur Vorbereitung internationaler Studierender auf ihren Studienaufenthalt in Deutschland brach ebenfalls ein. Zu Beginn der Pandemie konnte der starke Lernendenschwund in diesen Programmen durch eine verstärkte Nachfrage von DUO-Kursangeboten an deutschen Hochschulen und im Rahmen der Kooperation mit der Deutsch-Französischen Hochschule (DFH) zumindest teilweise aufgefangen werden, sodass sich die Teilnehmendenzahl 2020 auf knapp 5.300 Lernende belief. Nachdem 2021 viele Hochschulen angesichts des andauernden umfassenden Umstiegs von Präsenz- auf Onlinelehre eigene Lehrangebote auch für das Sprachenlernen zusammengestellt hatten, das Programm IsF vollständig eingestellt worden war und sich unter den Studierenden an den Hochschulen eine verbreitete Müdigkeit gegenüber dem Onlinelernen bemerkbar machte, sank die DUO-Teilnehmendenzahl auf einen Tiefstand von knapp 3.300 Lernenden. Dabei stand dem starken Rückgang bei den institutionellen Lernenden ein Zuwachs von rund 25% auf 845 Individuelllernende gegenüber. 2022 erholten sich die Zahlen auf Seiten der Institutionen (3.294). Auch auf Seiten der Individuelllernenden (1.036) konnten weitere Zu-

2.4.2 Development of user numbers

After peaking at just under 5,800 learners in 2019, DUO suffered a sharp decline in the number of participants in the coronavirus years from 2020 to 2022.

In 2020, the early termination of the “Idiomas sem Fronteiras (IsF) – Alemão” (Languages without Borders – German) programme in Brazil was the main factor here. The number of participants almost halved from 1,500 learners in 2019 to just under 800 learners in 2020. Demand in other larger programmes, such as the “Learning – Teaching – Helping” project run by the Institute for German as a Foreign Language at Ludwig-Maximilians-Universität for refugees in Bavaria and the cooperation with the Freshman Institute in Aachen to prepare international students for their studies in Germany, also collapsed. At the beginning of the pandemic, the sharp drop in student numbers in these programmes was at least partially offset by increased demand for DUO courses at German universities and as part of the cooperation with the Franco-German University (DFH), so that the number of participants in 2020 was some 5,300 students. After many universities had put together their own courses for language learning in 2021 in view of the ongoing comprehensive switch from face-to-face to online teaching, the IsF programme was completely discontinued and a widespread fatigue towards online learning became apparent among students at the universities, the number of DUO participants fell to a low of just under 3,300 learners. The sharp decline in institutional learners was offset by an increase of around 25% to 845 individual learners. In 2022, the number of institutional learners recovered (3,294). Further growth was also achieved among individual learners (1,036). In 2022, a total of 4,330 people learnt with DUO. This corresponds to an increase of around 30% compared to 2021.

wächse erzielt werden. Im Jahr 2022 lernten insgesamt 4.330 Personen mit der DUO. Dies entspricht einer Steigerung gegenüber 2021 um rund 30%.

2.4.3 DUO vor Ort – weltweit

Von Juli bis Dezember 2020 lief das vom Auswärtigen Amt geförderte Programm *DUO vor Ort – weltweit* (www.duo-weltweit.de). Das Ziel des Programms war, den Unterricht in Deutsch als Fremdsprache während der Covid-19-Pandemie aufrechtzuerhalten und die Entwicklung digital gestützter Unterrichtsformen zu fördern.

Angesprochen wurden Hochschullehrende im Ortslektorenprogramm des DAAD, denen g.a.s.t. mit der DUO zu diesem Zweck zwei digitale Angebote machte:

1. Schulung und Fortbildung zum digitalen Deutschlehren und -lernen
2. Online-Deutschkurse für Studierende

Die Schulungen und Fortbildungen beinhalteten eine Einführung in die Nutzung und das Unterrichten mit den DUO-Lernmaterialien sowie eine theoretische Fundierung und Reflexion der eigenen digitalen Lehre anhand von Fachtexten, Anwendungsbeispielen, Gruppendiskussion, Transferaufgaben, Unterrichtssimulation und Hospitationen. Das Deutschkursangebot umfasste 16 Module aus dem DUO-Kursportfolio und bediente damit alle Stufen des GER.

Insgesamt nutzten 150 Hochschullehrende (17% der Ortslektor*innen weltweit) mindestens eine der beiden Programmkomponenten. 86 Hochschullehrende nahmen das Schulungs- und Fortbildungsangebot der DUO wahr, 112 Lehrende bestellten DUO-Kursmodule für ihre Studierenden.

33 Länder waren im Programm *DUO vor Ort – weltweit* involviert. Überdurchschnittlich vertreten waren die Regionen Osteuropa/Zentralasien/Südkaucasus, Lateinamerika sowie Nahost, Nordafrika und Nordamerika. Die eher schwache Beteiligung aus Afrika Subsahara und Asien lässt sich zum Teil vermutlich mit einem Mangel an technischer Infrastruktur bzw. Zugangsbeschränkungen und curricularen Restriktionen im Falle von China (ausländische Lehr-/Lernmaterialien nicht erlaubt) erklären.

Die Nachfrage der Ortslektor*innen nach DUO-Kursen zur digitalen Unterstützung des Deutschunterrichts während der Pandemie war erfreulich hoch. Das zur Verfügung stehende Kurskontingent von 6.000 Kursplätzen wurde voll ausgeschöpft und von g.a.s.t. um mehr als 15% aufgestockt, um den Bedarf möglichst aller Ortslektor*innen bedienen zu können. Die DUO-Module wurden sehr unterschiedlich genutzt, vor allem als Zusatzmaterial für den eigenen Unterricht, aber auch als hauptsächliches (kurstragendes) Lernmaterial im Deutschkurs oder als freiwillige Zusatzoption für Studierende.

2.4.3 DUO vor Ort – weltweit

The DUO vor Ort – weltweit (www.duo-weltweit.de) programme funded by the Federal Foreign Office ran from July to December 2020. The aim of the programme was to maintain the teaching of German as a foreign language during the Covid-19 pandemic and to promote the development of digitally supported forms of teaching.

University lecturers in the DAAD's local lecturer programme were offered two digital product ranges with DUO for this purpose:

1. Training and further education on teaching and learning German digitally
2. Online German courses for students

The training and further training sessions included an introduction to using and teaching with the DUO learning materials as well as a theoretical foundation and reflection on their own digital teaching using specialised texts, application examples, group discussions, transfer tasks, lesson simulations and shadowing. The German course programme comprised 16 modules from the DUO course portfolio and thus covered all levels of the CEFR.

A total of 150 university lecturers (17% of local lecturers worldwide) used at least one of the two programme components. 86 university lecturers took advantage of the DUO training and further education programme, 112 lecturers ordered DUO course modules for their students.

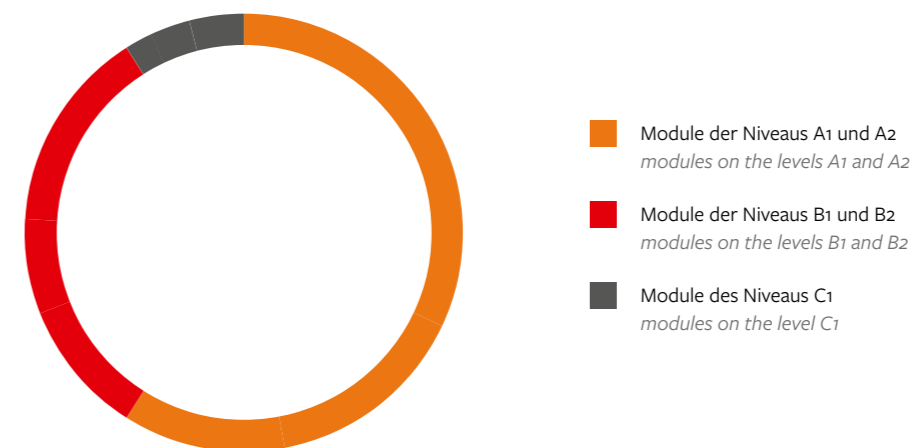
33 countries were involved in the DUO on site – worldwide programme. The regions of Eastern Europe/Central Asia/Southern Caucasus, Latin America, the Middle East, North Africa and North America had an above-average number of representatives. The rather low level of participation from Sub-Saharan Africa and Asia can be explained in part by a lack of technical infrastructure or access restrictions and curricular restrictions in the case of China (foreign teaching/learning materials not permitted).

The demand from local teachers for DUO courses to digitally support German lessons during the pandemic was high. The available course quota of 6,000 course places was fully utilised and increased by 15% by g.a.s.t. in order to meet the needs of as many local teachers as possible. The DUO modules were used in very different ways, primarily as additional material for their own lessons, but also as the main (course-supporting) learning material in the German course or as a voluntary additional option for students.

Knapp die Hälfte (47%) der Studierenden, die am Programm *DUO vor Ort – weltweit* teilnahmen, waren Anfänger*innen auf dem Niveau A1. Die wenigsten Studierenden (2%) lernten mit den DUO-Fachdeutsch-Modulen auf dem Niveau C1.

Just under half (47%) of the students who took part in the DUO vor Ort – worldwide programme were beginners at level A1. The lowest number of students (2%) learnt at C1 level with the DUO specialised German modules.

Verteilung der Lernniveaus im Programm DUO vor Ort – weltweit Distribution of learning levels in the programme DUO vor Ort – weltweit



3 Forschung und Entwicklung

3 Research and development

3.1 Bildungs- und Digitalisierungsprojekte

3.1.1 Digitaler Campus

Das vom BMBF geförderte Projekt „Digitaler Campus“ hat das Ziel, eine Online-Plattform vernetzter Services zu entwickeln, die internationalen Studienbewerbenden als zentrale Anlaufstelle für alle Informationen und Dienstleistungen zum Bewerbungsprozess sowie als Basis zur Studienvorbereitung dient. g.a.s.t ist seit Beginn (1. Februar 2020) an dem Projekt beteiligt. Nach Ende der ersten Projektphase wurde die Finanzierung seitens des BMBF im Mai 2021 bis zum 30.04.2024 verlängert und das Projekt in seinem Fortbestand gesichert.

Mit dem Übergang in die nächste Projektphase hat die TU Berlin das Projektkonsortium verlassen, da ihre Vorarbeit zur Entwicklung eines didaktisch-methodischen Konzepts für Inhalte des Campus in der ersten Phase abgeschlossen werden konnte. Es verbleiben damit neben g.a.s.t. der Deutsche Akademische Austauschdienst (DAAD) als mandatierte Organisation, das Goethe-Institut e.V., die Lernplattform Kiron – Open Higher Education, die Technische Hochschule Lübeck und die RWTH Aachen.

Die Aufgaben für g.a.s.t. haben sich mit Eintritt in die zweite Phase verändert. Diese lassen sich in zwei Bereiche gliedern: die Kuratierung des inhaltlichen Angebots und eine für die spätere technische Umsetzung geeignete Beschreibung der Nutzungspfade internationaler Studierender zum Erwerb von Sprachqualifikation, zur Berufswahl und Bewerbung sowie zur fachlichen Qualifizierung. Diese Aufgaben beinhalten u.a. eine Analyse der geplanten Angebote aus dem Portfolio aller Partner auf noch zu schließender Angebotslücken, Festlegung der Metadatenstrukturen für alle Typen von Angeboten (Fachkurse, Sprachkurse, High-Stakes-Tests, Beratungsangebote etc.), Weiterentwicklung des didaktisch-methodischen Konzepts, Entwicklung und Befüllung von Prozessbeschreibungsmodellen für die unterschiedlichen Nutzungspfade, Kontaktaufnahme zu anderen Bildungsinstitutionen zwecks Eruerung möglichen Datenaustauschs, Koordination mit dem höhergeordneten Projekt BIRD (Bildungsraum Digital) und Zusammenstellung von Qualitätskriterien. Die Umsetzung erfolgt in Arbeitsteams aus den entsprechenden Fachkräften aller Partner entlang etablierter Prozesse des Projektmanagements.

3.1 Education and Digitalisation projects

3.1.1 Digitaler campus

The BMBF-funded “Digital Campus” project aims to develop an online platform of linked services that serves as a central point of contact for international applicants for all information and services relating to the application process and as a basis for preparing for their studies. g.a.s.t has been involved in the project since the beginning (1 February 2020). After the end of the first project phase, funding from the BMBF was extended in May 2021 until 30 April 2024.

With the transition to the next project phase, TU Berlin has left the project consortium, as its preliminary work on developing a didactic-methodological concept for the content of the campus was completed in the first phase. The other consortium members are the German Academic Exchange Service (DAAD) as the mandated organisation, the Goethe-Institut e.V., the learning platform Kiron – Open Higher Education, the Technical University of Lübeck and RWTH Aachen University.

The tasks for g.a.s.t. have changed since entering the second phase. These can be divided into two areas: the curation of the content of the product range and a description of the utilisation paths for international students to acquire language qualifications, career choices and applications, as well as subject-specific qualifications, suitable for later technical implementation. These tasks include analysing the planned product ranges from the portfolios of all partners to identify gaps that still need to be closed, defining the metadata structures for all types of courses (specialist courses, language courses, high-stakes tests, advisory services, etc.), further developing the didactic and methodological concept, developing and filling in process description models for the different usage paths, contacting other educational institutions to determine possible data exchange, coordinating with the twin project BIRD (Bildungsraum Digital) and compiling quality criteria. Implementation is carried out in working teams made up of the relevant specialists from all partners along established project management processes.



3.1.2 Bildungsraum Digital (BIRD)

Im Projekt „Bildungsraum Digital (BIRD)“ wird seit April 2021 ein Prototyp für eine künftige Nationale Bildungsplattform durch einen Projektverbund entwickelt, dem g.a.s.t. angehört. Ein Verlängerungsantrag wurde Ende August 2022 bewilligt, wodurch sich das Projekt bis 31. März 2025 verlängert. Das Vorhaben wird durch die Universität Potsdam koordiniert und hat sich zum Ziel gesetzt, eine technische Infrastruktur zur Vernetzung von bereits bestehenden Bildungsplattformen und digitalen Lehr-, Lern- und Serviceangeboten aller Bildungssektoren zu entwickeln. Damit soll die Machbarkeit einer künftigen nationalen digitalen Plattform, die in allen Phasen des lebensbegleitenden Lernens den Zugang zu Bildung erleichtern soll, demonstriert werden. Diese Plattform wird das Testen von Strukturen für den Datenaustausch, die Interoperabilität von unterschiedlichen Plattformtypen und die Implementierung von Standards ermöglichen.

Die Aufgaben von g.a.s.t. betreffen die Technologie- und Infrastrukturentwicklung. Zudem hat g.a.s.t. die Product Ownership für die Entwicklung und Integration eines Workflow-Management-Systems. Dieses ermöglicht eine am Bedarf der Nutzenden orientierte Bereitstellung und Vermittlung von digitalen Angeboten. Die Grundlage dafür bildet das Metadaten-Management-System von g.a.s.t., das auf einer generischen Metadatenstruktur basiert und eine Vielzahl von Repositorien für unterschiedliche Metadaten (Modulinformationen, Kurse, Prüfungen, Lernobjekte, usw.) bereitstellen kann. Dadurch wird der standard-agnostische Ansatz beim Austausch von Metadaten ermöglicht. Weitere Aufgaben sind die Anbindung der Systeme an eine Data Wallet und die Realisierung von Single-Sign-On-Lösungen für die bestehenden Portallösungen.

In verschiedenen Präsentationen – u.a. im Bundestag vor Abgeordneten und Mitarbeitenden der Ampelfraktionen aus den Ausschüssen für „Bildung, Forschung und Technikfolgenabschätzung“ sowie „Digitales“ und auf der DAAD-Konferenz „Moving Target Digitalisation“ – wurden unter Beteiligung von g.a.s.t. die Grundfunktionen des prototypischen Portals bereits in einer lauffähigen und nutzbaren Version demonstriert.

3.1.3 Profis D – Fach.Deutsch.Digital für Studium und Beruf: ein Verbundprojekt

Die Brandenburgische Technische Universität Cottbus-Senftenberg (BTU) und g.a.s.t. arbeiten in dem von der Stiftung Innovation in der Hochschullehre (StIL) von 2021–2024 geförderten Verbundprojekt „Profis D – Fach.Deutsch.Digital für Studium und Beruf“ zusammen, um die Studienerfolgs-, Berufs- und Bleibechancen internationaler Studierender im englisch- und deutschsprachigen Studium zu fördern. Das Projekt adressiert digitale Lösungsansätze, um das jüngst vom DAAD in seinem Positions-

3.1.2 Digital education space (BIRD)

In the “Bildungsraum Digital (BIRD)” project, a prototype for a comprehensive future national education platform has been under development since April 2021 by a project consortium including g.a.s.t.. An extension application was approved at the end of August 2022, extending the project until 31 March 2025. The project is coordinated by the University of Potsdam and aims to develop a technical infrastructure for networking existing educational platforms and digital teaching, learning and service offerings from all educational sectors. The aim is to demonstrate the feasibility of a future national digital platform that will facilitate access to education in all phases of lifelong learning. This platform will enable the testing of structures for data exchange, the interoperability of different platform types and the implementation of standards.

The tasks of g.a.s.t. relate to technology and infrastructure development. In addition, g.a.s.t. has product ownership for the development and integration of a workflow management system. This enables digital product ranges to be provided and communicated in line with user needs. The basis for this is g.a.s.t.’s metadata management system, which is based on a generic metadata structure and can provide a variety of repositories for different metadata (module information, courses, exams, learning objects, etc.). This enables the standard-agnostic approach to the exchange of metadata. Further tasks are the connection of the systems to a data wallet and the realisation of single sign-on solutions for the existing portal solutions.

The basic functions of the prototype portal have already been demonstrated in an executable and usable version in various presentations – including in the Bundestag in front of members of parliament and staff of the traffic light parliamentary groups from the committees for “Education, Research and Technology Assessment” and “Digital Affairs” and at the DAAD conference “Moving Target Digitalisation” – with the participation of g.a.s.t.

3.1.3 Profis D – Fach.Deutsch.Digital für Studium und Beruf: a joint project

Brandenburg Technical University Cottbus-Senftenberg (BTU) and g.a.s.t. are working together in the joint project “Profis D – Fach.Deutsch.Digital für Studium und Beruf”, funded by the Stiftung Innovation in der Hochschullehre (StIL) from 2021 to 2024, to promote the academic success, career and employment opportunities of international students studying in English and German. The project addresses digital solutions in order to utilise the potential of universities as “migration magnets and

paper Internationale Studierende als Fachkräfte von morgen angesprochene Potenzial der Hochschulen als „Migrationsmagneten und -motoren für die Fachkräftesicherung“ zu nutzen.

Im Rahmen des Projekts entwickeln die beiden Partnerinnen Lehr- und Lernkonzepte sowie Lernmaterialien für den flexiblen Einsatz in der digitalen und digital-gestützten Sprachvermittlung. Diese werden im Projekt nicht nur an der BTU, sondern auch an weiteren Hochschulen erprobt, um sicherzustellen, dass ortsunabhängig einsetzbare Materialien entstehen. Die Erprobungen werden durch eine extern beauftragte Wirkungsstudie begleitet.

g.a.s.t. wird im Projekt bis Sommer 2024 mehrere hundert Stunden Lernmaterial mit Fokus auf akademische Fachsprachen und kommunikative Anforderungen im Beruf entwickeln. Eine Säule des Projekts ist die konsequente Orientierung am Bedarf dieser Handlungsfelder. Hierfür wurden im Wintersemester 2021/22 Studierende, Sprach- und Fachlehrende befragt und sprachliche und fachliche Lernmaterialien analysiert. Zusätzlich wurde in Lehrveranstaltungen verschiedener Fächer hospitiert. Die Ergebnisse dieser Bedarfsanalyse wurden auf der Tagung „Sprache als Schlüssel – wozu? Bildungserfolg im Fokus“ im März 2022 präsentiert und bilden die Grundlage für die Entwicklungen der DUO zu Studierstrategien mit fachlicher Orientierung. Für die Bereiche Technik und Wirtschaft wurde die Aufgabenentwicklung 2022 begonnen.

drivers for securing skilled workers”, as recently mentioned by the DAAD in its position paper International Students as the Skilled Workers of Tomorrow.

As part of the project, the two partners are developing teaching and learning concepts as well as learning material for flexible use in digital and digitally-supported language teaching. These will be piloted not only at BTU but also at other universities to ensure that materials can be used regardless of location. The tests will be accompanied by an externally commissioned impact study.

g.a.s.t. will develop several hundred hours of learning material focussing on academic terminology and communicative requirements in the workplace by summer 2024. One pillar of the project is the consistent orientation towards the needs of these fields of activity. To this end, students, language and subject teachers took part in a survey in winter semester 2021/22. The results of this needs analysis were presented at the conference “Sprache als Schlüssel – wozu? Educational success in focus” in March 2022. They form the basis for DUO’s development of study strategies with a subject-specific focus. The development of tasks for the fields of technology and economics began in 2022.

The development of tasks to prepare students for part-time jobs and internships has also already begun in 2022. These will be offered in the project as two modules “German for university,



Auch mit der Entwicklung von Aufgaben für die Vorbereitung auf Nebenjob und Praktika wurde 2022 bereits begonnen. Diese werden im Projekt als zwei Module „Deutsch für Uni, Job und Alltag“ auf den Stufen B1.1 und B1.2 angeboten und ergänzen das Angebot der DUO künftig durch sprachliche Vorbereitung auf den Berufseinstieg in Deutschland. Eine Bedarfserhebung mit Interviews mit Arbeitgeber*innen in der Region und internationalen Berufseinsteiger*innen zu den sprachlichen Herausforderungen beim Übergang in den Beruf ist die Grundlage der Konzept- und Aufgabenentwicklung für einen daran anschließenden, interdisziplinär gestalteten Brückenkurs für den Berufseinstieg. Die Lernmaterialien für dieses Angebot sollen sowohl die Entwicklung von Strategien für den erfolgreichen Berufseinstieg als auch Kompetenzen für die Kommunikation in berufsübergreifenden, allgemeinen Handlungsfeldern niveauübergreifend für Studierende in deutsch- und englischsprachigen Studiengängen fördern.

Neben Studierenden sind auch Sprach- und Fachlehrende Zielgruppen des Projekts. Sprachlehrende werden durch ein umfangreiches Austausch- und Fortbildungsprogramm nach dem Ad-hoc-Einsatz digitaler Lerntechnologien in den „Corona-Semestern“ für deren Einsatz methodisch qualifiziert. In diesem Rahmen fanden 2021/22 bereits Workshops zum digitalen Fremdsprachenlernen, zum handlungs- und kompetenzorientierten Lehren und Lernen sowie zum (auch digitalen) Prüfen und Testen mit verschiedenen Referent*innen sowie ein digitales Austauschformat für Sprachlehrende zu „Mediendidaktischen Häppchen“ statt.

Im Projekt werden außerdem ein Single-Sign-On-Login zu den Produkten von g.a.s.t. aus den Learning-Management-Systemen der Hochschulen erprobt und ein Webportal pilotiert, das Sprach- und Fachlehrende vernetzt, um Synergien für eine dauerhafte Ausrichtung am Bedarf zu schaffen, und den Grundstein für nachhaltige Zusammenarbeit in der Begleitung internationaler Studierender legt.

Mehr Informationen:
<https://www.b-tu.de/sprachen/projekte/profis-d>

job and everyday life” at levels B1.1 and B1.2 which will supplement the DUO programme in the future with language preparation for a career start in Germany. A needs assessment with interviews with employers in the region and international career starters on the linguistic challenges of the transition to work is the basis for developing the concept and tasks for a subsequent, interdisciplinary bridging course for career entry. The learning material for this product range is intended to promote both the development of strategies for a successful career start and skills for communication in general fields of action across different professions and levels for students on German and English-language degree programmes.

In addition to students, language and subject teachers are target groups of the project. Following the ad hoc use of digital learning technologies in the “coronavirus semesters”, language teachers will receive methodological training for their use through an extensive exchange and further training programme. As part of this programme, workshops on digital foreign language learning, activity- and skills-based teaching and learning as well as (also digital) assessment and testing with various speakers and a digital exchange format for language teachers on “Mediadic Morsels” have already taken place in 2021/22.

The project is also testing a single sign-on login to g.a.s.t. products from the universities’ learning management systems and piloting a web portal that puts together language and subject teachers in order to create synergies for a long-term focus on needs and lay the foundations for sustainable cooperation in supporting international students.

More information:
<https://www.b-tu.de/sprachen/projekte/profis-d>



3.2 TestDaF-Forschung

3.2.1 Online-Standard-Setting beim digitalen TestDaF

Der im Oktober 2020 eingeführte digitale TestDaF hat im Juni 2021 das für den Nachweis seiner Qualität erforderliche Standard-Setting absolviert. Ziel des Standard-Settings war es, die neu konzipierten Aufgaben des digitalen TestDaF daraufhin zu untersuchen, ob sie (a) das Spektrum der TestDaF-Niveaus (TDN) 3 bis 5 zuverlässig abzubilden erlauben und (b) den relevanten GER-Niveaus B2 und C1 zuzuordnen sind. Für die Beantwortung dieser Fragestellung wurden auch die Skalen des GER-Begleitbands (Council of Europe, 2020) herangezogen.

Coronabedingt wurde das Standard-Setting in einer mehrtägigen Veranstaltung von g.a.s.t.-Mitarbeitenden ausschließlich online durchgeführt. Für die Kommunikation im Plenum kam ein Videokonferenz-System zum Einsatz. Testmaterialien sowie Übungen, die darauf abzielten, die Teilnehmenden mit den

3.2 TestDaF research

3.2.1 Online standard setting for the digital TestDaF

The digital TestDaF, which was introduced in October 2020, completed the standard setting required to prove its quality in June 2021. The aim of the standard setting was to establish that the newly designed tasks of the digital TestDaF (a) reliably reflect the range of TestDaF levels (TDN) 3 to 5 and (b) can be assigned to the relevant CEFR levels B2 and C1. The scales of the CEFR companion volume (Council of Europe, 2020) were also used to answer this question.

Due to the coronavirus, the standard setting was carried out exclusively online by g.a.s.t. employees in an event lasting several days. A video conferencing system was used for communication in the plenary session. Test material and exercises aimed at familiarising participants with the CEFR scales were made available on the Moodle platform and in separate software applications.



GER-Skalen vertraut zu machen, wurden in der Moodle-Plattform und in eigenen Software-Anwendungen bereitgestellt.

Teilnehmende waren insgesamt 32 nationale und internationale Expert*innen aus den Bereichen der Testentwicklung und Leistungsbeurteilung. Nach einer Einführung in die neuen Aufgabenformate des digitalen TestDaF und in die verwendeten GER-Skalen wurden zwei Gruppen mit je 16 Expert*innen gebildet.

In zwei aufeinander folgenden Trainingsrunden übte Gruppe 1 die GER-Zuordnung ausgewählter Items im Lesen und Hören nach der Basket-Methode; Gruppe 2 übte die GER-Zuordnung typischer Leistungsbeispiele im Schreiben und Sprechen nach der Benchmarking-Methode (Council of Europe, 2009; Kecker, 2011; Kecker & Eckes, 2010)⁴. Beide Gruppen verwendeten dazu die neunstufige GER-Skala von A1 bis C2 inklusive der Plus-Stufen.

Psychometrische Analysen ergaben ein zufriedenstellendes Maß an Übereinstimmung innerhalb beider Gruppen nach Runde 2 der Trainingsphase. Im anschließenden Standard-Setting hatte Gruppe 1 (Lesen, Hören) Items aus der ersten digitalen TestDaF-Prüfung vom 22. Oktober 2020 wiederum nach der Basket-Methode zuzuordnen. Gruppe 2 (Schreiben, Sprechen) hatte aus dieser Prüfung stammende Leistungsbeispiele analog zum Training auf der GER-Skala einzuordnen.

Die Übereinstimmungs- und Reliabilitätswerte innerhalb beider Gruppen lagen wieder auf einem erfreulich hohen Niveau. Zudem stimmten die TDN-Einstufungen aus der ersten digitalen TestDaF-Prüfung mit den Einstufungen der Expert*innen weitgehend überein. In einer abschließenden Befragung konnten vereinzelt beobachtete Abweichungen auf die besonderen Merkmale des Online-Formats zurückgeführt werden.

A total of 32 national and international experts from the fields of test development and performance assessment took part. After an introduction to the new task formats of the digital TestDaF and the CEFR scales used, two groups of 16 experts each were formed.

In two consecutive training rounds, Group 1 practised the CEFR mapping of selected items in reading and listening using the basket method; Group 2 practised the CEFR mapping of typical performance examples in writing and speaking using the benchmarking method (Council of Europe, 2009; Kecker, 2011; Kecker & Eckes, 2010)⁴. Both groups used the nine-level CEFR scale from A1 to C2, including the plus levels.

Psychometric analyses revealed a satisfactory level of agreement within both groups after round 2 of the training phase. In the subsequent standard setting, group 1 (reading, listening) had to assign items from the first digital TestDaF exam from 22 October 2020, again using the basket method. Group 2 (Writing, Speaking) had to classify performance examples from this exam on the CEFR scale analogue to the training.

The agreement and reliability values within both groups were again at a pleasingly high level. In addition, the TDN classifications from the first digital TestDaF exam largely matched the experts' classifications. In a final survey, isolated deviations observed could be attributed to the special features of the online format.

3.2.2 Konstrukt integrierter Schreibaufgaben

Beim akademischen Schreiben müssen die Studierenden in der Regel Informationen aus verschiedenen Quellen verarbeiten und in ihre eigenen Texte integrieren. Daher werden in Sprachtests für die Hochschulzulassung zunehmend integrierte Schreibaufgaben verwendet. Dabei handelt es sich um Aufgaben, in denen Studierende umfassendes Quellenmaterial inhaltlich und sprachlich in ihren schriftlichen Leistungen verarbeiten müssen. Es ist jedoch vor allem aus der Perspektive der Bewertung unklar, welche Faktoren für die Leistung verantwortlich sind, d.h. inwieweit die Schreibfähigkeit oder Lesekompetenz zur Qualität der schriftlichen Texte und somit letztendlich zu den Testergebnissen beitragen.

Im Rahmen einer Validierungsstudie hat die Leiterin des Bereichs Produktentwicklung bei g.a.s.t., Sonja Zimmermann, daher in ihrer Dissertation⁵ das der integrierten Schreibaufgabe im digitalen TestDaF zugrundeliegende Konstrukt näher beleuchtet. In Anlehnung an neuere Ansätze aus der Sprachtestforschung bediente sie sich dabei eines Mehr-Methoden-Designs, und untersuchte (a) die beteiligten kognitiven Prozesse, (b) die Qualität der schriftlichen Produkte sowie (c) die Zuverlässigkeit der Bewertung der integrierten Leistungen.

Im ersten Teil wurden die kognitiven Prozesse von internationalen Studienbewerber*innen mit einer Kombination aus Eye-Tracking- und Stimulated-Recall-Verfahren erforscht. Die Ergebnisse der quantitativen und qualitativen Daten zeigten, dass die Testteilnehmenden bei der Bearbeitung der integrierten Schreibaufgabe eine Vielzahl von grundlegenden Prozessen des Lesens und Schreibens, aber auch spezifische Lese-Schreib-Prozesse einsetzten.

3.2.2 Construct of integrated writing tasks

Academic writing usually requires students to process information from different sources and integrate it into their own texts. For this reason, integrated writing tasks are increasingly used in language tests for admission to higher education. These are tasks in which students have to process comprehensive source material in terms of content and language in their written performance. However, it is unclear, particularly from an assessment perspective, which factors are responsible for performance, i.e. the extent to which writing ability or reading skills contribute to the quality of the written texts and therefore ultimately to the test results.

As part of a validation study, Sonja Zimmermann, Head of Product Development at g.a.s.t., therefore took a closer look at the construct underlying the integrated writing task in the digital TestDaF in her thesis⁵. In line with more recent approaches from language test research, she used a multi-method design and investigated (a) the cognitive processes involved, (b) the quality of the written products and (c) the reliability of the assessment of integrated performance.

In the first part, the cognitive processes of international applicants were investigated using a combination of eye-tracking and stimulated recall methods. The results of the quantitative and qualitative data showed that the test participants used a variety of basic reading and writing processes as well as specific reading and writing processes when completing the integrated writing task.

In the second part of the thesis, written performances of test takers were analysed to investigate the relevance and accuracy of the information taken from the two sources as well as the lin-

⁴ Council of Europe. (2009). *Relating language examinations to the "Common European Framework of Reference for Languages: Learning, Teaching, Assessment" (CEFR): A manual*. Council of Europe/Language Policy Division.

Council of Europe. (2020). *Gemeinsamer europäischer Referenzrahmen für Sprachen: Lernen, lehren, beurteilen. Begleitband*. Klett Sprachen.

Kecker, G. (2011). *Validierung von Sprachprüfungen: Die Zuordnung des TestDaF zum Gemeinsamen europäischen Referenzrahmen für Sprachen*. Peter Lang.

Kecker, G., & Eckes, T. (2010). Putting the manual to the test: The TestDaF-CEFR linking project. In W. Martyniuk (Ed.), *Aligning tests with the CEFR: Reflections on using the Council of Europe's draft manual* (pp. 50–79). Cambridge University Press.

⁵ Zimmermann, S. (2022). *Validating integrated writing tasks – A mixed-method approach to investigate the construct of summarization*. [Dissertation, Universität Bremen].

Im zweiten Teil der Dissertation wurden schriftliche Leistungen von Testteilnehmenden analysiert, um die Relevanz und Genauigkeit der aus den beiden Quellen aufgenommenen Informationen sowie die sprachliche Transformation des Eingangsmaterials zu untersuchen. Die Ergebnisse zeigten nicht nur Unterschiede zwischen Teilnehmenden auf verschiedenen Leistungsniveaus, sondern auch Unterschiede bei der Wiedergabe von Informationen zwischen dem schriftlichen Quelltext und dem grafischen Input.

Im dritten Teil der Dissertation wurden unter Anwendung der Multifacetten-Rasch-Analyse Daten von Prüflingen und Beurteilenden analysiert. Die Ergebnisse zeigten, dass die Beurteilenden in der Lage waren, die Bewertungsskala wie beabsichtigt anzuwenden und konsistent zu bewerten.

Die umfassende Datenanalyse sowie die Ergebnisse in allen drei empirischen Studien der Dissertation tragen somit zur Validierung der integrierten Schreibaufgabe im digitalen TestDaF bei.

3.2.3 Teilnehmendenbefragung

Die Einführung von digitalen Prüfungsformaten wird oftmals durch Untersuchungen zur Vergleichbarkeit der kognitiven Prozesse und der Prüfungsergebnisse (im Sinne von Test Scores) zwischen dem neuen digitalen Format und traditionellen papierbasierten Prüfungen begleitet. Solche Studien berücksichtigen zwar meist individuelle Faktoren wie Vertrautheit mit dem Computer oder Tippgeschwindigkeit, wie aber Prüfungsteilnehmende die Anforderungen der Testaufgaben oder die Benutzeroberfläche wahrnehmen, spielt eine untergeordnete Rolle – wenn sie überhaupt untersucht werden.

g.a.s.t. hat daher in der Entwicklungsphase des digitalen TestDaF sowie in den ersten beiden Jahren nach Einführung der Prüfung Teilnehmendenbefragungen durchgeführt, um mehr über die Perspektive der Lernenden beim Übergang von einem papierbasierten zu einem digitalen Testformat zu erfahren. In der Entwicklungsphase des digitalen TestDaF lag der Fokus der Befragung zunächst auf der Wahrnehmung der neuen Aufgabenformate und der digitalen Durchführung. Daten von über 1.400 Teilnehmenden zeigten, dass die digitale Testumgebung und die Darbietung der Aufgaben als sehr benutzerfreundlich wahrgenommen wurden. Die Rückmeldungen aus diesen Befragungen flossen teilweise in die Optimierung der Prüfungssoftware und -durchführung ein. Darüber hinaus wurde aber auch deutlich, dass die Anforderungen der neuen Aufgabenformate eine Herausforderung für die Teilnehmenden darstellten.

In den Teilnehmendenbefragungen nach Einführung der Prüfung wurde gezielt danach gefragt, wie Lernende sich auf den digitalen TestDaF vorbereiten und wie sie die unterschiedlichen Ressourcen für die Prüfungsvorbereitung bewerten. Eine Auswertung dieser Daten (bisher knapp 900 vollständige Datensätze) zeigte zum einen, dass sich viele Teilnehmende in relativ

guistic transformation of the input material. The results showed not only differences between participants at different proficiency levels, but also differences in the reproduction of information between the written source text and the graphical input. In the third part of the thesis, data from examinees and raters were analysed using multifaceted Rasch analysis. The results showed that the assessors were able to apply the rating scale as intended and rate it consistently.

The comprehensive data analysis as well as the results in all three empirical studies of the dissertation thus contribute to the validation of the integrated writing task in the digital TestDaF.

3.2.3 Participant survey

The introduction of digital examination formats is often accompanied by studies on the comparability of cognitive processes and examination results (in terms of test scores) between the new digital format and traditional paper-based examinations. Although such studies usually take into account individual factors such as familiarity with the computer or typing speed, how test takers perceive the requirements of the test tasks or the user interface plays a subordinate role – if they are analysed at all.

g.a.s.t. therefore conducted participant surveys during the development phase of the digital TestDaF and in the first two years following the introduction of the exam to understand learners' perspectives on the transition from a paper-based to a digital test format. During the development phase of the digital TestDaF, the survey initially focused on the perception of the new task formats and the digital implementation. Data from over 1,400 participants showed that the digital test environment and the presentation of the tasks were perceived as very user-friendly. The feedback was incorporated into the optimisation of the examination software and administration. Still, the requirements of the new task formats seem to pose a challenge for the participants.

The participant surveys after the introduction of the exam asked learners about their preparation for the digital TestDaF and the various resources for exam preparation. An analysis of this data (almost 900 complete data sets to date) showed that many participants prepare for the exam in a relatively short time (1–2 weeks: 26.1%, 3–4 weeks: 30.6%). Secondly, participants particularly favoured material that specifically prepare them for the exam format and are available as self-study material. Further systematic analysis of the data should provide information on how preparation for the digital TestDaF can be further improved.

kurzer Zeit auf die Prüfung vorbereiten (1–2 Wochen: 26,1%, 3–4 Wochen: 30,6%). Zum anderen bevorzugten Teilnehmende vor allem solche Materialien, die gezielt auf das Prüfungsformat vorbereiten und als Selbstlernmaterial zur Verfügung stehen. Die weitergehende systematische Analyse der Daten soll Aufschluss darüber geben, wie die Vorbereitung auf den digitalen TestDaF weiterhin verbessert werden kann.

3.2.4 Evaluation der Vorbereitungs-materialien für den digitalen TestDaF

Die Teilnehmendenbefragung während der Pilotierung für den digitalen TestDaF (s. Kapitel 3.2.3) zeigte u.a., dass die neuen Aufgabenformate von den Erprobungsteilnehmenden als herausfordernd wahrgenommen wurden. Die größte Schwierigkeit bestand laut Angaben der internationalen Studienbewerber*innen für viele darin, ihre bestehenden (Testbearbeitungs-)Strategien auf die neuen Aufgabentypen zu übertragen.

Um daher zukünftig Teilnehmende zielgerichtet auf die Prüfung und somit auf das Studium in Deutschland vorzubereiten, wurden bei g.a.s.t. ein Vorbereitungskonzept sowie umfassende Lehr- und Lernmaterialien entwickelt (s. Kapitel 2.2). Diese Materialien wurden von g.a.s.t. hinsichtlich ihrer Nützlichkeit für die Testvorbereitung untersucht. Dazu wurden Interviews mit Lehrkräften geführt, die in TestDaF-Vorbereitungskursen die von g.a.s.t. entwickelten Materialien eingesetzt hatten.

Die Auswertung der Interviews zeigte, dass die Lehrkräfte sich allein durch das neue Aufgabenformat des digitalen TestDaF eine positive Rückwirkung auf den Unterricht versprochen. Die noch stärkere Abbildung der Sprachverwendung in den ziel-

3.2.4 Evaluation of the preparation materials for the digital TestDaF

The participant survey during the piloting of the digital TestDaF (see chapter 3.2.3) showed, among other things, that the new task formats were perceived as challenging. According to the international applicants, the greatest difficulty was transferring their existing (test-taking) strategies to the new task types.

In order to prepare future participants for the test and thus for studying in Germany, g.a.s.t. has developed a preparation concept and comprehensive teaching and learning materials (see chapter 2.2). These resources were analysed by g.a.s.t. in terms of their usefulness for test preparation. To this end, interviews were conducted with teachers who had used the materials developed by g.a.s.t. in TestDaF preparation courses.

The evaluation of the interviews showed that the teachers expected the new task format of the digital TestDaF alone to have a positive impact on teaching. The even stronger mapping of language use in the target language communication situations at the university in the test format has the potential to interlink exam and study preparation even more closely and to no longer simply teach to the test. According to the teachers, the



sprachlichen Kommunikationssituationen an der Hochschule im Testformat habe das Potenzial, Prüfungs- und Studienvorbereitung noch enger miteinander zu verzahnen und kein reines teaching to the test mehr zu betreiben. Die von g.a.s.t. entwickelten Materialien zur Prüfungsvorbereitung unterstützen laut Aussage der Lehrkräfte diesen Ansatz und tragen ebenfalls dazu bei, Teilnehmende für die Anforderungen der Testaufgaben und den Zusammenhang mit den sprachlichen Anforderungen im Hochschulkontext zu sensibilisieren.

3.2.5 Simultane Messung von Strenge- und Zentraltendenzen

Bei Leistungsbeurteilungen, etwa bei Beurteilungen der schriftlichen und mündlichen Leistung von Teilnehmenden beim TestDaF, zeigt sich immer wieder, dass Beurteilende in unterschiedlichem Maße streng bzw. milde urteilen. Dies gilt auch dann, wenn sie umfangreiche Schulungen durchlaufen haben.

Eine andere Tendenz, die in der Forschung vermehrt Aufmerksamkeit auf sich gezogen hat, ist die Zentraltendenz (oder auch Zentralität). Damit ist die Tendenz gemeint, mittlere Kategorien einer Ratingskala bevorzugt zu verwenden. Beurteilende, die dieser Tendenz unterliegen, unterschätzen die Fähigkeit von leistungsstarken und überschätzen die Fähigkeit von leistungsschwachen Teilnehmenden.

Übliche psychometrische Modelle, hierzu zählen insbesondere Multifacetten-Rasch-Modelle, erlauben die direkte Messung von Strengetendenzen. Im Unterschied hierzu lassen sich Zentraltendenzen anhand solcher Modelle nur indirekt erfassen. Ein erweitertes Multifacetten-Rasch-Modell, das die simultane Messung von Strenge und Zentralität erlaubt, ist das „Facets Model – Severity and Centrality“ (FM-SC) von Jin und Wang (2018)⁶.

In einer Reihe von Studien haben Thomas Eckes und Kuan-Yu Jin (Hong Kong Examinations and Assessment Authority, Hongkong) das FM-SC zur Analyse von Strenge- und Zentraltendenzen bei TestDaF-Beurteilungen schriftlicher und mündlicher Leistungen eingesetzt (Eckes & Jin, 2021a, 2021b)⁷. Die Ergebnisse zeigten, dass die Beurteilenden präzise auf den separaten Dimensionen der Strenge und Zentralität angeordnet werden konnten. Zudem lieferten die Studien Belege dafür, dass die gemeinsame Berücksichtigung von Strenge- und Zentralitätseffekten zu einer größeren Fairness bei der Reihung der Teilnehmenden nach ihren Fähigkeiten führte.

exam preparation resources developed by g.a.s.t. support this approach and also help to sensitise participants to the requirements of the test tasks and the connection with the language requirements in the university context.

3.2.5 Simultaneous measurement of severity and central tendencies

When assessing performance, for example when assessing the written and oral performance of TestDaF participants, it has been shown that assessors judge with varying degrees of rigour or leniency. This is true even if they have undergone extensive training.

Another tendency that has attracted increased attention in research is the central tendency (or centrality) in marking. Assessors inclining to this tendency underestimate the ability of high-performing participants and overestimate the ability of low-performing participants.

Conventional psychometric models, in particular multifaceted Rasch models, allow the direct measurement of strictness tendencies. In contrast, central tendencies can only be measured indirectly using such models. The “Facets Model – Severity and Centrality” (FM-SC) by Jin and Wang (2018)⁶ is an extended multifaceted Rasch model that allows the simultaneous measurement of severity and centrality.

In a series of studies, Dr. Thomas Eckes and Kuan-Yu Jin (Hong Kong Examinations and Assessment Authority, Hong Kong) used the FM-SC to analyse rigour and centrality tendencies in TestDaF assessments of written and oral performance (Eckes & Jin, 2021a, 2021b)⁷. The results showed that the raters could be accurately placed on the separate dimensions of rigour and centrality. In addition, the studies provided evidence that the joint consideration of rigour and centrality effects led to greater fairness in the ranking of participants according to ability.

3.2.6 Statistische Indikatoren der Zentraltendenz

Unter der Zentraltendenz versteht man den Fehler von Beurteilenden, wenig zu differenzieren und die mittleren Kategorien einer Ratingskala zu bevorzugen. Beurteilende, die dieser Tendenz unterliegen, unterschätzen die Fähigkeit von leistungsstarken und überschätzen die Fähigkeit von leistungsschwachen Teilnehmenden.

Um festzustellen, ob und in welchem Maße Beurteilende einen Zentralitätseffekt aufweisen, sind in der Forschungsliteratur mehrere statistische Indikatoren vorgeschlagen worden. Der einfachste Indikator ist die Standardabweichung der Verteilung von beobachteten Werten (Ratings), die sich individuell ermitteln lässt. Eine ausgeprägte Zentraltendenz wäre mit einer sehr niedrigen Standardabweichung verbunden. Andere Indikatoren verlangen die Anwendung psychometrischer Modelle. Es handelt sich in der Regel um Modelle, die darauf ausgerichtet sind, mit der Strenge von Beurteilenden einen einzigen Beurteilungsparameter zu schätzen. Aus den Ergebnissen der psychometrischen Analyse werden Maße der Zentraltendenz abgeleitet.

Kuan-Yu Jin und Thomas Eckes haben die Eignung verschiedener Indikatoren für Zentraltendenzen in Simulationsstudien und in einer Studie mit realen TestDaF-Daten vergleichend untersucht (Jin & Eckes, 2022)⁸. Dazu haben sie das in Kapitel 3.2.5 skizzierte psychometrische Modell zur simultanen Messung von Strenge- und Zentraltendenzen zugrunde gelegt. Die Ergebnisse zeigten, dass neben zwei modellbasierten Indikatoren die oben beschriebene Standardabweichung vorliegende Zentraltendenzen zuverlässig aufdecken konnte. Allerdings ist dieser Indikator den modellbasierten Indikatoren deutlich unterlegen, wenn für einzelne Beurteilende nur eine vergleichsweise schwache Datenbasis vorliegt.

3.2.7 Analyse von differenziellen Beurteilungsfunktionen

Beurteilungen können Einflüssen von Faktoren unterliegen, die nicht mit der zu messenden Fähigkeit der Teilnehmenden in Verbindung stehen. Zu solchen fähigkeitsirrelevanten Faktoren können z.B. das Geschlecht, das Alter oder andere demografische Merkmale der Teilnehmenden zählen. Aber auch die Vertrautheit der Beurteilenden mit der Urteilsituation, ihre Strenge- oder Zentraltendenzen, motivationale Faktoren oder Stimmungseinflüsse fallen hierunter. Zusammen werden diese unerwünschten und die Validität bzw. Fairness der Beurteilungen gefährdenden Urteilstendenzen als differenzielle Beurteilungsfunktionen („differential rater functioning“, DRF) bezeichnet.

3.2.6 Statistical indicators of central tendency

Several statistical indicators have been proposed in the research literature to determine whether and to what extent raters exhibit a centrality effect. The simplest indicator is the standard deviation of the distribution of observed values (ratings), which can be determined individually. A pronounced central tendency would be associated with a very low standard deviation. Other indicators require the use of psychometric models. These are usually models that are designed to estimate a single assessment parameter with the rigour of assessors. Measures of central tendency are derived from the results of the psychometric analysis.

Kuan-Yu Jin and Dr. Thomas Eckes analysed the suitability of various indicators for central tendencies in simulation studies and in a study with real TestDaF data (Jin & Eckes, 2022)⁸. They used the psychometric model outlined in Chapter 3.2.5 for the simultaneous measurement of rigour and central tendencies. The results showed that, in addition to two model-based indicators, the standard deviation described above was able to reliably detect existing central tendencies. However, this indicator is clearly inferior to the model-based indicators if only a comparatively weak database is available for individual assessors.

3.2.7 Analysis of differential judgement functions

Assessments can be influenced by factors that are not related to the participant’s ability. Such factors may include the participant’s gender, age or other demographic characteristics. However, they also include the assessor’s familiarity with the judgement situation, their strict or central tendencies, motivational factors or mood influences. Together, these undesirable judgement tendencies that jeopardise the validity or fairness of the assessments are referred to as differential rater functioning (DRF).

⁶ Jin, K.-Y. & Wang, W.-C. (2018). A new facets model for rater’s centrality/extremity response style. *Journal of Educational Measurement*, 55(4), 543–563.

⁷ Eckes, T. & Jin, K.-Y. (2021a). Examining severity and centrality effects in TestDaF writing and speaking assessments: An extended Bayesian many-facet Rasch analysis. *International Journal of Testing*. Advance online publication.

Eckes, T. & Jin, K.-Y. (2021b). Measuring rater centrality effects in writing assessment: A Bayesian facets modeling approach. *Psychological Test and Assessment Modeling*, 63(1), 65–94.

⁸ Eckes, T. & Jin, K.-Y. (2022). Detecting illusory halo effects in rater-mediated assessment: A mixture Rasch facets modeling approach. *Psychological Test and Assessment Modeling*, 64(1), 87–111.

Auf der Grundlage des in Kapitel 3.2.5 skizzierten psychometrischen Modells (Jin & Wang, 2018)⁹ haben Kuan-Yu Jin und Thomas Eckes ein Multifacetten-Rasch-Modell entwickelt, das DRF-Effekte hinsichtlich Strenge- und Zentraltendenzen simultan zu messen erlaubt. Anders als bei bislang verwendeten Modellen ist es mit dem „Dual DRF Facets Model“ (DDRFM) möglich, für einzelne Beurteilende zu prüfen, ob und in welchem Umfang sie DRF-Effekte in ihrer Strenge, ihrer Zentralität oder gleichzeitig in beiden Arten von Tendenzen aufweisen.

In Simulationsstudien konnten die Autoren zeigen, dass das DDRFM vorhandene Dual-DRF-Effekte zuverlässig aufdeckte. In einer weiteren Studie mit Daten aus dem schriftlichen Teil einer TestDaF-Prüfung ließen sich anhand dieses Modells bei einzelnen Beurteilenden geschlechtsabhängige DRF-Effekte in beiden Urteilstendenzen nachweisen. So haben die betroffenen Beurteilenden z.B. männliche Teilnehmende strenger bewertet als weibliche Teilnehmende oder bei männlichen Teilnehmenden häufiger die mittleren Kategorien der TDN-Skala verwendet als bei weiblichen Teilnehmenden. Ferner zeigt sich, dass die Berücksichtigung der dualen DRF-Effekte zu einer größeren Fairness bei der Einstufung der Fähigkeit führte als bei alleiniger Betrachtung von DRF-Effekten hinsichtlich der Strenge.

Based on the psychometric model outlined in Chapter 3.2.5 (Jin & Wang, 2018)⁹, Kuan-Yu Jin and Dr. Thomas Eckes have developed a multifaceted Rasch model that allows DRF effects to be measured simultaneously with regard to severity and central tendencies. Unlike previously used models, the “Dual DRF Facets Model” (DDRFM) makes it possible to examine whether and to what extent individual judgements show DRF effects in their severity, their centrality or both types of tendencies simultaneously.

In simulation studies, the authors were able to show that the DDRFM reliably revealed existing dual-DRF effects. In another study with data from the written part of a TestDaF exam, this model was used to demonstrate gender-dependent DRF effects in both judgement tendencies for individual assessors. For example, the assessors concerned rated male participants more strictly than female participants or used the middle categories of the TDN scale more frequently for male participants than for female participants. Furthermore, it can be seen that taking the dual DRF effects into account led to greater fairness in the categorisation of ability than when considering DRF effects alone with regard to rigour.



3.3 TestAS-Forschung und -Entwicklung

3.3.1 Vorhersage von Studienerfolg internationaler Studierender (LIMES)¹⁰

Zentraler Untersuchungspunkt der LIMES-Studie war die Analyse der Vorhersagekraft unterschiedlicher Schulabschlüsse (auch Abitur 2020 vs. Abitur 2021) für den Studienerfolg im Vergleich zum Studieneignungstest TestAS. Einen weiteren Schwerpunkt stellte die Rolle nicht-leistungsbezogener Faktoren als Prädiktoren für Studienerfolg dar, auch im Zusammenspiel mit den kognitiven Faktoren (Schulabschlussnote und kognitive Tests). Das Hauptaugenmerk lag hierbei auf der Vorhersage von Änderungen des Studienstatus (bzw. den Überlegungen, eine Änderung zu vollziehen, fortan Änderungsgedanken) als Grundlage eines Monitoring-Tools zur Früherkennung von Schwierigkeiten im Studium. Zusätzlich

3.3 TestAS research and development

3.3.1 Predicting the academic success of international students (LIMES)¹⁰

The central focus of the LIMES study was to analyse the predictive power of different school-leaving qualifications (including Abitur 2020 vs. Abitur 2021) for academic success in comparison to the TestAS aptitude test. Another focus was on the role of non-performance-related factors as predictors of academic success, also in interaction with cognitive factors (school-leaving grade and cognitive tests). The main focus here was on the prediction of changes in study status (or thoughts of making a change, henceforth change thoughts) as the basis of a monitoring tool for the early detection of difficulties in studies. In addition, students' perceived pandemic-related restrictions in their studies and their effects were analysed.

⁹ Jin, K.-Y. & Wang, W.-C. (2018). A new facets model for rater's centrality/extremity response style. *Journal of Educational Measurement*, 55(4), 543–563.
Jin, K.-Y. & Eckes, T. (2022). Detecting rater centrality effects in performance assessments: A model-based comparison of centrality indices. *Measurement: Interdisciplinary Research and Perspectives*, 20(4), 228–247.

¹⁰ www.gast.de/limes



wurden die von den Studierenden empfundenen pandemiebedingten Einschränkungen im Studium und deren Auswirkungen untersucht.

Das Projekt LIMES wurde aus Mitteln des Bundesministeriums für Bildung und Forschung finanziert. Es wurde vor dem Hintergrund hoher Abbruchquoten insbesondere internationaler Studierender in Deutschland konzipiert. Es fügt sich so in eine Reihe ähnlicher aktueller Studien (insbesondere im Rahmen der Projekte „SeSaBa“ und „SpraStu“) und ergänzt diese um neue Aspekte: den Vergleich bzw. das Zusammenspiel von kognitiven Tests, verschiedenen Schulabschlussnoten und nicht-leistungsbezogenen Variablen als Prädiktoren von Studienerfolg.

Die Studienteilnehmenden (internationale und nationale Sekundarschulabsolvent*innen bzw. Studienanfänger*innen in zwei Kohorten) nahmen an drei Online-Befragungen im ersten Studienjahr (Beginn Semester 1, Ende Semester 1, Ende Semester 2) und an einem selbst-konzipierten digitalen kognitiven Test (LIMES-Test) teil. Zudem lagen für die internationale Gruppe die Ergebnisse der TestAS-Prüfung vor. Beide Tests dienten in der Studie als Maß kognitiver Fähigkeiten. Der Fragebogen ist inhaltlich angelehnt an Befragungen im Projekt SeSaBa (der SeSaBa-Fragebogen wurde zur Verfügung gestellt).

Die folgenden Ergebnisse zur Vorhersage von Studienerfolg bilden den Kern der Studie:

1. Zur Vorhersage von Semesternoten waren die kognitiven Maße entscheidend. Für die nationalen Studierenden war dies die Hochschulzugangsberechtigung (HZB). Für die internationalen Studierenden war dies teilweise (stärker) die HZB, teilweise ausschließlich LIMES-Test oder TestAS. Die Abbildung auf der nächsten Seite veranschaulicht, für die internationale Gruppe, die Prognosekraft der Ergebnisse im Kerntest des TestAS für die Noten im ersten Semester.

The LIMES project was funded by the Federal Ministry of Education and Research. It was conceived against the backdrop of high drop-out rates, particularly among international students in Germany. It thus fits into a series of similar current studies (in particular as part of the “SeSaBa” and “SpraStu” projects) and supplements these with new aspects: the comparison and interaction of cognitive tests, various school-leaving grades and non-performance-related variables as predictors of academic success.

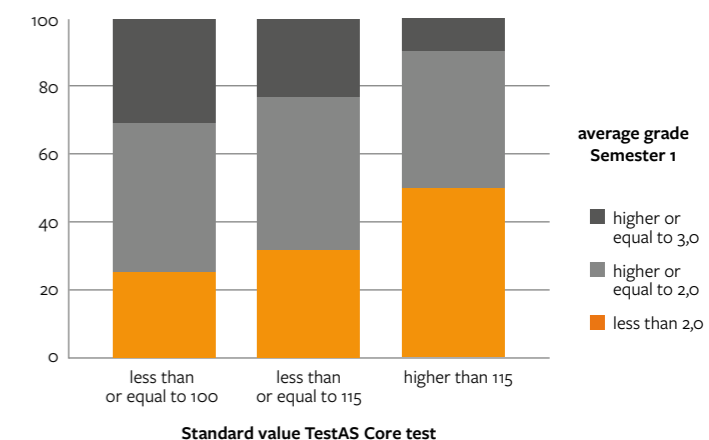
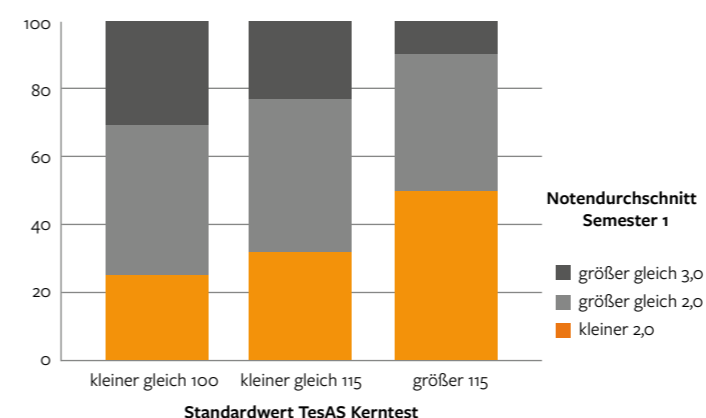
The study participants (international and national secondary school graduates and first-year students in two cohorts) took part in three online surveys in the first year of study (beginning of semester 1, end of semester 1, end of semester 2) and in a self-designed digital aptitude test (LIMES test). The results of the TestAS exam were also available for the international group. Both tests served as a measure of cognitive abilities in the study. The content of the questionnaire is based on surveys in the SeSaBa project (the SeSaBa questionnaire was made available).

The following results on the prediction of academic success form the core of the study:

1. *The cognitive measures were decisive for predicting semester grades. For national students, this was the higher education entrance qualification (Abitur). For the international students, this was partly (more strongly) the HZB, partly exclusively the LIMES test or TestAS. The figure on the next page illustrates, for the international group, the predictive power of the results in the core test of the TestAS for the grades in the first semester.*

2. Zur Vorhersage von Freude im Studium und selbst eingeschätzten Studienfähigkeiten spielten sowohl kognitive (für nationale die HZB; für internationale die HZB und/oder LIMES-Test) als auch nicht-leistungsbezogene Variablen (Anerkennung durch Mitstudierende und Dozierende, Identifikation mit der Hochschule; Kontakt zu Dozierenden und Mitstudierenden) eine Rolle.
3. Für die Vorhersage von Änderungsgedanken spielten die untersuchten nicht-leistungsbezogenen Variablen eine große (internationale Gruppe) oder alleinige (nationale Gruppe) Rolle. In der nationalen Gruppe setzte sich eindeutig die Anerkennung durch Dozierende und Mitstudierende durch. In der internationalen Gruppe war die Identifikation mit der Hochschule der entscheidende Faktor (zusätzlich lieferte hier auch der Kerntest des TestAS eine signifikante Varianzerklärung).

2. *Both cognitive (for national students the Abitur; for international students the HZB and/or LIMES test) and non-performance-related variables (recognition by fellow students and lecturers, identification with the university; contact with lecturers and fellow students) played a role in predicting enjoyment of studies and self-assessed study skills.*
3. *The non-performance-related variables analysed played a major (international group) or sole (national group) role in predicting thoughts of change. In the national group, recognition by lecturers and fellow students clearly prevailed. In the international group, identification with the university was the decisive factor (in addition, the core test of the TestAS also provided a significant explanation of variance here).*



Weitere Ergebnisse umfassten folgende Aspekte:

- Die Bedeutung standardisierter Tests wie dem TestAS für internationale Studierende wurde auch angesichts weniger aussagekräftiger Noten in ausländischen HZBen (verstärkt durch Verzerrungen bei der Umrechnung der Noten in das deutsche Notensystem) deutlich.
- Die wichtigsten Faktoren für Änderungsgedanken waren ein Mangel an Motivation, falsche Vorstellungen vom Studieneinfach, Zweifel an der eigenen Eignung oder Leistungsprobleme. Auch die Unzufriedenheit mit dem Studienort, der Wunsch nach einer praktischen Tätigkeit und persönliche Gründe spielten eine Rolle.
- Studierende im Frühjahr 2021 empfanden eine hohe Belastung durch die überwiegende Onlinelehre und mangelnde soziale Kontakte als Folge der Pandemie. Diese Einschränkungen hatten einen verstärkenden Einfluss auf Änderungsgedanken. Die empfundenen Belastungen nahmen im Laufe der Befragungszeitpunkte konsistent nach und nach ab.
- Internationale Studierende verfolgten offenbar gezielter Strategien zum Lernen und zur Motivierung. Das Bewusstmachen der Wichtigkeit von Noten und das Bewusstsein über Lernziele stand in klarem korrelativem Zusammenhang mit Studienerfolg für nationale und internationale Studierende.

Further results included the following aspects:

- *The relevance of standardised tests such as the TestAS for international students was proven in view of the less meaningful grades in foreign HZBs (exacerbated by distortions in the conversion of grades into the German grading system).*
- *The most important factors for a reconsideration of study choices were a lack of motivation, misconceptions of the degree programme, doubts concerning one’s suitability or performance problems. Dissatisfaction with the place of study, the desire for practical work and personal reasons also played a role.*
- *In spring 2021, students overall perceived a high level of stress due to the predominance of online teaching and a lack of social contact as a result of the pandemic. These restrictions had an increasing influence on thoughts of change. The perceived stress levels gradually and consistently decreased over the course of the survey.*
- *International students apparently pursued more targeted strategies for learning and motivation. Awareness of the importance of grades and awareness of learning goals was clearly correlated with academic success for national and international students.*



3.4 onSET-Forschung und -Entwicklung

3.4.1 Mobiler onSET

Teilnehmende legen den onSET bislang an einem Desktop-PC oder Notebook ab. Die Durchführung ist zwar bereits browserbasiert, doch nur mit ausreichend großem Bildschirm ist sichergestellt, dass die erzielten Ergebnisse valide sind. Teilnehmende sehen den gesamten Lückentext auf einen Blick und füllen die Lücken mit einer physikalischen Tastatur.

In den letzten Jahren ist Verbreitung und Akzeptanz von mobilen Endgeräten mit Touchscreen rasant gewachsen. Entsprechend stieg auch der Wunsch nach einer onSET-Version, die auf diesen Geräten nutzbar ist. g.a.s.t hat daher mit Arbeiten an einer ebenfalls browserbasierten mobilen Version begonnen.

3.4 onSET research and development

3.4.1 Mobile onSET

The onSET was designed to be taken on a desktop PC or notebook as a larger screen size was considered necessary to ensure validity of the results obtained. Only on larger screens, participants can see the entire cloze text at a glance and fill in the gaps using a physical keyboard.

Still, the use and acceptance of mobile devices with touchscreens has grown rapidly in recent years. The desire for an onSET version that can be used on these devices has also increased accordingly. g.a.s.t has therefore started work on a browser-based mobile version for handheld devices.

Zur vorläufigen mobilen Version des onSET für Smartphone oder Tablet stellen sich u. a. folgende Fragen:

- Unterscheiden sich die Messgenauigkeit (Reliabilität) und die Aussagekraft (Validität) der onSET-Ergebnisse an mobilen Endgeräten von denen an Desktop-PCs oder Notebooks?
- Welche besonderen Anforderungen an die technische Umsetzung sind bei der Entwicklung der mobilen onSET-Version zu beachten?

Das Beispiel der virtuellen Tastatur verdeutlicht die Herausforderungen der technischen Umsetzung: Smartphone-eigene Tastaturen schlagen Ergebnisse vor oder ergänzen eigenständig die eingegebenen Buchstaben zu ganzen Wörtern (Autovervollständigung bzw. Autokorrektur). Texte müssen so dargestellt werden können, dass Teilnehmende nicht durch den Text scrollen müssen. Denn vor allem bei höherer Sprachkompetenz erfassen Teilnehmende mehr als nur die jeweilige Lücke, sollten also den ganzen Text im Blick haben können. Und natürlich müssen Messgenauigkeit und Testsicherheit gewährleistet sein.

Ein Prototyp wurde entwickelt. Er erfüllt die genannten Kriterien und wurde in einer g.a.s.t.-internen Testphase erprobt. Eine wissenschaftliche Untersuchung der Messung von Reliabilität und Validität der Messung der Sprachkenntnisse steht noch aus.

3.4.2 Faktor Zeit beim C-Test

Im Mai 2022 startete das von der Deutschen Forschungsgesellschaft (DFG) geförderte Projekt *Zeit Lücken zu schließen: eine multimethodische Studie zur Rolle der Zeit im Konstrukt computergestützter C-Tests in drei Sprachen* (Projektnummer 462766474).

Beim C-Test müssen Sprachlernende Lücken in einem Text füllen. Zur Vervollständigung eines C-Test-Texts haben die Teilnehmenden fünf Minuten Zeit. Das Projektteam um Prof. Dr. Anastasia Drackert untersucht einen neuen zeitreduzierten *speeded C-Test* (S-C-Test), der einen effizienteren Einsatz verspricht, da sich die Bearbeitungszeit pro C-Test und so für den gesamten C-Test halbiert. Dieses Instrument wurde jedoch noch nicht ausreichend untersucht. Es ist daher nicht klar, welchen Einfluss die Verkürzung der Bearbeitungszeit auf das Konstrukt, also die zu messende Fähigkeit, hat. Mit längerer Bearbeitungszeit ausgestattet, misst der kanonische C-Test Grotjahn (2010) zufolge v.a. den Umfang deklarativen und prozeduralen Wissens, während der S-C-Test zudem den Grad der Automatisierung der Fähigkeiten und Effizienz der Informationsverarbeitung erfasst. Dadurch würde der S-C-Test mit Hörverstehen und Sprechen höher korrelieren als der C-Test, aber niedriger mit Schreiben und Lesen. So ist zu vermuten, dass dem S-C-Test ein z.T. abweichendes Konstrukt zugrunde liegt. Obwohl erste Erkenntnisse über die Rolle der Zeit im Konstrukt vorliegen (Fadaeipour & Zohoorian, 2017; Zimmermann, 2019), bleiben Fragen offen, insbesondere bei computergestützten C-Tests, z.B. bezüglich des Einflusses

The following questions arise regarding the provisional mobile version of onSET for smartphones or tablets:

- *Do the measurement accuracy (reliability) and significance (validity) of the onSET results on mobile devices differ from those on desktop PCs or notebooks?*
- *What special technical implementation requirements need to be taken into account when developing the mobile onSET version?*

The example of the virtual keyboard illustrates the challenges of technical implementation: smartphones' own keyboards suggest results or independently complete the letters entered to form whole words (autocomplete or autocorrect). Moreover, texts must be presented in such a way that participants do not have to scroll through the text. This is because, especially with higher language skills, participants recognise more than just the respective gap and should therefore be able to see the entire text. And, of course, measurement accuracy and test reliability must be guaranteed.

A prototype has been developed. It meets the aforementioned criteria and was tested in an internal g.a.s.t. test phase. A scientific study to measure the reliability and validity of the language skills assessment is still pending.

3.4.2 Time factor in the C-test

In May 2022, the German Research Foundation (DFG)-funded project Closing time gaps: a multi-method study on the role of time in the construct of computerised C-tests in three languages (project number 462766474) was launched.

In the C-test, language learners have to fill in carefully designed gaps in a text. Participants have five minutes to complete a C-test text. The project team led by Prof. Dr. Anastasia Drackert is investigating a new time-reduced speeded C-Test (S-C-Test) which promises to be more efficient as the processing time per C-text and thus for the entire C-Test is halved. However, this instrument has not yet been sufficiently analysed. It is therefore not clear what influence the reduction in processing time has on the construct, i.e. the ability to be measured. According to Grotjahn (2010), the canonical C-test, with its longer processing time, primarily measures the scope of declarative and procedural knowledge while the S-C test also measures the degree of automation of skills and efficiency of information processing. As a result, the S-C test would correlate higher with listening comprehension and speaking than the C-test, but lower with writing and reading. It can therefore be assumed that the S-C-test is based on a partially different construct. Although initial findings on the role of time in the construct are available (Fadaeipour & Zohoorian, 2017; Zimmermann, 2019), questions remain unanswered, particularly in the case of computer-based C-tests, e.g. regarding the influence of time on test reliability and learners' performance depending on their language level and regarding

der Zeit auf die Testreliabilität und die Leistung der Lernenden abhängig vom ihrem Sprachniveau sowie hinsichtlich der Rolle verschiedener Komponenten (z.B. des deklarativen Wissens) im Konstrukt zeitlimitierter C-Tests. Ferner fehlen systematische Belege der Auswirkung der Zeit auf kognitive Prozesse, die in den Strategien der Lernenden zum Ausdruck kommen.

Angesichts der Rolle, die Zeit im Konstrukt zu spielen scheint, untersucht das mehrsprachige Projekt dies systematisch, um einen höheren Grad an Verallgemeinerbarkeit der Ergebnisse in Bezug auf 1) unterschiedliche Niveaustufen, 2) verschiedene Sprachen (Englisch, Deutsch, Russisch) und 3) computergestützte C-Tests zu ermöglichen. Hierzu werden vergleichende Analysen von Lernendenleistungen im kanonischen und *speeded-C-Test* durchgeführt und mit weiteren Messungen u.a. des deklarativen und prozeduralen Wissens verglichen. Anhand einer Stichprobe (n=540) von Sprachlernenden an Hochschulen soll die Studie aufzeigen, was unterschiedlich getaktete C-Tests tatsächlich messen, indem sie auf Untersuchungen der Flüssigkeit (*fluidity*) des Konstrukts aufbaut und diese vertieft (Drackert & Timukova, 2020 a&b; Sigott, 2006). Es handelt sich um die erste Studie, die sich mit der Komponentenstruktur des Konstrukts befasst und diese mit Strukturgleichungsmodellen und einer prozessorientierten video-basierten Analyse von Teststrategien untersucht. Ein verbessertes Verständnis des Einflusses der Zeit auf das Konstrukt verspricht großen Nutzen für die Anwendung von C-Tests: Forschende und Lehrende können das Instrument für verschiedene Kontexte anpassen, um begründeter über die Sprachkompetenz von Lernenden entscheiden zu können. Das Instrument wird dadurch präziser, effizienter und ökonomischer.

the role of various components (e.g. declarative knowledge) in the construct of time-limited C-tests. Furthermore, there is a lack of systematic evidence of the effect of time on cognitive processes expressed in learners' strategies.

Given the role that time seems to play in the construct, the multilingual project investigates this systematically in order to enable a higher degree of generalisability of the results with regard to 1) different levels, 2) different languages (English, German, Russian) and 3) computer-based C-tests. To this end, comparative analyses of learner performance in the canonical and speeded C-test will be carried out and compared with other measures of declarative and procedural knowledge, among others. Using a sample (n=540) of language learners at universities, the study aims to show what differently timed C-tests actually measure by building on and deepening investigations into the fluidity of the construct (Drackert & Timukova, 2020 a&b; Sigott, 2006). This is the first study to look at the component structure of the construct and examine it using structural equation modelling and a process-oriented video-based analysis of test strategies. An improved understanding of the influence of time on the construct promises great benefits for the use of C-tests: researchers and teachers can adapt the instrument to different contexts in order to make more informed decisions about learners' language competence. This will make the instrument more precise, efficient and economical.



3.5 DUO-Lernplattform

Die Lernplattform der Deutsch-Uni Online (DUO) wurde in den vergangenen Jahren didaktisch überarbeitet und technisch verbessert, um sie fortlaufend an die Bedürfnisse und Erwartungen der Lernenden anzupassen. 2020 starteten die Planungen für eine komplette Neugestaltung der DUO-Lernplattform mit optimierten Benutzeroberflächen und erweiterten Möglichkeiten für das tutorierte Lernen in der Gruppe sowie die selbstständige Lernorganisation.

Die Benutzeroberfläche für die neue Lernplattform konzipierte g.a.s.t 2020 in Zusammenarbeit mit der Hochschule der Medien (HdM). Auf Grundlage von Interviews im Rahmen einer Online-Umfrage wurden hierfür Lernende der DUO sowie andere Nutzergruppen nach ihren Erlebnissen beim Sprachenlernen befragt. Als besonders wichtig wurden hierbei die Themen „Fortschritt- und Kompetenzerleben“, „soziale Zugehörigkeit“ und „Lernunterstützung“ herausgearbeitet.

3.5 DUO learning platform

The Deutsch-Uni Online (DUO) learning platform has been didactically revised and technically improved in recent years in order to adapt it continuously to the needs and expectations of learners. In 2020, planning began for a complete redesign of the DUO learning platform with optimised user interfaces and expanded options for tutored group learning and independent learning organisation.

The user interface for the new learning platform was designed by g.a.s.t 2020 in collaboration with the Hochschule der Medien (HdM). Based on interviews conducted as part of an online survey, DUO learners and other user groups were asked about their experiences of language learning. The topics “experience of progress and competence”, “social affiliation” and “learning support” were identified as particularly important.

Im Zuge der Konzeptentwicklung wurden zusätzlich prototypische Nutzer*innenprofile (sogenannte Personas) entworfen und typische Szenarien für deren Aktivitäten auf der DUO-Lernplattform beschrieben. Auf dieser Grundlage entwickelte die HdM zwischen Ende 2020 und März 2021 Designvorschläge für die Darstellung von Kursinhalten und Kursorganisation, des Lernens in der Gruppe und des Lernens mit Tutor*in.

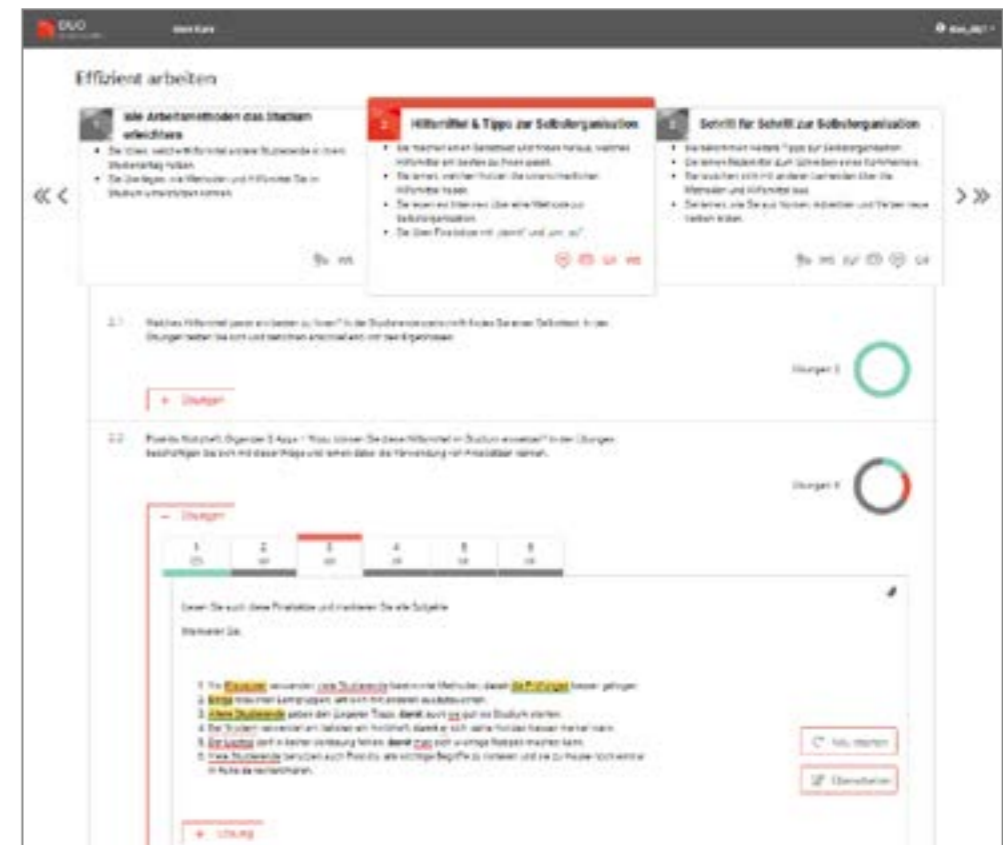
Zwischen April 2021 und Januar 2022 wurden die ersten Designvorschläge umgesetzt und die notwendigen Funktionen für das Lernen ohne Tutor*in technisch entwickelt. Konzeptionell lag der Schwerpunkt während dieser Zeit auf der Darstellung von Lernfortschritt: Welche quantitativen und qualitativen Daten sind für die Lernenden in unterschiedlichen Phasen eines Kurses von Interesse? Wie lassen sich diese Daten technisch ermitteln und an welchen Stellen sollen sie auf der Lernplattform dargestellt werden? Es wurden fünf Darstellungsebenen identifiziert, grafisch entworfen und technisch umgesetzt. Nach der Überprüfung durch Lernexpert*innen wurden die Fortschrittsanzeigen im Februar 2022 mit sechs Proband*innen (ehemaligen DUO-Lernenden) qualitativ erprobt. Die Ergebnisse der Erprobung fielen positiv aus: Die verschiedenen Ebenen der Fortschrittsanzeige wurden als leicht auffindbar, intuitiv verständlich, hilfreich für die Orientierung in allen Phasen eines Kurses und motivierend wahrgenommen.

Weitere Schwerpunkte der konzeptionellen und technischen Arbeit lagen 2022 auf der Optimierung und Standardisierung des Layouts und der Auswertung auf Übungsebene. Eine erste Live-Schaltung der neuen DUO-Lernplattform für das Lernen ohne Tutor*in wird zu Beginn des Jahres 2024 angestrebt.

As part of the concept development, prototypical user profiles were also designed and typical scenarios for their activities on the DUO learning platform were described. On this basis, the HdM developed design proposals for the presentation of course content and course organisation, group learning and learning with a tutor between the end of 2020 and March 2021.

Between April 2021 and January 2022, the first design proposals were implemented and the necessary functions for learning without a tutor were technically developed. Conceptually, the focus during this time was on the presentation of learning progress: What quantitative and qualitative data is of interest to learners in different phases of a course? How can this data be technically determined and where should it be displayed on the learning platform? Five display levels were identified, graphically designed and technically implemented. After being reviewed by learning experts, the progress indicators were tested with six test subjects (former DUO learners) in February 2022. The results of the trial were positive: The different levels of the progress indicator were perceived as easy to find, intuitively understandable, helpful for orientation in all phases of a course and motivating.

In 2022, the conceptual and technical work also focused on optimising and standardising the layout and evaluation at exercise level. The go live of the new DUO learning platform for learning without a tutor is planned for the beginning of 2024.





3.6 Forschungsk Kooperationen

3.6.1 Herder-Institut (Universität Leipzig)

In einem vom BMBF geförderten Projekt untersucht Dr. Katrin Wisniewski den Zusammenhang von Sprache und Studien-erfolg bei Bildungsausländer*innen (gemeinsam mit Prof. Dr. Wolfgang Lenhard, Universität Würzburg). g.a.s.t. beteiligt sich an diesem Projekt nicht nur in beratender Funktion, sondern auch mit dem onSET zur Einstufung der Deutschkenntnisse und mit dem Schriftlichen Ausdruck des TestDaF. Mittels einer Original-TestDaF-Schreibaufgabe sollen die schriftsprachlichen Kompetenzen der Teilnehmenden im Längsschnitt erfasst werden. Darüber hinaus sollen die korpuslinguistischen Auswertungen in intensivem Austausch mit g.a.s.t.-Mitarbeiter*innen für TestDaF-Forschungszwecke nutzbar gemacht werden.

3.6 Research cooperation

3.6.1 Herder Institute (University of Leipzig)

In a project funded by the BMBF, Dr Katrin Wisniewski is investigating the relationship between language and academic success among foreign students (together with Prof. Dr Wolfgang Lenhard, University of Würzburg). g.a.s.t. is involved in this project not only in an advisory capacity, but also with the onSET for the classification of German language skills and with the written expression of the TestDaF. Using an original TestDaF writing task, the written language competences of the participants are to be recorded longitudinally. In addition, the corpus analyses are to be used for TestDaF research purposes in intensive exchange with g.a.s.t. employees.

3.6.2 Sprachtechnologie-Labor (Universität Duisburg-Essen)

Die Kooperation zur automatischen Auswertung schriftlicher Texte im TestDaF mit dem *Language Technology Lab* (LTL; Leitung: Prof. Dr. Zesch) hat im November 2016 zunächst als Austausch von Forschungsdaten begonnen. Der erfolgreiche Verlauf der anschließend vorgenommenen Analysen zu Teilnehmendenantworten in den Prüfungsteilen Hörverstehen und Schriftlicher Ausdruck führte im Dezember 2018 zum Abschluss einer Kooperationsvereinbarung. Gegenstand dieser Vereinbarung war die Durchführung von Untersuchungen im Projekt „Automated and Assisted Scoring of Freetext Answers“. g.a.s.t. finanzierte hierzu eine Stelle für eine*n wissenschaftliche*n Mitarbeiter*in im LTL für zwei Jahre.

Im Rahmen des Projekts wurden z.B. Möglichkeiten einer maschinellen Unterstützung der Bewertung von Teilnehmendenantworten bei den halboffenen Aufgaben im Hörverstehen untersucht. Diese Unterstützung kann etwa in der Identifikation identischer Teilnehmendenantworten und deren Ausschluss aus der Menge der zu bewertenden Antworten bestehen, sobald die Bewertung einer dieser identischen Antworten erfolgt ist. Weitere Formen der Unterstützung betreffen ein Clustering (Vorklassifikation) der Antworten und die Möglichkeit, durch vorformulierte (automatisierte) Bewertungsregeln für Rechtschreibfehler eine Entlastung der Bewertenden und damit eine Beschleunigung des Bewertungsprozesses zu erreichen.

3.6.2 Language Technology Laboratory (University of Duisburg-Essen)

The cooperation with the Language Technology Lab (LTL; Head: Prof. Dr Zesch) for the automatic evaluation of written texts in the TestDaF began in November 2016, initially as an exchange of research data. The success of the subsequent analyses of test taker responses in the Listening Comprehension and Written Expression sections led to the conclusion of a cooperation agreement in December 2018. The purpose of this agreement was to conduct research in the “Automated and Assisted Scoring of Freetext Answers” project. g.a.s.t. was funding a position for a research assistant at LTL for two years.

As part of the project, for example, the possibilities of automated support for the assessment of participants’ answers in the semi-open listening comprehension tasks were being investigated. This support can consist of identifying identical participant responses and excluding them from the set of responses to be evaluated as soon as one of these identical responses has been evaluated. Other forms of support include clustering (pre-classification) of the answers and the possibility of using pre-formulated (automated) assessment rules for spelling errors to reduce the workload of the assessors and thus speed up the assessment process.



4 g.a.s.t.-Akademie

4 g.a.s.t. Akademie

4.1 Seminar für Prüfungsbeauftragte

Das Seminar für Prüfungsbeauftragte bietet jedes Jahr auf dem Campus der Ruhr-Universität Bochum ein zentrales Forum für den Wissens- und Erfahrungsaustausch zwischen g.a.s.t. und den Mitarbeiter*innen lizenzierter g.a.s.t.-Testzentren. Im Zuge der Pandemie hat g.a.s.t. als Alternative zu einem Präsenzseminar 2020 im Zeitraum von Februar bis Mitte März 2021 eine Webinarreihe für Prüfungsbeauftragte veranstaltet.

Die unter dem Titel „g.a.s.t.-Dialog“ durchgeführten Webinare beschäftigten sich mit unterschiedlichen Themenschwerpunkten und dienten der Information sowie der übergreifenden Beratung zu verschiedenen Themen und den g.a.s.t.-Produkten. Die Teilnehmenden erhielten Tipps und Hinweise zur Beratung von TestDaF-Teilnehmenden sowie Empfehlungen zur Prüfungsvorbereitung der Lernenden auf den papierbasierten und den digitalen TestDaF. Der Bereich Marketing und Kommunikation stellte zudem die neuen Informations- und Werbematerialien zu beiden Testformaten und zum Teilnehmendenmarketing an Testzentren vor.

Weitere Themenschwerpunkte waren die Einsatzmöglichkeiten und Erweiterung des onSET, das neue kompetenzerweiternde Modul Uni-Deutsch B1.1 sowie die Trainingsmodule zur TestDaF-Vorbereitung der DUO und die digitale Version des TestAS.

Jeder g.a.s.t.-Dialog bot ausgiebig Gelegenheit zum Austausch und zur Beratung sowie zur Klärung individueller Fragen mit den g.a.s.t.-Mitarbeiter*innen. Die Online-Veranstaltungen wurden gut angenommen und in Zeiten der Pandemie als willkommene Alternative zum jährlich stattfindenden Präsenzseminar gesehen.

4.2 Workshops und Webinare zum Testen, Evaluieren und E-Learning

Die Fortbildungen der g.a.s.t.-Akademie umfassen Workshops und Webinare zu verschiedenen Themenbereichen und vermitteln fachliche und praxisnahe Inhalte aus den Arbeitsschwerpunkten Testen, Evaluieren und E-Learning. Die Themen reichen

4.1 Seminar for examination officers

Every year, the seminar for examination officers offers a central forum for the exchange of knowledge and experience between g.a.s.t. and staff of licensed g.a.s.t. test centres on the campus of Ruhr University Bochum. During the pandemic, g.a.s.t. organised a webinar series for examination officers from February to mid-March 2021 as an alternative to a face-to-face seminar in 2020.

The webinars held under the title „g.a.s.t. dialogue“ focused on various topics and served to provide information and comprehensive advice on various topics and the g.a.s.t. products. Participants received tips and advice on counselling TestDaF participants as well as recommendations on how to prepare learners for the paper-based and digital TestDaF. The Marketing and Communication department also presented the new information and advertising materials for both test formats and for participant marketing at test centres.

Other key topics included the possible uses and expansion of onSET, the new competence-enhancing module Uni-Deutsch B1.1 as well as the DUO training modules for TestDaF preparation and the digital version of TestAS.

Each g.a.s.t. dialogue offered ample opportunity for exchange and advice as well as to clarify individual questions with the g.a.s.t. employees. The online events were well received and, in times of the pandemic, were seen as a welcome alternative to the annual face-to-face seminar.

4.2 Workshops and webinars for Testing, evaluation and e-learning

The training courses offered by g.a.s.t. Academy include workshops and webinars on various topics and convey specialised and practical content from the areas of testing, evaluation and e-learning. The topics range from exam preparation, test and

von der Prüfungsvorbereitung, Test- und Aufgabenerstellung sowie Leistungsmessung bis zur Aufgabenentwicklung für E-Learning-Programme (s. Tab. S. 59).

Im Zuge der Corona-Pandemie hat g.a.s.t. ihr Fortbildungsprogramm digital erweitert. Neben den 60- bis 90-minütigen Webinaren führten die g.a.s.t.-Referent*innen im Jahr 2021 Online-Workshops mit dem Schwerpunkt „Der digitale TestDaF“ durch. Die Workshops sind interaktiv, der Austausch mit den Teilnehmenden spielt eine bedeutende Rolle.

4.3 Seminare nach Maß im In- und Ausland

g.a.s.t. bietet für private und öffentliche Institutionen, die ihre Mitarbeiter*innen zu Themen der Prüfungsvorbereitung, Qualitätssicherung, Leistungsmessung oder g.a.s.t.-Produkten fortbilden möchten, maßgeschneiderte Workshops. Die Planung und Durchführung der Veranstaltungen erfolgt in enger Absprache zwischen den Referent*innen von g.a.s.t. und den anfragenden Institutionen. Von 2020 bis 2022 haben u.a. folgende individuell abgestimmte Veranstaltungen stattgefunden:

task creation and performance measurement to task development for e-learning programmes (see chart on p. 59).

In the wake of the coronavirus pandemic, g.a.s.t. has digitally expanded its training programme. In addition to the 60 to 90-minute webinars, g.a.s.t. instructors conducted online workshops focussing on “The digital TestDaF” in 2021. The workshops are interactive and the exchange with the participants plays an important role.

4.3 Customised workshops in Germany and abroad

g.a.s.t. offers customised training workshops relating to exam preparation, quality assurance, performance measurement or g.a.s.t. products. The events are planned and organised in close consultation between the g.a.s.t. instructors and the requesting institutions. From 2020 to 2022, the following customised events took place, among others:

| Seminare nach Maß Customized seminars | |
|--------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------|
| Leseverstehen: Aufgaben- und Itemerstellung für Sprachtests | Technische Universität Budapest (BME) - Fremdspracheninstitut (INYK) |
| Mehrteilige Fortbildungsreihe zur Erstellung von Testaufgaben und Testspezifikationen | Institut National des Langues Luxembourg |
| Der digitale TestDaF: Einblick in die Anforderungen unterschiedlicher Aufgabenformate der Prüfungsteile Schreiben und Sprechen | Goethe-Institut Moskau |
| Nutzung des GER für einen kompetenzorientierten und kommunikativen Sprachunterricht | Chinesisch-Deutsches Kolleg für Intelligente Produktion der Shanghai Dianji-Universität |
| How can CEFR scales and test tasks be related to each other? | Sprachenzentrum der Universität Leipzig |
| Validität von Aufgaben im Hörverstehen; Itemkonstruktion im Lese- und Hörverstehen | Institut zur Qualitätsentwicklung im Bildungswesen; Wissenschaftliche Einrichtung der Länder an der Humboldt-Universität zu Berlin e.V. |
| Datenanalyse mit FACETS | Hochschule für Musik, Theater und Medien, Hannover Music Lab |
| Online Sprachen unterrichten am Beispiel der Deutsch-Uni Online | Ludwig-Maximilians-Universität München |
| Einsatz von C-Tests am Beispiel des onSET | Brandenburgische Technische Universität Cottbus-Senftenberg, ZE Sprachen |
| Leistungsmessung: Prinzipien der Evaluation fremdsprachlicher Kompetenzen mit Bezug zum GER/Begleitband | Technische Universität München, Sprachenzentrum |
| Qualitätssicherung: Beurteilung von Textproduktion und produktiven Aufgaben in Lese- und Hörverstehen anhand des GER | Philipps-Universität Marburg, Studienkolleg Mittelhessen |
| GER: Deskriptoren und deren Anwendung in realen Kontexten/Bewertungskriterien schriftlicher und mündlicher Prüfungen | Universität Leipzig, Sprachenzentrum |
| Mündliche Sprachprüfungen: praktische Handreichungen zur Erstellung und Bewertung von mündlichen Prüfungsleistungen (A1–B2) | interculturals e.V., Bildungsakademie |

| Workshops 2020/2021/2022 Workshops 2020/2021/2022 | Monat/Jahr Month/year |
|------------------------------------------------------------------------------------------------------------------------------------------|---------------------------|
| Der digitale TestDaF: Ziele, Aufgaben und Anforderungen | 02/2020 |
| Der digitale TestDaF: Teilnehmende gut vorbereiten | 03/2022 |
| Der digitale TestDaF: Förderung zentraler Kompetenzen in Ihrem Prüfungsvorbereitungskurs – „Zusammenhänge erkennen und ausdrücken“ | 09/2021 05/2022 11/2022 |
| Der digitale TestDaF: Förderung zentraler Kompetenzen in Ihrem Prüfungsvorbereitungskurs – „Informationen verarbeiten“ | 10/2021 12/2022 |
| Der digitale TestDaF: Förderung zentraler Kompetenzen in Ihrem Prüfungsvorbereitungskurs – „Positionen erkennen und sich positionieren“ | 11/2021 |
| Der digitale TestDaF: Beurteilung mündlicher und schriftlicher Leistungen | 11/2021 03/2022 06/2022 |
| Webinare 2020/2021/2022 Webinars 2020/2021/2022 | Anzahl Number of dates |
| Der digitale TestDaF: Ziele, Aufgaben und Anforderungen | 39 |
| Der digitale TestDaF: Prüfungsteil Lesen | 3 |
| Der digitale TestDaF: Prüfungsteil Hören | 3 |
| Der digitale TestDaF: Prüfungsteil Schreiben | 4 |
| Der digitale TestDaF: Prüfungsteil Sprechen | 4 |
| Der digitale TestDaF: Beurteilung schriftlicher und mündlicher Leistungen | 1 |
| Vorbereitung auf den digitalen TestDaF: Kurzantworten im Prüfungsteil Hören | 1 |
| Der papierbasierte TestDaF: Aufbau, Ziele und Format | 5 |
| Deutsch-Uni Online: Deutsch lernen im Internet | 8 |
| Deutsch-Uni Online: Training digitaler TestDaF – Einsatz für Ihren Prüfungsvorbereitungskurs | 11 |
| Deutsch-Uni Online: Deutsch unterrichten mit der DUO – Lernmodule auf dem Niveau A1/A2 institutionell betreuen | 3 |
| Deutsch-Uni Online: Deutsch unterrichten mit der DUO – Lernmodule auf dem Niveau B1/B2 institutionell betreuen | 5 |
| Der Online-Spracheinstufungstest (onSET) im Überblick | 5 |
| Der digitale TestAS: Aufbau, Ziele, Format | 7 |
| Der digitale TestDaF: Organisation und Technik für Prüfungsbeauftragte | 13 |
| g.a.s.t.-Dialog für Prüfungsbeauftragte: Teilnehmendenberatung zum TestDaF – Prüfungsvorbereitung und Materialien auf der neuen Webseite | 2 |
| g.a.s.t.-Dialog für Prüfungsbeauftragte: Studienvorbereitende Module der DUO | 2 |
| g.a.s.t.-Dialog für Prüfungsbeauftragte: Der Online-Spracheinstufungstest onSET | 2 |
| g.a.s.t.-Dialog für Prüfungsbeauftragte: Der TestAS | 2 |
| g.a.s.t.-Dialog für Prüfungsbeauftragte: Deutschlernende effektiv auf den digitalen TestDaF vorbereiten | 1 |
| g.a.s.t.-Dialog für Prüfungsbeauftragte: TestDaF digital: Vorbereitungskurse mit g.a.s.t.-Materialien gestalten und durchführen | 3 |

5 Aufträge

5 *Third-party contracts*

5.1 Institut National des Langues Luxembourg (INL)

Das INL in Luxemburg, eine Einrichtung des luxemburgischen Ministeriums für Bildung, Kinder und Jugend, hat g.a.s.t. damit beauftragt, bei der Entwicklung der neuen Sprachprüfung *Lëtzebuergesch als Friemsprooch A2–B1 (Test LAF)* zu beraten und die dafür erforderliche Software bereitzustellen.

Der Test LAF ist eine allgemeinsprachliche Prüfung mit dem Ziel, die luxemburgische Sprache und die soziale Integration von erwachsenen Einwander*innen unterschiedlichen Alters zu fördern. Teilnehmende legen die Prüfung vor allem zur Verbesserung ihrer Berufsaussichten ab oder um aus persönlichem Interesse ihre Luxemburgisch-Kenntnisse zu überprüfen.

Im hybriden Testformat des Test LAF werden die Prüfungsteile Lesen, Hören und Schreiben online am Computer durchgeführt. Da im Prüfungsteil Sprechen die mündliche Interaktion im Mittelpunkt steht, wird dieser Prüfungsteil „face to face“ als Paarprüfung abgenommen. Teilnehmende haben die Möglichkeit, den Test LAF modular abzulegen, d.h. sich nur in einzelnen Prüfungsteilen prüfen zu lassen.

Die Beratungsleistungen von g.a.s.t. umfassen unter anderem die Begutachtung der Testspezifikationen (Testkonstrukt und Aufgabenformate), Workshops und Fortbildungen zur Vereinheitlichung und Objektivierung der Beurteilung in den produktiven Prüfungsteilen sowie die psychometrischen Analysen.

Nach der erfolgreichen Pilotierung von ersten, neu entwickelten Aufgabentypen im Juni 2021 mit Teilnehmenden am INL, konnten im Projekt bislang das Testformat festgelegt und Testspezifikationen erarbeitet werden. Weitere Erprobungen im Jahr 2022 dienten der Entwicklung von Aufgabenmaterial für die Itembank. Zudem wurde die Software für die digitale Prüfung kontinuierlich angepasst sowie Schulungskonzepte für Lehrkräfte, Beurteilende und Prüfende erarbeitet. Das Projekt endet zum derzeitigen Stand mit der Einführung der neuen Prüfung Ende des Schuljahres 2023/2024.

5.1 *Institut National des Langues Luxembourg (INL)*

*The INL in Luxembourg, an institution of the Luxembourg Ministry of Education, Children and Youth, has commissioned g.a.s.t. to advise on the development of the new language test *Lëtzebuergesch als Friemsprooch A2–B1 (Test LAF)* and to provide the necessary software.*

The LAF test is a general language test with the aim of promoting the Luxembourgish language and the social integration of adult immigrants. Participants take the test primarily to improve their job prospects or to test their Luxembourgish language skills out of personal interest.

In the hybrid test format of the LAF test, the reading, listening and writing sections are carried out online on the computer. As the speaking section of the test focuses on oral interaction, this section of the test is taken face-to-face by way of a paired test. Participants have the option of taking the LAF test in a modular format, i.e. they can only be tested in individual parts of the test.

The consulting services provided by g.a.s.t. include the assessment of test specifications (test construct and task formats), workshops and further training to standardise and objectify the assessment in the productive parts of the test as well as psychometric analyses.

Following the successful piloting of the first newly developed item types with participants at the INL in June 2021, the project has so far been able to determine the test format and develop test specifications. Further tests in 2022 were used to develop item material for the item bank. In addition, the software for the digital test was continuously adapted and training concepts for teachers, assessors and examiners were developed. The project currently ends with the introduction of the new examination at the end of the 2023/2024 school year.





5.2 Italienische Bildungsdirektion der Autonomen Provinz Bozen

Die *Italienische Bildungsdirektion der Autonomen Provinz Bozen* hat g.a.s.t. damit beauftragt, die Entwicklung von Sprachstandserhebungen im Fach Deutsch auf den Schulstufen 4 und 11 wissenschaftlich zu begleiten. Im Projekt, das von 2021 bis 2023 läuft, werden zwei Sprachstandstests für eine Erhebung an den Grundschulen und Oberschulen von Südtirol neu entwickelt. g.a.s.t. hat das Projektpersonal der Bildungsdirektion im ersten Halbjahr 2021 durch entsprechende Fortbildungsveranstaltungen auf diese Aufgabe vorbereitet. Seit Mitte 2021 begleiten Mitarbeiter*innen von g.a.s.t. aus den Bereichen Produktentwicklung und Psychometrie das Projekt. Dazu gehört die Begutachtung von Testspezifikationen und Testformat sowie die Beratung zu den Pilotierungen der neuen Tests und den psychometrischen Analysen. Eine Beratung zum Standard-Setting ist für das zweite Halbjahr 2023 geplant. Für den Auftraggeber ist eine Testentwicklung nach gültigen internationalen und wissenschaftlichen Qualitätsstandards eine essentielle Voraussetzung, um eine Akzeptanz der Sprachstandserhebung bei den Stakeholdern in Südtirol sicherzustellen.

5.3 Zentralstelle für das Auslandsschulwesen (ZfA)

Seit 2005 führt g.a.s.t. im Auftrag der ZfA psychometrische Auswertungen und Analysen zum Deutschen Sprachdiplom (DSD) der Kultusministerkonferenz (KMK) durch. Im Oktober 2012 wurde hierüber ein Rahmenvertrag geschlossen. 2021 wurde der Vertrag verlängert.

Die Aufträge betreffen insbesondere die Erfassung und Analyse von Daten aus Erprobungen und Echtläufen des DSD I (zielt auf die GER-Niveaus A2/B1), DSD II (B2/C1) und DSD I PRO (A2/

5.2 Italian Education Directorate of the Autonomous Province of Bolzano

The Italian Directorate of Education of the Autonomous Province of Bolzano has commissioned g.a.s.t. to provide scientific support for the development of language proficiency tests in German at school levels 4 and 11. In the project, which runs from 2021 to 2023, two new language proficiency tests will be developed for a survey at primary and secondary schools in Southern Tyrolia. g.a.s.t. prepared the project staff of the Directorate of Education for this task in the first half of 2021 by organising appropriate training events. Since mid-2021, g.a.s.t. staff from product development and psychometrics have been supporting the project. This includes reviewing the test specifications and test format as well as advising on the piloting of the new tests and the psychometric analyses. Consultancy on the standard setting is planned for the second half of 2023. For the client, test development in accordance with valid international and scientific quality standards is an essential prerequisite for ensuring acceptance of the language proficiency survey among stakeholders in Southern Tyrolia.

5.3 Zentralstelle für das Auslandsschulwesen (ZfA)

Since 2005, g.a.s.t. has been carrying out psychometric evaluations and analyses of the German Language Diploma (DSD) of the Standing Conference of the Ministers of Education and Cultural Affairs of the Länder (KMK) on behalf of the ZfA. A framework agreement was concluded for this in October 2012. The contract has been extended in 2021.

The tasks relate in particular to the collection and analysis of data from trial tests and runs of the DSD I (aimed at CEFR lev-

B1, berufsbezogen), jeweils mit den Prüfungsteilen Hörverstehen, Leseverstehen und Schriftliche Kommunikation. Dazu kommt die Erstellung von Beurteilungsprofilen im Prüfungsteil Schriftliche Kommunikation (Echtläufe) mit psychometrischen Indikatoren der Strenge und Konsistenz jedes einzelnen Beurteilenden. Die Analysen der Erprobungsdaten schließen die Verankerung der Prüfungsteile Hör- und Leseverstehen zum Zwecke des Test-Equating mittels C-Tests ein.

Bis Ende 2022 haben an DSD-Echtläufen insgesamt 473.106 Schüler*innen teilgenommen. Zusammen mit Erprobungen hat g.a.s.t. in diesem Zeitraum Prüfungsdaten von 492.402 DSD-Teilnehmenden maschinell erfasst, ausgewertet und psychometrisch analysiert.

5.4 Bundesagentur für Arbeit

Der Berufspsychologische Service der Bundesagentur für Arbeit in Nürnberg hat im Oktober 2020 den Auftrag zur Erstellung zweier C-Tests erteilt. Nach 2008 ist dies der zweite Auftrag zur C-Test-Erstellung.

Der erste C-Test zielt auf Erwachsene, der zweite auf Jugendliche ab ca. 15 Jahre. Beide Tests bestehen aus sechs Aufgaben (Lückentexten) mit je 20 Lücken. Die in den Texten behandelten Themen wurden in Rücksprache mit dem Auftraggeber festgelegt. Sämtliche Texte wurden nach der beim onSET bewährten Methodik auf der Grundlage psychometrisch-statistischer Analysen kalibriert. Die C-Tests erlauben es, in der jeweiligen Zielgruppe die allgemeine Sprachkompetenz in Deutsch als Fremdsprache oder Zweitsprache analog zu den GER-Niveaus A2 bis C1 (oder höher) zu erfassen.

els A2/B1), DSD II (B2/C1) and DSD I PRO (A2/B1, vocational), each with the listening comprehension, reading comprehension and written communication sections of the exam. In addition, assessment profiles are created in the Written Communication section of the exam with psychometric indicators of the rigour and consistency of each individual assessor. The analyses of the test data include the anchoring of the Listening and Reading Comprehension sections for the purpose of test equating using C-tests.

By the end of 2022, a total of 473,106 pupils had taken part in real DSD tests. During this period, g.a.s.t. collected, evaluated and psychometrically analysed examination data from 492,402 DSD participants.

5.4 Federal Employment Agency

In October 2020, the Vocational Psychology Service of the Federal Employment Agency in Nuremberg commissioned the creation of two C-tests. This is the second request for the creation of C-tests since 2008.

The first C-test is aimed at adults, the second at adolescents aged 15 and over. Both tests consist of six tasks (cloze texts) with 20 gaps each. The topics covered in the texts were determined in consultation with the client. All texts were calibrated on the basis of psychometric-statistical analyses in accordance with the proven methodology of the onSET. The C-tests make it possible to assess the general language competence in German as a foreign or second language in the respective target group analogue to CEFR levels A2 to C1 (or higher).

6 Weitere Aktivitäten von g.a.s.t. e.V.

6 Further activities of g.a.s.t. e.V.

6.1 Gremiensitzungen

| | |
|------------|------------------------------------------------|
| 15.05.2020 | 39. Vorstandssitzung und Mitgliederversammlung |
| 11.12.2020 | 40. Vorstandssitzung |
| 21.05.2021 | 41. Vorstandssitzung und Mitgliederversammlung |
| 19.11.2021 | 42. Vorstandssitzung |
| 29.04.2022 | 43. Vorstandssitzung und Mitgliederversammlung |
| 18.11.2022 | 44. Vorstandssitzung und Mitgliederversammlung |

6.1 Committee meetings

| | |
|------------|----------------------------------------|
| 15.05.2020 | 39. Board Meeting and General Assembly |
| 11.12.2020 | 40. Board Meeting |
| 21.05.2021 | 41. Board Meeting and General Assembly |
| 19.11.2021 | 42. Board Meeting |
| 29.04.2022 | 43. Board Meeting and General Assembly |
| 18.11.2022 | 44. Board Meeting and General Assembly |

6.2 Mitgliedschaften

- Association of Language Testers in Europe (ALTE; das TestDaF-Institut ist Vollmitglied)
- European Association for Language Testing and Assessment (EALTA; Institutionelles Mitglied)
- European Association for Quality Language Services (EAQUALS; Assoziiertes Mitglied)
- Gesellschaft für Medien in der Wissenschaft (GMW, korporatives Mitglied)
- GATE-Germany Internationales Hochschulmarketing
- International Test Commission (ITC; Assoziiertes Mitglied)

6.2 Memberships

- Association of Language Testers in Europe (ALTE; Full Member)
- European Association for Language Testing and Assessment (EALTA; Institutional Member)
- European Association for Quality Language Services (EAQUALS; Associate Member)
- Society for Media in Science (GMW, Corporate Member)
- GATE-Germany International University Marketing
- International Test Commission (ITC; Associate Member)

6.3 Beirats- und Gremientätigkeit

- Zentrales Netzwerk Deutsch des Auswärtigen Amtes (Althaus/Weingärtner)
- Beirat Sprache des Goethe-Instituts (Althaus/Drackert)
- Wissenschaftlicher Beirat der Zentralstelle für das Auslandsschulwesen (Althaus/Drackert)
- Editorial Board der Fachzeitschrift *Language Testing* (Eckes)
- Editorial Board der Open Access-Fachzeitschrift SAGE Open (Eckes)
- Executive Committee der EALTA (Zimmermann)
- Wissenschaftlicher Beirat im DAKODA-Projekt (Zimmermann)
- Wissenschaftlicher Beirat der Zeitschrift *Fremdsprachen und Hochschule* (Drackert)

6.3 Advisory board and committee activities

- Central German Network of the Federal Foreign Office (Althaus/Weingärtner)
- Language Advisory Board of the Goethe-Institut (Althaus, Drackert)
- Scientific Advisory Board of the Central Agency for Schools Abroad (Althaus, Drackert)
- Editorial Board of the journal *Language Testing* (Eckes)
- Editorial Board of the open access journal SAGE Open (Eckes)
- Executive Committee of EALTA (Zimmermann)
- Scientific Advisory Board in the DAKODA project (Zimmermann)
- Scientific advisory board of the journal *Fremdsprachen und Hochschule* (Drackert)



6.4 Gutachtertätigkeit

Peer-Review-Fachzeitschriften (Eckes, Drackert)

- Assessing Writing
- Behavior Research Methods
- Frontiers in Psychology
- Journal of Applied Measurement
- Language Assessment Quarterly
- Language Testing
- Psychological Test and Assessment Modeling
- Studies in Second Language Acquisition
- Fremdsprachen und Hochschule

Executive Committee der EALTA (Zimmermann)

Sonja Zimmermann gehört dem Executive Committee der European Association for Language Testing and Assessment (EALTA) an. Sie ist zugleich *Chair of Committee for Conference Organisation* der EALTA.

6.5 Bildungsmessen und Fachtagungen

g.a.s.t. als Sponsoring-Partner

| | Ereignis Event | Ort, Land Venue, country |
|------|------------------------------------------|----------------------------------------------|
| 2020 | Education and Career | Minsk, Belarus <i>Belarus</i> |
| | 31. AKS-Arbeitstagung | Frankfurt (Oder), Deutschland <i>Germany</i> |
| | DaFWEBKON | online |
| | Marburger FaDaF-Thementage | online |
| | DAAD-Leitertagung | online |
| 2021 | DaFWEBKON | online |
| | ALTE 1st International digital Symposium | online |
| | 48. FaDaF-Jahrestagung | online |
| | Lange Nacht der Ideen (DAAD) | online |
| | DAAD-Lektorensommertreffen | online |
| | Eaquals 30th Anniversary Conference | online |
| | EALTA | online |
| 2022 | 32. AKS-Arbeitstagung | online |
| | LTRC-Colloquium | online |
| | ALTE 8th International Conference | Madrid, Spanien <i>Spain</i> |
| | EALTA | Budapest, Ungarn <i>Hungary</i> |
| | IDT | Wien, Österreich <i>Austria</i> |
| | FaDaF-Jahrestagung | Kassel, Deutschland <i>Germany</i> |
| | DAAD-Lektorensommertreffen | online |

6.4 Reviewing activities

Peer-reviewing for journals (Eckes, Drackert)

- Assessing Writing
- Behaviour Research Methods
- Frontiers in Psychology
- Journal of Applied Measurement
- Language Assessment Quarterly
- Language Testing
- Psychological Test and Assessment Modelling
- Studies in Second Language Acquisition
- Foreign languages and university

Executive Committee of EALTA (Zimmermann)

Sonja Zimmermann is a member of the Executive Committee of the European Association for Language Testing and Assessment (EALTA). She is also Chair of the Committee for Conference Organisation of EALTA.

6.5 Education fairs and conferences

g.a.s.t. as sponsoring partner

6.6 g.a.s.t. auf der IDT in Wien, Motto: *mit.sprache.teil.haben

Vom 15. bis 20. August 2022 fand in Wien die XVII. Internationale Tagung der Deutschlehrer*innen statt. Am Sonntag davor trafen sich die Sektionsleiter*innen, die anschließend im Rathaus der Stadt offiziell begrüßt wurden, bevor am Montag das Fachprogramm der Tagung eröffnet wurde. g.a.s.t. war mit einem Stand und zahlreichen Mitarbeiter*innen im Fachprogramm vertreten. Günther Depner, Ines Paland-Riedmüller und Sonja Zimmermann leiteten mit unterschiedlichen Partner*innen jeweils eine Sektion in den Sektionsfeldern „Pädagogische, didaktische und spracherwerbstheoretische Konzepte“ sowie „Lehren und Lernen mit und in digitalen Umgebungen“. Sie waren dabei für die Auswahl der Beiträge, die Programmgestaltung und Moderation in ihren Sektionen zuständig. Daniela Marks und Anja Peters informierten in einem Workshop über Anforderungen des digitalen TestDaF und wie Lehrkräfte potentielle Teilnehmende optimal auf diese vorbereiten können. Ulrike Arras hielt einen Vortrag zu Sprachmittlung im Hochschulkontext.

Ziele des g.a.s.t.-Standes waren insbesondere folgende:

- Aufmerksamkeit der Besucher*innen erhalten
- Mit Besucher*innen ins Gespräch kommen/diskutieren am Stand
- Über g.a.s.t. und ihre Produkte informieren, Interesse wecken
- Besucher*innen „unterhalten“ und in Erinnerung bleiben
- g.a.s.t.-Akademie herausstellen
- Adressverteiler von Deutschlehrkräften ausbauen mit Einwilligung zur Nutzung der Daten

Neben persönlichen Ansprechpartnerinnen von g.a.s.t., Informations- und Werbematerialien erwartete die Besucher*innen des g.a.s.t.-Standes eine Online-Testversion des onSET sowie ein Kapitel eines DUO-Moduls, um Einblicke in das C-Testverfahren und die Arbeit mit der Lernplattform zu gewinnen. Über zwei g.a.s.t.-Laptops vor Ort sowie das eigene (auch mobile) Endgerät via QR-Code oder Link konnten Interessierte die Dienste ausprobieren und ihre Kontaktdaten hinterlassen.

Themenschwerpunkte am Stand waren der digitale TestDaF (Formatumstellung und Prüfungsvorbereitung), die Lizenzierung internationaler Sprachschulen/-instituten als g.a.s.t.-Testzentrum, die DUO als Ergänzung zum Präsenzunterricht sowie die g.a.s.t.-Akademie und das positive Feedback zu Schulungen aus dem Hause g.a.s.t.



6.6 g.a.s.t. at the IDT in Vienna, slogan: *mit.sprache.teil.haben

The XVII. International Conference of German Teachers took place in Vienna from 15 to 20 August 2022. The section heads met on the previous Sunday and were officially welcomed at Vienna City Hall before the conference's specialist programme opened on Monday. g.a.s.t. was represented with a stall and numerous staff members in the specialist programme. Günther Depner, Ines Paland-Riedmüller and Sonja Zimmermann, together with various partners, each served as section heads in the fields of "Pedagogical, didactic and language acquisition theory concepts" and "Teaching and learning with and in digital environments". They were responsible for programme design and moderation in their sections. Furthermore, Daniela Marks and Anja Peters provided information on the requirements of the digital TestDaF and how teachers can optimally prepare potential participants for it. Ulrike Arras gave a presentation on language mediation in the university context.

The main targets of the g.a.s.t. exhibition stand were as listed here:

- Attract the attention of visitors
- Engage in dialogue with visitors/discuss at the exhibition stand
- Provide information about g.a.s.t. and its products, generate interest
- "Entertain" visitors and stay in their memory
- Highlighting the g.a.s.t. academy
- Expand the mailing list of German teachers with consent to use data

In addition to personal contact with g.a.s.t., visitors to the g.a.s.t. stall were able to explore an online test version of onSET and a chapter of a DUO module to gain insight into the C-test procedure and working with the learning platform.

The main topics at the stall were the digital TestDaF (format revision and exam preparation), the licensing of international language schools/institutions as g.a.s.t. test centres, DUO as a supplement to face-to-face teaching, the g.a.s.t. Academy and the positive feedback on g.a.s.t. training courses.

7 Wirtschaftliche Entwicklung

7 Financial review

Die Gesellschaft für Akademische Studienvorbereitung und Testentwicklung e.V. (g.a.s.t.) hat in den vergangenen drei Jahren ihre satzungsmäßigen Aufgaben zu 84% aus Eigeneinnahmen und zu 16% durch öffentliche Projektförderung finanziert.

Die Gesamteinnahmen aus allen g.a.s.t.-Projekten und -Produkten lagen durchschnittlich pro Jahr bei 8,3 Mio. Euro.

Die Eigeneinnahmen berücksichtigen die TestDaF- und TestAS-Prüfungsentgelte, die Einnahmen aus der DUO und dem onSET sowie den Fremdaufträgen. Dazu gehören u.A.:

- Testmethodische Auswertungen der Deutschen Sprachdiplome der Kultusministerkonferenz im Auftrag der Zentralstelle für das Auslandsschulwesen (ZfA) in Bonn;
- Erstellung und Wartung einer Prüfungsmanagement-Software für das Europäische Patentamt (EPA) in München;
- Softwareentwicklungen und -pflege für die Akademische Prüfstelle (APS) in Peking;
- Entwicklung der digitalen Prüfung Lëtzebuergesch als Friemsprouch für das Institut National des Langues in Luxemburg;
- Beratungsaufträge zu den Themen Testen und Prüfen sowie E-Learning.

Die Einnahmen aus Zuwendungen beinhalten folgende Projekte:

- Verbundprojekt „Digitaler Campus – ein Portal vernetzter Plattformservices“, DAAD aus Mitteln des BMBF;
- Verbundprojekt „Bildungsraum Digital – BIRD“, BMBF;
- Verbundprojekt „Fach.Deutsch.Digital für Studium und Beruf“, Stiftung Innovation in der Hochschullehre in Trägerschaft der Toepfer Stiftung;
- Projekt der Deutschen Forschungsgemeinschaft „Zeit, Lücken zu schließen: eine multimethodische Studie zur Rolle der Zeit im Konstrukt computergestützter C-Tests in drei Sprachen“.

Die weltweiten Testzentren erhalten einen Teil der TestDaF- und TestAS-Prüfungsentgelte und finanzieren damit ihre technische Ausstattung und die Prüfungsdurchführung. Dieser Betrag lag durchschnittlich pro Jahr bei 1,8 Mio. Euro.

Die freiwilligen externen Wirtschaftsprüfungen der Jahresabschlüsse von g.a.s.t. haben in den letzten drei Jahren zu keinen Einwendungen gegen die Ordnungsmäßigkeit der Jahresabschlüsse geführt.

Over the past three years, the Society for Academic Study Preparation and Test Development e.V. (g.a.s.t.) has financed 84% of its statutory tasks from its own revenue and 16% by public project funding.

Total revenue from all g.a.s.t. projects and products averaged EUR 8.3 million per year.

Own revenue includes TestDaF and TestAS examination fees, from DUO and onSET as well as third-party orders. This includes, among other things

- *Test methodological analyses of the German Language Diplomas of the Standing Conference of the Ministers of Education and Cultural Affairs on behalf of the Zentralstelle für das Auslandsschulwesen (ZfA) in Bonn;*
- *Creation and maintenance of examination management software for the European Patent Office (EPO) in Munich;*
- *Software development and maintenance for the Academic Evaluation Centre (APS) in Beijing;*
- *Development of the digital test Lëtzebuergesch as Friemsprouch for the Institut National des Langues in Luxembourg;*
- *Consultancy assignments on the topics of testing and e-learning.*

Revenue from public funding was raised from for:

- *Joint project “Digital Campus – a portal of networked platform services”, DAAD with funding from the BMBF;*
- *Joint project “Bildungsraum Digital – BIRD”, BMBF;*
- *Joint project “Fach.Deutsch.Digital für Studium und Beruf”, Stiftung Innovation in der Hochschullehre in Trägerschaft der Toepfer Stiftung;*
- *Project of the German Research Foundation “Time to close gaps: a multi-method study on the role of time in the construct of computerised C-tests in three languages”.*

The global test centres receive a portion of the TestDaF and TestAS examination fees to cover their outlay for technical equipment and the administration of the examinations. This amount averages 1.8 million Euros per year.

The voluntary external audits of g.a.s.t.’s annual financial statements confirmed the correctness of the annual financial statements in the last three years.

8 Vorträge und Publikationen

8 Lectures and publications

8.1 Vorträge und Präsentationen (Auswahl)

8.1 Lectures and presentations (selection)

2020

Roche, J. (2020, Juli). *Lernen und Lehren im digitalen Zeitalter – Herausforderungen und Chancen*. Vortrag online gehalten beim Virtuellen Lektoren-Sommertreffen des DAAD.

Roche, J. (2020). *Deutsche Grammatik in Bewegung – Embodiment in der multimedialen Vermittlung der deutschen Grammatik*. German Teacher Virtual Conference, Goethe Institut, 20.-21.10.2020 (online).

Roche, J. (2020). *Sinn und Unsinn der digitalen Medien im FSU*. Virtueller jordanischer Deutschlehrertag, German Jordanian University 10.10.2020 (online).

2021

Althaus, H.-J., Krause, K., Paland-Riedmüller, I. (2021, Juli). *Masse oder Klasse? Zu Bedeutung und Nutzen von Tutoren- und digital erzeugtem Feedback beim Online-Sprachlernen*. Online-Vortrag gehalten auf dem XIV. Kongress der Internationalen Vereinigung für Germanistik (IVG), Palermo.

Arras, U. (2021, Juli). *Non lo so in italiano. Sprachmittlung in mehrsprachigen akademischen Kontexten*. Online-Vortrag gehalten auf dem XIV. Kongress der Internationalen Vereinigung für Germanistik (IVG), Palermo.

Depner, G. (2021, September). *Fremdsprachendidaktik: Der digitale TestDaF als Instrument zur Entwicklung kommunikativer Kompetenzen*. Vortrag gehalten auf dem 29. Kongress für Fremdsprachendidaktik der Deutschen Gesellschaft für Fremdsprachenforschung (DGFF), Online-Tagung

Marks, D. (2021, Oktober). *Der digitale TestDaF: Ziele, Aufgaben und Anforderungen*. Online-Vortrag gehalten auf der Internationalen digitalen Bildungskonferenz Usbekistan „DaF-Unterricht im digitalen Zeitalter“, Online-Tagung.

Roche, J. (2021). *On the Role of Language Policies in the Internationalisation of Higher Education Programmes*. Plenarvortrag, Políticas linguísticas e internacionalização do ensino superior. Diálogo entre Brasil e Europa. Tagung der UFPR, Bundesstaatsuniversitäten und Bundesstaatsregierung, Rio de Janeiro, Brasilien, 8. Dezember 2021 (online).

Roche, J. (2021). *Sprache als wissenschaftliches Kapital*. Tag der Lehre. Brandenburgische Technische Universität Cottbus, Cottbus, 11. November 2021 (online).

Roche, J. (2021). *Mit-Sprachen – Herausforderungen einer pragmatischen Mehrsprachenpolitik und ihre Folgen für die Sprachvermittlung und Didaktik*. Mehrsprachigkeitskonferenz, Goethe Institut Ljubljana, Slowenien, 20. Oktober 2021.

Roche, J. (2021). *Berufssprache Deutsch: Grundlagen – Didaktik – Potentiale*. Fokus DAF/DAZ – Gegenwärtige Tendenzen in Forschung und Lehre. Deutsch-Jordanische Universität, Madaba, Jordanien, 15. März 2021 (online).

Roche, J. (2021). *Zu den Mehrwerten digitaler Medien im Fremdsprachenunterricht – Über komplexe Theorien und eine (oft) minimalistische Praxis*. Vortrag und Workshop „Fremdsprachen digital unterrichten – mit Sinn, Freude und Verstand“, DAAD Tokio, Japan, 7. Mai 2021 (online).

Zimmermann, S. (2021, Mai). *Validating an L2 integrated writing assessment. Linking process, product and scoring data*. Paper presented at the L2 Writing Research Seminar virtual conference, Universidad des Murcia.

Zimmermann, S. (2021, Juni). *Die integrierte Schreibaufgabe im digitalen TestDaF. Ein Mehr-Methoden-Ansatz zur Validierung des Konstrukts*. Vortrag gehalten auf der 48. Jahrestagung DaF/DaZ, digital-hybrid an der PH Freiburg in Zusammenarbeit mit der Universität Kassel.

Zimmermann, S. (2021, Juni). *Linking process and product data in the context of an integrated writing assessment*. Paper presented at the 42nd Language Testing Research Colloquium, virtual.

Zimmermann, S. & Schwarz, L. (2021, September). *Sprachtechnologische Ansätze zur Auswertung fremdsprachlicher Prüfungsleistungen*. Vortrag gehalten auf dem 29. Kongress für Fremdsprachendidaktik der Deutschen Gesellschaft für Fremdsprachenforschung (DGFF), Online-Tagung [zusammen mit Andrea Horbach & Ronja Quante-Laarmann].

Zimmermann, S. (2021, September). *Test-takers' perception of the digital TestDaF*. Paper presented at the BAAL TEA SIG Online-Conference.

2022

Arras, U. & Schuhmann, K. (2022, August). *Sprachmittlung im Hochschulkontext: Mediation competences US-amerikanischer DaF-Lerner*. Vortrag gehalten auf der XVII. Internationalen Tagung der Deutschlehrerinnen und Deutschlehrer, Wien.

Drackert, A. (2022, März). *Sprachliche Anforderungen für die Zulassung internationaler Studierender zum Studium. Die Rolle zugelassener Sprachprüfungen für den Bildungserfolg*. Vortrag gehalten auf der Tagung „Sprache als Schlüssel – wozu? Bildungserfolg im Fokus“, Bamberg.

Korflür, D. (2022, September). *Der digitale TestDaF – Von den sprachlichen Anforderungen im Hochschulalltag zum Testformat*. Vortrag gehalten auf der 49. FaDaF Jahrestagung Deutsch als Fremd- und Zweitsprache, Kassel.

Marks, D. & Peters, A. (2022, August). *Der Digitale TestDaF: Anforderungen – Aufgaben – Vorbereitung*. Vortrag gehalten auf der XVII. Internationalen Tagung der Deutschlehrerinnen und Deutschlehrer, Wien.

Roche, J. (2022). *Von der Transferdifferenz zur Konzeptuellen Kompetenz – Visualisierte Metaphern als Paralleltex-te im Modell der Kognitiven Sprachdidaktik*. In: Burneva, Nicolina/Hepp, Marianne et al. (Hgg.). *Kulturrealia in Paralleltex-ten als didaktischer Fokus trans- und interkultureller Germanistik*. IVG (Internationale Gesellschaft der Germanistik) 2020/21, Università degli studi Palermo, Palermo, Italien, 28. Juli 2021. Frankfurt: Lang, 21–29.

Roche, J. (2022). *Das Prinzip der vollständigen Handlung im DAF-Unterricht: Linguistische, didaktische und lerntheoretische Grundlagen der Szenariendidaktik*. Ringvorlesung „Berufsorientiertes Sprachlernen und -lehren“. Universität Kassel, Kassel, 21. November (online).

Roche, J. (2022). *Lebenswelten und Grammatik: Was uns der Sport über die deutsche Grammatik verrät*. Experimenta Heilbronn, Heilbronn, 29. September 2022.

Roche, J. (2022). *Sprachen digital lehren und lernen – Impulse, Konzepte, Lern- und Kommunikationsmodelle, Werkzeuge*. HSU digital – Unterschiedliche Medien zielgerichtet nutzen. Qualitäts- und Unterstützungs-Agentur – Landesinstitut für Schule des Landes Nordrhein-Westfalen (QUA-LiS NRW), Soest, 31. August 2022 (online).

Paland-Riedmüller, I. (2022, August). *Deutsch-Uni Online (DUO). Online-Lernmodule für Ihren Deutschunterricht*. Vortrag gehalten auf der XVII. Internationalen Tagung der Deutschlehrerinnen und Deutschlehrer, Wien.

Podlaha, T., Wegener, D. (2022, Dezember). *The digital TestAS: Objective and fair selection of applicants through digital and standardised aptitude testing*. Online-Vortrag gehalten auf der Moving Target Digitalisation: Increasing the Impact of Internationalisation in Higher Education, Berlin.

Zimmermann, S., Marks, D. & Hukava, Y. (2022, Juni). *On common ground? – Challenges of test-taker accomodation in a digital high-stakes exam*. Paper presented at the EALTA EDI SIG Meeting, Budapest.

Zimmermann, S. (2022, Juni). *Test-takers' perception of a new digital high-stakes test for admission purposes*. Paper presented at the 18th EALTA Conference, Budapest.

Zimmermann, S. & Peters, A. (2022, Dezember). *The digital TestDaF: Preparing international applicants for university study in Germany*. Online-Vortrag gehalten auf der Moving Target Digitalisation: Increasing the Impact of Internationalisation in Higher Education, Berlin.



8.2 Publikationen

8.2 Publications

2020

Arras, U. & Fohr, T. (2020). Mitschreiben: Funktionen und didaktische Überlegungen zu Formen der Wissensverarbeitung an der Hochschule. In A. Gryszko, C. Lammers, K. Pelikan & T. Roelcke (Hrsg.), *DaFFür Berlin: Perspektiven für Deutsch als Fremd- und Zweitsprache in Schule, Beruf und Wissenschaft* (S. 131–149). Göttingen: Universitätsverlag.

Arras, U. & Widmann, B. (2020). Fachliteratur exzerpieren: Techniken und Strategien der visuellen Aufbereitung von Fachtexten im studienbegleitenden Fremdsprachenunterricht. In E. Bonetto, M. Ennis & D. Unterkofler (Eds.), *Teaching languages for specific and academic purposes in higher education* (pp. 141–167). Bozen, Italien: Bozen-Bolzano University Press.

Aryadoust, V., Eckes, T. & In'nami, Y. (Eds.). (2020). *Frontiers in language assessment and testing. Frontiers in Psychology*.

Eckes, T. (2020). Rater-mediated listening assessment: A facets modeling approach to the analysis of raters' severity and accuracy when scoring responses to short-answer questions. *Psychological Test and Assessment Modeling*, 65(4), 449–471.

Eckes, T. & Althaus, H.-J. (2020). Language proficiency assessments in higher education admissions. In M. E. Oliveri & C. Wendler (Eds.), *Higher education admissions practices: An international perspective* (pp. 256–275). Cambridge, UK: Cambridge University Press.

Heppt, B., Köhne-Fuetterer, J., Eglinsky, J., Volodina, A., Stanat, P. & Weinert, S. (2020). *BiSpra 2–4: Test zur Erfassung bildungssprachlicher Kompetenzen bei Grundschulkindern der Jahrgangsstufen 2–4*. Münster: Waxmann.

Roche, J. (2020). *Fremdsprachenerwerb - Fremdsprachendidaktik*. Tübingen, UTB (UTB 2691), 4. neu bearbeitete Auflage.

Roche, J. (2020). Methoden der Grammatik-Arbeit. In: Hallet, Wolfgang/Königs, Frank G./Martinez, Hélène (Hgg.). *Handbuch Methoden im Fremdsprachenunterricht*. Seelze, Klett Kallmeyer, 169–173.

Roche, J./EL-Bouz, K. (2020). *Zur Räumlichkeit temporaler Präpositionen – Ein kognitionsdidaktischer Ansatz*“. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 1/2020.

Roche, J. (2020). *Körper und Grammatik – Zu den Grundlagen der Kognitiven Sprachdidaktik im Aktuellen Grammatik-Studio*. In: Schaar, Torsten/Altal, Mahasen/Shi Wen, Chang (Hgg.). *Fokus DaF/DaZ. Gegenwärtige Tendenzen in Forschung und Lehre*. Münster: LIT Verlag, 59–76.

Zimmermann, S. (2020). „Das ist doch Leseverstehen!“ – Eine empirische Untersuchung zum Konstrukt von integrierten Schreibaufgaben. In A. Drackert, M. Mainzer-Murrenhof, A. Soltyska & A. Timukova (Hrsg.), *Testen bildungssprachlicher Kompetenzen und akademischer Sprachkompetenzen. Zugänge für Schule und Hochschule* (S. 187–213). Frankfurt am Main: Peter Lang.

Zimmermann, S. (2020). „Das ist Zeitverlust für mich, den Text wieder lesen“ – Einblicke in das Schreiben von Zusammenfassungen in der Fremdsprache Deutsch. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 25(2), 111–133.

2021

Arras, U. (2021). Warum sollen wir das übersetzen!? Ergebnisse aus einer empirischen Studie zu Sprachmittlung und plurilingualen Kompetenzen in mehrsprachigen akademischen Kontexten. In S. Kirchmeyer (Hrsg.), *Klassisch – modern – digital: Fremdsprachenunterricht an Hochschulen zwischen Tradition und Moderne* (S. 333–359). Weimar: Bauhaus-Universität.

Aryadoust, V., Eckes, T. & In'nami, Y. (2021). Editorial: *Frontiers in language assessment and testing*. *Frontiers in Psychology*, 12, 691614.

Depner, G. (2021). Entwurfsfassungen als Instrumente zur Förderung argumentativer Schreibkompetenzen in der Fremdsprache. In: Reitbrecht, S. (Hrsg.) *Schreiben in Kontexten – soziokulturelle Dimensionen der Textproduktion in Deutsch als Fremdsprache*. Berlin: Erich Schmidt Verlag, 147–160.

Eckes, T. & Jin, K.-Y. (2021). Examining severity and centrality effects in TestDaF writing and speaking assessments: An extended Bayesian many-facet Rasch analysis. *International Journal of Testing*, 21(3–4), 131–153.

Eckes, T. & Jin, K.-Y. (2021). Measuring rater centrality effects in writing assessment: A Bayesian facets modeling approach. *Psychological Test and Assessment Modeling*, 63(1), 65–94.

Jin, K.-Y. & Eckes, T. (2021). *Detecting differential rater functioning in severity and centrality: The dual DRF facets model*. *Educational and Psychological Measurement*. Advance online publication.

Kecker, G. & Eckes, T. (2021). Standard-Setting zum digitalen TestDaF: Ein Online-Verfahren für eine Online-Prüfung

Köhne-Fuetterer, J., Drenhaus, H., Delogu, F., & Demberg, V. (2021). The online processing of causal and concessive discourse connectives. *Linguistics*, 59(2):417–448, 2021. doi:10.1515/ling-2021-0011

Roche, J. (Hg.) (2021). *WIR in Deutschland – Zusammen Leben Lernen*. Tübingen, Gunter Narr Verlag.
E-Book: https://www.lernen-lehren-helfen.daf.uni-muenchen.de/wir2_buch_einzelseiten.pdf. Als App qrco.de/bbKzpg.

Roche, J. (2021). *Berufssprache Deutsch: Bedarf, Konzept und Programmatik*. In: Gretsche, Petra/Wulff, Nadja (Hgg.). *Deutsch als Zweit- und Fremdsprache in Schule und Beruf*. Paderborn, Schöningh, 304–316.

Roche, J./Finkbohner, A. (2021). *Berufssprache Deutsch als durchgängiges Konzept in der beruflichen Aus- und Weiterbildung. Zur fachpolitischen Bedeutung der neuesten KMK-Beschlüsse*. *Zeitschrift für Interkulturellen Fremdsprachenunterricht (ZIF)*, 1/2021.

2022

Arras, U. (2022). The Language of Europe is Mediation: Sprachmittlung in mehrsprachigen (Arbeits-)Situationen. In K. Aguado, S. Hoffmann & M. Hoshii (Hrsg.), *Jahrbuch für internationale Germanistik 2021*. Frankfurt am Main: Peter Lang, 513–525.

Arras, U. (2022). Wissensverarbeitung in Bildungskontexten: Mitschreiben und Exzerpieren als Sprachhandlungen zwischen Wissensvermittlung und Sprachmittlung. In P. Katelhön & P. Marecková (Hrsg.), *Sprachmittlung und Mediation im schulischen und universitären Fremdsprachenunterricht*. Berlin: Frank und Timme, 227–246.

Depner, G. & Peters, A. (2022) Sprachkompetenzen entwickeln und trainieren: Ein Konzept für eine kompetenzorientierte Prüfungsvorbereitung, *Informationen Deutsch als Fremdsprache* 49 (4), 325–345.

Drackert, A., Küplüce, C., Werner, S. (2022). Digitalisierungsbezogene Kompetenzen von Lehrkräften – Eine Konkretisierung fächerübergreifender Kompetenzmodelle aus fremdsprachendidaktischer Perspektive. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 27 (1), 261–283.

Drackert, A., Konzett-Firth, C., Stadler, W., & Visser, J. (2022). Studium und Referendariat – war's das? Zur Fort- und Weiterbildung von Lehrkräften romanischer Sprachen in Deutschland und Österreich. *Zeitschrift für romanische Sprachen und ihre Didaktik*, 16, 123–156.

Felberg, D., Hülsmann, N., Propp, K. & Drackert, A. (2022). Einsatz von Digital Storytelling im herkunftssprachlichen Russischunterricht an einer Samstagsschule. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 27: 2, 51–77. <https://doi.org/10.48694/zif.3497>

Eckes, T. & Jin, K.-Y. (2022). Detecting illusory halo effects in rater-mediated assessment: A mixture Rasch facets modeling approach. *Psychological Test and Assessment Modeling*, 64(1), 87–111.

Grotjahn, R. & Drackert, A. (2022). *The electronic C-Test bibliography: version October 2022*. https://www.gast.de/fileadmin/gast.de/GAST/4_PDF/2-Forschung-Entwicklung/Publikationen/Grotjahn_Drackert_Electronic_Ctest_Bibliography_10_2022.pdf

Jin, K.-Y. & Eckes, T. (2022). Detecting differential rater functioning in severity and centrality: The dual DRF facets model. *Educational and Psychological Measurement*, 82(4), 757–781.

Jin, K.-Y. & Eckes, T. (2022). Detecting rater centrality effects in performance assessments: A model-based comparison of centrality indices. *Measurement: Interdisciplinary Research and Perspectives*, 20(4), 228–247.

Kecker, G. & Eckes, T. (2022). Der digitale TestDaF: Aufbruch in neue Dimensionen des Sprachtestens. *Informationen Deutsch als Fremdsprache*, 49(4), 289–324.

Kecker, G., Zimmermann, S. & Eckes, T. (2022). Der Weg zum digitalen TestDaF: Konzeption, Entwicklung und Validierung. In P. Gretsche & N. Wulff (Hrsg.), *Deutsch als Zweit- und Fremdsprache in Schule und Beruf* (S. 393–410). Paderborn: Brill Schöningh.

Korflür, D., Marks, D., Weiss-Motz, F. (2022). Analytisch oder holistisch? Welchen Einfluss hat die Beurteilungsmethode auf das Verhalten von Beurteilenden und die Ergebnisse von Prüfungsteilnehmenden?, *Informationen Deutsch als Fremdsprache* 49 (4), 346–368.

Laarmann-Quante, R., Schwarz, L., Horbach, A., & Zesch, T. (2022). 'Meet me at the ribary'–Acceptability of spelling variants in free-text answers to listening comprehension prompts. *Proceedings of the 17th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2022)*, 173–182.

Mainzer, M. & Drackert, A. (2022). Erfassung von Sprachkompetenzen im Deutschen von Schülerinnen und Schülern mittels C-Tests im Arbeitsfeld Sprachbildung: Fragen der Validität revisited. *Informationen Deutsch als Fremdsprache*, 22, 470–492.

Roche, J. (2022). Der lange Weg zu einer pragmatischen Sprachenpolitik. In: Li, Yuan/Liu, Fang/Wang, Zhongxin (Hgg.). *Didactica, Cultura, Lingua – Perspektive des Deutschen*. München, iudicium, 40–50.

Roche, J./Zahn, P. (2022). *Digitale und analoge Best Practice in der Sprach- und Integrationsförderung. Erkenntnisse aus dem Projekt „Lernen, Lehren, Helfen“*. In: Birkner, Karin/Hufeisen, Britta/Rosenberg, Peter. *Spracharbeit mit Geflüchteten*. Berlin: Lang: 319–348.

Roche, J. (2022). *Mehrsprachigkeit und digitale Bildung – Plädoyer für Pragmatik*. ZMI Magazin Köln, 2022: 12–14.

Zimmermann, S. (2022). *Validating integrated writing tasks – A mixed-method approach to investigate the construct of summarization*. [Dissertation, Universität Bremen]. <https://doi.org/10.26092/elib/2207>

Zimmermann, S., Schwarz, L., Peters, A. & Depner, G. (in Druck). Enhancing teachers and test takers' assessment literacy? Insights from test preparation for the digital TestDaF. In B. Baker & L. Taylor (Eds.), *Learning about assessing language: a matter of literacy or competency development?* Cambridge: Cambridge University Press.

