

„Online-Assessments“ und „Self-Assessments“ in der Eignungsdiagnostik

10

Michael Ott, Anna-Sophie Ulfert und Martin Kersting

Zusammenfassung

„Online-Assessments“ finden zunehmende Verbreitung im Kontext der Eignungsdiagnostik. In diesem Beitrag werden sie zur Personalauswahl und Eignungsdiagnostik beschrieben. Dabei definieren wir „Online-Assessment“ als Oberbegriff und differenzieren anhand von drei Unterscheidungsmerkmalen relevante Varianten: Erstens anhand des primären Ziels (Selbst- oder Fremdselektion), zweitens bezüglich der Methode (Testen oder Informieren), drittens anhand des Ausmaßes an Kontrolle in der Durchführung (beaufsichtigt oder unbeaufsichtigt). Wir berichten über die Effekte der unbeaufsichtigten Testdurchführung. Auf der Basis dieser Differenzierungen definieren wir „Self-Assessments“ als diejenige Form des „Online-Assessments“, die mittels unbeaufsichtigter Testung auf die Selbstselektion von Bewerbern abzielen. Abschn. 10.2 dient der Erläuterung dieser drei Unterscheidungsmerkmale. Abschn. 10.3 thematisiert die Besonderheiten von „Online-Assessments“ (z. B. Usability, Datenschutz, Akzeptanz), die sich aus dem online-basierten Charakter ergeben. Diese Ausführungen bilden die Grundlage zur Qualitätsbeurteilung von „Online-Assessments“, wobei wir verdeutlichen, welche Konsequenzen sich aus den Besonderheiten für die praktische

M. Ott (✉) · A.-S. Ulfert · M. Kersting

Institut für Psychologie und Sportwissenschaft, Justus-Liebig-Universität Giessen,
Giessen, Deutschland

E-Mail: Michael.Ott@psychol.uni-giessen.de

A.-S. Ulfert

E-Mail: Anna-Sophie.Ulfert@psychol.uni-giessen.de

M. Kersting

E-Mail: Martin.Kersting@psychol.uni-giessen.de

diagnostische Anwendung ableiten. Abschn. 10.4 stellt Kriterien vor, anhand derer die Qualität von entsprechenden Online-Tools bewertet bzw. gesichert werden kann.

Inhaltsverzeichnis

10.1	Einleitung	216
10.2	Unterscheidungsdimensionen	219
10.2.1	Selbst- und Fremdselektion	220
10.2.2	Testen und Informieren	223
10.2.3	Durchführung mit und ohne „Aufsicht“ – Herausforderungen	224
10.2.4	Kombination der Gestaltungselemente	227
10.3	Spezifika der Durchführung als Folge des Online-Charakters	228
10.3.1	Usability und Attraktivität von Tools	229
10.3.2	Wahrgenommener Datenschutz und wahrgenommene Fairness	230
10.3.3	Interindividuelle Unterschiede bezüglich der Akzeptanz von Online-Testungen	230
10.4	Bewertung der Qualität von „Online-Assessments“ und „Self-Assessments“	231
10.5	Fazit	234
10.5.1	Was bei der Konstruktion und dem Einsatz von „Online-Assessments“ und „Self-Assessments“ zu beachten ist	234
10.5.2	Forschungsdesiderata	236
10.5.3	Schlusswort	238
	Literatur	238

10.1 Einleitung

Die Begriffe „Online-Assessment“, „online tools“, „Online-Tests“, „e-Assessment“, „Recruitment“, „Online-Self-Assessment“ oder auch nur „Self-Assessment“ werden im Kontext der Eignungsdiagnostik häufig genutzt. Da verschiedene diagnostische Praktiken und Ziele unter diese Begriffe subsumiert werden, verschwimmt mitunter die Prägnanz der Begriffe. Der folgende Beitrag verfolgt aus diesem Grund zwei Ziele: Zunächst nehmen wir begriffliche Setzungen vor. Wir schlagen vor, den Begriff „Online-Assessment“ als Oberbegriff zu nutzen, und definieren „Self-Assessments“ als Unterkategorie dieses Oberbegriffs. Nach welchen Gesichtspunkten „Online-Assessments“ differenziert werden können und wie „Self-Assessments“ hierin zu verorten sind, wird im ersten Teil des Beitrags verdeutlicht. Darüber hinaus berichten wir empirische Befunde zum Einsatz von „Online-Assessments“, wobei wir auch Befunde aus benachbarten Themenbereichen (z. B. Online-Bewerbungen) aufführen, soweit sie für die Thematik bedeutsam und für die Beurteilung, Sicherung und Optimierung der Qualität des (wachsenden) Angebots hilfreich sind.

Betrachten wir zunächst die bisherige Begriffsverwendung und beginnen mit dem Begriff „Online-Assessment“. Konradt und Sarges [36] nutzen diesen Begriff im Kontext

der Personalauswahl ganz allgemein für jede online durchgeführte Diagnostik, die das Ziel verfolgt, die berufliche Eignung von Bewerbern festzustellen. „Online-Assessments“ dienen somit der Selektion (oder Vorselektion) von geeigneten Bewerbern, wobei die diagnostischen Verfahren in der Regel nicht frei zugänglich sind, sondern vom Bewerber im Laufe des Auswahlprozesses erst dann durchgeführt werden können, wenn eine entsprechende Einladung durch die auswählende Organisation erfolgt [14]. Neben der Selektion kann das „Online-Assessment“ nach unserer Definition auch die Selbstselektion umfassen.

- ▶ „Online-Assessments“ bieten zahlreiche Vorteile:
 - Unabhängigkeit von Raum, Zeit und Ressourcen (Rund-um-die-Uhr-Angebot, von jedem Ort der Welt [mit Internetanschluss] erreichbar)
 - Anonyme Durchführung möglich (bei „Self-Assessments“)
 - Hohe Auswertungs- und Interpretationsobjektivität
 - Ökonomie
 - Passung zur Zielgruppe (das Internet ist das Medium der Zielgruppe, Angebote im Internet erreichen die Zielgruppe und tragen zu einem „modernen“ Image des Anbieters bei)

Der vorliegende Beitrag beschränkt sich auf die Nutzung von „Online-Assessments“ im Bereich der Eignungsdiagnostik. Es gibt aber weitere Anwendungsfelder von „Online-Assessments“ im Personalbereich, beispielsweise das Personalmarketing und die Personalentwicklung [25]. Kirbach und Wottawa [34] verweisen explizit darauf, dass „Online-Assessments“ (und „Self-Assessments“) nicht nur Potenziale für die Personalauswahl, sondern insbesondere auch für das Personalmarketing besitzen. In Bezug auf „Self-Assessments“ verweisen Kubinger et al. [39] mit Bezug auf den Hochschulsektor darauf, dass diese nicht nur der Selbstselektion, sondern auch der positiven Außendarstellung einer Organisation dienen. Dem breiten Anwendungsfeld entsprechend mannigfaltig sind die möglichen Zielgruppen des „Online-Assessments“, von Schülern und Studieninteressierten über Berufsanfänger bzw. -einsteiger bis hin zu Personen mit Berufserfahrung [25].

Der Begriff „Self-Assessment“ ist in der Personalauswahl relativ neu. In der wörtlichen Übersetzung einer „Selbsteinschätzung“ handelt es sich aber um nichts Neues. Neu hingegen ist, dass unter dem Begriff „Self-Assessment“ eine Reihe unterschiedlicher Konzepte aufkommen, die die Internet-Technologie nutzen und somit einen computergetriebenen, onlinebasierten diagnostischen Prozess ermöglichen.

Nach Kubinger [38] testet sich eine Person beim „Self-Assessment“ „... (zumeist über Internet) selbst und absolviert folglich den diagnostischen Prozess eigenverantwortlich und fachpsychologisch unkontrolliert“. Damit können „Self-Assessments“ sowohl im Sinne einer Selbsteinschätzung, aber auch als eine Form selbst (auch „an sich selbst“) durchgeführter Diagnostik verstanden werden (vgl. auch der Begriff des „Selbsttests“, z. B. Hell [24]).

Um eine Systematik zu etablieren, orientieren wir uns bei unseren Definitionen an dem eignungsdiagnostischen Prozess, den wir stark vereinfachend in drei Phasen untergliedern und am Beispiel der Selektion (Personalauswahl) erläutern (siehe Abb. 10.1). Zu Beginn der Selektion, in der ersten Phase, steht ein mehr oder minder konkretes Interesse der potenziellen Bewerber an einer Tätigkeit oder Position bzw. Organisation und/oder das Interesse der Organisation an bestimmten Bewerbern. In der zweiten Phase werden Informationen über die tatsächlichen Bewerber und/oder die Anforderungen generiert. Diese Informationen sind die Basis für die Entscheidung, der dritten Phase unseres vereinfachten Prozessmodells. Dabei kann die Entscheidung sowohl die Entscheidung des Bewerbers sein, die Bewerbung aufrecht zu erhalten oder ein Angebot anzunehmen als auch die Entscheidung der Organisation, einem Bewerber ein Angebot zu unterbreiten.

Der Begriff „Online-Assessment“ bezieht sich auf die zweite Phase des dargestellten Prozesses. „Online-Assessments“ können sowohl zur Personal(vor-)auswahl als auch zur Selbsterkenntnis bzw. Selbstselektion der Teilnehmer eingesetzt werden. Den Begriff „Self-Assessment“ sollte man hingegen, so unser Vorschlag, nur nutzen, wenn die Selbsterkenntnis des Teilnehmers und die darauf basierende Selbstselektion primäres Ziel sind. Eine Gemeinsamkeit aller hier unter „Online-Assessment“ subsumierten Verfahren ist die computergestützte und onlinebasierte Umsetzung. Um vorhandene Angebote sachgerecht zu beurteilen und neue Instrumente zielorientiert zu gestalten, schlagen wir drei zentrale Unterscheidungsdimensionen vor, anhand derer sich „Online-Assessments“ charakterisieren bzw. kategorisieren lassen:

- ▶ • Ziele: Selbst- und Fremdselektion
- Methoden: Testen und Informieren als Funktionsmechanismen zur Erreichung dieser Ziele
- Kontrolle: Unbeaufsichtigte und beaufsichtigte Testung als Eigenschaft der Darbietungsweise

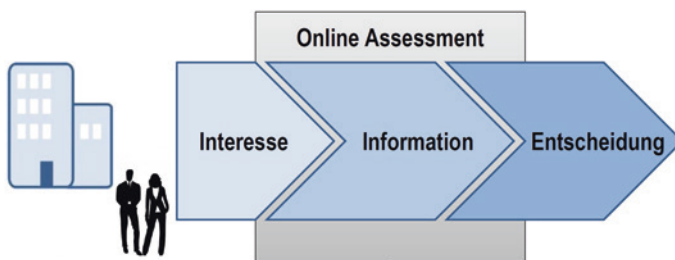


Abb. 10.1 Vereinfachtes Prozessmodell der Eignungsdiagnostik bei der Selbst- oder Fremdselektion

Neben diesen drei Unterscheidungskriterien existieren weitere Unterscheidungsdimensionen, auf die wir hier nicht eingehen. So könnte man auch

- das Ausmaß der Interaktivität (niedrig/hoch) sowie
- das Ausmaß der Einbindung spielerischer Elemente bei der Testung (niedrig/hoch) (Stichwort: „Recrutainment“, [13])

als Unterscheidungsdimensionen nutzen. Für den Hochschulsektor könnte man die von Hell [24] vorgeschlagenen Differenzierungen

- „hochschulübergreifend“ versus „hochschulspezifisch“,
- „fachübergreifend“ versus „fachspezifisch“

sowie weitere Einteilungsgesichtspunkte nutzen. Wir beschränken uns auf die zuerst genannten drei Dimensionen.

Nach einer näheren Beschreibung der drei oben genannten Unterscheidungsdimensionen thematisieren wir die Besonderheiten, die aus dem Online-Setting für den diagnostischen Prozess erwachsen. Dies betrifft die drei Hauptaspekte Usability, Datenschutz und Akzeptanz von „Online-Assessments“. Diese Ausführungen bilden die Grundlage zur Qualitätsbeurteilung von „Online-Assessments“, auf die wir im letzten Teil des Beitrags eingehen. Dabei nehmen wir eine Analyse von Erfolgs- sowie Risikofaktoren vor. Da „Online-Assessments“ häufig automatisiert erstellte Ergebnis- und (zum Teil) Entscheidungsrückmeldungen umfassen, gehen wir auch auf diese Form des Feedbacks ein und thematisieren die besonderen Anforderungen, die sich dabei stellen.

10.2 Unterscheidungsdimensionen

Zunächst differenzieren wir die unterschiedlichen Ziele der „Online-Assessments“ anhand des Entscheidungsträgers und unterscheiden zwischen Selbst- und Fremdselektion [48]. Die nächste Unterscheidungsdimension betrachtet zwei unterschiedliche Methoden, mit denen die Selbst- und Fremdselektion erreicht werden kann: Das Testen und das Informieren. Schließlich unterscheiden sich „Online-Assessments“ im Ausmaß der Kontrolle der diagnostischen Situation. Das „Online-Assessment“ kann beaufsichtigt oder unbeaufsichtigt erfolgen.

Diese drei Dimensionen dienen dazu, unterschiedliche „Online-Assessments“ zu charakterisieren. Alle drei Unterscheidungsdimensionen (siehe Abb. 10.2) stehen orthogonal zueinander, so dass theoretisch jede Kombinationsvariante der drei Dimensionen möglich ist.

Ein „Online-Assessment“ kann mit dem Ziel der Fremdselektion beaufsichtigt durchgeführt werden und informative Elemente umfassen oder aber auch mit dem Ziel der Selbstselektion unbeaufsichtigt administriert werden und sich auf das Testen beschränken. Ungeachtet dessen kommen in der Praxis bestimmte Kombinationen häufiger vor,

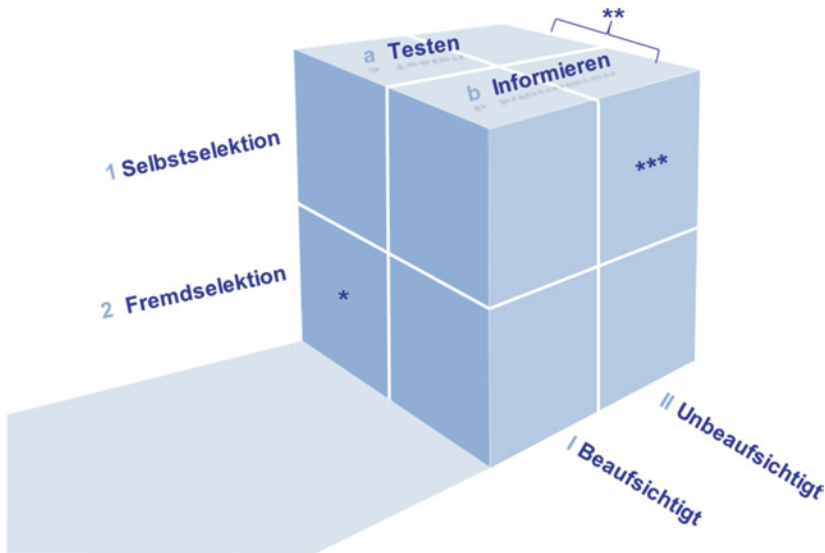


Abb. 10.2 Unterscheidungsdimensionen von „Online-Assessments“. *Fremdselektion (via beaufsichtigtem Testen); **Selbstselektion (via Testen und Informieren, unbeaufsichtigt bzw. frei zugänglich); ***Selbstselektion (via Informieren, unbeaufsichtigt bzw. frei zugänglich)

beispielsweise die beaufsichtigte Fremdselektion und die Koppelung von Selbstselektion einerseits und Information andererseits. Konkrete Beispiele für gängige Kombinationen geben wir nach der Erläuterung der drei Dimensionen. Mit einem „Online-Assessment“ können auch verschiedene Elemente, die sich in der Taxonomie gegenüber stehen, gleichzeitig realisiert werden, beispielsweise kann ein „Online-Assessment“ sowohl informative als auch diagnostische (testende) Elemente enthalten.

10.2.1 Selbst- und Fremdselektion

Die Frage nach der Selbst- bzw. Fremdselektion betrifft die Frage, wer die mit dem Test erhobenen Daten interpretiert und eine Entscheidung trifft. Der klassische Fall der Personalauswahl sieht eine Interpretation der Daten durch die einstellende Organisation vor. Person X bewirbt sich bei Organisation Y, absolviert einen Test, danach treffen Entscheidungsträger aus der Organisation Y eine Entscheidung über den weiteren Fortgang der Bewerbung. In der praktischen Anwendung läuft es häufig so, dass Bewerber bei selektierenden „Online-Assessments“, nachdem sie in Form einer gelungenen Bewerbung den ersten Schritt des Auswahlprozesses absolviert haben, eine individuelle Einladung zur Teilnahme am „Online-Assessment“ erhalten. Solche „Online-Assessments“ finden Anwendung beispielsweise bei der Auszubildendenauswahl bei der Deutschen Post, E.ON oder der Targobank [68, 69, 70] (für ein Beispiel siehe Abb. 10.3).

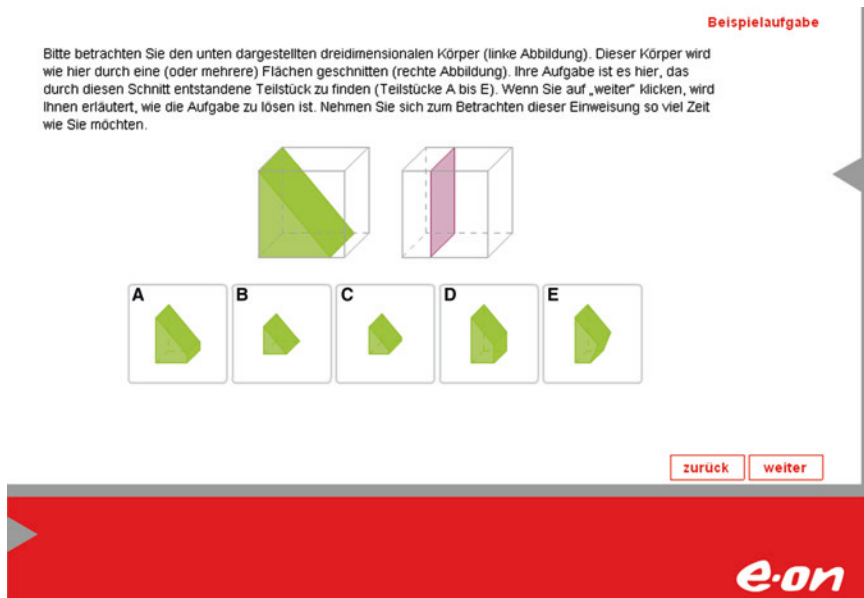


Abb. 10.3 Beispiel für eine Aufgabe aus einem Online-Assessment. Beispielaufgabe aus dem „E.ON Phasenprüfer“. (Quelle: Cyquest GmbH)

Mittels „Self-Assessments“ können „Interessenten auch ohne eine Bewerbung etwas über die eigenen Stärken und Schwächen erfahren“ [31]. Bei der Selbstselektion kann eine Person X vor der eigentlichen Bewerbung in einem Test ausprobieren, ob ihr Eignungsprofil zu dem Anforderungsprofil passt (vgl. dazu Kapitel Palmer und Kersting in diesem Band). Die Testung kann dabei anonym erfolgen, d. h. dass nur Person X ihr Testergebnis kennt. Person X entscheidet dann über ihr weiteres Bewerbungsverhalten, z. B. über Bewerbung, keine Bewerbung, Bewerbung bei einer anderen Organisation oder auch das Annehmen eines Stellenangebots. „Self-Assessments“ (also der Selbstselektion dienende „Online-Assessments“) eröffnen einem Bewerber die Möglichkeit, Informationen über die eigene Passung zu einer Ausbildung, einem Beruf, einer bestimmten Stelle und/oder einer Organisation zu erhalten. Dies soll die „Selbstreflexion“ über die Passung von persönlicher Eignung und Anforderungen anregen und eine stimmige Selbstselektion hervorrufen. Die zugrunde liegende Annahme lautet, dass geeignete Bewerber sich für eine Bewerbung entscheiden, während weniger geeignete Bewerber entweder von der Bewerbung absehen oder aber weitere Informationen sowie ggf. Kompensationsangebote nachfragen. Bei denjenigen, die sich für eine Bewerbung entscheiden, kann das „Self-Assessment“ zu einer ersten Identifikation mit der Organisation und zur Steigerung des „commitment“ führen.

- Das „Self-Assessment“ soll die absolute Zahl der Bewerber (und somit den Aufwand) minimieren und den relativen Anteil an geeigneten Bewerbern in der Gesamtgruppe der Bewerber maximieren.

Da das „Self-Assessment“, wie die meisten „Online-Assessments“, kostenfrei, zeit- und ortsunabhängig durchgeführt werden kann, handelt es sich um ein niederschwelliges Angebot: Es ist für potenzielle Nutzer leicht zugänglich und mit einem überschaubaren Aufwand verbunden. Die Teilnahme an „Self-Assessment“ Angeboten erfolgt darüber hinaus – anders als „Online-Assessments“ zur Fremdselektion – in der Regel anonym. Ob eine geeignete Person sich nach Teilnahme an dem „Self-Assessment“ für die Bewerbung bei der Organisation entscheidet, hängt wesentlich davon ab, wie attraktiv die Organisation von den (potenziellen) Bewerbern wahrgenommen wird. In Zeiten des Personalmangels wird dieser Aspekt der Entscheidung des Kandidaten immer relevanter, so dass die Organisationen viel Wert auf eine „positive candidate experience“ legen. Die Attraktivität und Professionalität des „Self-Assessments“ hat eine hohe Bedeutung dafür, wie attraktiv die Organisation eingeschätzt wird. Die unterschiedlichen Entscheidungsperspektiven der Selbst- und Fremdselektion sind in Abb. 10.4 dargestellt (Abb. 10.4).

Cronbach und Gleser [12] grenzen die Personen-Selektion von der Bedingungs-Selektion ab. Bei der Personenselektion entscheidet sich eine Organisation für einen Bewerber. Bei der Bedingungsselektion wird eine Wahl zwischen verschiedenen Bedingungen getroffen. So könnte eine Organisation überlegen und entscheiden, in welchen Bereich der Organisation ein Bewerber am besten passt. Bei der Selbst-Bedingungsselektion werden beispielsweise den Testteilnehmern im Anschluss an die Teilnahme verschiedene, individuell für sie „passende“ Ausbildungsberufe empfohlen (oder einem Studieninteressierten werden passende Studiengänge empfohlen).

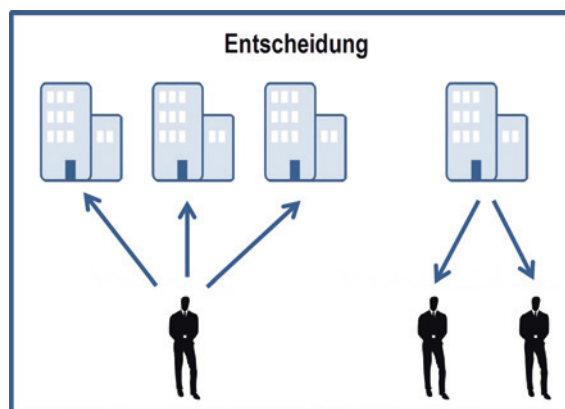


Abb. 10.4 Entscheidungsträger im „Online-Assessment“

10.2.2 Testen und Informieren

Neben dem Einsatz von Tests kann ein „Online-Assessment“ auch dazu genutzt werden, Informationen zu vermitteln. Dies kann unabhängig von der Testung geschehen, indem z. B. Test- und Informationsblöcke sich abwechseln. Testung und Information können aber auch auf vielfältige Art und Weise miteinander verknüpft werden. So können z. B. über verschiedene Medien (Text, Audio, Video) dargebotene Informationen über eine Organisation oder Position anschließend in einer Art „Wissensquiz“ abgefragt werden. Oder ein „Online-Assessment“ startet direkt (ohne Vorabinformationen) mit Wissensfragen und die Teilnehmer erhalten eine informative Rückmeldung darüber, ob ihre Antworten richtig oder falsch waren. In diesem Sinne funktionieren beispielsweise so genannte „Erwartungschecks“, bei dem die Teilnehmer etwa Aussagen zu der infrage stehenden Berufsausbildung als „richtig“ oder „falsch“ einstufen sollen. Anschließend erhalten sie eine Erklärung zu dem Sachverhalt. Auf diese Art und Weise kann man verbreitete Fehlannahmen zu einer Ausbildung oder einem Beruf (z. B. „Lehrer haben ab mittags frei“) korrigieren. Durch eine derartige Testung wird die Informationsvermittlung interaktiv. Dadurch erhöht sich die Wahrscheinlichkeit, dass Informationen, die auf einer Website oder in einer Broschüre unbeachtet geblieben wären, zur Kenntnis genommen werden. Bei der Information kann es sich auch um eine Information über geeignete Lernwege und -methoden handeln. Im Idealfall beziehen sich z. B. Informationen zu Lernmodulen passgenau auf die vorher durch das Testen diagnostizierten individuellen Wissenslücken und Fähigkeits-/Fertigkeitsdefizite und ermöglichen so ein selbstgesteuertes Lernen.

Ein Vorteil von „Online-Assessments“ mit Informationen ist, dass hier alle für die Teilnehmer interessanten Informationen an einem (virtuellen) Ort gesammelt präsentiert werden. Die Organisation, die das „Online-Assessment“ gestaltet, kann entscheiden, welche Informationen der Teilnehmer erhält. Durch die Interaktivität und Kontrollfragen erhöht sich die Wahrscheinlichkeit, dass die Informationen – anders als Informationen in Broschüren oder auf Websites – auch zur Kenntnis genommen werden. Man kann sogar die Informationsaufnahme durch Testfragen kontrollieren. Auf diese Weise kann eine Organisation auch auf Berufsfelder oder Aspekte der Organisation hinweisen, die eventuell noch unbekannt sind und nach denen der Teilnehmer deshalb auch nicht aktiv suchen würde. Schließlich können auch die Testaufgaben selbst Informationen vermitteln, sodass „Testen“ und „Informieren“ keine Gegensätze sind. Wenn die Testaufgaben nämlich die Aufgaben simulieren, die auch im „wirklichen“ Leben (in der Ausbildung, im Studium, im Beruf) auf Menschen zukommen, so bieten diese Testaufgaben einen „realistic job preview“. Außerdem sind solche Aufgaben „augenscheinvalide“, was sich positiv auf die Akzeptanz auswirkt [30].

Die Unterscheidungsdimension „Testen/Informieren“ ist häufig mit der Unterscheidungsdimension Selbst- und Fremdelektion verbunden. Während sich Tools zur Selbst- und Fremdelektion (Abschn. 10.2.1) in Aufbau und verwendeter Methodik häufig

ähnlich sind, unterscheiden sie sich oft bezüglich des Anteils an Informationen, wobei die Selbstselektion mit einem höheren Informationsanteil einhergeht (Abb. 10.2). Wir benennen die Pole dieser Unterscheidungsdimension hier vereinfachend „Testen und Informieren“, korrekt müssten die Pole aber „ausschließliches Testen“ einerseits und „Kombination von Testen und Informieren“ andererseits genannt werden. Verfahren, die keine Testanteile umfassen und lediglich der Information dienen zählen nicht zur Diagnostik und sollten nicht „Online-Assessment“ genannt werden.

10.2.3 Durchführung mit und ohne „Aufsicht“ – Herausforderungen

Der Begriff des „online“ Testens wird häufig synonym verwendet zu einer Testung „ohne Aufsicht“. „Online“ bedeutet aber lediglich, dass ein Computer, auf dem die Testung stattfindet, sich in einem Netzwerk mit anderen Computern (z. B. via Internet oder Intranet) befindet. Teilnehmer, die einen Test online bearbeiten, können dabei sehr wohl beaufsichtigt werden. Der relevante Einteilungsgesichtspunkt besteht also nicht in der Frage „online“ oder „offline“, sondern in der Frage, ob die Testungen unter „Aufsicht“ oder „ohne Aufsicht“ durchgeführt werden. Die International Test Commission [62] unterscheidet vier Modi der Testadministration, die sich im Ausmaß an Aufsicht bzw. Kontrolle voneinander unterscheiden und auf einer übergeordneten Ebene den Kategorien „unbeaufsichtigt“ und „beaufsichtigt“ zugeordnet werden können (Tab. 10.1).

Im sogenannten „open mode“ liegt am wenigsten Kontrolle vor. Hierzu zählen online frei zugängliche Tools. Der „controlled mode“ sieht, ebenso wie der „open mode“, zwar keine menschlichen Testleiter vor, die Tools sind jedoch insofern zugangsbeschränkt, dass sie beispielsweise nur bestimmten Personen mit speziellen Zugangsdaten für einen bestimmten Zeitraum zugänglich sind. Eine teilweise Kontrolle bzw. starke Kontrolle durch menschliche Testleiter liegt im „supervised (proctored) mode“ bzw. im „managed mode“ vor. Der Unterschied beider Modi liegt darin, dass im „managed mode“ zusätzlich zur Anwesenheit eines menschlichen Testleiters Rahmenbedingungen wie der Ort der Testung von den Testdurchführenden festgelegt werden. Ein Beispiel für den „managed mode“ ist die beaufsichtigte Bearbeitung von computerbasierten Tests in den Räumen der auswählenden Organisation. Die ITC [62] subsumiert die ersten beiden Typen unter unbeaufsichtigte („unproctored“) und die letzten beiden Typen unter beaufsichtigte („proctored“) Testung.

Tab. 10.1 Formen der Testbeaufsichtigung. (Quelle: [62])

Unbeaufsichtigte Testung („unproctored“)	<ul style="list-style-type: none"> • Open mode • Controlled mode
Beaufsichtigte Testung („proctored“)	<ul style="list-style-type: none"> • Supervised mode • Managed mode

Im Kontrast zu klassischen Auswahlverfahren finden „Online-Assessments“ und „Self-Assessments“ häufig in einer unbeaufsichtigten Umgebung (z. B. zu Hause) und ohne menschlichen Testleiter statt (d. h. im obigen Sinne „unproctored“). Daraus resultieren spezifische Herausforderungen für die Selbst- und Fremdselektion. Potenziell verzerrend auf die Ergebnisse kann sich auswirken, dass Teilnehmer ein „Online-Assessment“ unter ungünstigen oder, allgemeiner formuliert, unter ungleichen äußeren oder motivationalen Bedingungen bearbeiten. Hinzu kommt die Möglichkeit des Einsatzes unerlaubter Hilfsmittel, die nur bei beaufsichtigten Testungen vollständig kontrolliert werden kann. Im Folgenden berichten wir empirische Befunde, die auf die Äquivalenz (bzw. Divergenz) der Ergebnisse von beaufsichtigten und unbeaufsichtigten Testungen schließen lassen.

Effekte der unbeaufsichtigten Testdurchführung auf die Testergebnisse

„Online-Assessments“ finden im Internet und häufig ohne Aufsicht statt. In der Literatur wird das unkontrollierte Testen bzw. das Testen ohne Aufsicht mit dem Begriff „unproctored Internet-based testing“ (UIT) bezeichnet [42]. Da die meisten empirischen Untersuchungen zur Validität von Personalauswahlverfahren sich auf das „klassische“, beaufsichtigte Setting beziehen, drängt sich die Frage auf, inwiefern UIT im Rahmen der Personalauswahl äquivalente Ergebnisse liefert. Durch die (relativ) unkontrollierte Testung im Rahmen eines „Online-Assessments“ resultiert eine spezifische Verfälschbarkeit, die die Validität ungünstig beeinflussen kann. In diesem Kontext ist allerdings zu beachten, dass die Intention, positive Ergebnisse vorzutäuschen („faking good“) vor allem bei der Fremdselektion, nicht aber bei der Selbstselektion zu erwarten ist. Eine weitere Verfälschungsmöglichkeit besteht in einem unkontrollierten Ausmaß an „Übung“, indem sich Bewerber unter verschiedenen Identitäten immer wieder Zugang zu einem „Online-Assessment“ verschaffen, bevor sie das Assessment unter dem eigenen Namen bearbeiten. Dies ist besonders problematisch, wenn die Bewerber in stark unterschiedlichem Maße Gelegenheit zur Übung hatten [65]. Natürlich könnten Bewerber auch andere (geeignete) Personen bitten, die Testaufgaben zu lösen.

Inkonsistente Testwerte zwischen UIT und beaufsichtigter Testung stellen einen Indikator für Effekte der Verfälschung dar. Das Ausmaß lässt sich beispielsweise durch wiederholtes Testen quantifizieren. Kantowitz und Dainis [28] analysierten die Daten eines Fähigkeitstests, den über 4000 Bewerber im Rahmen von Bewerbungen bei verschiedenen Organisationen zunächst unbeaufsichtigt und danach beaufsichtigt durchführten. Inkonsistente Testwerte fanden sich bei 6,4 % der Bewerber. Templer und Lange [59] konnten anhand einer Stichprobe potenzieller Bewerber für einen kognitiven Fähigkeitstest sowie mehrere berufsbezogene Persönlichkeitsverfahren keine systematischen Unterschiede zwischen UIT und beaufsichtigter Testung feststellen. Ployhart et al. [51] verglichen bei realen Bewerbern Papier-Bleistift-Testungen mit beaufsichtigten Online-Testungen (jeweils $N > 2300$), wobei ein Persönlichkeitsfragebogen, ein biografischer

Fragebogen und ein Situational-Judgement-Test (SJT) zum Einsatz kamen. Zwar zeigten sich Divergenzen, jedoch sprachen diese aus statistischer und diagnostischer Sicht überwiegend für die Internet-basierte Testung (z. B. höhere Varianzen und interne Konsistenzen). Offen ist allerdings, inwiefern sich das auf „Online-Assessments“ zur Personalauswahl übertragen lässt.

Einen wichtigen Beitrag in diesem Bereich leisteten Beaty et al. [6]. In einer Metaanalyse verglichen sie Validitätskoeffizienten von Auswahlverfahren via UIT gegenüber beaufsichtigten Online-Testungen. Die Autoren stützten sich auf interne Daten eines Personaldienstleisters ($k = 52$), der für Stellen in verschiedenen Branchen und Berufsfeldern Auswahlverfahren durchgeführt hat. Als Prädiktoren wurden die erfasste Gewissenhaftigkeit und Vertriebsbeignung sowie ein biografischer Fragebogen herangezogen, als Kriterien standen verschiedene Indikatoren des Berufserfolgs zur Verfügung. Die ermittelten Unterschiede in den Validitätskoeffizienten waren geringfügig und überwiegend nicht signifikant. Inwiefern sich diese Ergebnisse generalisieren lassen muss noch beantwortet werden, da der Großteil der Studien sich auf Stellen mit eher niedriger Komplexität bezog (z. B. Tätigkeit im Callcenter). Ebenso wurden nur nicht-kognitive Maße untersucht. Zu kognitiven Maßen liegt u. a. eine Studie von Wright et al. [66] vor. Die Autoren verglichen die Daten eines Fähigkeitstests, den Bewerber in beaufsichtigter und unbeaufsichtigter Testung bearbeiteten. In zwei großen Stichproben ($N > 24.000$ bzw. 9000) zeigten sich zwar Testleistungsunterschiede (bezogen auf die theta-Werte) zwischen unbeaufsichtigter und beaufsichtigter Testung, die Effekte waren jedoch klein und nicht systematisch ($d = ,20$ bzw. $-,15$).

In der Literatur [49] wird ein Absinken von Testergebnissen von UIT zu einer nachfolgenden, beaufsichtigten Testung als Indiz für Verfälschung durch die Teilnehmer gewertet. Bedeutsame Leistungsverschlechterungen in diesem Sinne konnten jedoch Untersuchungen von Nye et al. [49] sowie Lievens und Burke [43] nicht bestätigen (stattdessen fanden sich teilweise Leistungsverbesserungen, womöglich aufgrund von Übungs- oder motivationalen Effekten).

Eine andere Frage, die sich bei der unbeaufsichtigten Testung stellt, ist die nach der Qualität des von der Testperson genutzten Gerätes zur Bearbeitung. Sofern keine Aufsicht über die Testung besteht, kann auch nicht kontrolliert werden, welches „Endgerät“ die Teilnehmer für die Testung nutzen (dies wäre nur im „managed mode“ möglich). Eine Herausforderung stellen diesbezüglich die mobilen Endgeräte dar. Arthur et al. [2] analysierten einen sehr großen Datensatz von über 3,5 Mio. Bewerbern, die über UIT Verfahren zur Erfassung der Persönlichkeit sowie kognitiver Fähigkeiten bearbeiteten. Eine Teilstichprobe von gut 69.000 Bewerber bearbeitete die Verfahren über Mobilgeräte. Für diese Gruppe konnte geprüft werden, ob die Testergebnisse in Abhängigkeit vom genutzten Gerät variieren. Die Autoren konnten in Bezug auf die Faktorstruktur der eingesetzten Skalen und die Skaleninterkorrelationen Hinweise auf Messinvarianz bzw. -äquivalenz finden. Bei den Skalenmittelwerten des Persönlichkeitsverfahrens gab es nur

geringfügige Unterschiede, beim Fähigkeitstest hingegen zeigten sich deutlich schlechtere Werte bei der mobilen gegenüber der nicht-mobilen Bearbeitung ($d = ,90$). Morelli et al. [46] führten mit einem großen Datensatz von Bewerbern, die via UIT unter anderem einen kognitiven Fähigkeitstest (siehe dazu Kapitel von Kersting und Palmer in diesem Band), biografische Fragebogen (siehe dazu Kapitel von Strobel in diesem Band) und einen SJT (siehe dazu Kapitel von Franczukowska in diesem Band) bearbeiteten, eine ähnliche Untersuchung durch. Die Messinvarianz konnte für alle Verfahren bestätigt werden, im SJT zeigten sich jedoch schlechtere Leistungen bei der mobilen Durchführung. Arthur et al. [2] äußern die Vermutung, dass eine bessere Darstellung von Verfahren über Smartphone-Applikationen dazu beitragen könnte, derartige Unterschiede abzubauen. Hier ist sicherlich noch weitere Forschung und auch technische Entwicklungsarbeit nötig.

Zur Kontrolle von Verfälschung empfiehlt die ITC [62] im Falle von UIT bei „high-stakes“-Testungen (also z. B. im Auswahlprozess) eine nachfolgende Prüfung unter kontrollierten Bedingungen. Für die Durchführung eines beaufsichtigten Tests nach einem UIT stellen Guo und Drasgow [20] sowie Tendeiro et al. [60] statistische Verfahren vor, die eine systematische Identifikation inkonsistenter Testwerte erlauben.

Darüber hinaus variiert das Verfälschungsrisiko in Abhängigkeit von der Gestaltung des „Online-Assessments“. Die öffentliche Verbreitung von Aufgabenmaterial aus Leistungstests (vgl. hierzu die „Testknacker“-Ratgeberliteratur) oder auch der Einfluss von Testtrainings ist geringer, wenn Items randomisiert aus großen Item-Datenbanken gezogen werden. Allerdings lässt sich von bestimmten Aufgabentypen (z. B. Matrizenaufgaben) leichter eine große Item-Menge erzeugen, als von anderen Aufgabentypen (z. B. Items mit verbalem Inhalt). Es ist außerdem aus Sicht des Teilnehmers wesentlich leichter, sich eine Wissensfrage zu merken als eine Matrizenaufgabe. Dies kann dazu führen, dass bei „Online-Assessments“ und „Self-Assessments“ häufiger auf Aufgaben zum schlussfolgernden Denken mit figuralem Material zurückgegriffen wird. Die Auswahl der Aufgaben muss sich aber aus der Anforderungsanalyse ergeben und nicht aus pragmatischen Gesichtspunkten. Der Hinweis, dass bei Aufgaben zum schlussfolgernden Denken mit figuralem Material ein höherer Testschutz besteht ist unerheblich, wenn der gedankliche Umgang mit figuralem Aufgabenmaterial für die infrage stehende Ausbildung resp. den Beruf irrelevant ist.

10.2.4 Kombination der Gestaltungselemente

Die hier behandelten drei Unterscheidungsdimensionen sind unabhängig voneinander. Diese prinzipielle Kombinierbarkeit von Gestaltungsmerkmalen stellt ein großes Potenzial von „Online-Assessments“ dar und ermöglicht deren Anpassung an den jeweiligen Bedarf (siehe die Anwendungsfelder nach Hiltmann [25]). Ungeachtet dessen sind einige Kombinationen sinnvoller und häufiger als andere. So sind die in Tab. 10.2 aufgeführten Kombinationen für bestimmte Anwendungen besonders gängig.

Tab. 10.2 Gängige Kombinationen von Gestaltungselementen des „Online-Assessments“ für bestimmte Einsatzziele. (Quelle: eigene Darstellung)

	Gestaltung					
	Selektion		Funktionsmechanismus		Darbietungsweise	
Ziel des „Online-Assessment“	Selbst	Fremd	Testen	Informieren	Beaufsichtigt	Unbeaufsichtigt
Personal-Vorauswahl (Screening)		●	●			●
Personal-Endauswahl		●	●		●	
Self-Assessment	●		●	●		●

Aus der Bezeichnung „Online-Assessment“ allein lässt sich weder ableiten, um welche Art der Diagnostik es sich handelt, noch lassen sich Rückschlüsse auf eine hohe oder niedrige Qualität ziehen. Um qualitativ hochwertige „Online-Assessments“ zu erkennen oder zu gestalten, muss man einige Besonderheiten berücksichtigen, die sich aus dem englischen Wort „online“ in der Bezeichnung „Online-Assessment“ ergeben. Eine dieser Besonderheiten ist die Frage nach der Beaufsichtigung der Testung, die wir weiter oben bereits thematisiert haben. In den folgenden Abschnitten gehen wir auf weitere Besonderheiten des „Online-Assessments“ ein.

10.3 Spezifika der Durchführung als Folge des Online-Charakters

Die Gründe, die zur zunehmenden Anwendung von „Online-Assessments“ (und spezifischer von „Self-Assessments“) führen sind vielfältig. Zum einen liegt es an dem immer weiter zunehmenden Stellenwert, den das Internet bei den heutigen potenziellen Bewerbern hat. Bei der Zielgruppe solcher Rekrutierungsverfahren ist das Internet die erste und zumeist auch die einzige Informationsquelle im Berufswahl- oder Bewerbungsprozess, wodurch die Präsenz von Organisationen in diesem Bereich inzwischen unabdingbar ist. In einer Umfrage der Universität Erfurt gaben insgesamt 85 % der Schüler einer gymnasialen Oberstufe an, das Internet genutzt zu haben, um Informationen bezüglich des Studiums oder Berufs zu recherchieren [17]. Das Internet stellt nach dieser Studie bei Gymnasiasten in der Oberstufe die wichtigste Informationsquelle im Berufsorientierungsprozess dar. Es ist davon auszugehen, dass diese Zahl in den vergangenen Jahren noch weiter angestiegen ist.

Zum anderen spielt die Attraktivität eines Arbeitgebers eine wichtige Rolle wenn es darum geht, Bewerber auf die Organisation aufmerksam zu machen. Dabei geht es nicht nur um die Attraktivität der tatsächlichen Position oder Karriere, sondern auch darum, wie attraktiv sich eine Organisation präsentiert, beispielsweise durch ihren Online-Auftritt. Wie zuvor dargestellt, müssen viele Organisationen verstärkt um die Aufmerksamkeit qualifizierter Bewerber kämpfen, wobei ein attraktiver Auftritt eine wichtige Rolle spielt. Viele Organisationen können es sich bei dem momentanen Bewerbermangel nicht mehr leisten, geeignete Bewerber zu übersehen [32]. Die Erweiterung von Fremdselektionsverfahren vom klassischen Papier-Bleistift-Modus auf den Computer und (darüber hinaus) ins Internet sowie die Ergänzung um die Möglichkeit zur Selbstselektion stellen aktuell eine große Herausforderung für die Personalauswahl dar.

Mittlerweile gibt es umfangreiche Forschungsarbeiten zu den Besonderheiten computerbasierter Diagnostik. Dies zu vertiefen würde den Umfang dieses Beitrags übersteigen (siehe dazu z. B. [4, 35]). Stattdessen soll die Aufmerksamkeit auf drei Aspekte gerichtet werden, die sich spezifisch aus dem online und häufig ohne Aufsicht stattfindenden diagnostischen Prozess ergeben: 1) Usability und Attraktivität von Online-Tools, 2) der wahrgenommene Datenschutz und die wahrgenommene Fairness sowie 3) interindividuelle Unterschiede in der Akzeptanz von Online-Tools.

10.3.1 Usability und Attraktivität von Tools

Wer im Internet eine „gute Figur“ abgeben möchte, muss für einen professionellen Internetauftritt sorgen. Ein intuitiv verständliches und optisch attraktives Tool wird Akzeptanz erfahren, optisch unzeitgemäße oder fehleranfällige Tools werden zu Frustration und Ablehnung führen. Dass sich eine ansprechende und benutzerfreundliche Website auch auf Bewerberreaktionen auswirkt konnten Braddy et al. [10] zeigen. Die Wahrnehmung des Unternehmens durch die Teilnehmer war dabei maßgeblich durch die Gestaltung der Website beeinflusst. Während einfach zu navigierende, attraktive Seiten zu positiven Einschätzungen in Bezug auf das Unternehmen führten, hatte der Umstand, ob die einschätzenden Personen das Unternehmen kannten oder nicht nur wenig Einfluss auf die Einschätzung.

Die Gestaltung der Online-Tools spielt auch eine Rolle wenn es darum geht, möglichst viele Bewerber für eine Stelle anzuwerben. Deshalb sind „Online-Assessments“ ungeachtet ihrer diagnostischen Funktion immer auch Instrumente des Personalmarketings. In der Literatur [11, 58] wird die Annahme vertreten, dass Merkmale eines Online-Auswahlsystems, wie es beispielsweise ein „Online-Assessment“ ist, dem Bewerber als Informationsbasis für ein Urteil über die Organisation dienen. Dabei fließen gerade leicht verfügbare Informationen, wie z. B. die Benutzerfreundlichkeit eines entsprechenden Tools, die nicht direkt mit der zu besetzenden Stelle in Bezug stehen, in die Urteilsbildung ein. Sinar et al. [56]

befragten über 23.000 Bewerber, die sich online bei verschiedenen Unternehmen bewarben und dabei auf der Webseite Fragen beantworten mussten (z. B. einen biografischen Fragebogen), die zur Vorselektion genutzt wurden. Es zeigten sich deutliche positive Korrelationen zwischen der Einstellung gegenüber Organisationen, die das Internet zur Personalauswahl nutzen und der subjektiven Geschwindigkeit und Benutzerfreundlichkeit der jeweiligen Webseite ($r = ,39$ bzw. $,41$; partielle Korrelationen, kontrolliert um subjektiven Stellenbezug der Fragen). Braddy et al. [10] konnten ebenfalls positive Effekte der Benutzerfreundlichkeit von Rekrutierungswebseiten mehrerer großer US-Unternehmen auf die Einstellung potenzieller Bewerber gegenüber den jeweiligen Unternehmen demonstrieren.

10.3.2 Wahrgenommener Datenschutz und wahrgenommene Fairness

Die wahrgenommene vertrauliche Behandlung der Daten ist ein entscheidender Punkt für die Akzeptanz eines „Online-Assessments“. Dieser Aspekt ist umso wichtiger, je größer der Testanteil in einem Tool wird. Harris et al. [21] fanden in einer Untersuchung mit belgischen und US-amerikanischen Studierenden eher schwache Zusammenhänge zwischen Bedenken bezüglich Datenschutz und -sicherheit sowie der ablehnenden Haltung gegenüber onlinebasierten Auswahlprozessen. Interessanterweise fand sich aber (bei den US-amerikanischen Studierenden) eine deutliche Korrelation ($r = ,50$) zwischen der ablehnenden Haltung gegenüber onlinebasierten Auswahlprozessen und der Sorge, man könne aufgrund technischer Probleme in einem Online-Auswahlverfahren benachteiligt werden. Bauer et al. [5] konnten sowohl anhand einer Labor- als auch anhand einer Feldstudie zeigen, dass bei Bewerbungen Bedenken bezüglich des Datenschutzes einen negativen Effekt auf die wahrgenommene Fairness des Auswahlprozesses (siehe dazu Kapitel von Kanning in diesem Band) haben, welche wiederum die wahrgenommenen Attraktivität der Organisation sowie die Zuversicht, bei guter Bewerbungsleistung die Stelle zu erhalten, prädiziert. Die Akzeptanz eines Online-Auswahlprozesses hängt nicht nur von äußeren Merkmalen der Webseite, wie der Geschwindigkeit des Online-Systems und dessen Benutzerfreundlichkeit, ab [58], sondern auch von Aspekten wie dem wahrgenommenen Datenschutz, der wahrgenommenen Fairness sowie der Berufsbezogenheit des Inhalts und der wertschätzenden Interaktion des Systems mit dem Bewerber (z. B. keine sittenwidrigen Fragen) (für die letzten beiden Aspekte siehe [37]).

10.3.3 Interindividuelle Unterschiede bezüglich der Akzeptanz von Online-Testungen

Hinsichtlich der Akzeptanz von „Online-Assessments“ ist zu beachten, dass die Reaktionen der Bewerber auch von interindividuell variierenden Merkmalen der Bewerber abhängen. Wiechmann und Ryan [64] führten eine Studie durch, in der Probanden in einer simulierten Auswahl-situation eine Postkorbübung bearbeiten mussten. Die selbst-

eingeschätzte, spezifische Erfahrung (z. B. Programmierkenntnisse) mit Computern hatte einen signifikant positiven Effekt auf die Leistung in der computerbasierten Darbietungsform ($b = ,31$). Die Scheu vor dem Umgang mit Computern hingegen hatte deutliche negative Effekte auf die Akzeptanz des Verfahrens, so auf die erlebte Fairness ($b = -,24$), die Augenscheinvalidität ($b = -,40$) und darauf, wie angenehm der Test empfunden wurde ($b = -,36$). Diese Aspekte sollten berücksichtigt werden, da sie einen Einfluss darauf haben, ob „Online-Assessments“ ihr Ziel erreichen oder nicht. So zeigte sich in der bereits zitierten Untersuchung von Harris et al. [21] – zumindest in der US-amerikanischen Teilstichprobe – ein negativer Zusammenhang zwischen dem selbstberichteten Wissen über das Internet und den Bedenken bezüglich des Datenschutzes bei onlinebasierten Auswahlverfahren ($r = -,25$).

Es ist davon auszugehen, dass sich Bewerber in ihrem Wissen über das Internet und in ihren spezifischen Erfahrungen mit Computern unterscheiden. Daraus erwächst die Gefahr, dass, wenn nicht die Ergebnisse selbst, so doch zumindest die Akzeptanz von „Online-Assessments“ in Abhängigkeit von Merkmalen der Bewerber unterschiedlich ausfallen. Sofern es sich bei der Erfahrung im Umgang mit Computern (oder auch der Scheu davor) um ein eignungsirrelevantes Merkmal handelt, welches die diagnostische Entscheidung (oder Empfehlung) beeinflusst, handelt es sich um einen „bias“ [53], der das Eignungsurteil verfälscht. Es darf allerdings die optimistische Vermutung geäußert werden, dass die Erfahrung im Umgang mit dem Internet und Computern bei den sogenannten „digital natives“ [52], zu denen zumindest ein Großteil der heutigen Berufseinsteiger zählen dürfte, als weitgehend gegeben erachtet und somit vorausgesetzt werden kann. Inwieweit dies bei anderen (z. B. lebensälteren) Bewerbergruppen ebenfalls angenommen werden kann, bedarf weiterer Forschungsaktivität. Untersuchungswert sind auch geschlechtsspezifische Erfahrungsunterschiede mit dem Internet sowie Unterschiede in der Einstellung gegenüber Computern und ihre Auswirkungen auf die Akzeptanz.

10.4 Bewertung der Qualität von „Online-Assessments“ und „Self-Assessments“

Neben den bereits behandelten Aspekten der Usability, des wahrgenommenen Datenschutzes und der wahrgenommenen Fairness kommt bei „Online-Assessments“ mit Testelementen auch und vor allem der Qualität dieser Testung eine hohe Bedeutung zu. Die Bewertung der Qualität der Testung kann im Sinne des Testbeurteilungssystems des Testkuratoriums erfolgen (TBS-TK, Testkuratorium der Föderation Deutscher Psychologinnenvereinigungen [61]). Das TBS-TK-System berücksichtigt bereits einige Qualitätsforderungen der DIN 33430, einer Norm zur berufsbezogenen Eignungsbeurteilung [16, 29]. Diese Norm formuliert aber auch Anforderungen an die Eignungsdiagnostik, die über die Qualität des Tests (siehe dazu Kapitel von Melchers in diesem Band) hinausgehen und bei der Bewertung von „Online-Assessments“ ebenfalls zu beachten sind. Hierzu zählt beispielsweise die Forderung nach einem eindeutigen Anforderungsbezug

oder nach einer regelgeleiteten Durchführung und Auswertung der Verfahren sowie einer regelgeleiteten Interpretation der Verfahrensergebnisse sowie Urteilsbildung.

Weitere Qualitätsstandards beziehen sich spezifisch auf computergestütztes Testen und thematisieren etwa den Datenschutz, die technischen Voraussetzungen und die Frage, wer eigentlich z. B. für „automatisch“ erstellte Diagnosen verantwortlich ist. Einschlägig sind hier beispielsweise die „International Guidelines on Computer-Based and Internet-Delivered Testing“ der International Test Commission [62] oder die „Guidelines for computer-based testing“, die von der Association of Test Publishers [3] herausgegeben wurden. Wir greifen hier aus der Vielzahl der Qualitätsaspekte für „Online-Assessments“ vier Teilbereiche heraus: die Validität, den Nutzen, die Akzeptanz sowie die Qualität der Rückmeldung, die die Teilnehmer erhalten.

Validität

Die Validität von Entscheidungen, die aus diagnostischen Verfahren abgeleitet wurden, ist in Bezug auf die Vorhersage des beruflichen Erfolgs vielfach belegt worden [55]. Auch die Vorhersagekraft von „Self-Assessments“ im Rahmen der Studienwahl für Studiennoten und Studienzufriedenheit ist durch eine wachsende Zahl an Untersuchungen belegt (eine Übersicht liefert [22]). Bislang liegen allerdings für „Online-Assessments“ in der Personalauswahl – anders als für den traditionellen Einsatz von Personalauswahlverfahren – noch keine umfangreichen Validitätsdaten vor. Die Validität wurde bislang vor allem bezüglich einzelner Tests überprüft, kaum jedoch für gesamte „Online-Assessments“. Die weiter oben geschilderten Befunde zur Testung ohne Aufsicht erlauben einen vorsichtigen Optimismus und geben Anlass zu der Vermutung, dass die Validität weniger eine Frage des Darbietungsmediums (Papier oder Bildschirm) ist, sondern – nicht überraschend – eine Frage der Qualität der diagnostischen Verfahren, z. B. des Fragebogens und Tests, selbst.

Nutzen

Den Nutzen von „Online-Assessments“ können zwei Praxisbeispiele aus dem Personalbereich veranschaulichen. In Bezug auf die „Online-Assessments“ der Deutschen Luftwaffe AG berichtet Wiedmann [65], dass sich in einem mehrstufigen Selektionsprozess durch die Vorselektion aufgrund der „Online-Assessments“ eine bessere Basisrate bei späteren Stufen im Auswahlprozess erzielen lässt. Dies führt zu deutlichen Kosteneinsparungen, da sich der Anteil der zu beurteilenden Bewerber in späteren (meist teureren) Auswahlstufen (z. B. Assessment-Center) erheblich reduziert. Wurow [67] beschreibt ein „Online-Assessment“ der Conergy AG, das zur internen Stellenbesetzung eingesetzt wird. Auf Basis eines darin administrierten verhaltensbasierten Fragebogens werden von den Bewerbern Profile ermittelt. Diese Profile werden in späteren Auswahlritten, wie strukturierten Interviews, aufgegriffen. Es wird von einer erhöhten Validität und reduzierten Kosten berichtet.

Während die beiden Beispiele den Nutzen aus der Perspektive der Organisation betrachten, zeigt die folgende Studie des „Centre of Human Resources Information Systems“ (CHRIS) der Universitäten Frankfurt am Main und Bamberg den subjektiven Nutzen von

„Self-Assessments“ im Bewerbungsprozess aus der Sicht der Nutzer (sprich den Bewerbern selbst) auf. Für ein „Self-Assessment“ ist zu verlangen, dass die Person nach der Teilnahme mehr über ihre eigene Passung zu den Anforderungen weiß als zuvor und sich zu einer fundierten Entscheidung für oder gegen eine Bewerbung befähigt fühlt. Zwischen 2008 bis 2012 wurden für die CHRIS-Studie über 26.000 (potenzielle) Bewerber u. a. nach dem subjektiven Nutzen von „Self-Assessments“ befragt [18]. Das Urteil über „Self-Assessments“ fiel überwiegend positiv aus. So gaben (für das Befragungsjahr 2008) jeweils über 40 % der Befragten an, ein „Self-Assessment“ biete eine gute Rückmeldung über die eigene Eignung für eine Position und erhöhe die Effektivität und Qualität der Bewerbung. Immerhin 33 % gaben an, dass „Self-Assessments“ die Bewerbungsentscheidung erleichtern. Kritisch zu werten ist das mangelhafte Vertrauen in den Datenschutz sowie in die Aussagekraft. Nur etwa 30 % der Befragten waren überzeugt, dass die Daten nicht an Dritte weitergeleitet würden und nur etwa ein Fünftel äußerte Vertrauen in das Ergebnis eines „Self-Assessments“.

Soziale Akzeptanz

Auf die hohe Bedeutung der sozialen Akzeptanz sind wir weiter oben bereits eingegangen. „Online-Assessments“ die unbeaufsichtigt dargeboten werden, sind aufgrund der allgemeinen Zugänglichkeit ein mediales Aushängeschild einer Organisation und beeinflussen somit nicht nur die Selbst- oder Fremdselektion, sondern auch das Ansehen jener Organisationen, welche sie einsetzen. Eine mangelnde Akzeptanz von Personalauswahlverfahren kann negative Konsequenzen für die auswählende Organisation und deren Reputation haben [23, 30]. Die Bewertung der Qualität von „Online-Assessments“ muss sich also nicht nur mit den Hauptgütekriterien, sondern auch mit dem Nebengütekriterium der sozialen Akzeptanz auseinandersetzen.

Rückmeldung

Mit den Reaktionen der Bewerber ist ein weiteres, spezifisches Qualitätsmerkmal von „Online-Assessments“ assoziiert, nämlich die Qualität der automatisiert erstellten Rückmeldung. Ilgen et al. [26] stellten in einer Literaturübersicht ein Modell vor, das eine Prozesskette zwischen Feedback und der Reaktion von Personen auf das Feedback beschreibt. Große Bedeutung kommt hier der Akzeptanz des Feedbacks zu, also dem subjektiven Urteil über die Richtigkeit der im Feedback enthaltenen Aussagen über die eigene Person. Nach den von Ilgen et al. [26] zusammengetragenen Befunden neigen Menschen dazu, positives Feedback eher zu akzeptieren als negatives. Dies kann durch das Selbstaufwertungsmotiv („self-enhancement“) erklärt werden [7, 47]. Anseel und Lievens [1] demonstrierten dies qua manipulierter Passungsrückmeldungen auf Basis eines Persönlichkeitsfragebogens: Der (positive) Zusammenhang zwischen der Rückmeldung (niedrige vs. hohe Passung) und der Einstellung gegenüber der auswählenden Organisation wurde durch die Akzeptanz des Feedbacks mediiert. Beim Einsatz von Persönlichkeitsskalen ist zusätzlich zu beachten, dass Menschen dazu tendieren, allgemeingültige und unkonkrete Beschreibungen ihrer eigenen Person als zutreffend zu erachten (der sog. Barnum-Forer-

Effekt [19, 33, 45]). Auch das Selbstbestätigungsmotiv („self-verification“ [47]) kann einer wirksamen Selbstselektion entgegenstehen, beispielsweise wenn eine negative Passungsrückmeldung mit stark ausgeprägten Ausbildungs- oder Berufswünschen kollidiert. In einem Versuch, diese sich widersprechenden Motive zu integrieren, postulierten Swann und Schroeder [57], dass mit zunehmender Verarbeitungstiefe – welche wiederum abhängt von aktueller Motivation und verfügbaren kognitiven Ressourcen – verstärkt die Plausibilität von selbstbezogenen Informationen überprüft wird (Selbsteinschätzungsmotiv oder auch „self-assessment“¹ [47]). Feedbacks in einem „Self-Assessment“ sollten möglichst konkret und anforderungsbezogen sein und darüber hinaus die Bewerber zu einer sorgfältigen Verarbeitung der Rückmeldungsinformationen ermutigen [1].

Diese Verarbeitung ist stets risikobehaftet. Aufgrund der automatisierten Befundung besteht keine Möglichkeit, Missverständnisse oder Fehlinterpretationen aufzufangen. Dieser Tatsache ist durch eine überaus sorgfältige und zurückhaltende Gestaltung der Feedbacks Rechnung zu tragen. So sollte bei Testungen ohne Aufsicht beispielsweise darauf hingewiesen werden, dass die erzielten Ergebnisse auch davon abhängen, unter welchen äußeren und motivationalen Bedingungen die Tests bearbeitet wurden. Darüber hinaus sollte bei „Self-Assessments“ darauf hingewiesen werden, dass neben dem Verfahrensergebnis auch weitere Informationen, wie z. B. die Schulnoten sowie weitere Beratungen (z. B. durch Eltern und Freunde) bei der Entscheidung (d. h. der Selbstselektion) berücksichtigt werden sollten. Sinnvoll ist auch der Hinweis auf konkrete externe Beratungsangebote. Die Teilnehmer sollten bei der Auseinandersetzung mit den Feedbacks in ihrer Reflexion geleitet und dazu angehalten werden, die Konsistenz oder Inkonsistenz verschiedener Informationen (Ergebnis im „Self-Assessment“, Schulnoten, Fremdeinschätzungen von z. B. Eltern, Freunden usw.) zu berücksichtigen. Insgesamt ist die Wirkung der aus „Online-Assessments“ abgeleiteten automatischen Feedbacks bislang zu wenig erforscht.

10.5 Fazit

10.5.1 Was bei der Konstruktion und dem Einsatz von „Online-Assessments“ und „Self-Assessments“ zu beachten ist

„Online-Assessments“ und „Self-Assessments“ bieten ein großes Potenzial für die Personalauswahl. Sie können, wie eingangs bereits erwähnt, orts- und zeitunabhängig bearbeitet werden, ermöglichen automatisierte und damit in hohem Maße objektive Testauswertungen und -rückmeldungen und bieten einen großen Spielraum in der Umsetzung von Testformaten [54]. Wie in Abschn. 1.2 dargelegt wurde, lassen sich Elemente von Selbst- und Fremdselektion kombinieren, sodass „Online-Assessments“ sowohl

¹Der Begriff ist in dieser Verwendung nicht zu verwechseln mit dem in diesem Beitrag beschriebenen Online-Tool.

in Funktion als auch bezüglich des Inhalts (Testen vs. Informieren) an die jeweiligen Bedürfnisse der Situation angepasst werden können. Angesichts der weiter wachsenden Bedeutung des Internets für alle Lebensbereiche ist davon auszugehen, dass der Einsatz von „Online-Assessments“ für Personalauswahl und -marketing noch zunehmen wird. Sowohl Bewerber als auch Organisationen können davon profitieren. „Online-Assessments“ und speziell „Self-Assessments“ können jedoch nur dann ihre Ziele erreichen, wenn bestimmte Anforderungen erfüllt sind. Diese lassen sich in drei Felder gliedern: Anforderungen an die eingesetzten Verfahren, an die Technik und den Datenschutz sowie an die sinnvolle Integration in den diagnostischen Prozess.

Die in „Online-Assessments“ und „Self-Assessments“ eingesetzten diagnostischen Verfahren müssen den Anforderungen der Norm DIN 33430 für die berufsbezogene Eignungsbeurteilung [16, 29] sowie des TBS-TK [61] und den weiter oben genannten computerspezifischen Qualitätsstandards genügen [3, 62]. Hervorzuheben ist die Forderung nach einem eindeutigen Anforderungsbezug sowie die Forderung nach möglichst hoher Objektivität, Reliabilität und Validität.

Ebenso wichtig ist die professionelle und bedarfsgerechte technische Umsetzung des „Online-Assessments“. Laier [41] führt als Maßstab für die technische Qualität von Online-Tests und damit auch von „Online-Assessments“ und „Self-Assessments“ die Norm DIN EN ISO 9241-110 [15] zu Grundsätzen der Dialoggestaltung auf und stellt vier zentrale Aspekte heraus, die erfüllt werden müssen: 1) die Aufgabenangemessenheit, also die effektive und effiziente Bedienbarkeit durch den Benutzer, 2) die Steuerbarkeit des Prozesses, 3) die Erwartungskonformität, also eine erkennbare logische Struktur des Tools, und schließlich 4) die Fehlertoleranz. Wie bereits erwähnt spielt der Datenschutz auch eine zentrale Rolle für die Akzeptanz der Bewerber. Organisationen sollten hier in jedem Fall die jeweiligen gesetzlichen Vorgaben einhalten (in Deutschland beispielsweise das Bundesdatenschutzgesetz [BDSG]).

Schließlich sind die Besonderheiten der onlinebasierten Diagnostik zu berücksichtigen. Spezielle Formen des „Online-Assessments“ finden in einem weitgehend unkontrollierten Setting statt. Da es hierbei keinen Testleiter gibt, der prüfen kann, ob ein Bewerber Hilfsmittel nutzt oder sich auf andere Weise Vorteile verschafft, bestehen verschiedene Risiken. Hierzu zählen die Verfälschung, Übungseffekte, die Verbreitung von Aufgabenmaterial und -lösungen sowie das Risiko, dass die Assessments mit ungeeigneten Endgeräten bearbeitet werden. In der Regel sollten „Online-Assessments“ ohne Aufsicht daher nur zur Vorselektion von Bewerbern eingesetzt werden (siehe hierzu [40]) (Tab. 10.3).

Ansätze, die zuerst genannten drei Problematiken zu entschärfen sind beispielsweise die randomisierte Darbietung von Items aus großen Itemdatenbanken [54], die Testwiederholung unter kontrollierten Bedingungen bzw. die entsprechende Ankündigung der selbigen sowie die Aufklärung über Konsequenzen von Verfälschung [62, 65].

- ▶ Falls die Teilnehmer die Testung unbeaufsichtigt absolviert haben, ist dies bei der Gestaltung des Feedbacks zu berücksichtigen. Es ist denkbar, dass die Teilnehmer das „Online-Assessment“ unter ungünstigen äußeren oder motivationalen

Tab. 10.3 Checkliste zu ausgewählten Anforderungen an „Online-Assessments“. (Quelle: eigene Darstellung sowie [41])

Sind die Anforderungen an diagnostische Verfahren gemäß TBS-TK und DIN 33430 erfüllt?	Insbesondere <ul style="list-style-type: none"> • Anforderungsbezug • Objektivität, Reliabilität und Validität • Prüfung der Gültigkeit der Normwerte • Akzeptanz
Ist das „Online-Assessment“ technisch bedarfsgerecht umgesetzt gemäß DIN EN ISO 9241-110?	Insbesondere <ul style="list-style-type: none"> • Bedienbarkeit • Steuerbarkeit • Erwartungskonformität • Fehlertoleranz
Ist der Datenschutz gewährleistet nach dem Bundesdatenschutzgesetz (BDSG)? Sind folgende Datenschutzerfordernungen erfüllt?	<ul style="list-style-type: none"> • Datenvermeidung und -sparsamkeit • Erfassung nur von erforderlichen Daten • Zweckbindung der erfassten Daten • Transparenz bezüglich der Verwendung der Daten
Wie ist das „Online-Assessment“ in den diagnostischen Prozess integriert?	<ul style="list-style-type: none"> • Fokus auf Selbst- vs. Fremdselektion • Wird das „Online-(Self-)Assessment“ zur Vorauswahl genutzt? • Ist die Rückmeldung konkret, weist sie Handlungsempfehlungen auf? • Vorkehrungen gegen und Kontrolle von Verfälschungstendenzen

Bedingungen bearbeiten, was die Ergebnisse verzerrt. Auf diesen Aspekt ist bei der Ergebnismeldung hinzuweisen. Außerdem besteht die Gefahr, dass Teilnehmer die „Diagnose“ überbewerten. Diesem Risiko ist durch eine sorgfältige und zurückhaltende Gestaltung der Feedbacks Rechnung zu tragen.

10.5.2 Forschungsdesiderata

Zukünftige Forschungsarbeiten sollten sich u. a. damit beschäftigen, unter welchen Bedingungen eine erfolgreiche Selbstselektion stattfinden kann. Dies betrifft die grundlegende Frage, ob Bewerber sich in Online-Tools zur Selbstselektion anders verhalten als in jenen zur Fremdselektion. Bewerber als „sich Bewerbende“ möchten naturgemäß einen positiven Eindruck bei einem potenziellen Arbeitgeber hinterlassen. Eine „erfolgreiche“ Selbstselektion, also eine die eigene Passung angemessen reflektierende Bewerbungsentscheidung, stützt sich indes auf eine zutreffende Selbsteinschätzung.

Ausgehend von Paulhus [50] ist sozial erwünschtes Antwortverhalten in zwei verschiedene Faktoren differenzierbar: Selbsttäuschung („self deception“) betrifft Selbstbeschreibungen, die unangemessen positiv sind, wobei die Person jedoch von der Richtigkeit der Selbsteinschätzung überzeugt ist. „Impression management“ bedeutet demgegenüber das

(gezielte) Bestreben, ein positives Bild von der eigenen Person nach außen zu transportieren. Während bei der Fremdselektion insbesondere mit „impression management“ seitens der Bewerber zu rechnen ist, spielt bei der Selbstselektion die Selbsttäuschung eine größere Rolle.

Zum Phänomen des „impression management“, insbesondere zur aktiven Verfälschung („faking“) sowie möglicher Gegenmaßnahmen liegt inzwischen umfangreiche Forschungsarbeit vor. Metaanalytisch zeigen sich beispielsweise unterschiedliche Mittelwerte in Skalen von Persönlichkeitsfragebogen, sowohl unter experimenteller Instruktion von Verfälschung („fake good“ bzw. „fake bad“ [63]) als auch beim Vergleich von Bewerberstichproben gegenüber Nicht-Bewerberstichproben [8]. „Impression management“ kann bei der Personalauswahl problematisch sein [27] und stellt zweifellos eine Herausforderung für die Fremdselektion dar.

Unklar ist aber, in welchem Ausmaß „impression management“ – und insbesondere auch „self deception“ – bei der Selbstselektion wirksam ist. Marcus [44] schlägt ein Prozessmodell vor, in dem die positive Selbstdarstellung in der Personalauswahl als Funktion u. a. von der Attraktivität einer Position, der Fähigkeit zur Selbstdarstellung und der subjektiven Instrumentalität der Selbstdarstellung verstanden wird. Dieser Prozess beginnt nach der Bewerbungsentscheidung – die Selbstselektion ist an dieser Stelle bereits abgeschlossen. Es ist möglich, dass „impression management“ und „self-deception“ insofern die Selbstselektion beeinträchtigen, indem sie eine valide Selbsteinschätzung behindern. Eine Beurteilung der eigenen Passung enthält Beurteilungen über die eigene Person oder, allgemeiner gesprochen, die eigene Identität oder das eigene Selbstkonzept. Bozeman und Kacmar [9] beschreiben ein Modell zur Erklärung von „impression management“ in Organisationen. Eine Implikation des Modells ist, dass beim Vergleich der aktuellen und der erwünschten sozialen Identität ein Bedarf besteht, die eigene Identität anzupassen, zu schützen oder auch aufrecht zu erhalten. Dieser Bedarf lässt sich durch „impression management“ decken. In dem Maße wie diese Prozesse am Werk sind wird sich die Selbstselektion zu einem großen Teil nur auf eine verzerrte Selbsteinschätzung stützen können.

Während zum Thema „faking“ und möglicher Gegenmaßnahmen in der Fremdselektion zahlreiche empirische Untersuchungen vorliegen, besteht bezüglich des Selbstkonzepts einerseits und der Selbstselektion andererseits noch Forschungsbedarf. Damit zusammen hängend ist zu klären, ob Bewerbern auch bei der Selbstselektion dieselbe grundlegende Motivation zur positiven Selbstdarstellung unterstellt werden kann wie bei der Fremdselektion – oder ob letzteres Phänomen erst dann auftritt, wenn der Rubikon überschritten wurde – also ein Bewerber die Bewerbungsentscheidung getroffen hat.

Die Wirkung von „Self-Assessments“ wird – wie oben dargestellt – wesentlich von der Gestaltung des Feedbacks beeinflusst – über entsprechende Erfolgsfaktoren und Risiken bei der Gestaltung ist aber nur wenig bekannt. Eine Forschungsfrage ist schließlich auch, wer freiwillige „Self-Assessment“ Angebote nutzt. Es ist zu befürchten, dass insbesondere solche Personen das Angebot nutzen, die ohnehin dazu neigen, ihre Entscheidung gut abzusichern (Matthäus-Effekt). Diesbezüglich müssten Strategien zur Steigerung der Anwendungshäufigkeit formuliert und überprüft werden.

10.5.3 Schlusswort

„Online-Assessments“ bergen große Potenziale, weisen aber auch Risiken auf. Das gleiche gilt für „Self-Assessments“ als Sonderform des „Online-Assessments“. Eine erfolgreiche, onlinebasierte Selbstselektion kann die Personalauswahl verbessern und gleichzeitig zu einer positiven Darstellung der Organisation und der zu besetzenden Position und deren Anforderungen beitragen. Die weitere Forschung hierzu kann und sollte von der wachsenden Erfahrung des Praxiseinsatzes von „Online-Assessments“ und „Self-Assessments“ sowohl in der Personalauswahl als auch in Berufs- und Studienorientierung profitieren (Tab. 10.3).

Fragen

1. Worin unterscheiden sich „Self-Assessments“ von „Online-Assessments“?
2. Nach welchen Unterscheidungsdimensionen lassen sich „Online-Assessments“ beschreiben?
3. Welche Vor- und Nachteile bietet der Einsatz von „Online-Assessments“?
4. Welche Möglichkeiten gibt es, die Gefahr von Verfälschung bei „Online-Assessments“ bzw. „Self-Assessments“ zu begrenzen?
5. Welche Prozesse können einer erwünschten Selbstselektion entgegenwirken?
6. Was ist bei der Gestaltung automatisierter Rückmeldungen in „Online-Assessments“ und speziell in „Self-Assessments“ zu beachten?

Literatur

1. Anseel, F., & Lievens, F. (2009). The mediating role of feedback acceptance in the relationship between feedback and attitudinal and performance outcomes. *International Journal of Selection & Assessment*, 17(4), 362–376.
2. Arthur, W., Doverspike, D., Muñoz, G. J., et al. (2014). The use of mobile devices in high-stakes remotely delivered assessments and testing. *International Journal of Selection & Assessment*, 22(2), 113–123.
3. Association of Test Publishers (ATP). (2002). *Guidelines for computer-based testing*. Washington: ATP.
4. Bartram, D., & Hambleton, R. K. (Hrsg.). (2005). *Computer-based testing and the internet*. West Sussex: Wiley.
5. Bauer, T. N., Truxillo, D. M., Tucker, J. S., et al. (2006). Selection in the information age: The impact of privacy concerns and computer experience on applicant reactions. *Journal of Management*, 32(5), 601–621.
6. Beaty, J. C., Nye, C. D., Borneman, M. J., et al. (2011). Proctored versus unproctored internet tests: Are unproctored noncognitive tests as predictive of job performance? *International Journal of Selection & Assessment*, 19(1), 1–10.
7. Bell, S. T., & Arthur, W. (2008). Feedback acceptance in developmental assessment-centers: The role of feedback message, participant personality, and affective response to the feedback session. *Journal of Organizational Behavior*, 29(5), 681–703.

8. Birkeland, S. A., Manson, T. M., Kisamore, J. L., et al. (2006). A Meta-Analytic investigation of job applicant faking on personality measures. *International Journal of Selection & Assessment*, 14(4), 317–335.
9. Bozeman, D. P., & Kacmar, K. (1997). A cybernetic model of impression management processes in organizations. *Organizational Behavior and Human Decision Processes*, 69(1), 9–30.
10. Braddy, P. W., Meade, A. W., & Kroustalis, C. M. (2008). Online recruiting: The effects of organizational familiarity, website usability, and website attractiveness on viewers' impressions of organizations. *Computers in Human Behavior*, 24(6), 2992–3001.
11. Cober, R. T., Brown, D. J., Levy, P. E., et al. (2003). Organizational web sites: Web site content and style as determinants of organizational attraction. *International Journal of Selection & Assessment*, 11(2–3), 158–169.
12. Cronbach, L. J., & Gleser, G. C. (1965). *Psychological tests and personnel decisions* (2. Aufl.). Urbana: University of Illinois Press.
13. Diercks, J., & Kupka, K. (Hrsg.). (2013a). *Recrutainment*. Wiesbaden: Springer Fachmedien.
14. Diercks, J., & Kupka, K. (2013b). Recrutainment – Bedeutung, Einflussfaktoren und Begriffsbestimmung. In J. Diercks & K. Kupka (Hrsg.), *Recrutainment* (S. 1–18). Wiesbaden: Springer Fachmedien.
15. DIN EN ISO 9241-110. (2006). *Ergonomie der Mensch-System-Interaktion – Teil 110: Grundsätze der Dialoggestaltung*. Berlin: Beuth.
16. DIN 33430. (2016). *Anforderungen an berufsbezogene Eignungsdiagnostik*. Berlin: Beuth.
17. Driesel-Lange, K., & Hany, E. (2005). Berufsorientierung am Ende des Gymnasiums: Die Qual der Wahl. In B. Kracke & E. Hany (Hrsg.), *Schriften zur Berufsorientierungsforschung* (Bd. 1). Erfurt: Universität Erfurt.
18. Eckhardt, A., Laumer, S., & Vornewald, K. (2013). Bewertung von Self- und E-Assessments durch Kandidaten und Unternehmen. In J. Diercks & K. Kupka (Hrsg.), *Recrutainment* (S. 19–32). Wiesbaden: Springer Fachmedien.
19. Forer, B. R. (1949). The fallacy of personal validation: A classroom demonstration of gullibility. *The Journal of Abnormal and Social Psychology*, 44(1), 118–123.
20. Guo, J., & Drasgow, F. (2010). Identifying cheating on unproctored internet tests: The z-test and the likelihood ratio test. *International Journal of Selection & Assessment*, 18(4), 351–364.
21. Harris, M. M., Hoye, G. van, & Lievens, F. (2003). Privacy and attitudes towards internet-based selection systems: A cross-cultural comparison. *International Journal of Selection & Assessment*, 11(2–3), 230–236.
22. Hasenberg, S. (2012). *Zur prädiktiven Validität von Self-Assessments für die Studienzufriedenheit*. Unveröffentlichte Dissertation, Philipps-Universität Marburg, Marburg.
23. Hausknecht, J. P., Day, D. V., & Thomas, S. C. (2004). Applicant reactions to selection procedures: An updated model and meta-analysis. *Personnel Psychology*, 57(3), 639–683.
24. Hell, B. (2009). Selbsttests zur Studienorientierung: nützliche Vielfalt oder unnützer Wildwuchs? In G. Rudinger (Hrsg.), *Self-Assessment an Hochschulen. Von der Studienfachwahl zur Profilbildung* (S. 9–20). Bonn: Bonn University Press.
25. Hiltmann, M. (2013). Online-Self-Assessments: Ein Impuls zur persönlichen und beruflichen Weiterentwicklung. *Wirtschaftspsychologie*, 1, 72–80.
26. Ilgen, D. R., Fisher, C. D., & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology*, 64(4), 349–371.
27. Kanning, U. P. (2003). Sieben Anmerkungen zum Problem der Selbstdarstellung in der Personalauswahl. *Zeitschrift für Personalpsychologie*, 2(4), 193–195.
28. Kantrowitz, T. M., & Dainis, A. M. (2014). How secure are unproctored pre-employment tests? Analysis of inconsistent test scores. *Journal of Business and Psychology*, 29(4), 605–616.

29. Kersting, M. (2008a). *Qualität in der Diagnostik und Personalauswahl – der DIN-Ansatz*. Göttingen: Hogrefe.
30. Kersting, M. (2008b). Zur Akzeptanz von Intelligenz- und Leistungstests. *Report Psychologie*, 33(9), 420–433.
31. Kersting, M. (2011a). *Personalgewinnung und Personalauswahl im öffentlichen Dienst: Potpourri der Methoden*, 2011(1), 10–12.
32. Kersting, M. (2011b). Schlaraffenland ist abgebrannt – warum sich die Bewerberauswahl ändern muss. *Wirtschaftspsychologie aktuell*, 4, 23–26.
33. Kersting, M. (2014). Zirkustricks entlarven. Qualität von Persönlichkeitsfragebogen. *Personalmagazin*, 1(14), 28–32.
34. Kirbach, C., & Wottawa, H. (2009). Online-Tools zur Gewinnung passender Mitarbeiter. In H. Steiner (Hrsg.), *Online-Assessments. Grundlagen und Anwendung von Online-Tests in der Unternehmenspraxis* (S. 68–82). Heidelberg: Springer.
35. Klinck, D. (2002). *Computergestützte Diagnostik. Beeinflusst das Medium der Testbearbeitung die Testcharakteristika, die Testfairness oder das Erleben der Testsituation?*. Göttingen: Hogrefe.
36. Konradt, U., & Sarges, W. (2003). *E-Recruitment und E-Assessment: Rekrutierung, Auswahl und Beurteilung von Personal im Inter- und Intranet*. Göttingen: Hogrefe.
37. Konradt, U., Warszta, T., & Ellwart, T. (2013). Fairness perceptions in web-based selection: Impact on applicants' pursuit intentions, recommendation intentions, and intentions to reapply. *International Journal of Selection & Assessment*, 21(2), 155–169.
38. Kubinger, K. D. (2009). *Psychologische Diagnostik. Theorie und Praxis psychologischen Diagnostizierens* (2., überarb. und erw. Aufl.). Göttingen: Hogrefe.
39. Kubinger, K. D., Frebort, M., Khorramdel, et al. (2013). Prinzipien und Verfahren der Self-Assessments vom „Wiener Autorenkollektiv Studienberatungstests“. *Wirtschaftspsychologie*, 1(1), 17–24.
40. Kubinger, K. D. (2013). Self-Assessment im Rahmen der Personalauswahl: Möglichkeiten und Grenzen. *Wirtschaftspsychologie*, 1, 81–87.
41. Laier, W. (2009). Einsatz von Online-Tests aus technischer Sicht. In H. Steiner (Hrsg.), *Online-Assessments. Grundlagen und Anwendung von Online-Tests in der Unternehmenspraxis* (S. 37–54). Heidelberg: Springer.
42. Landers, R. N., & Sackett, P. R. (2012). Offsetting performance losses due to cheating in unproctored internet-based testing by increasing the applicant pool. *International Journal of Selection & Assessment*, 20(2), 220–228.
43. Lievens, F., & Burke, E. (2011). Dealing with the threats inherent in unproctored internet testing of cognitive ability: Results from a large-scale operational test program. *Journal of Occupational and Organizational Psychology*, 84(4), 817–824.
44. Marcus, B. (2009). 'Faking' from the applicant's perspective: A theory of self-presentation in personnel selection settings. *International Journal of Selection & Assessment*, 17(4), 417–430.
45. Meehl, P. E. (1956). Wanted – a good cook-book. *American Psychologist*, 11(6), 263–272.
46. Morelli, N. A., Mahan, R. P., & Illingworth, A. J. (2014). Establishing the measurement equivalence of online selection assessments delivered on mobile versus nonmobile devices. *International Journal of Selection & Assessment*, 22(2), 124–138.
47. Morf, C. C., & Koole, S. L. (2014). Das Selbst. In K. Jonas, W. Stroebe, & M. Hewstone (Hrsg.), *Sozialpsychologie. Eine Einführung* (6., komplett neu verfasste und aktual. Aufl., S. 141–195). Berlin: Springer.
48. Nerdinger, F. W. (2014). Gravitation und organisationale Sozialisation. In F. W. Nerdinger, G. Blickle, & N. Schaper (Hrsg.), *Arbeits- und Organisationspsychologie* (S. 71–82). Berlin: Springer.

49. Nye, C. D., Do, B., Drasgow, F., et al. (2008). Two-step testing in employee selection: Is score inflation a problem? *International Journal of Selection & Assessment*, 16(2), 112–120.
50. Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46(3), 598–609.
51. Ployhart, R. E., Weekley, J. A., Holtz, B. C., et al. (2003). Web-based and paper-and-pencil testing of applicants in a proctored setting: Are personality, biodata, and situational judgement tests comparable? *Personnel Psychology*, 56(3), 733–752.
52. Prensky, M. (2001). Digital natives, Digital immigrants Part 1. *On the Horizon*, 9(5), 1–6.
53. Rudolph, C. W., Wells, C. L., Weller, M. D., et al. (2009). A meta-analysis of empirical studies of weight-based bias in the workplace. *Journal of Vocational Behavior*, 74(1), 1–10.
54. Schaper, N. (2009). Online-Tests aus diagnostisch-methodischer Sicht. In H. Steiner (Hrsg.), *Online-Assessments. Grundlagen und Anwendung von Online-Tests in der Unternehmenspraxis* (S. 17–36). Heidelberg: Springer.
55. Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124(2), 262–274.
56. Sinar, E. F., Reynolds, D. H., & Paquet, S. L. (2003). Nothing but ‘net? Corporate image and web-based testing. *International Journal of Selection & Assessment*, 11(2–3), 150–157.
57. Swann, W. B., & Schroeder, D. G. (1995). The search for beauty and truth: A framework for understanding reactions to evaluations. *Personality and Social Psychology Bulletin*, 21(12), 1307–1318.
58. Sylva, H., & Mol, S. T. (2009). E-Recruitment: A study into applicant perceptions of an online application system. *International Journal of Selection & Assessment*, 17(3), 311–323.
59. Templer, K. J., & Lange, S. R. (2008). Internet testing: Equivalence between proctored lab and unproctored field conditions. *Computers in Human Behavior*, 24(3), 1216–1228.
60. Tendeiro, J. N., Meijer, R. R., Schakel, L., et al. (2012). Using cumulative sum statistics to detect inconsistencies in unproctored internet testing. *Educational and Psychological Measurement*, 73(1), 143–161.
61. Testkuratorium der Föderation Deutscher Psychologenvereinigungen. (2009). TBS-TK. Testbeurteilungssystem des Testkuratoriums der Föderation Deutscher Psychologenvereinigungen. Revidierte Fassung vom 09. September 2009. *Report Psychologie*, 34, 470–478.
62. The International Test Commission. (2006). International guidelines on computer-based and internet-delivered testing. *International Journal of Testing*, 6(2), 143–171.
63. Viswesvaran, C., & Ones, D. S. (1999). Meta-Analyses of fakability estimates: Implications for personality measurement. *Educational and Psychological Measurement*, 59(2), 197–210.
64. Wiechmann, D., & Ryan, A. M. (2003). Reactions to computerized testing in selection contexts. *International Journal of Selection & Assessment*, 11(2–3), 215–229.
65. Wiedmann, J. (2009). Mehrstufiges Auswahlverfahren mit Online-Assessments bei der Luftansa. In H. Steiner (Hrsg.), *Online-Assessments. Grundlagen und Anwendung von Online-Tests in der Unternehmenspraxis* (S. 105–126). Heidelberg: Springer.
66. Wright, N. A., Meade, A. W., & Gutierrez, S. L. (2014). Using invariance to examine cheating in unproctored ability tests. *International Journal of Selection & Assessment*, 22(1), 12–22.
67. Wurow, A. (2009). Interne Stellenbesetzung innerhalb der Conergy Gruppe mit dem verhaltensbasierten Fragebogen »Predictive Index«. In H. Steiner (Hrsg.), *Online-Assessments. Grundlagen und Anwendung von Online-Tests in der Unternehmenspraxis* (S. 127–140). Heidelberg: Springer.

Internetverweise

68. Information zum E-Assessment der Deutschen Post AG. <http://www.eligo.de/de/files/2012/07/Personalmarketing-Azubis-eligo.pdf>. Zugegriffen: 13. März 2015.
69. Information zum E-Assessment der Targobank. <http://www.cyquest.net/projektliste/webbasiertes-eassessment-zur-rekrutierung-von-auszubildenden-bei-der-targobank/>. Zugegriffen: 13. März 2015.
70. Self-Assessment von E.ON. <http://www.cyquest.net/projektliste/e-on-testet-ausbildungsplatz-bewerber-online-mit-dem-e-on-phasenprufer/>. Zugegriffen: 13. März 2015.