

Methoden der internetbasierten Diagnostik in der Psychologie

Timo Gnams

Leibniz Institut für Bildungsverläufe, Bamberg, Germany

Version: 2023-12-11

Das Dokument ist ein Preprint von:

Gnams, T. (2024). Methoden der internetbasierten Diagnostik in der Psychologie. R. Dohrenbusch (Hrsg.), *Psychologische Begutachtung*. Springer. https://doi.org/10.1007/978-3-662-64801-8_47-1

Abstract

Computer und andere Informationstechnologien können viele Aufgaben in der psychologischen Diagnostik unterstützen oder überhaupt erst ermöglichen. Dies reicht von Erstgesprächen zur Exploration der Problemlage über Videokonferenzsysteme bis hin zu internetbasierten Persönlichkeits- und Leistungsmessungen oder der Erfassung biophysiologicaler Merkmale und Aktivitätsmuster mittels tragbarer Sensortechnik. Der vorliegende Beitrag gibt einen Überblick über internetbasierte Methoden in der psychologischen Diagnostik und diskutiert deren Möglichkeiten für verschiedene Anwendungsfelder.

Keywords

technologiebasierte Diagnostik, mobile Diagnostik, Computer, Internet, Smartphones

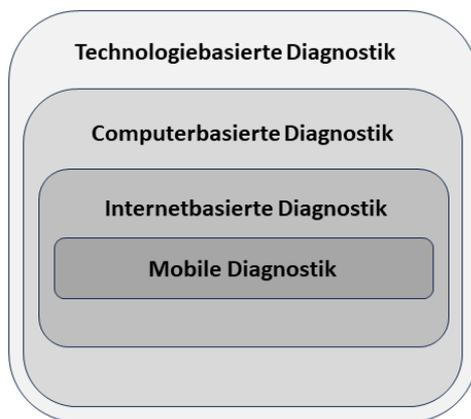
Einleitung

Die zunehmende Digitalisierung des privaten, beruflichen und öffentlichen Lebens hat in den letzten Jahrzehnten die Art und Weise, wie Menschen miteinander kommunizieren, interagieren und Informationen austauschen, nachhaltig verändert. Für die psychologische Praxis und insbesondere auch für die psychologische Diagnostik, die ein zentraler Bestandteil vieler psychologischer Tätigkeiten ist, haben internetbasierte Methoden zahlreiche neue Möglichkeiten eröffnet. Traditionell setzt die psychologische Diagnostik auf persönliche Interaktionen zwischen den beteiligten Personen, um im Gespräch, durch Beobachtung und mit standardisierten Erhebungsverfahren Informationen über relevante Merkmale einer Person zu sammeln, zu interpretieren und daraus Handlungsempfehlungen abzuleiten (Petermann & Wirtz, 2023). Moderne Technologien wie internetfähige Smartphones oder Videokonferenzsysteme können heute viele diagnostische Tätigkeiten wie Anamnesegespräche, Leistungstests oder klinische Diagnostik in vielfältiger Weise unterstützen. So ermöglichen digitale Anwendungen eine orts- und auch zeitunabhängige Durchführung psychologischer Arbeit und damit einen breiteren Zugang zu psychologischer Unterstützung auch für Personen, die diese sonst nicht in Anspruch nehmen könnten (z.B. in der schulpyschologischen Arbeit, Reupert et al., 2022, oder in der frühkindlichen Autismusdiagnostik, Dahiya et al., 2020). Computeradaptive Leistungstests ermöglichen eine individualisierte und präzisere Messung psychologischer Merkmale als viele herkömmliche lineare Tests (Frey, 2023). Mobile Technologien erschließen sogar neue Informationsquellen, beispielsweise für die Verlaufsdiagnostik, indem psychologische Merkmale wiederholt im natürlichen Lebensumfeld der Betroffenen gemessen werden (z.B. affektive Störungen im Wochen- oder Monatsverlauf; vgl. Grossman et al., 2020). Die Integration von Internettechnologien in die psychologische Diagnostik birgt jedoch auch Herausforderungen und erfordert eine intensive Auseinandersetzung mit ethischen und testtheoretischen Fragen. So ist beispielsweise häufig ungeklärt, wie die im Rahmen des Diagnoseprozesses erhobenen personenbezogenen Daten sicher und datenschutzkonform übertragen und gespeichert werden können. Zu berücksichtigen ist auch, inwieweit wenig standardisierte Erhebungsbedingungen die Validität und Normierung psychologischer Tests beeinträchtigen und damit die Interpretation der Ergebnisse erschweren. Dieses Kapitel versucht nun einen Überblick über Methoden der internetbasierten Diagnostik in der Psychologie zu geben und Forschende wie Praktikerinnen und Praktiker anzuregen, über deren Chancen und Herausforderungen zu reflektieren, in der Hoffnung, neue Perspektiven für die Zukunft der psychologischen Diagnostik zu eröffnen.

Merkmale internetbasierter Diagnostik

Internetbasierte Diagnostik (auch Online-Diagnostik) ist ein Teilbereich der technologiebasierten Diagnostik, die als Sammelbegriff verschiedene Ansätze beschreibt, psychologische Diagnostik mit Hilfe moderner Informationstechnologien wie Computern (z.B. Laptops) oder mobilen Endgeräten (z.B. Smartphones) zu unterstützen oder überhaupt erst zu ermöglichen. Um der Vielfalt unterschiedlicher Varianten technologiebasierter Diagnostik in der Praxis gerecht zu werden, ist es sinnvoll, eine hierarchische Struktur zugrunde zu legen (siehe Abbildung 1 in Anlehnung an Jurecka & Hartig, 2007).

Abbildung 1. Varianten technologiebasierter psychologischer Diagnostik



Computerbasierte Diagnostik beschreibt alle Anwendungen, die Computer und andere Technologien nutzen, um psychologische Erhebungsverfahren zu präsentieren und (in der Regel) die Reaktionen der Testpersonen automatisch zu protokollieren. Dies umfasst sowohl traditionelle Papier-Bleistift-Verfahren, die weitgehend unverändert am Computer administriert werden, als auch neue Verfahren wie computeradaptive oder reaktionszeitsensitive Tests, die ohne Technologieeinsatz nur schwer realisierbar wären. Moderne Testsysteme integrieren zudem häufig auch automatische Auswertungen und Unterstützung bei der (normorientierten) Interpretation der Ergebnisse. Während die computerbasierte Diagnostik auch den lokalen Einsatz digitaler Erhebungsverfahren umfasst, steht bei der internetbasierten Diagnostik insbesondere die Möglichkeit im Vordergrund, dass Beteiligte über mobile oder stationäre Endgeräte mit Internetzugang orts- sowie zeitunabhängig miteinander in Kontakt treten können. Die Teststeuerung erfolgt in der Regel über einen zentralen Testserver, der alle für den Testablauf relevanten Informationen über das Internet an die Geräte der Testpersonen sendet und deren Reaktionen speichert. Internetbasierte Verfahren sind in ihren Möglichkeiten gegenüber herkömmlichen computerbasierten Tests kaum

eingeschränkt. Bei entsprechender Gestaltung sind die meisten computerbasierten Verfahren auch als internetbasierte Varianten realisierbar, wie beispielsweise domänenspezifische Kompetenztests (Gnambs & Lenhard, 2023), reaktionszeitsensitive Tests (Gnambs, 2023) und objektive Persönlichkeitstests (Gnambs et al., 2015) zeigen. Mobile Diagnostik beschreibt internetbasierte Erhebungsformen, die mit Hilfe tragbarer elektronischer Geräte wie Smartphones durchgeführt werden. Da diese Geräte im Vergleich zu stationären Computern meist über eine eingeschränkte technische Ausstattung verfügen (z.B. kleinere Bildschirme, keine externe Tastatur), müssen psychologische Verfahren, insbesondere in der Leistungsdiagnostik, diesen Rahmenbedingungen Rechnung tragen. Tests, die zum Beispiel komplexe Interaktionen mit dem Itemmaterial erfordern (z.B. Zahlenverbindungstests), sind daher nur schwer als mobile Anwendungen zu realisieren.

Die Einsatzmöglichkeiten internetbasierter Verfahren in der diagnostischen Praxis werden wesentlich von zwei Kontextbedingungen beeinflusst, dem Grad der Kontrolle und der Standardisierung (vgl. Kroehne et al., 2019). Internetbasierte Tests und Fragebögen können prinzipiell unbeaufsichtigt ohne Anwesenheit einer Testleiterin oder eines Testleiters bearbeitet werden. Personen können dann ohne individuelle Identifikation (*offener Modus*; Bartram, 2006) oder nach Eingabe eines personalisierten Passworts (*kontrollierter Modus*) auf das Testmaterial zugreifen und es selbständig bearbeiten. Bei diesem Vorgehen kann jedoch nicht ausgeschlossen werden, dass auch unbefugte Dritte anstelle der vorgesehenen Personen den Test bearbeiten. Darüber hinaus kann nicht ausgeschlossen werden, dass Personen unzulässige Hilfsmittel verwenden, die bei einer herkömmlichen Anwendung der Verfahren in einer psychologischen Praxis nicht zur Verfügung stehen würden (vgl. Steger et al., 2020). Daher beschränkt sich der Nutzen internetbasierter Diagnostik im offenen und kontrollierten Modus auf ausgewählte Anwendungsfelder, wie beispielsweise als Selbsttest zur Unterstützung der Studienfachwahl (Stoll et al., 2022) oder zur Früherkennung kognitiver Beeinträchtigungen (Charalambous et al., 2020). Sind personenbezogene Einzelleistungen von Interesse, um individuelle Handlungsempfehlungen abzuleiten, wie es beispielsweise in der Personalauswahl der Fall ist, sollten überwachte Testszenarien bevorzugt werden (*angeleiteter Modus*; Bartram, 2006). Hierbei kontrollieren menschliche Testleiterinnen oder Testleiter mittels Video- und Bildschirmübertragung oder Systeme mit künstlicher Intelligenz nicht nur die Identität der teilnehmenden Personen, sondern überwachen auch den gesamten Testablauf und die Testsituation. Dieses Verfahren stellt allerdings höhere Anforderungen an die technische Infrastruktur aller Beteiligten.

Internetbasierte Datenerhebungen können selten so standardisiert durchgeführt werden wie in einem psychologischen Labor, in dem für alle Personen vergleichbare Rahmenbedingungen geschaffen werden. Die internetbasierte Diagnostik erfolgt in der Regel auf der Hard- und Software, die den Teilnehmenden gerade zur Verfügung steht. Variationen in der Computerausstattung, wie zum Beispiel unterschiedliche Bildschirmgrößen, Eingabemedien (z.B. Maus, Touchscreen) sowie die Geschwindigkeit der Internetverbindung, können unter Umständen die Ergebnisse der psychologischen Messungen beeinflussen. Zudem findet internetbasierte Diagnostik häufig in sehr unterschiedlichen Umgebungen statt, da sie je nach Präferenz der Teilnehmenden zu Hause, am Arbeitsplatz oder an beliebigen öffentlichen Orten wie Zügen oder Cafés durchgeführt werden kann, solange eine Internetverbindung zur Verfügung steht. Dadurch ist oft nicht gewährleistet, dass psychologische Verfahren in einer ähnlich geschützten und ruhigen Umgebung durchgeführt werden können, wie es in einer psychologischen Praxis möglich wäre. Je nach Situation können Lärm oder Ablenkungen durch andere Personen den diagnostischen Prozess stören. Unter Umständen kann auch die jeweilige soziale Situation das Testergebnis beeinflussen. Beispielsweise konnte gezeigt werden, dass psychologische Messungen variieren können, je nachdem, ob während der Erhebung andere Personen im Raum anwesend sind oder nicht (Becker et al., 2017; Gnambs & Kaspar, 2015).

Beide Kontextfaktoren, eine mangelnde Kontrolle und eine eingeschränkte Standardisierung, oder eine Kombination aus beiden, können bei internetbasierten Erhebungen zu systematischer, konstruktirrelevanter Variabilität führen und damit die Validität psychologischer Messungen gefährden. Bislang ist nicht eindeutig geklärt, welche Kontextfaktoren in welchen Situationen zu nicht zu vernachlässigenden Verzerrungen können (zur Diskussion vgl. Gnambs & Lenhard, 2023; Schroeders & Wilhelm, 2010; Steger et al., 2020).

Methoden internetbasierter Diagnostik

Das Internet bietet für die psychologische Diagnostik keine völlig neuen Methoden, die über Fragen, Testen und Beobachten hinausgehen würden. Die Digitalisierung eröffnet jedoch neue Möglichkeiten für deren Umsetzung. Im Folgenden werden daher ausgewählte Methoden der internetbasierten psychologischen Diagnostik kurz vorgestellt.

Internetbasierte Interviews

Psychologische Explorationsgespräche über das Internet wurden lange Zeit überwiegend in schriftlicher Form durchgeführt. Während synchrone Interviews über Chat-Systeme realisiert

Internetbasierte Diagnostik

werden, in denen der Austausch wie im persönlichen Gespräch unmittelbar ohne Zeitverzögerung erfolgt, sind asynchrone Interviews, beispielsweise per E-Mail, durch zeitversetzte Frage-Antwort-Runden gekennzeichnet. Ein Problem dieser Interviewformen ist das Fehlen nonverbaler Informationen wie Blickkontakt, Mimik oder Körperhaltung, die die Interpretation der Gesprächsinhalte mitbestimmen. Schriftliche oder grafische Darstellungen emotionaler Inhalte in Form von Emoticons oder Emojis können hier nur bedingt Abhilfe schaffen, da sie menschliche Emotionen nur selten so reichhaltig vermitteln können. Zudem dauern Chat-Interviews oft deutlich länger als vergleichbare persönliche Interviews, da der Dialog stark von der Lese- und Tippgeschwindigkeit der Beteiligten abhängt. Qualitativ unterscheiden sich internetbasierte Gespräche jedoch häufig nicht wesentlich von persönlichen Gesprächen (vgl. Lopez et al., 2019). Internetbasierte Interviews in schriftlicher Form sind heute jedoch nur noch Ausnahmefällen vorbehalten. Beispielsweise kann die internetbasierte Kommunikation für Personen mit Kontaktschwierigkeiten (z.B. bei sozialen Angststörungen) einen niedrighwelligen Zugang darstellen, um professionelle psychologische Hilfe in Anspruch zu nehmen.

Mit der zunehmenden Verbreitung von Videokonferenzsystemen, bei denen sich die Beteiligten sehen und hören können, werden internetbasierte Gespräche immer natürlicher. Psychologische Interviews können damit so ähnlich wie in einer psychologischen Praxis gestaltet werden. Zwar wird vereinzelt kritisiert, dass Personen in videobasierten Interviews manchmal Schwierigkeiten haben können, die sozialen Signale ihres Gegenübers zu interpretieren (Banbury et al., 2018), doch sind diese Probleme meist gering und haben in der Regel keinen bedeutsamen Einfluss auf den Gesprächsinhalt. Verschiedene Vergleichsstudien deuten darauf hin, dass videobasierte Interviews persönliche Interviews ersetzen können, ohne qualitative Einschränkungen in Kauf nehmen zu müssen (Ballejos et al., 2018; Endres et al., 2023). So sind selten Unterschiede in der Bereitschaft, intime Informationen über die eigene Person preiszugeben, zwischen den beiden Befragungsvarianten zu beobachten (Oates et al., 2022). Allerdings werden deutlich weniger persönliche Erfahrungen geteilt, wenn videobasierte Interviews an öffentlichen Orten durchgeführt werden, an denen andere Personen die Gesprächsinhalte mithören können (Jenner & Meyers, 2018). Daher sollte bei internetbasierten Interviews darauf geachtet werden, dass die Personen geschützte Orte für das Interview wählen (z.B. die eigene Wohnung).

Internetbasierte Fragebögen

Schriftliche Befragungen zur Messung psychologischer Merkmale wie Persönlichkeit, Motivation oder Interessen sind die mit Abstand am weitesten verbreiteten Erhebungsverfahren in der internetbasierten Diagnostik. Ihre Umsetzung ist technisch vergleichsweise einfach und erfordert häufig nur die Übertragung etablierter Papier-Bleistift-Fragebögen auf den Computer. Größere Anpassungen sind selten notwendig. Ein Vorteil ist, dass auch multimediales Stimulusmaterial wie kurze Videosequenzen sehr einfach in den Testablauf integriert werden können. Die Bearbeitung internetbasierter Fragebögen erfordert in der Regel wenig spezifische Vorkenntnisse. Verschiedene Vergleichsstudien zeigen, dass persönlichkeits- aber auch klinisch-diagnostische Instrumente nur selten durch die wenig standardisierten und kontrollierten Erhebungsbedingungen der internetbasierten Diagnostik beeinträchtigt werden. So sind psychometrische Eigenschaften wie Reliabilität und Faktorenstruktur meist nicht systematisch davon betroffen, ob ein Instrument vom Papier auf den Computer übertragen wurde oder ob es im offenen oder angeleiteten Modus eingesetzt wurde (vgl. Sawhaney & Cigularov, 2014; Vecchione et al., 2012). Es gibt auch wenig Hinweise darauf, dass sich Personen in internetbasierten Befragungen weniger Mühe geben, sorgfältig zu antworten (Clement et al., 2023) oder weniger dazu neigen sozial erwünschte Antworten zu geben (Gnambs & Kaspar, 2017). Auch die Nutzung mobiler Endgeräte mit kleinen Bildschirmen scheint die Messung von Selbstberichten nicht wesentlich zu beeinflussen (Menold & Toepoel, 2022). Manche Arbeiten weisen sogar darauf hin, dass häufig von einer Äquivalenz der Testscores für mobile und nicht-mobile Anwendungen ausgegangen werden kann (Arthur et al., 2018), was die Anwendung gemeinsamer Normen für unterschiedliche Endgeräte ermöglichen würde. Einschränkend muss jedoch betont werden, dass die wenigsten in der psychologischen Praxis eingesetzten Fragebögen explizit für die Anwendung im Internet validiert wurden. Trotz verschiedener vielversprechender Befunde zur psychometrischen Qualität internetbasierter Befragungen im Allgemeinen ist die Übertragung von Normen für spezifische Instrumente aus kontrollierten Papier-Bleistift-Befragungen auf weniger standardisierte Befragungsformen daher immer mit Unsicherheiten behaftet.

Schließlich erschließen internetbasierte Instrumente auch neue Datenquellen, die für diagnostische Zwecke genutzt werden können. Am Computer kann auch das Testbearbeitungsverhalten wie etwa die Bearbeitungszeit für jedes Item oder mögliche Antwortänderungen automatisch erfasst werden. Solche Prozessdaten stellen neben den eigentlichen Itemantworten eine zusätzliche Informationsquelle dar, die für die Merkmalsmessung genutzt werden kann. So wurden verschiedene Ansätze vorgeschlagen,

Bearbeitungszeiten zu verwenden, um Personen oder einzelne Itemantworten zu identifizieren, die auf unaufmerksames oder zufälliges Antwortverhalten schließen lassen (Schroeders et al., 2022; Ulitzsch et al., 2022). Diese Information kann herangezogen werden, um psychologische Merkmale präziser zu messen oder um die Ergebnisse einzelner Personen nur mit Vorsicht zu interpretieren.

Internetbasierte Tests

Computerbasierte Leistungstests zur Messung psychologischer Merkmale wie Intelligenz, bereichsspezifischer Kompetenzen aber auch der Persönlichkeit (im Sinne einer objektiven Persönlichkeitsdiagnostik) stellen mittlerweile in vielen Anwendungsbereichen einen *de-facto* Standard dar, da eine digitale Umsetzung häufig eine höhere Standardisierung und Präzision psychologischer Messungen erlaubt. Die meisten dieser Verfahren können prinzipiell auch über das Internet administriert werden. Dies umfasst einfache Auswahlaufgaben wie die bekannten Matrizenaufgaben zur Messung des logisch-schlussfolgernden Denkens (Nussenbaum et al., 2020) ebenso wie simulationsbasierte Tests zur Messung der Risikoneigung (Gnambs et al., 2015). Bei entsprechender Gestaltung können internetbasierte Tests sogar mit kleinen Kindern (Gnambs & Lenhard, 2023; Silver et al., 2021) und auch mit älteren Menschen (Papp et al., 2021) durchgeführt werden. Dennoch werden internetbasierte Tests bislang deutlich seltener eingesetzt als internetbasierte Befragungen.

Da Leistungstests darauf abzielen die maximale Leistung bei der Aufgabenbearbeitung zu erfassen, stellt die eingeschränkte Kontrolle und Standardisierung ein wesentliches Hindernis für eine valide Leistungsdiagnostik über das Internet dar. Wenig überraschend neigen Personen in unbeaufsichtigten Situationen dazu unlautere Hilfsmittel zu verwenden (z.B. die richtigen Antworten im Internet zu suchen), was zu systematisch höheren Testergebnissen führt als in angeleiteten Testsituationen (Steger et al., 2020). Zudem neigen Personen in unbeaufsichtigten Testsituationen häufig zu einem weniger sorgfältigen Antwortverhalten und zeigen eine höhere Tendenz zum schnellen Raten (Kroehne et al., 2020). Internetbasierte Leistungsdiagnostik sollte daher immer in irgendeiner Form beaufsichtigt werden, insbesondere wenn die Testergebnisse individuelle Konsequenzen für die Personen haben. Eine Alternative zur Überwachung sind sogenannte Verifikationstests (Lievens & Burke, 2011), die in der Selektionsdiagnostik im Anschluss an eine nicht-überwachte internetbasierte Testung durchgeführt werden. Dabei handelt es sich um eine (meist computeradaptive) Kurzversion des ursprünglichen Tests, die unter kontrollierten Bedingungen wiederholt wird, um die Plausibilität der unkontrollierten Testergebnisse zu bestätigen.

Da Leistungstests auf unterschiedlichen Endgeräten (z.B. Laptops, Tablets, Smartphones) bearbeitet werden können, wird häufig befürchtet, dass diese Unterschiede die Vergleichbarkeit der Messungen ungewollt einschränken. Solche Effekte lassen sich zwar für zeitlich begrenzte Speed-Tests nachweisen (Passell et al., 2021), bei nicht zeitlich begrenzten Power-Tests zeigen sich jedoch häufig nur geringe Unterschiede (Brown & Grossenbacher, 2017; Schroeders & Wilhelm, 2010). Verschiedene Vergleichsstudien legen nahe, dass die psychometrische Qualität von Leistungstests in der Regel nicht systematisch unter einer internetbasierten Vorgabe leidet (z.B. Gnambs & Lenhard, 2023; Zinn et al., 2021). Die Frage, inwieweit eine mangelnde Standardisierung die Validität internetbasierter Messungen beeinträchtigt, kann jedoch nicht unabhängig vom eingesetzten Testverfahren und der untersuchten Person beantwortet werden. Beispielsweise ist ein kleiner Bildschirm bei einfachen Matrizenaufgaben weniger problematisch als bei Kompetenztests, die längere Texte zur Erfassung des Leseverständnisses verwenden. Ebenso können unterschiedliche Erfahrungen mit technischen Geräten und damit die digitale Kompetenz von Personen beeinflussen, wie präzise Informationen mit bestimmten Endgeräten (z.B. Smartphones) erhoben werden können.

Internetbasierte Beobachtungen

Mobile digitale Geräte sind aus dem Alltag vieler Menschen nicht mehr wegzudenken. Unter dem Begriff *Wearables* werden verschiedene elektronische Geräte zusammengefasst, die am Körper getragen werden und in der Regel mit Sensoren und drahtloser Konnektivität ausgestattet sind. Sie können unterschiedlichste Formen annehmen und umfassen unter anderem Smartphones, Smartwatches und Fitness-Tracker. Die eingebauten Sensoren ermöglichen die automatische Überwachung der Aktivitäten des Trägers oder der Trägerin und die Messung verschiedener physiologischer Parameter. Der Einsatz von Wearables für psychologische Anwendungen hat in den letzten Jahren, insbesondere in der klinisch-psychologischen Diagnostik, stark zugenommen, da sie eine Vielzahl von Daten erfassen können, die die Messung individueller Merkmale unterstützen können. Beispielsweise können physiologische Parameter wie die Herzfrequenzvariabilität oder die Hautleitfähigkeit ein Indikator für Stress sein (Klimek et al., 2023). Eine plötzliche Abnahme der körperlichen Aktivität, gemessen durch das Zählen von Schritten oder zurückgelegten Kilometern, kann den Beginn einer depressiven Episode anzeigen (Abd-Alrazaq et al., 2023). Die Überwachung der Atemfrequenz kann Hinweise auf emotionale Zustände liefern (Ba & Hu, 2023). Darüber hinaus gibt es Versuche, auch Informationen für diagnostische Anwendungen nutzbar zu machen, die auf den ersten Blick keine konzeptionelle Nähe zu den gemessenen Eigenschaften

Internetbasierte Diagnostik

aufweisen. Beispielsweise gibt es Befunde, die nahelegen, dass auf Basis von Nutzungsdaten des Endgerätes (z.B. wann der Bildschirm eingeschaltet ist, Ladestand des Akkus), Standortdaten (z.B. Längen- und Breitengrad) und sozialen Interaktionen (z.B. Anzahl und Zeitpunkt von Nachrichten und Anrufen) mit Methoden des maschinellen Lernens Rückschlüsse auf Persönlichkeitseigenschaften von Personen gezogen werden können (Marengo et al., 2023).

Allen Anwendungen gemeinsam ist jedoch, dass die Reliabilität und Validität dieser Messungen im Vergleich zu herkömmlichen Verfahren oft deutlich eingeschränkt ist oder bisher nicht systematisch untersucht wurde. Auch die Kontextualisierung der erhobenen Daten ist oft schwierig, da häufig nicht klar ist, in welchen sozialen Situationen zum Beispiel auffällige physiologische Muster beobachtet wurden. Daher sollten Sensordaten, die mit Wearables erhoben werden, derzeit vor allem als unterstützendes Instrument von Fachkräften im Rahmen eines umfassenden diagnostischen Ansatzes angesehen werden und nicht als Ersatz für traditionelle diagnostische Methoden. Insbesondere als Screening-Instrument zur Früherkennung von Anzeichen klinisch-psychologischer Symptome wie Depressionen können mit Wearables erhobene Daten dazu beitragen, Betroffene rechtzeitig auf psychologische Unterstützungsangebote hinzuweisen.

Ethische Herausforderungen internetbasierter Diagnostik

Die internetbasierte psychologische Diagnostik ist mit verschiedenen ethischen Herausforderungen konfrontiert, die einer sorgfältigen Reflexion bedürfen. Beispielsweise birgt die Übertragung sensibler psychologischer Informationen über das Internet das Risiko, dass Unbefugte Zugang zu persönlichen Informationen der Betroffenen erhalten. Auch bei der Nutzung scheinbar sicherer technischer Systeme (z.B. Videokonferenzsysteme bekannter Anbieter) muss sichergestellt werden, dass Informationen nicht automatisch in Länder außerhalb Deutschlands mit weniger strengen Datenschutzbestimmungen übermittelt und dort verarbeitet werden. Darüber hinaus wird das Problem der Standardisierung der Erhebungsbedingungen den Betroffenen aufgebürdet. Sie sind dafür verantwortlich, die für die jeweilige Situation als optimal erachtete technische Infrastruktur und die räumliche Umgebung bereitzustellen. Personen, die dazu nicht in der Lage sind, zum Beispiel weil sie kein Endgerät mit ausreichend großem Bildschirm oder keinen ruhigen Raum organisieren können, müssten dann von der psychologischen Diagnostik ausgeschlossen werden oder sich mit den möglicherweise erschwerten Kontextbedingungen arrangieren. Dies kann die Fairness internetbasierter Diagnostik einschränken. Internetbasierte Erhebungsformen sollten daher eher

Internetbasierte Diagnostik

als zusätzliches Angebot für Personen verstanden werden, die die notwendigen Voraussetzungen erfüllen, und nicht als Verpflichtung für Personen, die dies nicht können oder möglicherweise nicht wollen.

Insbesondere bei nicht überwachten Varianten internetbasierter Diagnostik kann es auch schwierig sein, Personen ausreichend über den Zweck und mögliche Risiken einer Teilnahme aufzuklären (*informed consent*), da schriftliche Informationen häufig nicht richtig gelesen werden (Perrault & Keating, 2018). Während im persönlichen Gespräch durch gezieltes Nachfragen Unklarheiten meist schnell erkannt und geklärt werden können, ist dies bei internetbasierten Erhebungsformen ohne Anwesenheit einer Testleiterin oder eines Testleiters mitunter schwierig. Ähnliches gilt für automatisierte Rückmeldungen von Testergebnissen, die den Betroffenen meist nur standardisierte Interpretationshilfen bieten, aber nicht auf die individuelle Problemlage des oder der Einzelnen eingehen. Selten kann eine Zusammenfassung schriftlicher Testergebnisse ein Gespräch mit Fachleuten ersetzen, um die Aussagekraft und Grenzen von Testergebnissen für bestimmte Situationen und Entscheidungen abzuwägen.

Für den Fall von Notfällen wie akuten psychischen Krisen oder Selbstmordgefährdung, beispielsweise während eines Videointerviews, muss sichergestellt sein, dass effektive Interventionsmechanismen vorhanden sind, um den Betroffenen unkompliziert professionelle Hilfe zukommen zu lassen. Daher müssen im Vorfeld Prozesse vorbereitet werden, wie mit Notfällen umgegangen wird, auch wenn die Beteiligten räumlich voneinander getrennt sind. Die Berücksichtigung dieser ethischen Herausforderungen ist entscheidend, um sicherzustellen, dass internetbasierte psychologische Diagnostik ethisch vertretbar und für die Betroffenen sicher ist. Daher ist es wichtig, dass sich Forschende und Fachkräfte kontinuierlich mit den sich entwickelnden ethischen und datenschutzrechtlichen Standards auseinandersetzen, um den bestmöglichen Schutz aller Beteiligten angesichts sich verändernder Technologien zu gewährleisten.

Fazit

Die Digitalisierung vieler Lebensbereiche bietet zweifellos neue Möglichkeiten für die psychologische Diagnostik, bringt aber auch Herausforderungen mit sich, die besonderer Aufmerksamkeit bedürfen. So können viele diagnostische Aufgaben wie das psychologische Explorationsgespräch, Befragungen und Tests auch über das Internet durchgeführt werden. Die Qualität der erhobenen Informationen wird jedoch stark von Kontextbedingungen wie dem Grad der Kontrolle und Standardisierung beeinflusst. Testtheoretische Überlegungen sind daher unerlässlich, um sicherzustellen, dass internetbasierte Erhebungen die gleiche psychometrische

Internetbasierte Diagnostik

Qualität aufweisen wie traditionelle Methoden. Angewandte Psychologinnen und Psychologen müssen sich bewusst sein, dass technologische Unterschiede die Vergleichbarkeit beeinträchtigen können und daher Strategien entwickeln, um solche Verzerrungen zu minimieren. Die Digitalisierung bringt auch verschiedene ethische Herausforderungen mit sich, die die Umsetzung klarer Richtlinien und Maßnahmen zum Schutz persönlicher Informationen der Betroffenen unabdingbar machen. Insgesamt zeigt sich, dass der Schlüssel zu einer erfolgreichen internetbasierten psychologischen Diagnostik in einem ausgewogenen Ansatz liegt, der die Stärken digitaler Technologien nutzt und gleichzeitig ethische Standards und psychometrische Prinzipien respektiert. Kontinuierliche Forschung in diesem Bereich wird entscheidend sein, um die Qualität internetbasierter psychologischer Diagnostik zu gewährleisten.

Literatur

- Abd-Alrazaq, A., AlSaad, R., Shuweihdi, F., Ahmed, A., Aziz, S., & Sheikh, J. (2023). Systematic review and meta-analysis of performance of wearable artificial intelligence in detecting and predicting depression. *NPJ Digital Medicine*, 6(1), 84. <https://doi.org/10.1038/s41746-023-00828-5>
- Arthur, W. Jr., Keiser, N., Hagen, E., & Traylor, Z. (2018). Unproctored Internet-based device-type effects on test scores: The role of working memory. *Intelligence*, 67, 67–75. <https://doi.org/10.1016/j.intell.2018.02.001>
- Ba, S., & Hu, X. (2023). Measuring emotions in education using wearable devices: A systematic review. *Computers & Education*, 200:104797. <https://doi.org/10.1016/j.compedu.2023.104797>
- Banbury, A., Nancarrow, S., Dart, J., Gray, L., & Parkinson, L. (2018). Telehealth interventions delivering home-based support group videoconferencing: systematic review. *Journal of Medical Internet Research*, 20(2):e25. <https://doi.org/10.2196/jmir.8090>
- Bartram, D. (2006). Testing on the internet: Issues, challenges and opportunities in the field of occupational assessment. In D. Bartram & R. K. Hambleton (Hrsg.), *Computer-Based Testing and the Internet: Issues and Advances* (S. 13-37). Wiley. <https://doi.org/10.1002/9780470712993.ch1>
- Ballejos, M. P., Oglesbee, S., Hetteima, J., & Sapien, R. (2018) An equivalence study of interview platform: Does videoconference technology impact medical school acceptance rates of different groups?. *Advances in Health Sciences Education* 23, 601–610. <https://doi.org/10.1007/s10459-018-9817-2>
- Becker, N., Koch, M., Schult, J., & Spinath, F. M. (2017). Setting doesn't matter much: A meta-analytic comparison of the results of intelligence tests obtained in group and individual settings. *European Journal of Psychological Assessment*, 35(3), 309–316. <https://doi.org/10.1027/1015-5759/a000402>
- Brown, M. I., & Grossenbacher, M. A. (2017). Can you test me now? Equivalence of GMA tests on mobile and non-mobile devices. *International Journal of Selection and Assessment*, 25(1), 61-71. <https://doi.org/10.1111/ijasa.12160>
- Charalambous, A. P., Pye, A., Yeung, W. K., Leroi, I., Neil, M., Thodi, C., & Dawes, P. (2020). Tools for app-and web-based self-testing of cognitive impairment: systematic search and evaluation. *Journal of Medical Internet Research*, 22(1):e14551. <https://doi.org/10.2196/14551>

- Clement, S. L., Severin-Nielsen, M. K., & Shamshiri-Petersen, D. (2023). Satisficing behaviour in web surveys. Results from a comparison of web and paper mode across four national survey experiments. *Survey Methods: Insights from the Field*. <https://doi.org/10.13094/SMIF-2023-00007>
- Dahiya, A. V., McDonnell, C., DeLucia, E., & Scarpa, A. (2020). A systematic review of remote telehealth assessments for early signs of autism spectrum disorder: Video and mobile applications. *Practice Innovations*, 5(2), 150–164. <https://doi.org/10.1037/pri0000121>
- Endres, K., Hillygus, D. S., DeBell, M., & Iyengar, S. (2023). A randomized experiment evaluating survey mode effects for video interviewing. *Political Science Research and Methods*, 11(1), 144-159. <https://doi.org/10.1017/psrm.2022.30>
- Frey, A. (2023). Computerized adaptive testing and multistage testing. In R. J. Tierney, F. Rizvi und K. Ercikan (Hrsg.), *International Encyclopedia of Education* (4. Ausgabe, S. 209-216). Elsevier. <https://doi.org/10.1016/B978-0-12-818630-5.10028-4>
- Gnambs, T. (2023). The web-based assessment of mental speed: An experimental study of testing mode effects for the Trail-Making Test. *European Journal of Psychological Assessment*, 39(5), 349-353. <https://doi.org/10.1027/1015-5759/a000711>
- Gnambs, T., Appel, M., & Oeberst, A. (2015). Color red and risk-taking in online environments. *PLOS ONE*, 10(7): e0134033. <https://doi.org/10.1371/journal.pone.0134033>
- Gnambs, T., & Kaspar, K. (2015). Disclosure of sensitive behaviors across self-administered survey modes: A meta-analysis. *Behavior Research Methods*, 47, 1237-1259. <https://doi.org/10.3758/s13428-014-0533-4>
- Gnambs, T., & Kaspar, K. (2017). Socially desirable responding in web-based questionnaires: A meta-analytic review of the candor hypothesis. *Assessment*, 24, 746-762. <https://doi.org/10.1177/1073191115624547>
- Gnambs, T., & Lenhard, W. (2023). Remote testing of reading comprehension in 8-year-old children: Mode and setting effects. *Assessment*. Advance online publication. <https://doi.org/10.1177/10731911231159369>
- Grossman, J. T., Frumkin, M. R., Rodebaugh, T. L., & Lenze, E. J. (2020). mHealth assessment and intervention of depression and anxiety in older adults. *Harvard Review of Psychiatry*, 28(3), 203-214. <https://doi.org/10.1097/HRP.0000000000000255>
- Jenner, B. M., & Myers, K. C. (2019). Intimacy, rapport, and exceptional disclosure: A comparison of in-person and mediated interview contexts. *International Journal of Social Research Methodology*, 22(2), 165-177. <https://doi.org/10.1080/13645579.2018.1512694>

- Jurecka, A., & Hartig, J. (2007). Anwendungsszenarien computer- und netzwerkbasierter Assessments. In J. Hartig & E. Klieme (Hrsg.), *Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik* (S. 69-79). Bundesministerium für Bildung und Forschung.
- Klimek, A., Mannheim, I., Schouten, G., Wouters, E. J., & Peeters, M. W. (2023). Wearables measuring electrodermal activity to assess perceived stress in care: a scoping review. *Acta Neuropsychiatrica*. Advance online publication. <https://doi.org/10.1017/neu.2023.19>
- Kroehne, U., Deribo, T., & Goldhammer, F. (2020). Rapid guessing rates across administration mode and test setting. *Psychological Test and Assessment Modeling*, 62(2), 147-177. <https://doi.org/10.25656/01:23630>
- Kroehne, U., Gnams, T., & Goldhammer, F. (2019). Disentangling setting and mode effects for online competence assessment. In H.-P. Blossfeld & H.-G. Roßbach (Hrsg.), *Education as a lifelong process* (2. Auflage, S. 171-193). Springer VS. https://doi.org/10.1007/978-3-658-23162-0_10
- 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, 817–824. <https://doi.org/10.1348/096317910X522672>
- Lopez, A., Schwenk, S., Schneck, C. D., Griffin, R. J., & Mishkind, M. C. (2019). Technology-based mental health treatment and the impact on the therapeutic alliance. *Current Psychiatry Reports*, 21:76. <https://doi.org/10.1007/s11920-019-1055-7>
- Marengo, D., Elhai, J. D., & Montag, C. (2023). Predicting Big Five personality traits from smartphone data: A meta-analysis on the potential of digital phenotyping. *Journal of Personality*, 91(6), 1410-1424. <https://doi.org/10.1111/jopy.12817>
- Menold, N., & Toepoel, V. (2022). Do different devices perform equally well with different numbers of scale points and response formats? A test of measurement invariance and reliability. *Sociological Methods & Research*. Advance online publication. <https://doi.org/10.1177/00491241221077237>
- Nussenbaum, K., Scheuplein, M., Phaneuf, C. V., Evans, M. D., & Hartley, C. A. (2020). Moving developmental research online: comparing in-lab and web-based studies of model-based reinforcement learning. *Collabra: Psychology*, 6(1). <https://doi.org/10.1525/collabra.17213>

- Oates, M., Crichton, K., Cranor, L., Budwig, S., Weston, E. J., Bernagozzi, B. M., & Pagaduan, J. (2022). Audio, video, chat, email, or survey: How much does online interview mode matter?. *PloS one*, *17*(2):e0263876. <https://doi.org/10.1371/journal.pone.0263876>
- Papp, K. V., Samaroo, A., Chou, H. C., Buckley, R., Schneider, O. R., Hsieh, S., ... & Amariglio, R. E. (2021). Unsupervised mobile cognitive testing for use in preclinical Alzheimer's disease. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, *13*(1), e12243. <https://doi.org/10.1002/dad2.12243>
- Passell E., Strong R. W., Rutter L. A., Kim H., Scheuer L., Martini P., Grinspoon L., Germine L. (2021). Cognitive test scores vary with choice of personal digital device. *Behavior Research Methods*, *53*(6), 2544–2557. <https://doi.org/10.3758/s13428-021-01597-3>
- Perrault, E. K., & Keating, D. M. (2019). Seeking ways to inform the uninformed: Improving the informed consent process in online social science research. *Journal of Empirical Research on Human Research Ethics*, *13*(1), 50-60. <https://doi.org/10.1177/1556264617738846>
- Petermann, F. & Wirtz, M. A. (2023). Psychologische Diagnostik. In M. A. Wirtz (Hrsg.), *Dorsch Lexikon der Psychologie. Hogrefe*. Verfügbar unter: <https://dorsch.hogrefe.com/gebiet/psychologische-diagnostik>
- Reupert, A., Schaffer, G. E., Von Hagen, A., Allen, K. A., Berger, E., Büttner, G., ... & May, F. (2022). The practices of psychologists working in schools during COVID-19: A multi-country investigation. *School Psychology*, *37*(2), 190-201. <https://doi.org/10.1037/spq0000450>
- Sawhney, G., & Cigularov, K. P. (2014). Measurement equivalence and latent mean differences of personality scores across different media and proctoring administration conditions. *Computers in Human Behavior*, *36*, 412-421. <https://doi.org/10.1016/j.chb.2014.04.010>
- Schroeders, U., Schmidt, C., & Gnams, T. (2022). Detecting careless responding in survey data using stochastic gradient boosting. *Educational and Psychological Measurement*, *82*(1), 29-56. <https://doi.org/10.1177/00131644211004708>
- Schroeders, U., & Wilhelm, O. (2010). Testing reasoning ability with handheld computers, notebooks, and paper and pencil. *European Journal of Psychological Assessment*, *26*(4), 284-292. <https://doi.org/10.1027/1015-5759/a000038>
- Silver, A. M., Elliott, L., Braham, E. J., Bachman, H. J., Votruba-Drzal, E., Tamis-LeMonda, C. S., ... & Libertus, M. E. (2021). Measuring emerging number knowledge in toddlers. *Frontiers in Psychology*, *12*:703598. <https://doi.org/10.3389/fpsyg.2021.703598>

- Steger, D., Schroeders, U., & Gnams, T. (2020). A meta-analysis of test scores in proctored and unproctored ability assessments. *European Journal of Psychological Assessment, 36*, 174-184. <https://doi.org/10.1027/1015-5759/a000494>
- Stoll, G., Weis, S., Beauducel, A., Becker, A., Brazel, C., Delzepich, R., ... & Zay, K. (2022). Ein Rahmenmodell zu Entwicklung, Wirkmechanismen und Bewertung von Online-Self-Assessments. In G. Stoll & S. Weis (Hrsg.), *Online-Self-Assessments zur Studienfachwahl: Entwicklung-Konzepte-Qualitätsstandards* (S. 1-23). Springer. https://doi.org/10.1007/978-3-662-63827-9_1
- Ulitzsch, E., Yildirim-Erbasli, S. N., Gorgun, G., & Bulut, O. (2022). An explanatory mixture IRT model for careless and insufficient effort responding in self-report measures. *British Journal of Mathematical and Statistical Psychology, 75*(3), 668-698. <https://doi.org/10.1111/bmsp.12272>
- Vecchione, M., Alessandri, G., & Barbaranelli, C. (2012). Paper-and-pencil and web-based testing: The measurement invariance of the Big Five personality tests in applied settings. *Assessment, 19*(2), 243-246. <https://doi.org/10.1177/1073191111419091>
- Zinn, S., Landrock, U., & Gnams, T. (2021). Web-based and mixed-mode cognitive large-scale assessments in higher education: An evaluation of selection bias, measurement bias, and prediction bias. *Behavior Research Methods, 53*, 1202-1217. <https://doi.org/10.3758/s13428-020-01480-7>