



**Bericht**  
**Machbarkeitsanalyse zur Entwicklung von On-**  
**line-Sprachtests**

07.02.2022

Liana Konstantinidou & Karin Madlener-Charpentier

Unter Mitarbeit von: Curtis Gautschi, Hrisztalina Hrisztova-Gotthardt, Elsa Liste Lamas,  
Urs Egli

Unterstützt durch die PraktikantInnen: Yannik Aebi, Aline Criblez, Maria Ilcenko

## Inhaltsverzeichnis

<b>1. ZIELSETZUNG .....</b>	<b>5</b>
<b>2. MANAGEMENT SUMMARY.....</b>	<b>6</b>
<b>3. METHODISCHES VORGEHEN DER MACHBARKEITSANALYSE.....</b>	<b>9</b>
3.1 Desk-Research .....	9
3.2 Interviews .....	12
<b>4. ÜBERSICHT ÜBER AUFGABENTYPEN, AUFGABENGESTALTUNG UND NUTZERINTERFACE BEI DIGITALEN SPRACHTESTS.....</b>	<b>14</b>
4.1 Typische digitale Aufgabenformate: Multiple-Choice- und Zuordnungsaufgaben.....	14
4.2 Hör- und Lesekompetenzen prüfen .....	21
4.3 Schreibkompetenzen prüfen .....	23
4.4 Automatische Auswertungsmöglichkeiten .....	26
4.5 Sprechkompetenzen prüfen .....	26
4.5.1 Monologisches Sprechen (mit eigener Aufnahme).....	26
4.5.2 Dialogisches Sprechen (im Video-Call) .....	29
4.6 Digitale Umsetzungsmöglichkeiten der Aufgabentypen im aktuellen fide-Test .....	30
4.7 Arbeitsanweisungen und Nutzerinterface .....	31
4.8 Adaptivität und Testkonstrukt.....	33
4.9 Medien, Infrastruktur und Sicherheit .....	34
4.10 Support .....	36
4.10.1 Support vor und während der Prüfung .....	36
4.10.2 Herausforderungen .....	38
<b>5. TECHNISCHE TOOLS FÜR DIE DIGITALE DURCHFÜHRUNG UND DURCHFÜHRUNGSSICHERHEIT (DATENSICHERHEIT, TÄUSCHUNGSSICHERHEIT, DATENSCHUTZ) .....</b>	<b>39</b>
5.1 Durchführungssicherheit (Termin- und Datensicherheit).....	39
5.2 Täuschungssicherheit (1): Identitätsprüfung .....	40
5.3 Täuschungssicherheit (2): Proctoring und Datenschutz .....	40
5.4 Technische Tools für automatische Korrektur .....	44

<b>6. AKZEPTABILITÄT UND ZUMUTBARKEIT .....</b>	<b>45</b>
6.1 Integrationsbeauftragte, Geschäftsstelle fide.....	45
6.2 Anbieterinstitutionen .....	46
6.3 Akzeptabilitätsbedingung „Papierversion“? .....	46
<b>7. PASSUNG FÜR DIE ZIELGRUPPE(N) UND HERAUSFORDERUNGEN.....</b>	<b>48</b>
7.1 Digitale Kompetenzen .....	48
7.2 Schriftsprachliche Kompetenzen .....	49
7.3 Den Prüfungsteilnehmenden zur Verfügung stehende technische Infrastruktur.....	50
7.4 (Lern-)Psychologische Funktionsfähigkeiten .....	50
<b>8. POTENZIALE DIGITALER PRÜFUNGEN.....</b>	<b>51</b>
8.1 Angebotsoptimierung.....	51
8.2 Realitätsnähe .....	51
8.3 Digitale Grundkompetenzen.....	51
<b>9. KOSTEN / AUFWAND – SPARPOTENZIALE.....</b>	<b>53</b>
9.1 Digitaler Test vor Ort: Aufwand / Kosten.....	53
9.2 Digitaler Test vor Ort: Sparpotenziale.....	54
9.3 Digitaler ortsunabhängiger Test: Aufwand / Kosten für die Entwicklung.....	54
9.3.1 Beispiel E-TEF / DFP (digital vor Ort).....	55
9.3.2 Beispiel TestDaF digital (vor Ort).....	55
9.3.3 Beispiel ECL, Ungarn (vor Ort und remote).....	56
9.3.4 Beispiel Goethe-Institut (aktuell: vor Ort: geplant: remote).....	56
9.4 Digitaler ortsunabhängiger Test: Aufwand / Kosten in der Durchführung .....	58
9.5 Digitaler ortsunabhängiger Test: Sparpotenziale .....	58
<b>10. HANDLUNGSOPTIONEN, SPIELRÄUME UND SZENARIEN .....</b>	<b>59</b>
10.1 Tools und Plattform/en .....	59
10.1.1 Content Management System (Item-Datenbank, Nutzerinterface) und Administration .....	59
10.1.2 Proctoring.....	60
10.2 Szenarien (1): Papier weiter als Option.....	60
10.2.1 Szenario 1: Nur Papierversion .....	60
10.2.2 Szenario 2: Parallel-Versionen Papier und digital vor Ort.....	61
10.2.3 Szenario 3: Parallel-Versionen Papier und digital remote .....	62
10.2.4 Szenario 4: Parallel-Versionen Papier, digital vor Ort und digital remote .....	64

<b>10.3 Szenarien (2): Ausschliesslich digitales Angebot .....</b>	<b>64</b>
10.3.1 Szenario 5: Nur digitale Vor-Ort-Version .....	64
10.3.2 Szenario 6: Parallelversionen digital vor Ort und digital Remote .....	65
10.3.3 Szenario 7: Nur digitale Remote-Version .....	66
<b>10.4 Überblick der Szenarien im Vergleich: Merkmale, Pro-/Contra-Argumente, Gelingensbedingungen .....</b>	<b>67</b>
<b>10.5 Fazit und Empfehlung .....</b>	<b>69</b>
10.5.1 Abschliessende Bewertung Szenario 3 .....	69
10.5.2 Abschliessende Bewertung Szenario 6 .....	70
10.5.3 Handlungsempfehlungen .....	71
<b>LITERATURVERZEICHNIS .....</b>	<b>72</b>

## 1. Zielsetzung

Der fide-Test ist als Nachweis von Sprachkompetenzen in ausländer- und bürgerrechtlichen Verfahren in der Schweiz anerkannt. Er ist spezifisch auf die Situation der Schweiz ausgerichtet und soll die Evaluierung sprachlicher Kompetenzen im Kontext der Integration erlauben. Der fide-Test existiert aktuell nur in einer Papierform und kann nur vor Ort durchgeführt werden.

Angestossen durch die Covid-19-Pandemie und die Einschränkungen bei der Prüfungsdurchführung soll nun geprüft werden, welche Chancen/Potenziale sich durch eine Digitalisierung des fide-Tests, insbesondere in einer ortsunabhängigen Variante, ergeben würden.

Die vorliegende Studie analysiert und evaluiert auf Basis einer Literaturrecherche und einer Reihe von ExpertInnen-Interviews die Optionen für

- a. die technischen Möglichkeiten zur Erstellung, Einspielung, Darstellung und Auswertung von verschiedenen Aufgabentypen, Optionen für Item-Datenbanken und die Gestaltung des Nutzerinterface,
- b. die Durchführung und Täuschungs-/Durchführungssicherheit von digitalen ortsabhängigen bzw. digitalen ortsunabhängigen bzw. hybriden Varianten der Sprachkompetenzevaluierung, inklusive Datensicherheit und Datenschutz,
- c. die Eignung existierender Tools in Bezug auf die spezifischen Bedarfe und Herausforderungen im fide-Kontext (angesichts u.a. der Spezifik der Teststruktur, der Testziele, des Test-impacts, der Zielgruppen).

Kapitel 2 bietet eine Zusammenfassung der wichtigsten Aspekte der Analyse. Kapitel 3 beschreibt das methodische Vorgehen der Analyse. Im Kapitel 4 werden technische Möglichkeiten und Anforderungen für die Darstellung und Einspielung von Aufgaben sowie das Nutzerinterface dargestellt. Kapitel 5 analysiert technische Optionen für die Durchführungssicherheit im Sinne der Täuschungs- und Datensicherheit sowie des Datenschutzes. Kapitel 6 und 7 beleuchten die Akzeptabilität digitaler Tests und die Eignung bestehender Möglichkeiten im spezifischen Kontext des fide-Tests. Kapitel 8 analysiert die Potenziale einer Digitalisierung des fide-Tests; Kapitel 9 den Aufwand, die Kostenfolgen und Sparpotenziale. In Kapitel 10 werden die Befunde in neun Szenarien zusammengefasst, Handlungsoptionen identifiziert und Handlungsempfehlungen skizziert. Im letzten Kapitel werden die Erkenntnisse aus der Analyse aus Sicht des Projektteams diskutiert.

## 2. Management Summary

### **Machbarkeitsanalyse für die Digitalisierung des fide-Tests: Auftrag und methodisches Vorgehen**

Im Rahmen der Machbarkeitsanalyse für die Digitalisierung des fide-Tests beauftragt das Staatssekretariat für Migration die ZHAW, Zürcher Hochschule für Angewandte Wissenschaften, Möglichkeiten für die Digitalisierung des fide-Tests zu prüfen, dabei Potenziale und Herausforderungen darzustellen und Szenarien für Handlungsoptionen zu skizzieren. Dazu wird eine Desk Research durchgeführt, welche der Analyse bestehender digitaler Sprachtests in Hinblick auf Zielgruppen, Testzwecke, Testformate, Aufgabentypen, Datensicherheit und Datenschutz sowie Täuschungs- und Durchführungssicherheit dient. Gleichzeitig finden Interviews mit internationalen TestexpertInnen und Stakeholdern im Bereich der sprachlichen Integration in der Schweiz statt, welche die genannten Aspekte vertiefen und/oder Fragen der Akzeptanz und der Sinnhaftigkeit einer Digitalisierung des fide-Tests beantworten sollen.

### **Aufgabentypen, Aufgabengestaltung und Nutzerinterface bei digitalen Sprachtests**

Die Analyse bestehender digitaler Tests zeigt, dass es gute technische Möglichkeiten gibt, Aufgabentypen so zu gestalten, dass schulungsgewohnte Teilnehmende nicht benachteiligt werden. Das Nutzerinterface und die Usability müssten jedoch mit verschiedenen Zielgruppen gut erprobt werden. Insbesondere scheint eine digitale Prüfung eher unproblematisch zu sein im Hinblick auf die Komponenten Hörverstehen und Leseverstehen. Eher wenig problematisch scheint in einer digitalen Prüfung auch die Umsetzung des Prüfungsteils "Sprechen" zu sein, da dieser Prüfungsteil entweder weiter live als Face-to-face-Gespräch durchgeführt werden kann (digital vor Ort) oder relativ unproblematisch als Video-Call (digital remote). Die notwendige, qualitativ hochwertige und zuverlässige Infrastruktur wären entweder institutions- oder teilnehmerseitig Voraussetzung.

Am ehesten problematisch erscheint in einer digitalen Prüfung die Umsetzung des Anteils "Schreiben" – und dies, obwohl das Schreiben mit digitalen Medien den Prüfungsszenarien mehr entspricht und lebensnäher als das Schreiben auf Papier ist (z.B. für Mails, SMS). Die grösste Herausforderung stellt das Schreiben auf einer Tastatur dar. Mögliche Lösungsansätze wären: (a) eine Erhöhung der Bearbeitungszeit; (b) das Tippen auf einer digitalen Tastatur wie auf dem Smartphone; (c) im Falle einer Vor-Ort-Version die Möglichkeit, den geforderten Text auf Papier zu schreiben; und (d) das Schreiben mit einem digitalen Stift.

### **Digitale Durchführung und Durchführungssicherheit**

Je nach Durchführung digital vor Ort bzw. digital remote ergeben sich teilweise unterschiedliche Anforderungen hinsichtlich der Daten- und Täuschungssicherheit. In Bezug auf die Datensicherheit steht einerseits der Schutz vor Datenverlusten im Fokus. Für die Durchführung vor Ort müssen die Testzentren die Zuverlässigkeit der Internetverbindung für den Abruf der Aufgaben aus der Aufgaben-Datenbank sowie die Speicherung der Eingaben der Testteilnehmenden garantieren. Im Falle der Remote-Durchführung liegt die Verantwortung für die Verbindungszuverlässigkeit bei den Testteilnehmenden, es muss daher ein klares Protokoll für den Umgang mit Verbindungsabbrüchen geben (Zeitkompensation, Rekursgründe etc.). Andererseits muss der Schutz der Daten vor unerlaubtem Zugriff gewährleistet sein, v.a. in Bezug auf sensible persönliche Daten, die in der Remote-Variante für die Identitätsprüfung (Gesichtserkennung) und maschinelles Proctoring (Blickbewegungen etc.) notwendig sind.

Die grössten Herausforderungen sind im Hinblick auf die Täuschungssicherheit zu erwarten. Im Falle einer Durchführung vor Ort kann die Identitätsprüfung und Prüfungsaufsicht wie bisher durch Aufsichtspersonen vor Ort durchgeführt werden. Es ist allerdings damit zu rechnen, dass dafür tendenziell mehr Aufsichtspersonen nötig sind und diese eine entsprechende technische Ausbildung bzw. Schulung brauchen; auch muss jederzeit ein spezifischer technischer Support online oder vor Ort abrufbereit sein. Im Falle einer Remote-Durchführung müssen Identitätsprüfung und Prüfungsaufsicht online durchgeführt werden, und zwar entweder – personell aufwändig – durch menschliches Proctoring oder – datenschutzrechtlich aufwändig und für die Testteilnehmenden technisch komplex – durch maschinelles Proctoring. Die entsprechenden Entscheidungen müssen frühzeitig getroffen und die datenschutzrechtlichen Folgen frühzeitig in die Abklärung gegeben werden, da dies erfahrungsgemäss viel Zeit in Anspruch nimmt und spätere Prozesse erschweren kann. Die Datenschutzabklärung muss von ExpertInnen begleitet werden.

## Akzeptabilität und Zumutbarkeit

Konsens besteht bei allen Schweizer Stakeholdern, dass eine digitale Prüfungsversion nicht die fide-Kernzielgruppe der schul- und schriftungewohnten Lernenden benachteiligen oder gar ausschliessen darf. Besonders anspruchsvoll in Hinblick auf schriftsprachliche und digitale Kompetenzen wird die Remote-Variante eingeschätzt, insbesondere bei der Option eines maschinellen Proctoring-Verfahrens, welches ein aufwändiges Prozedere der Identitätsprüfung sowie die Notwendigkeit zum eigenständigen Herunterladen und Installieren verschiedener Applikationen umfasst. Zudem betonen sowohl die nationalen als auch die internationalen ExpertInnen die Wichtigkeit einer menschlichen Begleitung aus lernpsychologischer Perspektive (Stichwort «Prüfungsangst»).

Schulungsgewohnte Lernende zeigen sich gegenüber einer digitalen Version offen, wenn sie die Möglichkeit haben, sich mit dem Prüfungsformat vertraut zu machen. Sie nennen explizit das Schreiben auf Tastatur als die grösste Herausforderung und bevorzugen eine Durchführung auf Smartphone oder Tablet gegenüber einer auf PC oder Laptop; ausserdem fühlen sie sich sicherer, wenn sie den Test in Testzentren und nicht remote absolvieren können.

Seitens der nationalen AkteurInnen der sprachlichen Integration werden darüber hinaus generellere Fragen bezüglich der Etablierung eines neuen Testsystems gestellt, das eventuell neue Strukturen, neue Verantwortlichkeiten der Anbieter-Institutionen und der Geschäftsstelle fide, neue PrüferInnen-Kompetenzprofile, neue Qualitätsstandards etc. bedingt. Im Falle eines Entscheids für die Entwicklung eines digitalen Tests empfiehlt es sich deshalb von Anfang an, ein gutes Kommunikationskonzept und Begleitmassnahmen zu entwickeln, die erlauben, die verschiedenen Zielgruppen rechtzeitig in den Prozess zu involvieren und die oben genannten Fragen zielführend zu klären.

## Potenziale digitaler Prüfungen

Neben den Risiken und Herausforderungen digitaler Prüfungen ergeben sich Chancen und Potenziale. Dazu gehört, im Falle einer Remote-Prüfung, die Durchführungssicherheit, Ortsunabhängigkeit und Termiflexibilität. In diesem Sinne kann eine Digitalisierung zur Angebotsoptimierung beitragen und den fide-Test auch für Testteilnehmende jenseits der fide-Kernzielgruppe attraktiv machen.

Weiterhin kann in einer digitalen Prüfung, bei tatsächlicher Ausnutzung des digitalen Mediums und seiner Spielräume, die Realitätsnähe der Prüfungsaufgaben erhöht werden; z.B. würden dann digitale Textsorten wie SMS oder E-Mails tatsächlich auch digital gelesen oder verfasst.

Schliesslich ergibt sich im Hinblick auf das zentrale Stichwort „digitale Fertigkeiten“ neben der Herausforderung, die Kernzielgruppe des fide-Tests und ihre Bedürfnisse angemessen zu berücksichtigen, auch ein möglicher positiver Nebeneffekt; dies in Form einer potenziellen Beschleunigung des digitalen Wandels beim Lernen und Lehren von Sprachen, u.a. in der Aus- und Weiterbildung der Lehrpersonen.

## Kosten/Aufwand – Sparpotenziale

Allgemein ist damit zu rechnen, dass die Digitalisierung personell und finanziell aufwändig ist und Zeit beansprucht, v.a. der Anteil der Programmierung (von Nutzerinterface, Aufgaben-Datenbank, Aufgaben-Darstellung, Proctoring, Prüfungsadministration etc.) und ggf. die Neukonzeptualisierung von (Teilen der) Aufgaben. Je nach Szenario entstehen unterschiedliche Kosten und Sparpotenziale für ortsabhängige bzw. ortsunabhängige digitale Prüfungen.

Die unterschiedlichen Aufwände beziehen sich z.B. auf die technische Infrastruktur, ihre Wartung und Aktualisierung im Falle der Vor-Ort-Prüfung, auf die Entwicklung und Durchführung des Proctoring und den entsprechenden Datenschutz im Falle der Remote-Version. Allgemein ist mit Aufwänden u.a. für Kommunikation, Schulungen, Datensicherheit, Support und Begleitforschung zu rechnen. Um den Aufwand bei der Programmierung allgemein zu reduzieren, kann die Möglichkeit geprüft werden, bestehende Plattformen zu lizenzieren und ggf. teilweise anzupassen.

Die Variante “digital remote” verschiebt das Problem der (Kosten für die) Infrastruktur auf die Testteilnehmenden, die mindestens über (a) zwei reibungslos funktionierende Geräte; (b) einen ruhigen Ort; und (c) eine sichere, stabile Internetverbindung verfügen müssen. Hardwarebedingte Benachteiligungen einzelner Testteilnehmender sind unbedingt auszuschliessen.

Auch die Sparpotenziale unterscheiden sich je nach Prüfungsart und Szenario, z.B. im Hinblick auf Anreise, Infrastruktur (Geräte, Räume) und Personal. Allgemein ist vorrangig mit Sparpotenzialen bei der Prüfungsverwaltung zu rechnen, teilweise auch bei der Prüfungskorrektur. Eine Remote-Variante hat institutionsseitig das grösste langfristige Sparpotenzial. Eine gewisse Rentabilität ist je nach Szenario eventuell längerfristig bei konstant hoher Nachfrage der digitalen Prüfungen zu erreichen.

### **Handlungsoptionen, Spielräume und Szenarien**

Die sieben erarbeiteten Szenarien, die von der Weiterführung nur der aktuellen Papier-Version bis zur Entwicklung einer Remote-Version ohne weitere Parallelversionen reichen, weisen Chancen und Risiken auf (siehe Tabelle 6, S. 66ff.), die im Hinblick auf die Zielsetzungen des Auftraggebers genauer zu prüfen sind. Angesichts der fortschreitenden Digitalisierung in der Gesellschaft und der Entwicklungen im Bereich des Lehrens/Lernens und Prüfens empfehlen wir insbesondere die Szenarien zu prüfen, die eine Digitalisierung des fide-Tests vorsehen. Diese gewährleistet zumindest die Attraktivität des Angebots (Ortsunabhängigkeit, Terminflexibilität), die Positionierung des fide-Test auf dem sich rasant verändernden internationalen Prüfungsmarkt und, im Falle einer Remote-Version, die Durchführungssicherheit.

### 3. Methodisches Vorgehen der Machbarkeitsanalyse

Im Rahmen der Machbarkeitsanalyse zur Entwicklung von Online-Sprachtests wurden zwei parallele Rechercheansätze verfolgt, einerseits eine umfassende Desk Research zu 24 aktuellen Sprachtests im Bereich allgemeine Sprachkenntnisse, Hochschulzugang und Aufenthalt / Einbürgerung, mit einem Fokus auf digitalen ortsabhängigen und ortsunabhängigen Tests; andererseits eine Serie von 16 Einzel- und Fokusgruppeninterviews mit verschiedenen Stakeholdern der sprachlichen Integration und des digitalen Testens und Prüfens.

#### 3.1 Desk-Research

In der Desk Research wurden für 24 nationale und internationale Sprachtests (Zielsprachen: Deutsch, Englisch, Französisch, Italienisch, Spanisch, Norwegisch, Bulgarisch, Ungarisch) grundlegende Informationen in Bezug auf folgende Aspekte erhoben: (a) Zielgruppe, (b) Testziele (Zertifizierung, Placement), (c) Durchführungsformate (Papier, digital vor Ort, digital remote), (d) Durchführungsbedingungen, (e) Aufgabenformate und Optionen der Einbindung von Bild-, Audio- und Videodateien, (f) Adaptivität, (g) Darstellungsoptionen und Gestaltung des Nutzerinterface, (h) Durchführungsmedien (PC, Laptop, Tablet, Smartphone), (i) Durchführungsanforderungen (Hardware, Software, *digital skills*, *literacy skills*), (j) Auswertungsoptionen, (k) Täuschungssicherheit, (l) Datensicherheit, (m) Datenschutz, (n) Validierung und Begleitforschung. Eine Übersicht der berücksichtigten Tests findet sich in Tabelle 1.

**Tab. 1: Übersicht über die in der Desk Research berücksichtigten Tests**

Name des Tests	Zielsprache	Testtyp	Testzweck	Format
TestDaF digital	Deutsch	Zertifizierung B2-C1	Hochschulzugang	digital vor Ort
SDS	Deutsch (CH)	Zertifizierung A1-B2	Aufenthalt/ Einbürgerung	digital vor Ort
Deutschtest für Zuwanderer	Deutsch	Zertifizierung A2-B1	Aufenthalt/ Einbürgerung	Papier
Goethe digital	Deutsch	Zertifizierung, bisher B1-B2	allgemeine Kenntnisse	digital vor Ort
ÖSD-Zertifikat	Deutsch	Zertifizierung A1-B1	Aufenthalt/ Einbürgerung	Papier
Deutschtest für Österreich (Österreichischer Integrationsfonds, Telc, Goethe-Institut)	Deutsch	Zertifizierung A2-B1	Aufenthalt/ Einbürgerung	Papier
E-TEF	Französisch (F, CAN, CH)	Zertifizierung	Hochschulzugang, Aufenthalt/ Einbürgerung	digital vor Ort
TEF express	Französisch (CAN)	Zertifizierung	Hochschulzugang	remote
TCF	Französisch (F, CAN, CH)	Zertifizierung	Hochschulzugang, Aufenthalt/ Einbürgerung	Papier, digital vor Ort
CELI Immigrati	Italienisch	Zertifizierung A2-B1	Aufenthalt/ Einbürgerung	Papier
Duolingo	Englisch	Placement	allgemeine Kenntnisse, Hochschulzugang	remote

EFSET	Englisch	Placement	Hochschule, Arbeitsmarkt	remote: Desktop, Tablet, mobiles Gerät
Pearson Versant	Englisch (USA)	Placement	VISA, Placement, Hochschulzugang	digital vor Ort; remote seit 30. April 2020
Cambridge Linguaskills Anywhere	Englisch	Placement	Hochschule, Arbeitsmarkt	remote
Language Testing International (ACTFL)	Englisch	Zertifizierung	Arbeitsmarkt	remote
TOEFL iBT / TOEFL iBT at home	Englisch	Zertifizierung	Hochschulzugang	digital vor Ort; remote
TOEFL ITP Home Edition	Englisch	Classroom Achievement, A2-C1	allgemeine Kenntnisse	Papier, digital vor Ort, remote
IELTS Migration	Englisch (UK, CAN, AUS, NZ)	Zertifizierung	Hochschulzugang, Aufenthalt/ Einbürgerung	Papier, digital vor Ort; in Pandemie: auch remote, aber nur für HSZ
IELTS Life Skills for UKVI (UK)	Englisch	Zertifizierung A1-C2	Aufenthalt/Einbürgerung	Papier, digital vor Ort
CELP	Englisch (CAN)	Zertifizierung	Hochschulzugang, Aufenthalt/ Einbürgerung	digital vor Ort
EuroExam	Deutsch, Englisch	Zertifizierung B1-C1	allgemeine Kenntnisse	digital vor Ort, remote
ECL Sprachprüfungen	Deutsch, Englisch, Ungarisch	Zertifizierung A2-C1	allgemeine Kenntnisse	digital vor Ort, remote
Digital Test of Bulgarian as a foreign language	Bulgarisch	Zertifizierung A2-C2 (B1-B2 ALTE-anerkannt)	allgemeine Kenntnisse	digital vor Ort
SIELE	Spanisch (SPA, MEX, ARG)	Zertifizierung A1-C1	allgemeine Kenntnisse, Hochschulzugang	digital vor Ort oder remote
Norwegian Language Test	Norwegisch (Bokmål, Nynorsk)	Zertifizierung A1-B2	u.a. Aufenthalt/ Einbürgerung	digital vor Ort

Insgesamt sind aktuell auf dem Markt verfügbare **digitale ortsabhängige und ortsunabhängige Tests** in erster Linie Placement- oder Zertifizierungstests für allgemeine Sprachkenntnisse, die u.a. für den Zugang zum Arbeitsmarkt bzw. zur tertiären Bildung (Hochschulzugang) eingesetzt werden. Eine Auswahl von digital ortsabhängig (in zertifizierten Testzentren) bzw. ortsunabhängig (remote) verfügbaren Placement- und Zertifizierungstests ist in Tabelle 2 zusammengefasst.

**Aufenthalts- und einbürgerungsrelevante High-stakes-Tests** werden tendenziell immer noch in Papierform oder digital vor Ort durchgeführt; das gilt selbst für das Englische als Zielsprache (mit der Ausnahme einiger neuer Entwicklungen wie TOEFL iBT Home und einigen ACTFL-Tests in den USA).

Ausgeschlossen ist eine Remote-Durchführung zumindest technisch nicht, allerdings anspruchsvoll im Hinblick auf die Täuschungssicherheit und datenschutztechnische Aspekte (s.u.). Letzteres gilt vor allem,

wenn der Test von verschiedenen Ländern aus zugänglich sein soll (v.a. von Drittstaaten jenseits der EU und assoziierter Staaten aus).

Aktuelle Tests im Kontext ausländer- und bürgerrechtlicher Anwendungen (Aufenthalt, Niederlassung, Einbürgerung) sind in Tabelle 3 zusammengefasst. Zum *Norwegian Language Test* muss angemerkt werden, dass bereits die entsprechenden Sprachkurse explizit zwei Ziele verfolgen, nämlich die Entwicklung von Sprachkompetenzen und die Entwicklung von Grundkompetenzen inkl. digitaler Skills.

**Tab. 2: Übersicht digital ortsabhängige und ortsunabhängige Placement- und Zertifizierungstests**

	Digital ortsabhängig (in Testzentren)	Digital ortsunabhängig (remote)
<b>Zielsprache Deutsch</b>	<ul style="list-style-type: none"> <li>• TestDaF digital</li> <li>• Goethe digital B1-B2</li> <li>• EuroExam</li> <li>• ECL</li> </ul>	<ul style="list-style-type: none"> <li>• EuroExam Online</li> <li>• ECL Online</li> <li>• <i>in Entwicklung</i>: Goethe-Zertifikat A1 Start Deutsch 1</li> </ul>
<b>Zielsprache Englisch</b>	<ul style="list-style-type: none"> <li>• TOEFL iBT</li> <li>• TOEFL ITP</li> <li>• Pearson Versant</li> <li>• EuroExam</li> <li>• ECL</li> </ul>	<ul style="list-style-type: none"> <li>• Duolingo</li> <li>• EFSET</li> <li>• Cambridge Linguaskills Anywhere</li> <li>• Language Testing International</li> <li>• IELTS Migration (nur für Hochschulzugang)</li> <li>• TOEFL iBT</li> <li>• TOEFL ITP</li> <li>• Pearson Versant (seit April 2020)</li> <li>• EuroExam Online</li> <li>• ECL Online</li> </ul>
<b>Zielsprache Französisch</b>	<ul style="list-style-type: none"> <li>• DFP (Diplôme de Français Professionnel)</li> </ul>	<ul style="list-style-type: none"> <li>• TEF Express (nur Kanada, nur für Hochschulzugang; temporäre Version, es wird aber u.a. für weit abgelegene Regionen an einer längerfristigen Remote-Version gearbeitet)</li> </ul>
<b>Andere Zielsprachen</b>	<ul style="list-style-type: none"> <li>• SIELE (Spanisch)</li> <li>• ECL (Ungarisch)</li> <li>• Digital Test of Bulgarian as a Foreign Language</li> </ul>	<ul style="list-style-type: none"> <li>• SIELE (Spanisch)</li> </ul>

**Tab. 3: Überblick analoger und digitaler Tests im Kontext Aufenthalt und Einbürgerung**

	Papier-Format	Digital ortsabhängig (in Testzentren)	Digital ortsunabhängig (remote)
<b>Zielsprache Deutsch</b>	<ul style="list-style-type: none"> <li>• fide-Test (Schweiz)</li> <li>• Deutshtest für Zuwanderer (Deutschland)</li> <li>• Deutshtest für Österreich</li> </ul>	<ul style="list-style-type: none"> <li>• SDS (Schweiz)</li> </ul>	<ul style="list-style-type: none"> <li>• <i>in Entwicklung</i>: Goethe-Zertifikat A1 Start Deutsch 1</li> </ul>
<b>Zielsprache Englisch</b>	<ul style="list-style-type: none"> <li>•</li> </ul>	<ul style="list-style-type: none"> <li>• IELTS for Migration</li> <li>• IELTS Life Skills UKVI</li> <li>• CELPIP</li> </ul>	<ul style="list-style-type: none"> <li>•</li> </ul>
<b>Andere Zielsprachen</b>	<ul style="list-style-type: none"> <li>• CELI Immigrati</li> </ul>	<ul style="list-style-type: none"> <li>• E-TEF</li> <li>• TCF</li> <li>• Norwegian Language Test</li> </ul>	

### 3.2 Interviews

In den Interviews wurden verschiedene ergänzende, komplementäre und weiterführende Aspekte der Fragestellung bearbeitet; dazu wurden verschiedene nationale und internationale Stakeholder- und ExpertInnengruppen befragt, nämlich

- VertreterInnen der Geschäftsstelle fide
- kantonale Integrationsbeauftragte
- Anbieterinstitutionen
- TestexpertInnen, inkl. führende ExpertInnen der *Association of Language Testers in Europe (ALTE)* und des *American Council on the Teaching of Foreign Languages (ACTFL)*
- EntwicklerInnen und AnbieterInnen von digitalen ortsabhängigen und ortsunabhängigen Tests
- ProctoringentwicklerInnen und ProctoringanbieterInnen,
- DatenschutzexpertInnen
- Lernende (Tab. 4)

Sie wurden, je nach Expertisedomäne, u.a. befragt zu

- Potenzialen und Herausforderungen digitaler Sprachtests (z.B. in Bezug auf Entwicklung, Durchführung, Adaptivität, IT- und Datensicherheit, Datenschutz, Täuschungssicherheit)
- Aufwand und Kosten
- Akzeptabilität digitaler Tests und Zielgruppenspezifik
- Optionen und Best Practices

**Tab. 4: Übersicht über InterviewpartnerInnen und Interviewdaten**

Institution	Funktion	Interviewdatum
<b>ExpertInnen für Testen im Bereich Migration / Integration</b>		
Geschäftsstelle fide	VertreterInnen der aktuellen und künftigen Leitung der Geschäftsstelle	19.10.2021
Integrationsbeauftragte	2 deutschsprachige, 1 französischsprachige/r	12.10.2021 21.10.2021 02.11.2021
Bürgergemeinde BS	Juristin / Anbieterin	03.11.2021
Anbieterinstitutionen	Leitung; 1 deutschsprachige., 1 französischsprachige	13.10.2021 29.10.2021
Goethe-Institut	Leitung Prüfungsentwicklung; Leitung digitale Deutschprüfung	20.01.2022
Goethe-Prüfungszentrum ZHAW	Leiter Prüfungszentrum	05.11.2021
ALTE	Ko-Leitung der Special Interest Group <i>Language Assessment for Migration and Integration</i> , C. Hamnes / L. Rocca	22.10.2021
<b>ExpertInnen für digitales Prüfen</b>		
ALTE	Ko-Leitung der Special Interest Group <i>Technology in Language Assessment</i> , D. Casanova	09.11.2021
TestDaF	Ko-Leitung, G. Kecker, T. Eckes; Entwickler, Anbieter	02.11.2021

ACTFL	Director of the Assessment and Evaluation Language Resource Center & Center for Assessment, Research and Development, M. Malone	18.10.2021
ECL Internationales Prüfungszentrum an der Universität Pécs (Ungarn)	Leitung, R. Werner	25.10.2021
Educational Testing Service (ETS) / TOEFL	Senior Research Director and Principal Research Scientist at ETS, J. Norris	schriftliche Anfrage und Auskunft
<b>ExpertInnen für Remote-Testen, Proctoring, Rechts- und Datenschutz, Datensicherheit etc.</b>		
MeazureLearning / ProctorU	Proctoring-Anbieter, S. Clark	schriftliche Anfrage und Auskunft
Cambridge Assessment English	Data Protection Manager, I. Hanahoe (Rechtsanwalt)	04.11.2021
SEM	Datenschutzbeauftragter, P. Dubois	26.11.2021 / schriftlich
Lernende, zukünftige Testteilnehmende		
Lernende	Zielsprache Deutsch, Zürich	25.11.2021

## 4. Übersicht über Aufgabentypen, Aufgabengestaltung und Nutzerinterface bei digitalen Sprachtests

Im folgenden Kapitel werden Optionen für die Darstellung verschiedener Aufgabentypen und für die Gestaltung des Nutzerinterface dargestellt. Abschnitt 4.1 gibt einen generellen Überblick über typische digitale Aufgabenformate. Abschnitte 4.2 bis 4.5 spezifizieren die Möglichkeiten für die verschiedenen Fertigkeiten (Hören, Lesen, Schreiben, Sprechen). Abschnitt 4.6 zeigt exemplarisch auf, wie die aktuellen fide-Aufgabentypen digital umgesetzt werden können. Abschnitte 4.7 bis 4.10 beschreiben weitere grundlegende Aspekte im Hinblick auf Darstellungsoptionen und Nutzerinterface, Adaptivität, die Wahl der Durchführungsmedien, Sicherheitsaspekte und Support.

Im Prinzip ist technisch alles möglich, was Aufgabentypen, Aufgabengestaltung und Nutzerinterface angeht. Es ist dementsprechend möglich, für alle fide-Aufgabentypen eine gute digitale Entsprechung zu finden (u.a. als Multiple-Choice-, Richtig-Falsch-, Drag&Drop-, Lückentext oder freie Schreibaufgaben, s. 4.7).

### 4.1 Typische digitale Aufgabenformate: Multiple-Choice- und Zuordnungsaufgaben

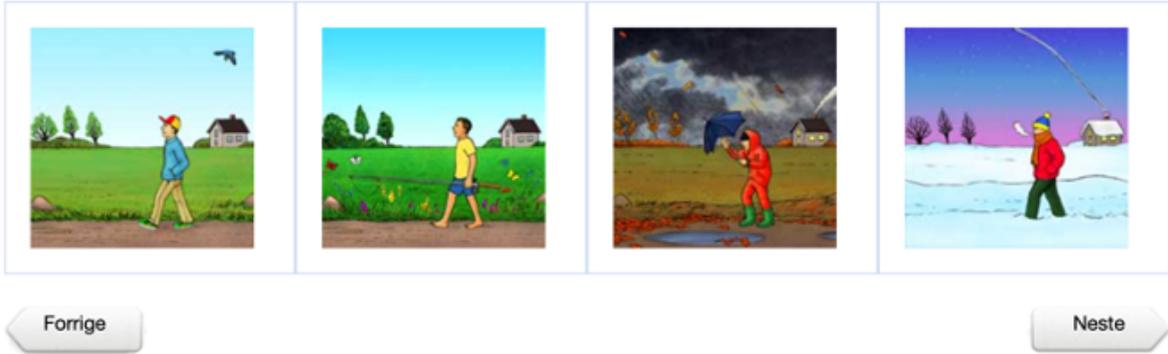
In erster Linie werden in digitalen Sprachtests typischerweise Multiple-Choice-, Auswahl- und Zuordnungs- (Drag&Drop-) Aufgaben gestellt, die automatisiert (und unkontrovers) korrigierbar und auch auf niedrigen Niveaustufen einsetzbar (Goethe-Institut) sind. Diese Aufgabentypen können unkompliziert eingesetzt werden, u.a. für Hör- und Leseverstehen, sie ermöglichen u.a. folgende, in den folgenden Beispielen 1 bis 14 illustrierten Arbeitsformate.

#### Beispiele 1-2:

Auswahl des passenden Bildes, passenden Gegenstandes, passenden Datums etc. aus einer Serie vorgegebener Stimuli zu einer mündlichen oder schriftlichen Äusserung / zu einem kurzen Text

*Les og klikk på riktig bilde.*

1. Det er varmt og fint, og sola skinner. Knut går med t-skjorte og korte bukser. Han skal på fisketur.



Forrige Neste

Bsp. 1, Norwegian Language Test

Practice Test A - Listening Part 1: Listening to Problem Solving Time remaining: 22 seconds

**1 Listen to the question. You will hear it only once.**

Playing...

0.00 / 0.05

This playbar will not appear in the official test.

**Question 1 of 8**

**1 Choose the best answer to each question.**

- 
- 
- 
- 

Bsp. 2, CELPIP

**Beispiel 3:**

Auswahl der passenden Person, der passenden Aktion, des passenden Gegenstandes etc. aus einem komplexen Bild entsprechend einer vorgegebenen mündlichen oder schriftlichen Äusserung

*Les og klikk på riktig sted i bildet.*

**3. Victor skal ut og reise. Nå er bussen sen, og han er redd for å komme for sent til flyet. Klikk på Victor.**



The illustration shows a bus stop with a shelter. A man in a suit (Victor) stands with a suitcase and a dog. Other people are engaged in various activities: a woman with a red skirt, a man with a guitar, a woman with a child, and others walking or sitting. A signpost is visible in the background.

Bsp. 3, Norwegian Language Test

**Beispiele 4-5:**

Auswahl der passenden Aussage / Antwort aus einer Serie von mehreren Aussagen zu einem vorgegebenen Bild oder Kurztext

COMPRÉHENSION ORALE 89

Écoutez les 4 propositions. Choisissez celle qui correspond à l'image et cliquez sur le bouton correspondant.



Cliquez sur votre réponse

- A Réponse A
- B Réponse B
- C Réponse C
- D Réponse D

Bsp. 4, TCF

Practice Test A - Listening Practice Task Time remaining: 11 seconds

**i Listen to a short statement. You will hear it only once.**

Playing...

0:00 / 0:06

This playbar will not appear in the official test.

**i Choose the sentence that is closest in meaning to the statement.**

- I want to see the concert again.
- I didn't attend the concert.
- The concert was cancelled.
- The concert wasn't good.

Bsp. 5, CELPIP

**Beispiele 6-10:**

Beantworten kurzer Fragen zu einem Text/Bild; verschiedene Antwortoptionen sind vorgegeben, die richtige Antwort muss ausgewählt werden.

**Les og klikk på riktig svar.**

Emilia er ute for å kjøpe noen nye klær. Det hun har mest bruk for, er en jakke. Hun finner ikke jakke, men hun prøver en fin genser. Hun kjøper den, selv om den er litt dyr. Det er også flotte kjoler i butikken, men det får bli en annen gang.

**3. Hva kjøper Emilia?**

- Kåpe
- Kjole
- Jakke
- Genser

← Forrige
→ Neste

Bsp. 6, Norwegian Language Test

Les og klikk på riktig tekst.

4. Hvilken tekst passer til bildet?



Arild bor litt utenfor Oslo sammen med kona si. Han har to barn, men begge er store og har flyttet hjemmefra. Arild synes ofte at det er veldig stille hjemme nå.

Birger jobber som sjåfør. Han kjører varer mellom Sør- og Nord-Norge. Hvis han får barn, vil han ikke reise så mye.

Fillip jobber i et firma som selger datautstyr. Han er veldig glad i jobben sin. Men han er også glad for at han kan være hjemme med sønnen sin, som snart blir ett år.

Gunnar er glad i barn og holder på å utdanne seg til lærer. Han har lyst til å jobbe med de minste barna på skolen.

Forrige

Neste

Bsp. 7, Norwegian Language Test

Logg ut

Les og klikk på riktig tekstmelding.

2. Hvem har ikke penger?

<p>Hei, tante!</p> <p>Nå er jeg i Stavanger. Kan du ringe når det passer for deg å møte meg? Jeg blir til søndag kveld.</p> <p>Hilsen Oleg</p>	<p>Hei, Sara</p> <p>Jeg kan dessverre ikke gå på kino i dag likevel. Sønnen min er blitt syk. Kanskje vi kan gå neste uke?</p> <p>Hilsen Anna</p>	<p>Hei, Bea</p> <p>Kan du låne meg 200 kroner? Du får dem igjen når jeg får lønn. Hvis det passer, kan vi treffes utenfor Bystasjonen klokken 12.</p> <p>Hilsen Ray</p>	<p>Hei, Conny</p> <p>Kan du møte meg på stasjonen klokken 16? Toget går 16.30, men vi må kjøpe billetter først. Det koster 340 kroner hver vei.</p> <p>Hilsen Anny</p>
--	---	---	--

Bsp. 8, Norwegian Language Test

Les og klikk på riktig svar. Denne teksten har flere spørsmål.

**Åpningstider tannlegevakten:**

Mandag til onsdag: 17-20  
Torsdag: 20-24  
Lørdag: 10-20  
Søndag: 17-24

5. Når åpner tannlegevakten på tirsdager?

Klokkå 10

Klokkå 12

Klokkå 17

Klokkå 20

Forrige
Neste

Bsp. 9, Norwegian Language Test

**VOUS ÊTES NOTRE MEILLEURE CLIENTE !**

**RIVIERIS VOUS OFFRE :**  
[Un séjour de thalasso à Mont-Tremblant, au Québec](#)

**2 nuits pour 2 personnes**

- bassins d'eau extérieurs et baignade en rivière
- bains scandinaves, hammam, eaux thermales



**+ 1 soin au choix**  
Périodes de validité :  
de mars à juin, hors vacances scolaires et jours fériés.

**Dans ce document**

- on décrit un hôtel.
- on décrit une région.
- on annonce un cadeau.
- on annonce une promotion.

Bsp. 10, E-TEF

**Beispiel 11:**

Auswahl einer angemessenen Zusammenfassung zu einem vorgegebenen mündlichen oder schriftlichen Text aus einer vorgegebenen Liste

<p><b>Au royaume du foot, les mafias sont reines</b></p> <p>Vous aimez le sport, ses beaux gestes, ses exploits, ses efforts hors du commun ? Oubliez !</p> <p>Vous aimez l'argent facile, les martingales improbables et les arnaques de haut vol ? Misez sur le sport de haut niveau !</p> <p>Ainsi pourrait se résumer la chronique de ces dernières années. On croyait avoir touché le fond avec les aveux de Lance Armstrong. Après des années de soupçons et de dénégations courroucées, le « champion » cycliste reconnaissait s'être dopé pendant ses années de gloire, avec un cocktail EPO-transfusions sanguines-</p>	<p><b>Selon lui, les scandales se multiplient car</b></p> <p><input type="radio"/> la pression médiatique et policière s'intensifie.</p> <p><input type="radio"/> certaines pratiques dévoyées ne sont plus taboues.</p> <p><input type="radio"/> des gouvernements sont régulièrement mis en cause.</p> <p><input type="radio"/> certains sportifs protègent moins bien leur vie privée.</p>
--	---

Bsp. 11, E-TEF

**Beispiel 12:**

Drag&Drop-Aufgabe: Textteile in die richtige Reihenfolge bringen

*Avsnittene nedenfor skal bli en sammenhengende tekst. Klikk og dra avsnittene slik at de kommer i riktig rekkefølge.*

1.

**Til alle foresatte ved Storhaugen skole!**

Som dere vet, finnes det allerede en abonnementsordning for skolemilk. Nå har Storhaugen skole vært så heldig å få bli med i en tilsvarende prøveordning med skolefrukt. Alle skoleelever vil fra neste uke få tilbud om å få to frukter per dag. Bestilling og betaling skjer på internett.

Se bruksanvisningen nedenfor:

Når du har gjort dette, blir du logget inn i selve bestillingssystemet. Her kan du velge hvilken frukt barnet ditt skal få, og om han/hun skal få én eller to frukter per dag.

Dersom du har valgt nettbetaling, blir du sendt videre til en sikker betalingside, hvor du betaler for frukten på samme måte som du betaler regninger i nettbanken.

Når du har valgt produktene du vil bestille, kan du velge om du vil få tilsendt en faktura eller om du vil betale direkte på nettet.

Lenken tar deg til en nettside hvor du kan registrere navnet til barnet ditt og hvilken klasse han/hun går i. Her må du opprette et brukernavn og et passord for å komme videre i prosessen.

Når du skal registrere deg, må du gå inn på [www.proveordningfrukt.no](http://www.proveordningfrukt.no) og oppgi epostadressen din. Når du er blitt registrert i systemet, får du en epost med en lenke som du må klikke på for å melde barnet ditt på fruktordningen.

Forrige Neste

Bsp. 12, Norwegian Language Test

**Beispiel 13:**

Komplexe Auswahl-Aufgabe,; es muss entschieden werden, ob eine Aussage zu einem von zwei präsentierten Lesetexten passt (d.h. zu Text 1, zu Text 2, zu beiden Texten, zu keinem Text)

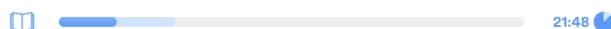


Read both passages and the 6 statements. Decide if each statement is supported by:

- Passage A
- Passage B
- Both Passages

Passage A

The Notting Hill Carnival is one of the world's largest street festivals. It is held in the Notting Hill area of West London. Over one million people attend the carnival, which takes place over three days every August. The event started in 1966 as a celebration of Caribbean culture, with people from the West Indies, who had immigrated to London, organizing the festival. Nowadays, the event is very multi-cultural and people from all the different communities of London come together at the festival. However, the music, costumes, floats and food still have a strong Caribbean influence.



Text anzeigen

The passage says that people have to pay for parts of the carnival.

Passage A
  Passage B
  Both Passages

The passage describes what people see and do at the carnival.

Passage A
  Passage B
  Both Passages

Bsp. 13, EFSET

**Beispiel 14:**

Auswahl des passenden Wortes oder der passenden Satzergänzung an einer Stelle in einem Lückentext

Practice Test A - Listening Part 4: Listening to a News Item Time remaining: 3 minutes

**1** Choose the best way to complete each statement from the drop-down menu ( ▾ ).

1. The news item is about

2. Before the surgery, Dianne was  an Australian doctor.  
 a medical breakthrough.  
 an eye disease.  
 a computer program.

3. With the prototype, Dianne was able to

4. Dianne's prototype will

5. The two other patients

Bsp. 14, CELPIP

Weitere Auswahl- und Zuordnungsformate wie z.B. die Zuordnung von Satzhälften, die Zuordnung von Wörtern aus einer Liste zu den richtigen Positionen in einem Lückentext etc. sind denkbar (siehe Kapitel 4.7), müssten in ihrer Eignung auf sehr niedrigen Niveaustufen und für schulungsgewohnte Testteilnehmende allerdings geprüft werden (ebenso wie die Optionen 11-14).

## 4.2 Hör- und Lesekompetenzen prüfen

Hör- und Leseverstehen können mit dieser Art von Multiple-Choice- und Zuordnungsaufgabentypen ggf. auch adaptiv evaluiert werden (vgl. Norwegian Language Test, SIELE, EFSET, Duolingo). Adaptiv ist auch der Goethe-Test PRO, der computerbasiert Hör- und Lesekompetenzen am Arbeitsplatz evaluiert; allerdings setzt das nach Aussage der Expertinnen des Goethe-Instituts einen sehr grossen Aufgabenpool voraus.

In den gängigen Tests werden Hören und Lesen häufig verschränkt, da z.B. die Fragen zu den Hörverstehenstexten oft schriftlich gestellt werden (wie in Bsp. 14 oben, Bsp. 15 unten).

**Sélectionnez la bonne réponse dans chaque menu déroulant.**



1- Ce message est diffusé dans

2- Dans ce message, on annonce

Bsp. 15, E-TEF

Es gibt aber durchaus auch sehr attraktiv umsetzbare, weitgehend schriftfreie, z.B. bild- und audiobasierte Aufgabentypen auch für niedrige Kompetenzlevel. Im *Norwegian Language Test* z.B. werden die Inputs zwar schriftlich präsentiert, aber gleichzeitig auch vorgelesen, d.h. im Audioformat präsentiert, sobald eine neue Seite aufgerufen wird; im Beispiel 3 (s.o.) muss z.B. auf Basis einer Personenbeschreibung die passende Person im komplexen Bild identifiziert werden.

Aufgaben (inkl. der Audios für das Vorlesen der Aufgabenstellung bzw. die Hörverstehensitems) können entweder direkt nach einer gewissen Zeit (z.B. direkt bei Seitenwechsel, 3 Sekunden nach Seitenwechsel etc.) eingespielt werden (und ggf. über einen Button durch die Testteilnehmenden beschleunigt angefordert) oder aber mit freiem Zeitmanagement durch die Testteilnehmenden selbst aufgerufen (über die entsprechenden Play-Buttons); Letzteres stellt deutlich höhere Ansprüche an die Navigationskompetenzen und an das Selbstmanagement der Testteilnehmenden, auch wenn jeweils eine Angabe (im Sinne eines Countdowns) zur noch verfügbaren Zeit pro Aufgabe bzw. für den Gesamttest gemacht wird (dies ist allerdings auch bei der Papier-Version der Fall).

Gerade Remote-Sprachtests bestehen aber häufig eben nur aus Lese- und Hörverstehensaufgaben, z.B. EFSET, Duolingo, TEF express. Diese sind dann direkt automatisiert auswertbar (und dadurch potenziell auch relativ eindeutig adaptiv zu gestalten), messen aber nicht alle Fertigkeiten und gerade nicht die im bisherigen fide-Testkonstrukt zentralen pragmatischen und interaktionalen Kompetenzen.

Beim Leseverstehen ist eine zentrale Herausforderung die Darstellung längerer Texte bzw. die Darstellung der Texte zusammen mit den entsprechenden Aufgaben, auch in Abhängigkeit von der Bildschirmgröße. Je nach Anzahl der Informationseinheiten und Darstellungsform – Text und Aufgaben untereinander bzw. nebeneinander – entsteht potenziell zusätzlicher kognitiver Aufwand durch das Zurechtfinden in einem komplexen Nutzerinterface. Bei kleinem Bildschirm, d.h. gerade die Aufgaben beispielsweise nicht auf PC, Laptop oder einem grösseren Tablet, sondern auf einem Smartphone aufgerufen werden, entsteht zusätzlicher Gedächtnisaufwand durch die Notwendigkeit, mehr oder weniger häufig und/oder mehr oder weniger weit nach unten/oben bzw. seitlich zu scrollen und sich entsprechend gelesene Textinhalte und/oder Fragestellungen zu merken (siehe Bsp. 16-17); dies betonen auch die Expertinnen des Goethe-Instituts. Die oben angeführten digitalen Tests für Spanisch, Norwegisch und Französisch sind bspw. nur auf Computern abrufbar.

Eine kontinuierliche Optimierung des Nutzerinterface wäre auf der Basis von Usability-Studien (z. B. Eye-Tracking) möglich und machbar (s.u.), dieser Aufwand sollte mitgedacht und geplant werden.

Instructions ▲
Exit

PRACTICE QUESTIONS

For these questions, choose the correct answer.

A visit to Japan

I'm an 18-year-old girl from Sweden, called Susanna. Last summer, I visited my Japanese friend Nori in Tokyo. When I got there we drove through the city. I knew there'd be lots of people and it was great to see the bright lights, but the best thing was listening to the Japanese language for the first time.

I had a surprise when I saw Nori's house. I already knew that Japanese people often sit on the floor, and that some people sleep in their living room, but I didn't know that you can move walls to make rooms bigger or smaller.

One day we went to some clothes shops, but I didn't get anything. I wanted to visit a famous comic shop, but it was closed, so we went to a strange café where you can drink tea and play with lots of cats!

Usually we ate in Nori's home and once I tried sashimi, the famous fish dish. I learned that Japanese food is very good and keeps people healthy. Once we went to a restaurant and Nori said 'Try the noodles'. They were much better than the ones that people in my country eat.

We went to an onsen, too, a large bath which is outside. The water comes from the ground and it's hot. It was too hot for

**21** How did Susanna feel when she arrived in Tokyo?

pleased to see places she knew

worried about how busy it was

interested in what she could hear

**22** Susanna was surprised that in some Japanese houses

**23** When Nori and Susanna went shopping, they

**24** What does Susanna say about the food she ate in Japan?

**25** What does 'ones' in the fourth paragraph mean?

**26** What did Susanna like about the onsen?

Reading ▼

9 ●
10 ●
11 ●
12 ●
13 ●
14 ●
15 ●
16 ●
17 ●
18 ●
19 ●
20 ●
21 ●
22 ●
23 ●
24 ●
25 ●
26 ●

< >

Bsp. 16, Cambridge Linguaskills Anywhere

Practice Test A - Reading Part 1: Reading Correspondence

Time remaining: 10 minutes

NEXT

Just when we thought things would slow down, Marco broke the news that he had accepted a job offer in Tokyo. He leaves at the end of the month! He was invited to work at a top engineering firm that specializes in rebuilding cities after large disasters. Needless to say Marco is very excited about it. The firm has been really active in the reconstruction efforts following the 2011 earthquake in Japan, and it's a great first step into his career, not to mention the opportunity to experience a new culture, and learn a new language. He's a little apprehensive about communicating in Japanese but the firm has a translator and a tutor to help him.

Marco is over the moon, but Jack and I are having a tougher time with it. We think he's so young to be so far away, but we know it's for the best. Just the other day Jack and I were talking about that trip all of us took to Disneyland when the kids were little, do you remember that? It's been 20 years! Time flies, doesn't it? Well, get prepared, Cindy is next!

In any event, we are planning a farewell party for friends next weekend, and Marco cannot imagine not having you, Jason and Cindy there. The party will be at his apartment. He really does not want to leave without saying goodbye to you all. We hope you can make it, after all we haven't seen you since last Christmas. Let me know so I can get the rooms ready.

We'll be in touch,

Love,  
Maria

response by filling in the blanks. Select the best choice for each blank from the drop-down menu ( ▾ ).

Hi Maria,

This is such wonderful news! Count us in, we would hate to miss

7. ▾ We'll leave 8. ▾ at 6am Saturday

afternoon. That way if you need

▾ you'll have some extra

ions.

duation gift. Initially we

couver apartment, but I guess

ve any suggestions? Does he

have everything he needs for the 11. ▾ Winter

clothes, perhaps?

Let me know, and see you on Saturday!

Love,  
Mea

the leaving party.

the graduation.

the trip.

the flight.

Bsp. 17, CELPIP

### 4.3 Schreibkompetenzen prüfen

Schreibaufgaben können auf verschiedenen Schwierigkeitsniveaus gestellt werden, u.a. als

- Lückentexte mit Freitexteingabe, auch als Beschriftung von Skizzen
- Satzvervollständigungsaufgaben
- Vervollständigen von Tabellen
- Schreiben von Teilsätzen oder Sätzen, kürzeren oder längeren Texten

Diese Aufgaben können durchaus attraktiv und lebensnah gestaltet werden, z.B. als das Verfassen einer WhatsApp o.ä. (Bsp. 18), einer SMS oder einer E-Mail (Bsp. 19), mit dem entsprechenden Layout und entsprechenden Wiedererkennungseffekt. Auch Hilfen können in diesen Formaten leicht eingebaut werden, z.B. eine Wortzählungsfunktion.

1.

Du har bursdag. Skriv en invitasjon til vennene dine.



Bsp. 18, Norwegian Language Test

**Skriv en e-post til busselskapet. Skriv 80 ord eller mer.**

3.

Busselskapet vil legge ned noen av bussrutene der du bor. Dette skaper problemer for deg og andre i nærmiljøet.

Skriv en e-post til busselskapet og forklar hvorfor det er viktig at bussene går like ofte som før.

0 ord

Bsp. 19, Norwegian Language Test

Durchaus authentisch kann hier auch dialogisches / interaktives Schreiben inszeniert werden (z.B. in Form einer Mail-Antwort auf eine erhaltene Mail, Bsp. 20):

Tarea 1 Tarea 2

Tiempo prueba: 00:49:52

Usted ha recibido un mensaje de una amiga y va a responderlo.

En el mensaje a su amiga usted debe:

- saludarla;
- decir cómo es su casa en Granada, qué medios de transporte hay y cómo es la zona donde está;
- decir qué cosas le gustan y qué cosas no le gustan de su nueva vida;
- contar qué cosas ha hecho en el extranjero durante estos años;
- despedirse.

**CORREO ELECTRÓNICO**

CORREO | CALENDARIO | CONTACTOS | AGENDA

Hola, ¿qué tal?

Me ha dicho Miguel que estás viviendo en Granada porque has encontrado un trabajo aquí. ¡Qué bien! ¿En qué zona estás? ¿Cómo es tu casa? ¿Te gusta la ciudad? Dice también que has estado en el extranjero. ¿Qué has hecho allí? ¿Cómo te ha ido?

Escribeme y cuéntamelo todo.

Nuria

Número de palabras recomendado: entre 100 y 150

Escriba aquí

Teclado español

Número de palabras escritas: 0

Bsp. 20, SIELE

Auch längere Schreibaufgaben können z.B. durch Bildstimuli eingeleitet und gesteuert werden, siehe Bsp. 21:

Beskriv bildet. Skriv cirka 50–80 ord.

2.

Hva ser du på bildet? Hva gjør personene?



0 ord

Forrige

Neste

Bsp. 21, Norwegian Language Test

Für die Testteilnehmenden hat das Schreiben im Rahmen einer digitalen Prüfung nach Aussage von TestanbieterInnen potenzielle Vorteile, u.a. in Bezug auf

- schnelle Korrektur von Fehlern im eigenen Text
- Zählen der bereits geschriebenen (bzw. noch erlaubten/noch geforderten) Wörter
- einfache Ergänzung von Textteilen
- Überarbeitung von Textteilen,

allerdings nur unter der Bedingung, dass die Tastatur alle benötigten Schrift- und Sonderzeichen enthält (ggf. problematisch im Remote-Fall), die Testteilnehmenden ans Tastaturschreiben gewöhnt sind, und zwar mit dem Tastaturtyp, der für die Prüfung verwendet wird (QWERTZ, QWERTY, AZERTY).

Problematisch hingegen ist die potenzielle Bevorteilung / Benachteiligung individueller Lernender oder Lernendengruppen, je nach ihrer Gewohnheit ans Tastaturschreiben; manche Testteilnehmenden seien ggf. Schreiben auf Papier gar nicht mehr gewohnt, andere seien Schreiben an der Tastatur nicht gewohnt (Casanova, ALTE).

Die für diese Machbarkeitsstudie interviewten Lernenden an der AÖZ schätzen, dass sie für einen Text, den sie in 5 Minuten handschriftlich verfassen würden, ca. 20 Minuten an einer PC- oder Laptop-Tastatur, ca. 10 Minuten an einer Smartphone-/Touchscreen-Tastatur brauchen würden. Im Falle einer Digitalisierung müsste als ihr Zeitbudget für die Schreibaufgabe(n) bzw. für den Anteil Lesen & Schreiben deutlich erhöht werden. Empirisch erhoben werden müsste durch Begleitforschung der tatsächliche durchschnittliche Mehraufwand, so dass das Zeitbudget insgesamt fair angepasst werden kann.

Denkbar sind hier ebenfalls drei Alternativszenarien zum klassischen Schreiben an einer PC- oder Laptop-Tastatur:

1. Grundsätzlich schreiben die Testteilnehmenden am Computer (Tastatur), aber man kann individuell einen Antrag stellen, dass man seine Texte auf Papier schreiben will; der handschriftliche Text wird dann eingescannt und hochgeladen (erprobt und funktional für E-TEF und TCF, nach Aussage von Casanova, ALTE).
2. Alternativ zum Schreiben auf einer klassischen Tastatur (PC, Laptop) wird das Tippen/Schreiben auf einer Touchscreen-Tastatur (Tablet, Smartphone) angeboten, woran gerade schulungsgewohnte Lernende nach eigener Aussage medial am ehesten gewöhnt sind, da sie z.B. WhatsApp für das Versenden kurzer Nachrichten auf dem Smartphone benutzen. Technisch umsetzbar wäre dies (vgl. Goethe digital) durch eine generelle Durchführung auf Tablets mit der Option der Nutzung einer externen Tastatur. Für tastaturgewohnte Testteilnehmende könnte eine externe Tastatur mit Maus zur Verfügung gestellt werden.
3. Es wird mit einem digitalen Stift (Smartpen, iPad-Stift o.ä.) direkt auf den (herunterklappbaren) Bildschirm eines entsprechenden Gerätes geschrieben; diese Option kann gerade für kürzere Texte in Betracht gezogen werden, es müsste allerdings wohl vorher eine Übungsgelegenheit geben.

#### 4.4 Automatische Auswertungsmöglichkeiten

Im Unterschied zu den Aufgabentypen Multiple-Choice, Richtig/Falsch und Drag&Drop ist aktuell nicht klar, ob bzw. inwiefern auch Schreibaufgaben automatisch ausgewertet werden können; von einigen Experten wird das im Hinblick auf das Deutsche und vor allem niedrigere Niveaustufen bezweifelt (u.a. TestDaF); in dem Sinne wären die Konsequenzen eines Angebots des Schreibens auf Papier (s.o.) für den Korrekturaufwand nicht gravierend. Bei ECL wird auf diese Art bereits erfolgreich bei den papierbasierten Tests gearbeitet, d.h. die Texte werden eingescannt und am Computer ausgewertet, und zwar von geschulten Bewerberinnen und nicht automatisch.

In Bezug auf erwartbare Entwicklungen werden entsprechend Fortschritte bei der (teil-)automatischen Korrektur von Texten genannt, neben Fortschritten beim adaptiven digitalen Testen im Allgemeinen und schliesslich auch Fortschritte bei der Entwicklung interaktionaler Evaluierungsformen wie z.B. durch Chat-Plattformen (interaktives Schreiben) (Casanova, ALTE-Experte).

#### 4.5 Sprechkompetenzen prüfen

In Bezug auf die Komponente Sprechen gibt es zwei grundlegende Möglichkeiten:

- monologisch: Die Testteilnehmenden sehen, lesen oder hören einen Stimulus (ein Bild, einen Text, eine Situationsbeschreibung, eine Problembeschreibung etc.) und reagieren mündlich monologisch darauf (d.h. beantworten Fragen, geben einen Ratschlag, hinterlassen eine Nachricht, diskutieren ein Problem etc.) und nehmen sich beim Sprechen auf (s. 4.5.1).
- dialogisch: Die Testteilnehmenden führen ein Gespräch mit 1-2 Prüfenden, und zwar in einem Video-Call, der in Gesprächsform mündliche Interaktion evaluiert (s. 4.5.2).

##### 4.5.1 Monologisches Sprechen (mit eigener Aufnahme)

Die Option, das eigene (monologische) Sprechen aufzunehmen (Abb. 22-24), entspricht nicht dem aktuellen fide-Testkonstrukt der Evaluierung interaktionaler Sprechkompetenzen.

Es ermöglicht allerdings weiterführend gedacht eine integrierte Leistungsevaluierung, z.B. eine Kombination aus Leseverstehen, Hörverstehen und Sprechen. In Beispielaufgabe 25 (SIELE) wird zum Beispiel eine Anzeige gelesen (Leseverstehen), dazu wird eine Frage gehört (Hörverstehen) und eine Antwort gegeben (Sprechen). Das integrierte Testen der unterschiedlichen Fertigkeiten, wie es im fide-Test zum Teil umgesetzt wird, berücksichtigt die neuesten Entwicklungen im Lehren, Lernen und Beurteilen. Im Begleitband des Gemeinsamen Referenzrahmens für Sprachen wurden nämlich die einzelnen Fertigkeiten durch die sogenannten kommunikativen Sprachaktivitäten und Strategien (Rezeption, Produktion, Interaktion, Mediation) ersetzt. Die neue Systematik soll besser den natürlichen lebensweltlichen Sprachgebrauch erfassen (vgl. Europarat, 2020, S. 38f.).

Auch hier ergibt sich allerdings wieder die Herausforderung der Darstellbarkeit der verschiedenen Komponenten auf kleineren Bildschirmen (Tablet, Smartphone). Zudem stellt sich auch hier die Frage der Anforderungen an die digitalen Grundkompetenzen und der Verunsicherung durch die Aufgabenstellung; gerade für schulungsgewohnte Lernende mit wenig digitaler Erfahrung kann die hohe Anforderung in Bezug auf das technische Verständnis (Wo muss ich in welcher Reihenfolge klicken, um was zu tun? Wann beginne ich zu sprechen? Wie kann ich meine Produktion ggf. korrigieren? etc.) ggf. benachteiligend wirken; daher müssten in solchen komplexen Aufgaben die Arbeitsschritte besonders klar nachvollziehbar dargestellt und vorher ggf. geübt worden sein. Begleitforschung müsste zeigen, dass die Anforderungen von verschiedenen Zielgruppen ähnlich gut bewältigt werden.

In Bezug auf die Täuschungssicherheit wird in Remote-Tests häufig gefordert, dass nicht nur eine Audio-, sondern eine Video-Datei aufgenommen und hochgeladen/ingereicht wird, damit zu sehen ist, wer spricht (Abb. 22). Die kann ggf. einen zusätzlichen Stress- bzw. Ablenkungsfaktor darstellen.

2:52

Abb. 22, Duolingo

Bsp. 23, SIELE

Tarea 1 Tarea 2 Tarea 3 Tarea 4 Tarea 5

Opción 2

Describe la fotografía. Estos son los aspectos que tiene que comentar:

- Cuántas personas hay y qué relación existe entre ellas.
- Cómo son esas personas y qué ropa llevan.
- Qué están haciendo esas personas y cómo cree que se sienten.
- En qué lugar se encuentran y cómo es.
- Qué objetos hay, cómo son y dónde están.

Tiene dos minutos para preparar la tarea. Si lo desea, puede tomar notas.

Después, grabe su respuesta. La grabación empieza automáticamente. Tiene de uno a dos minutos para realizar la tarea.



Grabación automática en: 02:25



Escriba aquí sus notas. NO EVALUABLE

Bsp. 24, SIELE

Tarea 1 Tarea 2 Tarea 3 Tarea 4 Tarea 5

Opción 1.

Lea el texto. A continuación, pulse para escuchar cada pregunta (🔊) y grabe su respuesta (🎤). Tiene usted 1 minuto para responder cada pregunta.

**EL RECICLAJE**

A nivel mundial, el manejo de la basura es uno de los factores que más afecta al medio ambiente. El incremento de basureros altera el ecosistema y pone en peligro la flora y la fauna del planeta. En una visión ecológica del mundo, el reciclaje es la medida principal para conseguir la disminución de residuos sólidos.

Pregunta 1



El audio empezará en: 00:54



Grabar

Pregunta 2



Reproducir



Grabar

Bsp. 25, SIELE

#### 4.5.2 Dialogisches Sprechen (im Video-Call)

Die Option "Video-Call", d.h. die Inszenierung dialogischen Sprechens im digitalen Medium mit einem/r konkreten menschlichen GesprächspartnerIn, hat folgende Vorteile:

- Der Video-Call ist lebensnah, die Teilnehmenden kennen Video-Calls privat und aus dem Online-Unterricht (über Skype, Zoom, Facebook, WhatsApp); sich selbst dabei aufzunehmen, wie man z.B. 2 Minuten lang über ein Thema spricht, ist hingegen im Sinne des szenariobasierten Ansatzes zumindest auf niedrigen Niveaus schwierig zu rechtfertigen.
- Die Bedienungsanforderungen sind im Vergleich zur monologischen Option mit Aufnahme der eigenen Produktion (4.6.1) niedrig; die Testteilnehmenden, auch die schulungsgewohnten, kennen grundsätzlich das Prinzip Video-Call, müssen lediglich in den entsprechenden "Raum" eingeladen werden oder einem Link folgen.
- Die monologische Option, sich selbst beim Sprechen aufzunehmen - und zwar eigentlich als Video, nicht nur Audio, im Sinne der Täuschungssicherheit (s. Abb. 24 oben) - dürfte (technisch und sprachlich) für manche Testteilnehmenden einen zusätzlichen Stressfaktor darstellen.

Die Option "Video-Call" ruft allerdings auch einige Einschränkungen auf, und zwar digital vor Ort wie digital remote:

- Es werden genauso viele Prüfende gebraucht wie in der Papier-Version, d.h. es ergibt sich an der Stelle keine Kostenersparnis.
- Die Terminflexibilität ist stark eingeschränkt; der Teil "Sprechen" müsste ggf. in einen Extra-Termin ausgelagert werden; es müsste sichergestellt sein, dass das Testkonstrukt dadurch nicht (oder nur in explizit gewollter Form) verändert wird.

Teilweise (z.B. von einer Vertreterin einer Bürgergemeinde) wird gefordert, es müsse sichergestellt sein, dass digital / analog derselbe Aufbau der Komponente "Sprechen" besteht, u.a. genau die gleichen Fragen gestellt werden. Hier stellt sich also die Frage der Gleichwertigkeit verschiedener, potenziell parallel angebotener Prüfungsformate; Begleitforschung bestehender digitaler Test zeigt, dass Video-Calls zumindest für Testteilnehmende mit relativ hohen digitalen und schriftsprachlichen Kompetenzen vergleichbare Anforderungen stellt und ähnlich gut akzeptiert wird wie eine dialogische Prüfung vor Ort.

Über sehr gute Erfahrungen mit einer digitalen Sprechen-Komponente während der Pandemie berichtet u.a. ein ALTE-Testexperte. Gearbeitet wurde in diesem Fall mit Zoom-Calls mit Aufnahme des Gesprächs für Dokumentationszwecke; entsprechend arbeite das Team aktuell an einer integrierten Lösung für Sprechen via Video-Call innerhalb der Aufgabenplattform, nicht über eine Zusatzapplikation.

Allerdings muss für die Option Video-Call die Proctoring-Frage geklärt werden; u.W. gibt es für Video-Calls aktuell keine integrierten Proctoring-Angebote (u.a. schliesst der Grossanbieter MeazureLearning / ProctorU das aus, s.u.). Skype beispielsweise läuft nicht auf Safe-Exam-Browsern. Das Gespräch müsste also im Falle der Remote-Version mit menschlichem Proctoring durchgeführt werden, daher auch im Sinne der Rekursicherheit wie aktuell entweder von zwei Prüfenden geführt oder durch eine Aufnahme dokumentiert für potenzielle Überprüfungen post hoc. Längerfristig müsste wohl ggf. wie oben angesprochen an der Entwicklung einer integrierten Lösung gearbeitet werden.

An dieser Stelle sei kurz gesagt, dass z.B. die ExpertInnen von TestDaF nicht zu einer Gesamteinbettung der gesamten Prüfung in ein Zoom-Gespräch raten, was insgesamt einer Face-to-Face-Durchführung am nächsten käme; dies würde bedeuten, dass der/die Prüfende im Rahmen eines Video-Calls nicht nur den Anteil „Sprechen“ evaluieren würden, sondern die Testteilnehmenden gesamthaft in einem Video-Call durch die ganze Prüfung begleiten, wie sie es aktuell im Face-to-face-Gespräch tun. Dabei würden sie nach und nach auch die einzelnen Testaufgaben für die Testteilnehmenden starten; dabei könnte nach Ansicht der ExpertInnen potenziell zu Missverständnissen kommen (Sind die Testteilnehmenden bereit? Ging es zu schnell?) und entsprechende Beschwerden geben; damit könne ggf. ein Rekursgrund vorliegen, weil dann eine andere Person die Prüfung steuere (s. auch ECL).

#### 4.6 Digitale Umsetzungsmöglichkeiten der Aufgabentypen im aktuellen fide-Test

Die folgende Tabelle 5 zeigt auf, wie die aktuellen Aufgabentypen des fide-Tests in einer digitalen Testversion potenziell umsetzbar wären. Teilweise sind mehrere Optionen denkbar.

**Tab. 5: Digitale Umsetzungsmöglichkeiten für die Aufgabentypen im aktuellen fide-Test**

Aufgabe	Digitale Umsetzungsoptionen
<b>Hören &amp; Sprechen</b>	
A1.1: Bild / Situation beschreiben	monologisch: Aufnahme dialogisch: Video-Call
A1.2: Dialog rekonstruieren in dieser Situation	Video-Call
A2.1: Sprechen zu Bild - Erzählen Sie, was Sie sehen	monologisch: Aufnahme dialogisch: Video-Call
A2.2: Telefonanruf spielen, basierend auf der bildlich dargestellten Situation, z.B. Terminabsprache Arzt	monologisch: Aufnahme dialogisch: Video-Call
A2.3: freies Gespräch zu der bildlich dargestellten Situation	Video-Call
B1.1: Sprechen über eine Situation, mehrere Fragen zum Thema	monologisch: Aufnahme dialogisch: Video-Call
<b>Lesen &amp; Schreiben</b>	
A1/A2.1: 1: Kurztext/Annonce lesen und kurze W-Fragen beantworten	Kurztext-Antwort (Bsp. 26) oder Multiple-Choice (Bsp. 6, 9-10)
A1/A2.2: : Anmeldeformular ausfüllen	Lückentext (Bsp. 27, 28)
A1/A2.3: etwas längeren Text/Info-Brief lesen und W-Fragen beantworten	Kurztext-Antwort (Bsp. 26) oder Multiple-Choice (Bsp. 6, 9-10)
A1/A2.4: komplexeres Anmeldeformular ausfüllen, inkl. Kommentare / Rückfragen	Lückentext (Bsp. 27, 28) und/oder Kurztext-Antwort (Bsp. 26)
A1/A2.5 bzw. A2/B1.1: zwei Annoncen lesen, vergleichen, Informationen in eine Tabelle eintragen	Lückentext (Bsp. 27) und/oder Kurztext-Antwort (Bsp. 26)
A1/A2.6 bzw. A2/B1.2: Mail schreiben - Termin verschieben (mind. 30 Wörter)	Freitext (Bsp. 19-21)
A2/B1.3: : Informationsbrief lesen und ankreuzen, welche Aussagen richtig sind	Multiple-Choice (Bsp. 6, 9-10)
A2/B1.4: Brief schreiben - z.B. Wohnungskündigung (mind. 30 Wörter)	Freitext (Bsp. 19-21)
A2/B1.5: 5 Kurztexte lesen und zu jedem eine MC-Frage beantworten	Multiple-Choice (Bsp. 6, 9-10)
A2/B1.6: Mail schreiben - z.B. Geschenkidee, Buchempfehlung	Freitext (Bsp. 19-21)

Wofür steht die Abkürzung fide?

Antwort:

Bsp. 26, Kurztext-Antwort, Moodle

Grundlegende Aufgabe von Lehrenden ist es, den  und die Interaktionskontexte so aufzuwerten, dass die  der jeweils relevanten Zielstruktur erhöht wird; und damit auch die Wahrscheinlichkeit, dass die LernerInnen die Zielstrukturen  und dabei den für eine Erweiterung oder Restrukturierung ihrer in Entwicklung befindlichen Lernergrammatik notwendigen  generieren.

Bsp. 27, Lückentext, Moodle

Um welches Prinzip handelt es sich?

«Das Prinzip der  peilt das Erreichen von (selbst)gesetzter Ziele an.»

Personalisierung

Lernerorientierung

Kompetenzorientierung

Kontextualisierung

Interaktionsorientierung

Erfolgsorientierung

Aufgabenorientierung

Handlungsorientierung

Bsp. 28, Lückentext mit Drag&Drop-Auswahl, Moodle

## 4.7 Arbeitsanweisungen und Nutzerinterface

Generell sind in digitalen Sprachtests viele Erklärungen (in Bezug u.a. auf technische Funktionen oder die Navigation), Arbeitsanweisungen und Aufgabentypen schriftbasiert, es müssen z.B. Aufgabenbeschreibungen gelesen werden, die im aktuellen fide-Durchführungsformat mündlich kommuniziert werden, auch Fragen zum Hörverstehen werden häufig schriftlich dargeboten. Dieses Problem kann allerdings entschärft werden, wenn z.B. die Arbeitsanweisungen und/oder Items beim Hörverstehen auch mündlich angeboten werden (s.o., z.B. *Norwegian Language Test*). Auch die Einbindung von Videos (als Erklärvideos zur Navigation, zu einzelnen Aufgabentypen oder als Stimuli) kann ggf. Teile des Problems lösen; die an der AOZ befragten Lernenden präferieren jedoch deutlich Erklärungen bzw. Support und Prüfungsbegleitung durch eine reale Aufsichtsperson, Erklärvideos halten sie nicht für ausreichend hilfreich.

Dabei muss bedacht werden, dass in vielen gängigen digitalen Tests das Nutzerinterface komplex ist und Aufgabenscreens teilweise sehr komplex, informationsdicht bis überladen sind, s. Bsp. 29 unten; es finden sich typischerweise auf engem Raum u.a.

- Aufgabenanweisungen
- Aufgabeninhalte wie Lesetexte
- die eigentlichen Aufgabenitems mit Auswahl-, Zuordnungs- oder Schreibmöglichkeiten
- Hilfsfunktionen
- Navigationsfunktionen
- Navigationsübersichten
- ggf. auch eine Angabe zur verbleibenden Zeit pro Aufgabe / insgesamt

Instructions ▲ PRACTICE QUESTIONS Exit

Click on each gap then choose the correct answer.

**Dancing**

I really enjoy dancing **11** [ because ✕ ] it's fun and it's also very good for your body. I go to two lessons **12** [ ✕ ] week, on Tuesdays and Saturdays. **13** [ ] favourite kind of dancing is called salsa - it's great. Salsa **14** [ ] from countries in Latin America, for example Cuba and Puerto Rico. **15** [ ] is also a kind of music called salsa. I love that too!

every                      one                      all

Reading ▼ 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 < >

🔊

Bsp. 29, Cambridge Linguaskills Anywhere

Der Navigationsprozess kann eine Herausforderung gerade für schrift- / schul- bzw. digital ungewohnte Testteilnehmende sein, wie auch die Expertinnen des Goethe-Instituts betonen; es sei essenziell für die Testteilnehmenden, auch im digitalen Format die Übersicht zu wahren: Wie weit bin ich? Was kommt noch? Wo bin ich gerade? Habe ich versehentlich etwas übersprungen? Wie komme ich wohin?

Die digitale Darstellung könne gerade für schrift- und digital ungewohnte Testteilnehmende potenziell sehr verunsichernd sein. Das Benutzerinterface muss daher für einen digitalen fide-Test für schrift- und digital ungewohnte Testteilnehmende und für niedrige Kompetenzstufen in jedem Fall medienunabhängig im Vergleich zu gängigen Placement- und Zertifizierungstests wohl deutlich entlastet werden (s. Goethe-Institut). Dies betrifft potenziell die Anzahl verschiedener Aufgabentypen sowie die Anzahl von Inhalts-, Hilfe- und Navigationsfeldern, die gleichzeitig / nacheinander auf dem Bildschirm angezeigt werden (Aufgabeninhalt, z.B. Lesetext oder Play-Button für den Hörtest, Aufgabenstellung, konkrete Frage-Items, ggf. Feld zum Schreiben; Seite zurück, Seite nach vorn, Übersicht Gesamtnavigation, Gesamthilfefunktion, aufgabenspezifische Hilfen etc.).

Bei der Orientierung helfen können Einblendungen von navigationsbezogenen Informationen (u.a. zur verbleibenden Prüfungszeit, zur Übersicht über die bisher gelösten und noch offenen Items und Aufgaben), wenn diese durchdacht und im Kontext der inhaltlichen Komponenten übersichtlich angeordnet sind. Überlegt werden könnte nach Ansicht der Expertinnen des Goethe-Instituts ggf. auch, ob eine digitale Prüfungsversion wirklich parallel zur Papierversion sein müsse und freies Hin- und Herspringen ermöglichen; ggf. sei es einfacher im Digitalen, wenn die Führung strikter sei, d.h. die Testteilnehmenden keine/weniger Wahl hätten, was sie zuerst und in welcher Reihenfolge bearbeiten wollen. Eine striktere Führung nehme den Teilnehmenden Verantwortung ab (Zeitmanagement) und zeige die Struktur und Navigationsoptionen besser auf – allerdings müsste in diesem Fall sichergestellt werden, dass die striktere Führung keine Benachteiligung in anderer Hinsicht darstelle.

In Ausnahmefällen könnte z.B. für motorisch beeinträchtigte oder sehr schulungsgewohnte Testteilnehmende auf den niedrigsten Kompetenzstufen eine weitere Möglichkeit in Betracht gezogen werden, falls die Papier-Version des Tests nicht mehr existieren sollte, nämlich eine Durchführung vor Ort, bei der nicht der/die TestnehmerIn das Programm bedient, sondern ein/e ausgebildete/r PrüferIn (Vorschlag von Malone, ACTFL, s. aber die Kritik von TestDaF).

In jedem Fall müsste genau geprüft und gut pilotiert werden (in sog. Usability-Studien), welche Art der Darstellung insgesamt und für einzelne Aufgabentypen sinnvoll, überschaubar und zielführend ist. Es müsste geprüft werden, wie das Nutzerinterface im Gesamtlayout schlank, fokussiert und verständlich angelegt werden kann und welche Darstellungsoptionen sich für spezifische Aufgabentypen auch für die schrift- und schulungsgewohnten Testteilnehmenden als verständlich und händelbar erweisen (ohne gleichzeitig für schul- und schriftgewohnte Testteilnehmende unattraktiv zu sein).

Die befragte Expertin von ACTFL berichtet in dieser Hinsicht von extensiven und aufwändigen Pilotierungen mit Hilfe von aufgezeichneten Probedurchgängen mit Testteilnehmenden, die entweder während des Probedurchgangs direkt live ihre Probleme kommentieren (sog. Think-aloud-Protokoll) oder aber im Rückblick auf Basis der Aufzeichnung kommentieren, welches Problem aufgetreten ist (sog. stimulated recall); an den jeweiligen Stellen wurde dann nachgefragt, was den Testteilnehmenden geholfen hätte (z.B. ein grosser gelber Pfeil für "Klicken Sie hier"); diese Optionen wurden dann jeweils in neuen Probedurchläufen evaluiert. Auch das Goethe-Institut berichtet von extensiven Pilotierungen bezüglich der Usability verschiedener Medien (Laptops, Tablets, Tablets mit externer Tastatur/Maus) und Darstellungsoptionen für die Navigationsführung; z.B. sei der ursprünglich eingesetzte grosse Pfeil als Startbutton für die Testteilnehmenden in dieser Funktion nicht verständlich gewesen.

Es muss in jedem Fall sichergestellt werden, dass die Teilnehmenden bei Bedarf eine Einführung und eine bzw. mehrere Übungsgelegenheiten bekommen, um z.B. die Navigationsoptionen und Anforderungen einzelner Aufgabentypen (Anklicken / Auswählen, Verschieben, Schreiben, Korrigieren etc.) kennen zu lernen und sich mit ihnen ausreichend vertraut zu machen.

#### 4.8 Adaptivität und Testkonstrukt

Im Zusammenhang mit der Aufgabengestaltung müsste frühzeitig die Frage der Adaptivität geklärt werden. Bisher beruht die Adaptivität des fide-Tests auf menschlicher Beurteilung (und zwar jeweils separat für die Komponente Sprechen / Hören und für die Komponente Lesen / Schreiben).

Eine adaptive Gestaltung beim Lese- und Hörverstehen wäre in einer digitalen Prüfung, bei entsprechender Ausgestaltung der Item-Datenbank und der Entscheidungswege für den Aufruf von Aufgaben-Items spezifischer Schwierigkeitsgrade, wohl durchaus möglich (s. EFSET, Duolingo, Norwegian Language Test, Goethe PRO). Allerdings stünde diese Herangehensweise im Kontrast zum bisherigen Testkonstrukt, das umgekehrt das Niveau für das Hörverstehen auf Basis der Kompetenzen im Sprechen festlegt.

Für Sprechen (in Option 1, Aufnahme) und Schreiben wäre hingegen in der digitalen Prüfung wohl eine Anmeldung für ein spezifisches Niveau nötig.

Im Sinne des bestehenden Testkonstrukts digital ortsabhängig gedacht wären folgende Abläufe möglich:

1. persönliche Anmeldung mit Identitätskarte im Testzentrum
2. Video-Call für Sprechen (oder aber ein persönliches Gespräch, wenn nicht gerade Lockdown ist), mit Aushandlung des Zielniveaus, mit einer/einem Prüfenden und einer/einem Beisitzenden für die Qualitätssicherung und ggf. Aufnahme für die Dokumentation und Rekursicherheit; damit entfielen ein grosser Teil der Datenschutzproblematik (s.u.)
3. weiterer Prüfungsverlauf (Hörverstehen, Lesen und Schreiben) in derselben Abfolge wie bisher, ausgelöst durch die/den Prüfende/n, ggf. mit Wahlmöglichkeit zwischen PC oder Tablet und ggf. Antragsmöglichkeit für Schreiben auf Papier bzw. Schreiben mit digitalem Stift auf Bildschirm

Im Sinne des bestehenden Testkonstrukts digital ortsunabhängig gedacht wären folgende Abläufe möglich:

1. Anmeldung persönlich im Testzentrum oder online mit Remote-Identitätsprüfung (mit den entsprechenden Konsequenzen für den Datenschutz, s.u.)
2. Video-Call für Sprechen, ggf. mit Aushandlung des Zielniveaus
3. der/die Testteilnehmende bekommt einen Code, der ins entsprechende Niveau führt (für Hörverstehen und Lesen/Schreiben; Alternativantragsmöglichkeit für Schreiben auf Papier ist in diesem Fall unklar, das müsste für die Täuschungssicherheit zumindest mit einem weiteren Gerät gefilmt werden oder mit menschlichem Proctoring stattfinden)

Wenn der Prüfungsteil Sprechen nicht als Video-Call durchgeführt wird, ist im Sinne des aktuellen Testkonstrukts eine menschliche Bewertung der Sprechleistung nach Prüfung notwendig, mit anschliessender Niveauempfehlung (und ggf. entsprechender Zusendung des Codes für A1/A2 oder A2/B1). Das braucht Zeit und zwei getrennte Termine (von Anfang getrennt gebucht, das erhöht den administrativen Aufwand). Im Falle einer Durchführung digital vor Ort wäre das für die Testteilnehmenden aufwändig (2x Anreise etc.; das ist allerdings ggf. aktuell auch der Fall).

Möglich wäre aber ggf. ein Video-Call im Remote-Modus, für die weitere Niveauempfehlung, und dann das Absolvieren von Hörverstehen, Leseverstehen und Schreiben vor Ort. Das führt jedoch wiederum zu den Herausforderungen, die mit Proctoring und Datenschutz für sensible, biometrische Daten für die Identitätsprüfung verbunden sind, wenn die Anmeldung online möglich sein soll; für Überlegungen zu einer datenschutztechnisch weniger kritischen, aber personell sehr aufwändigen Variante siehe unten Kapitel 5.4.3 zum Proctoring.

Andere Optionen würden eine Änderung des Testkonstrukts notwendig machen, was ggf. eine längere inhaltliche Neukonzeptualisierungsphase und entsprechende Absprachennotwendigkeiten zur Folge hätte, z.B.

- die Teilnehmenden melden sich (z.B. entsprechend ihres aktuellen Kursniveaus, der Empfehlung ihres/ihrer KursleiterIn, eines groben Online-Einstufungstests in eigener Verantwortung oder auch auf Basis eines kurzen Einstufungsgesprächs) selbst für ein bestimmtes Niveau an;
- die Einstufung erfolgt über die geschlossenen Aufgaben Hörverstehen / Leseverstehen; das ist problematisch in dem Sinne, dass in der Kernzielgruppe der schulungsgewohnten Testteilnehmenden die mündlichen Kompetenzen typischerweise höher entwickelt sind; ggf. könnte man nur das Hörverstehen als Einstufungsoption in Betracht ziehen (auch das würde aber das Testkonstrukt und die Reihenfolge der geprüften Fertigkeiten verändern).

#### 4.9 Medien, Infrastruktur und Sicherheit

Was die mediale Umsetzung betrifft, so stellen sich zwei grundlegende Herausforderungen, je nach Durchführungsart (digital vor Ort, digital remote):

Im Falle der **Durchführung vor Ort** ist zu überlegen, ob es gerade für schrift- und schulungsgewohnte Lernende, die digital tendenziell am ehesten an den Umgang mit dem Smartphone gewöhnt sind, eine Wahlmöglichkeit geben soll zwischen der Bearbeitung des digitalen Tests am PC / Laptop (mit Maus und Tastatur) und am Tablet (Touchscreen); um die Kosten für die Infrastruktur zu reduzieren und eine Doppelausstattung zu vermeiden, könnte allgemein eine Durchführung am Tablet angestrebt werden, mit optionaler externer Tastatur und Maus für die Testteilnehmenden, die dies wünschen. In jedem Fall müssen alle Testzentren mit genau demselben Material (Arbeitsplätze, Headsets etc.) ausgestattet sein, damit die Durchführungsbedingungen für alle Testteilnehmenden gleich sind.

Dies stellt auch angesichts der abzusehenden Anforderungen an Wartung und Aktualisierung der Hardware ggf. einen erheblichen finanziellen Aufwand dar. Dass dieser Aufwand von den Testzentren selbst geleistet werden kann, wird von einigen Testanbietern deutlich bezweifelt; sie fordern eine Finanzierung und Ausstattung, inkl. Wartung und Aktualisierung der Geräte, durch das SEM (s.u).

Im Falle einer **Remote-Durchführung** liegt die Verantwortung für die technische Ausstattung bei den Testteilnehmenden. Es dürfte der Fall sein, dass zwar nicht alle Lernenden die notwendigen Geräte zuhause

haben, aber sich ggf. eines ausleihen können (s.u.). Es könnte in Betracht gezogen werden, den Testteilnehmenden (zumindest im Inland) eine Geräteausleihe zu ermöglichen (z.B. über die lokale Gemeindeverwaltung).

Es ist allerdings nicht davon auszugehen, dass die Lernenden über einen bestimmten Gerätetyp bzw. über das aktuellste Betriebssystem verfügen. Es muss bei der Entwicklung der Testumgebung sowie der Proctoring-Plattform daher darauf geachtet werden, dass die Applikationen gleichermaßen einfach bedienbar und gleichermaßen zuverlässig auf MacOS, Windows- und ggf. auch Linux-Geräten und/oder im Falle einer Tablet- (oder Smartphone-)Durchführung zudem auf den entsprechenden mobilen Betriebssystemen (Android, iOS etc.) laufen, und zwar zumindest für die letzten 5-6 Generationen Betriebssystem (und perspektivisch auf kommenden Generationen der verschiedenen Betriebssysteme, s. z.B. die erst kürzlich aufgetretenen Schwierigkeiten bei der Umstellung von den 32- auf die 64-Bit-Versionen), s. auch Malone/ACTFL.

Schriftungsgewohnte Lernende präferieren deutlich einen Test auf dem Smartphone (oder höchstens auf dem Tablet), nicht auf Laptop / PC, wenn es keine Papierversion mehr geben sollte (s.u.). Technisch ist die Durchführung auf dem Smartphone nach Aussage verschiedener TestexpertInnen (u.a. Goethe-Institut, TestDaF) allerdings in mehrfacher Hinsicht problematisch:

- Die reduzierte Bildschirmgrösse selbst grösserer Smartphones macht es schwierig, längere Texte bzw. mehrere Inhalte gleichzeitig darzustellen, z.B. im Kontext Leseverstehen (s.o.).
- Dies darf ebenfalls als problematisch für das Schreiben längerer Texte gelten (hier: Tastatur- und Bildschirmgrösse), was bis Niveau B1 allerdings nicht gravierend sein dürfte.
- Schliesslich sind Smartphones deutlich anfälliger für Aufgabendiebstahl bzw. allgemein für Proctoring-Lücken; so kann deutlich weniger gut überprüft werden, wer das Gerät gerade bedient, auch Safe-Exam-Browser laufen typischerweise nicht auf Smartphones, man kann also z.B. ungehindert Screenshots machen (und damit Aufgaben veröffentlichen, so dass die Item-Datenbank deutlich häufiger aktualisiert werden müsste), aber auch anderweitig schummeln.

Allerdings gibt M. Malone (ACTFL) zu bedenken, die Durchführbarkeit auf dem Smartphone sei eine potenzielle zeitnahe Entwicklung, zu der sich wohl auch ein digitaler fide-Test positionieren müsste:

“I think the next thing is going to be whether or not we allow learners to take tests on their phone, and I think that is a really hard thing”.

Allgemein muss die Item-Datenbank regelmässig aktualisiert werden, um Overexposure zu vermeiden<sup>1</sup>; TestexpertInnen gehen davon aus, dass ein Item höchstens 2000 bis 5000 mal ausgespielt werden sollte, um die Wahrscheinlichkeit gering zu halten, dass es bereits öffentlich zirkuliert. Das Update der Item-Datenbank ist ein nicht zu unterschätzender Aufwand (vgl. ECL), vor allem, wenn der Test adaptiv sein soll; allerdings ist es in einer digitalen Version möglich, neue Items für die Evaluierung (Gleichwertigkeit) zeit- und kostensparend im Rahmen der eigentlichen Testdurchführungen mitlaufen zu lassen (als nicht bewertete Sonderitems, wobei die Testteilnehmenden natürlich nicht wissen, welche Items bewertet und welche gerade noch erprobt werden, u.a. TCF).

Es muss jedenfalls so gut wie möglich sichergestellt werden, dass Aufgaben nicht “entwendet” werden, z.B. durch Screenshots. Dies ist bei der Durchführung digital vor Ort potenziell kein Problem, durch Proctoring auch weitgehend vermeidbar bei Durchführung remote am PC oder Laptop, ggf. auch am Tablet; Je weniger gut die Sicherheit der Items garantiert werden kann, desto weniger häufig kann jedes Item ausgespielt werden, d.h. desto häufiger muss die Itemdatenbank aktualisiert werden.

---

<sup>1</sup> Der fide-Test in seiner aktuellen Form ist nicht item-, sondern aufgaben- bzw. modulbasiert, d.h. mehrere Aufgaben gehören jeweils zu einem Modul, das wiederum einem Szenario bzw. Handlungsfeld zugeordnet ist; daher gibt es für die aktuelle Form des fide-Tests keine Item-Datenbank im traditionellen Sinne, sondern eher eine Aufgaben- bzw. Modulsammlung. Seit seinem Bestehen wurde der fide-Test jährlich einer Überarbeitung unterzogen, letztmalig 2020. Daher gibt es im Moment keine substanzielle Sammlung von Aufgaben, die den aktuellen Testspezifikationen entsprechen und nach einer gewissen „Ruhezeit“ wieder aktiviert werden könnten. 2021 wurden mehrere Module entwickelt, die miteinander zu Sets kombiniert werden können; dabei werden normalerweise neue mit älteren Modulen gemischt, keine ganz neuen Sets für die Prüfungsdurchführungen eingesetzt. Ab Frühjahr 2022 sollen nach Möglichkeit regelmässig neue Module entwickelt werden, die mit älteren Modulen zu (neuen) Prüfungssets kombiniert werden.

## 4.10 Support

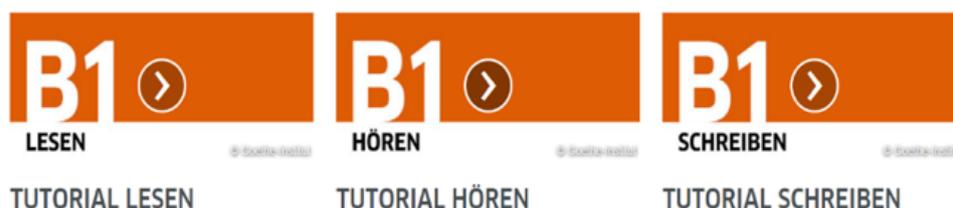
### 4.10.1 Support vor und während der Prüfung

Allgemein müssen (technischer) Support und (Navigations-)Hilfsoptionen für digital und/oder schrift- und schulungsgewohnte Testteilnehmende spezifisch gedacht werden, und zwar im Hinblick auf Formate, Auffindbarkeit und Verständnissicherung. Auch hier ist technisch alles möglich, z.B. die Einbindung von (weitgehend schriftfreien und in einfacher Sprache formulierten) Erklärvideos bzw. Tutorials. Diese könnten vor dem Test als Informationsressource zur Verfügung stehen (z.B. Erklärung der Gesamtnavigation und/oder einzelner Aufgabentypen), wie dies z.B. beim Goethe-Test (Bsp. 31), CELPIP (Bsp. 32) und dem Norwegian Language Test der Fall ist.

Es wäre z.B. möglich, eine Anmeldung bzw. ein Login schon 24 Stunden vor Teststart zu erlauben, zum Ausprobieren und Kennenlernen der Testumgebung. So können potenzielle Fragen zur Navigation und Testumgebung vorab gestellt und beantwortet werden. Möglich wäre auch (gerade im Fall einer Remote-Durchführung), dass der Einstieg in den Test über ein Erklärvideo gestaltet ist, das die einzelnen Testteile und spezifischen Aufgabentypen kurz vorstellt (z.B. Goethe-Test). Spezifische Erklärvideos könnten auch während des Tests bei Problemen mit einzelnen Aufgabenformaten zur Verfügung stehen (Wo muss ich klicken? In welcher Reihenfolge? Wie komme ich zum letzten Item zurück, zum nächsten Item? etc.).

Die Formulierung der entsprechenden Erklärungen muss selbstverständlich dem Kompetenzniveau der Testteilnehmenden angepasst sein, für die Stufe A1/A2 ist besonders auf einfache, klare Sprache zu achten, ebenso auf eine reduzierte Informationsdichte; dies stelle eine grosse Herausforderung dar (s. Goethe-Institut); auch die Verfügbarkeit von Tutorials bzw. Erklärvideos in der Mittlersprache Englisch (und/oder in den jeweiligen Erstsprachen?) kann ggf. in Betracht gezogen, aber wohl nicht von allen Testteilnehmenden sinnvoll in Anspruch genommen werden.

## GOETHE-ZERTIFIKAT B1 DIGITAL



Bsp. 31, Goethe-Institut

## Left-CLICK on the mouse to select your answer.

Reading Part 1: Reading Correspondence Time remaining: 9 minutes 31 seconds **NEXT**

**1** Read the following message.

Hi Mea,

I'm sorry for taking so long to reply to your email. As you know, I've been very busy this summer with Marco's university graduation and my family's visit from Chile. The graduation ceremony was great! Too bad you guys couldn't make it, but we understand it's a bit of a drive from Calgary. You were truly missed. My Mum remembers you well from when we were kids. After the celebrations, we took the family sightseeing here in Vancouver. They loved it! They left yesterday for Victoria, and will catch a plane back home from there in three days.

Just when we thought things would slow down, Marco broke the news that he had accepted a job offer in Tokyo. He leaves at the end of the month! He was invited to work at a top engineering firm that specializes in rebuilding cities after large disasters. Needless to say Marco is very excited about it. The firm has been really active in the reconstruction efforts following the 2011 earthquake in Japan, and it's a great first step into his career, not to mention the opportunity to experience a new culture, and learn a new language. He's a little apprehensive about communicating in Japanese but the firm has a translator and a tutor to help him.

Marco is over the moon, but Jack and I are having a tougher time with it. We think he's so young to be so far away, but we know it's for the best. Just the other day Jack and I were talking about that trip all of us took to Disneyland when the kids were little, do you remember that? It's been 20 years! Time flies, doesn't it? Well, get

**1** Using the drop-down menu ( ▾ ), choose the best option according to the information given in the message.

- Maria's mother is now 
  - in Chile.
  - in Canada.
  - in Vancouver.
  - in Victoria.
- In a few weeks, Maria's
- Marco is feeling
- Marco's employers are
- Maria and Jack are worried about
- Mea and Maria

**1** Here is a response to the message. Complete the response by filling in the blanks. Select the best choice for each blank from the drop-down menu ( ▾ ).

Hi Maria,

This is such wonderful news! Count us in, we would hate to

### Bsp. 32, CELPIP

Daneben sollte gerade für die Zielgruppe der schulungsgewohnten Testteilnehmenden überlegt werden, in welcher Form Übungs- und Trainingsmöglichkeiten nicht nur in einer Online-Übungslandschaft, sondern ggf. auch in einer begleiteten Form mit Aufsicht und/oder Erklärung in den Testzentren zur Verfügung gestellt werden können. Bei einer Durchführung vor Ort ist es auch vorstellbar, dass eine kurze erklärende Einführung durch die Aufsichtsperson den Einstieg in den Test darstellt.

Bei allen Aufgaben sollte das Nutzerinterface klar darstellen, welche Arbeitsschritte in welcher Reihenfolge absolviert werden müssen (z.B. beim Hörverstehen: Anklicken des Play-Buttons, Auswahl der korrekten Antwortmöglichkeiten, Weiterklicken zum nächsten Item etc.) und welche Hilfe-Ressourcen zur Verfügung stehen.

In der digitalen Durchführung vor Ort kann bei Navigationsproblem im Zweifelsfall die Aufsichtsperson zugezogen werden (siehe aber TestDaF: Eingreifen nur bei Hardware-Problemen), für genuin technischen Support muss sehr kurzfristig ein Programmierer / IT-Experte zugezogen werden können. Zu klären ist,

- wie weit die Unterstützung bei Navigations- und technischen Problemen gehen darf (z.B. zeigen, erklären);
- wie durch Navigationsprobleme, technische Bugs, Internetverbindungsprobleme etc. verlorene Zeit genau gemessen und ggf. kompensiert werden kann (wird z.B. die Arbeitszeit angehalten und nach Fehlerbehebung der Countdown wieder eingesetzt? oder wird am Ende Zeit zugegeben? wie wird dieses Zeitvolumen genau berechnet?);
- was in diesem Kontext als Rekursgrund zugelassen ist und was nicht.

Im Falle einer Remote-Durchführung braucht es neben den oben genannten Hilfe- und Erklärressourcen eine Möglichkeit, entweder die Online-Aufsichtsperson zur Hilfe zu rufen (im Falle menschlichen Proctorings) oder aber einen spezifischen Technik-Support, über einen Support-Button oder eine zuverlässige und im Sinne der Kundenorientierung kostenlose Telefonhotline. Erfahrungen gibt es z.B. bei ECL mit Technik-Support in Break-out-Räumen in Zoom. Der Technik-Support müsste neben dem Live-Support

auch Zugriffsrechte auf zwischengespeicherte Versionen haben und ggf. versehentlich Gelöschtes wiederherstellen können.

Unklar ist u.E. aktuell, inwiefern

- sich Stress, der durch technische Probleme ausgelöst wird, auch auf Folgeaufgaben überträgt und die Leistungsfähigkeit der Testteilnehmenden über die konkret problematische Aufgabe hinaus einschränkt;
- eine Durchführung des Tests mit komplett eigenem Zeitmanagement ggf. noch mehr als bei der bisherigen Papier-Version dazu führt, dass einzelne Testteilnehmende sich bei einzelnen Aufgaben-Items zu lange aufhalten (weil die Übersichtlichkeit in der digitalen Version gegenüber der Papierversion eingeschränkt ist, s. Goethe-Expertinnen);
- eine laufende Countdown-Uhr die Testteilnehmenden eher zu gutem Zeitmanagement verhilft oder sie unter Stress setzt;
- umgekehrt vorgetaktetes Zeitmanagement, d.h. das automatische Aufrufen des jeweils nächsten Items nach einer bestimmten Zeit, ggf. zu Irritationen bei einzelnen Testteilnehmenden führt;
- eine stärkere Führung in der Durchführung digitaler Tests (z.B. keine Wahlfreiheit bei der Reihenfolge der Bearbeitung verschiedener Aufgaben) die digitale Navigation entlasten kann (ohne gleichzeitig die Chancengleichheit zu beeinträchtigen).

#### 4.10.2 Herausforderungen

Unerlässlich ist es nach Ansicht der verschiedenen Stakeholder im Hinblick auf die Entwicklung, Darstellung und Bewertung der Aufgaben eines digitalen fide-Tests und des entsprechenden Nutzerinterface, folgende Aspekte sicherzustellen:

- die sozio-kognitive Validität des neuen digitalen Tests bzw. seiner Versionen im Sinne des (ggf. angepassten) Testkonstrukts: Was genau misst der Test bzw. messen die einzelnen Aufgaben (und die einzelnen Items? Entspricht das dem grundsätzlichen Anspruch und Konstrukt des fide-Tests (z.B. auch in Bezug auf die Adaptivität)?
- die Gleichwertigkeit der parallel existierenden Testversionen (Papier, digital vor Ort, digital remote) durch entsprechende Begleitforschung; geklärt werden muss dazu vorab, was genau "Gleichwertigkeit" in diesem Kontext bedeuten soll, also
  - tatsächliche genaue Gleichheit der Aufgaben und Unterlagen in den parallel existierenden Testversionen (Papier, digital vor Ort, digital remote), wie gefordert z.B. von einer Vertreterin einer Bürgergemeinde)
  - oder Gleichwertigkeit der Anforderungen und gemessenen Kompetenzen in den parallel existierenden Testversionen (Papier, digital vor Ort, digital remote) im Sinne des Testkonstrukts (Messen die Tests genau dieselben Kompetenzen, auch wenn die Aufgaben nicht absolut identisch sind?)
- die Rekursicherheit, und zwar durch ein entsprechendes rechtssicheres Reglement und die Etablierung einer exhaustiven Liste der Probleme, die als nicht von den Teilnehmenden verschuldet gelten; dies beinhaltet die Klärung u.a. folgender Fragen:
  - Was wäre bei einem digitalen Test Grund für einen Rekurs?
  - Was passiert mit Verbindungsunterbrüchen etc. (s.o.), wie werden diese Schwierigkeiten nachgewiesen?
  - Was genau wären dann die Rechte der Testteilnehmenden (Testwiederholung, ...)?
- den Umfang des Einspruchsrechts bei automatisch korrigierten Texten (s.u., Datenschutzexperte Cambridge Assessment English)
- die Beratungssicherheit der Testteilnehmenden (z.B. die Möglichkeit, ein individuelles Anliegen oder Problem zu schildern, die aktuell als hoch eingeschätzt wird)
- die benachteiligungsfreie Durchführbarkeit für schul- und schriftungewohnte bzw. digital weniger kompetente Testteilnehmende
- die benachteiligungsfreie Durchführbarkeit für Personen mit Beeinträchtigungen wie z.B. Sehbehinderungen, Hörbehinderungen, motorische Behinderungen

## 5. Technische Tools für die digitale Durchführung und Durchführungssicherheit (Datensicherheit, Täuschungssicherheit, Datenschutz)

Je nach Durchführung digital vor Ort bzw. digital remote (ortsunabhängig) ergeben sich teilweise unterschiedliche Anforderungen hinsichtlich der Durchführungs- und Täuschungssicherheit.

Im Folgenden werden die Anforderungen und Optionen aufgeschlüsselt nach

1. Durchführungssicherheit (Termin- und Datensicherheit), siehe 5.1
2. Täuschungssicherheit (1): Identitätsprüfung, siehe 5.2
3. Täuschungssicherheit (2): Proctoring und Datenschutz, siehe 5.3

Allgemein und insgesamt ist eine zweifelsfreie Durchführungssicherheit sicherzustellen, insbesondere muss eindeutig ausgeschlossen sein,

- dass es Täuschungsmöglichkeiten gibt (dies auch in Bezug auf die politische Ausnutzung des Gedankens, der digitale Test wäre aufgrund von Täuschungsoptionen praktisch wertlos für die Einbürgerung, wie die Vertreterin einer Bürgergemeinde zu bedenken gibt);
- dass die Testteilnehmenden an der Hardware scheitern (z.B. an den Anforderungen, für die Remote-Durchführung über 2 reibungslos funktionierende Geräte zu verfügen);
- dass die Testteilnehmenden an der Software scheitern (z.B. an der Identitätsprüfung für die Remote-Version);
- dass sich durch die Digitalisierung starke Kostenerhöhungen für die Testteilnehmenden ergeben.

Zu klären ist diesbezüglich u.a. die rechtliche und finanzielle Verantwortung / Verantwortlichkeit, insbesondere

- im Falle der Durchführung digital vor Ort für:
  - Infrastruktur, d.h. Hardware (PCs/Tablets mit externer Tastatur, Headsets sowie Wartung, regelmässige Erneuerung und Support),
  - Internetverbindung und Verbindungssicherheit,
  - Server und Datenspeicherung
- im Falle der Remote-Durchführung für:
  - Server und Datenspeicherung,
  - Identitätsprüfung und Täuschungssicherheit (Proctoring)
  - Datenschutz, vor allem im Falle von Proctoring durch künstliche Intelligenz.

### 5.1 Durchführungssicherheit (Termin- und Datensicherheit)

Terminsicherheit, d.h. eine Prüfungsdurchführung auch bei Lockdown und ähnlichen Situationen ebenso wie klare Kontaktreduktionen erlaubt nur eine Remote-Durchführung. Allerdings stellt sich gerade für High-stakes-Tests in zentraler Weise die Herausforderung der Prüfungsüberwachung und Täuschungssicherheit (s.u.).

In Bezug auf die Durchführungssicherheit darf auch die Datensicherheit nicht vergessen werden, also der Schutz vor Datenverlusten. Für die Durchführung vor Ort müssen die Testzentren die Zuverlässigkeit der Internetverbindung garantieren; bei schlechter Verbindung kann es sein, dass ggf. beim Abruf aus der Aufgaben-Datenbank Items übersprungen werden (Hinweis von Casanova, ALTE-Experte).

Im Falle der Remote-Durchführung liegt die Verantwortung für die Verbindungszuverlässigkeit bei den Testteilnehmenden, es muss daher ein klares Protokoll für den Umgang mit Verbindungsabbrüchen geben (z.B. TestDaF: Re-Login möglich mit Fortfahren an der Stelle, wo aufgehört wurde) und für die potenziellen Ansprüche auf Zeitkompensation bei Internetverbindungs-, Hardware- oder Software-Problemen. Bei einer Remote-Variante ist es gerade nicht möglich, wie z.B. beim TCF, eine Sondervariante anzubieten, wo alles vorab heruntergeladen werden kann, da dies eine lokale Speicherung voraussetzt und die Aufgaben dabei potenziell veröffentlicht werden.

Zu prüfen ist in diesem Zusammenhang allgemein, prioritär aber für die Remote-Variante, ob die Speicherung direkt bei jeder Eingabe auf dem Server erfolgen soll oder erst lokal, dann bei Abgabe der Prüfung eine Überspielung auf den Server erfolgt. Bei ECL z.B. findet eine automatische Speicherung alle 10 Sekunden statt; auf Wunsch des/der Testteilnehmenden, z.B. im Falle versehentlichen Löschens, kann die Aufsichtsperson die letzte Version wiederherstellen. Auch ein weiterer Experte plädiert für eine rein webbasierte Speicherungslösung ohne lokale Datenkopien, so dass keine unerlaubten Zugriffe über lokale Geräte möglich sind. Allerdings ruft diese Option u.a. die Frage nach der Stabilität der Internetverbindung auf. Allgemein ist zu prüfen, in welcher Weise ein Back-up-Server bzw. eine Back-up-Version Datenverluste verhindern helfen kann; für das Goethe-Institut erwies sich das offenbar als grosse Herausforderung.

## 5.2 Täuschungssicherheit (1): Identitätsprüfung

Im Falle einer digitalen Durchführung vor Ort kann die Identitätsprüfung bei der Prüfungsanmeldung und beim Prüfungstermin persönlich durchgeführt werden. Dies ermöglicht eine Prüfungsdurchführung ohne Speicherung sensibler, biometrischer Daten z.B. (Foto, ID-Karte, Gesichtserkennung).

Im Falle einer Remote-Durchführung muss die Identitätsprüfung bei Prüfungsanmeldung und beim Login zum Prüfungstermin online durchgeführt werden. Dazu wird typischerweise verlangt, dass die Testteilnehmenden ein Foto von sich und ein Foto von ihrer Identitätskarte hochladen; diese beiden Fotos müssen typischerweise nach dem Login in der eigentlichen Testumgebung live vor der Kamera gemacht und dann hochgeladen werden. Dazu müssen die Testteilnehmenden fähig sein, sich selbständig einzuloggen, mit Hilfe der Kamerafunktion des PC, Laptops oder Tablets ein Foto von sich aufzunehmen (wobei das Gesicht in einem Rahmen auf dem Bildschirm platziert werden muss), ein Foto ihrer Identitätskarte aufzunehmen (die neben dem Gesicht in den vorgesehenen Rahmen gehalten muss), beide Fotos zu hinterlegen und anschliessend ein Formular mit grundlegenden Angaben zu ihrer Person und zur Einwilligung in die Datenspeicherung und -verwendung (s.u.) auszufüllen, bevor sie den Test beginnen können.

Es kann erwogen werden, ein vorgezogenes Login zu ermöglichen, damit die Testteilnehmenden vorab die notwendigen Applikationen herunterladen, sich unter weniger Stress mit den Prozeduren für das Login und mit dem Aufbau der Testumgebung vertraut machen können (vgl. ECL: Login 24 Stunden vor der Prüfung möglich, inkl. vorgezogene Identitätsprüfung).

## 5.3 Täuschungssicherheit (2): Proctoring und Datenschutz

Im Falle einer Durchführung vor Ort kann die Prüfungsaufsicht wie bisher durch Aufsichtspersonen vor Ort durchgeführt werden. Es ist allerdings damit zu rechnen (s. ECL), dass für die Aufsicht bei einer digitalen Prüfung tendenziell mehr Aufsichtspersonen nötig sind und diese eine entsprechende technische Ausbildung / Schulung brauchen (z.B. schnell müssen entscheiden können, ob ein Systemfehler oder ein Navigationsproblem bei der/dem Testteilnehmenden vorliegt). Zusätzlich muss jederzeit ein spezifischer technischer Support (durch eine/n TechnikerIn bzw. InformatikerIn) online oder vor Ort abrufbereit sein (z.B. bei Verbindungsabbrüchen, Hardware- oder Softwareproblemen).

Im Falle einer Remote-Durchführung müssen Identitätsprüfung und Prüfungsaufsicht online durchgeführt werden (sog. Proctoring); dazu gibt es grundsätzlich zwei Möglichkeiten, nämlich

- menschliches Proctoring (durch eine Online-Aufsichtsperson)
- maschinelles Proctoring (durch künstliche Intelligenz)

Proctoring setzt grundsätzlich voraus, dass die Testteilnehmenden über zwei einwandfrei funktionierende Geräte (inkl. Kameras) und ein Headset (Hören, Sprechen) verfügen. Auf einem Laptop wird z.B. die eigentliche Prüfung durchgeführt, während die interne Kamera den/die Testteilnehmende/n filmt. Gleichzeitig filmt ein zweites Gerät den Testteilnehmenden und das erste Gerät aus einem anderen Winkel (seitlich,

von hinten), so dass ausgeschlossen ist, dass sich weitere Personen oder Hilfsmittel im Raum befinden; das zweite Gerät muss dazu typischerweise in eine Video-Call-Funktion eingeloggt werden.

Normalerweise darf die/der Testteilnehmende während der Remote-Prüfung nicht aufstehen, nicht den Blick vom Bildschirm abwenden, es darf niemand den Raum betreten, es darf nichts auf dem Tisch liegen, die Ohren müssen sichtbar sein etc. Es wäre in diesem Zusammenhang sicherzustellen, dass allen Testteilnehmenden ein ruhiger Ort zur Verfügung steht, an dem sie ungestört die Prüfung ablegen können, so dass keine Missverständnisse bei Störungen u.a. durch Kinder entstehen. Auch eine stabile Internetverbindung muss gesichert sein, denn bricht die Internetverbindung auf beiden Geräten gleichzeitig für mehr als 2 Minuten ab, wird die Prüfung abgebrochen (ECL).

Bei **maschinell**em Proctoring wird der gesamte Prüfungsverlauf (Video) aufgezeichnet. Es werden auffällige Verhaltensweisen (z.B. vom Bildschirm abgewandter Blick) bereits während der Prüfung "geflaggt" und müssen nachträglich manuell gesichtet werden; alternativ kann die Prüfung bei auffälligem Verhalten auch direkt abgebrochen werden, diese Option dürfte allerdings rekursiv problematisch sein. Bei **menschlichem** Proctoring kann ggf. auf eine Speicherung der Prüfungsdurchführungsvideos verzichtet werden; da eine Aufsichtsperson aber üblicherweise mehrere Prüflinge beaufsichtigt, wird aber typischerweise der Prüfungsverlauf für eine spätere Überprüfung auf auffälliges Verhalten hin doch aufgezeichnet. Im Falle einer High-Stakes-Prüfung müsste geklärt werden, ob eine menschliche 1-zu-1-Aufsicht notwendig ist bzw. was die maximale Anzahl der von einer Aufsichtsperson überwachten Testteilnehmenden wäre, so dass die Beaufsichtigung weiterer Testteilnehmender auch gewährleistet ist, während die Aufsichtsperson punktuell mit Support beschäftigt sein sollte.

Die notwendigen Applikationen / Programme müssen vor Prüfungsbeginn heruntergeladen und installiert werden. Sie übernehmen, ähnlich wie Safe-Exam-Browser, zusätzlich zur Film- und Aufzeichnungsfunktion weitgehend die Kontrolle über den Rechner des/der Testteilnehmenden. Das heisst, sie blockieren z.B. alle Internetbrowser, verhindern das Öffnen weiterer Fenster oder Dokumente ausser der Prüfungsapplikation etc. Dies soll sicherstellen, dass keine unerlaubten Hilfen in Anspruch genommen werden.

Zu den Herausforderungen gehört einerseits sicher die Anforderung, dass alle Komponenten der digitalen Remote-Prüfung (Item-Datenbank, Nutzerinterface, Datenübermittlung, Proctoring, ...) auf allen denkbaren Betriebssystemen gleichermaßen zuverlässig laufen (s.o.) und insbesondere auch dieselbe Darstellung haben müssen, und zwar für eine ganze Serie von verschiedenen (aktuellen, vergangenen und zukünftigen) Generationen dieser Betriebssysteme. Andererseits gehört zu den Herausforderungen der Anspruch, Testteilnehmenden mit weniger guten Sprachkompetenzen, weniger Schriftgewohnheit und/oder weniger ausgeprägten digitalen Skills ggf. benötigte Hilfe beim Download und der Installation der verschiedenen Applikationen auf den beiden Geräten, bei den geforderten Uploads von Dokumenten (Foto des eigenen Gesichts, Foto der Identitätskarte, Einwilligungsfomular etc.) und/oder beim Verständnis für die spezifischen Funktionen der einzelnen Geräte sowie für verschiedene Aspekte der Navigation zukommen zu lassen; eine Testanbieterin weist z.B. darauf hin, dass Testteilnehmende häufig das Handy nicht richtig befestigten, Applikationen und/oder Funktion/en von Handy und Computer verwechselten. Die VertreterInnen von Goethe-Institut digital weisen darauf hin, dass die Entscheidung, zuerst die B1- und B2-Prüfung digital anzubieten u.a. auch deswegen gefallen sei, weil auf/ab dieser Niveaustufe vorausgesetzt werden können, dass das Deutsche als Erklärsprache ausreichend gut beherrscht werde.

Generell gilt für alle Proctoring-Optionen die Datenschutzgesetzgebung der Schweiz. Da das Thema Integration und die entsprechenden Anforderungen in Bezug auf Sprachnachweise bei den einzelnen Kantonen liegen, müssten ggf. auch die kantonalen Datenschützer bzw. die Dachorganisation in die Entwicklung einbezogen werden (P. Dubois, SEM); dies wäre frühzeitig zu klären.

Soll der Test bereits vor Einreise in die Schweiz absolvierbar sein, d.h. auch von anderen Ländern aus abgenommen werden können, gelten zusätzlich die Datenschutzverordnungen der Länder, in denen sich die Testteilnehmenden zum Prüfungszeitpunkt befinden. Da die entsprechenden Verordnungen komplex

sind und sich auch beständig ändern (Hanahoe, Cambridge Assessment English), bräuchte es sicher in der Entwicklungsphase von Anfang an eine/n DatenschutzexpertIn im Team, voraussichtlich, aber auch längerfristig und ggf. unterstützt durch ExpertInnen vor Ort in den anderen Ländern. Der Datenschutzexperte von Cambridge Assessment English betont, dass es sich um eine zentrale und schwierige Aufgabe handelt, die von Anfang an mitgedacht werden muss, da sich Folgen für die Datenspeicherung (welche Daten dürfen gespeichert werden? wo/wie und für welchen Zeitraum?) ergeben:

“It’s very difficult to do, specially there’s new laws coming up all the time [...] You just need to make sure you’re considering them from the offset because it could have implications on data storage. For example, you might have to set up servers in-country to [...] store the data”

Er weist darauf hin, dass manche Länder (z.B. China, Russland) eine Datenspeicherung vor Ort im Land fordern (d.h. dass Serverplatz aufgebaut oder von bereits etablierten Drittanbietern gemietet werden muss), manche fordern sogar einen Vertreter vor Ort:

“some of the jurisdictions do require that you have a representative with the country.”

Die Testexpertin von ECL weist übrigens darauf hin, dass die von ihnen angebotenen Tests im Inland grundsätzlich als Remote-Version durchgeführt werden (angesichts der hohen Kosten für die Infrastruktur in Testzentren), im Ausland hingegen aus Sicherheitsgründen vor Ort in Testzentren.

Kritisch ist in Bezug auf Datenschutz die Erhebung und Speicherung sog. sensibler Daten, also vor allem biometrischer Daten für die Identitätsprüfung (u.a. Fotos der Testteilnehmenden und ihrer Identitätskarten) und für die Prüfungsüberwachung (Gesichtserkennung, Eyetracking-/Blickbewegungsdaten), wie sie in erster Linie beim maschinellen Proctoring entstehen. Diese Daten unterliegen besonderen Verwendungs- und Speicherungsbedingungen, der Zugang zu dieser Art Daten muss besonders gut geschützt sein (Schutz vor unerlaubtem Zugriff, vor Hackern etc.). In Pécs werden z.B. alle persönlichen und sensiblen Daten in einer Cloud gespeichert, Zugriff haben nur Befugte und es wird permanent überprüft, wer worauf zugegriffen hat, um unerlaubte Zugriffe zu identifizieren. Zu klären wäre, wer dafür im Fall von fide zuständig wäre (die Testzentren, die Geschäftsstelle fide, das SEM?).

In Teilen kann die Problematik der sensiblen Daten bei einer Remote-Durchführung ggf. umgangen werden, wenn die Identitätsprüfung live durch eine menschliche Aufsichtsperson stattfindet, da in diesem Fall ggf. kein Foto der Testteilnehmenden, zumindest kein Foto ihrer Identitätskarte notwendig wäre; falls das Proctoring während der Prüfung auch durch eine menschliche Aufsichtsperson gewährleistet wird (mit nahtlosem Anschluss an die Identitätsprüfung, um einen Personenwechsel zwischen Identitätsprüfung und Prüfungsdurchführung zu verhindern); soll das Proctoring (teilweise) maschinell gewährleistet werden, bleiben die Herausforderungen bezüglich des Datenschutzes bestehen. Menschliches Proctoring ist jedoch mit entsprechend hohem personellem Aufwand verbunden, vor allem im Falle einer 1-zu-1-Aufsicht; in dem Fall erhöht sich der Aufwand für die (auch technisch geschulten) Aufsichtspersonen erheblich.

Angesichts des öffentlichen Interesses an der Täuschungssicherheit bei Sprachnachweisen für Zwecke des Aufenthalts und der Einbürgerung geht der Datenschutzexperte Hanahoe (Cambridge Assessment English) davon aus, dass mit einer generellen Bewilligung für die Datenerhebung und -speicherung im Falle des fide-Tests gerechnet werden dürfte. Zusätzlich müssen aber die Testteilnehmenden jeweils explizit ihre Zustimmung geben; ein entsprechendes Einwilligungsfomular (“Ich stimme der Speicherung und Verwendung meiner biometrischen Daten im Rahmen der Prüfungsdurchführung zu”) wäre also noch vor Prüfungsbeginn digital zu unterschreiben. Kritisch ist dabei die Tatsache, dass die Testteilnehmenden diese Zustimmung jederzeit widerrufen können; ob die Prüfung im Falle später zurückgezogener Zustimmung gültig bliebe, wäre zu klären.

Wenn einzelne Testteilnehmende keine Zustimmung zu maschinellen Proctoring geben möchten, muss es eine Alternative geben, dass sie den Sprachnachweis trotzdem ablegen können. Für diese Personen könnte z.B. menschliches Proctoring eingerichtet werden, alternativ könnten diese Person auch digital vor Ort in einem Testzentrum geprüft werden oder die Papier-Version wählen, falls diese weiter besteht.

Maschinelles Proctoring durch künstliche Intelligenz ist laut Hanahoe von Cambridge Assessment English kostengünstig, schnell und höchst zuverlässig im Vergleich zu menschlichem Proctoring:

“human examiners [...] are not necessarily as accurate as proctors and as the AI”

Maschinelles Proctoring sei seiner Ansicht nach auch der nächste zentrale, essenzielle Schritt für die Entwicklung digitaler Prüfungen, allerdings ist maschinelles Proctoring wegen der besonderen Sensibilität der Daten (für die Gesichtserkennung, die Blickbewegungsmessungen) eben datenschutztechnisch auch besonders anspruchsvoll:

“I do think proctoring is the way forward. I think it's in the society we live now. It's going to be a necessity, but there is [...] some of the highest risk processing that you can do because you're using biometric data, you often use third party providers to do that [...] when you're using AI proctoring, this is where the data protection concerns really come in”

TestDaF hat angesichts des entsprechend tiefen Eingriffs in die Privatsphäre noch keine Entscheidung für einen Remote-Test getroffen (hingegen arbeitet das Goethe-Institut an der zeitnahen Einführung einer digitalen Remote-Version des A1-Tests). Generell ist früh im Entwicklungsprozess zu klären, was ein Proctoring-System im Kontext eines High-stakes-Tests einerseits mindestens können / leisten / überwachen muss, was ein solches Proctoring-System andererseits im Einklang mit der Schweizer Datenschutzgesetzgebung nicht darf.

Cambridge Assessment English verwendet angesichts der beschriebenen Herausforderungen in Bezug auf maschinelles Proctoring aktuell nur menschliches Proctoring, teilweise ergänzt durch maschinelles Proctoring, weil eine Aufsichtsperson mehrere Prüflinge überwacht (bis zu 10) und nicht immer überall gleichzeitig hinschauen kann.

Menschliches Proctoring verringert zwar die datenschutzrechtlichen Herausforderungen (falls tatsächlich keine biometrischen Daten erhoben und gespeichert werden), erhöht aber gleichzeitig den finanziellen und personellen Aufwand. Eine Expertin weist darauf hin, dass eine digitale Prüfung potenziell zusätzlichen Stress verursache (dies ist im Verlauf der Pandemie im Übrigen auch für Studierende festgestellt worden, dürfte also nicht nur für besonders schul- und digital ungewohnte Testteilnehmende gelten); man brauche daher in besonderem Masse erfahrene, ruhige, geduldige Aufsichts-/Support-Personen, die aufgetretene Fehler schnell identifizieren, das Problem besprechen und gemeinsam den Fehler beheben können (dies gilt im Übrigen für die digitale ortsabhängige wie die ortsunabhängige Variante). Dabei muss selbstverständlich eindeutig geklärt sein, (1) was genau die Aufsichtsperson im Zweifelsfall tun, wobei sie helfen darf (oder eben nicht) und (2) wie durch Navigations- oder Verständnis-, Software- oder Hardwareprobleme verlorene Zeit im Einzelfall kompensiert werden soll / kann; bei ECL z.B. wird Extra-Zeit beim Lese- und Hörverstehen gegeben, wenn die Fehlerbehebung ein bestimmtes Zeitlimit überschreitet; das Hörverstehen kann ggf. mit neuer Zeitberechnung neu gestartet werden.

Auch Hanahoe (Cambridge Assessment English) betont, dass Aufsichtspersonen umfassend technisch geschult werden müssen, ob durch das jeweilige Testzentrum oder durch die entsprechenden Drittanbieter, bei denen die Proctoringleistungen eingekauft werden. Wenn Proctoring von Drittanbietern eingekauft wird, muss im Vertrag eindeutig festgehalten werden, wer eingestellt werden darf und welche Art der Schulung und des Schulungsnachweises erbracht werden muss:

“make sure the liability flows down to the Proctor provider and you've got control of that”

Voraussetzung für die Einstellung von entsprechenden Aufsichtspersonen ist neben des Trainings-/Schulungsnachweises der Nachweis, dass keine Vorverurteilungen vorliegen.

Insbesondere weist Hanahoe darauf hin, dass es während der aktuellen Pandemie beispielsweise Fälle der Belästigung Minderjähriger durch solche Aufsichtspersonen gegeben habe.

MeasureLearning weist darauf hin, dass Proctoring für den Anteil “Sprechen” im Falle eines Video-Calls generell (noch) nicht möglich ist; diese Verantwortung müsste also der/die Prüfende bzw. Beisitzende übernehmen.

## 5.4 Technische Tools für automatische Korrektur

Im Kontext der Abklärung der technischen Möglichkeiten wurden ebenfalls die Optionen einer automatischen Korrektur geprüft. Die automatische Auswertung geschlossener Items (z.B. Multiple-Choice- und Zuordnungsaufgaben) ist technisch möglich und auch rekurstechnisch unproblematisch, da in diesen Fällen die Maschine / künstliche Intelligenz lediglich einen Abgleich zwischen der Lösung des/der Testteilnehmenden und klar vorprogrammierten Parameter-Settings durchführt.

Die automatische Auswertung offener Aufgaben, insbesondere freier Schreibaufgaben, ist hingegen sowohl technisch als auch rekurstechnisch höchst herausfordernd.

In Bezug auf die technische Umsetzung wäre zu klären, inwiefern bzw. unter welchen Bedingungen mehr oder weniger freie Schreibaufgaben in den vier Nationalsprachen auf niedrigen Kompetenzniveaus zuverlässig automatisch ausgewertet werden können. Es müsste grundlegend nachgewiesen werden, dass die automatische Korrektur zu mindestens 90–95% korrekt arbeitet, und zwar für alle Testsprachen gleichermaßen (einzelsprachspezifisch oder sprachübergreifend).

In Bezug auf die Rekursicherheit weist Hanahoe (Cambridge Assessment English) darauf hin, dass Testteilnehmende bei einer vollautomatischen Textauswertung eine manuelle Nachkorrektur fordern können:

“if you’re making a purely computer-based decision, then the individual has a right to object to that, and this really stems from consumer law”

Wie viele Testteilnehmende von dieser Möglichkeit Gebrauch machen würden, ist unklar. Eindeutig geklärt werden müsste nichtsdestotrotz, welche Rechte die Testteilnehmenden haben, falls im Falle einer solchen menschlichen Nachkorrektur die Note/Bewertung tatsächlich abweiche.

Insgesamt ist der personelle und zeitliche Aufwand für die Klärung der rechtlichen Fragen (in Bezug auf Datenschutz, Datenspeicherung, Proctoring, maschinelle Auswertung etc.) erfahrungsgemäss hoch (Hanahoe, Cambridge Assessment English):

“That can take time, so it’s very difficult to quantify how long these things take, but they do take a long time, they really do”

Die entsprechenden Überlegungen zu den rechtlichen Fragen müssen daher bereits früh im Entscheidungs- und Entwicklungsprozess unter Zuzug entsprechender ExpertInnen berücksichtigt werden, andernfalls kann es im weiteren Verlauf zu zeitaufwändigen Blockierungen (z.B. für Nachverhandlungen, Restrukturierungen, Programmierung, Serverinstallation, ...) und zu hohen Zusatzkosten kommen (Hanahoe, Cambridge Assessment English). Zu prüfen ist dabei auch, für welche Aspekte der digitalen Prüfungsdurchführung es Verordnungs- oder Gesetzesanpassungen braucht (Dubois, SEM).

## 6. Akzeptabilität und Zumutbarkeit

Im Folgenden fassen wir grundlegende Punkte zusammen, die in den Gesprächen mit den Stakeholdern (Geschäftsstelle fide, Integrationsbeauftragte, Testteilnehmende) in Bezug auf die Akzeptabilität eines digitalen fide-Tests aufgerufen wurden: Welche Aspekte der Digitalisierung erscheinen den Stakeholdern besonders problematisch bzw. aussichtsreich? Unter welchen Bedingungen könnte Akzeptabilität für eine Digitalisierung geschaffen werden?

Die befragten AkteurInnen im Bereich der sprachlichen Integration in der Schweiz (Geschäftsstelle fide, Integrationsbeauftragte, fide-Testanbieter-Institutionen) äussern eine gewisse Überraschung und Skepsis hinsichtlich der Erwägung des SEM, den fide-Test zu digitalisieren. Diese Überraschung und Skepsis ist nicht grundsätzlich mit einer negativen Haltung gegenüber digitalen Sprachprüfungen verbunden. Sie bezieht sich vor allem auf die aktuelle Hauptzielgruppe des fide-Tests, die eher schulungsgewohnt ist und häufig nicht über digitale Basiskompetenzen und/oder die für eine digitale Prüfung nötige technische Ausstattung verfügt. Die Vertreterin einer Test-Anbieter-Institution schildert ihre Reaktion auf die Nachricht der Machbarkeitsanalyse:

“Ich war schon etwas erstaunt und überrascht, dass es [die Digitalisierung des fide-Tests] so schnell kommt. Ich finde es aber wichtig, dass man auf dem Zug der Digitalisierung bleibt. Für uns ist es einfach wirklich das Problem der schulungsgewohnten Teilnehmenden mit den fehlenden Basiskompetenzen in der Digitalisierung [...]”.

### 6.1 Integrationsbeauftragte, Geschäftsstelle fide

Ähnlich reagierte eine Interaktionsbeauftragte, die die Kompatibilität eines digitalen fide-Tests mit dem ganzen fide-Testkonzept in Frage stellt:

“Einerseits im Zuge der Digitalisierung und aufgrund der Pandemie-Situation war ich nicht sehr erstaunt, dass diese Überlegungen jetzt, zu diesem Zeitpunkt kommen, also das passt ja noch; und gleichzeitig kam mir in den Sinn: aber wie ist es kompatibel mit dem ganzen fide-Testkonzept der Niederschwelligkeit und der Ausrichtung auf eher lernungsgewohnte Personen?”

Die Gefahr, dass durch die Digitalisierung des fide-Tests “vieles, was den fide-Test ausmacht, verloren gehen sollte oder könnte”, sehen auch die interviewten VertreterInnen der Geschäftsstelle fide (für weitere Ergebnisse bezüglich der Passung eines digitalen fide-Test für die aktuelle Hauptzielgruppe s. Kapitel 7).

Die Skepsis der befragten AkteurInnen im Bereich der sprachlichen Integration bezieht sich ebenfalls auf die Frage, ob eine so grosse Veränderung das noch nicht wirklich etablierte System vor neue Herausforderungen stellen würde. Zurückhaltend ist z.B. ein Integrationsbeauftragter aus der französischsprachigen Schweiz, dessen Einschätzung nach zunächst die bestehende Papier-Version verbessert werden sollte:

[...] “il faudrait déjà qu'ils améliorent le test actuel et puis les supports, tous ces éléments-là, avant de se lancer dans une version numérique”. [...]

Mit der Verbesserung der aktuellen Version meint der Integrationsbeauftragte hier die digitalen Administrationsprozesse, aber auch die vorhandenen Audiodateien.

In Zusammenhang mit eventuellen Änderungen im erst kürzlich etablierten System der Sprachkompetenzprüfung wird auch der Aufwand für die Testanbieter-Institutionen sowie für die Prüfenden angesprochen. Eine digitale Version würde die erst kürzlich etablierten und noch nicht perfekt funktionierenden Strukturen und Prozesse in Frage stellen: Bei einer digitalen ortsunabhängigen Version bräuchte es ggf. die Anbieter-Institutionen nicht mehr, so die Einschätzung der Anbieter; allerdings muss die Prüfungsverwaltung natürlich personell abgesichert sein. Bei einer digitalen ortsabhängigen Version mit Testanbietern als Testzentren verändern sich die Rollen und Zuständigkeiten, ebenfalls die Aufgaben und Kompetenzprofile der Prüfenden. Ein Integrationsbeauftragter befürchtet, dass die Aufgaben für die Prüfenden kaum interessant und attraktiv sein werden, wenn diese sich nur auf die Prüfungsaufsicht beschränken.

Die Geschäftsstelle fide ihrerseits schätzt den Aufwand der Organisation eines digitalen ortsunabhängigen oder ortsabhängigen Tests mit den aktuellen Ressourcen als kaum zu bewältigen ein. Man müsste personell aufstocken (die Schätzung geht bis zu einer Verzehnfachung des Personals). Im Falle eines digitalen Tests vor Ort bräuchte es Testzentren in allen grösseren Städten.

## 6.2 Anbieterinstitutionen

Die Test-Anbieter-Institutionen stellen sich ihrerseits die Frage, wer im Falle einer digitalen Vor-Ort-Version die Kosten der technischen Ausrüstung und Wartung übernehmen würde. Kleinere Institutionen sehen keine Möglichkeit einer Selbstübernahme, die Kantone hätten ebenfalls Schwierigkeiten, Mittel zur Verfügung zu stellen. Ein Teil der Institutionen verfügt bereits jetzt über die technische Infrastruktur, vor allem, weil sie in der Corona-Zeit für den Online-Unterricht die nötige Infrastruktur aufbauen musste.

Herausforderungen sehen die Testanbieter nicht nur in Hinblick auf die Infrastruktur, sondern auch auf das fehlende technische Know-how. Sie verfügen nicht über ausreichend Personal, das die technische Unterstützung gewährleisten könnte (z.B. im Testzentrum oder als Technical-Support-Hotline). Auch die Prüfenden selbst müssten eventuell über andere Kompetenzen verfügen, bzw. die Prüfenden müssten für eine eventuell Abnahme von digitalen Prüfungen vor Ort geschult werden.

Die Geschäftsstelle fide unterstreicht, dass die Akzeptanz eines digitalen Tests nur dann gelingt, wenn es keine Kostensteigerung für die Anbieter gibt. Lieferung, Installation und Wartung von Hardware sollten weder die Anbieterinstitutionen noch die Testteilnehmenden finanziell belasten.

## 6.3 Akzeptabilitätsbedingung „Papierversion“?

Konsens besteht bei allen Schweizer Stakeholdern, dass eine digitale Prüfungsversion nicht die fide-Kernzielgruppe der schul- und schriftungewohnten Lernenden benachteiligen oder gar ausschliessen darf. Sämtliche AkteurInnen der sprachlichen Integration nennen u.a. deswegen als Bedingung für die Akzeptanz eines digitalen fide-Tests die Co-Existenz einer Papier-Version, so dass die schulungsgewohnte Zielgruppe des fide-Tests nicht ausgeschlossen oder benachteiligt wird. Dabei stellt sich die Frage der Gleichwertigkeit der zwei dann parallelen Versionen, die durch Begleitforschung geprüft werden müsste. (Andererseits stellt sich die Frage, ob ein Beharren auf der Papier-Version für die schulungsgewohnten Teilnehmenden diese nicht gerade in einer neuen Lockdown-Phase ressourcen- und alternativlos zurücklassen würde, s.u. *Szenarien*.)

Die Co-Existenz einer Papier- und einer digitalen Version des fide-Tests unterstützen auch Lernende, die eine Teilnahme an einer fide-Prüfungssession geplant haben. Auch wenn die Papier-Version deutlich bevorzugt wird, lehnen die Lernenden eine digitale Prüfung nicht per se ab bzw. zeigen sich einige bereit, "das auszuprobieren". Bedingung dafür wäre, dass die digitale Prüfung im Testzentrum und auf dem Handy, eventuell auch auf Tablets, stattfindet. Testzentren werden insbesondere von Lernenden bevorzugt, die mit ihren Familien zusammenwohnen, und deshalb unsicher sind, ob sie während der Prüfung zuhause ungestört arbeiten können.

Eine digitale Prüfung auf dem Smartphone können sich die Lernenden am ehesten vorstellen, da sie das Smartphone bereits privat und zunehmend auch im Kontext des Deutschlernens verwenden. Tablets werden gegenüber Computer/Laptop favorisiert, weil sie ähnlich wie Smartphones zu bedienen sind (Touchscreen). Geschlossene Items (Multiple Choice, Drag and Drop), die rezeptive Fertigkeiten testen, sind aus Sicht der Teilnehmenden mit entsprechender Vorbereitung bewältigbar. Auch den Teil "Sprechen" muten sie sich eher zu, wenn sie zum Beispiel in Zoom geprüft würden. Sie seien aufgrund des Online-Unterrichts in der Pandemie-Situation vertraut mit Video-Calls. Ausserdem nutzen sie Video-Calls im Privaten (Videoanrufe mit Familie und FreundInnen). Hingegen wird das Schreiben insbesondere auf Tastatur von der Gesamtheit der interviewten Lernenden für kaum bewältigbar gehalten. Digitale Tastaturen auf Tablets wären zwar etwas einfacher zu bedienen, aber für den Teil Schreiben ist die Papier-Version deutlich bevorzugt.

Das deckt sich nicht mit der Einschätzung der ExpertInnen im Bereich sprachliche Integration, die bei einer möglichen Teildigitalisierung gerade an den schriftlichen Teil des fide-Test denken. Sie sehen in der Digitalisierung des schriftlichen Teils die Möglichkeit einer Erhöhung der Testauthentizität (s. oben Kapitel 4 und unten Kapitel 7), da die aktuelle Papier-Version Aufgaben wie das Schreiben einer SMS oder einer Mail umfasst.

Jedenfalls braucht es für die Herstellung von Akzeptanz für einen digitalen fide-Test von Anfang an ein gutes **Kommunikationskonzept** für die verschiedenen Zielgruppen, insbesondere müssten folgende Parteien früh involviert werden:

- Kantone und kantonale Integrationsbeauftragte.
  - Mit ihnen müssen Verantwortlichkeiten geklärt werden.
  - Es muss aufgezeigt werden, dass eine digitale Prüfung täuschungssicher ist.
  - Es muss aufgezeigt werden, dass in einer digitalen Prüfung die schriftfernen, schulungsgewohnten Lernenden nicht benachteiligt werden.
- Testanbieter und Geschäftsstelle fide; mit ihnen müssen Rollen und Zuständigkeiten geklärt werden (Testadministration etc.).
- Testteilnehmende, vor allem die Gruppe der schul- und schriftungsgewohnten Lernenden; mit ihnen müssen Chancen und Risiken diskutiert, individuelle Bedürfnisse und potenzielle Anpassungsmöglichkeiten und Optionen abgewogen werden.

Zusätzlich zur Kommunikation ist eine Reihe von weiteren **Begleitmassnahmen** vorzusehen, für die Variante digital vor Ort und/oder digital remote, u.a.

- Schulungen in Testzentren
- Schulung von Aufsichtspersonen (vor Ort, online)
- Schulung und Bereitstellung von Technik-Support in auf A1-Niveau verständlicher Sprache (vor Ort, online)
- Betreuung der Testteilnehmenden (Support vor der Prüfung, u.a. für Anmeldung, Erklärungs- und Übungsoptionen; Support während der Prüfung, u.a. für Login, Identitätsprüfung, bei technischen Problemen, jeweils in verständlicher Sprache)

## 7. Passung für die Zielgruppe(n) und Herausforderungen

In Bezug auf Herausforderungen und Risiken der Testdigitalisierung und entsprechend auch der Akzeptabilität beschäftigt die Frage der Passung eines digitalen Tests für die aktuelle Kern-Zielgruppe des fide-Tests – d.h. die schulungsgewohnten, schriftfernen Testteilnehmenden – die befragten ExpertInnen der sprachlichen Integration in der Schweiz am meisten. Sie bezieht sich auf

1. die digitalen Kompetenzen der Testteilnehmenden
2. die schriftsprachlichen Kompetenzen der Testteilnehmenden
3. die Verfügbarkeit der Infrastruktur
4. lernpsychologische Aspekte, die einen Einfluss auf die Prüfungsleistung haben können

### 7.1 Digitale Kompetenzen

Die ExpertInnen im Bereich der sprachlichen Integration schätzen die digitalen Kompetenzen der aktuellen Haupt-Zielgruppe so ein, dass sie von einem erheblichen Einfluss des Testformats auf das Testergebnis ausgehen. Dabei würde es sich um eine Verletzung des Qualitätskriteriums der Interaktivität handeln. Das heisst, dass die Testergebnisse nicht aufgrund der zu messenden sprachlich-kommunikativen Kompetenzen, sondern aufgrund der digitalen Kompetenzen variieren würden (vgl. Tschirner 2000).

Auf diese Gefahr weisen ebenfalls die internationalen ExpertInnen der ALTE oder des ACTFL (American Council on the Teaching of Foreign Languages) hin. Es muss sichergestellt werden, dass die Testleistung nur aufgrund der sprachlichen Kompetenz und nicht aufgrund des Testformats variiert. Eine Lösung könnte sein, die Förderung digitaler Kompetenzen in die Curricula der Sprachkurse zu integrieren; dies wurde laut ALTE-Experten zum Beispiel in Norwegen gemacht. Es gäbe aber dennoch auch in diesem Fall eine Verschiebung des Testkonstrukts, die transparent dargelegt werden müsse. Diese transparente Verschiebung des Testkonstrukts wäre auch für ExpertInnen im Inland denkbar, insbesondere in Hinblick auf die Sprachhandlungen im schriftlichen Teil, die mit digitalen Medien stattfinden (Mail, SMS, Formular). Die Triage sei aber kompliziert, betont eine Integrationsbeauftragte, denn einige Testteilnehmende verfügten bereits über digitale Kompetenzen und möchten nur die Sprache lernen.

Die Frage, wie sie die Entwicklung der digitalen Kompetenzen der aktuellen Hauptzielgruppe des fide-Tests einschätzen, spaltet die Gruppe der ExpertInnen: Während einige der Meinung sind, dass ein digitaler Test auch in 10 Jahren noch eine wichtige Anzahl von Prüfungsteilnehmenden (u.a. Ältere, Frauen, Nichterwerbstätige, Erwerbstätige in handwerklichen Berufen, Schulungsgewohnte) benachteiligen würde, gehen andere davon aus, dass sich die Problematik mit der Zeit abschwächen bzw. erledigen würde, da bereits jetzt viele jüngere Teilnehmende über ausreichende digitale Kompetenzen verfügen.

Wie oben bereits beschrieben zeigt die Desk-Research, dass Aufgaben so dargestellt werden können (s. z.B. Goethe digital, Duolingo, SIELE, Norwegian Language Test), dass sie auch Teilnehmende mit tieferen digitalen Kompetenzen eventuell weniger beanspruchen würden.

Insbesondere im Hinblick auf die Evaluierung der Hör- und Lesekompetenzen sowie des Sprechens erscheint eine digitale Prüfung insgesamt eher unproblematisch zu sein; dies unter der Voraussetzung, dass die Aufgabenformate, Aufgabenformulierungen und das Nutzerinterface gut durchdacht sind und es eine (ggf. begleitete) Vorab-Trainingsmöglichkeit gibt.

Im Falle einer ortsunabhängigen Version (v.a. bei maschinellem Proctoring) sind zusätzliche Prozesse und Aktivitäten notwendig, die höhere digitale Kompetenzen voraussetzen (z.B. Identitätsprüfung mit Hochladen von Identitätskarte, Installation der Apps auf den verschiedenen Geräten, Installation des Handys, so dass das Filmen möglich ist etc.). Sämtliche ExpertInnen sind sich deshalb einig, dass eine Remote-Version noch anspruchsvoller wäre (und wahrscheinlich eine Überforderung eines Teils der Teilnehmenden darstellen würde).

## 7.2 Schriftsprachliche Kompetenzen

Wie die Desk Research zeigte, können digitale Tests von den Aufgabenformaten und der Darstellung her so gestaltet werden, dass sie keine hohen digitalen Kompetenzen seitens der Teilnehmenden voraussetzen, sie bleiben aber in der Regel stark schriftbasiert (s. oben Kapitel 4). Die aktuelle Hauptzielgruppe des fide-Tests ist häufig aber wenig literalisiert, eine Navigation durch schriftliche Anweisungen könnte sie deshalb überfordern, meinen VertreterInnen der Geschäftsstelle fide, der Integrationsbeauftragten und der Testanbieter-Institutionen.

Es wurde von den EntwicklerInnen des fide-Tests sehr viel Wert darauf gelegt, den mündlichen Teil der Prüfung (Hören und Sprechen) möglichst schriftfrei zu gestalten, so dass wenig literalisierte Personen die Möglichkeit haben, ihre tatsächlichen mündlichen Sprachkompetenzen nachzuweisen; dieses zentrale Testmerkmal soll auch in einer digitalen Variante unbedingt erhalten bleiben. In diesem Teil ist die Unterstützung der Prüfenden, welche die Prüfungsteilnehmenden durch die einzelnen Aufgaben führen, von grosser Bedeutung. Sollte diese Unterstützung fehlen und durch weitere technische Lösungen (z.B. aufgezeichnete mündliche Anweisungen, Anweisungen durch Videos) nicht gewährleistet werden können, wären diese Personen benachteiligt.

Die schriftlichen Anweisungen, die sich auf die Identifikationsprozesse und Täuschungssicherheit im Rahmen eines digitalen ortsunabhängigen Tests beziehen, sind sprachlich und inhaltlich noch deutlich komplexer. Die ExpertInnen der sprachlichen Integration weisen deshalb nochmals explizit auf die Gefahr hin, wenig literalisierte Personen zu benachteiligen. Die längeren schriftlichen Anweisungen verbunden mit der Prüfungsangst seien definitiv eine grosse Herausforderung für die aktuelle Kernzielgruppe des fide-Tests; auch die Goethe-ExpertInnen weisen mehrfach auf die Schwierigkeit hin, auf niedrigen Niveaustufen verständliche Anleitungen und auch Hilfefunktionen anzubieten.

Am ehesten problematisch erscheint in einer digitalen Prüfung in Bezug auf die literalen Kompetenzen die Umsetzung des Anteils "Schreiben" (s.o. Kapitel 4.3). Dieser Prüfungsteil könnte am ehesten die eigentliche Kerngruppe des fide-Tests benachteiligen, die wenig schul- und schriftgewohnt, aber vor allem nicht ans Schreiben auf einer (zudem auch noch deutschen) Tastatur gewöhnt sind. Im Kapitel 4.3 sind Optionen beschrieben, mit denen potenziell eine Benachteiligung der schul- und schriftungewohnten Testteilnehmenden verhindert werden kann. Begleitforschung müsste zeigen, ob eine der skizzierten Alternativen bzw. welche tatsächlich effektiv ist.

In den Interviews stellte sich auch die umgekehrte Frage, ob das digitale Schreiben auch eine Entlastung für wenig literalisierte Personen sein könnte. Die könnte unter Umständen der Fall sein, wenn minimale digitale Kompetenzen bereits in den Sprachkursen vorgespurt werden (z.B. durch WhatsApp-Gruppen zum Austausch mit der Lehrperson). In diesem Fall könnten beispielsweise für SchreibanfängerInnen kognitiv und motorisch aufwändige hierarchieniedrige bzw. basale Schreibfertigkeiten wie leserliche Handschrift in einer digitalen Umgebung (Touchscreen) entlastet werden; auch das eigene Produzieren der Buchstabenformen würde durch ein Wiedererkennen und Auswählen von der Tastatur ersetzt werden. Ob die Rechtschreibung zusätzlich durch ein reduziertes Korrekturprogramm innerhalb der digitalen Testumgebung unterstützt werden soll, müsste geprüft werden. Dies würde es schriftungewohnten Lernenden auf niedrigen Niveaus erlauben, dem Textproduktionsprozess und den hierarchiehöheren Schreibfertigkeiten (Planen, Formulieren und Überarbeiten) mehr Aufmerksamkeit zu widmen (Sturm 2014).

Obwohl für die schrift- und wohl auch tastaturgewohnteren Testteilnehmenden hier potenziell das Argument greift, das digitale Schreiben sei in vielen Fällen (z.B. für Mails, SMS) lebensnäher als das Schreiben auf Papier, müssen jedoch aktuelle Forschungsbefunde berücksichtigt werden, die selbst für (schul-, schrift- und tastaturschriftgewohnte) BA-Studierende aufzeigen, dass das Schreiben an der Tastatur im Vergleich mit handschriftlichem Schreiben eine kognitive Zusatzbelastung darstellt (erhöhte Reaktionszeiten, geringere Gedächtnisleistung):

"This implies that, even in undergraduates, typing is still more effortful than handwriting and therefore has a negative impact on performance on an ongoing activity." (Bouriga & Olive 2021)

### 7.3 Den Prüfungsteilnehmenden zur Verfügung stehende technische Infrastruktur

Wie oben erwähnt wären für die Teilnahme an einer digitalen remote Prüfung, die von einem Proctoring-System überwacht wird, zwei Geräte nötig. Sämtliche interviewte Lernende verfügen zwar über ein Smartphone, wenige aber über ein Tablet. Bei Teilnehmenden mit Familie sind Computer / Tablets zwar oft im Haushalt vorhanden, gehören aber den Kindern. Dieses Bild deckt sich mit der Einschätzung der ExpertInnen in Hinblick auf die aktuelle Kernzielgruppe des fide-Tests. Die Vertreterin einer Testanbieter-Institution spricht die Möglichkeit der Geräteausleihe an, sei es vom Freundeskreis, von den Sprachkursanbietern oder von den Testzentren.

Die Qualität der Internetverbindung der potenziellen Prüfungsteilnehmenden beschäftigt die ExpertInnen ebenfalls. Je nach Gebiet (städtisch/ländlich) und Vertrag (Debit) kann diese variieren. Die Lernenden selbst berichten von Verbindungsproblemen, wenn sie mit ihren Verwandten Video-Calls durchführen, auch wenn diese nicht allzu häufig auftraten. Angesichts u.a. des Hinweises darauf, dass bei schlechter Verbindung ggf. Aufgaben-Items vom System selbst übersprungen werden, ist dies aber kein geringes Problem.

### 7.4 (Lern-)Psychologische Funktionsfähigkeiten

Bei lern- und schriftgewohnten Lernenden scheint die Akzeptanz einer digitalen Prüfung (vor Ort, ab B-Niveaus) hoch zu sein (Goethe-Institut); für Lernende auf niedrigeren Niveaus und allgemein eher lernungewohnte Testteilnehmenden ist hingegen eher mit Bedenken und Prüfungsangst zu rechnen (Goethe-Institut; Casanova, ALTE), insbesondere bei einer ortsunabhängigen digitalen Prüfung.

Casanova (ALTE) ebenso wie VertreterInnen von Test-Anbieter-Institutionen betonen die Relevanz der Lehrperson beim Abbau von Prüfungsstress. Diese Unterstützung würde bei einer ortsunabhängigen Prüfung wegfallen – ausser es wird ggf. eine personell aufwändige Variante menschlichen 1-zu-1-Proctorings im Sinne einer Online-Prüfungsbegleitung in Betracht gezogen.

Gerade für die schulungsgewohnten Testteilnehmenden sei jedenfalls ein gutes Kommunikationskonzept für den Umgang mit diesen Ängsten notwendig, ausserdem ein sehr klarer Aufbau der Aufgabenumgebung und eine sehr klare Darstellung und Formulierung der Erklärungen (Goethe-Institut).

Bei einer digitalen Prüfung in Testzentren wird davon ausgegangen, dass die Prüfungen vor allem in grösseren Städten stattfinden. Die Interviewten sind sich einig, dass dies ebenfalls eine zusätzliche Belastung für die Prüfungsteilnehmenden sein könnte. Dass die Prüfung bislang häufig an einem Ort stattfindet, mit dem die Prüfungsteilnehmenden vertraut sind, sei laut Integrationsbeauftragten und VertreterInnen der Geschäftsstelle fide hilfreich.

## 8. Potenziale digitaler Prüfungen

Trotz erster zurückhaltender Reaktionen nennen ExpertInnen ebenfalls Vorteile, die eine Digitalisierung des fide-Tests mit sich bringen könnte. Als Erstes wird im Falle einer digitalen Remote-Version die Ortsunabhängigkeit und deshalb auch die Durchführungssicherheit genannt. Die Erfahrungen aus der Pandemie (Absage von Prüfungen, Raumkapazitätseinschränkungen, Maskenpflicht, Distanzregelungen) führen dazu, dass diese Aspekte als sehr wichtig erachtet werden.

### 8.1 Angebotsoptimierung

Darüber hinaus verändere die Pandemiesituation die Erwartungen der Testteilnehmenden. Insbesondere werde heutzutage Flexibilität im Hinblick auf die Prüfungstermine erwartet, so die Testanbieter. Eine Digitalisierung, insbesondere eine ortsunabhängige Version, könnte die aktuelle Hauptzielgruppe des fide-Tests erweitern und den Test stärker auf dem Markt positionieren. Aktuell wird der Test insbesondere von schulungsgewohnten Teilnehmenden abgelegt; Testanbieter und Integrationsbeauftragte gehen aber davon aus, dass eine digitale Remote-Version die Attraktivität des fide-Tests für schulungsgewohnte MigrantInnen erhöhen würde. Es gehe dabei auch um das Testimage, sagt die Vertreterin einer Anbieter-Institution. Aktuell leide dieses Image durch die mangelnde internationale Anerkennung des fide-Tests und/oder durch die Tatsache, dass der fide-Test hauptsächlich für eine schulungsgewohnte Zielgruppe konzipiert worden ist.

Allein die Digitalisierung würde für die Markterweiterung allerdings nicht reichen, ergänzt eine Integrationsbeauftragte aus einer städtischen Region der Deutschschweiz; der fide-Test solle auch höhere Sprachniveaus prüfen, denn die Niveaustufen B2 und C1 sind für gewisse Berufsfelder und das Studium Mindestvoraussetzungen. Gleichzeitig erwähnt sie die Zielgruppe, die eventuell vor der Einreise in die Schweiz ihre Sprachkompetenz nachweisen muss. In diesem Fall könne eine digital Remote-Version durchaus attraktiv sein, und zwar für schulungsgewohnte und schulungsgewohnte Teilnehmende.

In diesem Zusammenhang weisen ExpertInnen auf die Gefahr des Prüfungsmissbrauchs hin. Aktuell reiche für die Einreise in die Schweiz die Anmeldung für einen Sprachkurs; es könne hingegen sein, dass der Nachweis der Niveaustufe A1 vor der Einreise tatsächlich verlangt werde, wenn eine digitale Remote-Version vorhanden sei.

### 8.2 Realitätsnähe

Weiterhin wird von den ExpertInnen die Möglichkeit genannt, mithilfe der Technologie realitätsnähere Anlässe zu schaffen. Dabei handelt es sich – wie oben bereits erwähnt – primär um die Möglichkeit, die Authentizität der Schreibaufgaben zu erhöhen. In den Prompts wird nämlich das Schreiben als digitale Aktivität beschrieben (Mail / Nachricht schreiben), aber das eigentliche Schreiben in der Prüfung findet auf Papier statt.

Ferner geht es aber auch um Ideen, die den Einsatz von Videos oder von Virtual Reality zur Visualisierung der Szenarien anregen. Falls dies eine Möglichkeit wäre, sollte diese Visualisierung jedoch die unterschiedlichen Realitäten der Lernenden berücksichtigen, betont ein Integrationsbeauftragter aus der Westschweiz. Er geht davon aus, dass MigrantInnen, die in ländlichen Regionen wohnen, Schwierigkeiten hätten, zum Beispiel zu erkennen, dass es um eine Post geht, wenn diese mit der Sihlpost in Zürich illustriert werden sollte.

### 8.3 Digitale Grundkompetenzen

Schliesslich sehen die ExpertInnen auch eine Chance bei einer möglichen Erweiterung des Testkonstrukts auf digitale Grundkompetenzen. Heutzutage seien nicht nur Sprachkompetenzen, sondern auch digitale Kompetenzen eine Bedingung für die Teilnahme am gesellschaftlichen und beruflichen Leben. Bei einer

Erweiterung des Testkonstrukts sollte aber unbedingt die Förderung der digitalen Kompetenzen berücksichtigt werden. Ob und wie diese (wie z.B. in Norwegen) bereits im Rahmen der Sprachkurse stattfinden soll, wäre genauer zu prüfen.

Ähnliche Überlegungen betreffen ggf. auch die digitale Qualifikation der Lehr- und Testpersonen; ein wünschenswertes *wash-back* der Digitalisierung des fide-Tests (d.h. ein positiver Effekt des Testens auf das Lehren und Lernen von Sprachen, Curriculumsentwicklung etc.) wäre an dieser Stelle eine Beschleunigung der Aus- und Weiterbildung der Lehrpersonen in Bezug auf die Anwendung digitaler Kompetenzen beim Lehren von Sprachen sowie in Bezug auf die Vermittlung digitaler Kompetenzen im Kontext des Sprachenlernens, die in absehbarer Zeit sowieso notwendig sein wird, auch nach der Pandemie. Dies könnte z.B. die Form eines weiteren Moduls „Digitales Lernen, Lehren und Testen“ in der Zertifizierung der Sprachkursleitenden im Integrationsbereich annehmen.

## 9. Kosten / Aufwand – Sparpotenziale

Allgemein ist damit zu rechnen (s. u.a. TestDaF, Goethe-Institut, ECL), dass die Digitalisierung personell und finanziell aufwändig ist, v.a. der Anteil der Programmierung (Malone/ACTFL), aber ggf. auch die Neukonzeption des Testkonstrukts (in Bezug auf die Rolle von Sprach- und/oder Grundkompetenzen) und die Neukonzeptualisierung von (Teilen der) Aufgaben. Eine genaue Schätzung des Aufwands für eine Digitalisierung ist auf Basis der uns verfügbaren Informationen nicht möglich, da weder detaillierte Aussagen zu Kosten noch genaue Angaben zum Sparpotenzial vorliegen; auch bezüglich der Rentabilität können angesichts der Vielfalt der möglichen Szenarien (siehe Kapitel 10) und Optionen (sowie der teilweise klar ausweichenden Antworten unserer InterviewpartnerInnen aus dem Bereich der EntwicklerInnen und AnbieterInnen digitaler Prüfungen) keine konkreten Aussagen gemacht werden. Eine Annäherung an die erwartbaren Kosten und Aufwände findet sich in den vier Beispielen unter 9.3.1 bis 9.3.4. Genauere Schätzungen können u.E. erst vorgenommen werden, wenn u.a. klar ist, welches der möglichen Szenarien genau angestrebt und wo/wie die Leistung genau eingekauft wird.

### 9.1 Digitaler Test vor Ort: Aufwand / Kosten

Für eine digitale Durchführung vor Ort müssen folgende Systemkomponenten entwickelt (bzw. bei einem Drittanbieter lizenziert und ggf. angepasst) werden:

- ein Verwaltungstool / -interface für die Prüfungsadministration, u.a. für
  - die Vergabe der Prüfungstermine
  - die Organisation der Aufsichtspersonen (und ggf. die KorrektorInnen für die nicht automatisch korrigierten Prüfungsteile)
  - die Ergebnismitteilung (Integrationsinstitutionen, Testteilnehmende)
- eine Item-Datenbank und ein Nutzerinterface, u.a. für
  - die Darstellung verschiedener Aufgabentypen (inkl. Einbettung von Audio-, Bild- und Video-Dateien)
  - die Verwaltung und Aktualisierung der Aufgabenitems
  - die (pseudo-) randomisierte Einspielung und Darstellung der Aufgaben unter Berücksichtigung (1) ggf. der Adaptivität von Testteilen und (2) des Ausschlusses von sog. Overexposure (d.h. eine Limitation der Ausspielung einzelner Items auf maximal 2000–5000 Prüfungen)
  - Hilfen/Erklärungen zur Navigation, zu einzelnen Aufgabentypen etc. (in einfacher Sprache, mit Bild-Unterstützung)
- ggf. eine Übungslandschaft und weitere Trainings- und Support-Optionen wie z.B. Erklärvideos

Die Entwicklung / Programmierung der eben genannten Software stellt nach Auskunft verschiedener TestexpertInnen den grössten finanziellen, personellen und zeitlichen Aufwand dar (“one of the biggest costs is programming [...] it’s the reprogramming”, Malone/ACTFL), Kosten entstehen aber auch für die Ausstattung der Testzentren mit Hardware, u.a.

- Arbeitsgeräte (PCs, Laptops oder Tablets; Headsets): Anschaffung, Wartung, regelmässige Erneuerung
- Internetverbindung, Serverkapazitäten und IT- / Datensicherheits- und Datenschutzmassnahmen

Ausserdem entstehen wie bisher bzw. neu Kosten für

- die Aufsichtspersonen, inkl. technischem Support (tendenziell mehr Aufsichtspersonen als bei der Papier-Version, ECL)
- weiterhin zwei Prüfende (1 Prüfende/n und 1 Beisitzende/n) pro TestnehmerIn, falls der Anteil “Sprechen” im Video-Call durchgeführt werden soll
- die Räume
- die Anreise aller Beteiligten
- die Schulung(en) der Aufsichtspersonen und Prüfenden

- das regelmässige Update der Item-Datenbank, d.h. die Entwicklung und Erprobung neuer, genau gleichwertiger Aufgaben-Items

Weiterhin müssten für folgende Aspekte die entsprechenden technischen und rechtlichen Klärungen und Begleitmassnahmen in die Wege geleitet werden (ggf. unter Einbezug externer ExpertInnen):

- Datensicherheit und Datenschutz sowie die entsprechenden Zuständigkeiten  
Bisher gebe es laut Aussage einer Vertreterin einer Bürgergemeinde für die Anbieterinstitutionen keine besonderen gesetzlichen Verpflichtungen bzgl. Datenschutz, das würde sich mit der Einführung eines digitalen Sprachnachweises wohl ändern; ggf. würde dies einen erheblichen zusätzlichen Aufwand für die Anbieterinstitutionen darstellen.
- Umfang und Art des (technischen) Supports während der Prüfung sowie Schulung der Aufsichtspersonen: Was muss die Aufsichtsperson können? Wobei muss/darf die Aufsichtsperson helfen? Welche Rechte/Möglichkeiten haben die Testteilnehmenden im Falle eines technischen Problems? Wie werden Zeitverluste durch technische Probleme kompensiert?  
Beim TestDaF digital z.B. darf die Aufsichtsperson nur in Ausnahmefällen wie kompletten Computercrashes eingreifen, ansonsten sind lediglich Erklärvideos verfügbar; die Testteilnehmenden können sich im Falle technischer Probleme jedoch kurz aus- und wieder einloggen und nach dem erneuten Login weitermachen, wo sie stehen geblieben sind. Bei den digitalen Goethe-Prüfungen kann die Aufsichtsperson die Prüfung für alle oder eine Person pausieren (z.B. bei Feueralarm oder Hardware-Problemen), sie darf auch Navigationsoptionen erklären (bzw. zeigen, wo man klicken kann bzw. muss), aber nicht anstelle der Testteilnehmenden den Klick ausführen.

## 9.2 Digitaler Test vor Ort: Sparpotenziale

Verschlinkungs- und Sparpotenzial entsteht bei der digitalen Prüfung vor Ort vor allem in Bezug auf

- Druck- und Versandkosten (kein Versand der Materialien an die Testzentren)
- Administration (digitale Ausstellung des Sprachnachweises, ggf. digitale Anmeldung)
- Korrektur
  - kein Versand der abgelegten Prüfungen an die KorrektorInnen und zurück an die Testzentren (bzw. keine Anreise von KorrektorInnen in die Geschäftsstelle für zentrale Korrektur)
  - teilweise automatisierte Korrektur zumindest der geschlossenen Aufgabenitems; es könnte zumindest eine hybride Evaluierung angestrebt werden – eine vollständig automatisierte Korrektur wirft die oben genannten datenschutz- und rekurstechnischen Komplikationen auf und wäre in der Entwicklung erst einmal sicher finanziell sehr aufwändig.
- Entwicklung, Erprobung, Validierung und Optimierung von neuen Aufgaben/Items

## 9.3 Digitaler ortsunabhängiger Test: Aufwand / Kosten für die Entwicklung

Für die digitale Remote-Version werden ebenfalls die drei oben bereits genannten grundlegenden Komponenten (Verwaltungstool, Item-Datenbank und Nutzerinterface, Übungslandschaft und Supportmedien) benötigt; dabei ist sicherzustellen, dass die Software auf verschiedenen Betriebssystemen (s.o.) gleichermaßen zuverlässig läuft und dieselbe Darstellung des Nutzerinterface erfolgt; weiterhin werden folgende Komponenten benötigt:

- Teilnehmerverwaltung und Identitätsprüfung
  - Anmeldung zur Prüfungsteilnahme, Identitätsnachweis, Eröffnung eines Zeitslots für die Prüfungsdurchführung
  - Anmeldung zur Prüfungsdurchführung: Download und Installation der Test-Applikation (mit Safe-Exam-Browser), Download und Installation der Proctoring-Applikation, Einrichtung der beiden Geräte, Identitätsprüfung
  - Einwilligung hinsichtlich Daten und Datenschutz
- Online-Support, 24/7 bei maximaler angestrebter Terminflexibilität für die Testteilnehmenden
  - Anmelde- und Login-Probleme

- Probleme beim Herunterladen der Applikationen und der Einrichtung der Geräte
- technische Probleme während der Prüfung (z.B. Internetverbindungsprobleme, technische Fehler u.a. bei der Einspielung und Darstellung der Aufgaben bzw. einzelner Items)
- nutzerbezogene Probleme (z.B. versehentliches Löschen von Text, Navigationsschwierigkeiten)
- Prüfungsüberwachung während/nach der Prüfung: Proctoring-Tool
  - entweder menschlich oder maschinell (menschliches Proctoring ist deutlich teurer und aufwändiger als maschinelles Proctoring, dafür datenschutzrechtlich weniger komplex und aufwändig)
  - Live-Prüfung und Aufzeichnung
  - Entscheidung: Abbruch der Prüfung bei suspektem Verhalten oder “flagging” für post-hoc-Auswertung durch geschulte menschliche Proctors (mit dem entsprechenden Aufwand für die menschliche Auswertung der maschinell identifizierten Fälle potenziell suspekten Verhaltens)

In Bezug auf die Entwicklungs- und Durchführungskosten ist laut ExpertInnen (z.B. TestDaF) mit einem hohen Aufwand zu rechnen, wenn zwei Tests für denselben Zweck parallel eingesetzt werden (Papier, digital). Die Kosteneffizienz digitaler Tests ist abhängig von der Zahl der Testteilnehmenden (TestDaF). Die Konkurrenz auf dem Markt sei aktuell gross in Bezug auf digitale Tests; konkret arbeitet z.B. das Goethe-Institut aktuell an einer digitalen Remote-Variante des einwanderungs- und einbürgerungsrelevanten A1-Tests.

Die konkreten Kosten und Aufwände für die Entwicklungs- und Pilotierungsphase sind schwer zu schätzen, aber es dürfte die Einschätzung von Dr. Malone (ACTFL) gelten:

“it will cost you twice as much as you think, and it will take you three times as long.”

Im Folgenden vier exemplarische Aufwandsschätzungen (personeller und zeitlicher Aufwand) von digitalen Tests (ortsabhängig bzw. remote) für die Zertifizierung:

### 9.3.1 Beispiel E-TEF / DFP (digital vor Ort)

- Als Ausgangspunkt für die Entwicklung des E-TEF bestand bereits eine Vorab-Version auf CD-Rom.
- Von dieser CD-Rom-Variante ausgehend dauerte es 2 Jahre bis hin zu einer ersten digitalen Version, bei der das gesamte Datenpaket vorab heruntergeladen wird und lokal gespeichert, dann übertragen wurde.
- Beteiligt waren 2 EntwicklerInnen / InformatikerInnen (mit Erfahrung aus der Entwicklung der Variante auf CD-Rom); 1 Projektleitung IT, 3 ProjektleiterInnen, 4 Item-EntwicklerInnen.
- Ausgangspunkt für die Entwicklung des DFP (Diplôme de Français Professionnel) war wiederum der E-TEF.
- Pro berufsspezifische Domäne (z.B. Hotelfachleute, ...) dauerte die Entwicklung – mit einem ähnlichen Team – ca. 1 Jahr.

### 9.3.2 Beispiel TestDaF digital (vor Ort)

- Entwicklung / Designphase 2010–2016: Entwicklung von Aufgabentypen, die vom Testkonstrukt aus gesehen relevant schienen; Prüfung, ob digital gut, zuverlässig und vor allem eindeutig abbildbar; Begleitforschung: Erprobung in “kleineren” Gruppen mit 100–150 Personen
- ab 2016 Erprobungsphase des Gesamttests weltweit (mit Begleitforschung)
- Start Oktober 2020
- Standardsetting Juni 2021
- höchster Aufwand: Software-Entwicklung der Testumgebung → interne Entwicklungsabteilung (Vorteil: extern wäre sehr teuer; externer Anbieter macht lediglich die Umsetzung eines festen Produkts; interne EntwicklerInnen hingegen können in die Designphase selbst eingebunden werden, z.B. für die Erprobung verschiedener Darstellungsmöglichkeiten von Aufgaben)

- zusätzlicher Aufwand: Entwicklung der Testadministrationsumgebung

### 9.3.3 Beispiel ECL, Ungarn (vor Ort und remote)

- Ausgangspunkt war ein bestehendes Content-Management-System, das lizenziert werden konnte, es mussten lediglich ein paar zusätzliche Optionen entwickelt werden (z.B. Einbinden von Bildern)
- auch das Proctoring-Tool wird extern eingekauft (bezahlt wird dafür eine Summe von 4 Euro pro TestnehmerIn)
- Entwicklungsdauer: 1 Jahr, für Anpassungen, Entwicklung der Aufgabentypen und Items, Funktionstests etc.
- Team: 3 externe MitarbeiterInnen (IT) und 10 interne MitarbeiterInnen (Testentwicklung, Administration)
- Zwischenfazit: der Aufwand habe sich (noch) nicht gelohnt, denn
  - die Umstellung auf digital bringe viele (Anfangs-)Schwierigkeiten mit sich: Unsicherheit der Testteilnehmenden, Schulung der Aufsichtspersonen (digital, technisch), Ausstattung der Testzentren, technisches Support-Personal etc.
  - 90% der Testteilnehmenden wählen weiterhin die Papier-Version, obwohl 90% der dortigen Testteilnehmenden Gymnasiasten und Studierende sind, d.h. schriftsprachgewohnt und mit relativ hohen digitalen Kompetenzen;
  - die überwiegende Mehrheit der knapp 10% Testteilnehmenden, die eine digitale Prüfung wählen, wählen die Remote-Version;
  - die Nachfrage nach dem digitalen Test vor Ort ist also sehr klein (ca. 2%), wenn es Alternativen (Papier, Remote) gibt, aber die Umsetzung ist institutionsseitig sehr teuer (u.a. durch den Bedarf an Infrastruktur) und aufwändig (in Bezug auf Aufsicht, Support etc.).
- Daher wird vorläufig weiter auch die Papier-Version angeboten, das Testzentrum beobachtet die Entwicklungen

### 9.3.4 Beispiel Goethe-Institut (aktuell: vor Ort: geplant: remote)

#### Etappe 1: digitale B1-/B2-Prüfung vor Ort

- 2015 Start der Konzeptionalisierung der digitalen Prüfung vor Ort (B1, B2) und Entscheidung für Lizenzierung einer bestehenden Plattform statt eigene Entwicklung
- 2016 europaweite Ausschreibung, Auswahl des externen Anbieters für die Itemdatenbank und das Nutzerinterface sowie die Schnittstelle zu den bestehenden Testverwaltungstools
- 2017 Start der Programmierung, d.h. der Anpassung des lizenzierten Systems an die existierenden Goethe-Prüfungen und die Bedarfe
  - Dauer der Anpassungen für Darstellung und Ausspielung sowie Kompatibilität mit den anderen Systemen: ca. 1 Jahr
  - Sprechen: weiterhin face-to-face in Präsenz
  - Herausforderung: Back-up-Server für die Vermeidung von Datenverlusten bei Abbrüchen der Internetverbindung
- 2019, Februar: Start der digitalen B1- und B2-Prüfung an den ersten Testzentren; die einzelnen Testzentren entscheiden, ob sie (vorerst) weiterhin nur die Papier-Version anbieten, die digitale Version oder beide
- Team:
  - Grund-Team mit 2.5 Stellen (Projektleitung, im Rückblick gesehen hätte es aber ein größeres Team gebraucht, u.a. für den Aufwand an Kommunikation und Koordination)
  - interne IT-Abteilung (stark beansprucht)
  - interne Rechtsberatung (stark beansprucht)
  - diverse interne Item-EntwicklerInnen, inkl. KollegInnen im Ausland etc.
  - Personal des externen Anbieters des Content Management Systems, bei bereits bestehendem System, das nur angepasst wird,
  - externe Agentur für die Pilotierungen und die Usability-Studie

## Etappe 2: Remote-A1-Prüfung *Start Deutsch 1*

- 2021: europaweite Ausschreibung für die Remote-A1-Prüfung
- 2022: Start der Entwicklung der Remote-A1-Prüfung, wiederum mit externem Anbieter
  - Papier-Version soll parallel bestehen bleiben
  - Digitale Version soll wiederum eine 1-zu-1-Umsetzung des bestehenden Tests sein, da die bestehende Prüfung anerkannt ist, es zu aufwändig wäre, eine neue, abweichende digitale Version erneut anerkennen zu lassen
  - Papier-Version vor Ort sei auch stark präferiert für lernungewohnte Testteilnehmende, gegenüber einer Remote-Durchführung alleine zuhause, ohne Betreuung, Support, Erklärungen, Infrastruktur
  - Problem: Online-Support; wie/in welcher Sprache erklärt man den Testteilnehmenden, was das Problem ist, was sie wann wie tun müssen?
- geschätzter Zeithorizont der Entwicklung, auf Basis der vorliegenden digitalen B1-/B2-Prüfung vor Ort: unter 5 Jahre
- Team: deutlich grösser als für die digitale B1-/B2-Prüfung
- Rentabilität erreichbar, ...
  - wenn nur noch die digitale Version angeboten wird, nicht parallel eine Papier- und eine digitale Version
  - bei konstant hohen Teilnehmendenzahlen (bei einer Testteilnehmendenzahl von ca. 20'000 pro Jahr könne das „eine ganze Weile“ dauern, wenn sich die Testgebühren für die Teilnehmenden nicht erhöhen sollten)

Unsere schriftliche Anfrage bei ETS (Educational Testing Service, Entwickler TOEFL-Tests), mit welchen Kosten zu rechnen wäre, sollte ETS die Umsetzung der Digitalisierung des fide-Tests übernehmen, führte zu einem klaren negativen Bescheid. Offenbar schätzt ETS den Aufwand als so gross ein, dass es sich für sie nicht einmal lohnt, ein detailliertes Angebot zu erstellen:

“I’ve been working with colleagues across the organization as well as one of our subsidiary companies for the past week to determine how we can best meet your needs and provide a quote to digitize the Swiss Language Test for Immigration Purposes. Unfortunately, we don’t have the capacity to support such an effort currently and therefore cannot provide a quote, sorry for disappointing news.”

Allein für die Designphase, also die inhaltliche Entwicklung und Erprobung der Aufgabentypen und Aufgabenitems, schätzt die ATFL-Expertin den Aufwand für 50 neue Items für bestehende Aufgabenformate folgendermassen ein: 3 Wochen bei einem Team von 2 Item-EntwicklerInnen, 1 ProjektmanagerIn, 1 Leitung, 1 ReviewerIn, mehrere SoftwareingenieurInnen, 1 KoordinatorIn; danach folgen mehrere Schleifen von Review durch erfahrene Prüfende (mit einer detaillierten Check-Liste); es ist allerdings unklar, in welchem Verhältnis diese Aufwandsschätzung zum Aufwand für das fide-Test-Format steht, in dem ja nicht einzelne Aufgaben-Items entwickelt werden müssen, sondern ganze Aufgaben bzw. Module, die zu Sets kombiniert werden (s. Fussnote 1).

Insgesamt berichtet die ACTFL-Expertin über die Notwendigkeit einer ganzen Serie von Pilotierungen bzgl. Der Gesamtusability und einzelner Items. ACTFL setzt hier Zoom-Testungen mit menschlicher Beobachtung und Aufnahme sowie folgende Methoden ein:

- *Think-aloud*-Protokolle während der Testbearbeitung (Wie gehen Sie vor? Wo stossen Sie auf Schwierigkeiten?)
- *Stimulated recalls* auf Basis der Aufzeichnungen: Was war in dem Moment schwierig? Warum? Was bräuchte es, um das zu lösen? (z.B. “I didn’t know how or where to respond” → “would a big yellow arrow help?”)

Eine solche Usability-Begleitforschung wäre gerade für die fide-Kernzielgruppe schulungsgewohnter Testteilnehmender zentral.

Zusätzlich müsste durch Begleitforschung nachgewiesen werden, dass die Papier- und die digitale Version gleichwertig sind, d.h. die gleichen Kompetenzen messen, die gleichen Anforderungen stellen; darauf weist auch die ACTFL-Expertin hin:

“the first thing I would do would be to make sure that any digital test was yielding similar results to the paper test”

#### 9.4 Digitaler ortsunabhängiger Test: Aufwand / Kosten in der Durchführung

In der Durchführung entstehen bei Remote-Tests zusätzliche Kosten für maschinelles oder menschliches Proctoring:

- Der Aufwand für menschliches Proctoring bemisst sich in der Arbeitszeit der geschulten (!) Aufsichtspersonen; bei einer 1-zu-1-Aufsicht während des gesamten Prüfungsablaufs inkl. Identitätsprüfung wäre der personelle Aufwand erheblich.
- Der Preis für maschinelles bzw. menschliches Live-Proctoring unterscheidet sich je nach Anbieter stark, z.B. beläuft sich der Preis pro Testnehmeur bestehenderIn in Pécs auf 4 Euro, hingegen ruft MeazureLearning (ProctorU) folgende Preise auf:
  - 10–12\$ pro Identitätsprüfung
  - 20\$ für die erste Stunde Live-Proctoring, 10\$ für jede weitere angebrochene Stunde
- Bei maschinellm Proctoring kommt die menschliche Post-hoc-Auswertung der in den Aufzeichnungen “geflaggt” Anzeichen / Momente suspekten Verhaltens hinzu.

Angesichts des oben beschriebenen hohen Aufwandes für die Klärung der datenschutzrechtlichen Aspekte für das Proctoring, die Speicherung und Verwendung biometrischer Daten, ist davon auszugehen, dass sich hier eine nicht unerhebliche Kostenquelle während der Entwicklungsphase ergibt, angesichts der potenziellen dauernden Änderungen in der Rechtslage, aber auch nach Abschluss der Entwicklungsphase im laufenden Betrieb eine rechtliche Begleitung durch eine/n DatenschutzberaterIn notwendig ist.

In Bezug auf das regelmässige Update der Item-Datenbank ist damit zu rechnen, dass im Vergleich zur digitalen Durchführung vor Ort eher häufiger / mehr neue Items entwickelt, erprobt und auf Gleichwertigkeit geprüft werden müssen, da bei der Durchführung zuhause trotz aller Vorsichtsmassnahmen ggf. doch ab und zu Aufgabenitems veröffentlicht werden.

#### 9.5 Digitaler ortsunabhängiger Test: Sparpotenziale

Dafür entstehen Sparpotenzial und Benefits an folgenden Stellen:

- siehe oben: Druck, Versand, Korrektur, Administration, Item-Entwicklung
- Infrastruktur: Räume, Hardware (bestehen bleiben jedoch die Kosten für den Server-Speicherplatz, ggf. auch in Drittländern, s.o.)
- Anreisen (Prüfende, Testteilnehmende)

Insgesamt ergeben sich durch eine digitale Version, ob vor Ort oder remote, folgende weitere Vorteile für die Testentwicklung, aber auch für die Forschung:

- kostengünstige, effiziente Erprobung neuer Aufgaben-Items im laufenden Betrieb, als Zusatzaufgaben in den regulären Prüfungen (Eckes, TestDaF; Casanova, ALTE)
- bei Durchführung vor Ort weniger Täuschungsmöglichkeiten durch verschiedene individuelle Testversionen (statt dieselbe Papierversion für alle in einer Klasse) (Casanova, ALTE)
- unkomplizierte Möglichkeit der Korpuserstellung, d.h. der Sammlung grosser Mengen von Lernerdaten u.a. für die Spracherwerbsforschung, aber auch für die Entwicklung automatischer Bewertungsoptionen beim Schreiben (und Sprechen) (Stichwort *machine learning*; vgl. Casanova, ALTE)

## 10. Handlungsoptionen, Spielräume und Szenarien

Nach der Darstellung der technischen Möglichkeiten bezüglich Testaufgaben, Durchführungs- und Täuschungssicherheit sowie der Reflexion über Potentiale und Herausforderungen hinsichtlich des Einsatzes digitaler Tests im sensiblen Kontext der Integration werden im Folgenden Handlungsoptionen zusammengefasst und geprüft. Bei der Evaluierung der Optionen steht im Vordergrund das vom Auftraggeber verfolgte Ziel der Angebotsoptimierung für die Gesamtheit der MigrantInnen, die für ausländer- und bürgerrechtliche Verfahren ihre sprachlichen Kompetenzen nachweisen müssen. Weitere wichtige Aspekte bei der Entwicklung und Evaluierung der Szenarien sind die Sicherstellung der Durchführbarkeit auch in Krisensituationen (vgl. Lockdown in der Corona-Pandemie) und die Gewährleistung der Zugänglichkeit für die schulungsgewohnten Teilnehmenden.

Die hier entwickelten Szenarien berücksichtigen nicht nur die aktuelle Situation, sondern auch aufgrund der digitalen Transformation erwartbare Entwicklungen. PrüfungsexpertInnen gehen davon aus, dass digitales Prüfen zur Normalität werden wird. Die Digitalisierung hat aber nicht nur einen Einfluss auf das Prüfungsformat, sondern auch auf die kommunikativen Situationen im Alltag, die wiederum in den Sprachtests abgebildet werden sollen. Es wird deshalb eine Veränderung des Testkonstrukts, der Aufgabenformate und des Testens der Fertigkeiten erwartet. Multimodalität und Komplexität der kommunikativen Situationen sprechen – wie bereits im Begleitband des Gemeinsamen Referenzrahmens reflektiert – eher für ein integriertes Testen der Fertigkeiten (z.B. Lesen und Schreiben; Hören und Sprechen) und entsprechend für neue multimodale Aufgabentypen.

Darüber hinaus werden bei der Evaluierung der Szenarien die erwarteten technologischen Entwicklungen (inkl. *Machine Learning*) reflektiert. Sie bringen weitere Möglichkeiten hinsichtlich der Gestaltung des Interface, der Navigation, der automatischen Korrektur etc. mit sich.

### 10.1 Tools und Plattform/en

#### 10.1.1 Content Management System (Item-Datenbank, Nutzerinterface) und Administration

Im Prinzip gibt es für diesen aufwändigen Entwicklungsaspekt zwei grundlegende Möglichkeiten: Es kann entweder eine bestehende Plattform "eingekauft" werden (d.h. lizenziert, ggf. angepasst, mit einem entsprechenden Service-Paket). Es kann alternativ die entsprechende Plattform selbst entwickelt und programmiert werden (mit den entsprechenden Ansprüchen an Pflege, Update etc.).

Eine Lizenz-Variante wäre gerade bei hoher angestrebter Komplexität des Tools mit hoher Wahrscheinlichkeit (dies ist allerdings abhängig von der Anzahl der pro Jahr durchgeführten Prüfungen und von der Nachhaltigkeit der Plattform) in einer ersten Phase schneller (s. Goethe-Institut) und günstiger und könnte je nach Vertrag ein grundlegendes Service-Paket enthalten (z.B. Anpassungen, Support). Dabei reicht die Bandbreite möglicher Anbieter (und der erwartbaren Preise) potenziell von etablierten SprachprüfungsanbieterInnen über kostengünstige Lernplattformen wie Moodle bis hin zu Anbietern von Videospiel-Plattformen; auch die Wahl eines Anbieters im In- oder Ausland dürfte Konsequenzen für die Grössenordnung der entstehenden Kosten haben.

Erwartbar wären allerdings auch einige potenzielle Nachteile, z.B. in Bezug auf den Datenschutz und Anpassungsmöglichkeiten für bestimmte Aufgabentypen. Vor allem müsste gut abgewogen werden, inwiefern langfristig ungünstige Abhängigkeiten entstehen könnten, z.B. bei den Möglichkeiten der Aktualisierung der Item-Datenbank (Zugang, Einpflegen von Items etc.) oder in Bezug auf Preissteigerungen. Beachtet werden muss, dass es voraussichtlich problematisch wäre, sich zuerst für eine Lizenzierung und dann zu einem späteren Zeitpunkt für eine eigene Entwicklung zu entscheiden, wenn z.B. die Zahlen der Testteilnehmenden steigen, da das durch Lizenzierung (und Anpassung) Erreichte nicht unbedingt „mitgenommen“ werden könnte.

Im Falle einer eigenen Entwicklung müsste sichergestellt werden, dass das (interne) Team ExpertInnen aus den Bereichen Prüfen / Testen und aus der Informatik umfasst, so dass theoretische Konstrukte und Aspekte direkt in Bezug auf ihre Umsetzbarkeit geprüft und das sich entwickelnde CMS jeweils schnell angepasst werden kann. In jedem Fall wäre eine Entwicklung/Lizenzierung einer gemeinsamen Plattform

(Item-Datenbank, Nutzerinterface, Navigation, Support, Aufgabentypen, ...) nötig, die allen Sprachversionen zugrunde liegt; das Entwickeln und Einpflegen der spezifischen Items zu den einzelnen Aufgaben müsste dann sprachspezifisch stattfinden.

### 10.1.2 Proctoring

Auch in Bezug auf das Proctoring gibt es die oben genannten zwei grundsätzlichen Möglichkeiten, d.h. einerseits die Lizenzierung einer bestehenden, zuverlässigen Applikation (für sehr unterschiedliche Preise pro Prüfung, je nach Anbieter, s.o.) mit einem bestimmten Service-Paket (Support vor und während der Prüfung etc.), andererseits eine eigene Entwicklung / Programmierung.

Der erwartete Hauptvorteil einer Lizenzierung wäre die Schnelligkeit, mit der eine Lösung zur Verfügung stehen würde (ggf. mit verhandelbaren Varianten und Anpassungen). Zudem könnte auch hier mit einem Service-Paket gerechnet werden, bis zu einer bestimmten Anzahl von Teilnehmenden pro Jahr wäre die Lösung zu Beginn wohl günstiger.

Mögliche Nachteile sind auch im Hinblick auf diese Komponente die Einschränkungen in Bezug auf spezifische Aspekte des Datenschutzes (z.B. Welche Daten werden erhoben? Wo werden die Daten gespeichert?) und der technischen und damit auch finanziellen längerfristigen Abhängigkeit.

## 10.2 Szenarien (1): Papier weiter als Option

### 10.2.1 Szenario 1: Nur Papierversion

Das Szenario sieht (im Moment) keine Digitalisierung des fide-Tests vor. Dabei steht allerdings angesichts der aktuellen Entwicklungen im Lernen, Lehren und Prüfen die Nachhaltigkeit einer solchen Entscheidung in Frage: Nach Einschätzung der befragten ExpertInnen verändern sich die Bedürfnisse der Testteilnehmenden, u.a. Terminflexibilität und Ortsunabhängigkeit werden von Testanbietern zunehmend erwartet. Nach den Erfahrungen der Pandemie wird zudem der Aspekt der Durchführungssicherheit ebenfalls immer wichtiger (z.B. für Behörden, Arbeitgeber).

Auf diese veränderten Bedürfnisse reagieren die grossen Testanbieter bereits. TestDaF beispielsweise prüft aktuell die Möglichkeit, auch eine (digitale) A2-Prüfung zu entwickeln; das Goethe-Institut arbeitet an einer digitalen Remote-Version der besonders nachgefragten einwanderungsrelevanten A1-Prüfung, die international absolvierbar sein soll (und kann dabei sowohl auf die bestehende digitale B1-/B2-Prüfung als auch auf das bestehende Netzwerk von Testzentren zurückgreifen); grössere internationale Anbieter verfügen über Plattformen, die ihnen die schnelle Entwicklung digitaler Tests in mehreren Sprachen erlauben könnten. Diese Marktentwicklungen werfen die Frage der Konkurrenzfähigkeit des fide-Tests ohne digitale Version auf. Die Goethe-Expertinnen werfen ganz explizit die Frage auf, ob eine papierbasierte Testversion in 10 Jahren nicht unattraktiv bzw. unprofessionell wirken werde.

Ferner bleibt in diesem Szenario das Thema der Durchführungssicherheit ungelöst. Wenn zukünftig weitere oder sogar längere Lockdowns vorkommen sollten, sind zwei Szenarien möglich:

1. eine Gesetzesänderung, welche die Pflicht für den Sprachnachweis für Aufenthalt/Einbürgerung unter Lockdown aussetzt
2. eine Gesetzesänderung, welche das Absolvieren einer Sprachprüfung für Aufenthalts-/Einbürgerungszwecke mit einem besonderen Schutzkonzept auch unter Lockdown-Bedingungen erlaubt (als Grundbedürfnis)
3. eine kurzfristige ad-hoc-Digitalisierung der bestehenden Papier-Version, die möglicherweise aufgrund des Zeitdrucks bei der Entwicklung gerade die sensible Kernzielgruppe des fide-Tests benachteiligen könnte

Aufgrund der hier genannten datengestützten Überlegungen gehen wir davon aus, dass die Digitalisierung des fide-Tests höchstens aufgeschoben werden kann. Das Aufschieben hiesse dabei eher ein paralleles Aufgleisen der digitalen Version, das die Zeit des Aufschubs nutzt, um

- a. die vorhandene Papier-Version zu optimieren und international anerkennen zu lassen
- b. Gespräche mit den unterschiedlichen Stakeholdern weiterzuführen und sie früh genug in die Entwicklungen einzubeziehen

- c. die Entwicklungen hinsichtlich des digitalen Prüfens und des Prüfungsmarkts besser einschätzen zu können

Dieses Abwarten könnte auch beinhalten, dass der fide-Test zum Beispiel keine Plattformen und/oder digitale Aufgabentypen für sich entwickelt, sondern bestehende auf ihre Zuverlässigkeit und Eignung prüft sowie bereits die notwendigen Anpassungen skizziert und den Umfang des Anpassungsbedarfs evaluiert (s.o.).

### 10.2.2 Szenario 2: Parallel-Versionen Papier und digital vor Ort

Im Szenario zwei wird der fide-Test weiterhin auf Papier und zudem digital vor Ort angeboten. Folgende Bedingungen sollten dabei gelten:

- a. Die Papier- und die digitale Version sind gleichwertig. Dies ist nachgewiesen durch Begleitforschung und entsprechend dokumentiert. Beiden Versionen liegt das gleiche Testkonstrukt zugrunde; eine Berücksichtigung von digitalen Kompetenzen im Testkonstrukt wäre dadurch ausgeschlossen.  
Durch die Entwicklung der digitalen ortsabhängigen Version könnten sich umgekehrt ggf. Änderungen in der Papier-Version ergeben (Teststruktur, Aufgabentypen). Deshalb empfehlen wir in diesem Fall nicht, wie von einigen ExpertInnen vorgeschlagen, zunächst die Optimierung der Papier-Version abzuschliessen. Auch die Goethe-Expertinnen weisen explizit darauf hin, dass eine 1-zu-1-Umsetzung papierbasierter Tests im digitalen Medium lediglich einen ersten Schritt in der digitalen Prüfungsentwicklung darstellen kann, dass letztlich nach und nach die spezifischen Möglichkeiten und Formate des Digitalen genutzt werden müssen.
- b. Der fide-Test berücksichtigt weiterhin und explizit die Zielgruppe der schulungsgewohnten MigrantInnen. Das heisst, es werden die technologischen Möglichkeiten genutzt, um das Nutzerinterface, die Aufgaben sowie die Navigation so zu gestalten, dass die Niederschwelligkeit weiterhin gewährleistet ist. Die Anweisungen sind in allen Prüfungsteilen, insbesondere aber im mündlichen Teil (Hören und Sprechen), schriftfrei bzw. multimodal (Videos, Audiodateien). Erklärvideos unterstützen die Teilnehmenden bei der Prüfungsvorbereitung (und ggf. während der Prüfung), digitale Modellprüfungen und entsprechende Prüfungstrainings stehen zur Verfügung. Zudem besteht für den schriftlichen Teil die Option, den Text entweder mit einer Art von SmartPen handschriftlich auf einem Touchscreen zu verfassen oder handschriftlich auf Papier (mit anschliessendem Einscannen durch die Aufsichtsperson).
- c. Die Stakeholder sind adäquat informiert und zielführend in die Entwicklungen einbezogen.

Da sowohl die Papier- als auch die digitale Vor-Ort-Version des fide-Tests in Testzentren abgewickelt wird, verändern sich die Rolle und Zuständigkeiten der Geschäftsstelle fide und der Test-Anbieter nur geringfügig. Dennoch entsteht – über die Testentwicklung und Klärung der juristischen Fragen hinaus – ein wichtiger Aufwand, der mit den Kosten der Infrastruktur (Hardware, Software, Server) und ihrer Wartung verbunden ist. Darüber hinaus sind Schulungen für PrüferInnen und Aufsichtspersonen, welche die digitale Version vor Ort abnehmen, ebenfalls aufwändig, aber notwendig, genauso die Verfügbarkeit des Technik-Supports vor Ort.

Die Abwicklung der zwei Versionen in Testzentren kann weder die Durchführungssicherheit gewährleisten noch die Ortsunabhängigkeit und die Terminflexibilität. Damit verbunden sind die im Szenario 1 bereits dargestellten Problematiken: Nicht-Berücksichtigung der Bedürfnisse der Testteilnehmenden und der Stakeholder, Konkurrenzfähigkeit, und schliesslich das Risiko, im Falle eines längeren Lockdowns eine digitale Version einzusetzen, die wenig durchdacht ist.

Zudem ist es fraglich, ob es eine Nachfrage nach digitalen Prüfungen vor Ort gibt bzw. wie hoch diese ist. Schulungsgewohnte Lernende ziehen aktuell die Papier-Version vor, schulungsgewohnte Lernende schätzen die Ortsunabhängigkeit und Terminflexibilität, die eine Vor-Ort-Prüfung nicht anbietet. In Testzentren wählen auch schulungsgewohnte Lernende die Papier Version aus, so die Erfahrung des ungarischen Testanbieters. Um dieser Unsicherheit entgegenzuwirken, wäre es denkbar, die digitale Vor-Ort-Version zur Standardversion zu benennen. Die Papier-Version könnte nur auf Antrag der Testteilnehmenden und/oder in einer Übergangszeit abgelegt werden. Vorstellbar wäre ggf. sogar eine Teilnahme an der digitalen Prüfung mit der

Unterstützung einer prüfenden Person, welche die Navigation übernehmen könnte (Malone/ACTFL; siehe dazu aber die Bedenken von TestDaF).

Der Aufwand für die Entwicklung der digitalen Vor-Ort-Version ist beträchtlich, dazu kommt der Aufwand für die Aktualisierung und laufende Optimierung der zwei Versionen. Der Aufwand für die Parallelführung einer Papierversion und einer digitalen Version wird als sehr hoch eingeschätzt (TestDaF, Goethe-Institut).

Dennoch wäre diese digitale ortsabhängige Version der erste Schritt für die Platzierung am digitalen Prüfungsmarkt. Weitere Vorteile wären die Nutzung der Technologien für eine bessere Berücksichtigung von Menschen mit Beeinträchtigung sowie die Integration der automatischen Korrektur, zumindest für die geschlossenen Items.

Ob dieses Szenario nachhaltig genug ist, hängt u.a. mit der Antwort auf die Frage zusammen, wie schnell die digitale Vor-Ort-Version eingesetzt werden kann. Sollte die Erarbeitung der digitalen ortsabhängigen Prüfung mehr als drei Jahre in Anspruch nehmen, gehen wir davon aus, dass die Entwicklungen im Kontext des digitalen Prüfens bis dahin so weit sind, dass die digitale Version vor Ort keine innovative Lösung darstellt. Die Goethe-Expertinnen gehen davon aus, dass für die Goethe-Prüfungen die Papier-Version noch einige Jahre lang neben der digitalen Version existieren wird, aber die längerfristige Entwicklung für den Zeithorizont ab 2040 sei unklar und stark abhängig einerseits von der Marktentwicklung, andererseits von weiteren Entwicklungen der Unterrichtsformen (und damit verbunden auch der Entwicklung der digitalen Kompetenzen der Testteilnehmenden).

### 10.2.3 Szenario 3: Parallel-Versionen Papier und digital remote

Im Szenario drei sind eine Papier-Version und eine digitale Remote-Version vorgesehen. Auch in diesem Szenario sollten die drei oben genannten Bedingungen erfüllt werden: Gleichwertigkeit der zwei Versionen, Niederschwelligkeit, Einbezug der Stakeholder.

Eine Remote-Variante ermöglicht generell

- eine Prüfungsdurchführung auch in besonderen, lockdown-ähnlichen Situationen
- eine Prüfungsdurchführung ohne Anreisaufwand
- eine maximale, von vielen Testteilnehmenden zunehmend gewünschte bzw. erwartete Terminflexibilität, in Bezug auf die Kurzfristigkeit der Anmeldung und die Durchführungszeiten (reduziert, falls die Komponente "Sprechen" als Video-Call mit 1-2 Prüfenden durchgeführt wird, s.u.).

In diesem Szenario ist daher die Durchführungssicherheit gewährleistet, die veränderten Bedürfnisse der Zielgruppe hinsichtlich Terminflexibilität und Ortsunabhängigkeit sind erfüllt, der fide-Test ist auf dem Prüfungsmarkt für eine breite Zielgruppe attraktiv und konkurrenzfähig.

Kostenersparnisse ergeben sich aus der automatischen Korrektur von geschlossenen Items in der digitalen Version. Gegebenenfalls ist längerfristig auch mit einer weiteren Kostenreduktion zu rechnen, wenn die Nachfrage für die Remote-Version über eine längere Zeit gross genug ist, denn der Initialaufwand für eine digitale ortsunabhängige Version wird zumindest so hoch eingeschätzt wie für die digitale Vor-Ort-Version. Hier fallen im Vergleich zur digitalen Vor-Ort-Version zwar Teile der Kosten der Infrastruktur für die Ausstattung der Testzentren weg (hingegen müssen Räume, Anreisen etc. für die Papier-Version weiter vorgesehen werden, auch die Server-Infrastruktur ist notwendig), aber die Entwicklung von Item-Datenbank, Nutzerinterface, Administrationssystem etc. sowie vor allem des Proctoring-Systems sowie die rechtlichen Abklärungen hinsichtlich Datenschutzes sind mit grossem zeitlichem und finanziellen Aufwand verbunden.

Die Variante "digital remote" verschiebt allerdings das Problem der Infrastruktur auf die Testteilnehmenden, die mindestens verfügen müssen über

- zwei reibungslos funktionierende Geräte (PC, Laptop oder Tablet für die Durchführung, Smartphone fürs Proctoring)
- ein Headset
- einen ruhigen Ort
- eine sichere, stabile Internetverbindung.

Zusätzlich beinhaltet die Remote-Variante nach Einschätzung u.a. der befragten Integrationsbeauftragten und der Lernenden selbst, aber auch aus Sicht verschiedener TestexpertInnen, die grössten Störquellen in Bezug auf das Testkonstrukt und die sozio-kognitive Validität des Tests, da neben sprachlichen Kompetenzen potenziell auch immer gewisse andere Grundfertigkeiten und *digital skills* mit erfasst werden (u.a. Navigation, Zurechtfinden in einer komplexen digitalen Umgebung, Tastatur-Schreiben, s.o. Kap. 7.1). Nach Einschätzung der Befragten betrifft das digitale Remote-Prüfungen, wo der/die Testteilnehmende alleine zuhause vor dem Bildschirm sitzt, potenziell stärker als digitale Vor-Ort-Tests, wo zumindest das aufwändige Prozedere für die Identitätsprüfung (Anmeldung, Login etc., s.u.) sowie die Notwendigkeit zum eigenständigen Herunterladen und Installieren verschiedener Applikationen entfallen – und wo der Vorstellung nach im Zweifelsfall auch eine Aufsichtsperson bei gravierenden technischen Problemen direkt helfen könnte.

Eine digitale (Remote-)Prüfung kann zudem einen zusätzlichen Stressfaktor für Teilgruppen der Testteilnehmenden darstellen (s.o. Kap. 7.4).

Auch kann es im Remote-Fall herausfordernder sein, gerade auf niedrigen Kompetenzniveaus und für schulungsgewohnte Lernende, in verständlicher Sprache Online-Support bezüglich der Navigation (Wo bin ich? Wie komme ich zurück/nach vorn zur nächsten Aufgabe?), technischer Probleme oder der Anforderungen des Proctoring (Anmeldung, Login, Identitätsprüfung, Installation des Mobiltelefons mit der zweiten Kamera etc.) zu bieten. Die Goethe-Expertinnen weisen darauf hin, dass erst ab der B1-Stufe sei mit hoher Wahrscheinlichkeit davon auszugehen sei, dass die Testteilnehmenden über ausreichende Deutschkenntnisse verfügten, um komplexe technische Erklärungen zum Set-up (Vorgehen bei Login, Identitätsprüfung etc.), zur Navigation und zum Umgang mit potenziellen Problemen zu verstehen. Gerade auch bei schulungsgewohnten Lernenden ist nicht davon auszugehen, dass sie beispielsweise das Englische als Mittlersprache ausreichend gut beherrschen, dass Erklär-Videos auf Englisch zielführend wären. Es ist durchaus zu erwarten, dass diese Problematik auch für niedrigere Kompetenzstufen gelöst werden kann; in Betracht gezogen werden könnten z.B. Untertitelungen oder Synchronisierungen der Erklärvideos in der Erstsprache der Testteilnehmenden; dies zumindest für ausgewählte grössere Migrationssprachen.

In diesem Szenario ist insgesamt anzunehmen, dass die Remote-Version in der Praxis eher die schulungsgewohnten MigrantInnen anspricht und die Papier-Version weiterhin die Hauptoption für die schulungsgewohnten MigrantInnen bleibt. Die hier befragten Lernenden der fide-Kernzielgruppe schrift- und schulungsgewohnter TestnehmerInnen, denen zentrale digitale Kompetenzen fehlen, äussern eine starke Präferenz für die Durchführung vor Ort (gegenüber der Remote-Variante), wenn es keine Papierversion mehr geben sollte; dies vor allem aufgrund der Durchführung in Anwesenheit einer Aufsichtsperson, die ggf. bei technischen oder Navigationsproblemen eingreifen könnte. Die entsprechende Präferenz für digital vor Ort bzw. digital remote scheint abhängig einerseits von den digitalen Kompetenzen der Testteilnehmenden (bzw. deren diesbezüglicher Selbsteinschätzung) und andererseits von der individuellen Familien- und Wohnsituation (Störpotenzial durch Kinder, Verfügbarkeit eines ruhigen Raumes etc.).

Das könnte heissen, dass bei neuen Lockdown-Situationen gerade die aktuelle Kernzielgruppe des fide-Tests indirekt benachteiligt wäre; nicht von der Prüfung selbst, denn diese sollte die Bedingung der Niederschwelligkeit erfüllen, aber weil schulungsgewohnte MigrantInnen nicht gut genug bzw. rechtzeitig genug auf eine digital Remote-Prüfung vorbereitet wären.

Was das Proctoring betrifft, muss im Rahmen dieses Szenarios früh entschieden werden, ob dieses automatisch (maschinelles Proctoring) oder durch Online-Aufsichtspersonen durchgeführt werden soll (menschliches Proctoring). Im Falle vor allem maschinellen Proctorings müssen Datenspeicherungs- und Datenschutz-Aspekte von Beginn an geklärt und berücksichtigt werden. Die Klärung dieser datenschutzrechtlichen Aspekte nimmt erfahrungsgemäss (s. Cambridge Assessment English) viel Zeit in Anspruch und hat Konsequenzen für viele spätere Planungsetappen und Handlungsoptionen; sie muss daher vom frühestmöglichen Zeitpunkt an mitbedacht und von ExpertInnen begleitet werden.

Insgesamt stellt eine Remote-Variante institutionsseitig hohe Anforderungen an Datensicherheit und Datenschutz, u.a. im Hinblick auf

- die Verhinderung von Darstellungsproblemen (z.B. systembedingtes Überspringen von Aufgaben-items)

- die Verhinderung von Datenverlust durch Verbindungsabbrüche (z.B. Optionen für Backup-Server bzw. Back-up-Version)
- den Schutz gegen unerlaubten Zugriff auf sensible Daten (s.u. biometrische Daten)
- den Datenschutz im Falle der Verarbeitung besonders sensibler, biometrischer Daten (s.u. Identitätsprüfung, Proctoring mit künstlicher Intelligenz)

#### 10.2.4 Szenario 4: Parallel-Versionen Papier, digital vor Ort und digital remote

Im Szenario 4 wird der fide-Test auf Papier, digital in Testzentren und digital remote angeboten. Auch für dieses Szenario gelten die oben genannten Bedingungen der Gleichwertigkeit der drei Versionen, der Niederschwelligkeit, und der Involvierung der Stakeholder.

In diesem Szenario stellt sich wie im Szenario 2 sich die Frage, ob die digitale Vor-Ort-Version, die mit hohen Kosten für die Infrastruktur verbunden ist, überhaupt nachgefragt wird. Zur Beantwortung dieser Frage empfehlen wir eine grössere NutzerInnen-Umfrage, um einschätzen zu können, welcher Anteil auf die verschiedenen Formate entfallen würde.

### 10.3 Szenarien (2): Ausschliesslich digitales Angebot

#### 10.3.1 Szenario 5: Nur digitale Vor-Ort-Version

Im Szenario fünf existiert der fide-Test digital und wird in Testzentren absolviert. Als Voraussetzungen gelten weiterhin die Bedingungen der Zugänglichkeit für die schulungsgewohnten MigrantInnen sowie des Einbezugs der Stakeholder bei der Entwicklung der neuen Version.

Für den Prüfungseinstieg (Willkommen, "Entstressen" und grundlegende Erklärungen) und den Teil "Sprechen", auf dem die Adaptivität des fide-Tests basiert, sehen wir drei Möglichkeiten:

- a. Beides findet wie in der aktuellen Papier-Version als persönliches Gespräch statt.
- b. Der Teil Sprechen findet in einem Video-Call statt, in dem Prüfende die gleiche Rolle übernehmen wie in der aktuellen Papier-Version. Dabei muss die Übertragungs- und Aufzeichnungsqualität gewährleistet werden.

Auch schulungsgewohnte Lernende sind mit Video-Calls vertraut, die sie häufig im privaten Kontext nutzen, um mit Freunden und Familie im Ausland zu kommunizieren. Neu werden Video-Calls ebenfalls im Kontext des Lehrens und Lernens (Sprachkurse, RAV-Kurse, Elternkurse) immer häufiger eingesetzt.

Als Nachteil dieser Option dürfen die mangelnde Terminflexibilität sowie die weiterhin hohen Personalkosten gelten.

- c. Der Teil Sprechen wird neu konzipiert und besteht aus Stimuli, welche die Testteilnehmenden als Hör- oder Lesetexte angeboten bekommen und auf die sie am Computer/Tablet reagieren, indem sie ein Wort, eine Äusserung oder einen längeren Text produzieren. Diese Reaktionen auf die Stimuli, d.h. die Lernerproduktionen, werden aufgenommen und anschliessend durch Prüfende bewertet. Möglichkeiten der automatischen Korrektur wären hier zu prüfen, wir gehen aber davon aus, dass es sehr schwierig ist, auf den tieferen Niveaus eine zufriedenstellende Messgenauigkeit zu erreichen.

In diesem Fall kann die Adaptivität des fide-Tests nicht mehr auf dem Teil Sprechen basieren. Ausserdem verschiebt sich der Fokus des Konstrukts von der Interaktion weg auf ein eher monologisches Sprechen. Dieses monologische Sprechen kann aber auch weiterhin lebensnah gestaltet werden (z.B. Sprachnachrichten via WhatsApp; Vorstellungsvideo im Rahmen eines Bewerbungsgesprächs; Darstellung von Symptomen beim Arztbesuch oder in der Apotheke).

Für die Gestaltung und Entwicklung der Teile Lesen und Hören sind die Herausforderungen kleiner. Die Frage, die sich hier vor allem stellt, ist, welche neue Aufgabentypen gut die heutzutage üblichen multimodalen kommunikativen Situationen abbilden können. Man könnte hier noch mehr auf ein integriertes Testen der Fertigkeiten setzen. Zum Beispiel könnte eine Annonce gelesen werden, auf die die Testteilnehmenden

mit einer Sprachnachricht reagieren (Lesen und Sprechen). Oder es können, ggf. in einer Annäherung an eine *virtual reality*, in einem etablierten situativen Kontext (Szenario) Fragen gestellt werden (z.B. Simulation eines Gesprächs am Bank-/Postschalter), auf die man mündlich reagiert und dabei die eigenen Sprachproduktionen aufnimmt (Hören und Sprechen).

Für den Teil Schreiben ist es, wie oben bereits erwähnt, für eine Teilgruppe der Testteilnehmenden wichtig, dass weiterhin die Möglichkeit besteht, den Text auf Papier zu verfassen. Durch die Digitalisierung steigt aber an sich potenziell die Authentizität der Schreibaufgaben bzw. des Tests, wenn die Prompts sich auf digitale Textsorten beziehen.

Bei einer digitalen Vor-Ort-Version muss frühzeitig die Frage des Mediums geklärt werden. Schulungsgewohnte Lernende zeigen eine Präferenz für das Smartphone als Testmedium. Aufgrund der Bildschirmgröße können dabei aber das Schreiben und auch die Evaluierung des Leseverstehens problematisch werden. Tablets funktionieren ähnlich wie Smartphones, verfügen aber über grössere Bildschirme und könnten eine gute Kompromisslösung darstellen. Die meisten schulungsgewohnten Lernenden geben an, mit Tablets vertraut zu sein. Häufig haben ihre Kinder solche Geräte. Hingegen werden Computer und Laptops von den schulungsgewohnten Teilnehmenden eher abgelehnt; die Navigation mit der Maus könnte für manche eine Herausforderung sein, vor allem aber behindert das Schreiben mit der Tastatur die Schreibgeschwindigkeit und Schreibflüssigkeit. Letzteres kann einen negativen Einfluss auf die Leistung der Testteilnehmenden haben (s. oben).

In diesem Szenario erwarten wir langfristig und bei konstant hohen Teilnehmerzahlen ebenfalls eine Kostenreduktion. Nicht zu unterschätzen sind aber nach dem Initialaufwand (Software, Hardware für die Testzentren) die Kosten für die Wartung und regelmässige Erneuerung der Infrastruktur (inkl. Räume) sowie für das Aufsichts- und technische Personal (Support). Dieses muss auch entsprechend geschult werden. Überhaupt braucht es für die Durchführung der digitalen Vor-Ort-Prüfung ein Durchführungskonzept, das im Sinne der Qualitätssicherung und der Rekursicherheit (a) Zuständigkeiten klärt (Was darf die Aufsichtsperson? Was darf das technische Personal?); (b) Fälle definiert, in denen Hilfe in Anspruch genommen werden kann; (c) praktische Fragen klärt wie z.B. die der Lärmbelastung im Raum oder der Zeitkompensation bei Verstehens-, Navigations- oder technischen Problemen.

Obwohl dieses Szenario keine Durchführung bei Lockdown sichert, schätzen wir es als durchaus erstrebenswert an. Es erlaubt die Konzeption einer digitalen, aber niederschweligen Prüfung für die gesamte Zielgruppe, die schrittweise (auch) in eine Remote-Version umgewandelt werden kann. In diesem Sinne kann die digitale Vor-Ort-Version als erster Schritt zu einer digitalen ortsunabhängigen Prüfung verstanden werden (s. Szenario 6).

### 10.3.2 Szenario 6: Parallelversionen digital vor Ort und digital Remote

In diesem Szenario wird der fide-Test als digital ortsabhängige und digital ortsunabhängige Prüfung angeboten. Dabei gelten wiederum die oben genannten Bedingungen: Gleichwertigkeit der zwei Versionen, Niederschwelligkeit, Einbezug der Stakeholder.

Die digitale Vor-Ort-Version kann eine Alternative für schulungsgewohnte MigrantInnen oder auch für Personen mit Beeinträchtigungen darstellen, die auf die Unterstützung in Testzentren angewiesen sind. Auch Personen, die aus Datenschutzgründen skeptisch hinsichtlich vor allem maschineller Proctoring-Verfahren sind, können die digitale Vor-Ort-Prüfung absolvieren. Die digitale Remote-Version gewährleistet die Durchführung in Krisensituationen und erfüllt die Bedürfnisse der hochqualifizierten Migration wie Ortsunabhängigkeit und Terminflexibilität.

Vorteilhaft ist, dass dieselben "digitalisierten" Aufgaben, eingespielt aus derselben Item-Datenbank, für beide Versionen eingesetzt werden können. Das Schreiben kann auch in diesem Szenario im Testzentrum auf Papier stattfinden (oder in der Variante „digitaler Stift“ oder auf einer Touchscreen-Tastatur statt auf einer klassischen Tastatur).

Der beträchtliche Zusatzaufwand für die digitale Remote-Version besteht in der Entwicklung und Durchführung des Proctoring sowie in der Klärung der datenschutzbezogenen und rechtlichen Fragen (inkl. Rekursicherheit). Es wird aber davon ausgegangen, dass bei hoher stabiler Nachfrage längerfristig mit einer Kostenreduktion zu rechnen ist. Insbesondere wird es weniger Testzentren und demzufolge eine schlankere Infrastruktur brauchen, wenn zum Beispiel die Hälfte der Zielgruppe den fide-Test remote absolviert. Auch hier empfehlen wir eine grössere Nutzer-Umfrage zur Akzeptabilität der digitalen Version und zur Präferenz zwischen digital vor Ort und digital remote, damit eingeschätzt werden kann, welcher Anteil auf die einzelnen Formate entfallen würde.

In Bewusstsein, dass dieses Szenario mit einem grossen Aufwand verbunden ist, schätzen wir es als zukunftsfähig ein. Es berücksichtigt die Veränderungen im Kontext des Sprachprüfens und antizipiert die Entwicklungen im Lehren und Lernen, bei der Zielgruppe und in der Gesellschaft überhaupt. Wie im Szenario 6 erwähnt, empfehlen wir, dass die zwei Versionen nacheinander entwickelt werden: die digitale ortsabhängige Version sollte die Basis für die digitale ortsunabhängige Version bilden. In der Zwischenzeit würde voraussichtlich auch das Know-How in Bezug auf Proctoring (inkl. technologischer Mittel dazu) wachsen.

### 10.3.3 Szenario 7: Nur digitale Remote-Version

Im Szenario sieben existiert der fide-Test nur als digitale Remote-Version. Die Bedingungen der Zugänglichkeit für die schulungsgewohnten MigrantInnen sowie des Einbezugs der Stakeholder bei der Entwicklung der neuen Version müssen auch hier erfüllt werden.

Es wird davon ausgegangen, dass es konzeptionell und technisch möglich ist, den Test so zu entwickeln (im Hinblick auf digitale Aufgabentypen, Benutzerinterface, Navigation, Erklärvideos), dass die schulungsgewohnten Teilnehmenden nicht benachteiligt werden. Allerdings bleiben gewisse Herausforderungen bestehen, deren Bewältigung uns schwierig erscheint:

- a. die technische Ausstattung, die den Teilnehmenden zur Verfügung steht  
Für die Durchführung des Proctoring sind zwei einwandfrei funktionierende Geräte sowie eine maximal stabile Internetverbindung notwendig. Aktuell, so die Aussage der Lernenden und die Einschätzung der IntegrationsexpertInnen, verfügen die schulungsgewohnten Teilnehmenden nicht alle über diese Infrastruktur.  
In Bezug auf die Programmierung stellt sich hier die zusätzliche Herausforderung, dass die Testapplikation nicht nur mit allen möglichen Betriebssystemen kompatibel sein muss, sondern auch eine gleichwertige Darstellung des Nutzerinterface und aller Funktionen darstellen (s.o.).
- b. die hohen digitalen Kompetenzen, welche die Teilnahme an einer digitalen Remote-Version voraussetzt  
Dabei geht es nicht um das Absolvieren der Testaufgaben selbst, sondern um all jene Prozesse, die durch die Anforderungen des Remote-Proctoring mit der Installation der verschiedenen Applikationen, der Anmeldung, der Identifikation, des Login-Verfahrens oder des Troubleshootings verbunden sind. Es musste eine entsprechende Unterstützung für diese Prozesse gewährleistet werden; der Aufwand für den/die Anbieter, aber auch für die Teilnehmenden wäre weiterhin hoch respektive signifikant höher.  
Möglich wäre es hier, ein Vorab-Login (z.B. 24 Stunden vor Testbeginn) zu ermöglichen, das es den Testteilnehmenden erlaubt, die Downloads, Anmeldung, Login etc. einmal noch weitgehend stressfrei zu absolvieren. Eventuell würde an der Stelle auch ein Zugang zu einer virtuellen Übungslandschaft freigegeben, in der man Erklärvideos zum Gesamtest (Aufbau, Navigation, Hilfe-Optionen, Support etc.) und Beispiele für die einzelnen Aufgabentypen zum Ausprobieren findet.
- c. die psychologische Unterstützung der Teilnehmenden, die bei menschlicher Präsenz gewährleistet ist und bei einer digitalen Remote-Version wegfallen würde  
Der ALTE-Experte für Testen im Kontext von Migration und Integration rät genau aus diesem Grund von einer digitalen Remote-Version als einziger Variante ab.

## 10.4 Überblick der Szenarien im Vergleich: Merkmale, Pro-/Contra-Argumente, Gelingensbedingungen

In der folgenden Tabelle 6 werden die oben skizzierten Szenarien in Kürze zusammengefasst in Bezug auf ihre Potenziale/Chancen, Herausforderungen/Risiken und Gelingensbedingungen. Einschätzungen zu Aufwänden und Kosten finden sich in Kapitel 9 und werden hier nicht wiederholt.

**Tab. 6: Übersicht über die Szenarien im Vergleich**

Szenario	Potenziale / Chancen	Herausforderungen / Risiken	Gelingensbedingungen
<b>1: nur Papier</b>	keine Kostenfolgen Weiterführung des bestehenden Systems	Durchführungssicherheit nicht garantiert Ortsunabhängigkeit nicht garantiert Zukunftsfähigkeit nicht garantiert Marktplatzierung nicht garantiert Kerngruppe des fide-Tests bei einer notfallmässigen Umstellung auf digitalen Test potentiell benachteiligt	
<b>2: Papier und digital vor Ort</b>	Marktplatzierung digitale Vor-Ort-Version nutzbar für spätere Remote-Version Kostensparnis durch automatische Korrekturoptionen in der digitalen Version kein Aufwand für Proctoring	Durchführungssicherheit nicht garantiert Ortsunabhängigkeit nicht garantiert Terminflexibilität nicht garantiert hoher Entwicklungsaufwand hoher Aufwand für Infrastruktur in Testzentren und Schulungen für Aufsichtspersonen hoher Aufwand für Parallelführung (Itemdatenbank; Gleichwertigkeitsstudien; Testadministration)	hohe Entwicklungsgeschwindigkeit Gleichwertigkeit der Versionen ggf. als Zwischenlösung ggf. Rentabilität erreichbar bei hoher Nachfrage nach der digitalen Version
<b>3: Papier und digital remote</b>	Durchführungssicherheit Ortsunabhängigkeit Terminflexibilität Marktplatzierung kein Aufwand für Infrastruktur in Testzentren Kostensparnis durch automatische Korrekturoptionen in der digitalen Version digitale Remote-Version nutzbar für spätere digitale Vor-Ort-Version	hoher Entwicklungsaufwand hoher Aufwand für Parallelführung (Itemdatenbank; Gleichwertigkeitsstudien; Testadministration) Aufwand für Schulungen für Aufsichtspersonen hoher Personalaufwand für menschliches Proctoring oder hoher Aufwand für Klärung Datenschutz bei maschinellem Proctoring Polarisierung: Papier-Version für Schulungsgewohnte – digital remote für Schulungsgewohnte Kerngruppe des fide-Tests bei einer notfallmässigen Umstellung auf digitalen Test potentiell benachteiligt	Konzeption der Remote-Version mit zentraler Berücksichtigung der schulungsgewohnten, schriftfernen Testteilnehmenden, um eine Benachteiligung auszuschliessen ggf. menschliche Begleitung für schulungsgewohnte Testteilnehmende in einem Video-Call Tutorials und Online-Support in verständlicher Sprache auch für niedrige Kompetenzstufen zunehmender Einbezug digitaler Kompetenzen im Lehr-, Lern- und Testprozess: eventuell Veränderung des Testkonstrukts Gleichwertigkeit der Versionen ggf. Rentabilität erreichbar bei hoher Nachfrage nach der digitalen Version ggf. als Zwischenschritt hin zu Szenario 6

<p><b>4: Papier, digital vor Ort und digital remote</b></p>	<p>Durchführungssicherheit Ortsunabhängigkeit Terminflexibilität Marktplatzierung Kostensparnis durch automatische Korrekturoptionen in den digitalen Versionen Aufwandsersparnis bzgl. Prüfungsadministration in den digitalen Versionen</p>	<p>hoher Entwicklungsaufwand hoher Aufwand für Parallelführung (Itemdatenbank; Gleichwertigkeitsstudien; Testadministration) unklare Nachfrage nach der digitalen Vor-Ort-Version, die mit hohen Kosten für die Infrastruktur verbunden ist Aufwand für Schulungen für Aufsichtspersonen hoher Personalaufwand für menschliches Proctoring in der Remote-Version oder hoher Aufwand für Klärung Datenschutz bei maschinell Proctoring in der Remote-Version relativ hoher Aufwand für den Nachweis der Gleichwertigkeit der drei Versionen</p>	<p>Konzeption der digitalen Versionen mit zentraler Berücksichtigung der schulungsgewohnten, schriftfernen Testteilnehmenden, um eine Benachteiligung auszuschliessen Tutorials und Online-Support in verständlicher Sprache auch für niedrige Kompetenzstufen zunehmender Einbezug digitaler Kompetenzen im Lehr-, Lern- und Testprozess: eventuell Veränderung des Testkonstrukts Gleichwertigkeit der Versionen NutzerInnen-Umfrage zu Format-Präferenzen, insbesondere Nachfrage digital vor Ort ggf. Rentabilität erreichbar bei hoher Nachfrage nach der digitalen Remote-Version</p>
<p><b>5: nur digital vor Ort</b></p>	<p>Marktplatzierung keine Kosten für Doppelführung Papier-digital Kostensparnis durch automatische Korrekturoptionen Aufwandsersparnis bzgl. Prüfungsadministration kein Aufwand für Proctoring digitale Vor-Ort-Version nutzbar für spätere Remote-Version menschlicher Support für schulungsgewohnte Testteilnehmende</p>	<p>Durchführungssicherheit nicht garantiert Ortsunabhängigkeit nicht garantiert Terminflexibilität nicht garantiert hoher Entwicklungsaufwand hoher Aufwand für Infrastruktur in Testzentren und Schulungen für Aufsichtspersonen Aufwand für Schulungen für Aufsichtspersonen Akzeptabilität herstellen</p>	<p>auch für schriftferne Testteilnehmende angemessenes Medium Konzeption mit zentraler Berücksichtigung der schulungsgewohnten, schriftfernen Testteilnehmenden, um eine Benachteiligung auszuschliessen Tutorials in verständlicher Sprache auch für niedrige Kompetenzstufen schnelle, gute Kommunikation mit den Stakeholdern</p>
<p><b>6: digital vor Ort und digital remote</b></p>	<p>Durchführungssicherheit Ortsunabhängigkeit Terminflexibilität Marktplatzierung Zukunftsfähigkeit keine Kosten für Doppelführung Papier-digital hohe Rentabilität des Entwicklungsaufwands durch doppelte Nutzung Kostensparnis durch automatische Korrekturoptionen Aufwandsersparnis bzgl. Prüfungsadministration Vor-Ort-Version mit menschlichem Support für schulungsgewohnte Testteilnehmende</p>	<p>hoher Entwicklungsaufwand Aufwand für Schulungen für Aufsichtspersonen hoher Aufwand für Infrastruktur in Testzentren und Schulungen für Aufsichtspersonen hoher Personalaufwand für menschliches Proctoring in der Remote-Version oder hoher Aufwand für Klärung Datenschutz bei maschinell Proctoring in der Remote-Version Akzeptabilität herstellen</p>	<p>auch für schriftferne Testteilnehmende angemessenes Medium Konzeption mit zentraler Berücksichtigung der schulungsgewohnten, schriftfernen Testteilnehmenden, um eine Benachteiligung auszuschliessen Tutorials und Online-Support in verständlicher Sprache auch für niedrige Kompetenzstufen ggf. Rentabilität erreichbar bei hoher Nachfrage nach der Remote-Version NutzerInnen-Umfrage zu Akzeptabilität und Format-Präferenzen schnelle, gute Kommunikation mit den Stakeholdern</p>
<p><b>7: nur digital remote</b></p>	<p>Durchführungssicherheit Ortsunabhängigkeit Terminflexibilität</p>	<p>hoher Entwicklungsaufwand sehr hohe Herausforderung für</p>	<p>auch für schriftferne Testteilnehmende angemessenes Medium</p>

	Marktplatzierung keine Kosten für Doppel- führung Papier-digital Kostenersparnis durch automatische Korrekturo- ptionen Aufwandsersparnis bzgl. Prüfungsadministration	schulungsgewohnte Testteilneh- mende Aufwand für Schulungen für On- line-Aufsichtspersonen hoher Personalaufwand für menschliches Proctoring (und Online-Support) oder hoher Aufwand für Klärung Datenschutz bei maschinellm Proctoring Akzeptabilität herstellen	Konzeption mit zentraler Be- rücksichtigung der schulunge- wohnten, schriftfernen Testteil- nehmenden, um eine Benach- teiligung auszuschliessen ggf. menschliche Begleitung für schulungsgewohnte Testteil- nehmende in einem Video-Call Tutorials und Online-Support in verständlicher Sprache auch für niedrige Kompetenzstufen schnelle, gute Kommunikation mit den Stakeholdern
--	---	---	--

## 10.5 Fazit und Empfehlung

Wie Kapitel 10.2 bis 10.4 zeigen, weisen alle denkbaren Szenarien Vor- und Nachteile auf und sie bringen Potenziale oder Risiken mit sich. Die Wahl bzw. Priorisierung des einen oder anderen Szenarios ist daher stark abhängig von den konkreten Zielen des Auftraggebers und der Priorisierung dieser Ziele:

Ist beispielsweise die Durchführungssicherheit zentrales Ziel, braucht es ein Szenario mit einer Remote-Version (Szenarien 3, 4, 6, 7).

Ist hingegen die Marktpositionierung zentrales Ziel, kommt nur ein Szenario mit einer digitalen, potenziell einer Remote-Variante in Frage (Szenarien 2-7).

Sollte die Berücksichtigung der schulungsgewohnten Testteilnehmenden, d.h. die Niedrigschwelligkeit des Tests, zentrales Anliegen sein, dann ist potenziell eine Variante mit menschlicher 1-zu-1-Begleitung notwendig, ob als Papier- oder digitale Version, ob vor Ort oder remote (Szenarien 1-7, mit jeweils unterschiedlichen Herausforderungen).

### 10.5.1 Abschliessende Bewertung Szenario 3

Da Szenario 3 (Papier und Remote), das die Durchführung des Tests gewährleistet und gleichzeitig die aktuelle fide-Kernzielgruppe (unter normalen Umständen) nicht benachteiligt, eine denkbare Option für den Auftraggeber darstellt, gehen wir im Folgenden nochmals fokussiert auf die Potenziale, Herausforderungen und Gelingensbedingungen von Szenario 3 ein, insbesondere im Vergleich zum u.E. besonders zukunftsfähigen Szenario 6 (digital vor Ort und remote).

Als besondere Potenziale von Szenario 3 sind hervorzuheben erstens die Wahlmöglichkeit – zumindest unter normalen Umständen – zwischen der etablierten Papier-Version vor Ort, die potenziell gerade schulungsgewohnte Lernende anspricht, und einer digitalen Remote-Version, die potenziell besonders für schulungsgewohnte Testteilnehmende attraktiv ist; zweitens das Bewahren etablierter Strukturen; und drittens die Durchführungssicherheit, Ortsunabhängigkeit und Terminflexibilität der Remote-Version. Zusätzlich kann angenommen werden, dass der für die Remote-Version entwickelte Test nach und nach auch in einer Vor-Ort-Version eingesetzt werden kann, das Szenario enthält also Entwicklungspotenzial.

Als herausfordernd bzw. aufwändig dürfte sich in Szenario 3 einerseits der erwartbar hohe Aufwand (personell und zeitlich) sowohl für die Parallelführung einer digitalen und einer Papierversion als auch für die Lösung der Proctoring-Frage (mit entweder hohem personellem Aufwand für menschliches Proctoring oder hohem datenschutzrechtlichem Klärungsaufwand für das maschinelle Proctoring) erweisen. Schwierig zu führen wäre auch der Nachweis, dass die Testversionen tatsächlich gleichwertig sind; dies würde voraussetzen, dass eine substantielle Subgruppe schulungsgewohnter Testteilnehmender die Papier-Version absolviert, umgekehrt eine substantielle Subgruppe der schulungsgewohnten Testteilnehmenden die Remote-Version absolviert (so dass nachweisbar ist, dass keine Gruppe in einem der Kontexte signifikant besser abschneidet). Zudem stellt sich die Frage, ob angesichts des allgemeinen digitalen Wandels die Papier-Version in den kommenden Jahren noch attraktiv und marktfähig sein wird, auch für schulungsgewohnte Lernende. Diesbezüglich ist vor allem darauf zu achten, dass in Szenario 3 in speziellen Situationen wie z.B. bei einem neuen Lockdown, die schulungsgewohnten Testteilnehmenden nicht benachteiligt werden, weil

ihre Bedürfnisse bei der Entwicklung der Remote-Version nicht angemessen berücksichtigt wurden und/oder sie nicht ausreichend auf das Absolvieren der Remote-Version vorbereitet wurden.

Szenario 3 bedingt daher in besonderem Masse die Berücksichtigung der schulungsgewohnten Testteilnehmenden, ihrer Bedürfnisse und besonderen Herausforderungen u.a. in Bezug auf Schriftsprachlichkeit und Prüfungsangst, bei der konzeptionellen Entwicklung der digitalen Remote-Version; gleichzeitig darf dies nicht dazu führen, dass der Test weniger attraktiv oder gar unattraktiv wird für schulungsgewohnte Testteilnehmende. Eine positive Rückwirkung (sog. *washback*) von Szenario 3 könnte darin bestehen, dass die digitalen Fertigkeiten bereits in Sprachkursen für Schulungsgewohnte stärker berücksichtigt werden (vgl. Norwegen) und dass parallel auch die entsprechende Aus- und Weiterbildung der Sprachlehrpersonen im Kontext Integration in Bezug auf die Anwendung digitaler Kompetenzen und ihre Förderung Fortschritte macht (Stichwort „digitales Lernen, Lehren, Testen“).

### 10.5.2 Abschliessende Bewertung Szenario 6

In Szenario 6 ist demgegenüber besonders positiv, dass der Aufwand für die Parallelführung der Papier- und einer digitalen Version entfallen (gleiche Itemdatenbank, gleiches Verfahren bei der Testadministration). Gleichzeitig darf dieses Szenario als besonders zukunftsfähig in Bezug auf die schnelle Positionierung am Markt gelten. Denn die digitale Vor-Ort-Version könnte relativ schnell entwickelt und angeboten werden, ohne zunächst die datenschutzrechtlichen Aspekte, die mit der Remote-Version verbunden sind, klären zu müssen.

Ausserdem erlaubt dieses Szenarios eine teilweise Neukonzeption des fide-Tests für eine digitale Version, ohne Gleichwertigkeitsproblem (s.o.). Und schliesslich ergibt sich für den erwartbar hohen Initialaufwand für die Entwicklung des digitalen Tests (und die Schulung der Aufsichtspersonen) eine relativ hohe Rentabilität, da dieselbe digitale Version vor Ort und remote genutzt wird, woraus sich reduzierte Kosten und eine einfache Handhabung derselben Plattform (Item-Datenbank, Nutzerinterface etc.) ergeben. Zur Kostenreduktion tragen auch eine teilweise vereinfachte Korrektur und reduzierte Aufwände für die Prüfungsadministration bei.

Als zentrale Herausforderung in Szenario 6 dürfte sich die Akzeptabilitätsfrage erweisen: Wie kann man die verschiedenen Stakeholder überzeugen, dass eine Papierversion des fide-Tests nicht essenziell sein muss? Für schulungsgewohnte Testteilnehmende ergeben sich die oben bereits angesprochenen Herausforderungen in Bezug auf Schriftsprachlichkeit und digitale Fertigkeiten; diese wären in einer digitalen Vor-Ort-Version aber deutlich reduziert gegenüber einer Remote-Version, so dass bei angenommener Wahlmöglichkeit voraussichtlich wenig Probleme entstehen, da die Vor-Ort-Version ein relativ stressfreies Setting mit menschlicher Begleitung und Hilfe bieten würde. Auch hier gilt, dass die Bedürfnisse der schulungsgewohnten Testteilnehmenden bei der Entwicklung der digitalen Version, aber auch der Erklärvideos und des Online-Supports zentral berücksichtigt werden müssen. Die Lösung der Proctoring-Frage (s.o.) wäre in Szenario 6 ebenso aufwändig wie in Szenario 3.

Gelingen kann das Szenario 6 bei einer strikten Berücksichtigung der schulungsgewohnten Testteilnehmenden (wie in Szenario 3), ggf. mit der Entscheidung, trotz des hohen personellen Aufwands das Online-Proctoring im Sinne einer Online-Begleitung mit einer menschlichen Aufsicht zu gestalten. Auch die gute, frühzeitige Kommunikation mit den Stakeholdern ist in diesem Szenario sicher essenziell. Bezüglich der Entwicklungsgeschwindigkeit gilt sicher für beide Szenarien: so bald wie möglich, mit der entsprechenden personellen Ausstattung des Entwicklungsteams.

Eine Wahl von Szenario 1, die angesichts der aufgeführten Herausforderungen bei der Entwicklung und Durchführung von digitalen Tests ggf. als einfachster Weg erscheint, ist zu bedenken, dass damit die Zukunftsfähigkeit des fide-Tests ggf. nicht gesichert ist, da sowohl der Markt als auch die Erwartungen (und Bedürfnisse) der Testteilnehmenden sich aktuell stark verändern und auch weiterhin verändern werden. Bei konstant hohen bzw. noch steigenden Teilnehmendenzahlen ist – je nach Szenario – mittel- und längerfristig durchaus mit einer hohen Effizienz durch digitales Testen zu rechnen.

### 10.5.3 Handlungsempfehlungen

Grundlegend muss in Hinblick auf die Beantwortung der Frage, ob die Digitalisierung des fide-Tests überhaupt stattfinden soll und wie schnell diese erfolgen soll, eine strategische Entscheidung getroffen werden: Soll es einen spezifischen, eigenen Schweizer Sprachtest geben?

Wenn es einen solchen geben soll, muss sich der fide-Test angesichts des digitalen Wandels in der Gesellschaft und im Lehr-/Lernkontext entsprechend entwickeln, und zwar hinsichtlich der Marktpositionierung bald; die erwartbaren beträchtlichen Kosten müssen entsprechend getragen werden.

Alternativ ist es angesichts der aktuellen Marktentwicklungen (digitales Goethe-Zertifikat B1, kommender *Start Deutsch 1*-Remote-Test, Norwegian Language Test etc.) allerdings auch möglich, auf eine in allen Aspekten eigenständige Schweizer Testentwicklung zu verzichten und stattdessen zeitnah einen bestehenden (oder aktuell in Entwicklung befindlichen) digitalen Test (mit Datenbank für Items, Nutzerinterface, Prüfungsadministration, ggf. Proctoring-Lösung) zu lizenzieren und anzupassen bzw. anpassen zu lassen (bzgl. Schweiz-Bezug der Aufgaben, Schriftfreiheit beim Hörverstehen etc.).

Wenn es allerdings einen eigenständigen Schweizer Test geben soll, empfehlen wir ein kleinschrittiges Vorgehen, das es erlaubt, die zeitlichen, personellen und finanziellen Ressourcen besser einschätzen zu können. Dieses könnte folgende Schritte umfassen:

- eine zeitnahe Durchführung einer Umfrage bei aktuellen/zukünftigen NutzerInnen zu Format-Präferenzen (u.a. Szenario 3, 4, 6);
- die Bildung eines konstanten Kernteams, das hohe Expertise aufweist und den ganzen Prozess leitet. Dadurch soll die Kontinuität zwischen den verschiedenen Arbeitsschritten gewährleistet und der Koordinationsaufwand minimiert werden;
- eine Überprüfung der bestehenden Aufgabentypen und -formate in Hinblick auf ihre Eignung für die digitale Version mit unverändertem Testkonstrukt;
- eine Diskussion bezüglich einer potentiellen Weiterentwicklung des Testkonstrukts für eine optimale Ausnutzung des digitalen Mediums;
- die Entwicklung oder Lizenzierung einer Plattform mit entsprechendem Nutzerinterface (Nachhaltigkeit der Zusammenarbeit sicherstellen);
- eine allmähliche Digitalisierung der bestehenden Aufgabentypen und -formate in Zusammenarbeit zwischen dem Kernteam, Item-EntwicklerInnen, IT-Verantwortlichen;
- frühzeitige Abklärungen von datenschutzrechtlichen Fragen;
- die Entwicklung eines Konzepts für die Kommunikation mit den Stakeholdern und deren frühzeitigen Einbezug in Entscheidungs- und Entwicklungsprozesse
- die Weiterbildung von PrüferInnen und Kursleitenden im Bereich des digitalen Lehrens, Lernens und Prüfens
- die technische Schulung der Aufsichtspersonen bzw. Support-Verantwortlichen
- die Entwicklung von Informations-, Übungs- und Vorbereitungsressourcen (u.a. Erklärvideos, Einführung durch eine Aufsichtsperson, digitale Übungsumgebung)
- die Etablierung einer klaren Liste von zulässigen Rekursgründen im Falle digitalen Testens (Rekursicherheit)
- je nach Szenario eine Organisation der notwendigen Begleitforschung bezüglich Usability und Gleichwertigkeit.

## Literaturverzeichnis

Bouriga, S. & Olive, T. (2021), Is typewriting more resources-demanding than handwriting in undergraduate students? *Reading & Writing* 34, 2227-2255.

Europarat (Hrsg.) (2020). *Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen*. Begleitband. Stuttgart: Ernst Klett Sprachen GmbH.

Sturm, A. (2014). Basale Lese- und Schreibfertigkeiten bei BerufsschülerInnen und die Notwendigkeit kompensatorischer Fördermassnahmen. *leseforum.ch* 1.

Tschirner, Erwin. (2000). Interaktion und Interaktivität in Testwissenschaft und Sprachlehrforschung. In K. R. Bausch, H. Christ, F. Königs & H. J. Krumm (Hrsg.), *Interaktion des Lehrens und Lernens fremder Sprachen* (S. 241–248). Arbeitspapiere der 20. Frühjahrestagung zur Erforschung des Fremdsprachenunterrichts. Tübingen: Narr.