

Masterstudiengang  
Beginn: 18. September 2023 | Abgabe: 3. Juli 2024

**Masterthesis**  
**Zum Einsatz von Leistungs- und Persönlichkeitstests im Rahmen von  
Personalauswahl und -entwicklung in der Schweiz**  
**Verwendungshäufigkeiten – Zwecke – Zielgruppen – Einflussgrößen**  
2., überarbeitete Version



**Autorin:**  
Tanja Senn  
tanja.senn@students.fhnw.ch

**Betreuung:**  
Prof. Dr. Benedikt Hell  
benedikt.hell@fhnw.ch

**Praxispartnerin:**  
Fachhochschule Nordwestschweiz (FHNW), Hochschule für angewandte Psychologie,  
Institut Mensch in komplexen Systemen (MiKS)

**Kontaktperson bei FHNW / MiKS:**  
Prof. Dr. Benedikt Hell  
benedikt.hell@fhnw.ch

## Zusammenfassung

Die vorliegende Masterthesis beschäftigt sich mit dem Einsatz von psychologischen Tests zur Personalauswahl und -entwicklung in der Schweiz. Dabei werden zum einen Verwendungshäufigkeit, Testart, deren Qualität, Zweck, Zielgruppen und Informationsquellen untersucht. Zum anderen besteht das Forschungsinteresse im Entdecken von Einflussgrössen, die den Testeinsatz in einem Unternehmen begünstigen oder hemmen.

Mittels sequenziellem Mixed Methods Design, bestehend aus einer vorgelagerten qualitativen Interviewstudie und einer quantitativen Fragebogenstudie, wurde das Forschungsinteresse bearbeitet. Der Fokus liegt auf den quantitativen Daten, die durch einen selbst entwickelten Fragebogen bei HR-Mitarbeitenden erhoben wurden.

Die Daten offenbaren, dass insb. qualitativ mangelhafte Persönlichkeitstests, wie der MBTI oder der DISC, eingesetzt werden. Auch der wissenschaftlich fundierte BIP findet vermehrt Anwendung findet. Psychologische Tests dienen laut Untersuchung vor allem dem Zweck der Personalauswahl, wobei mehrheitlich Bewerbende für eine Führungs- oder Fachfunktion getestet werden.

Als Informationsquellen dienen laut quantitativer Befragung überwiegend der Austausch mit Fachexpertinnen und -experten, sowie der Austausch mit Testanbietenden. Testrezensionen oder Fachliteratur werden hingegen kaum genutzt, was in der vorliegenden Thesis als eine mögliche Ursache hinsichtlich des Einsatzes mangelhafter Tests diskutiert wird.

Eine binär logistische Regression zeigte drei statistisch relevante Zusammenhänge für den Testeinsatz in Unternehmen: je positiver die persönliche Einstellung der HR-Mitarbeitenden ggü. psychologischen Tests und je weniger Herausforderungen und Risiken in Bezug auf den Testeinsatz durch die HR-Mitarbeitenden wahrgenommen werden, desto eher werden psychologische Tests in einem Unternehmen eingesetzt. Zudem hängt die Einsatzwahrscheinlichkeit auch davon ab, ob die HR-Mitarbeitenden ihre Meinung in Bezug auf psychologische Tests bei der Entscheidung einbringen dürfen.

Schlagnote: Personalauswahl, Personalentwicklung, psychologische Tests, Persönlichkeitstests, Leistungstests

Anzahl Worte der Zusammenfassung (ohne Schlagworte): 243

Anzahl Zeichen der Masterthesis (inkl. Leerzeichen): 239'990 Zeichen

## **Abstract**

This Master's thesis examines the use of psychological tests for personnel selection and development in Switzerland. It investigates the frequency of use, types of tests, their quality, purposes, target groups, and sources of information. Additionally, the research aims to identify factors influencing test use within companies.

Using a sequential mixed methods design involving qualitative interviews and a quantitative questionnaire survey, the study focuses on quantitative data collected via a self-developed questionnaire administered to HR professionals. The findings reveal a prevalent use of qualitatively inadequate personality tests, such as MBTI and DISC, alongside the scientifically grounded BIP test. Psychological tests primarily serve the purpose of personnel selection, secondarily assessing applicants for leadership or specialist roles.

According to the survey, HR professionals primarily rely on exchanges with subject matter experts and test providers as sources of information. Test reviews and academic literature are underutilized, potentially contributing to the prevalence of inadequate tests discussed in this thesis.

Binary logistic regression identified three statistically significant correlations for test usage in companies: positive attitudes of HR professionals toward psychological tests, perceptions of less challenges or risks associated with test deployment, and the extent to which HR professionals' opinions influence test decisions.

Keywords: personnel selection, personnel development, psychological tests, personality tests, aptitude tests

Number of words of abstract (without keywords): 196

# Inhaltsverzeichnis

1	Ausgangslage und Zielsetzung .....	7
1.1	Orientierung .....	9
2	Im Zentrum stehende Theorien, Konzepte und Modelle .....	10
2.1	Psychologische Tests .....	10
2.1.1	Persönlichkeitseigenschaften im Interesse der Eignungsdiagnostik .....	11
2.1.1.1	Kategoriale-typisierende vs. dimensionale Persönlichkeitsabfrage .....	13
2.1.2	Fähigkeiten im Interesse der Eignungsdiagnostik .....	13
2.1.2.1	Allgemeine Intelligenz .....	13
2.1.2.2	Emotionale Intelligenz .....	15
2.1.3	Interesse .....	17
2.2	Qualität – Gütekriterien .....	18
2.2.1	Hauptgütekriterien .....	19
2.2.1.1	Objektivität .....	19
2.2.1.2	Reliabilität .....	20
2.2.1.3	Validität .....	20
2.2.1.3.1	Inhaltsvalidität .....	21
2.2.1.3.2	Konstruktvalidität .....	21
2.2.1.3.3	Prognostische Kriteriumsvalidität .....	21
2.2.1.3.4	Inkrementelle Validität .....	23
2.2.2	Nebengütekriterien .....	24
2.2.3	Qualitätssicherung .....	26
2.3	Testeinsatz in der Praxis .....	27
2.4	Hypothesen .....	31
3	Datenerhebungs- und Auswertungsmethoden .....	32
3.1	Untersuchungsablauf .....	32
3.2	Qualitative Datenerhebung .....	34
3.2.1	Leitfadenentwicklung .....	34
3.2.1.1	Aufbau des Leitfadens .....	35
3.2.2	Stichprobenplanung und tatsächliche Stichprobe .....	35
3.2.3	Durchführung der qualitativen Datenerhebung .....	37
3.2.4	Qualitative Datenauswertung .....	37
3.2.5	Qualitative Ergebnisse .....	39
3.3	Quantitative Datenerhebung .....	49
3.3.1	Fragebogenentwicklung .....	49
3.3.1.1	Erstellung des Fragebogenentwurfs .....	49
3.3.1.1.1	Entwicklung der Fragen zu den unabhängigen, subjektiven Variablen .....	51
3.3.1.1.1.1	Entwicklung der Antwortskalen .....	51

3.3.1.1.1.2	Reihenfolge der Fragen .....	52
3.3.1.1.2	Entwicklung der Fragen zur abhängigen Variablen und Folgefragen .....	53
3.3.1.1.3	Entwicklung der Fragen zu den unabhängigen, objektiven Variablen .....	53
3.3.1.1.4	Entwicklung der Fragen zu den demografischen Daten .....	54
3.3.1.1.5	Fertigstellung des Fragebogenentwurfs .....	54
3.3.1.2	Empirische Überprüfung des Entwurfs mittels Pretest .....	54
3.3.1.2.1	Auswertung des Pretests hinsichtlich Inhaltsvalidität .....	55
3.3.1.2.2	Auswertung des Pretests hinsichtlich User Experience .....	56
3.3.1.2.3	Auswertung des Pretests hinsichtlich Retest-Reliabilität .....	56
3.3.1.3	Erstellung der endgültigen Testversion .....	57
3.3.1.3.1	Massnahmen betreffend Ziele 1 und 2 .....	57
3.3.1.3.2	Übersetzung und Einpflegen der Sprachversionen.....	57
3.3.2	Schätzung der Stichprobengrösse .....	58
3.3.3	Durchführung der Fragebogenstudie .....	59
4	Quantitative Datenauswertung und Ergebnisdarstellung .....	60
4.1	Vorbereitung der Datensets .....	60
4.2	Stichprobenzusammensetzung .....	61
4.3	OF1 – Häufigkeit .....	67
4.4	OF2 – Testart / Testnamen .....	68
4.5	UF1 – Qualität .....	71
4.6	UF2 – Zielgruppen .....	72
4.7	UF3 – Testzweck .....	73
4.8	UF4 – Informationsbasis .....	74
4.9	UF5 bis UF7 – Hintergründe .....	75
4.9.1	Berechnung der Korrelationen auf Itemebene .....	75
4.9.2	Faktorenbildung zur Datenreduktion.....	76
4.9.2.1	Durchführung der explorativen Faktorenanalyse (EFA).....	76
4.9.2.2	Durchführung der Reliabilitätsanalyse .....	79
4.9.2.3	Durchführung der konfirmatorischen Faktorenanalyse (CFA).....	81
4.9.2.4	Berechnung der Korrelationen auf Faktorebene .....	82
4.9.2.5	Berechnung binär logistischer Regressionen.....	82
4.9.2.6	Berechnung linearer Regressionen zur Exploration .....	84
5	Diskussion .....	87
5.1	Interpretation und Beurteilung der quantitativen Ergebnisse .....	87
5.1.1	Stichprobe .....	87
5.1.2	Häufigkeit und Testarten (OF1 und OF2).....	87
5.1.3	Qualität (UF1) .....	88
5.1.4	Zielgruppe (UF2).....	89

5.1.5	Zweck (UF3).....	89
5.1.6	Informationsbasis (UF4).....	89
5.1.7	Hintergrund (Einflussfaktoren, UF5 bis 7) .....	90
5.1.8	Beantwortung der Hypothesen .....	92
5.2	Gestaltungshinweise .....	93
5.2.1	Zentralisieren von Testrezensionen .....	93
5.2.2	Informieren und Sensibilisieren der Testanwendenden.....	94
5.3	Limitationen in der Methodik.....	95
5.3.1	Repräsentativität hinsichtlich Stichprobenzusammensetzung.....	95
5.3.2	Repräsentativität betreffend die Stichprobengröße.....	95
5.3.3	Unterschiede in Regressionskoeffizienten: binär vs. linear .....	95
5.3.4	Verzicht auf weitere explorative Auswertungen .....	96
5.4	Schlusswort.....	96
6	Literaturverzeichnis.....	97
Anhang A	.....	104
Anhang B	.....	106
Anhang C	.....	108
Anhang D	.....	109
Anhang E	.....	110
Anhang F	.....	111
Anhang G	.....	112
Anhang H	.....	114
Anhang I	.....	115
Anhang J	.....	127
Anhang K	.....	132
Anhang L	.....	134
Anhang M	.....	135
Anhang N	.....	137
Anhang O	.....	139
Anhang P	.....	141
Anhang Q	.....	142
Anhang R	.....	144
Anhang S	.....	147

# 1 Ausgangslage und Zielsetzung

Aktuelle Zahlen des Bundesamtes für Statistik (BFS) zeigen, dass im Jahr 2022 588'893 Personen in der Schweiz einen neuen Arbeitsvertrag mit einem Unternehmen eingegangen sind. Das bedeutet, dass 588'893 Personen einen Personalauswahlprozess in ihrem Unternehmen durchlaufen haben<sup>1</sup>, bzw. 588'893 Personalauswahlprozesse für diese Personen positiv verlaufen sind.

Personalauswahlprozesse bzw. Personalentscheide werden von Unternehmen in der Schweiz vielfältig gestaltet. Studien zeigen, dass Interviews und das Analysieren von Bewerbungsunterlagen beinahe immer durchgeführt werden. In 90% aller Personalauswahlentscheiden werden Referenzen eingeholt (Berchtold, 2005; König, Kehle, Berchtold & Kleinmann, 2010). Die Rolle psychologischer Tests, also von Persönlichkeitsinventaren und Leistungstests, ist demgegenüber eher untergeordnet. Die Studie von König et al. (2010) zeigt auf, dass in der Schweiz die Einsatzhäufigkeit von Persönlichkeitstests bei 32% liegt, diejenige von Leistungstests bei 18.6%.

Psychometrische Testverfahren, wie Intelligenztests oder Persönlichkeitsinventare, werden zur Personaldiagnostik eingesetzt. Das Ziel von Personaldiagnostik im Bereich der Personalauswahl ist die Sicherstellung der Passung zwischen den individuellen Kompetenzen und Eigenschaften des Bewerbenden, und den unternehmensspezifischen Arbeitsanforderungen. Im Bereich der Personalentwicklung dient die Personaldiagnostik der Optimierung der Passung zwischen den individuellen Kompetenzen und Eigenschaften des Mitarbeitenden und den unternehmensspezifischen Anforderungen. Kurz gesagt: durch den gezielten Einsatz psychometrischer Testverfahren sollen individuelle Kompetenzen beobachtbar gemacht und ungenutztes Potenzial entdeckt werden (Schuler, 2000).

Das Angebot für solche Tests ist mannigfaltig. Das führt dazu, dass Fachexpertinnen und -experten den Testmarkt als weitgehend intransparent und die Orientierung darin als kaum bewältigbar empfinden (Hell, 2013) – geschweige denn Kundinnen und Kunden ohne fachlichen Hintergrund betreffend Eignungsdiagnostik (Hossiep & Mühlhaus, 2005). Zweifelhafte Testangebote mit eindrücklichen Websites und einem guten Marketing profitieren von diesem Umstand (Hell, 2013).

Durch Testrezensionen wird versucht, der geschilderten Intransparenz entgegenzuwirken (Hell, 2013; Kersting, 2006). Es gibt eine Vielzahl an Beurteilungen psychologischer Tests in diversen Datenbanken und Fachzeitschriften. Deren Qualität ist aufgrund unsystematischer Bewertungsprozesse und damit unstandardisierter Beurteilungen, als eher kritisch zu betrachten (Kersting, 2006), was deshalb zu weiterer Intransparenz führt. Abhilfe soll das *Test-Bewertungssystem* (TBS-DTK) schaffen, welches durch das *Testkuratorium der Föderation Deutscher Psychologinnenvereinigungen* erarbeitet wurde. Jenes wird zur standardisierten Bewertung psychologischer Tests durch das eben genannte Testkuratorium seit 2010 verwendet (Hell, 2013; Beauducel & Leue, 2014). Das Problematische: es wird kaum eingesetzt (Hell, 2013). Aktuell befinden sich 100 Rezensionen im TBS-DTK. Die Intransparenz wächst mit jedem weiteren Testangebot auf dem Markt.

---

<sup>1</sup> Die aktuellen Zahlen des BFS legen dar: im Jahr 2022 waren 5'181'000 Personen in der Schweiz erwerbstätig (BFS, 2023b). Abgezogen von den 11.2% der selbstständig Erwerbenden ergibt dies 4'600'728 unselbstständig Erwerbende. Im Jahr 2022 haben 8.9% eine Arbeitstätigkeit in einem neuen Unternehmen aufgenommen. 3.9% haben intern ihre Stelle gewechselt (BFS, 2024). Das bedeutet, dass im Jahr 2022 12.8% aller unselbstständig Erwerbenden, demnach 588'893 Personen einen neuen Arbeitsvertrag mit einem Unternehmen eingegangen sind.

Im deutschsprachigen Raum gibt es schätzungsweise 3'000 Persönlichkeitstests (Hell, 2013), aber nicht alle sind stichhaltig und fundiert (Urech, 2020). Sucht man auf PubPsych – einer frei zugänglichen Informationsplattform, welche u.a. auch psychologische Tests führt – nach Persönlichkeitstests, erhält man 3'440 Ergebnisse. Sucht man auf ebendieser Plattform nach Intelligenztests, erhält man 519 Ergebnisse (ZPID, 2024).

Zweifelhafte Persönlichkeitsinventare halten sich schon über Jahrzehnte am Markt (Hell, 2013). Bspw. setzten 2005 in der Schweiz 33% derjenigen Unternehmen, die zur Personalauswahl auch Persönlichkeitstests einsetzen, den umstrittenen *Master Person Analysis (MPA)*<sup>2</sup> ein. 17% verwendeten den *Myers-Briggs Typ Indicator (MBTI)*<sup>3</sup> (Urech, 2020). Aktuelle Daten über die Einsatzhäufigkeit anderer Persönlichkeitstests oder Leistungstests zur Personalauswahl und -entwicklung in Schweizer Unternehmen fehlen. Damit fehlen auch Daten, für welche Zielgruppen und zu welchem Zweck Tests eingesetzt werden oder welche Einflüsse den Testeinsatz in einem Unternehmen begünstigen oder hemmen. Denkbar sind hier drei verschiedene Einflussgrößen, welche nachfolgend genannt und definiert werden:

### 1. persönliche Beweggründe

Der Duden umschreibt den Begriff Beweggrund als *einen Menschen zu einer Handlung bewogender Grund* und nennt den Begriff *Motiv* als Synonym (Duden, 2023). So führen bspw. vergangene Erfahrungen oder persönliche Vorlieben zur Bildung eines persönlichen Beweggrundes.

### 2. Abwägung von Entscheidungskriterien

Rationale Entscheidungen werden in der Regel entlang eines Kosten-Nutzen-Kalküls vollzogen. Man wägt also ab zwischen dem zu erwartenden Nutzen und den dafür zu investierenden Kosten (Myers, 2014). Damit unterscheidet sich der Begriff der Entscheidungskriterien von den persönlichen Beweggründen, die vergleichsweise einen weniger rationalen, dafür eher emotionalen Hintergrund haben.

### 3. kontextuelle Einflussgrößen

Beweggründe und Entscheidungskriterien sind persönlich gefärbt. Einzelne Personen bzw. Personengruppen treffen Entscheidungen anhand von Beweggründen und Entscheidungskriterien. Kontextuelle Einflussgrößen unterscheiden sich davon. Der Begriff spricht externale Einflussgrößen an, die z.B. organisational oder branchenspezifisch begründet sind. Typische Beispiele hierfür sind u.a. die Grösse des Unternehmens oder finanzielle Ressourcen (Holtbrügge, 2004).

Die Zielsetzung dieser Masterarbeit besteht also darin, einen Überblick über eingesetzte Tests gewinnen zu können. Ebenfalls soll geklärt werden, welche Zwecke durch den Einsatz von Tests verfolgt werden und für welche Personen bzw. Positionen diese eingesetzt werden. Zudem sollen mögliche Beweggründe, Entscheidungskriterien und Einflussgrößen eruiert werden, die den Einsatz von Tests in einem Unternehmen begünstigen oder hemmen. Dazu wurden Fragestellungen formuliert, welche für die vorliegende Masterthesis leitend sind (siehe Tab. 1). Diese wurden in Zusammenarbeit mit der betreuenden bzw. auftraggebenden Person, Prof. Dr. Benedikt Hell, erstellt.

---

<sup>2</sup> Der MPA beruht auf veralteten psychoanalytischen Annahmen, womit eine theoretische Fundierung gänzlich fehlt (Urech, 2020).

<sup>3</sup> Auch der MBTI – in den 1940er Jahren entwickelt – beruht auf veralteten psychoanalytischen Annahmen (Urech, 2020). Hinzukommend: Getestete hier werden in Gruppen eingeteilt (Hell, 2013).

Tabelle 1  
Fragestellungen, die in der vorliegenden Thesis bearbeitet werden.

<b>Oberfragestellungen (OF)</b>
<b>Oberfragestellung OF 1 - Verwendungshäufigkeit</b>
Wie häufig werden Leistungstests und Persönlichkeitsinventare in Schweizer Unternehmen zur Personalauswahl und zur Personalentwicklung eingesetzt?
<b>Oberfragestellung OF2 - Testart</b>
Welche Tests werden zur Personalauswahl und zur Personalentwicklung verwendet?
<b>Unterfragestellungen (UF)</b>
<b>Unterfragestellung UF1 - Qualität eingesetzter Tests</b>
Zu welchen der eingesetzten Verfahren liegt eine Testrezension nach TBS-DTK oder Schweizerische Dienstleistungszentrum Berufsbildung (SDBB) vor?
<b>Unterfragestellung UF2 - Einsatzzweck</b>
Für welchen Zweck werden Tests eingesetzt – Personalauswahl und / oder Personalentwicklung?
<b>Unterfragestellung UF3 - Zielgruppen</b>
Für welche Zielgruppen werden Tests in der Personalauswahl eingesetzt?
<b>Unterfragestellung UF4 - Informationsbasis</b>
Auf Basis welcher Informationen wird die Entscheidung für den Einsatz eines psychologischen Tests gefällt?
<b>Unterfragestellungen UF5 bis UF7 - Hintergründe in Bezug auf Testeinsatz</b>
UF5: Welche Beweggründe der Anwendenden stehen hinter der Wahl von Tests?
UF6: Welche Entscheidungskriterien der Anwendenden stehen hinter der Wahl von Tests?
UF7: Welche externalen Einflussgrössen auf die Wahl bzw. den Einsatz von Tests zur Personalauswahl oder -entwicklung sind entscheidend?

Die entsprechenden Hypothesen werden am Ende des Theoriekapitels präsentiert.

## 1.1 Orientierung

Der Aufbau dieser Thesis folgt der Chronologie der durchgeführten Arbeitsschritte. In Kap. 2 wird die theoretische Grundlage dargelegt. Darauf folgt in Kap. 3 eine ausführliche Beschreibung der Methodik der Datenerhebung. Hier werden auch die qualitativen Zwischenergebnisse besprochen. Kap. 4 widmet sich der Darstellung der quantitativen Hauptergebnisse. Den Abschluss bildet Kap. 5, in dem die Ergebnisse und Limitationen diskutiert und Gestaltungshinweise formuliert werden.

Der Anhang dieser Thesis ist in verschiedene Abschnitte gegliedert. Anhang A enthält ein Abkürzungsverzeichnis, das das Lesen der Thesis erleichtern soll. Auf weitere Anhänge wird im Laufe der Thesis verwiesen.

## 2 Im Zentrum stehende Theorien, Konzepte und Modelle

Dieses Kapitel beschäftigt sich mit Modellen und Studien, die als theoretische Grundlage für die vorliegende Thesis dienen. Zunächst folgt eine Definition psychologischer Tests im Allgemeinen und deren Inhalt. Danach werden Konstrukte besprochen, die im Rahmen der Personalauswahl durch Tests operationalisiert werden.

Daran anschliessend folgt die Besprechung der Gütekriterien, die psychologische Tests erfüllen müssen. Ein besonderer Fokus liegt hier auf den Gütekriterien Kriteriumsvalidität und inkrementeller Validität, die aufzeigen, welchen Nutzen man aus dem Einsatz psychologischer Tests im Rahmen von Personalauswahl ziehen kann.

Dann werden aktuelle Studienergebnisse dargelegt, die sich mit der Verwendungshäufigkeit psychologischer Test und den Gründen für bzw. gegen den Testeinsatz auseinandersetzen. Im letzten Unterkapitel werden dann die aus der Theorie abgeleiteten Hypothesen präsentiert, die für die vorliegende Masterthesis handlungsleitend sind.

### 2.1 Psychologische Tests

«Psychologische Tests sind Instrumente, die der systematischen Beobachtung und Beschreibung von Erleben und Verhalten mithilfe von Skalen (numerische Beschreibung) oder Kategorien (klassifizierende Beschreibung) dienen.»  
(Krohne & Hock, 2015, S.28).

Ein psychologischer Test dient demnach der standardisierten Messung von individuellen Eigenschaften, die nicht direkt beobachtbar sind. Durch Tests lassen sich Eigenschaften, wie Persönlichkeit oder Leistungsfähigkeit operationalisieren, also messbar machen. Die Resultate lassen Schlüsse auf individuelle Eigenschaftsausprägungen zu und erlauben Prognosen für Verhalten ausserhalb der Testsituation (Schuler, 2013). Für die berufliche Eignungsdiagnostik ist das Schliessen auf zukünftige berufliche Leistung von besonderem Interesse (Neyer & Asendorpf, 2018; Schuler, 2013; Krohne & Hock, 2015).

In der eigenschaftsorientierten Eignungsdiagnostik unterscheidet man grundsätzlich zwei Arten von Tests: Leistungstests und Fragebogen. Bei Leistungstests ist man am maximalen Verhalten interessiert (Krohne & Hock, 2015). Bewerbende müssen in vorgegebener Zeit so viele Aufgaben wie möglich korrekt lösen und so ihre Fähigkeiten unter Beweis stellen (Kanning, 2015). Bei Fragebogen ist die Zeitvorgabe sekundär. Der oder die Teilnehmende nimmt eine Selbsteinschätzung in Bezug auf Aussagen zum eigenen Charakter auf einer mehrstufigen Skala vor (Kanning, 2015). Hier ist das typische Verhalten von Interesse, wobei keine formale Vorgabe besteht, was als richtig und als falsch zu werten ist (Krohne & Hock, 2015). Insbesondere Persönlichkeitseigenschaften und Interessen werden mittels Fragebogen erfasst.

Psychologische Tests, die zum Zweck beruflicher Eignungsdiagnostik eingesetzt werden, widmen sich insb. den Persönlichkeitseigenschaften, den Fähigkeiten und den Handlungseigenschaften von Bewerbenden (Neyer & Asendorpf, 2018).

Persönlichkeitseigenschaften lassen sich durch Persönlichkeitsmodelle fassbar machen. Dazu wird in Kap. 2.1.1 das etablierte und wissenschaftlich fundierte Big Five-Modell besprochen.

Zu den Fähigkeiten lassen sich neben allgemeiner und emotionaler Intelligenz auch Fachwissen, sensumotorische Fähigkeiten und Konzentrationsfähigkeit zählen (Neyer & Asendorpf, 2018). Die vorliegenden Masterthesis beschäftigt sich mit den Fähigkeiten allgemeiner und emotionaler

Intelligenz (siehe Kap. 2.1.2). Fachkenntnistests werden trotz ihrer hohen Prognosekraft für berufliche Leistung nicht weiter besprochen. Ähnlich wie bei sensumotorischen Tests oder Tests zu Konzentration werden diese nur zu spezifischen Auswahlzwecken eingesetzt. Zu den Handlungseigenschaften zählen u.a. Interessen und Motivation (Neyer & Asendorpf, 2018). Während das Konstrukt der beruflichen Interessen in Kap. 2.1.3 noch vertieft wird, wird das Konzept der Motivation ausgeklammert, da deren theoretische Grundlage umstritten ist. Zudem bestehen Herausforderungen bei der Anwendung von Motivationstests, da sozial erwünschtes Antwortverhalten oder Faking die Nützlichkeit im Vergleich zu anderen Fragebogen einschränkt (Ryan & Ployhart, 2013).

Die Abb. 1 bietet einen Überblick über diejenigen Konstrukte, entsprechenden Tests und Testformen, denen sich die folgenden Unterkapitel widmen.

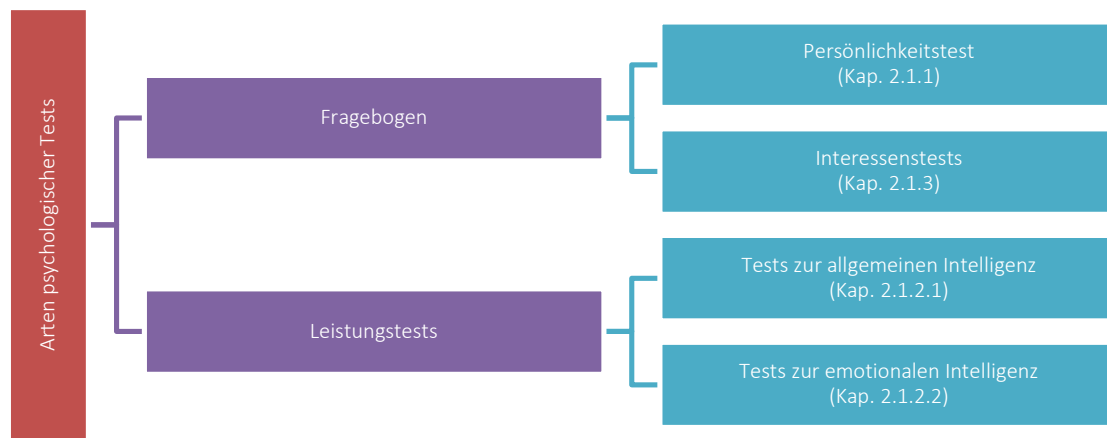


Abbildung 1. Übersicht über die verschiedenen Arten psychologischer Tests, welche zur Personalauswahl eingesetzt werden. Orientiert an Schuler, 2004; Krings, 2017 und Bühner, 2011.

### 2.1.1 Persönlichkeitseigenschaften im Interesse der Eignungsdiagnostik

Um Persönlichkeit beschreiben zu können, hat sich das Big Five-Modell durchgesetzt. Dieses ist wissenschaftlich überprüft und gilt als das meistverwendete Persönlichkeitsmodell (Myers, 2014). Das Big Five-Modell umfasst die fünf Persönlichkeitsfaktoren *Offenheit für Erfahrungen*, *Gewissenhaftigkeit*, *Extraversion*, *Verträglichkeit* und *emotionale Stabilität*. Jeder dieser Faktoren wird in je weitere sechs Facetten unterteilt. Jeder der fünf Faktoren und jede untergeordnete Facette entspricht einem Kontinuum, auf welchem sich Personen hinsichtlich ihrer persönlichen Ausprägungen einteilen lassen (Neyer & Asendorpf, 2018). Die fünf Persönlichkeitsfaktoren werden nachfolgend beschrieben:

**Offenheit für Erfahrungen** umfasst Merkmale, wie das Interesse an neuen Eindrücken und Erfahrungen (Rammstedt, Kemper, Klein, Beierlein & Kovaleva, 2013). Personen mit einer niedrigen Ausprägung auf diesem Persönlichkeitsfaktor sind praktisch veranlagt, sachorientiert, unflexibel, faktenorientiert und traditionsbewusst. Personen mit einer hohen Ausprägung lassen sich als kreativ, feinfühlig, experimentierfreudig, neugierig und aufgeschlossen beschreiben (Neyer & Asendorpf, 2018).

**Gewissenhaftigkeit** entspricht der Eigenschaft, kontrolliert zu handeln. Personen mit einer hohen Ausprägung zeigen sich umsichtig, organisiert, zuverlässig, fleissig und reflektiert. Personen mit einer niedrigen Ausprägung neigen zu Planlosigkeit, sind unordentlich, ziellos, chaotisch und unreflektiert (Neyer & Asendorpf, 2018).

**Extraversion** beschreibt interpersonelles Verhalten, welches sich durch Geselligkeit, Gesprächigkeit und Durchsetzungsfähigkeit auszeichnet. Personen mit einer hohen Ausprägung zeigen sich lebhaft und fröhlich, sind risikofreudig und dominant. Personen mit einer niedrigen Ausprägung nennt man auch introvertiert. Introvertierte Personen neigen zur Vorsicht, zeigen sich ernst und distanziert, sind zurückhaltend und unterwürfig (Neyer & Asendorpf, 2018).

**Verträglichkeit** beschreibt ebenfalls interpersonelles Verhalten (Rammstedt et al., 2013). Personen mit einer hohen Ausprägung sind gutgläubig, bescheiden und hilfsbereit, verhalten sich ehrlich und entgegenkommend. Personen mit einer niedrigen Ausprägung hingegen sind misstrauisch, berechnend, unbarmherzig, verhalten sich unehrlich und egoistisch (Neyer & Asendorpf, 2018).

**Emotionale Stabilität** korreliert stark mit der Häufigkeit und der Intensität positiver Emotionen. Der Gegenpol der emotionalen Stabilität – Neurotizismus – korreliert hingegen stark mit der Häufigkeit und Intensität negativer Emotionen. Neurotische Personen neigen zu Angst und Nervosität, zeigen sich impulsiv, schwermütig, schüchtern und sensibel. Emotional stabile Personen hingegen zeigen sich als unerschütterlich, ausgeglichen, selbstsicher und stressresistent (Neyer & Asendorpf, 2018).

Es gibt eine Vielzahl von Tests auf dem Markt, die sich mit der Operationalisierung der Big Five Persönlichkeitsfaktoren beschäftigen, wie bspw. das NEO-Persönlichkeitsinventar (NEO-PI-R), das Big Five Inventar 10 (BFI-10) oder das Big Five Inventar 2 (BFI-2). Diese genannten Inventare unterscheiden sich in der Menge der Items und damit der Ausfülldauer.

Neuere Strömungen sehen im *HEXACO-Modell* die Zukunft zur Einordnung von Persönlichkeitseigenschaften. Das HEXACO-Modell umfasst die sechs Persönlichkeitsfaktoren *Offenheit, Gewissenhaftigkeit, Verträglichkeit, Extraversion, Emotionalität und Ehrlichkeit-Bescheidenheit* (Neyer & Asendorpf, 2018).

Das HEXACO-Modell ist aber keineswegs eine Erweiterung der Big Five-Modells. Grundsätzlich sind im HEXACO-Modell dieselben Facetten vertreten wie im Big Five-Modell, nur werden diese unter sechs und nicht unter fünf Faktoren aufgeteilt. Stellt man das HEXACO-Modell dem Big Five-Modell gegenüber, entsprechen die gleichbenannten Facetten *Extraversion, Gewissenhaftigkeit* und *Offenheit* einander grösstenteils. Big Five-Facetten der *emotionalen Stabilität* werden der HEXACO-Facette *Verträglichkeit* zugeordnet und Big Five-Facetten der *Verträglichkeit* der HEXACO-Facette *Ehrlichkeit-Bescheidenheit* (Iller, Mäder & Schreiber, 2021).

Auch zur Operationalisierung des HEXACO-Modells bestehen Instrumente. So bspw. der *HEXAPRO* oder das HEXACO-Persönlichkeitsinventar (HEXACO-PIR), welches durch die Verfechter des HEXACO-Modells Michael Ashton und Kibeom Lee entwickelt wurde (Neyer & Asendorpf, 2018).

Eine Persönlichkeitsdimension, die für die Personalauswahl ebenfalls von Bedeutung ist, ist die *Integrität*, also die Vertrauenswürdigkeit. Doch Integrität kann nicht als eigenständiges Persönlichkeitsmerkmal verstanden werden (Neyer & Asendorpf). Bereits die Definitionen der Big Five Faktoren *Gewissenhaftigkeit* und *Verträglichkeit* lassen erahnen, dass wohl Assoziationen zu Integrität bestehen. Die Metaanalyse von Marcus, Funke und Schuler (1997, in Neyer & Asendorpf, 2018) zeigt moderate bis hohe Korrelationen von Integrität zu vier der fünf Big Five Persönlichkeitsfaktoren. Integrität korreliert insb. mit Gewissenhaftigkeit ( $\rho = .45$ ) und Verträglichkeit ( $\rho = .48$ ), aber auch mit emotionaler Stabilität ( $\rho = .31$ ) und Extraversion ( $\rho = .23$ ) (Marcus et al., 1997; zitiert nach Neyer & Asendorpf, 2018). Die Konstruktvalidität von Integrität bzw. von Integritätstests muss demnach in Frage gestellt werden. Die Kriteriumsvalidität, also die Prognosekraft betreffend die berufliche Leistung hingegen, ist moderat bis hoch, wie das noch folgende Kap. 2.2.1.3.3 aufzeigen wird.

Integritätsfragebogen, die in der Praxis eingesetzt werden, sind z.B. der psychologische Integritätstest (PIT) oder das Persönlichkeitsinventar zur Integritätsabschätzung (PIA) (Neyer & Asendorpf, 2018).

### 2.1.1.1 Kategoriale-typisierende vs. dimensionale Persönlichkeitsabfrage

Die Unterscheidung zwischen einer Taxonomie von Personen und einer Taxonomie von Dimensionen widerspiegelt die Unterscheidung zwischen dem typologischen und dem dimensionalen Ansatz von Persönlichkeitstests, so Robins, John und Caspi (1998).

Während die vorgestellten Persönlichkeitsmodelle und -faktoren einen dimensionalen Ansatz verfolgen und Persönlichkeit als Kontinuum betrachten, gibt es auch Persönlichkeitstests, die einen kategorialen oder typisierenden Ansatz verwenden (Robins et al., 1998). Bei diesen Verfahren werden Personen in Gruppen eingeteilt (Robins et al., 1998). Tests wie der *MBTI*, der *DISC*<sup>4</sup> oder der *Insights* nutzen Taxonomien, die ihren Ursprung in der Psychoanalyse, speziell bei C.G. Jung, haben. Diese verwendeten Taxonomien und Theorien sind jedoch nicht wissenschaftlich fundiert, weshalb die genannten Tests und deren Nutzung in der Kritik stehen (Urech, 2020; Hell, 2013; Hossiep, Weiss & Schecke, 2015).

### 2.1.2 Fähigkeiten im Interesse der Eignungsdiagnostik

«Fähigkeiten sind Persönlichkeitseigenschaften, die Leistungen ermöglichen.»  
(Asendorpf, 2019, S. 93).

Dieses Kapitel widmet sich den Fähigkeiten der allgemeinen und der emotionalen Intelligenz. Während die allgemeine Intelligenz in der beruflichen Eignungsdiagnostik seit mehr als 100 Jahren beforscht wird, handelt es sich bei der emotionalen Intelligenz um ein eher neueres Konstrukt, welches seit einigen wenigen Jahrzehnten in die Forschung beruflicher Eignungsdiagnostik miteinbezogen wird (Ryan & Ployhart, 2013).

#### 2.1.2.1 Allgemeine Intelligenz

«Intelligenz ist die Fähigkeit des Individuums, zweckvoll zu handeln, vernünftig zu denken und sich mit seiner Umgebung wirkungsvoll auseinanderzusetzen.»  
(Wechsler, 1944; zitiert nach Petermann und Daseking, 2015, S. 99).

Intelligenz ist ein komplexes Konstrukt, das verschiedene kognitive Fähigkeiten umfasst, die es einem Menschen ermöglichen, sich erfolgreich an seine Umwelt anzupassen und Probleme adäquat zu lösen (Petermann und Daseking, 2015). Das Cattell-Horn-Carroll-Modell (CHC-Modell) durchbricht diese Komplexität und unterteilt Intelligenz in unterschiedliche Faktoren. Das CHC-Modell ist ein wissenschaftlich validiertes und das bis heute wohl am besten untersuchte Intelligenzmodell (Neyer & Asendorpf, 2018). Es vereint Theorien, Modelle und Erkenntnisse verschiedener Forschenden. Neben den Namensgebern Cattell, Horn und Carroll sind auch die Forschungen von Spearman, Binet oder Guilford darin berücksichtigt worden. Das CHC-Modell beschreibt die allgemeine Intelligenz als Zusammenspiel bzw. als Summe zehn verschiedener Intelligenzfaktoren. Diese Intelligenzfaktoren bestehen wiederum aus zahlreichen spezifischen kognitiven Fähigkeiten, die über verschiedenste Testaufgaben erfasst werden können (Petermann & Daseking, 2015; Neyer & Asendorpf, 2018). Die nachfolgende Abb. 2 stellt das CHC-Modell vereinfacht dar und bietet einen Überblick über gängige Aufgaben, die der Operationalisierung der spezifischen kognitiven Fähigkeiten dienen.

---

<sup>4</sup> DISC ist ein Akronym für diejenigen Facetten, die im Instrument erfasst werden: Dominance, Influence, Steadiness, Conscientiousness.

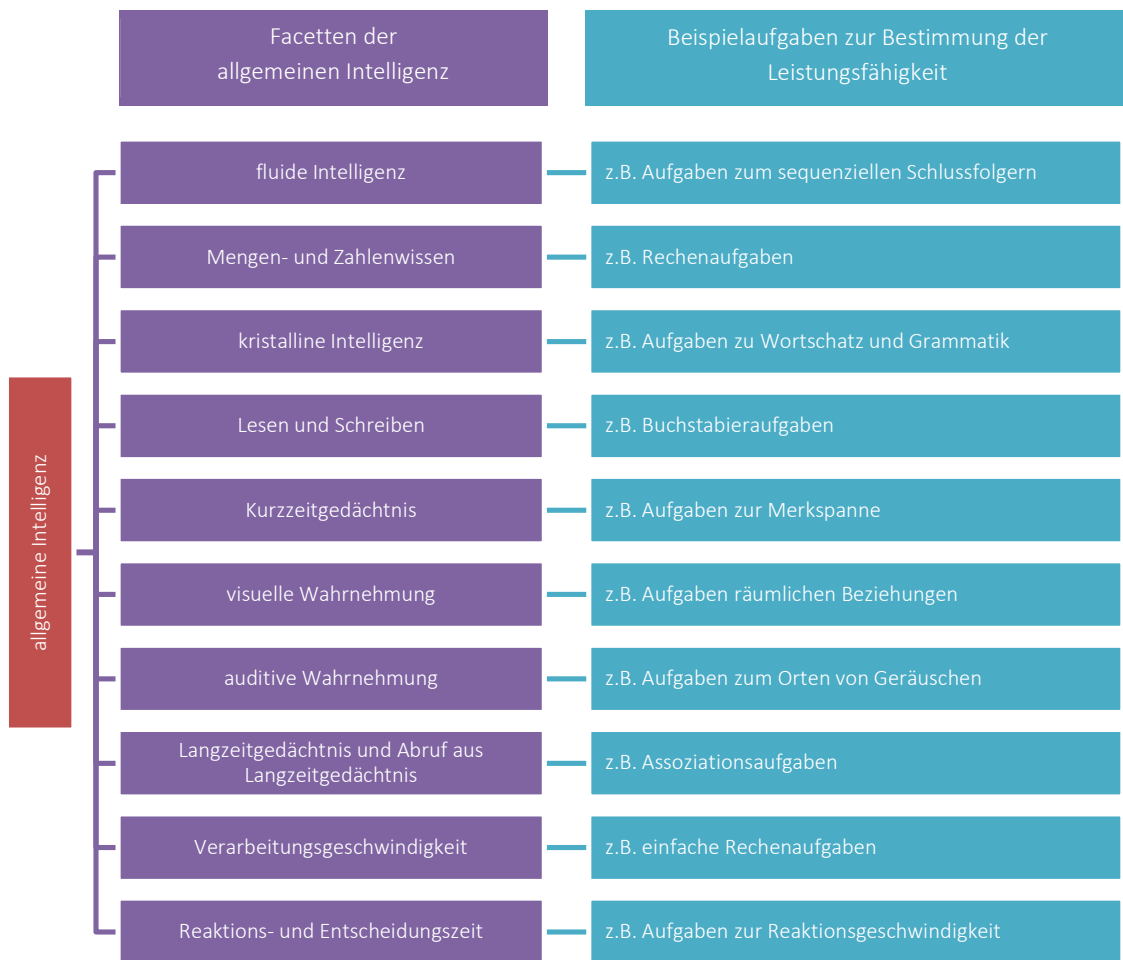


Abbildung 2. CHC-Modell. Aus Diagnostische Erhebungsverfahren von F. Petermann & M. Daseking, 2015 (S. 104). Göttingen: Hogrefe.

Die allgemeine Intelligenz kann durch standardisierte Tests gemessen werden, die sich dazu u.a. die in Abb. 2 genannten Aufgaben zunutze machen. Deren Ergebnis wird als Intelligenzquotient (IQ) ausgedrückt. Der IQ bietet Auskunft über die Position der getesteten Person innerhalb einer Population gleichen Alters (Petermann und Daseking, 2015). Eine solche Einteilung ist möglich, weil die Intelligenz einer Normalverteilung folgt und eine entsprechende Normierung vorgenommen wurde (Neyer & Asendorpf, 2018). Die Abb. 3 zeigt die Normalverteilung der IQ-Werte in der Population. Die daran abgebildeten Werte werden daran anschliessend erläutert und in mit dem Thema der Normierung verknüpft.

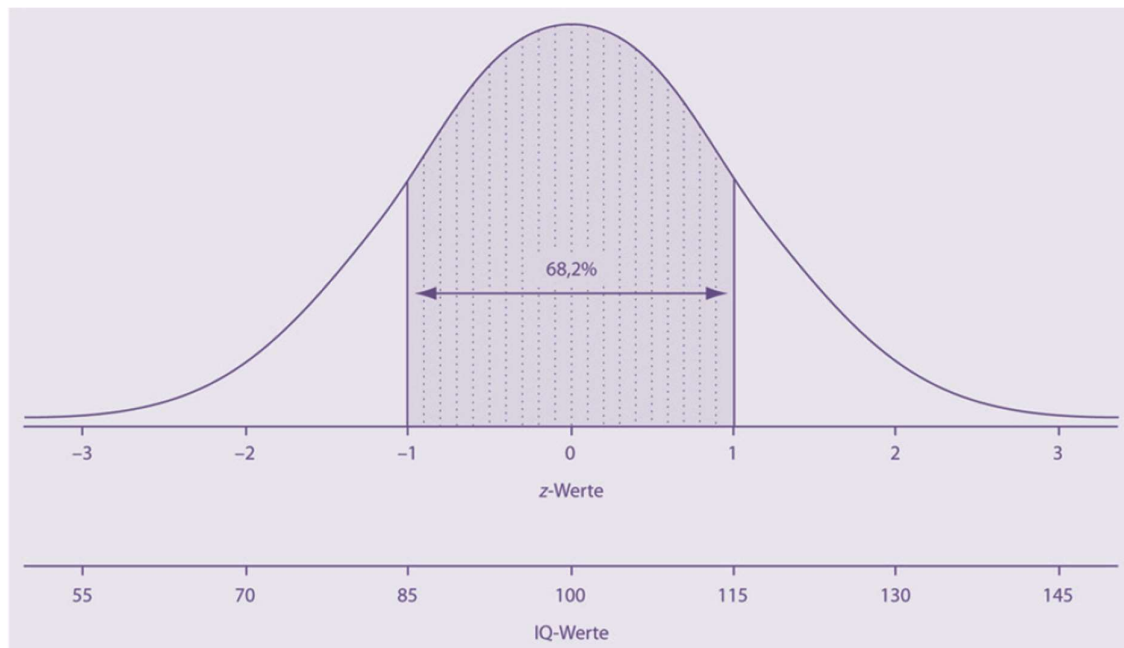


Abbildung 3. Normalverteilungskurve der IQ-Werte. Aus Psychologie der Persönlichkeit von F. Neyer & J. Asendorpf, 2011 (S. 153). Berlin: Springer.

Durch Normierung können IQ-Werte einfach interpretiert werden. IQ-Werte zwischen 85 und 115 entsprechen dem Normbereich der Intelligenz, welchem sich 68.2% der Population zuordnen lässt (Neyer & Asendorpf, 2018). IQ-Werte unter 85 entsprechen unterdurchschnittlicher Intelligenz, Werte über 115 entsprechen überdurchschnittlicher Intelligenz, wobei man ab IQ-Werten von 130 von Hochbegabung spricht (Krohne & Hock, 2015).

Bekannte Tests zur Messung der Intelligenz sind der Wechsler Adult Intelligence Scale (WAIS), der Intelligenz-Struktur-Test 2000R (IST-2000R) oder dessen entsprechender Kurztest *IST-Screening* (Petermann & Daseking, 2015).

### 2.1.2.2 Emotionale Intelligenz

Emotionale Intelligenz bezieht sich auf die Fähigkeit zum Umgang mit den eigenen Emotionen und den Emotionen anderer. Die emotionale Intelligenz entspricht dabei eher einer unbewussten als einer bewussten Leistung. So werden erhaltene Informationen zu Emotionen automatisch verarbeitet (Neyer & Asendorpf, 2018).

In der nun folgenden Abb. 4 werden die von Mayer, Salovey, Caruso und Sitarenios (2001) definierten Facetten der emotionalen Intelligenz dargelegt und definiert. Daran anschließend wird das Konzept der emotionalen Intelligenz diskutiert.

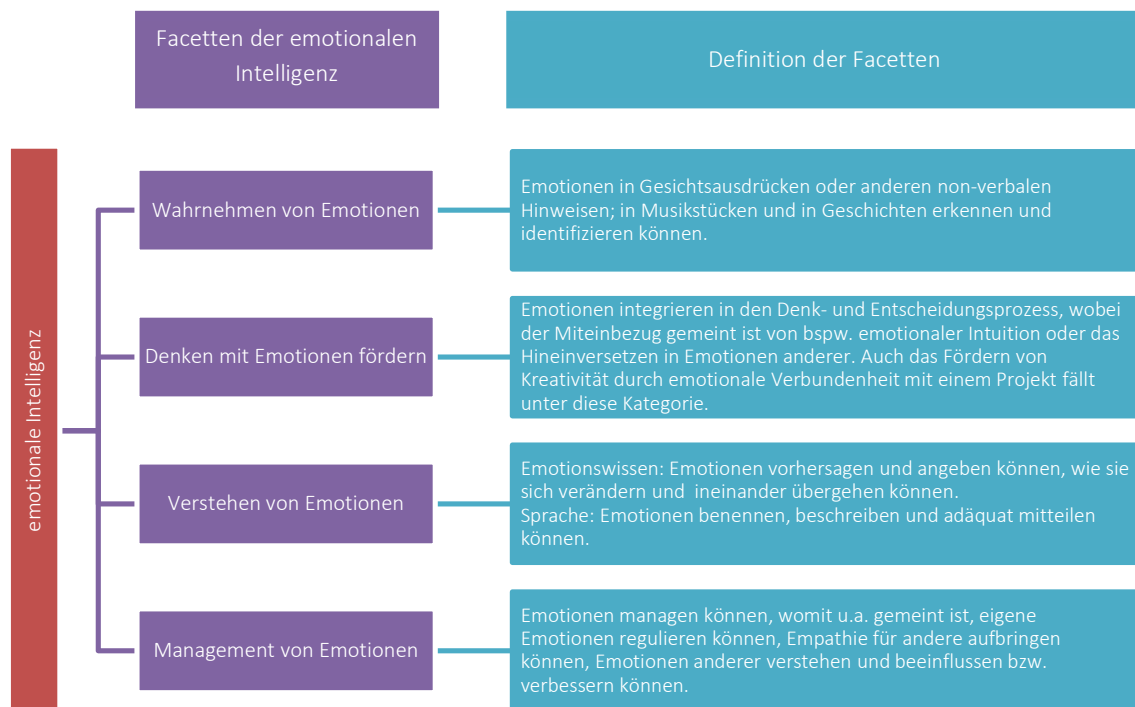


Abbildung 4. Die vier Fähigkeitsbereiche emotionaler Intelligenz. Aus Emotional intelligence as a standard intelligence von D. Mayer, P. Salovey, R. Caruso & G. Sitarenios (2001). *Emotion*, 1(3), 232–242.

Auf Basis der vier Facetten emotionaler Intelligenz (siehe Abb. 4) wurde der Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) entwickelt. Der MSCEIT entspricht einem Leistungstest, der, ähnlich einem Intelligenztests, den *Emotional Intelligence Quotient* (EIQ) als Resultat herausgibt. Der MSCEIT ist der heute am meisten eingesetzte Test zur Messung der emotionalen Intelligenz (Hogrefe Testzentrale, 2024). Im TBS-DTK ist der MSCEIT ebenfalls geführt, wobei die Rezension darauf verweist, dass Weiterentwicklungen notwendig und Untersuchungen bzw. Vergleiche mit ähnlichen Tests durchzuführen seien.

Van der Linden, Pekaar, Bakker, Schermer, Vernon, Dunkel, et al. (2017) haben in ihrer Metaanalyse eine moderate Korrelation von  $\rho = .27$  zwischen der Fähigkeit emotionaler Intelligenz und dem Generals Factor of Personality (GFP) gefunden. Ein hoher GFP entspricht hohen Ausprägungen in den Big Five-Persönlichkeitsmerkmalen *Verträglichkeit*, *Extraversion*, *emotionale Stabilität*, *Gewissenhaftigkeit* und *Offenheit*. Van der Linden et al. (2017) erklären, dass das Konzept der emotionalen Intelligenz kaum über Konstruktvalidität verfügt, aber Kriteriumsvalidität betreffend die berufliche Leistung sehr wohl vorhanden ist. Die emotionale Intelligenz gilt deshalb als moderater Prädiktor für berufliche Leistung, weil sie wichtige Prädiktoren für berufliche Leistung – die GFP – in sich vereint. Die Kriteriumsvalidität der emotionalen Intelligenz wird in dem noch folgenden Kap. 2.2.1.3.3 dargelegt.

### 2.1.3 Interesse

Interessen entsprechen den individuellen Bewertungen von gewissen Handlungen bzw. Tätigkeiten. Menschen unterscheiden sich darin, ob sie gewisse Tätigkeiten als anziehend und damit als interessant oder im Gegenteil, als abstossend und somit uninteressant empfinden (Neyer & Asendorpf, 2018). Entlang diesem Paradigma orientiert sich auch das Berufsinteressenmodell nach Holland. Menschen suchen nach beruflichen Umwelten, bei denen eine hohe Übereinstimmung mit ihren individuellen Interessensausprägungen vorherrschen (Schuler, Höft & Hell, 2014).

Das Berufsinteressenmodell nach Holland unterscheidet sechs Faktoren beruflichen Interesses, welche sich als Sechseckstruktur darstellen lassen. Dabei liegen ähnliche Interessenbereiche nebeneinander. Die Abb. 5 zeigt das Modell. Die sechs Interessensbereiche sind mit Buchstaben gekennzeichnet, deren Bedeutungen ebenfalls in Abb. 5 erläutert werden.

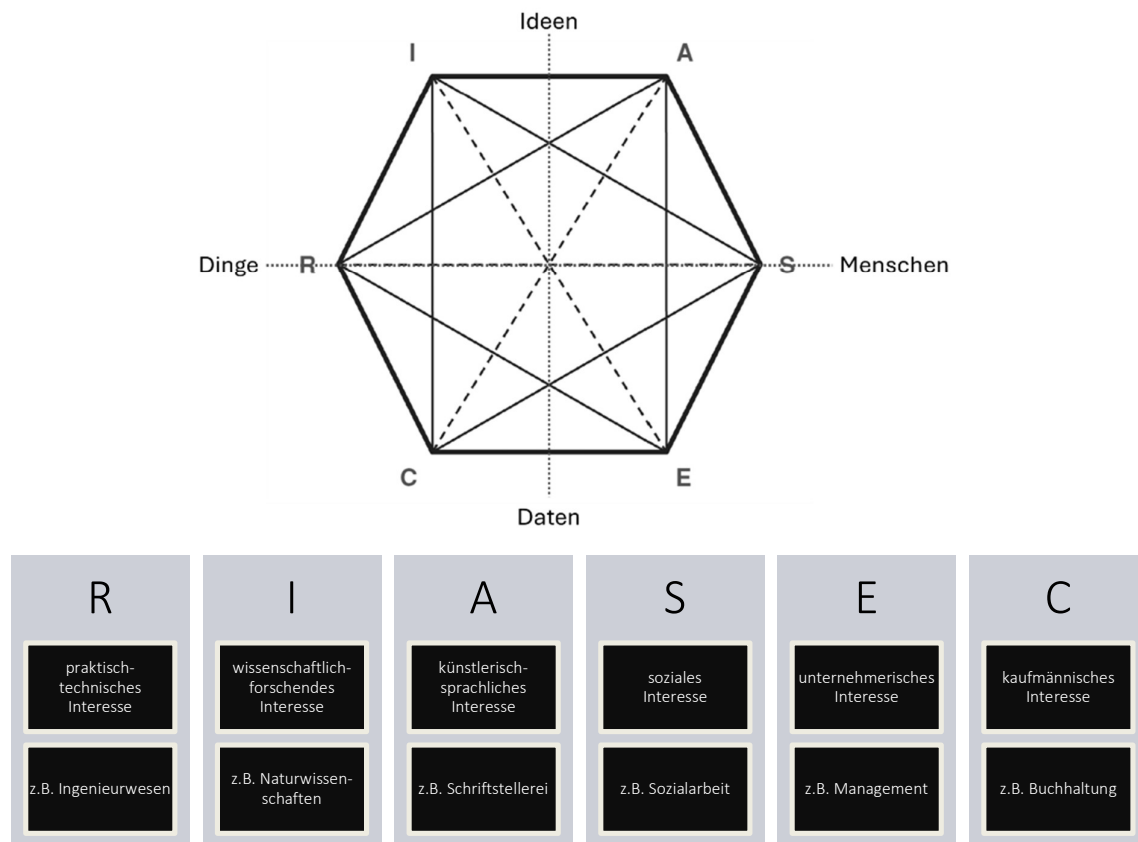


Abbildung 5. Berufsinteressenmodell nach Holland Aus Eigenschaftsorientierte Verfahren in der Personalauswahl, 2014 (S. 188). In H. Schuler & U. Kanning (Hrsg.), *Lehrbuch der Personalpsychologie* (S. 149 – 213). Göttingen: Hogrefe.

Das Modell nach Holland wurde empirisch bestätigt (Neyer & Asendorpf, 2018) und bildet das theoretische Fundament für das Testverfahren *EXPLORIX*. Dieses wird insb. bei Berufs- und Laufbahnberatungen eingesetzt, wo Interessensabfragen üblich sind. In der Personalauswahl werden Interessensabfragen kaum berücksichtigt (Schuler et al., 2014), obwohl diese über eine moderate bis hohe prognostische Kriteriumsvalidität für berufliche Leistung besitzen, wie dies auch noch im Kap. 2.2.1.3.3 dargelegt wird.

## 2.2 Qualität – Gütekriterien

Um Qualität beurteilen zu können, werden Kriterien benötigt, an welchen ein Abgleich und eine daran anschließende Beurteilung stattfinden kann. Im Bereich psychologischer Tests zieht man dazu die sogenannten Gütekriterien herbei.

Um von einem psychologischen Test sprechen zu können, muss ein solcher die Hauptgütekriterien *Objektivität*, *Reliabilität* und *Validität* erfüllen. Zudem gibt es auch Nebengütekriterien, die sich insb. mit dem Umstand der Brauchbarkeit eines Tests beschäftigen (Stemmler & Margraf-Stiksrud, 2015). Hauptgütekriterien geben also Auskunft über Genauigkeit und Prognosekraft eines Tests (Bühner, 2011). Nebengütekriterien besitzen einen dazu ergänzenden Charakter und sind nicht für alle Tests gleichermaßen bedeutsam (Krohne & Hock, 2015).

Die Abb. 6 bietet einen Überblick über die Struktur der Gütekriterien. Der Abb. 6 anschließend werden die darin genannten Kriterien definiert.

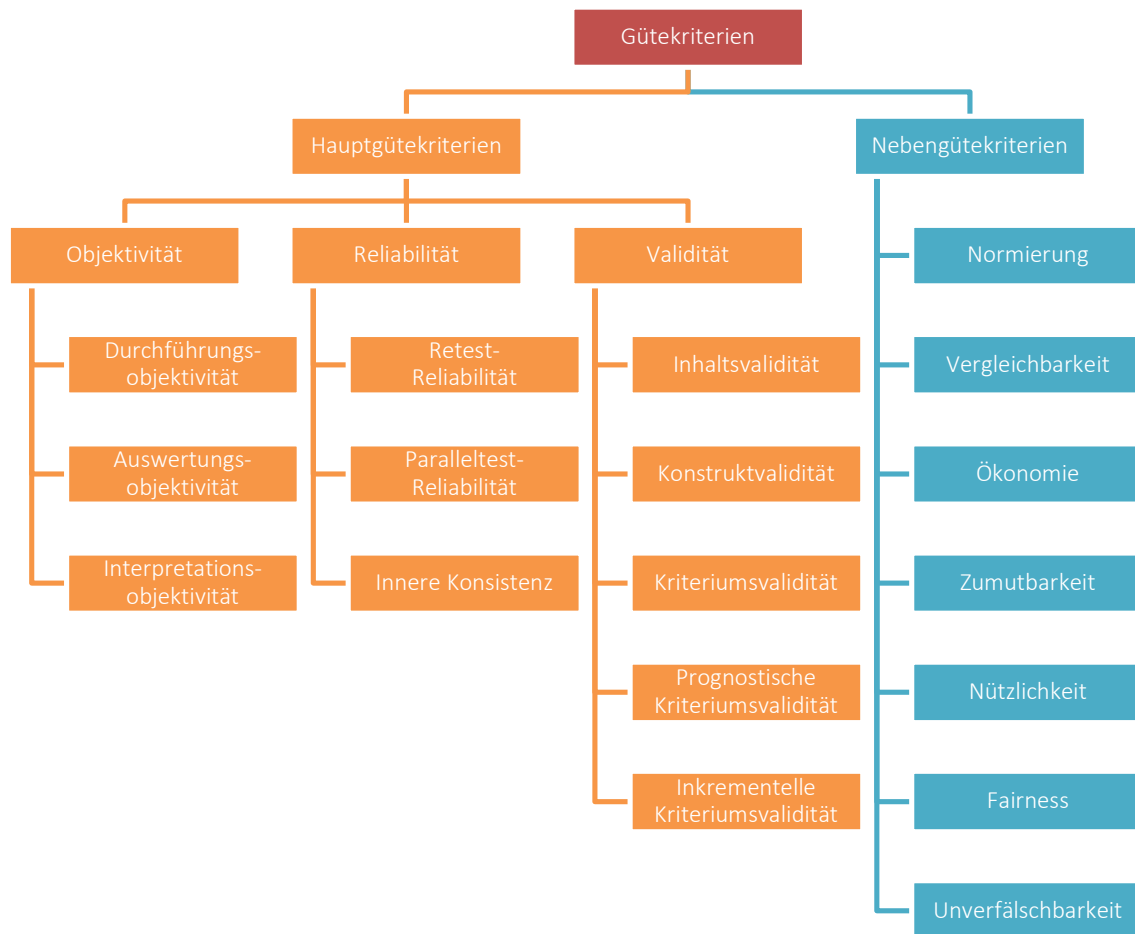


Abbildung 6. Gütekriterien. Aus Einführung in die Test- und Fragebogenkonstruktion von M. Bühner, 2011 (S. 74). Hallbergmoos: Pearson.

## 2.2.1 Hauptgütekriterien

Die Beziehung der drei Hauptgütekriterien ist von besonderer Art. So ist Objektivität zwingendes Kriterium zur Erlangung von Reliabilität, aber genügt allein nicht, um Reliabilität zu erreichen. Ebenso verhält sich die Reliabilität ggü. der Validität: Reliabilität ist zwingendes Kriterium, um Validität zu erreichen, genügt allein aber nicht, um Validität zu erreichen (Bühner, 2011). Diese drei Begriffe werden in den drei folgenden Unterkapiteln erläutert.

### 2.2.1.1 Objektivität

Objektivität bezeichnet das Ausmass, in welchem Testergebnisse unabhängig von der anwendenden Person sind (Bühner, 2011; Kanning, 2015). Das bedeutet, dass man von Objektivität sprechen kann, wenn der testleitenden Person keinerlei Handlungsspielraum in den Prozessschritten Durchführung, Auswertung oder Interpretation eingeräumt wird. Das verlangt ein hohes Mass an standardisierten Vorgaben, die typischerweise im Testmanual dokumentiert sind (Moosbrugger & Kelava, 2020). Entlang der genannten Prozessschritte unterscheidet man zwischen Durchführungs-, Auswertungs- und Interpretationsobjektivität. Welche Anforderungen diese Objektivitätsarten stellen und wie diese gewährleistet werden können, wird nachfolgend beschrieben:

#### Durchführungsobjektivität

Die Durchführung von Tests muss bei jedem Durchgang gleichbleiben, auch bei Wechsel von Testleitenden. Die Standardisierung der Durchführung wird typischerweise durch schriftliche Instruktionen zur Bearbeitung des Tests und durch eine fixierte Zeitvorgabe gewährleistet (Bühner, 2011).

#### Auswertungsobjektivität

Bei der Auswertung desselben Tests muss jeder und jede Testleitende dasselbe Resultat ermitteln. Die Auswertungsobjektivität wird durch eindeutige Auswertungsanweisungen unterstützt, welche üblicherweise im Testmanual enthalten sind (Moosbrugger & Kelava, 2020). Oftmals werden im Testmanual auch Hilfen wie Auswertungsschablonen oder Links zu Auswertungssoftware bereitgestellt, die die Auswertungsobjektivität zusätzlich unterstützen (Bühner, 2011).

#### Interpretationsobjektivität

Bei der Interpretation desselben Testresultats müssen verschiedene Testleitende zur selben Beurteilung gelangen. Interpretationsobjektivität verlangt eindeutige Regeln für die Testwertinterpretation (Bühner, 2011). Idealerweise bestehen hierzu ein Normsystem (vgl. Kap. 2.1.2.1) oder Grenzwerte, anhand deren die Testergebnisse eingeordnet und entsprechend interpretiert werden können (Moosbrugger & Kelava, 2020). Damit spricht die Interpretationsobjektivität auch das Nebengütekriterium *Normierung* an, welches später ebenfalls besprochen wird (siehe Kap. 2.2.2).

Insbesondere für den klinischen Bereich und die gesetzlich vorgeschriebene Berufseignungsdiagnostik (wie sie bspw. in Sicherheitsbereichen zu finden ist) hat sich die kriteriumsorientierte Testwertinterpretation durchgesetzt. Hier müssen gewisse Grenzwerte erfüllt sein, um vom eindeutigen Vorhandensein bzw. von eindeutiger Abwesenheit einer bestimmten Merkmalsausprägung sprechen zu können (Moosbrugger & Kelava, 2020).

### 2.2.1.2 Reliabilität

Reliabilität entspricht dem Ausmass an Messgenauigkeit bzw. an Messzuverlässigkeit, mit dem ein Test ein bestimmtes Persönlichkeits- oder Leistungsmerkmal einer Person misst. Eine einhundertprozentige Reliabilität entspräche der gänzlichen Abwesenheit von Messfehlern (Bühner, 2011). In der Praxis muss aber von Messfehlern ausgegangen werden.

Um die Reliabilität eines Tests schätzen zu können, dienen unterschiedliche korrelative Ansätze, wie die Retest-Reliabilität, Paralleltestreliabilität oder die Berechnung der internen Konsistenz (Bühner, 2011), die nachfolgend kurz erklärt werden und für die Methodik (siehe Kap. 3) besondere Relevanz besitzen.

#### Retest- und Paralleltestverfahren

Beim Retest-Verfahren bittet man dieselbe Testperson, denselben psychologischen Test zweimal, mit einem gewissen zeitlichen Abstand, ausfüllen. Die Testergebnisse der beiden Testdurchläufe werden anschliessend korreliert, um die Reliabilität des Tests feststellen zu können (Bühner, 2011).

Das Paralleltestverfahren macht sich ein ähnliches Prozedere zunutze. Bei der Berechnung der Paralleltestreliabilität zieht man zum überprüfenden psychologischen Test einen anderen psychologischen Test hinzu, der dieselbe Fähigkeit oder Eigenschaft zu erfassen beansprucht. Dieselbe Testperson füllt die beiden Tests aus, wobei die Ergebnisse der beiden Tests korreliert werden, um die Reliabilität feststellen zu können (Bühner, 2011). Moosbrugger und Kelava (2020) nennen das Paralleltestverfahren als Königsweg der Reliabilitätsbestimmungen, denn das Retest-Verfahren birgt vergleichsweise einige Tücken. Schliesslich bleibt das Testsetting beim Paralleltestverfahren dasselbe, was die Durchführungsobjektivität hochhält. Zudem können reliabilitätsmindernde Einflüsse, wie Erinnerungs- oder Übungseffekte durch das Paralleltestverfahren verhindert werden.

#### Interne Konsistenz

Die interne Konsistenz, zu deren Berechnung zumeist der Korrelationskoeffizient *Cronbachs Alpha* dient, prüft den korrelativen Zusammenhang zwischen den einzelnen Fragebogenitems, die demselben Faktor zugeordnet wurden (Bühner, 2011; Moosbrugger & Kelava, 2020). Es besteht der Anspruch, dass die Faktoren durch ihre enthaltenen Items das Merkmal konsistent und damit widerspruchsfrei zu erfassen vermögen.

Es gilt zu beachten, dass aus reliablen Tests trotz aller Vorkehrungen unreliable Ergebnisse resultieren können. Gerade Selbstberichtsverfahren wie Persönlichkeitstests, bieten Potential für bewusstes Verfälschen (Faking). Auch unbewusstes Verfälschen aufgrund unreflektierter Selbsteinschätzungen, Verständnisschwierigkeiten oder Übungseffekten mindert die Zuverlässigkeit von Testergebnissen (Bühner, 2011).

### 2.2.1.3 Validität

Validität entspricht dem Ausmass, mit dem von einem Testergebnis auf die tatsächliche Ausprägung des getesteten Persönlichkeits- oder Leistungsmerkmals geschlossen werden kann. Bei Validität wird zwischen Inhalts-, Konstrukt- und Kriteriumsvalidität differenziert. Die genannten Begriffe werden in den nun folgenden Unterkapiteln behandelt.

### 2.2.1.3.1 Inhaltsvalidität

Inhaltsvalidität besteht dann, wenn ein Test bzw. dessen Items ein Merkmal oder eine Fähigkeit repräsentativ und hinreichend erfassen (Bühner, 2011). Der Test soll also das messen, was er zu messen beansprucht – nicht mehr und nicht weniger.

Um Inhaltsvalidität zu beurteilen, müssen Expertinnen- und Expertenurteile eingeholt werden. Sie müssen fachliche Überlegungen anstellen, um die Repräsentativität und Vollständigkeit der im Test berücksichtigten Items bewerten zu können (Bühner, 2011).

### 2.2.1.3.2 Konstruktvalidität

Konstruktvalidität entspricht der wissenschaftlichen Fundierung desjenigen Merkmals, das ein Test zu messen beansprucht. Ist demnach der Rückschluss eines Testergebnisses auf ein psychologisches Konstrukt wie bspw. Leistungsfähigkeit oder eine Persönlichkeitseigenschaft wissenschaftlich fundiert, kann man von einem konstruktvaliden Test sprechen (Moosbrugger & Kelava, 2020). Die Überprüfung der Konstruktvalidität erfolgt, ähnlich dem Paralleltest-Verfahren zur Reliabilität, durch korrelative Vergleiche mit anderen Erhebungsverfahren, die dieselbe Eigenschaft zu messen beanspruchen (Bühner, 2011).

### 2.2.1.3.3 Prognostische Kriteriumsvalidität

Die Beschäftigung mit der prognostischen Kriteriumsvalidität<sup>5</sup> psychologischer Tests ist von besonderer Wichtigkeit für die vorliegende Thesis. Denn aufgrund bestehender Kriteriumsvalidität lässt sich der Einsatz psychologischer Tests Rahmen von Personalauswahl begründen.

Kriteriumsvalidität entspricht dem Zusammenhang von Testleistungen mit äusseren Kriterien. In der vorliegenden Masterthesis entspricht die *berufliche Leistung* diesem äusseren Kriterium. Durch Testergebnisse sollen also Prognosen über die spätere berufliche Leistung vorgenommen werden können. Dieser Umstand erklärt, weshalb psychologische Tests zur Personalauswahl eingesetzt werden. Während bei Fähigkeitstest plakativ gesagt gilt: je besser desto besser, gibt es bei Fragebogen keine formalen Kriterien für richtige oder falsche Antworten bzw. Ausprägungen (siehe auch Kap. 2.1). Doch die Ergebnisse jahrzehntelanger Forschungen zeigen, dass eine hohe Ausprägung auf gewissen Persönlichkeitseigenschaften der Big Five für die berufliche Leistung von Vorteil sind. Die nachfolgende Tab. 2 entspricht einer Sammlung verschiedener Metaanalysen und entsprechender Updates, die sich mit dem Zusammenhang von Testresultaten zur Arbeitsleistung auseinandersetzen. An die Tab. 2 anschliessend werden die darin enthaltenen Korrelationen besprochen.

---

<sup>5</sup> Weitere Unterarten von Kriteriumsvalidität, wie etwa die retrospektive Kriteriumsvalidität oder die Übereinstimmungsvalidität sind für diese Masterthesis nicht von Relevanz, weshalb darauf nicht eingegangen wird.

Tabelle 2

Zusammenstellung von Metaanalysen, die sich mit der prognostischen Kriteriumsvalidität von Persönlichkeits-, Interessens- und Leistungstests auf berufliche Leistung beschäftigen.

	Hunter & Hunter (1984)	Schmidt & Hunter (1998)	Update nach Schuler, Höft & Hell (2014)	Salgado & Táuriz (2014)	Update nach Schmidt, Oh & Shaffer (2016)	Lee, Berry, Gonzalez-Mulé (2019)	Sackett, Zhang, Berry & Lievens (2022)	Sackett, Zhang, Berry & Lievens (2023)
Allgemeine Intelligenztests	.27 - .61 <sup>6</sup>	.51	.62 (Salgado et al. 2003)		.65 (Schmidt, Shaffer & Oh, 2008)		.31	.23 <sup>7</sup>
Tests zu emotionaler Intelligenz <sup>8</sup>					.23 (Joseph, Jin, Newman & O'Boyle, 2015)		.30	
Interessenstests	.10	.10	.23 (Van Iddekinge, Roth, Putka & Lanivich, 2011)		.31 <sup>9</sup> (Van Iddekinge, Roth, Putka & Lanivich, 2011)		.24	
Integritätstests		.41	.18 <sup>10</sup> (Van Iddekinge, Roth, Raymark & Odle-Dusseau, 2012)		.46 (Ones, Viswesvaran & Schmidt, 1993)		.31	
Gewissenhaftigkeitstests		.31	.31 (Barrick, Mount & Judge, 2011)	.24	.22 (Schmidt, Shaffer & Oh, 2008)		.21	
Tests zu Offenheit für Erfahrungen				.14	.04 (Schmidt, Shaffer & Oh, 2008)		.06	
Tests zu Extraversion				.10	.09 (Schmidt, Shaffer & Oh, 2008)		.11	
Verträglichkeitstests				.06	.08 (Schmidt, Shaffer & Oh, 2008)		.10	
Emotionale Stabilität				.11	.12 (Schmidt, Shaffer & Oh, 2008)		.09	
Ehrlichkeit-Bescheidenheit						.15		

Die Kriteriumsvaliditäten von *Offenheit*, *Extraversion*, *Verträglichkeit*, *emotionaler Stabilität* und *Ehrlichkeit-Bescheidenheit* sind als eher klein einzuordnen. Zum einen ist denkbar, dass diese Eigenschaften für den beruflichen Erfolg nicht von hoher Wichtigkeit sind. Zum anderen bietet die Metaanalyse von Birkeland, Manson, Kisamore, Brannick und Smith (2006) eine Erklärung. Sie konnten zeigen, dass willentliche Verfälschung (Faking) in Persönlichkeitstests betrieben wird, wenn diese der Personalauswahl dienen. Insbesondere bei den Big Five-Konstrukten *Emotionale Stabilität* und *Gewissenhaftigkeit* betreiben Bewerbende *Faking-Good*, verfälschen also ihre Antworten, um sich besser darzustellen (Birkeland et al., 2006). Wie bereits in Kap. 2.2.2 angeschnitten, bedeutet das also, dass die Reliabilität und schliesslich die Kriteriumsvalidität entsprechender Persönlichkeitstests durch die Testteilnehmenden negativ beeinflusst wird.

Persönlichkeitstests zu *Interessen*, *Integrität*, *Motivation* und *Gewissenhaftigkeit* zeigen sich als moderate Prädiktoren für Berufserfolg. Gerade Integrität und Gewissenhaftigkeit sind als Eigenschaften konstruktiven Arbeitsverhaltens zu verstehen und besitzen daher Bedeutung für beruflichen Erfolg.

<sup>6</sup> Hunter & Hunter (1984) haben in ihrer Metaanalyse zwischen verschiedenen Berufsgruppen unterschieden, wobei keine Berechnung einer durchschnittlichen Kriteriumsvalidität vorgenommen wurde. Die tiefste Validität ( $\rho = .27$ ) fand sich für Verkaufsmitarbeitende, die höchste Validität ( $\rho = .61$ ) für Vertriebsmitarbeitende.

<sup>7</sup> Sackett et al. (2023) haben Überkorrekturen bei der Berechnung der Korrelationen festgestellt, welche sie 2022 veröffentlicht hatten. Die hier berichtete Korrelation wurde entsprechend korrigiert.

<sup>8</sup> Emotionale Intelligenz wird in der vorliegenden Masterthesis als Fähigkeit und nicht als Persönlichkeit verstanden. Damit folgt diese Thesis den Ansichten von Mayer, Salovey, Caruso und Sitarenios (2001).

<sup>9</sup> Der durch das Update nach Schuler, Höft und Hell (2014) berichtete Wert ist niedriger als derjenige von Schmidt, Oh und Shaffer (2016), obwohl sie sich auf dieselbe Quelle beziehen. Schmidt, Oh und Shaffer (2016) berichten einen durch sie korrigierten Wert, der dadurch zuverlässiger sei. Schuler, Höft und Hell (2014) verwenden den in der Studie von Van Iddekinge, Roth, Putka und Lanivich (2011) berichteten, originalen Wert.

<sup>10</sup> Ones et al. (2012) kritisierten in ihrem Kommentar die Meta-Analyse von Van Iddekinge et al. (2012) aufgrund methodischer Mängel und dem Miteinbezug irrelevanter Skalen. Der niedrige Korrelationskoeffizient von  $\rho = .18$  sei mit Vorsicht und eher als Ausreisser zu betrachten.

Neben der Perspektive, Stärken des menschlichen Verhaltens hervorzuheben, bietet sich auch die Einnahme einer gegenteiligen Perspektive an: die Metaanalyse von O'Boyle, Forsyth, Banks, Story & White (2015) zeigt, dass niedrige Gewissenhaftigkeits- und Verträglichkeitswerte stark korrelieren mit *Machiavellismus* und *Psychopathie* – zwei der drei Faktoren der *dunklen Triade*. Eine weitere Metaanalyse von O'Boyle, Forsyth, Banks & McDaniel (2012) konnte aufzeigen, dass *Machiavellismus* und *Psychopathie* nicht nur negativ mit beruflicher Leistung korrelieren, sondern auch erwähnenswerte positive Korrelationen zu destruktivem Arbeitsverhalten, wie etwa Diebstahl, Sabotage, ungerechtfertigtem Absentismus und böartigem Verhalten zu Mitarbeitenden aufweisen. Die Einnahme dieser Perspektive untermauert ebenfalls die Relevanz von Gewissenhaftigkeits- und Verträglichkeitstests und bestätigt gewissermassen auch die in Tab. 2 dargelegten Kriteriumsvaliditäten für die beiden genannten Konstrukte.

Fähigkeitstests zu allgemeiner und emotionaler Intelligenz besitzen laut Tab. 2 eine moderate bis hohe Prognosekraft für berufliche Leistung.

Besonders auffällig sind die schwankenden Kriteriumsvaliditäten zur allgemeinen Intelligenz und auch zu den Interessenstests. Grundsätzlich gilt, dass über verschiedene Berufe und deren Anforderungen verallgemeinernde Metaanalysen berechnet werden dürfen. Dabei werden die Kriteriumsvaliditäten von Einzelstudien um die Kriteriums-, manchmal auch um die Prädiktorunreliabilität und mögliche Varianzeinschränkungen korrigiert und anhand der jeweiligen Stichprobengrösse gewichtet (Schuler, 2004). Die Vermutung liegt nahe, dass nicht bei allen in der Tab. 2 aufgeführten Metaanalysen die genannten Korrekturrechnungen in derselben Weise vorgenommen wurden, was möglicherweise zu den Schwankungen geführt hat.

Gerade bei der allgemeinen Intelligenz, die laut Tab. 2 über die Jahrzehnte mehr als die Hälfte ihrer Prognosekraft eingebüsst haben soll, ist die Erklärung betreffend die Korrekturrechnungen wohl kaum ausreichend. Ryan und Ployhart (2014) verweisen darauf, dass in den letzten Jahrzehnten kaum mehr Forschung zum Zusammenhang von allgemeiner Intelligenz und entsprechenden Tests betrieben wurde, was wohl zu einer löchrigen Basis für neuere Metaanalysen geführt haben könnte.

Abschliessend klären lässt sich dies aber nicht. Dennoch wird in der Forschung die Prädiktorkraft von Intelligenz auf berufliche Leistung kaum bestritten und gilt nach wie vor als der beste Prädiktor für berufliche Leistung (Ryan & Ployhart, 2014).

#### **2.2.1.3.4 Inkrementelle Validität**

Auch unter Bezug der inkrementellen Validität lässt sich der Einsatz psychologischer Tests bzw. die Berücksichtigung verschiedener Testverfahren im Personalauswahlprozess begründen. Inkrementelle Validität bedeutet, dass die Hinzunahme weiterer Personalauswahlinstrumente weitere Varianz eines Kriteriums aufklärt. In diesem Falle entspricht die berufliche Leistung diesem Kriterium. Das bedeutet, dass die berufliche Prognose durch Berücksichtigung weiterer psychologischer Tests genauer ausfällt, als wenn man darauf verzichten würde (Bühner, 2011; Kanning, 2015).

In der Forschung um Personalauswahlverfahren hat sich etabliert, die inkrementelle Validität mit demjenigen Auswahlverfahren zu berechnen, welches über die höchste Prognosekraft verfügt. Demnach wird jede Testmethode daran gemessen, ob sie zusätzlich zum allgemeinen Intelligenztest noch weitere Kriteriumsvalidität zur Prognose beruflicher Leistung beitragen kann (Bühner, 2011). In der nachfolgenden Tab. 3 sind inkrementelle Validitäten von Leitungsfähigkeits- und Persönlichkeitstests aus gefundenen Metaanalysen bzw. Updates zusammengetragen. Der Tab. 3 folgt eine kurze Besprechung der darin enthaltenen Werte.

Tabelle 3

Zusammenstellung von Metaanalysen, die sich mit der inkrementellen Validität von Persönlichkeits- und Interessenstests zu Intelligenztests auf berufliche Leistung beschäftigen.

	Schmidt & Hunter (1998)	Update nach Schmidt, Oh & Shaffer (2016)
	Inkrementelle Validität	Inkrementelle Validität
<b>Interessenstests</b>	.01	.062 (Van Iddekinge, Roth, Putka & Lanivich, 2011)
<b>Integritätstests</b>	.14	.130 (Ones, Viswesvaran & Schmidt, 1993)
<b>Gewissenhaftigkeitsstests</b>	.09	.053 (Schmidt, Shaffer & Oh, 2008)
<b>Emotionale Intelligenz</b>		.004 (Joseph, Jr., Newman & O'Boyle, 2015)
<b>Offenheit</b>		.039 (Schmidt, Shaffer & Oh, 2008)
<b>Extraversion</b>		.006 (Schmidt, Shaffer & Oh, 2008)
<b>Verträglichkeit</b>		.002 (Schmidt, Shaffer & Oh, 2008)
<b>Emotionale Stabilität</b>		.000 (Schmidt, Shaffer & Oh, 2008)

Wie sich an den Werten der Tab. 3 erkennen lässt, besitzen insbesondere Integritätstests inkrementellen Charakter, wenn man diese neben Intelligenztests ebenfalls zur Personalauswahl einsetzt. Die niedrigen inkrementellen Validitäten, die u.a. emotionale Stabilität oder Extraversion aufweisen, lassen sich durch deren niedrige Kriteriumsvalidität (vgl. Tab. 2), gepaart mit der hohen Kriteriumsvalidität der allgemeinen Intelligenz erklären. Inkrementelle Effekte gehen dadurch schliesslich verloren.

## 2.2.2 Nebengütekriterien

Die Nebengütekriterien dienen ebenfalls der Beurteilung von psychologischen Tests. Sie sind insb. für die Testdurchführung von Bedeutung (Bühner, 2011), beschäftigen sie sich doch mit der praktischen Einsatzfähigkeit.

### Normierung

Die Normierung bezieht sich auf die Festlegung sinnvoller Normwerte durch den Vergleich mit Werten, die aus der Testung einer repräsentativen Stichprobe (auch Eichstichprobe genannt) erhalten wurden. Dadurch kann man den Wert einer Testperson mit der Eichstichprobe vergleichen, um die Position der Testperson im Verhältnis zu anderen festlegen und entsprechend interpretieren zu können (Myers, 2014; Bühner, 2011; Moosbrugger & Höfling, 2007). Ein Beispiel, das die Normierung veranschaulicht, ist die Normierung der allgemeinen Intelligenz durch den Intelligenzquotienten (vgl. Kap. 2.1.2.1).

### Ökonomie

Ein Test ist ökonomisch, wenn er wenig Bearbeitungszeit verlangt, wenig materielle Ressourcen verbraucht, Gruppentestungen möglich sind und die Auswertung bequem und schnell von statten geht (Bühner, 2011). Damit spricht die Ökonomie also auch die Beanspruchung der Testleitenden bzw. die praktische Anwendbarkeit eines Tests an. Ein ökonomischer Test muss demnach einfach anwendbar sein. Insbesondere Tests, die am Computer bearbeitet werden können, erfüllen das Kriterium der Ökonomie leichter (Moosbrugger & Kelava, 2020).

In der Praxis ist die Beurteilung der Ökonomie lediglich durch den Vergleich mit anderen Auswahlverfahren oder Tests möglich (Bühner, 2011; Moosbrugger & Kelava, 2020).

### **Nützlichkeit**

Ein Test gilt als nützlich, wenn er ein Merkmal operationalisiert, für dessen Untersuchung ein praktisches Bedürfnis besteht. Ein entwickelter Test sollte demnach Vorteile ggü. anderen Tests zeigen, die dasselbe Merkmal erfassen. Solche Vorteile können bspw. in der besseren Erfüllung der Hauptgütekriterien bestehen (z.B. Verbesserung der Validität) oder in der besseren Erfüllung der Nebengütekriterien, wie der Verringerung der Anzahl der Testitems durch die Erstellung einer Kurzversion (Bühner, 2011).

### **Zumutbarkeit**

Ein Test gilt als zumutbar, wenn der oder die Getestete durch den Test nicht unnötig belastet wird. Es besteht die Anforderung, dass der Test geringe zeitliche, psychische und physische Belastung für die Testperson mit sich bringt (Bühner, 2011). Dabei gilt es sorgfältig abzuwägen, ob die Beanspruchung der Testperson in einem angemessenen Verhältnis zum Nutzen des Tests steht, hinsichtlich dessen, was er über die Person aussagen kann (Moosbrugger & Kelava, 2020). Bspw. wird in sicherheitsrelevanten Branchen wie dem Personentransport das Durchführen zahlreicher Tests als zumutbar angesehen. Der Nutzen entspricht hier der Sicherheit durch geeignetes Personal.

### **Fairness**

Testfairness liegt vor, wenn Tests bzw. Testresultate keine systematische Benachteiligung bestimmter Gruppen oder Personen verursachen (Hartig, Frey & Jude, 2008). Ein Test darf keine Teilpopulationen (z.B. Geschlecht, ethnische Herkunft, Generation) systematisch benachteiligen. Personen aus unterschiedlichen Gruppen mit derselben Fähigkeitsausprägung müssen dieselbe Lösungswahrscheinlichkeit bei Fragebogenitems aufweisen (Bühner, 2011). Bspw. muss eine Mitarbeitendenumfrage für alle Stufen der Belegschaft gut verständlich sein, um Verzerrungen zu vermeiden.

Weiter lässt sich die Testfairness in die Testdurchführungsfairness differenzieren, welche die systematische Benachteiligung bestimmter Gruppen durch die Testdurchführungsmodalitäten thematisiert. Soll bspw. ein Test online ausgefüllt werden, müssen die Computerkenntnisse der jüngeren Teilnehmenden im Vergleich zu den älteren Teilnehmenden berücksichtigt werden. Bestehen generationsabhängige Unterschiede hinsichtlich der Kenntnisse, müssen die Durchführungsmodalitäten zugunsten der Fairness für die älteren Teilnehmenden entsprechend angepasst werden (Moosbrugger & Kelava, 2020).

### **Unverfälschbarkeit**

Die Unverfälschbarkeit ist insb. für Persönlichkeitstests relevant. Unverfälschbarkeit besteht, wenn eine Testperson nicht die Möglichkeit hat, dass sie durch gezieltes Antwortverhalten die Ausprägungen der Testwerte zu ihren Gunsten steuern oder verzerren kann (Moosbrugger & Kelava, 2020). Bühner (2011) nennt den Anspruch an Unverfälschbarkeit eines Tests unrealistisch, da für Testpersonen immer die Möglichkeit besteht, ihre Testergebnisse zu beeinflussen. Gerade im Bereich der Personalauswahl ist sozial erwünschtes Antwortverhalten oder auch Faking zu erwarten. Betreffend die Unverfälschbarkeit sollten im Testmanual Angaben enthalten sein, wie sehr ein Test verfälscht werden kann und ob in der Testauswertung Möglichkeiten bestehen, Verfälschungen aufzudecken (Bühner, 2011). Demnach bleibt lediglich die Möglichkeit zur Kontrolle von verfälschendem Antwortverhalten, wobei Bogner & Landrock (2015) den Einbezug einer Kurzskala für soziale Erwünschtheit (KSE-G) vorschlagen. Diese besteht aus sechs Items, bietet also eine gewisse Testökonomie und vermag verfälschendes Antwortverhalten aufzudecken.

## 2.2.3 Qualitätssicherung

Im ersten Kapitel dieser Thesis wurde die Problematik um den Einsatz qualitativ mangelhafter Tests kurz dargelegt. Qualitätssicherung im Rahmen von Testentwicklung und -einsatz ist daher unumgänglich. Um Qualitätssicherung betreiben zu können, müssen Qualitätsstandards herangezogen werden, anhand derer psychologische Tests auf ihre Qualität hin beurteilt werden können. Die zuvor besprochenen Gütekriterien sind als solche Qualitätsstandards zu verstehen. Im deutschsprachigen Raum bildet die Deutsche Industrienorm *DIN33430*, welche sich an den internationalen Standards zu Qualitätsmanagement [ISO 9000ff] orientiert, die Qualitätsstandards für Verfahren beruflicher Eignungsdiagnostik ab (Kersting, 2014).

Das Diagnostik- und Testkuratorium (DTK) sieht seine Aufgabe darin, die Öffentlichkeit über die Güte von psychologischen Tests zu informieren (DTK, 2018) und sie vor unzureichenden diagnostischen Verfahren und unsachgemäßer Anwendung zu schützen (Schuler, 2000). Dazu beurteilt es Testverfahren, wobei es sich an der *DIN33430* orientiert. Der Beurteilungsprozess des Testkuratoriums verläuft in zwei Schritten:

### 1. Prüfung der Verfahrenshinweise im Testmanual

Verfahrenshinweise entsprechen der Diskussion der Gütekriterien durch die Testentwickelnden selbst. Testentwicklerinnen und -entwickler bestimmen bspw. die Durchführungsmodalitäten für ihren Test und besprechen diese neben ersten Erkenntnissen hinsichtlich Reliabilität und Validität in ihren Verfahrenshinweisen. Die Rezensenten und Rezensentinnen des DTK prüfen dann, ob die genannten Verfahrenshinweise vorhanden sind und den vorgegebenen Anforderungen der *Checkliste 1* des *DIN Screen* entsprechen. Gerade Testverfahren mit hohem Qualitätsanspruch legen die notwendigen Informationen, die durch die Checkliste 1 gefordert werden, bereits in ihrem Testmanual entsprechend dar (Kersting 2014, in Schuler & Kanning, 2014).

### 2. Beurteilung durch Fachpersonen

Nur wenn der eben genannte erste Schritt positiv verlaufen ist, gilt ein Test als prüffähig. Nur prüffähige Tests werden rezensiert, wobei mindestens zwei Fachpersonen entlang des Leitfadens «Besprechungs- und Beurteilungskategorien des neuen TK-Systems» einen Test beurteilen (Kersting, 2006; Kersting, 2014, in Schuler & Kanning, 2014). Die Tab. 4 zeigt die Bewertungskategorien und -möglichkeiten des Leitfadens.

Tabelle 4  
*Testbeurteilungskriterien nach TBS-DTK. Aus Zur Beurteilung der Qualität von Tests: Resümee und Neubeginn von M. Kersting, 2006 (S. 251). Psychologische Rundschau, 57, S. 243 – 253.*

Bewertungskategorien	Bewertung der jeweiligen Kategorie	
	Vierstufige Likertskala, um Erfüllung des Kriteriums einzuschätzen	Freitext-Bewertung mit maximal 1'000 Zeichen
Allgemeine Informationen über den Test (Beschreibung und diagnostische Zielsetzung des Tests)	Ja	Ja
Theoretische Grundlagen als Basis der Testkonstruktion (Konstruktvalidität)	Nein	Ja
Objektivität	Ja	Ja
Normierung	Nein	Ja
Reliabilität	Ja	Ja
Validität (Fairness soll hier berücksichtigt werden)	Ja	Ja
Weitere Nebengütekriterien, wie z.B. Unverfälschbarkeit	Nein	Ja
Abschlussbewertung und Empfehlung	Nein	Ja, wobei maximal 2'000 erlaubt sind.

Wie die Tab. 4 erkennen lässt, zeugt das TBS-DTK von hoher Standardisierung. Dadurch lässt sich Vergleichbarkeit zwischen verschiedenen Testrezensionen erreichen, was der Problematik unsystematisch aufbereiteter Testbeurteilungen entgegenwirken soll (Kersting, 2006).

Zurzeit verfügt das TBS-DTK über 60 Rezensionen psychologischer Tests (Stand: 19.06.24). Etwa die Hälfte dieser rezensierten Tests lässt sich auch für die Personalauswahl bzw. -entwicklung einsetzen. Die andere Hälfte beschäftigt sich mit Rezensionen psychopathologischer oder entwicklungspsychologischer Diagnoseinstrumente.

In der Schweiz nimmt das *Schweizerische Dienstleistungszentrum Berufsbildung* (SDBB) Testrezensionen vor. Die *Fachgruppe Diagnostik* des SDBB hat den bundesweiten Leistungsauftrag, psychologische Tests zu bewerten, die in der Berufs-, Studien- oder Laufbahnberatung eingesetzt werden. Über 100 Testrezensionen sind in der Datenbank des SDBB gelistet (SDBB, 2023). Ein Grossteil dieser rezensierten Tests lässt sich auch für die Personalauswahl und -entwicklung einsetzen. Das SDBB bezieht sich bei der Erstellung von Testrezensionen nicht direkt auf die DIN33430 oder eine entsprechende Checkliste, sondern folgt nach eigenen Angaben den wichtigsten bzw. gängigen Gütekriterien (pers. Mitteilung, M.B., 21.05.2024). Das hat zur Folge, dass sich die durch das SDBB erstellten Rezensionen formal voneinander unterscheiden.

Qualitätssicherung wird neben Checklisten und Testrezensionen auch durch Personenlizenzierung nach DIN33430 vorgenommen. Das *DIN33430-Portal* vergibt Lizenzen an Personen, die berufsbezogene Eignungsbeurteilungen durchführen und sich nachweislich mit der DIN33430 auseinandergesetzt haben. Der Nachweis erfolgt durch eine Prüfung (Fellner, 2019; Kersting, 2008). Darüber hinaus setzen sich Vereine in Deutschland und der Schweiz für hochstehende Eignungsdiagnostik ein. Bspw. engagieren sich der Verein *SwissAssessment* und der Arbeitskreis *Assessment Center Deutschland* sowie auch das DIN33430-Portal für Sensibilisierung und Wissensvermittlung in Bezug auf eignungsdiagnostische Qualitätsstandards (Fellner, 2019).

## 2.3 Testeinsatz in der Praxis

Dieses Kapitel beschäftigt sich mit dem Einsatz psychologischer Tests in der Praxis. Aufgrund der begrenzten Studienlage in der Schweiz werden auch Studien aus Deutschland berücksichtigt. In einem ersten Teil beschäftigt sich dieses Kapitel mit der Einsatzhäufigkeit psychologischer Tests im Rahmen von Personalauswahl und -entwicklung. Anschliessend werden Studien besprochen, die sich mit Einflussgrössen beschäftigen, die den Testeinsatz in Unternehmen begünstigen oder verhindern.

### **Studienlage zur Einsatzhäufigkeit psychologischer Tests**

Die in Deutschland durchgeführte Längsschnittstudie von Arnoneit, Schuler und Hell (2020) zeigt, dass Persönlichkeitstests über die Jahrzehnte hinweg eine Einsatzhäufigkeit von etwa 20% aufweisen. Hingegen verzeichnen Leistungstests einen Abwärtstrend: während 1993 noch 40% der befragten Unternehmen angaben, Leistungstests im Rahmen von Personalauswahl einzusetzen, waren es 2017 noch aufgerundet 24%. In der Schweiz sind die Einsatzhäufigkeiten von Persönlichkeitstests höher und die Nutzung von Leistungstests tiefer. Hell, Vögeli und Hermann (2021) konnten in ihrer Studie feststellen, dass 32% aller Unternehmen in der Deutschschweiz im Rahmen von Personalauswahl Persönlichkeitstests einsetzen. Knapp 7% setzen Leistungstests ein. Véron, Mitrovic und Hell (2024) konnten diese Ergebnisse bestätigen: sie stellten ihrer Studie fest, dass knapp 30% der Deutschschweizer Unternehmen Persönlichkeitstests und 7% Leistungstests einsetzen.

Leistungstests finden laut Studie von Armoneit et al. (2020) insb. bei der Auswahl von Lernenden Anwendung. 17.9% der befragten Unternehmen gaben hier an, Leistungstests bei angehenden Lernenden einzusetzen. Persönlichkeitstests werden eher bei höheren Hierarchiestufen eingesetzt. Durchschnittlich 13% der befragten Unternehmen gaben an, dass sie bei Führungskräften Persönlichkeitstests einsetzen. Vergleichsweise finden Persönlichkeitstests bei Mitarbeitenden ohne Führungsfunktion nur in 7% der befragten Unternehmen statt (Armoneit et al., 2020).

Hossiep et al. (2015) stellten in ihrer Studie, welche unter den grössten deutschen Unternehmen ( $n = 120$ ) durchgeführt wurde, fest, dass Personalauswahl und -entwicklung die Haupteinsatzfelder für Persönlichkeitstests sind. 41% der befragten Unternehmen gaben an, die Tests zu Personalauswahl- und Personalentwicklungszwecken einzusetzen. 48% setzen Tests nur zu Entwicklungszwecken und 11% nur zu Auswahlzwecken ein.

Die Studie von Berchtold (2005) bietet einen Überblick über die in der Schweiz eingesetzten Personalauswahlverfahren in der externen Personalauswahl. Die Studie zeigt, dass insb. zweifelhafte Persönlichkeitstests wie etwa der *MBTI*, der *MPA* oder der *Insights* eingesetzt werden. Qualitative Persönlichkeitstests wie der *NEO-PIR* oder der *Occupational Personality Questionnaire (OPQ)* finden hingegen kaum Einsatz. Hingegen zeigt sich jedoch das *Bochumer Inventar zur berufsbezogenen Persönlichkeitseigenschaften (BIP)* mit moderater Verwendungshäufigkeit. Knapp 6% der befragten Unternehmen setzen den BIP ein (Berchtold, 2005), der über positive Rezensionen des SDBB und des DTK verfügt.

Die Studie von Hossiep et al. (2015) zeigt ein ähnliches Bild. 67% der Unternehmen gaben hier an, Persönlichkeitstest im Rahmen von Personalauswahl und / oder -entwicklung einzusetzen. Von denjenigen Unternehmen, die Tests einsetzen, nutzen 43% den MBTI, 40% den DISC und 35% den BIP. Hossiep et al. (2015) konnten zudem feststellen, dass der Insights, der 16PF oder der MPA kaum bis gar nicht in Deutschland eingesetzt werden. Das lässt eine regionale Abhängigkeit erahnen. Hossiep wiederholte zusammen mit Weiss, Netzer und Hossiep 2022 die Befragung unter den grössten Unternehmen ( $n = 115$ ). Dabei befragten sie neben Unternehmen in Deutschland auch Unternehmen aus Österreich und der Schweiz, wobei mehrheitlich deutsche Unternehmen teilgenommen hatten. 75% der befragten Unternehmen gaben an, Persönlichkeitstests im Rahmen von Personalauswahl und / oder -entwicklung einzusetzen. Hossiep et al. (2022) befürchten, dass die Anzahl an Testanwendenden in ihrer Studie überrepräsentiert ist. Von denjenigen Unternehmen, die Tests einsetzen, nutzen 58% den BIP. Im Vergleich zu 2015 hat die Nutzungshäufigkeit des BIP also stark zugenommen. 37% setzen den DISC ein, wobei kaum eine Veränderung zum Jahr 2015 festzustellen ist. 28% der befragten Unternehmen setzen den Insights ein, welcher im Jahr 2015 noch kaum genutzt wurde. 2015 wurde der MBTI mit 43% vergleichsweise am häufigsten eingesetzt. Die Daten von 2022 zeigen, dass sich die Nutzungshäufigkeit zwischenzeitlich mehr als halbiert hat, auf 17%.

Obwohl nur wenige Schweizer und Österreichische Unternehmen an der Befragung teilgenommen haben, konnten Hossiep et al. (2022) feststellen, dass der MPA in Deutschland nicht Fuss gefasst hat, in der Schweiz und in Österreich aber Verwendung findet.

### **Studienlage zu Einflussgrössen, die Testeinsatz fördern oder hindern**

Die Längsschnittstudie von Armoneit et al. (2020) untersuchte neben den Einsatzhäufigkeiten psychologischer Tests auch mögliche Einflussfaktoren auf den Testeinsatz, nämlich Akzeptanz, Praktikabilität und Validität. Der Faktor Akzeptanz meint die Einschätzung der HR-Mitarbeitenden hinsichtlich der durch die Bewerbenden aufgebrauchten Akzeptanz ggü. dem Testeinsatz (Armoneit et al. 2020). Praktikabilität kann definiert werden als Ausmass, in dem der Testeinsatz von den HR-Mitarbeitenden als geeignet empfunden wird, ihre Ziele zu erreichen (Latham & Finnegan, 1993). Validität entspricht der prädiktiven Kriteriumsvalidität betreffend die berufliche Leistung (Armoneit et al., 2020).

Über die Jahrzehnte hinweg konnten die Längsschnittstudie von Armoneit et al. (2020) wiederholt darlegen, dass die Einschätzung von Validität, Akzeptanz und Praktikabilität bei testanwendenden HR-Mitarbeitenden höher ausfällt als bei Nichtanwendenden. Zudem fanden Armoneit et al. (2020) heraus, dass, je höher HR-Mitarbeitende die Validität und Praktikabilität von online Persönlichkeitstests einschätzen, diese desto eher eingesetzt werden.

König et al. (2010) konnten ähnliche Ergebnisse hinsichtlich der Einschätzung der Validität finden, unternahmen aber keine Unterteilung in Persönlichkeits- und Leistungstests, bzw. online und Paper-Pencil-Format.

König et al. (2010) haben untersucht, welche Einflussfaktoren den Testeinsatz vorhersagen können. Ihre Studie führten sie in der Deutschschweiz unter 506 HR-Expertinnen und -Experten durch. Dabei fanden sie heraus, dass Verbreitung von Tests, Einschätzung der Akzeptanz und der Kriteriumsvalidität und die Kosten einflussreiche Prädiktoren für den Testeinsatz in einem Unternehmen sind: Je verbreiteter ein Test im Feld ist (also, je mehr ein Test auch durch andere Unternehmen eingesetzt wird), desto eher wird dieser auch im eigenen Unternehmen eingesetzt. Ein Test wird eher eingesetzt, je höher die HR-Mitarbeitenden die Bewerbendenakzeptanz und die prädiktive Kriteriumsvalidität eines Tests einschätzen. Einen umgekehrten Zusammenhang fanden König et al. (2010) betreffend die Kosten: je höher die Kosten für ein Testverfahren sind, desto unwahrscheinlicher ist deren Einsatz.

Hossiep et al. (2015) widmeten sich in ihrer Studie neben der Einsatzhäufigkeit von Persönlichkeitstests auch der Wichtigkeit verschiedener Verfahrensaspekte. So fanden sie heraus, dass bestimmte Gütekriterien und Aspekte der Anwendbarkeit von grosser bis sehr grosser Wichtigkeit sind für den Einsatz von Persönlichkeitstests. Hinsichtlich der Testgüte gaben befragte Unternehmen an, dass die Belastbarkeit der Ergebnisse, die wissenschaftliche Fundierung und die Transparenz der Tests wichtige Verfahrensaspekte seien. Betreffend die Anwendbarkeit nannten sie die Akzeptanz der Teilnehmenden, die Berücksichtigung der Privatsphäre, das Stellen von berufsnahen Fragen und ein dennoch breites Anwendungsspektrum als wichtig bis sehr wichtig.

Hossiep et al. (2015) konnten Gründe ausmachen, die gegen den Einsatz von Persönlichkeitstests im Rahmen von Personalauswahl und / oder -entwicklung sprechen. Insbesondere begründeten Unternehmen den Verzicht auf Tests dadurch, dass der zusätzliche Nutzen durch Testeinsatz zu gering sei, der berufliche Bezug fehle, andere Methoden geeigneter wären oder die Gefahr zu gross sei, durch Testeinsatz Bewerbende abzuschrecken. Ebenfalls wurden auch die Kosten als ausschlaggebend bezeichnet.

Auch Berchtold (2005) fand durch seine Studie heraus, dass Kosten, gemessen als finanzieller und auch zeitlicher Aufwand, ein wesentlicher negativer Einflussfaktor für den Testeinsatz ist.

Obwohl die Studien von Arnoneit et al. (2020), Hossiep et al. (2015), König et al. (2010) und auch Bechtold (2005) aufzeigen, dass die Validität als wichtiger Einflussfaktor für den Testeinsatz fungiert, überrascht es doch, dass eine Grosszahl eingesetzter psychologischer Tests qualitativ mangelhaft ist. Die Diskrepanz zwischen Anwendungshäufigkeit von Intelligenztests und der prognostischen Kriteriumsvalidität, lässt sich zum einen dadurch erklären, dass die Validität von Intelligenztests durch HR-Personen unterschätzt wird (Arnoneit et al., 2020; Hell, Pässler und Nido, 2023). Zum anderen besteht die Befürchtung, dass insb. Intelligenztests kaum durch Bewerbende akzeptiert werden (Hogrefe, 2022).

Die Diskrepanz zwischen Anwendungshäufigkeit und Qualität von Persönlichkeitstests ist schwieriger zu beantworten. Zum einen spielt hier wohl die Tatsache eine Rolle, dass sich über 3'000 verschiedene Persönlichkeitstest auf dem Markt befinden (siehe Kap. 1), was zu einer gewissen Entscheidungsschwierigkeit führt (Hell, 2013). Zum anderen wird im Zusammenhang mit dem Einsatz mangelhafter psychologischer Tests der Science-Practitioner-Gap diskutiert (z.B. Kehle, 2004; König et al., 2010, Arnoneit et al., 2020; Hossiep et al., 2022). Der Science-Practitioner-Gap, auch Research-Practice-Gap genannt, beschreibt die Kluft zwischen der personalpsychologischen Wissenschaft und der Praxis der Personalarbeit. Kehle (2004) erklärt das Zustandekommen dieser Kluft grob durch zwei verschiedene Punkte:

**1. Wissenschaftliche Erkenntnisse gelangen kaum in die Praxis**

Personen, die nicht in der Wissenschaft tätig sind, haben kaum Einblick in aktuelle Forschung und v.a. in aktuelle, fundierte Erkenntnisse. Es wäre aber zu einfach, den Einsatz mangelhafter Verfahren darauf zurückzuführen, dass Praktikerinnen und Praktiker sich nicht mit wissenschaftlichen Studien beschäftigen (Kehle, 2004). Viel eher ist der wohl nachfolgende 2. Punkt ausschlaggebend für den Science-Practitioner-Gap.

**2. Widersprüchliche Anforderungen durch Wissenschaft und Praxis**

Kehle (2004) gibt zu verstehen, dass sich Unternehmen konfrontiert sehen mit ökonomischen Anforderungen oder kontextuellen Erwartungen. So werden bspw. Personalauswahlverfahren nicht hinsichtlich ihrer Qualität, sondern bspw. aufgrund der vergleichsweise niedrigeren Kosten oder aufgrund von Branchenstandards gewählt.

## 2.4 Hypothesen

In der nun folgenden Tab. 5 werden Hypothesen genannt, die der Beantwortung der Fragestellungen (siehe Kap. 1) dienen sollen. Die Hypothesen wurden aus vorliegenden Theorien und Studien abgeleitet (siehe Kap. 2).

Tabelle 5

*Erstellte Hypothesen, die zur Beantwortung von der Fragestellungen dienen – aus vorliegenden Theorien und Studien abgeleitet.*

<b>Hypothesen zur Bearbeitung der Oberfragestellungen (OF)</b>
<b>Hypothesen zur Bearbeitung der Oberfragestellung OF 1 - Verwendungshäufigkeit</b>
H1.1: In der Schweiz tätige Unternehmen setzen zum Zweck der Personalauswahl und -entwicklung psychologische Tests ein (orientiert u.a. an Berchtold, 2005; Hell et al., 2021; Véron et al., 2024).
H1.2: Vergleichsweise werden mehr Persönlichkeitstests als Leistungstests eingesetzt (orientiert u.a. an Hell et al., 2021; Véron et al., 2024).
<b>Hypothese zur Bearbeitung der Oberfragestellung OF2 - Testart</b>
H1.3: In der Schweiz tätige Unternehmen setzen verschiedenste psychologische Tests im Rahmen von Personalentscheidungen ein (orientiert u.a. an Berchtold, 2005; Hossiep et al., 2015; Hell, 2013).
<b>Hypothesen zur Bearbeitung der Unterfragestellungen (UF)</b>
<b>Hypothese zur Bearbeitung der Unterfragestellung UF1 - Qualität eingesetzter Tests</b>
H1.4: Auch Tests, die nicht über eine Testrezension durch das TBS-DTK oder SDBB verfügen, bzw. betreffend ihrer Qualität in Kritik stehen, werden im Rahmen von Personalentscheidungen eingesetzt (orientiert u.a. an von Hell, 2013; Hossiep et al. 2015; Hossiep et al. 2022; Robins et al., 1998).
<b>Hypothese zur Bearbeitung der Unterfragestellung UF2 - Einsatzzweck</b>
H1.5: Psychologische Tests werden eher zum Zweck der Personalentwicklung als zum Zweck der Personalauswahl eingesetzt (orientiert an Hossiep et al., 2015).
<b>Hypothese zur Bearbeitung der Unterfragestellung UF3 - Zielgruppe für Tests</b>
H1.6: Je höher die vorgesehene Hierarchiestufe der zu rekrutierenden Person, desto eher werden Tests zur Personalauswahl eingesetzt (orientiert u.a. an Armoneit et al., 2020).
<b>Hypothese zur Bearbeitung der Unterfragestellung UF4 - Informationsbasis</b>
H1.7: Durch HR-Mitarbeitende werden verschiedenste Informationsquellen genutzt, um sich über einzusetzende Test zu informieren (grob orientiert u.a. an Kehle, 2004).
<b>Hypothese zur Bearbeitung der Unterfragestellungen UF5 bis UF7 - Hintergründe in Bezug auf Testeinsatz</b>
H1.8: Persönliche Beweggründe der HR-Mitarbeitenden begünstigen oder hemmen den Einsatz psychologischer Tests im Unternehmen, in welchem sie tätig sind (orientiert u.a. an Hossiep et al. 2015, Berchtold, 2005).
H1.9: Entscheidungskriterien, anhand derer HR-Mitarbeitende abwägen, begünstigen oder hemmen den Einsatz psychologischer Tests im Unternehmen, in welchem sie tätig sind (orientiert u.a. an Hossiep et al. 2015, Berchtold, 2005).
H1.10: Äussere Einflussgrössen begünstigen oder hemmen den Einsatz psychologischer Tests in einem Unternehmen (orientiert u.a. an Kehle, 2004; Hossiep et al., 2015; Berchtold, 2005).

### 3 Datenerhebungs- und Auswertungsmethoden

Dieses Kapitel beschäftigt sich mit den Datenerhebungs- und Auswertungsmethoden, die zur Untersuchung der entwickelten Forschungsfragen herangezogen wurden. Die Struktur des nachfolgenden Kapitels entspricht der Chronologie der durchgeführten Datenerhebung und Auswertung.

#### 3.1 Untersuchungsablauf

Das methodische Vorgehen entlang der in Kap. 1 formulierten Zielsetzung sieht ein Mixed Methods Design vor. Die nachfolgende Abb. 7 soll einen Überblick über das angestrebte methodische Vorgehen bieten. Daran anschliessend folgt die Einordnung der strukturellen Eigenschaften des geplanten Datenerhebungsdesigns.

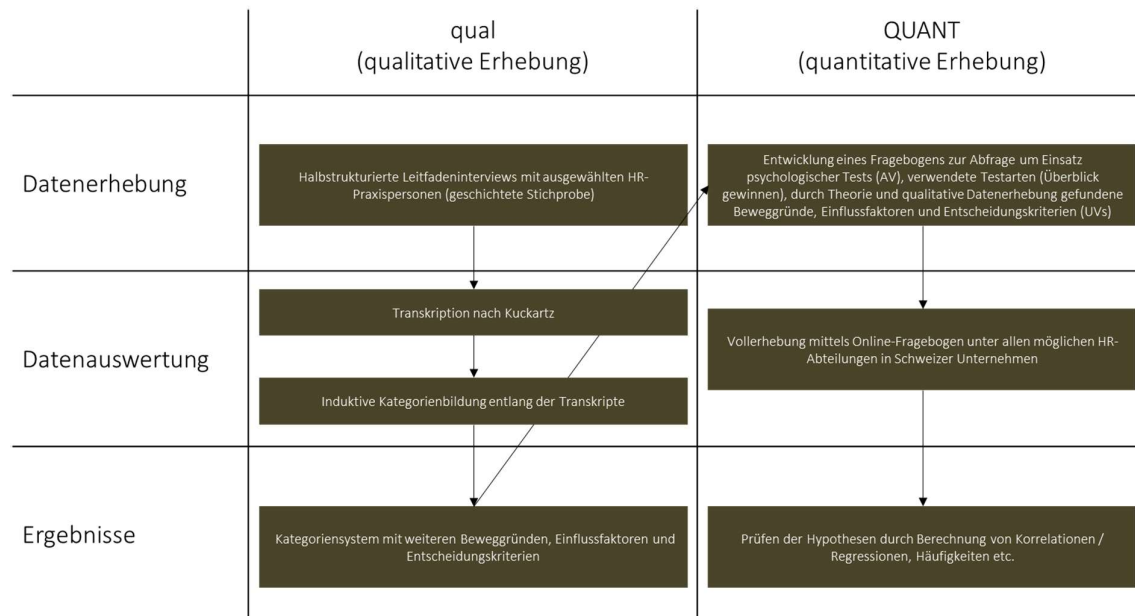


Abbildung 7. Geplantes Mixed Methods Design zur Untersuchung der Forschungsfragen und der Hypothesen.

## Eigenschaften des Mixed Method Designs

Die strukturelle Eigenschaften des gewählten Mixed Method Forschungsdesigns werden entlang der von Creswel (2003, in Kuckartz, 2014) vorgeschlagenen Struktur in der Abb. 8 beschrieben:

### Implementation

Das geplante Mixed Methods Design kann als sequenziell eingestuft werden. Zuerst erfolgt eine explorative qualitative Datenerhebung in Form von Expertinnen- und Experteninterviews. Dann folgt die darauf aufbauende explanative quantitative Datenerhebung durch einen online Fragebogen. Die qualitativen Daten sollen also quantifiziert werden.



### Priorität

Die Priorität liegt auf den Ergebnissen der quantitativen Studie. Das Forschungsinteresse liegt nicht in den Einzelfällen, sondern auf der gängigen Praxis hinsichtlich des Testeinsatzes in Schweizer Unternehmen (qual → QUANT).



### Integration

Die Datenintegration betreffend die qualitative Datenerhebung findet bei der Dateninterpretation statt. Die induktive Kategorienbildung entlang der Transkripte soll weitere mögliche Beweggründe, Entscheidungskriterien und Einflussgrößen hinsichtlich des Testeinsatzes in Unternehmen an den Tag bringen. Dies soll die Entwicklung eines online Fragebogens ermöglichen. Dieser hat eine quantitative Vollerhebung zum Ziel bei allen verfügbaren HR-Abteilungen von Unternehmen, die in der Schweiz tätig sind. Die Datenintegration betreffend die quantitative Datenerhebung erfolgt durch eine statistische Auswertung der erhobenen Fragebogendaten. Die qualitativ erhobenen Kategorien werden dadurch auf ihre Relevanz überprüft, folglich also quantifiziert.



### Theoretische Perspektive

Die qualitative Datenerhebung ist als explorativ einzustufen. Die Wortmeldungen der Interviewpartnerinnen und -partner bestimmen das darauffolgende quantitative Forschungsdesign. Daher kann die Perspektive übergeordnet als implizit bezeichnet werden. Des Weiteren werden die qualitativen Ergebnisse auch theoretisch untermauert, wodurch zusätzlich auch die Einnahme einer expliziten Perspektive erfolgt. Diese Perspektive ist aber vergleichsweise als untergeordnet zu bewerten.

Abbildung 8. Die strukturellen Eigenschaften des Mixed Method Design nach Creswel (2003, in Kuckartz, 2014).

## 3.2 Qualitative Datenerhebung

An den in Kap. 1 genannten Forschungsfragen wurde die qualitative Datenerhebung ausgerichtet. Der Prozess der qualitativen Datenerhebung umfasst fünf wesentliche Teilschritte. Die nachfolgende Abb. 9 bietet dazu einen Überblick und verweist auf die jeweiligen Unterkapitel, die die Teilschritte erläutern.

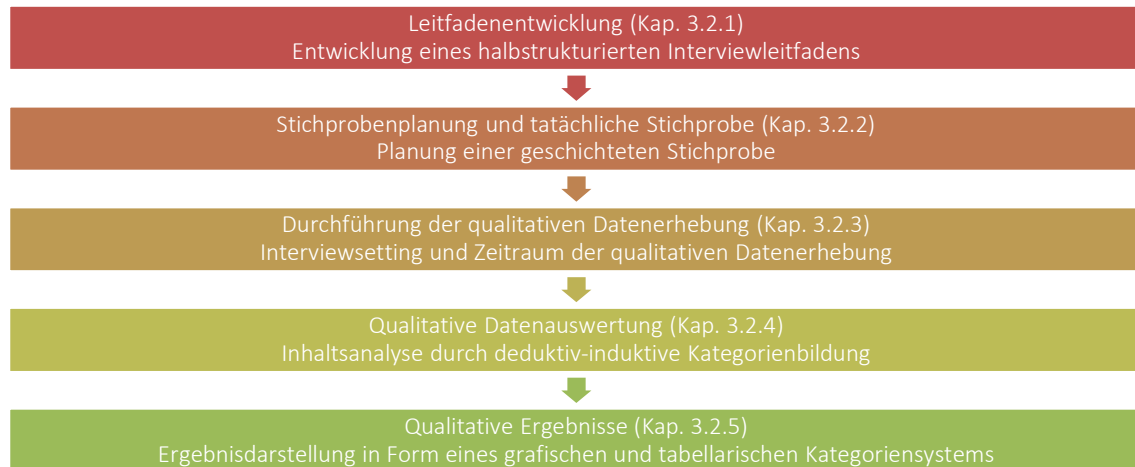


Abbildung 9. Fünf wesentliche Teilschritte der qualitativen Datenerhebung.

### 3.2.1 Leitfadententwicklung

Entlang der formulierten Forschungsfragen wurde ein halbstandardisierter Interviewleitfaden mit offenen Fragen konzipiert. Die Interviewpartnerinnen und -partner sollen zum Einsatz psychologischer Tests in ihrem Unternehmen befragt werden. Zudem werden Beweggründe, Einflussgrößen und Entscheidungskriterien erfragt, die sie bei ihrer Entscheidung für oder gegen den Testeinsatz geleitet haben.

Die Leitfadentkonzeption orientiert sich am «SPSS-Prinzip bei der Leitfadenterstellung» nach Helfferich (2011). Die nachfolgende Abb. 10 soll einen Einblick in die Erstellung des Leitfadens erlauben.



Abbildung 10. Erstellung des Interviewleitfadens entlang des SPSS-Prinzips bei der Leitfadenterstellung, orientiert an Helfferich (2011).

### 3.2.1.1 Aufbau des Leitfadens

In Anhang B kann der finalisierte Interviewleitfaden eingesehen werden. Der Interviewleitfaden beginnt mit dem Themenblock *Einstieg*. Neben der Abfrage des beruflichen Werdegangs enthält er auch sogenannte Eisbrecher-Fragen, die die Interviewteilnehmenden auf die Thematik der psychologischen Tests einstimmen sollen. Im Themenblock *Oberfragestellungen* wird die aktuelle Handhabung betreffend Testeinsatz im Unternehmen erfragt. Die hier erhaltenen Antworten entscheiden über den weiteren Verlauf des Interviews: berichten Interviewpartnerinnen oder -partner über den tatsächlichen Einsatz psychologischer Tests zu Personalentscheidungen, folgt der Themenblock *Unterfragestellungen 1.1 (UF1.1)*. Berichten Interviewpartnerinnen oder -partnern, dass sie in ihrem Unternehmen keine psychologischen Tests einsetzen, folgt der Themenblock *Unterfragestellungen 1.2 (UF1.2)*. Beide genannten Themenblöcke UF1.1 und UF1.2 erfragen die einflussnehmenden Gründe auf die Entscheidung für bzw. gegen den Testeinsatz.

Der darauffolgende Themenblock *Unterfragestellungen 2* ist an keine Bedingungen geknüpft. Dieser erfragt die Handhabung psychologischer Tests an früheren Arbeitsorten der Interviewteilnehmenden. Der letzte Themenblock *Abschluss* soll den Teilnehmenden ermöglichen, Anmerkungen zu machen oder Fragen zu stellen. Daran anschliessend folgt die Abfrage von Kontaktdaten, um das jeweilige Transkript und die Masterarbeit nach Abschluss per E-Mail zusenden zu können.

Helfferrich (2011) empfiehlt eine Gliederung der Interviewfragen in *erzählauffordernde Leitfragen*, *konkrete Fragen*, *Aufrechterhaltungs- und Steuerungsfragen* sowie den Einbau einer *Checkliste*.

Sich an dieser Empfehlung orientierend, enthält der entwickelte Leitfaden neben den erzählauffordernden Leitfragen auch eine Checkliste, welche bei der Interviewdurchführung helfen soll, Kontrolle über die erwähnten Inhalte zu behalten. Sollte sich herausstellen, dass über gewisse Inhalte nichts oder zu wenig erzählt wurde, können die Punkte der Checkliste auf konkrete Fragen verweisen, die dann als Nachfragen gestellt werden sollen.

Um vorbereitet zu sein auf ins Stocken geratende oder zu wenig ausführliche Erzählungen, wurden einige Aufrechterhaltungs- und Steuerungsfragen formuliert. Diese können während der Interviewdurchführung sporadisch und bei Bedarf eingesetzt werden können.

### 3.2.2 Stichprobenplanung und tatsächliche Stichprobe

Das Interesse liegt in der Reichhaltigkeit der Aussagen der Probandinnen und Probanden. Daher wird eine geschichtete Stichprobe hinsichtlich derjenigen Wirtschaftsabteilungen der Schweizer Wirtschaft angestrebt, die die meisten Mitarbeitenden (in Vollzeitäquivalenten) beschäftigen.

Weichbrodt (2022) empfiehlt bei halbstrukturierten Leitfadeninterviews eine Stichprobengrösse von zehn bis 20 Probanden und Probandinnen. Entlang der Anteile des jeweiligen Umfangs der Wirtschaftsabteilungen wurde die notwendige Stichprobengrösse mehr oder minder proportional berechnet. Bei Wirtschaftsabteilungen mit über 200'000 Mitarbeitenden im Vollzeitäquivalent ist je ein Interview und bei Wirtschaftsabteilungen mit über 500'000 Mitarbeitenden im Vollzeitäquivalent sind je ein bis zwei Interviews geplant. Die nachfolgende Tab. 6 soll dies aufzeigen. Im Anhang C befindet sich die entsprechende Tabelle mit allen Rohdaten.

Tabelle 6

Planung einer geschichteten Stichprobe für die qualitative Datenerhebung entlang derjenigen Wirtschaftsabteilungen der Schweiz, die die meisten Mitarbeitenden (in Vollzeitäquivalenten) aufweisen, orientiert an BFS (2023b).

Wirtschaftsabteilungen mit über 200'000 Mitarbeitenden in Vollzeitäquivalenten im Jahr 2023	aktuelle Quartalswerte in Vollzeitäquivalenten nach Wirtschaftsabteilungen		geplante Anzahl an Interviews pro Wirtschaftsabteilung	tatsächlich durchgeführte Interviews pro Wirtschaftsabteilung	Interviewpartnerschaften
	1. Quartal	2. Quartal			
<b>Total</b>	4'205'983	4'234'590	10 – 12	8	
<b>Verarbeitendes Gewerbe / Herstellung von Waren</b>	633'047	638'344	2	2	1: Produktion Spezialitätenchemie 2: Maschinenbau und -konstruktion
<b>Baugewerbe / Bau</b>	341'089	344'066	1	1	Bau- und Generalunternehmen
<b>Handel; Instandhaltung und Reparatur von Kraftfahrzeugen</b>	517'403	518'025	1 – 2	0	-
<b>Verkehr und Lagerei</b>	224'135	219'149	1	1	Logistikunternehmen
<b>Gastgewerbe / Beherbergung und Gastronomie</b>	197'931	200'248	1	0	-
<b>Erbringung von Finanz- und Versicherungsdienstleistungen</b>	219'273	220'097	1	1	Versicherungs- und Finanzdienstleisterin
<b>Erbringung von wissenschaftlichen und technischen Dienstleistungen</b>	397'381	397'942	1	1	Pharmakonzern
<b>Erziehung und Unterricht</b>	253'161	250'411	1	1	Berufsfach- und Maturitätsschule
<b>Gesundheits- und Sozialwesen</b>	575'471	579'454	1 – 2	1	Krankenhaus
<b>Erbringung von sonstigen wirtschaftlichen Dienstleistungen</b>	238'508	256'608	0	0	-

Anmerkung: Die Wirtschaftsabteilung *Erbringung von sonstigen wirtschaftlichen Dienstleistungen* zählt ebenfalls mehr als 200'000 Mitarbeitende in Vollzeitäquivalenten. Die Zusammensetzung der Wirtschaftsabteilung ist aber zu diffus, um einen eindeutig zugehörigen Interviewpartner bzw. -eine eindeutig zugehörige Interviewpartnerin finden zu können. Daher wurde diese Wirtschaftsabteilung ausgeklammert.

Um Interviewteilnehmende finden zu können, wurde u.a. auf das eigene Netzwerk zurückgegriffen. So konnten zwei der acht Interviewtermine vereinbart werden. Die Suche im Web, Recherchen auf Karriereportalen wie LinkedIn und das Durchforsten von Stellenanzeigen bestimmter Unternehmen nach HR-Kontakten dienten dazu, weitere Interviewteilnehmende zu finden.

Wie die Tab. 6 zeigt, weicht die Anzahl tatsächlich durchgeführter Interviews von der geplanten Anzahl ab. Die Mindestanzahl an Interviews, wie diese durch Weichbrodt (2022) empfohlen wird, konnte knapp nicht erreicht werden. Absagen und negative Bescheide auf Interviewanfragen führten dazu. Guest, Namey und Chem (2020) plädieren auf das Beobachten der Datensättigung zur Bestimmung einer angemessenen Interviewanzahl und verweisen darauf, dass insb. die Diversität der Interviewteilnehmenden und deren Erfahrung zu beachten sei.

Die geschichtete Stichprobe weist eine grosse Variabilität hinsichtlich der Wirtschaftsabteilungen auf. Zudem liess sich Datensättigung bei den durchgeführten Interviews beobachten, was auch die Häufigkeiten der vergebenen Kategorien in gewisser Weise zu unterstreichen vermag.

### 3.2.3 Durchführung der qualitativen Datenerhebung

Bereits bei der Terminvereinbarung wurden die Teilnehmenden aufgeklärt über die auditive Aufnahme, den Datenschutz und die Möglichkeit, das Interview zu widerrufen. Am Interviewtermin selbst haben die Teilnehmenden vorgängig eine Einverständniserklärung (siehe Anhang D) erhalten. Darin wurden über den Zweck und die Verwendung des Interviews aufgeklärt. Auf eine vorgängige Zusendung des Leitfadens wurde verzichtet, um die Unvoreingenommenheit der Teilnehmenden zu begünstigen.

Die Experten- und Expertinneninterviews wurden in einem Zeitraum von sechs Wochen durchgeführt (vom 27.11.2023 bis 09.01.2024). Fünf der acht Interviews wurden online mittels Microsoft Teams durchgeführt. Die anderen drei Interviews fanden am Arbeitsort der Interviewteilnehmenden statt. Entscheidend für die Durchführungsmodalitäten waren die jeweilige Wegdistanz zwischen Interviewerin und Expertin bzw. Experte und die Präferenz der interviewten Personen.

Die Dauer der Interviews reichte von knapp 30 Minuten bis zu mehr als einer Stunde. Die Interviews wurden auditiv aufgezeichnet und währenddessen zusätzlich handschriftliche Notizen gemacht. Die auditiven Aufnahmen dienten der Erstellung der Transkripte (siehe Anhang vB) und der späteren Datenanalyse, die handschriftlichen Notizen der Erstellung der jeweiligen Memos (siehe Anhang vA). Die ethische Unbedenklichkeit der geplanten Interviews und der noch folgenden quantitativen Datenerhebung wurde mit der Begleitperson Prof. Dr. Benedikt Hell besprochen und die Ethik-Checkliste der FHNW entsprechend bearbeitet (siehe Anhang vC).

### 3.2.4 Qualitative Datenauswertung

Die digital aufgezeichneten Expertinnen- und Experteninterviews wurden zum Zweck der qualitativen Datenauswertung mit der Software MAXQDA transkribiert und dann ausgewertet.

Die Transkription erfolgte nach den Regeln von Kuckartz, Dresing, Rädiker und Stefer (2007), die durch Einfachheit und damit schneller Erlernbarkeit überzeugten. Anhand dieser Regeln wurden die in schweizerdeutschem Dialekt geführten Interviews ins Hochdeutsche übersetzt. Sprache und Interpunktion wurden dabei geglättet. Nonverbale Äusserungen, die eine Aussage verdeutlichten oder unterstützten, wurden ebenfalls ins Transkript aufgenommen. Demnach wurden längere Pausen durch Auslassungspunkte (...) markiert und Lautäusserungen wie Lachen oder Seufzen in Klammern notiert. Die Absätze im Transkript entsprechen einem Sprecher- bzw. Sprecherinnenwechsel. Um dies zu verdeutlichen, wurden Passagen der Interviewerin mit einem «I», die Passagen der befragten Personen mit einem «B» gekennzeichnet. Die Transkripte wurden im Anschluss einer Überprüfung auf Rechtschreibfehler unterzogen und dann anonymisiert. Zur Anonymisierung wurden Namen von Personen und Firmen durch ein X ersetzt.

Die Transkripte wurden anschliessend der qualitativen Inhaltsanalyse unterzogen. Das Vorgehen orientierte sich am Ablaufmodell zur *deduktiv-induktiven Kategorienbildung und -anwendung* nach Mayring und Brunner (2006). Zusammenfassend lässt sich sagen, dass die Oberkategorien auf deduktivem Weg erhalten wurden. Diese lagen bereits durch die Fragestellungen UF5 bis UF7 vor. Die Unterkategorien wurden auf induktivem Weg entwickelt. Die nachfolgende Abb. 11 bietet einen Überblick über die durchgeführte Inhaltsanalyse. Die darin enthaltenen Punkte werden daran anschliessend detailliert besprochen.

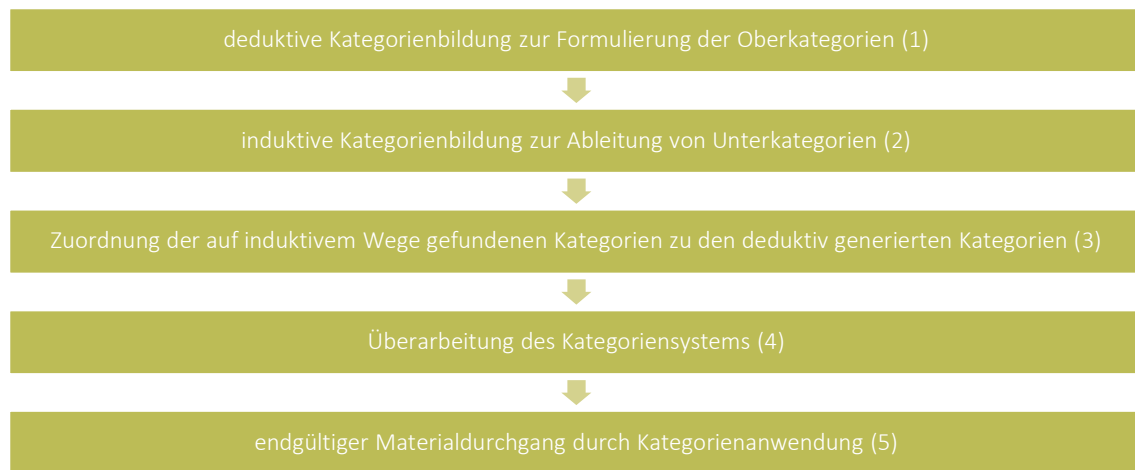


Abbildung 11. Ablaufmodell zur deduktiv-induktiven Kategorienbildung und -anwendung, orientiert an Mayring und Brunner (2006).

1. Die deduktive Kategorienbildung wurde bereits durch die Definition der Fragestellungen vorgegeben (siehe Kap. 1). Demnach werden den Testeinsatz unterstützende oder hinderliche Beweggründe, Entscheidungskriterien und äussere Einflussfaktoren in dieser Inhaltsanalyse als Oberkategorien angewendet.
2. Die Kategorienbildung erfolgte in einem zweiten Schritt auf induktivem Wege. So wurden die acht Transkripte nach einflussnehmenden Faktoren durchsucht. Hilfreich hierzu waren auch die angefertigten Memos (siehe Anhang vA), die bereits eine erste Idee gegeben hatten, welche einflussnehmenden Faktoren in den Transkripten vorliegen werden.
3. Die Sammlung dieser induktiv gefundenen Faktoren wurde anschliessend einer strukturgebenden Inhaltsanalyse unterzogen. Die einflussnehmenden Faktoren wurden den deduktiv gebildeten Oberkategorien «unterstützend» und «hinderlich» zugeordnet. Dann folgte innerhalb dieser beiden Oberkategorien die Zuordnung zu den Kategorien «Beweggründe», «Entscheidungskriterien» und «äussere Einflussgrössen». Dazu wurden systematische Kodierregeln entworfen, wie dies Kuckartz (2018) empfiehlt. Die nachfolgenden Fragen dienten als Kodierregeln und erwiesen sich als äusserst hilfreich, um die Zuordnung vorzunehmen:
  - **Beweggrund:** handelt es sich um eine persönliche Erfahrung oder eine persönliche Meinung?
  - **Entscheidungskriterium:** handelt es sich um eine Abwägung von Chancen zu Risiken, bzw. von Aufwand zu Nutzen?
  - **Äusseres Einflusskriterium:** handelt es sich um Situationen, Tatsachen oder Bedingungen, mit welchen die befragten Expertinnen oder Experten sich konfrontiert sehen, bzw. durch die Befragten nicht selbst verändert oder beeinflusst werden können?
4. Die erhaltene Kategorienstruktur wurde schliesslich überarbeitet, um das Kategoriensystem übersichtlicher und greifbarer zu gestalten. Dazu wurden zum einen sehr ähnliche Unterkategorien zusammengefasst. Zum anderen wurden thematisch verwandte Unterkategorien gruppiert.
5. Mit dem finalisierten Kategoriensystem wurden die acht Interviewtranskripte erneut durchforstet, um fehlende Transkriptpassagen zu kodieren. Das Kategoriensystem erwies sich beim erneuten Materialdurchgang als passend für die Forschungszwecke. In dem nun folgenden Unterkapitel 3.2.5 werden die qualitativen Ergebnisse dargestellt und besprochen.

### 3.2.5 Qualitative Ergebnisse

Das Ergebnis der durchgeführten Interviews und deren qualitative Auswertung ist ein Kategoriensystem. Dieses wird nachfolgend dargestellt, wobei zuerst die begünstigenden Faktoren (siehe Abb. 12) und dann die hemmenden Faktoren (siehe Abb. 13) folgen. Die daran anschließende Tab. 7 zeigt die Definitionen der deduktiv-induktiv gebildeten Kategorien, welche mit einigen Ankerzitate untermauert werden. Wie bereits in Kap. 3.1 zur theoretischen Perspektive des Mixed Method Designs angekündigt, werden die Unterkategorien mit relevanter Literatur gestützt. Da es sich um eine eher kleine Stichprobe handelt, soll dies den gebildeten Kategorien neben den Ankerzitate zusätzliches Gewicht verleihen.

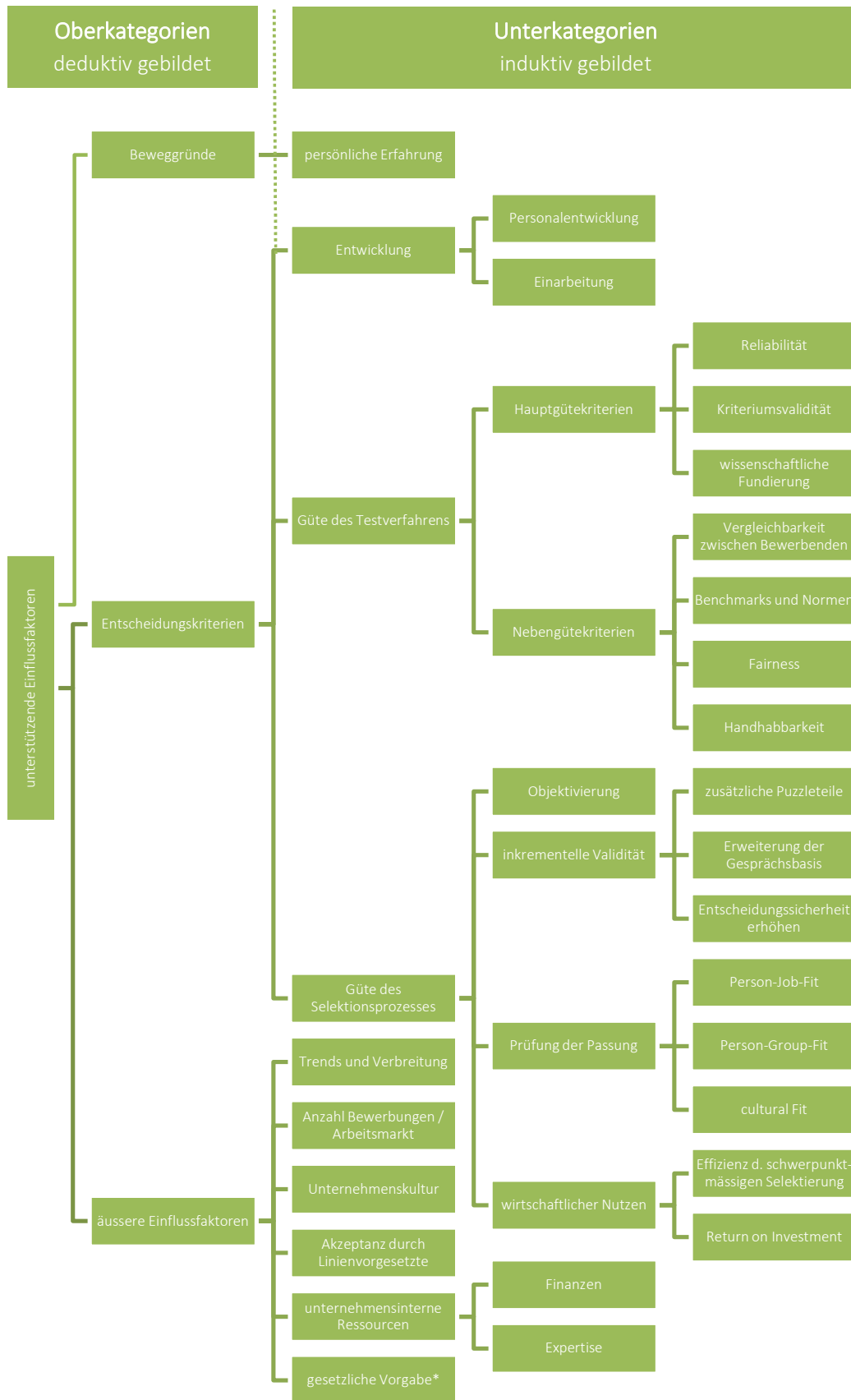


Abbildung 12. Kategoriensystem. Unterstützende Einflussfaktoren zum Testeinsatz. \* = äusserer Einflussfaktor ohne gegenteiliges Pendant.

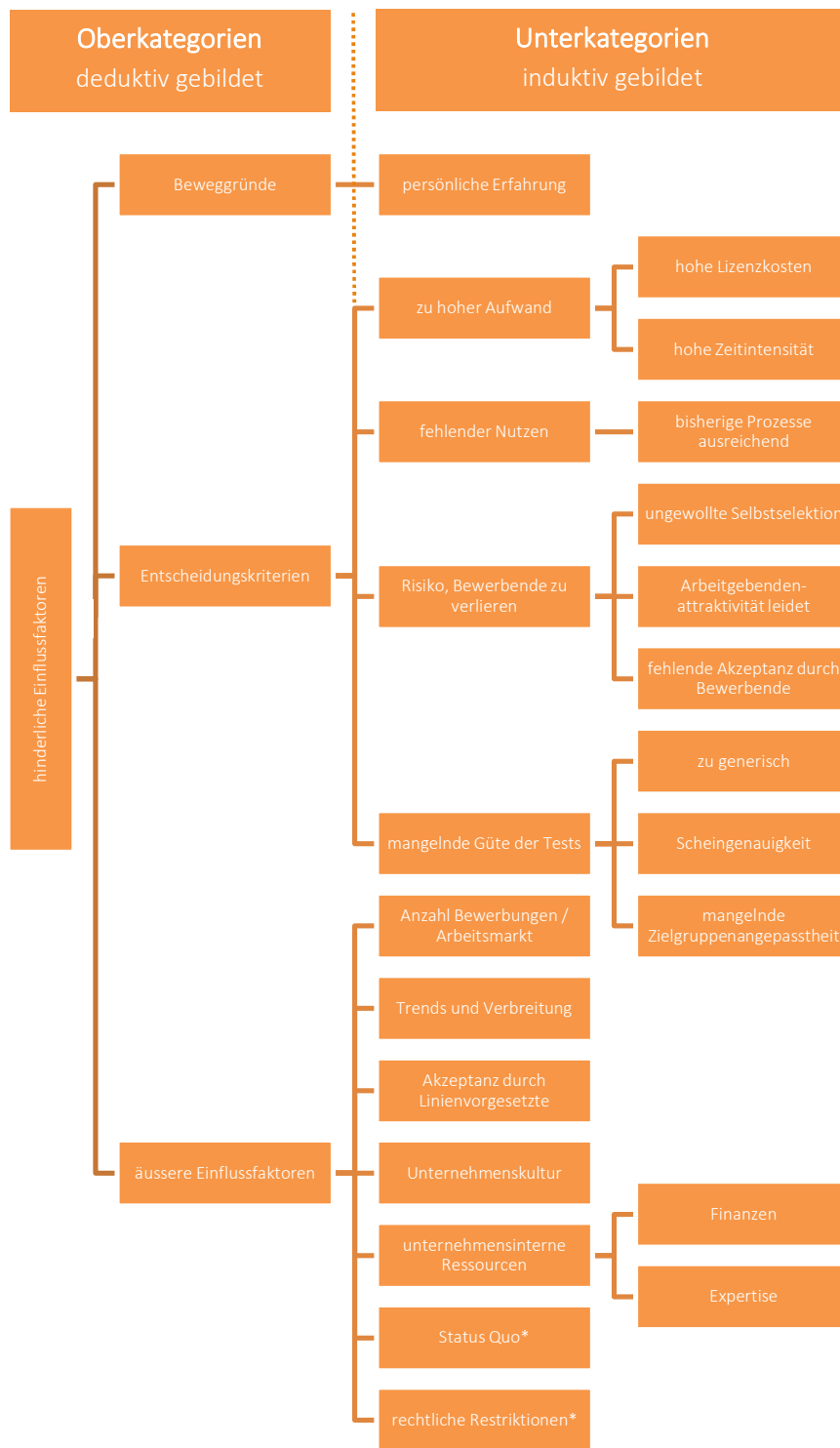


Abbildung 13. Kategoriensystem. Hinderliche Einflussfaktoren zum Testeinsatz. \* = äussere Einflussfaktoren ohne gegenteiliges Pendant.

Tabelle 7

Kategoriensystem mit Beispielzitate, Definitionen und Untermauerung durch Literatur.

begünstigende Kategorien		Beispielzitate	Definition der Kategorien und Untermauerung durch Literatur				
persönliche Beweggründe	positive persönliche Erfahrung	Es ist doch auch ein bisschen der Hintergedanke, zu schauen, bleibst Du wirklich bei der Sache. Was gibst Du unter Druck. Aber mega spannend im Nachhinein. Wenn ich jetzt so schaue, wenn wir das den Kandidaten, den Kandidatinnen vorlegen. Ich bin Fan, nein, ich bin Fan. Auch noch nach drei Monaten. Int. 4, Pos. 19	Positive persönliche Erfahrungen unterstützen eine positive Einstellung ggü. psychologischer Tests und unterstützen schliesslich auch den Einsatz psychologischen Tests im Unternehmen.  Dieser Umstand kann durch die Theory of Reasoned Action (TRA) von Fishbein und Ajzen (1975) erklärt werden: Verhaltensintention begünstigt tatsächliches Verhalten in hohem Masse. Verhaltensintention wiederum wird u.a. begünstigt durch die Einstellung der Person ggü. dem Verhalten. Persönliche Einstellungen wiederum sind geprägt von persönlichen Erfahrungen.				
		Entwicklungsmöglichkeiten	<table border="1"> <tr> <td>Personalentwicklung</td> <td>Also, jetzt einmal in Bezug auf Persönlichkeitsfragebogen. Das finde ich es mega interessant. Wenn Du nicht einen psychologischen Hintergrund hast, da würde ich behaupten (.) ich würde nicht behaupten, dass Du dann nicht sozial erwünscht antworten würdest, aber Du, Du beschäftigst Dich mal mit Dir selbst, mit Deinem Selbstbild, mit Deiner Selbsteinschätzung. Int. 4, Pos. 21</td> <td>Psychologische Tests werden eingesetzt, um Fähigkeiten und Potenziale bei Mitarbeitenden operationalisieren zu können. Dadurch lassen sich Möglichkeiten zur internen Personalentwicklung ableiten oder geeignete interne Kandidatinnen und Kandidaten für Beförderungen auswählen. Zudem lässt sich durch den Einsatz von Persönlichkeitstests auch Personalentwicklung initiieren: Mitarbeitende können durch die Beschäftigung mit ihren Testresultaten Selbstexploration bzw. -reflexion betreiben.  Diese Kategorie lässt sich u.a. auch durch die Studie von Hossiep et al. (2022) stützen. Sie konnten durch ihre Erhebung unter den grössten deutschen Unternehmen aufzeigen, dass Persönlichkeitstests insb. zu Personalentwicklungszwecken wie Selbstexploration der Mitarbeitenden, Teamentwicklungen und Coachings eingesetzt werden.</td> </tr> <tr> <td>Einarbeitung</td> <td>Es kann aber auch gut sein, dass gerade für bspw. eine Einarbeitung, weil Du siehst dann vielleicht gewisse Punkte, die rausgekommen sind bei dem Ergebnis, wo man vielleicht noch den Fokus legen müsste bei einer Einarbeitung. Int. 6, Pos. 17</td> <td>Psychologische Tests werden bei der Personalauswahl eingesetzt, um bei Kandidatinnen und Kandidaten auch individuelle Einarbeitungsbedarfe feststellen zu können.  Einhgehend mit der gerade eben genannten Kategorie <i>Personalentwicklung</i> lässt sich diese Kategorie als frühe Personalentwicklung verstehen. Personalentwicklung findet bereits «into-the-job» durch Einarbeitungsmassnahmen statt (Scholz, 2000).</td> </tr> </table>	Personalentwicklung	Also, jetzt einmal in Bezug auf Persönlichkeitsfragebogen. Das finde ich es mega interessant. Wenn Du nicht einen psychologischen Hintergrund hast, da würde ich behaupten (.) ich würde nicht behaupten, dass Du dann nicht sozial erwünscht antworten würdest, aber Du, Du beschäftigst Dich mal mit Dir selbst, mit Deinem Selbstbild, mit Deiner Selbsteinschätzung. Int. 4, Pos. 21	Psychologische Tests werden eingesetzt, um Fähigkeiten und Potenziale bei Mitarbeitenden operationalisieren zu können. Dadurch lassen sich Möglichkeiten zur internen Personalentwicklung ableiten oder geeignete interne Kandidatinnen und Kandidaten für Beförderungen auswählen. Zudem lässt sich durch den Einsatz von Persönlichkeitstests auch Personalentwicklung initiieren: Mitarbeitende können durch die Beschäftigung mit ihren Testresultaten Selbstexploration bzw. -reflexion betreiben.  Diese Kategorie lässt sich u.a. auch durch die Studie von Hossiep et al. (2022) stützen. Sie konnten durch ihre Erhebung unter den grössten deutschen Unternehmen aufzeigen, dass Persönlichkeitstests insb. zu Personalentwicklungszwecken wie Selbstexploration der Mitarbeitenden, Teamentwicklungen und Coachings eingesetzt werden.	Einarbeitung
Personalentwicklung	Also, jetzt einmal in Bezug auf Persönlichkeitsfragebogen. Das finde ich es mega interessant. Wenn Du nicht einen psychologischen Hintergrund hast, da würde ich behaupten (.) ich würde nicht behaupten, dass Du dann nicht sozial erwünscht antworten würdest, aber Du, Du beschäftigst Dich mal mit Dir selbst, mit Deinem Selbstbild, mit Deiner Selbsteinschätzung. Int. 4, Pos. 21	Psychologische Tests werden eingesetzt, um Fähigkeiten und Potenziale bei Mitarbeitenden operationalisieren zu können. Dadurch lassen sich Möglichkeiten zur internen Personalentwicklung ableiten oder geeignete interne Kandidatinnen und Kandidaten für Beförderungen auswählen. Zudem lässt sich durch den Einsatz von Persönlichkeitstests auch Personalentwicklung initiieren: Mitarbeitende können durch die Beschäftigung mit ihren Testresultaten Selbstexploration bzw. -reflexion betreiben.  Diese Kategorie lässt sich u.a. auch durch die Studie von Hossiep et al. (2022) stützen. Sie konnten durch ihre Erhebung unter den grössten deutschen Unternehmen aufzeigen, dass Persönlichkeitstests insb. zu Personalentwicklungszwecken wie Selbstexploration der Mitarbeitenden, Teamentwicklungen und Coachings eingesetzt werden.					
Einarbeitung	Es kann aber auch gut sein, dass gerade für bspw. eine Einarbeitung, weil Du siehst dann vielleicht gewisse Punkte, die rausgekommen sind bei dem Ergebnis, wo man vielleicht noch den Fokus legen müsste bei einer Einarbeitung. Int. 6, Pos. 17	Psychologische Tests werden bei der Personalauswahl eingesetzt, um bei Kandidatinnen und Kandidaten auch individuelle Einarbeitungsbedarfe feststellen zu können.  Einhgehend mit der gerade eben genannten Kategorie <i>Personalentwicklung</i> lässt sich diese Kategorie als frühe Personalentwicklung verstehen. Personalentwicklung findet bereits «into-the-job» durch Einarbeitungsmassnahmen statt (Scholz, 2000).					
Entscheidungskriterien	Hauptgüte des Testverfahrens selbst	Reliabilität	(...) obwohl es natürlich im Prinzip wenig Bias hat. Int. 8, Pos. 19  Ich glaube, IQ-Tests sind das, das am besten reproduzierbare Instrument in der Psychologie. Int. 8, Pos. 13	Der Einsatz psychologischer Tests wird dadurch gefördert, dass Tests über Reliabilität verfügen.  Reliabilität bedeutet, dass die durch die Testpersonen erreichten Ergebnisse als zuverlässig und genau verstanden werden können (Bühner, 2011). Dadurch lassen sich schliesslich zuverlässige Aussagen treffen oder Massnahmen ableiten.			
		Kriteriumsvalidität	Da hat man bei Hunderten, nein eher Tausenden Mitarbeitenden geschaut, die durch die Selektions-ACs am Anfang gegangen sind. Wie sehen diese Resultate aus, wie korrelieren diese Resultate, mit den drei oder fünf Jahren in der Rolle mit der Performance. Gibt es eine Übereinstimmung, eine positive Übereinstimmung zwischen guten Selektionsresultaten, zieht sich das durch. (...) Die Korrelation war sichtbar und sie war positiv gewesen. Zurück zu der Frage, hat es sich gelohnt: ja. Int. 1, Pos. 118-120	Psychologische Tests überzeugen, weil damit Leistung der getesteten Personen vorhergesagt werden kann. Damit wird das Gütekriterium der prognostischen Kriteriumsvalidität angesprochen.  König et al. (2010) konnten durch ihre Studie einen positiven Einfluss von <i>Wahrnehmung der prognostischen Validität eines Testinstruments</i> und dem tatsächlichen Einsatz des Testinstruments darlegen (vgl. Kap. 2.3).			
		wissenschaftliche Fundierung	Ich sehe da, wie soll ich sagen. Vorteile, ja, die lügen nicht. Int. 4, Pos. 32  Ja, also einerseits, dass es wissenschaftlich anscheinend fundiert ist. Int. 6, Pos. 62	Der Einsatz psychologischer Tests lässt sich durch deren Konstruktvalidität begründen. Die wissenschaftliche Fundierung spricht für den Einsatz von psychologischen Tests.  Die Ergebnisse der Studie von Hossiep et al. (2015) stützt dies. Der Verfahrensaspekt der wissenschaftlichen Fundierung ist bei Anwendenden von besonders hoher Wichtigkeit (vgl. Kap. 2.3).			

begünstigende Kategorien		Beispielzitate	Definitionen der Kategorien und Untermauerung durch Literatur			
Entscheidungskriterien	Nebengüte des Testverfahrens selbst	Vergleichsmöglichkeiten				
		Vergleichbarkeit zwischen Bewerbenden	Man kann vergleichen, vor allem, wenn man gleich mehrere Bewerbende auf dieselbe Position hat. Int. 3, Pos. 21	Der Testeinsatz begründet sich darin, dass man unter verschiedenen Bewerbenden für dieselbe Arbeitsstelle durch deren Testresultate Vergleiche ziehen kann. Man kann hier von einem unstandardisierten Vergleich sprechen, der die Personalentscheidung erleichtert.		
		Benchmarks und Normen	Und durch das, dass man auch international vernetzt ist mit unterschiedlichen Normen, glaube ich, hilft das auch noch. Int. 7, Pos. 93	Der Einsatz von psychologischen Tests bzw. die Auswahl eines bestimmten Instruments ist davon abhängig, ob Benchmarks oder Normen, also standardisierte Vergleichsmassstäbe, vorhanden sind, mit denen man eigens erhobene Ergebnisse vergleichen kann.  Im Vergleich zu der obengenannten Kategorie <i>Vergleichbarkeit zwischen Bewerbenden</i> handelt es sich hier um eine standardisierte Vergleichsmöglichkeit. An den standardisierten Vergleichsmassstäben kann man die erhaltenen Testresultate einordnen und schliesslich objektiv interpretieren (Krohne & Hock, 2015).		
		Fairness	Was ich eigentlich damit sagen will, wir gehen mit einem standardisierten Fragebogen durch das Ganze. Um einfach objektiv zu bleiben und allen Kandidatinnen und Kandidaten die gleichen Voraussetzungen zu schaffen. Int. 4, Pos. 62	Psychologische Tests werden aufgrund der durch sie geschaffenen Fairness unter allen Teilnehmenden befürwortet. Durch den Testeinsatz werden für verschiedene Kandidatinnen und Kandidaten gleiche Voraussetzungen geschaffen.  Stereotypisierung, persönliche Vorlieben oder Befangenheit der rekrutierenden Person haben also keinen Einfluss auf die Interpretation der Testresultate. Währenddessen sind andere Personalauswahlverfahren, wie (unstandardisierte) Bewerbungsgespräche anfälliger für genannte kognitive Verzerrungen (Moosbrugger und Kelava, 2020).		
	Qualität / Güte des Selektionsprozesses	inkrementelle Validität	Handhabbarkeit	Aber auch, weil es ein sehr intuitiver (..) also, die Ergebnisberichte, die Du nachher bekommst, die sind sehr intuitiv zu verstehen. Für die HR-Mitarbeitenden, aber auch für die Bewerber selbst. Auch wirklich gut aufgebaut, also, sehr übersichtlich. Und einfach verständlich. Und sie sind wirklich sehr gut aufgebaut. Int. 6, Pos. 62	Die empfundene Handhabbarkeit bzw. Praktikabilität entscheidet über den Einsatz psychologischer Tests im Rahmen von Personalentscheidungen.  Die Studie von Hossiep et al. (2015) unterstützt dies: Sie konnten in ihrer Studie aufzeigen, dass <i>die Wahrnehmung der Handhabbarkeit eines Tests</i> als wichtiger Verfahrensaspekt einzustufen ist. Arnoneit et al. (2020) konnten in ihrer Studie einen signifikanten Wirkzusammenhang zwischen <i>wahrgenommener Handhabbarkeit</i> und dem tatsächlichen Einsatz von online Persönlichkeitstests im Rahmen von Personalauswahl darlegen (vgl. Kap. 2.3).	
			Objektivierung	Es gibt bei vielen Führungskräften die Tendenz, einfach Leute, mit denen man selber gut kann oder mit denen man irgendwie sozusagen Chemie hat, einzustellen, vielleicht auch Leute, die einem sehr ähnlich sind und so, und weniger, was das Team zu Erfolg bringt. Ja, also alles, was mit der Objektivität zu tun hat, vom Selektionsprozess. Int. 8, Pos. 25	Dass durch psychologische Tests der Bewerbungsprozess objektiviert wird, unterstützt deren Einsatz.  Objektivität ist dann gegeben, wenn dem Testdurchführenden kein Handlungsspielraum bei Durchführung, Auswertung oder Interpretation gegeben ist (Moosbrugger & Kelava, 2020). Folglich werden dadurch auch kognitive Verzerrungen, Stereotypen, vorgefertigte Meinungen des Testdurchführenden ausgeschlossen bzw. im Selektionsprozess abgemildert (vgl. Kap. 2.2.1).	
			zusätzliche Puzzleteile	Aber meiner Meinung nach ist das halt noch ein Zusatzinstrument, um die Validität der Bewerbung auch sicherzustellen. Und auch noch ein anderes Bild zu bekommen von der Person, als das eigene Gesagte. Int. 2, Pos. 19	Dass das Bild des Bewerbenden, welches man durch CV-Check und Bewerbungsgesprächen erhalten hat, durch die Ergebnisse psychologischer Tests erweitert werden kann, spricht für deren Einsatz.  Damit wird die inkrementelle Validität angesprochen. Sie beschreibt in diesem Zusammenhang den zusätzlichen Nutzen durch psychologische Tests bei Personalentscheidungen. Psychologische Tests ergänzen in diesem Beispiel den CV-Check und das Bewerbungsgespräch und verbessern damit die Vorhersage der Eignung und Leistung des Bewerbenden (Bühner, 2011; Schmidt et al., 2016; Schmidt und Hunter, 1998) (vgl. Kap. 2.2.1.3.3).	
		Erweiterung der Gesprächsbasis	Erweiterung der Gesprächsbasis	Es ist aber auch ein gutes Instrument für die Gespräche selbst. Ich finde, wenn man die Ergebnisse bespricht mit den Bewerbern, dann entwickeln sich auch durch die Besprechung relativ gute Gespräche, wodurch Du auch wieder zusätzliche Sachen erfahren kannst. Int. 6, Pos. 17	Aus den Ergebnissen der durchgeführten psychologischen Tests lassen sich Interviewfragen ableiten. Damit lässt sich die Gesprächsbasis nicht nur erweitern, sondern auch vertiefen. Diese Möglichkeit begünstigt den Einsatz psychologischer Tests. Auch damit wird die inkrementelle Validität angesprochen (vgl. ebengenannte Unterkategorie <i>zusätzliche Puzzleteile</i> oder Kap. 2.2.1.3.3).	
			Entscheidungssicherheit erhöhen	Entscheidungssicherheit erhöhen	Der Einsatz von solchen Instrumenten, gerade, wenn man ein Stellen-, bzw. Anforderungsprofil hinterlegen kann, hat für mich auch ganz viel damit zu tun, dass man die Entscheidungssicherheit erhöhen kann. Int. 7, Pos. 55	Die zusätzlichen Informationen, die durch den Einsatz psychologischer Tests erhalten werden können, können die Entscheidungssicherheit bei Personalauswahl oder -entwicklung erhöhen.  Personalentscheidungen können falsch sein. Entweder, man lehnt einen passenden Bewerbenden ab (Fehler 1. Art, Alphafehler) oder man stellt einen unpassenden Bewerbenden ein (Fehler 2. Art, Betafehler). Aus statistischer Sicht ist es so, dass auch Personalentscheidungen genauer getroffen werden, also weniger Fehler begangen werden, je mehr Daten bzw. Datenpunkte verwendet werden können (Zöfel, 2003). Führt man neben den Rekrutierungsverfahren CV-Checks und Bewerbungsgesprächen zusätzlich noch einen psychologischen Test durch, können damit falsche Entscheidungen minimiert werden. Damit wird auch die inkrementelle Validität angesprochen (vgl. Kap. 2.2.1.3.3).

begünstigende Kategorien		Beispielzitate	Definitionen der Kategorien und Untermauerung durch Literatur	
Entscheidungskriterien	Qualität / Güte des Selektionsprozesses	Prüfung der Passung	<p><b>Person-Job-Fit</b></p> <p>(...) es geht darum, ein Abbild von Deiner Persönlichkeit zu erschaffen, um zu gucken, passt das zu dem, was wir suchen oder eben nicht. Und falls nicht, dann würde die Person in dem Beruf wahrscheinlich auch nicht glücklich werden, weil sie dann den Erwartungen gar nicht entsprechen kann. Int. 2, Pos. 27</p>	<p>Psychologische Tests werden eingesetzt, um die Eignung zwischen den Eigenschaften eines Bewerbenden und den Anforderungen der ausgeschriebenen Stelle zu klären.</p> <p>Ziel der Personalauswahl ist es, einen möglichst grossen Person-Job-Fit (P-J-Fit) zwischen Bewerbenden und Arbeitsstelle zu erreichen (Wilhelmy &amp; Kleinmann, 2016). Der P-J-Fit ist von besonderer Bedeutung, da dieser u.a. mit der Arbeitsleistung einer Person zusammenhängt: die Metaanalyse von Kristof-Brown, Zimmerman und Johnson (2005) konnte hierzu einen moderaten korrelativen Zusammenhang von <math>\rho = .20</math> feststellen, was die Wichtigkeit des P-J-Fit unterstreicht.</p>
			<p><b>Person-Group-Fit</b></p> <p>Weil letztendlich geht es doch darum, herauszufinden, passt die Person von ihren Eigenschaften in das Team. Wie ticken die Personen im Team, bzw. möchte man mehr Homogenität schaffen in den Teams oder Heterogenität? Also was ist schon vorhanden, was fehlt noch? Int. 2, Pos. 85</p>	<p>Dass durch psychologische Tests die Passung zwischen Bewerbenden und dem bestehenden Team festgestellt werden kann, begünstigt deren Einsatz.</p> <p>Die Metaanalyse von Kristof-Brown et al. (2005) verleiht auch dieser Kategorie Gewicht: die Metaanalyse zeigt eine moderate Korrelation zwischen dem Person-Group-Fit (P-G-Fit) und der Leistung des Mitarbeitenden auf (<math>\rho = .19</math>).</p>
			<p><b>cultural Fit</b></p> <p>Die Idee ist, die ich entwickle, ist, dass man die Werte des Unternehmens, das Konzept von value based, dass man das vom VP bis zum Lehrling überall miteinbezieht. Das Konzept also bei allen hat. Ich denke, es geht halt darum, dass man die Werte und was ist uns wichtig, der berühmte cultural fit, dass man das auf allen Stufen hat. Int. 1, Pos. 46</p>	<p>Der Einsatz psychologischer Tests ermöglicht die Prüfung der Passung zwischen Bewerbendem und der vorherrschenden Organisationskultur.</p> <p>Ziel der Personalauswahl ist es nicht nur, einen guten P-J-Fit, sondern auch einen Person-Organization-Fit (P-O-Fit) zu erreichen (Wilhelmy &amp; Kleinmann, 2016). Die Höhe des P-O-Fit korreliert nicht nur mit Leistung, sondern auch mit dem Verbleib im Unternehmen. Die Metaanalyse von Kristof-Brown et al. (2005) zeigt eine kleine Korrelation zwischen dem P-O-Fit und der beruflichen Leistung (<math>\rho = .07</math>) und zudem eine grosse, negative Korrelation zwischen dem P-O-Fit und der Kündigungsintention (<math>\rho = -.35</math>).</p>
	Wirtschaftlicher Nutzen		<p><b>Effizienz d. schwerpunktmässige Selektierung</b></p> <p>Man kann, wenn man das Testsetting gut wählt, auch nach Schwerpunkten selektieren. Int. 3, Pos. 21</p>	<p>Der Einsatz von psychologischen Tests zur Personalauswahl ermöglicht eine schwerpunktmässige Selektierung. Dadurch wird der Selektionsprozess effizient(er).</p> <p>Diese Kategorie wird historisch gestützt: zu Beginn des 20. Jahrhunderts, zu Zeiten des ersten Weltkrieges, konnten sich psychologische Tests zur Personalauswahl durchsetzen aufgrund ihrer Effizienz. Die USA setzten den Army Alpha- und den Army Beta-Test ein, um in kürzester Zeit geeignetes Militärpersonal auswählen zu können. Schnell haben psychologische Tests aufgrund dessen auch Einzug in die Privatwirtschaft und öffentlichen Verwaltungen gefunden (Schwarb, 1996).</p>
			<p><b>Return on Investment</b></p> <p>Jede Fehlbesetzung alleine spricht für eine Jahreslizenz. Int. 7, Pos. 103</p> <p>Ich sehe das aber immer als Return on Investment. Man müsste halt, wenn man z.B., keine Ahnung ein paar 100 Franken ausgeben müsste für ein Selektionsverfahren, aber dafür eine, ich weiss nicht, 50% höhere Chance hat das es ein guter Hire ist. Dann würde es sich es z. B. lohnen. Int. 8, Pos. 49</p>	<p>Der Nutzen, der durch den Einsatz psychologischer Tests entsteht, überwiegt die Kosten. Die Wahrnehmung eines «Return on Investment» fördert den Einsatz psychologischer Tests.</p> <p>Im Bereich von Personalauswahlverfahren dient das <i>Brodgen-Cronbach-Gleser-Modell</i> der Kosten-Nutzen-Analyse. Es ermöglicht die Berechnung des finanziellen Zusatznutzens bspw. durch den Einsatz psychologischer Tests. Diese Berechnung bzw. die Formel fusst auf folgendem Paradigma: im Vergleich zur Zufallsauswahl bringt der Einsatz von kriteriumsvaliden Personalauswahlverfahren wie psychologischen Tests immer einen finanziellen Zusatznutzen (Brodgen, 1949; zitiert nach Kauffeld &amp; Grohmann, 2011).</p>

hemmende Kategorien		Beispielzitate	Definition der Kategorien und Untermauerung durch Literatur	
persönliche Beweggründe	negative persönliche Erfahrung	Das habe ich völlig jenseitig gefunden, ehrlich gesagt. Und zwar, ein Teil dieser Tests hat immer auch so mit Dreidimensionalitäten zu tun. Formenerkennungen und Zeug und Sachen. (...) Das ist für mich der absolute Albtraum. (...) Das ist dann Stress. Das muss ich sagen. Sonst, analytisches Denken, das ist nicht das Problem gewesen. So die ganzen Formengeschichten, die habe ich mühsam gefunden. Int. 7, Pos. 45	Negative persönliche Erfahrungen begünstigen eine negative Einstellung ggü. psychologischen Tests, was gegen deren Einsatz im Unternehmen spricht.  Wie bereits bei der Kategorie <i>positive persönliche Erfahrung</i> dargelegt, wird die Verhaltensintention und damit das tatsächliche Verhalten durch persönliche Erfahrungen geprägt (Fishbein & Ajzen, 1975). Hier prägen also negative persönliche Erfahrungen die Einstellung, welche zur Verhaltensintention und schliesslich zum Verhalten führen, den Testeinsatz im Unternehmen abzulehnen.	
		Entscheidungskriterien	zu hoher Aufwand	hohe Lizenzkosten
Entscheidungskriterien	fehlender Nutzen	hohe Zeitintensität	Sie brauchen Zeit. In einem Bewerbungsprozess hast Du heutzutage keine Zeit mehr. Da musst Du zuschlagen. Int. 5, Pos. 31	Das Durchführen, Auswerten, Interpretieren und Besprechen der psychologischen Tests und deren Ergebnisse entspricht einem erheblichen Zeitaufwand, der zusätzlich zu anderen gängigen Auswahlverfahren anfällt. Neben den Lizenzkosten ist der zeitliche Aufwand der zweite ökonomische Faktor, den es ebenfalls zu betrachten gilt (Moosbrugger & Kelava, 2020; Berchtold, 2005) (vgl. Kap. 2.3).
		bisherige Prozesse ausreichend	Ich sage mal, wenn jemand eine Lehrbefähigung für das Gymnasium hat oder für BM, Maturitätsfächer. Dann kann man davon ausgehen, dass das passt. Das Ganze ist ja eh sehr formalisiert. Ich denke, an der Berufsfachschule, da müssen wir uns eh an die kantonalen Vorgaben halten bezüglich der Anstellung. Da ist es relativ klar. Int. 3, Pos. 56	Psychologische Tests werden nicht im Rahmen der Personalauswahl durchgeführt, weil das Prüfen formeller Qualifikationen, wie Ausbildungen und Berufserfahrung, als ausreichend erachtet werden. Die bisher eingesetzten Personalauswahlmethoden haben sich bewährt, wodurch kein zusätzlicher Nutzen durch den Einsatz zu erwarten ist.  Hossiep et al. (2015) fanden in ihrer Studie heraus, dass der Hauptgrund, der gegen den Einsatz von Persönlichkeitstests im Rahmen von Personalentscheidungen spricht, der fehlende (wahrgenommene) Nutzen ist (vgl. Kap. 2.3).
Entscheidungskriterien	Risiko, Bewerbende zu verlieren	ungewollte Selbstselektion	(...) (lacht) Dann finden wir ja gar keinen mehr! (...) Ja, wir haben es halt jetzt schon schwer, Fachpersonal zu finden. Int. 6, Pos. 88  Es kann sehr leicht dazu führen, dass Leute sich deselektieren vom Prozess, die vielleicht super wären für den Job, aber vielleicht Angst haben vor Tests. Int. 8, Pos. 19	Es besteht die Befürchtung, dass sich geeignete Bewerbende selbst aus dem Bewerbungsprozess selektieren, aufgrund ihrer negativen Einschätzung hinsichtlich ihrer eigenen Fähigkeiten und Fertigkeiten, den psychologischen Test bestehen zu können.  Selbstselektion wird zwar von Unternehmen gefördert, um unpassende Bewerbende vom Personalauswahlprozess ausschliessen zu können (Wilhelmy & Kleinmann, 2016). Doch Selbstselektion ist nicht gewollt, wenn sich geeignete Bewerbende ihrer Eignung nicht bewusst sind und deshalb auf eine Bewerbung verzichten bzw. das Bewerbungsverfahren deshalb abbrechen.
		Arbeitgebendenattraktivität leidet	Es kann sehr leicht dazu führen, dass Leute sich (...) denken, ein Arbeitgeber, der sowas durchführt, ist nicht besonders attraktiv und so weiter. Int. 8, Pos. 19	Es besteht die Angst, dass die Attraktivität der arbeitgebenden Unternehmen durch den Einsatz psychologischer Tests zum Zweck der Personalauswahl leidet.  Der <i>war for talents</i> verlangt, dass Personalauswahlprozesse für die Bewerbenden so angenehm wie möglich zu gestalten sind. Im Zusammenhang mit dem <i>war for talents</i> hat sich der Begriff Recruitment entwickelt, wobei es sich um ein Wortspiel aus Recruitment und Entertainment handelt. Personalauswahlprozesse sollen also einen gewissen Unterhaltungsfaktor für Bewerbende beinhalten, um die Arbeitgebendenattraktivität zu unterstützen (Fellner, 2019). Das unterstreicht die Relevanz der Bewerbendenakzeptanz. Mit den häufig geäusserten Bedenken um Bewerbendenakzeptanz bezüglich psychologischer Tests (vgl. Kap. 2.3), muss es als logische Konsequenz betrachtet werden, dass auf den Einsatz ebendieser verzichtet wird.
		fehlende Akzeptanz der Bewerbenden	Wir haben das mal eingesetzt. Bei (...) online-Suchen, die wir haben. Weisst Du, die, die über (...) Instagram oder Facebook ausgeschrieben werden. Den Leuten ist das in der jetzigen Situation, wo man es nicht so schwer hat, eine Stelle zu finden, ist das recht egal. Sie machen es nicht. Sie machen es nicht. Sie finden es doof. (...) Also, die Folge davon ist, wenn Du mehr als eine Frage stellst (...) uns hat man gesagt, vier kann man in etwa stellen so, wenn wir schon nur eine Frage stellen, bekommen wir keine Antwort, keine Bewerbung. Int. 5, Pos. 117	Fehlende Akzeptanz der Bewerbenden (im Sinne fehlender Offenheit und Bereitschaft) der Bewerbenden ggü. dem psychologischen Test wird befürchtet.  Nicht umsonst wurde die antizipierte Bewerbendenakzeptanz in unterschiedlichen Studien wiederholt untersucht. Arnoneit et al. (2020) konnten in ihrer in Deutschland durchgeführten Längsschnittstudie feststellen, dass die Einschätzung von Bewerbendenakzeptanz betreffend die psychologischen Tests in den letzten Jahrzehnten stetig gestiegen ist. HR-Mitarbeitende schätzen inzwischen die Bewerbendenakzeptanz neutral bis eher positiv ein (Arnoneit et al., 2020).

hemmende Kategorien		Beispielzitate	Definition der Kategorien und Untermauerung durch Literatur
Entscheidungskriterien	Mangelnde Güte	<b>zu generisch</b> Und sie sind halt sehr generisch, würde ich behaupten. Int. 4, Pos. 36  Weil, sie sind halt schon auch sehr allgemein. Und ja, da musst Du dann schon auch schauen, spezifisch, auf die Stelle, was braucht es da denn. Int. 4, Pos. 32	Der Umstand, dass psychologische Tests eher allgemeine als spezifische Abfragen tätigen, hemmt deren Einsatz. Der Test beschäftigt sich demnach zu wenig mit den Anforderungen, die eine bestimmte Arbeitstätigkeit bzw. -stelle mit sich bringt.  Ein generischer Test, der nicht spezifisch auf die zu messenden Konstrukte oder Fähigkeiten ausgerichtet ist, beeinträchtigt die Inhaltsvalidität (Moosbrugger & Kelava, 2020). Der fehlende berufliche Bezug zeigte sich auch in der Studie von Hossiep et al. (2015) als ausschlaggebenden Grund für den Verzicht auf Testeinsatz im Unternehmen (vgl. Kap. 2.3).
		<b>Scheingenauigkeit</b> Wichtig ist auch einfach. Es ist eine Momentaufnahme. Man weiss dann halt auch nie, ist die Person fit gewesen. Int. 4, Pos. 32  Wenn sie unsorgfältig vorgehen, wenn sie oberflächlich vorgehen, wenn sie so ein Testergebnis einfach so als bare Münze nehmen. Weil, so darf man nicht. Int. 7, Pos. 71	Obwohl das Ergebnis eines psychologischen Tests als absolute Zahl dargestellt wird, und damit eine gewisse Genauigkeit suggeriert, handelt es sich dabei lediglich um eine Scheingenauigkeit. Auch äussere Einflüsse, Messfehler etc. haben Einfluss auf das Ergebnis.  Eckardt (1973) verweist darauf, dass die Gefahr von Überschätzung der Testergebnisse hinsichtlich ihrer Aussage- und Prognosekraft besteht. Testergebnisse in numerischen Ausdrücken haben eine gewisse Suggestivwirkung, selbst wenn es sich um ausgewiesene Wahrscheinlichkeitsangaben handelt. Gerade bei psychologischer Diagnostik sei zu beachten, dass es sich um einen Momentaufnahme handelt, wodurch die Sicherheit einer Prognose auf zukünftige Leistung minimiert wird (Eckardt, 1973).
		<b>mangelnde Zielgruppen- angepasstheit</b> Zudem hast Du viel handwerkliche Jobs, wo auch zum Teil weniger qualifizierte Leute oder Leute mit Sprachbarriere arbeiten, wo man das halt auch gar nicht umsetzen kann. Int. 4, Pos. 30	Die Adressatengerechtigkeit von psychologischen Tests ist nicht für alle Anforderungsstufen gegeben. Es besteht die Gefahr, dass Fragen von gewissen Bewerbenden oder Mitarbeitenden aufgrund von Sprachschwierigkeiten nicht verstanden werden. Das hemmt den Testeinsatz.  Da durch den Test keine Testdurchführungsfairness erreicht werden könnte und der Nutzen dadurch eingeschränkt ist, wird auf Testeinsatz verzichtet (vgl. Kap. 2.2.2).

äussere, begünstigende und hemmende Faktoren	Beispielzitate	Definition der Kategorien und Untermauerung durch Literatur
<b>äussere Einflussgrössen</b>	<p>+: Sicher noch keine Entscheidungsgrundlage, aber so eine Basis, um darauf aufzubauen und natürlich auch, um eine erste Selektion zu machen. Weil, Du musst Dir vorstellen, was meinst Du, wie viele Bewerbungen pro Jahr reinkommen. Int. 4, Pos. 21</p> <p>-: Dann könnte man es auch ggü. den Kandidaten und Kandidatinnen verkaufen, dass es also zur Unterstützung im Auswahlprozess ist, auch für sie. Dann könnte ich mir das gut vorstellen. Aber ich muss noch sagen, wir haben eh zu kämpfen, dass wir überhaupt genug Bewerbende haben. Der Fachkräftemangel trifft auch uns. Int. 3, Pos. 52</p>	<p>Der Fachkräftemangel führt zu einem Mangel an Bewerbungen. Das Durchführen von psychologischen Tests zur Personalauswahl hat keinen Nutzen, da keine Auswahlmöglichkeiten bestehen. Eine hohe Anzahl an Bewerbungen begünstigt hingegen den Einsatz psychologischer Tests aufgrund der Möglichkeit, schwerpunktmässig zu selektieren.</p> <p>Bewerben sich viele Personen für dieselbe Stelle, kann davon ausgegangen werden, dass sich darunter Bewerbende befinden, die über eine hohe Eignung verfügen (Fellner, 2019). Die Bewerbendenakzeptanz kann hier als sekundär betrachtet werden. Hingegen verlangt eine niedrige Anzahl Bewerbungen auf eine Arbeitsstelle andere Überlegungen: fehlender Nutzen des Testeinsatzes und Befürchtungen betreffend die Bewerbendenakzeptanz sprechen gegen den Testeinsatz.</p>
	<p>+ : Wir sind auf jeden Fall dabei, unseren Prozess zu überdenken. Da freuen wir uns auch über die Ergebnisse aus der Empirie oder auch von anderen Unternehmen. Um zu wissen, was denn momentan auch State of the Art ist. Int. 2, Pos. 37</p> <p>- : Es ist halt wirklich schwierig, weil ich befürchte, dass man da möglicherweise auch ein Kompletionsproblem hat. Also der Arbeitsmarkt ist halt momentan ein Arbeitnehmermarkt und nicht ein anderer Markt, und ich kann mir vorstellen, dass das weniger attraktiv ist. Also deswegen weiss ich nicht, ob ich Tests einführen würde. Int. 8, Pos. 47</p>	<p>Wenn psychologische Tests eine gewisse Bekanntheit haben, begünstigt das ihren Einsatz. Trends und Bekanntheit beeinflussen demnach nicht nur den Testeinsatz an sich, sondern auch die Wahl des Tests. Hingegen werden Tests mit wenig Bekanntheit nicht bzw. kaum eingesetzt. Wenn Unternehmen derselben Branche auf Tests verzichten, tun ihnen das andere Unternehmen gleich.</p> <p>König et al. (2010) erklären das in ihrer Studie so, dass Unternehmen in komplexen Situationen andere Organisationen nachahmen und ähnliche strategische Entscheidungen treffen. Unternehmen entscheiden sich, orientiert an anderen Unternehmen, gegen Tests oder für (bestimmte) Tests (vgl. Kap. 2.3).</p>
	<p>- : Ich würde da noch gerne wissen, wie unsere Bereichsleiter reagieren würden, wenn ich sagen würde, wir machen zukünftig für alle ein Profil. Das würde mich noch interessieren. Wie das ankommen würde. (lacht). Int. 3, Pos. 54</p> <p>+ : Letztes Jahr hat es insgesamt X gegeben von den Acs, und das Jahr, da rede ich jetzt bis Q3, einfach dass Du siehst, die Popularität und die Akzeptanz, wie die am steigen ist, da haben wir schon, Ende September X gehabt. Es wird wohl auf die X zugehen dieses Jahr. Tendenz steigend. Int. 4, Pos. 82</p>	<p>Fehlende Akzeptanz der Linienvorgesetzten ggü. psychologischen Tests verhindert deren Einsatz. Hingegen begünstigt vorhandene Akzeptanz der Linienvorgesetzten den Einsatz psychologischer Tests im Rahmen von Personalentscheidungen.</p> <p>Auch Kanning (2015) verweist darauf, dass die direkten Linienvorgesetzten Einfluss auf den Testeinsatz nehmen. Insbesondere, wenn die Linienvorgesetzten eine dominante Rolle im Personalauswahlprozess einnehmen, können sie den Einsatz von psychologischen Tests verhindern oder durchgeführte Tests und deren Resultate zu wenig in die Entscheidung mitaufnehmen (Kanning, 2015). Hingegen können positive Meinungen der Linienvorgesetzten hinsichtlich des Testeinsatzes diesen auch fördern.</p>
	<p>- : Nicht nur wir möchten die Person kennenlernen, sondern die Person soll auch uns kennenlernen. Das soll wirklich ein Gespräch auf Augenhöhe sein und kein Verhör. So ein Test könnte halt schon den Eindruck erwecken, als dass man die Person so (.) hinters Licht führen wollte. Int. 2, Pos. 41</p> <p>+ : Dort haben wir in der Rekrutierung (.) aber da ist noch eine 19 vorndran. (...) So die Strukturiertheit, also jeden Kandidaten das genau Gleiche zu fragen. Das haben die technischen Grossunternehmen dazumal schon gemacht. Das ist noch spannend gewesen. Gerade im Nachhinein. Die waren dazumal den Banken schon so weit voraus in der Personalselektion. Die Industrie ist da sehr technisch auch. Int. 1, Pos. 141</p>	<p>Der Einsatz psychologischer Tests wird durch die Passung zur eigenen Unternehmens- oder Branchenkultur beeinflusst. Bspw. setzen technische und damit datenorientierte Unternehmen eher psychologische Tests ein, stützen sich somit bei Personalentscheidungen gerne auch auf Testdaten. In anderen Branchen ist der Einsatz psychologischer Tests so unüblich, dass man von fehlender Branchenangemessenheit sprechen muss.</p> <p>Ruhdorffer-Ritt (2013) verweist darauf, dass die Personalauswahl der Unternehmenskultur entsprechen muss, damit sich zukünftige Mitarbeitende mit dem Unternehmen auch identifizieren können. Als Bewerbende bzw. Bewerbender tritt man schliesslich durch die Personalauswahl mit dem Unternehmen ein erstes Mal in direkten Kontakt. Wie wichtig die Passung zwischen Organisation und Mitarbeitenden ist, wurde bereits unter der Kategorie zum P-O-Fit dargelegt.</p>

äussere Einflussgrössen	unternehmensinterne Ressourcen	Finanzen	<p>- : Ah, bei der Rekrutierung in dem Bereich macht es Sinn, eine Jahreslizenz zu nehmen oder nicht. Aber, ein KMU, die können sich das schlicht nicht leisten. Int. 4, Pos. 74</p> <p>+ : Und sie sind halt auch für uns rein von der Wirtschaftlichkeit, sei es Zeit, Geld und so. Wenn man bedenkt, dass man als Grossunternehmen, wie wir das sind, mit Lizenzen, also Jahreslizenzen arbeiten kannst, kann das schnell einmal interessant werden. Int. 4, Pos. 34</p>	<p>Der Einsatz psychologischer Tests ist mit zusätzlichem finanziellen Aufwand verbunden. Bestehen ausreichend finanzielle Ressourcen für die Personaldiagnostik, begünstigt das den Testeinsatz zu Personalentscheiden. Bestehen hingegen keine finanziellen Ressourcen, hemmt dies im Umkehrschluss den Einsatz psychologischer Tests im Unternehmen.</p> <p>Holtbrügge (2004) erklärt dies durch die Grösse des Unternehmens. Je grösser ein Unternehmen ist, desto eher verfügt es über finanzielle Ressourcen, die in die Personalauswahl investiert werden. Zwischen finanziellen Ressourcen und der Wahrscheinlichkeit zum Einsatz psychologischer Tests konnte durch Holtbrügge (2004) ein positiver Zusammenhang gefunden werden.</p>
		Expertise	<p>- : Einfach zu wenig KnowHow hat in Bezug auf das Testverfahren. Dann finde ich es (..) nicht nur, dass es ein Nachteil ist, dann finde ich auch, dass es fahrlässig ist. Int. 7, Pos. 71</p> <p>+ : Und da merkst Du auch, wie dankbar eigentlich das Umfeld ist, das direkt, indirekt mit uns zu tun hat, auch wissen, dass sie auf uns zurückgreifen können, auf unsere Expertise. Int. 4, Pos. 74</p>	<p>Fehlende interne Expertise zur Auswahl, Durchführung, Auswertung und / oder Interpretation der psychologischen Tests bzw. deren Ergebnisse führt dazu, dass keine psychologischen Tests eingesetzt werden. Hingegen ist vorhandene interne Expertise eine positive Einflussgrösse hinsichtlich des Einsatzes von Tests im Unternehmen.</p> <p>Einhergehend mit der eben genannten Kategorie der finanziellen Ressourcen muss auch die Kategorie der <i>internen Expertise</i> verstanden werden, die schliesslich von den finanzielle Ressourcen abhängig ist.</p>
Äussere Faktoren ohne gegenpoliges Pendant		Beispielzitate	Definition der Kategorien und Relevanz	
begünstigende äussere Einflussgrösse	gesetzliche Vorgabe	<p>Also, es ist gesetzlich geregelt, was muss so eine Person mitbringen. Sprich, sie geben die Mindestanforderungen an. Sie stellen die Mindestanforderungen, die erfüllt sein müssen von der Person. Durch das wissen wir ja, was wir messen müssen, sprich, operationalisieren. Int. 4, Pos. 48</p>	<p>Der Einsatz psychologischer Tests ist für manche berufliche Tätigkeiten gesetzlich vorgeschrieben. Bspw. müssen Mitarbeitende im Bereich des Personentransports oder Sicherheitskräfte gewisse Grenzwerte in Tests zu psychomotorischen, kognitiven oder klinisch-psychologischen Merkmalen erreichen.</p> <p>Für die Arbeit im Personentransport hat das Bundesamt für Verkehr (BAV, 2024) Richtlinien zur Tauglichkeit von Mitarbeitenden im Personentransport erlassen. Art. 10 des Bundesgesetzes über die Armee und die Militärverwaltung (SR 510.10) regelt den Inhalt der Rekrutierung für Militär- und Schutzdienst und schreibt Grenzen der medizinischen und psychologischen Tauglichkeit vor.</p>	
hemmende äussere Einflussgrössen	Status Quo	<p>Ich glaube, ich glaube nicht, dass das ein bewusster Entscheid ist. (..) Man hat es einfach nie gemacht. Int. 3, Pos. 42</p>	<p>Psychologische Tests werden nicht eingesetzt, weil das im Unternehmen schon immer so gehandhabt wurde. Demnach kann man nicht von einem bewussten Entscheid gegen den Einsatz von Tests sprechen.</p> <p>Diese Kategorie lässt sich mit dem <i>Default-Effekt</i> erklären, welcher insb. in der Marketingpsychologie von Interesse ist. Der Default-Effekt kann als Option verstanden werden, die durch eine Person unterhalten wird. Die Person entscheidet sich dabei aber nicht willentlich gegen eine andere Option, wodurch eigentlich keine wirkliche Entscheidung getroffen wird (Herrmann, Goldstein, Stadler, Landwehr, Heitmann, Hofstetter et al., 2011). Man verharrt also im Status Quo, weil andere Optionen – in diesem Falle also der Testeinsatz – nicht in Betracht gezogen werden.</p>	
	rechtliche Restriktionen	<p>Also, es gibt in manchen Ländern auch einfach rechtliche Restriktionen, solche metrischen Tests zu machen mit Kandidaten. Deswegen, da, da muss man sich genau anschauen, welchen dann. Und auch dann, wie man mit den Daten umgeht und so weiter. Es ist einfach erstmal eine rechtliche Frage. Int. 8, Pos. 19</p>	<p>In manchen Ländern gibt es rechtliche Restriktionen, die den Einsatz psychometrischer Tests einschränken. Schliesslich spielt die Ethik eine übergeordnete Rolle für das psychologische Testen.</p> <p>In der Schweiz ist dies durch das Humanforschungsgesetz (HFG) geregelt, welches auf Art. 118b1 der schweizerischen Bundesverfassung (BV) (HFG, 2023) fusst. Die BV und das HFG (SR 810.30, 2023) nennen hier als primäres Ziel den «Schutz des Menschen, seiner Würde, Persönlichkeit und Gesundheit». König et al. (2010) untersuchten in ihrer Studie den Einfluss wahrgenommener Gesetzmässigkeit auf den tatsächlichen Einsatz psychologischer Tests, erhielten aber ein vernachlässigbares Ergebnis. Sie erklären das dadurch, dass der rechtliche Druck in der Schweiz hinsichtlich psychologischem Testen im Vergleich zu Kanada oder den USA vergleichsweise klein ist. Demnach schränkt das HFG den Einsatz psychologischer Test im Rahmen von Personalentscheidungen nicht ein.</p>	

## 3.3 Quantitative Datenerhebung

Um die Erkenntnisse aus Literaturrecherche und qualitativer Datenerhebung bzw. -analyse auf Allgemeingültigkeit überprüfen zu können, wird ein umfassender Fragebogen eingesetzt.

Der Fragebogen entspricht dem Kernstück der Datenerhebung. Die quantitativen Daten sollen schliesslich die Fragestellungen dieser Masterthesis beantworten.

Mit besonderer Sorgfalt wurde der Fragebogen entwickelt. Es wurden zahlreiche Überlegungen und Abwägungen vorgenommen, um den Gütekriterien, die auch schon in Kap. 2 besprochen wurden, gerecht zu werden. Die nachfolgenden Unterkapitel widmen sich der quantitativen Datenerhebung, beginnend mit der Fragebogenentwicklung in Kap. 3.3.1. Danach folgen die Schätzung der notwendigen Stichprobengrösse in Kap. 3.3.2 und die Durchführung der Umfrage in Kap. 3.3.3. Die Auswertung und die Ergebnisdarstellung folgen in Kap. 4.

### 3.3.1 Fragebogenentwicklung

Die Fragebogenkonstruktion lässt sich laut Bühner (2011) auf übergeordneter Ebene in drei Teilschritte untergliedern: 1. Erstellung des Fragebogenentwurfs, 2. empirische Überprüfung des Entwurfs und 3. Erstellung der endgültigen Fragebogenversion. Die nachfolgenden Unterkapitel sind entlang dieser Teilschritte gegliedert.

Das Ziel der Fragebogenstudie liegt in der Erhebung von subjektiven und objektiven Eigenschaftsausprägungen (nämlich Einstellungen und Verhalten). Die Methode zur Testkonstruktion ist als deduktiv zu bezeichnen. Das ausgearbeitete Kategoriensystem (siehe Kap. 3.2.5), welches durch qualitative Datenerhebung und -auswertung erhalten wurde, dient hier als Grundlage. Zur Erstellung der quantitativen Fragen wurde der erfahrungsgeleitet-intuitive Ansatz verfolgt: die Erfahrungen und Perspektiven der befragten Expertinnen und Experten wurden entsprechend in den Fragebogen integriert. Nach Bühner (2011) begünstigt dieser Top-Down-Ansatz die Inhaltsvalidität des Fragebogens ungemein, denn dadurch kann man davon ausgehen, dass die entwickelten Fragen praktische Relevanz besitzen.

#### 3.3.1.1 Erstellung des Fragebogenentwurfs

Der Fragebogenentwurf bzw. die endgültige Version des Fragebogens sehen vor, zuerst die unabhängigen Variablen betreffend die Beweggründe und Entscheidungskriterien abzufragen. Diese entsprechen subjektiven Eigenschaftsausprägungen. Anschliessend folgt die Abfrage der abhängigen Variable – ob im Unternehmen Tests eingesetzt werden oder nicht. Diese Frage fungiert als Filterfrage. Je nachdem, ob dieser Frage zugestimmt oder ob diese abgelehnt wird, folgen andere Fragen. Wird diese Frage zum Testeinsatz bejaht, folgen spezifische Abfragen bezüglich des Testeinsatzes im Unternehmen. Ebenfalls werden positive äussere Einflussfaktoren abgefragt, welche objektiven Eigenschaftsausprägungen entsprechen.

Wird die Frage zum Testeinsatz verneint, folgt unmittelbar die Abfrage hinsichtlich negativer äusserer Einflussfaktoren, welche als Pendant zu den positiven äusseren Einflussfaktoren formuliert wurden. Zum Abschluss des Fragebogens werden alle Teilnehmende zu demografischen Daten über sich und über das Unternehmen, in dem sie tätig sind, befragt.

Die nachfolgende Abb. 14 zeigt den geplanten Aufbau des Fragebogens und visualisiert auch die alternativen Verläufe in Abhängigkeit der Beantwortung der Filterfrage.

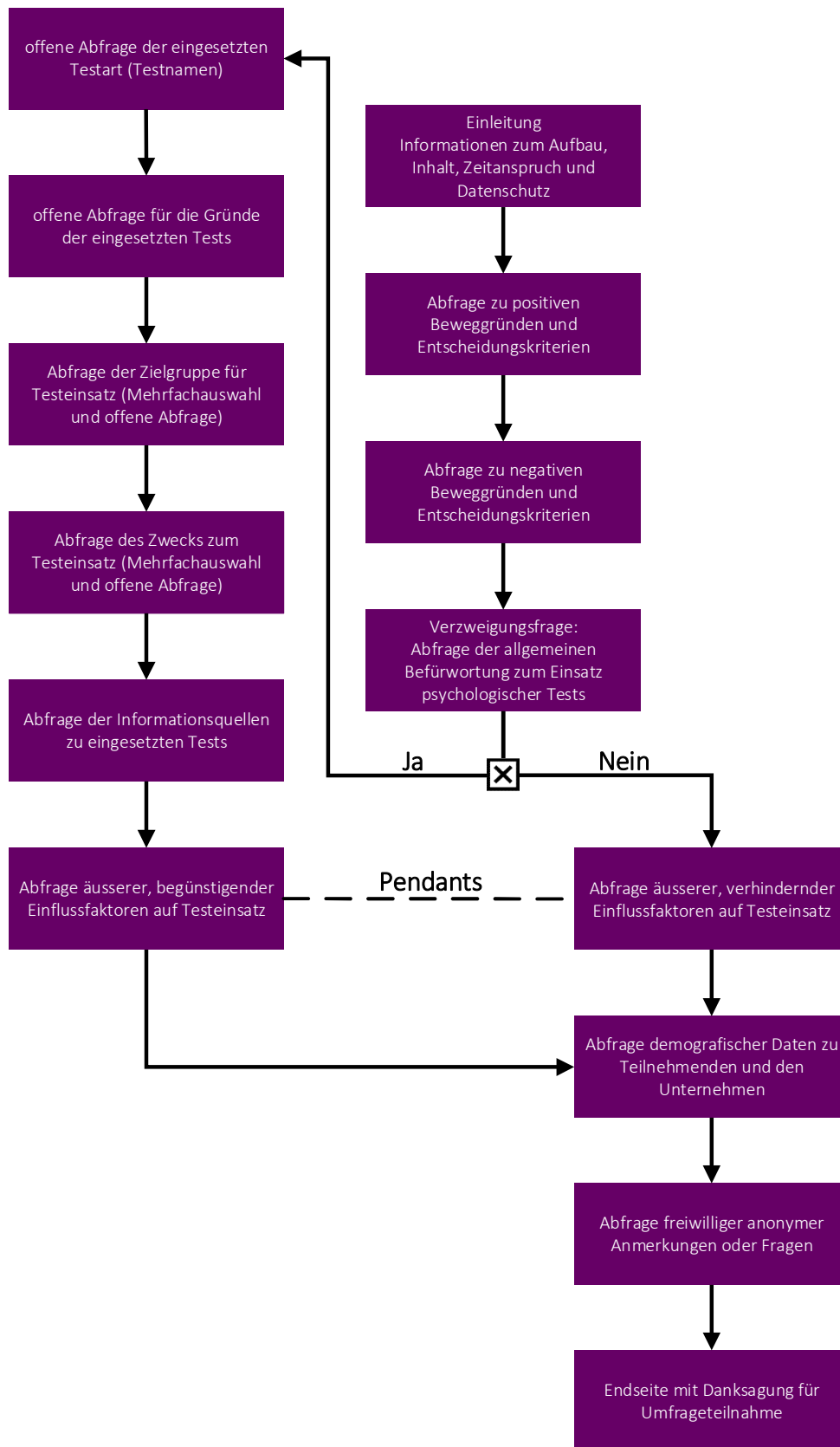


Abbildung 14. Geplanter Fragebogenaufbau mit Visualisierung der alternativen Verläufe in Abhängigkeit der Beantwortung der Filterfrage.

Der Fragebogen enthält aufgrund der verschiedenen Frageformate auch verschiedene Skalenformate, welche unterschiedliche Zwecke verfolgen. Die nachfolgenden Unterkapitel berichten von der Entwicklung der verschiedenen Frageblöcke. Die Reihenfolge richtet sich dabei entlang des gerade eben beschriebenen Fragebogaufbaus (vgl. Abb. 14).

### 3.3.1.1.1 Entwicklung der Fragen zu den unabhängigen, subjektiven Variablen

Der Grossteil der Fragebogenitems entspricht unabhängigen Variablen. Diese wurden durch geschlossene Fragen in Form von Aussagen operationalisiert. Mittels der Likert-Antwortskala können die Teilnehmenden bewerten, inwieweit die Aussagen für sie zutreffen.

Der Formulierung dieser Fragebogenitems, und insb. der Formulierung der Antwortskala, wurden bei der Entwicklung grosse Beachtung geschenkt.

Die befragten subjektiven Eigenschaftsausprägungen – persönliche Beweggründe und Entscheidungskriterien – werden in der Ich-Form formuliert. Um die Lesbarkeit zu erhöhen bzw. die Bearbeitungszeit des Fragebogens zu verkürzen, beginnen die Aussagen jeweils in der gleichen Formulierung, während das Satzende variiert. In der nachfolgenden Tab. 8 sind dazu Beispiele aufgeführt.

Tabelle 8

Beispiele für die Abfrage der Beweggründe und Entscheidungskriterien mit den entsprechenden Antwortskalen und der Einordnung des Skalenniveaus. Die kursiv gedruckten Satzanfänge sind in den jeweiligen Frageblöcken identisch.

Abfrage der subjektiven Eigenschaftsausprägungen hinsichtlich der Beweggründe und Entscheidungskriterien		
Variablengruppe	Beispiel-Item	Antwortskala
fürsprechende Beweggründe (subjektive Eigenschaftsausprägung)	<i>Den Einsatz von psychologischen Tests im Rahmen von Personalentscheidungen empfinde ich als sinnvoll, weil ich selbst positive Erfahrungen damit gemacht habe.</i>	Likertskalen, wobei die beiden Pole mit verbalen Anker versehen wurden: <i>trifft überhaupt nicht zu</i> und <i>trifft voll und ganz zu</i> .
gegensprechende Entscheidungskriterien (subjektive Eigenschaftsausprägung)	<i>Den Einsatz von psychologischen Tests im Rahmen von Personalentscheidungen empfinde ich als problematisch, weil die Kosten dafür zu hoch sind.</i>	

Das Ziel der vorliegenden Arbeit liegt in der Exploration möglicher Einflussgrössen, die den Einsatz psychologischer Tests begünstigen oder hemmen. Die durch die qualitative Datenerhebung gesammelten und durch die Literatur ergänzten Einflussgrössen könnten trotz aller Aufwände unvollständig sein. Daher werden den Teilnehmenden nach der Abfrage der geschlossenen Fragen offene Textfelder angeboten, in welchen sie weitere Vor- und Nachteile nennen können, die nicht im Fragebogen abgefragt wurden. Das soll der Inhaltsvalidität der Erhebung dienen.

#### 3.3.1.1.1.1 Entwicklung der Antwortskalen

«Die optimale Antwortskala gibt es nicht.» so Porst (2014, S.94). So entspricht die Festlegung des Antwortformates dem Ergebnis des Abwägens von Vor- und Nachteilen betreffend die verschiedenen Skalenbreiten und -beschriftungen. Die diesbezüglich getroffenen Entscheidungen werden nachfolgend dargelegt:

##### Geradzahlige Antwortskala

Um der Antworttendenz zu mittleren Urteilen entgegenzuwirken, wurde eine geradzahlige Antwortskala gewählt. Ein weiterer Vorteil, der sich laut Bühner (2011) dadurch ergeben soll, ist, dass durch das Weglassen der Mittelkategorie auch die unterschiedliche Auffassung ebendieser durch die Teilnehmenden ausgeschlossen werden kann. Die Mittelkategorie wird nicht immer als mittlere Ausprägung verstanden, sondern oftmals auch als neutral und nicht mit den beiden Endpolen der Antwortskala in Beziehung gesetzt, so Bühner (2011).

### **Sechsstufige Antwortskala (Differenzierungsgrad)**

Die Übersichtsarbeit von Eichhorn (2019) kann keine generellen Unterschiede zwischen vier- und sechsstufigen Antwortskalen aufdecken. Dennoch kann man sich darüber streiten, ob die Erwartung an das Abstraktionsvermögen der Teilnehmenden bei einer sechsstufigen Antwortskala nicht etwas zu hoch ist. Die sechsstufige Ratingskala soll aber den Erhalt möglichst differenzierter Information betreffend die Ausprägung der Zustimmung bzw. Ablehnung ermöglichen. Bühner (2011) empfiehlt eine Antwortskalenbreite von mindestens fünf Stufen, um die Reliabilität zu maximieren. Porst (2014) verweist darauf, dass mehr als neun Antwortskalenpunkte zu intellektueller Überforderung der Teilnehmenden und somit zu Scheinpräzision aufgrund unzureichender Differenzierung zwischen den Antwortpunkten führten. Die Wahl einer sechsstufigen Antwortskala geht demnach mit den Empfehlungen von Bühner (2011) und Porst (2014) einher.

Ob vier- oder sechs Stufen: es kann nicht verhindert werden, dass die Abstufungen der Ratingskala durch die Teilnehmenden eventuell unterschiedlich aufgefasst werden. Da es sich aber bei den Fragen um die Aufforderung subjektiver Bewertungen hinsichtlich der Fragen handelt, kann dieser Umstand wohl vernachlässigt werden.

### **Beschriftung / Verbalisierung der Skalenpunkte**

Die Antwortskalen werden an den beiden Endpolen mit verbalen Ankern versehen. Der linke Endpol wurde mit *trifft überhaupt nicht zu* beschriftet, der rechte Endpol mit *trifft voll und ganz zu*. Bühner (2011) empfiehlt zwar die Benennung aller Antwortalternativen. Zwei wesentliche Gründe führten dazu, dass dieser Empfehlung nicht gefolgt wurde:

1. Durch die Benennung aller sechs Stufen besteht die Gefahr, dass dadurch die Abstände zwischen den Antwortalternativen nicht mehr als gleich gross empfunden werden. Das hätte zur Folge, dass die Daten als ordinalskaliert und nicht intervallskaliert behandelt werden müssten (Zöfel, 2003). Die Übersichtsarbeit von Eichhorn (2019) bespricht die Beeinflussung des Antwortverhaltens durch die verbale Beschriftung der Antwortalternativen, ob es also ratsam ist, nur die Endpole oder eben alle Stufen zu beschriften. Im Grunde genommen lässt Eichhorn (2019) aber offen, inwiefern eine Beeinflussung angenommen bzw. befürchtet werden muss.
2. Die Darstellung im online Fragebogentool *Tivian* mit den Beschriftungen aller sechs Antwortalternativen hat nicht überzeugt. Die Darstellung zeigte sich als chaotisch, denn die Beschriftungen der einzelnen Skalenpunkte verliefen ineinander. Eine entsprechende Anpassung der Schriftgrösse wäre der Lesbarkeit und damit der User Experience wiederum abträglich gewesen.

#### **3.3.1.1.1.2 Reihenfolge der Fragen**

«Bad is stronger than good.», so der Leitsatz des Reviews von Baumeister, Bratslavsky, Finkenauer & Vohs, S.1). Damit sprechen sie auf Priming-Effekte an, wobei Negatives im Allgemeinen eine stärkere kognitive (oder emotionale) Beeinflussung bewirkt als Positives. Im Wissen um die Gefahr, dass die Abfrage negativer Indikatoren das anschliessende Antwortverhalten primen könnte, wurde die Reihenfolge der Frageblöcke entsprechend aufgegleist: zuerst werden positive subjektive Eigenschaftsausprägungen befragt, daran anschliessend die Negativen.

Um in den beiden genannten Frageblöcken die Gefahr von Reihenfolgeeffekten vorzugreifen, werden die Fragen randomisiert.

### 3.3.1.1.2 Entwicklung der Fragen zur abhängigen Variablen und Folgefragen

Die Frage, ob im Unternehmen psychologische Tests eingesetzt werden, entspricht der Abfrage der abhängigen Variablen. Die Antwort auf diese Frage bestimmt auch den weiteren Verlauf der Umfrage. Sie fungiert also als Filterfrage. Eine positive Antwort führt zur spezifischen Abfrage betreffend die eingesetzten Tests. Eine negative Antwort hingegen führt im Fragebogenverlauf an diesen spezifischen Nachfragen vorbei (vgl. Abb. 14).

Die Formulierungen der spezifischen Abfragen wurden in den meisten Fällen aus der qualitativen Datenerhebung übernommen. Diese haben sich bereits dort in Bezug auf die Formulierung als zielführend erwiesen.

Die Abfrage der abhängigen Variablen wurde nicht randomisiert, denn schliesslich folgen sie einer Erzähllogik, welche den Testeinsatz zuerst oberflächlich erfragt und anschliessend in die Tiefe geht: welche Tests werden eingesetzt, Gründe für die Auswahl, für wen, für welchen Zweck?

In Anhang E können die entwickelten Items zur Abfrage der abhängigen Variablen und der spezifischen Folgefragen eingesehen werden.

### 3.3.1.1.3 Entwicklung der Fragen zu den unabhängigen, objektiven Variablen

Während die subjektiven Eigenschaftsausprägungen – persönliche Beweggründe und Entscheidungskriterien – in der Ich-Form formuliert wurden, werden die objektiven Eigenschaftsausprägungen – äussere Einflussfaktoren – in der Wir-Form verfasst. Das soll den Teilnehmenden suggerieren, dass hier nicht mehr die eigene Meinung oder Einstellung befragt wird, sondern das tatsächlich stattfindende Prozedere in dem Unternehmen, in welchem sie tätig sind. Identisch zur Abfrage der unabhängigen subjektiven Variablen, beginnen auch bei diesem Frageblock die Aussagen in derselben Formulierung, während das Satzende variiert.

Die Abfrage der äusseren Einflussfaktoren ist abhängig von der Filterfrage. Daher wurden parallele Formulierungen geschaffen, wie auch das Beispiel in der nachfolgenden Tab. 9 zeigt.

Tabelle 9

Beispiele für die Abfrage der äusseren Einflussfaktoren mit den entsprechenden Antwortskalen und der Einordnung des Skalenniveaus. Die kursiv gedruckten Satzanfänge sind in den jeweiligen Frageblöcken identisch.

Abfrage der objektiven Eigenschaftsausprägungen hinsichtlich der äusseren Einflussfaktoren		
Variablengruppe	Beispiel-Item	Antwortskala
Wenn Testeinsatz bejaht wird: begünstigende äussere Faktoren (objektive Eigenschaftsausprägungen)	<i>In meinem Unternehmen werden psychologische Tests im Rahmen von Personalentscheidungen eingesetzt, weil sie unseren Selektionsprozess ineffizient machen würden.</i>	Likertskalen, wobei die beiden Pole mit verbalen Ankern versehen wurden: <i>trifft überhaupt nicht zu</i> und <i>trifft voll und ganz zu</i> .
Wenn Testeinsatz verneint wird: hemmende äussere Faktoren (objektive Eigenschaftsausprägungen)	<i>In meinem Unternehmen werden psychologische Tests im Rahmen von Personalentscheidungen eingesetzt, weil sie unseren Selektionsprozess effizient machen.</i>	

Identisch zur Abfrage der subjektiven unabhängigen Variablen werden auch diese Fragen randomisiert (vgl. Kap. 3.3.1.1.1.2), um Reihenfolgeeffekten vorzubeugen. Auch hier wird den Teilnehmenden nach der Abfrage der geschlossenen Fragen ein offenes Textfeld angeboten, in welchem sie weitere äussere Einflussfaktoren nennen können, die nicht abgefragt wurden.

#### **3.3.1.1.4 Entwicklung der Fragen zu den demografischen Daten**

Die Abfrage der demografischen Daten dient in erster Linie dazu, Aussagen über die Stichprobe machen zu können. Es werden Fragen zu den Teilnehmenden und zu deren Unternehmen gestellt. Es folgen Fragen zu Berufserfahrung, Ausbildung, Unternehmensgrösse, Wirtschaftsabteilung und ob ein Psychologe oder eine Psychologin im HR tätig ist. Bei diesem Frageblock speziell: die Antwortskalen sind an der Literatur orientiert. In Anhang F können die Items, die der Abfrage der demografischen Daten dienen, eingesehen werden. Ebenfalls sind die Quellen, an welchen sich die Antwortalternativen orientieren, dargelegt.

#### **3.3.1.1.5 Fertigstellung des Fragebogenentwurfs**

Der Fragebogenentwurf wurde der Begleitperson zur Durchsicht abgegeben und in einem Meeting diskutiert. Insbesondere die Verständlichkeit und die Formulierung der Fragen wurden kritisch beäugt, und gemeinsam redundante Items identifiziert. Die entsprechend angepassten Fragen wurden dann ins online Fragebogentool *Tivian* eingepflegt, um mit dem nächsten Schritt, der empirischen Überprüfung, weiterfahren zu können.

#### **3.3.1.2 Empirische Überprüfung des Entwurfs mittels Pretest**

Die Qualität der zu erhebenden Daten steht und fällt mit der Qualität des entwickelten Fragebogens. Die Durchführung eines Pretests zur Evaluation des entwickelten Fragebogens ist daher unerlässlich. Dazu wurden drei HR-Mitarbeitende gebeten, den Fragebogen zweimal, im Abstand von je ca. einer Woche, auszufüllen. Durch den Pretest sollen verschiedene Ziele verfolgt werden:

- **Ziel 1: Abklärung von Inhaltsvalidität und Konstruktvalidität**  
Die Fragen durch unvoreingenommene Personen auf Verständlichkeit, Relevanz, Vollständigkeit und Formulierung bewerten zu lassen.
- **Ziel 2: User Experience**  
Durch die drei Testpersonen das Layout und die Benutzungsfreundlichkeit beurteilen, sowie die genannte Bearbeitungszeit von 15min bestätigen zu lassen.
- **Ziel 3: Reliabilität schätzen**  
Retest-Reliabilität der drei Testsets berechnen und bewerten.

Um diese Ziele erreichen zu können, diene eine abgewandelte Form des Standard-Pretests, welcher sich an *den GESIS Survey Guidelines* von Lenzer, Neuert und Otto (2015) orientiert: Ein Standard-Pretest wird unter realistischen Bedingungen durchgeführt, um so die anstehende Haupterhebung simulieren zu können. So kann der gesamte Fragebogenablauf auch auf technische Mängel hin überprüft werden. Normalerweise handelt es sich bei einem Standard-Pretest um ein passives Vorgehen, in welchem die Fragebogenentwicklerinnen und -entwickler lediglich eine beobachtende Rolle einnehmen (Lenzer et al., 2015). Daher werden nach der simulativen Durchführung der Umfrage in *Tivian* auch Fragen zu Verständlichkeit und Formulierung der Items und zur User Experience eingepflegt. Diese Fragen können in Anhang G eingesehen werden.

Entgegen der Vorgabe des Standard-Pretests nach Lenzer et al. (2015), wird der Fragebogen durch die drei Teilnehmenden zweimal im Abstand von ca. einer Woche ausgefüllt. Dadurch wird erst die Berechnung der Retest-Reliabilität ermöglicht.

Die Auswertung der Pretests entlang der drei formulierten Ziele erfolgt in den Unterkapiteln 3.3.1.2.1 bis 3.3.1.2.3. Die entsprechenden Rohdaten befinden sich in Anhang H.

### 3.3.1.2.1 Auswertung des Pretests hinsichtlich Inhaltsvalidität

Die Testpersonen empfanden die Fragen des Fragebogenentwurfs als verständlich, relevant und vollständig, wie dies auch die nachfolgende Abb. 15 zeigt.

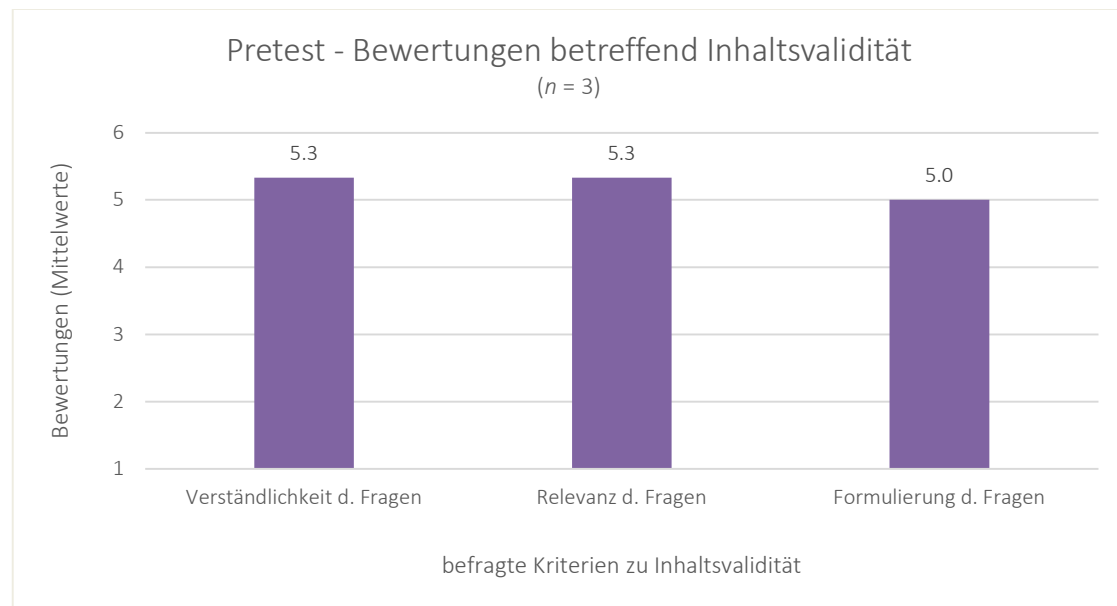


Abbildung 15. Auswertung Pretest zu Inhaltsvalidität.

Die Bewertungen liegen pro befragtes Kriterium durchschnittlich bei  $M \geq 5.0$ . Die Standardabweichung beträgt bei allen drei Kriterien  $SD = 0.7$ , was auf eine geringe Streuung hinweist.

Bei der offenen Antwortmöglichkeit haben die Pretest-Teilnehmenden folgende Punkte als verbesserungswürdig angemerkt:

1. Begriffe, wie *Einarbeitungsbedarfe* oder *Arbeitgebendenattraktivität* seien schwer lesbar.
2. Die Fragen seien schwer zu beantworten, wenn man sich mit der Thematik um psychologische Tests bisher noch nicht auseinandergesetzt habe.

### 3.3.1.2.2 Auswertung des Pretests hinsichtlich User Experience

Die User Experience wurde ebenfalls positiv bewertet, wie dies die nachfolgende Abb. 16 zeigt.

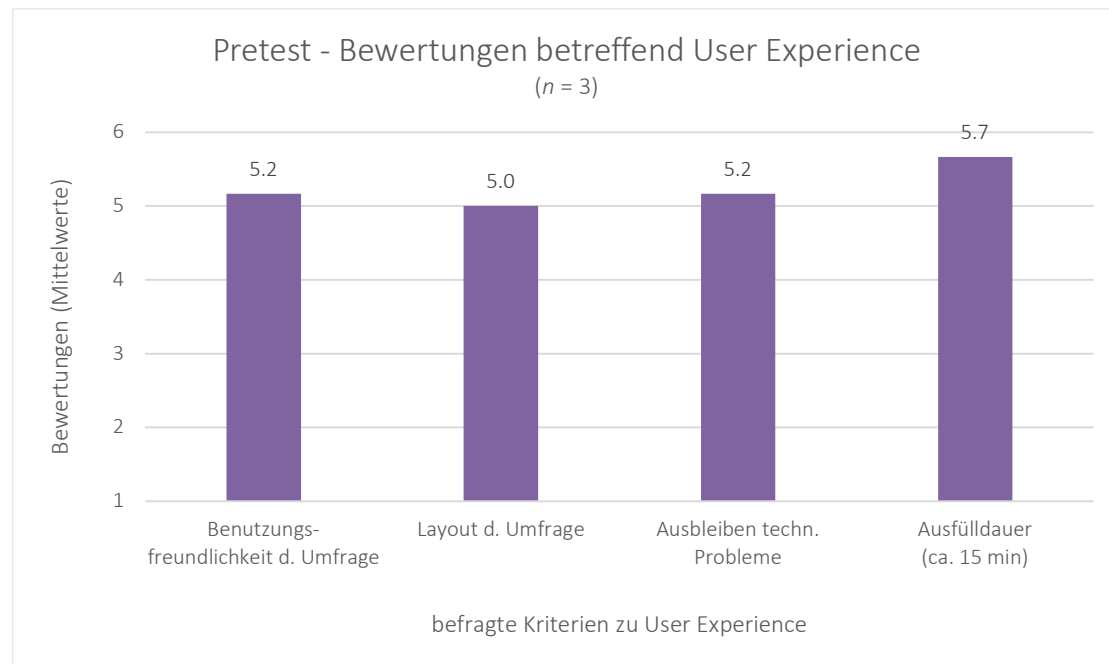


Abbildung 16. Auswertung Pretest zu User Experience.

Die Bewertungen liegen auch hier pro befragtes Kriterium durchschnittlich bei  $M \geq 5.0$ . Die Standardabweichung beträgt bei den Kriterien *Benutzungsfreundlichkeit* und *Layout*  $SD = 1.0$ , bei den Kriterien *Ausbleiben technischer Probleme* und *Ausfülldauer* bei  $SD = 1.2$ . Diese Werte weisen ebenfalls auf eine tiefe Varianz und damit auf Meinungsgleichheit hin.

Im offenen Antwortfeld haben die Pretest-Teilnehmenden drei weitere verbesserungswürdige Punkte genannt:

3. Die Schriftgrösse sei beim Ausfüllen der Umfrage am Computer zu klein.
4. Die Einleitung soll den Prozess der Umfrageentwicklung kurz aufgreifen.
5. Technische Probleme seien beim Ausfüllen auf dem Handy entstanden: das Layout und die Formatierung seien hier verzerrt.

### 3.3.1.2.3 Auswertung des Pretests hinsichtlich Retest-Reliabilität

Die Reliabilitätsprüfung der ordinalen bzw. quasi-intervallskalierten Fragen des entwickelten Fragebogens wurde mittels Retest-Reliabilität bzw. Intrarater-Reliabilität durchgeführt. So sollte die *Stabilität* der zu testenden Items herausgestellt werden. Drei HR-Fachpersonen sind der Bitte gefolgt, die Umfrage zweimal, im Abstand von ungefähr einer Woche, auszufüllen. Die erhaltenen Werte aus den beiden Durchgängen wurden pro Pretest-Teilnehmendem korreliert. Die Voraussetzungen der Daten – zwei ordinalskalierte Datensätze, ohne Prüfung auf Normalverteilung – erlaubten die Berechnung einer Rangkorrelation. Khamis (2008) empfiehlt aufgrund der kleinen Samplegrösse ( $n = 3$ ), Walther (2022) aufgrund der gleichen Anzahl an möglichen Ausprägungen pro Variable, die Berechnung des Korrelationskoeffizienten *Kendall*  $\tau_b$ .

Die Berechnung wurde mittels SPSS durchgeführt, um festzustellen, ob ein Zusammenhang zwischen den Datensätzen des ersten und des zweiten Durchgangs pro Pretest-Teilnehmendem besteht (siehe Anhang H). Zwischen den beiden Durchgängen liegt mit einem durchschnittlichen *Kendall*  $\tau_b = .561$  ( $p < 0.001$ ) – interpretiert nach Cohen (1988) – ein positiver starker Zusammenhang vor.<sup>11</sup> Demnach kann davon ausgegangen werden, dass der entwickelte Fragebogen reliabel ist.

### 3.3.1.3 Erstellung der endgültigen Testversion

Laut Pretest darf von Reliabilität, Inhaltsvalidität und User Experience ausgegangen werden. Einige verbesserungswürdige Punkte wurden durch die Pretest-Teilnehmenden geäußert, welche es durch geeignete Massnahmen zu beheben galt. Nachfolgend werden diese Massnahmen besprochen. Daran anschliessend wird das Finalisieren der Umfrage hinsichtlich der Übersetzung erläutert.

#### 3.3.1.3.1 Massnahmen betreffend Ziele 1 und 2

Die Punkte 1, 2 und 5 wurden gleich nach dem Pretest bearbeitet: die Schriftgrösse wurde erhöht und die Einleitung ergänzt mit einem kurzen Abriss der Umfrageentwicklung. Zudem wurde die Ausrichtung der Antwortskalen bei Handynutzung geändert – von horizontal zu vertikal. Der Anblick ist zwar etwas gewöhnungsbedürftig, die Antwortskala aber dadurch funktional.

Die Punkte 3 und 4, von derselben Testperson genannt, wurden hinsichtlich Verbesserungsmöglichkeiten mit ebendieser Testperson durchgesprochen. Verschiedene Formulierungsoptionen wurden diskutiert, welche aber der Lesbarkeit und Verständlichkeit abträglich gewesen wären. Die Formulierungen wurden daher beibehalten, was schliesslich auch im Sinne des betreffenden Pretesters war.

#### 3.3.1.3.2 Übersetzung und Einpflegen der Sprachversionen

Gerade am Arbeitsplatz wird in der Schweiz neben Deutsch, Französisch und Italienisch auch Englisch gesprochen (BFS, 2023a). Die Übersetzung des Fragebogens ins Französische und Englische wurden dem Übersetzungsbüro *SemioticTransfer AG*<sup>12</sup> anvertraut. Die Übersetzungen können in Anhang I eingesehen werden.

Anmerkung: In Absprache mit der Betreuungsperson wurde keine Übersetzung ins Italienische vorgenommen. Der Anteil an möglichen Fragebogenteilnehmenden aus dem italienischsprachigen Teil der Schweiz wird als vernachlässigbar eingeschätzt. Der Übersetzungsaufwand kann durch den möglichen Nutzen nicht gerechtfertigt werden.

Die Umfrage wurde anschliessend ins Fragebogentool *Tivian* eingepflegt. Der darin enthaltene Spracheneditor ermöglicht die Umfragedurchführung mit drei verschiedenen Sprachsets.

---

<sup>11</sup> Einteilung der Korrelationskoeffizienten nach Cohen (1988) – gilt für lineare und nicht-parametrische Korrelationskoeffizienten:  
 $r > .10$  = kleine Effektstärke |  $r > .30$  = mittlere Effektstärke |  $r > 0.50$  = grosse Effektstärke

<sup>12</sup> Die SemioticTransfer AG überzeugt mit ihrer ISO17100 Zertifizierung, welche qualitativ hochstehende und vertrauliche Übersetzungen garantiert.

### 3.3.2 Schätzung der Stichprobengrösse

Zumeist verfügen Kleinst- und Kleinunternehmen in der Schweiz nicht über eine eigene HR-Abteilung oder Mitarbeitende, die eigens für das Personalmanagement zuständig sind. Selbst bei mittleren Unternehmen mit 50 bis 99 Mitarbeitenden verfügen kaum 60% über eine eigene Personalabteilung (Mücke & Pekruhl, 2006). 80% der in der Schweiz tätigen mittleren Unternehmen mit 100 bis 249 Mitarbeitenden verfügen laut der Studie von Mücke und Pekruhl (2006) über eine eigene Personalabteilung. Daher wird die Schätzung der theoretischen Stichprobengrösse an der Summe der Anzahl mittlerer und grosser Unternehmen in der Schweiz im Jahr 2021 festgemacht (siehe Tab. 10), wobei eine Ad Hoc-Stichprobe<sup>13</sup> angestrebt wird.

Tabelle 10

Schätzung der theoretischen Stichprobe der quantitativen Datenerhebung entlang der Anzahl von mittleren und grossen Unternehmen in der Schweiz im Jahr 2021, orientiert am BFS (2023c).

	KMU-Betriebe			Gross- unternehmen	Total
	Kleinst- unternehmen	Kleine Unternehmen	Mittlere Unternehmen		
Grösse des Unternehmens (in Mitarbeitenden / Unternehmen)	1 – 9	10 – 49	50 – 249	> 249	
Anzahl Unternehmen	547'074	51'412	9'334	1'698	609'518
Anzahl Unternehmen (in %)	89.8%	8.4%	1.5%	0.3%	100.0%
geschätzte theoretische Stichprobengrösse für quantitative Studie	X	X	<b>11'032</b>		

Damit die Resultate als repräsentativ einzustufen sind, ist laut Berechnungen bei einer Fehlerspanne von 5% und einem Konfidenzniveau von 95% eine Stichprobe von mindestens 385 Teilnehmenden anzustreben. Die Berechnung wird nachfolgend dargestellt.

- Z-Wert für 95%-Konfidenzintervall (z) = 1.96
  - Fehlermarge (e) = 0.05
  - Standardabweichung (p) = 0.5
  - gesucht: Stichprobengrösse (n)
- $$n = \left( \frac{z^2 \times p \times (1 - p)}{e^2} \right) \Rightarrow \left( \frac{1.96^2 \times 0.5 \times (1 - 0.5)}{0.05^2} \right) \approx 385$$

<sup>13</sup> Bei einer Ad Hoc-Stichprobe werden alle verfügbaren Probandinnen und Probanden in das Untersuchungsdesign aufgenommen.

### 3.3.3 Durchführung der Fragebogenstudie

Die Durchführung der Umfrage begann am 25.03.2024 in Kalenderwoche 13 und endete am 26.04.2024, in Kalenderwoche 17. Verschiedene Kanäle haben während diesem Zeitraum dazu gedient, in der Schweiz tätige HR-Mitarbeitende mit der Einladung zum Ausfüllen der Umfrage zu erreichen. Diese werden in der Abb. 17 in chronologischer Reihenfolge aufgezählt.

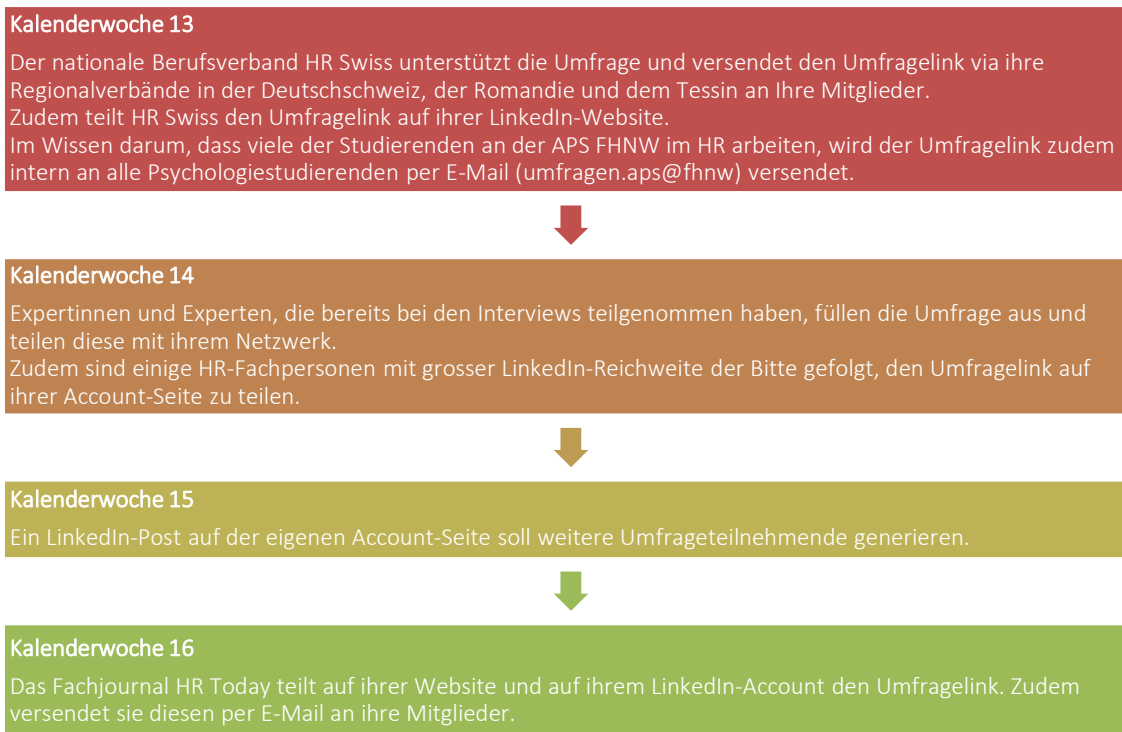


Abbildung 17. Umfrageverlauf in Kalenderwochen betreffend die genutzten Kanäle, um HR-Mitarbeitende in der Schweiz zu erreichen.

Während des genannten Zeitraums wurden auch Schulen angeschrieben, die auf Personalwirtschaft ausgerichtete Aus- und Weiterbildungen anbieten (z.B. Anbietende von Weiterbildungen zu HR-Assistentinnen und Assistenten, Personalfachleuten oder diplomierten HR-Leitenden oder Anbietende von MAS oder CAS). Diese Anfragen blieben leider erfolglos.

Das Nutzen der genannten Kanäle (siehe Abb. 17) führte in einem Zeitraum von knapp fünf Wochen dazu, dass der Umfragelink von 603 Personen geöffnet wurde. 418 dieser 603 Personen haben den Fragebogen jedoch nicht zu Ende ausgefüllt. 284 Personen haben bereits nach der Sprachauswahl bzw. der Begrüssungsseite ihre Teilnahme abgebrochen. Weitere 108 Personen haben die Umfrage nach der ersten oder zweiten Frageseite abgebrochen. 28 Personen haben die Umfrage im weiteren Verlauf abgebrochen.

Unvollständig ausgefüllte Fragebogendaten wurden in der Auswertung nicht berücksichtigt, wodurch eine Stichprobengrösse von  $n = 183$  erzielt werden konnte. Der zeitliche Verlauf der Umfrage und der darin berücksichtigten Teilnahmen ist in der nachfolgenden Abb. 18 dargestellt.

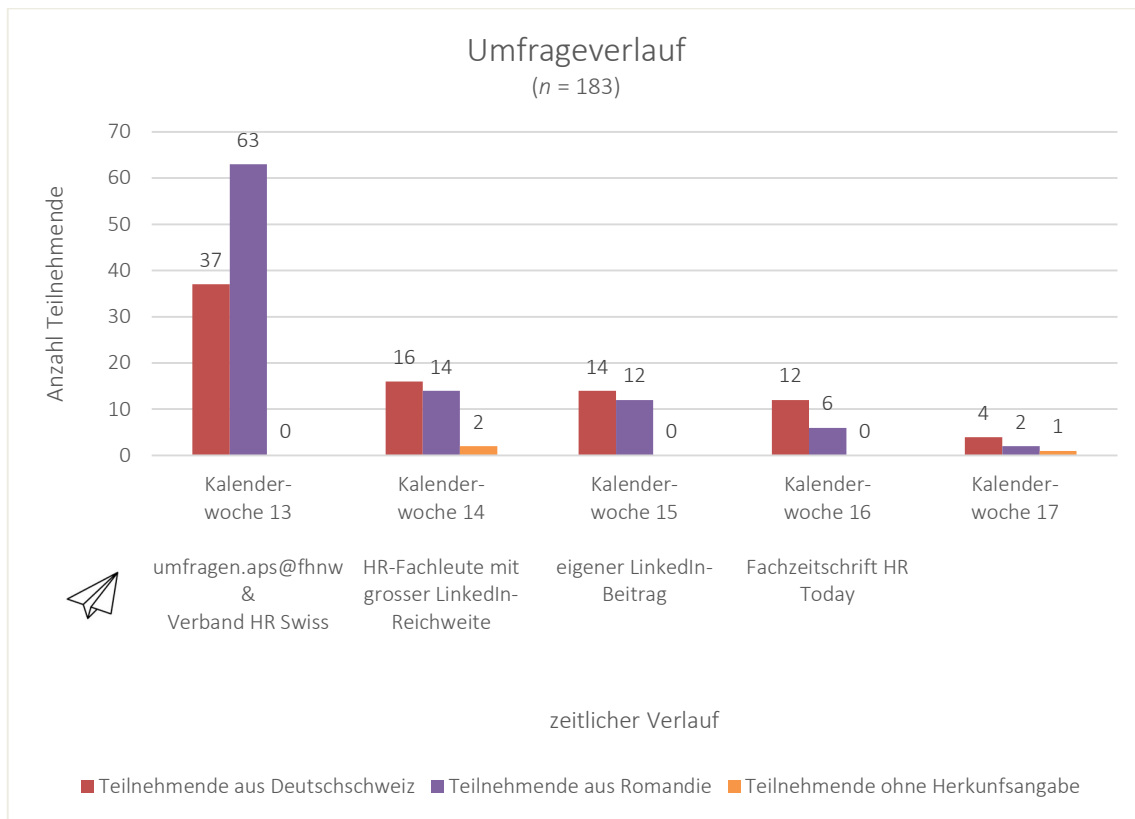


Abbildung 18. Umfrageverlauf in Kalenderwochen.

## 4 Quantitative Datenauswertung und Ergebnisdarstellung

Dieses Kapitel widmet sich der quantitativen Datenauswertung und den entsprechenden Ergebnisdarstellungen. Zunächst werden die Datenaufbereitung und die Stichprobenzusammensetzung beschrieben. Daran anschliessend werden die Fragestellungen beantwortet, beginnend mit den Fragestellungen OF1, OF2 und UF1 bis UF4. Deren Ergebnisse konnten durch einfache Auszählung gewonnen werden. Danach folgt die Bearbeitung der Fragestellungen UF5 bis UF7, die eine aufwendigere statistische Auswertung erforderten.

### 4.1 Vorbereitung der Datensets

Die durch das Umfragetool Tivian erhobenen Daten wurden in eine sav-Datei transformiert, um diese mit der Software SPSS bearbeiten und auswerten zu können. Zunächst wurden alle Datensets gesichtet, um unvollständig ausgefüllte Fragebogen identifizieren und ausschliessen zu können. 183 Datensets sind vollständig und wurden zur Auswertung verwendet. Des Weiteren wurde vorbereitend auf die Auswertung die Übersetzung der offenen Fragen vorgenommen, welche in französischer oder englischer Sprache formuliert wurden.

## 4.2 Stichprobenzusammensetzung

Die Fragebogenteilnehmenden haben sich zu demografischen Fragen zu ihrer Person und in Bezug auf das Unternehmen, in welchem sie tätig sind, geäußert. Diese Fragen wurden ausgewertet, um die Stichprobenzusammensetzung darstellen zu können. Nachfolgend werden zuerst die demografischen Daten betreffend die Teilnehmenden dargestellt. Daran anschliessend folgen die Daten zu den Unternehmen, in denen die Teilnehmenden tätig sind.

### Region

Die Mehrheit der Umfrageteilnehmenden stammt aus der Romandie ( $n = 98$ , 53%). Aus der Deutschschweiz haben 82 Personen (45%) teilgenommen. Bei drei Personen (2%) fehlt eine Angabe zur regionalen Herkunft. Die nachfolgende Abb. 19 veranschaulicht die Stichprobenzusammensetzung hinsichtlich der regionalen Herkunft der Teilnehmenden.

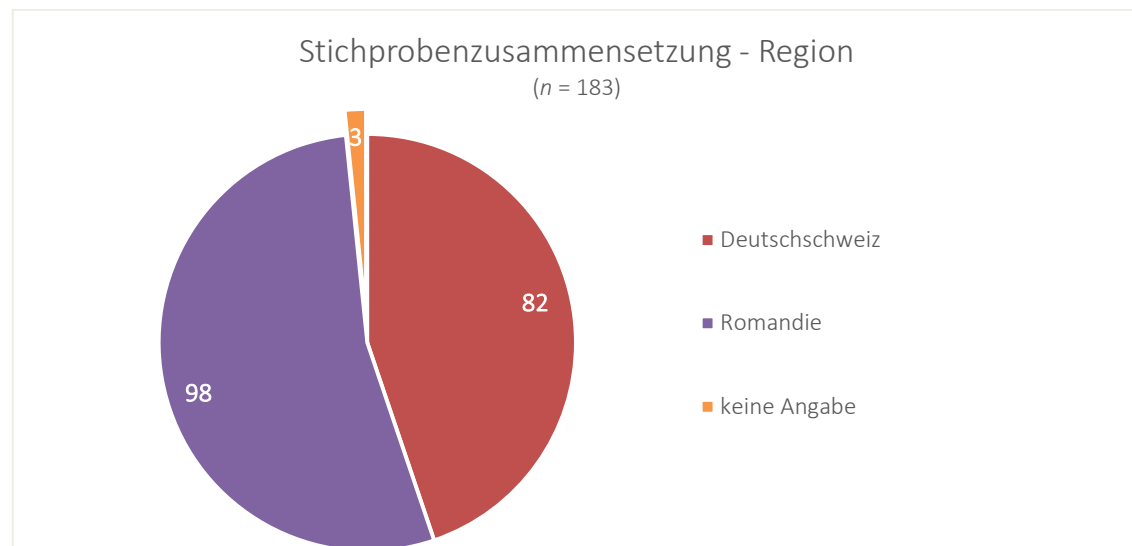


Abbildung 19. Stichprobenzusammensetzung – Region.

## Ausbildung

165 (90%) der HR-Personen, die an der Umfrage teilgenommen haben, verfügen über einen Ausbildungsabschluss auf Tertiärstufe, also einen Abschluss einer Hochschule ( $n = 98$ , 53%) oder einer berufsspezifischen Weiterbildung ( $n = 67$ , 37%). 14 (8%) der Befragten verfügen über einen Ausbildungsabschluss auf Sekundarstufe, haben also eine Berufslehre und / oder die Matura absolviert. Vier Personen (2%) haben keine Angabe zu ihren Ausbildungsabschlüssen gemacht. Die Abb. 20 visualisiert die Stichprobenzusammensetzung betreffend die Ausbildungen. Die daran anschliessende Abb. 21 zeigt ein differenzierteres Bild.

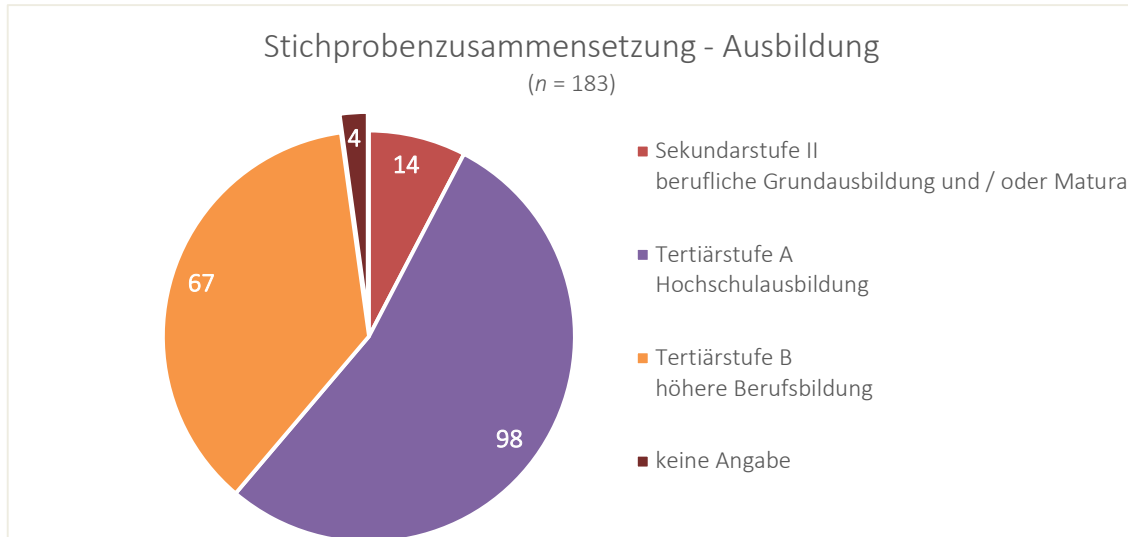


Abbildung 20. Stichprobenzusammensetzung – Ausbildung – grobe Auswertung.

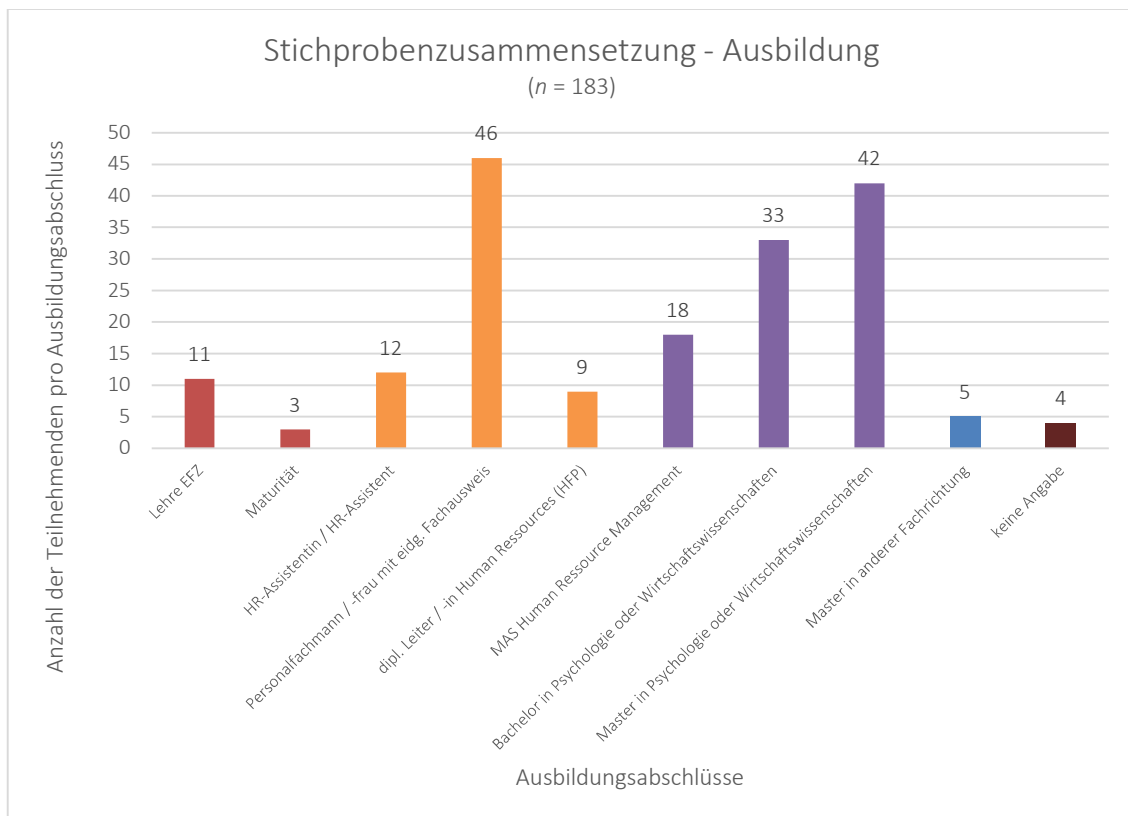


Abbildung 21. Stichprobenzusammensetzung – Ausbildung – spezifischere Auswertung.

Wie die Abb. 21 zeigt, haben fünf Personen einen Masterabschluss in anderen Fachrichtungen als Wirtschaft oder Psychologie. Drei der Befragten verfügen über einen Masterabschluss in sozialer Arbeit, eine Person hat einen Masterabschluss in Recht und eine Person einen Master in Geschichte.

### Berufserfahrung im Personalbereich

Die 183 Teilnehmenden haben im Durchschnitt 15 Jahre Berufserfahrung im Personalbereich. Der Range reicht von null bis 39 Jahre. Demnach ist die Streuung recht hoch: die Standardabweichung beträgt 9.8 Jahre, was einem Variationskoeffizienten von 65% entspricht. Die nachfolgende Abb. 22 visualisiert die hohe Variabilität betreffend die Berufserfahrung.

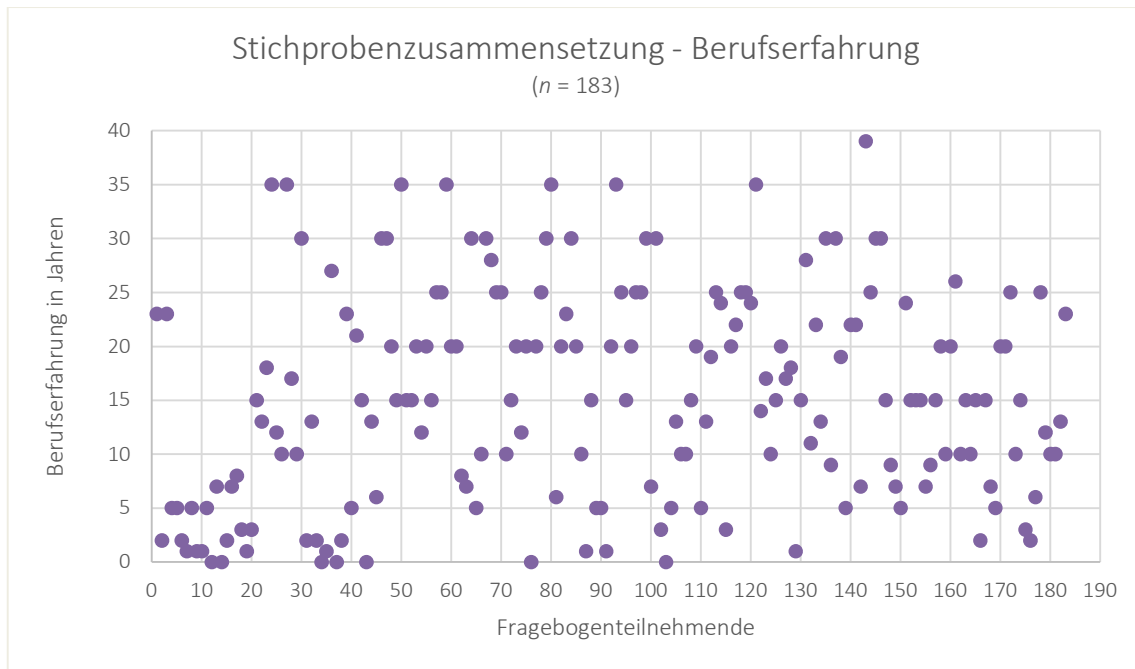


Abbildung 22. Stichprobenszusammensetzung – Berufserfahrung.

## Wirtschaftsabteilung

16 verschiedene Wirtschaftsabteilungen (orientiert am BFS, 2023c) standen den Umfrageteilnehmenden zur Auswahl. Die meisten Teilnehmenden haben sich den Wirtschaftsabteilungen *Erbringung wirtschaftlicher und sonstiger Dienstleistungen* ( $n = 24$ ), *Gesundheits- und Sozialwesen* ( $n = 26$ ), *verarbeitendes Gewerbe / Herstellung von Waren* ( $n = 28$ ) und *öffentliche Verwaltung* ( $n = 31$ ) zugeordnet. Andere Wirtschaftsabteilungen sind kaum bis moderat vertreten, wie dies die unterstehende Abb. 23 aufzeigt.



Abbildung 23. Stichprobenzusammensetzung – Wirtschaftsabteilung.

## Grösse des Unternehmens

95 (52%) der Umfrageteilnehmenden arbeiten für ein Grossunternehmen, 88 (48%) für ein kleines bis mittelgrosses Unternehmen (KMU). Durch die Daten war eine weitere Differenzierung betreffend die Unternehmensgrösse möglich, was veranschaulicht, dass die meisten Umfrageteilnehmenden in einem Grossunternehmen mit 1'000 bis 9'999 Mitarbeitenden arbeiten. Die nachfolgende Abb. 24 veranschaulicht die Stichprobenszusammensetzung hinsichtlich der groben Unterteilung und Abb. 25 hinsichtlich der differenzierten Unterteilung betreffend die Unternehmensgrösse.

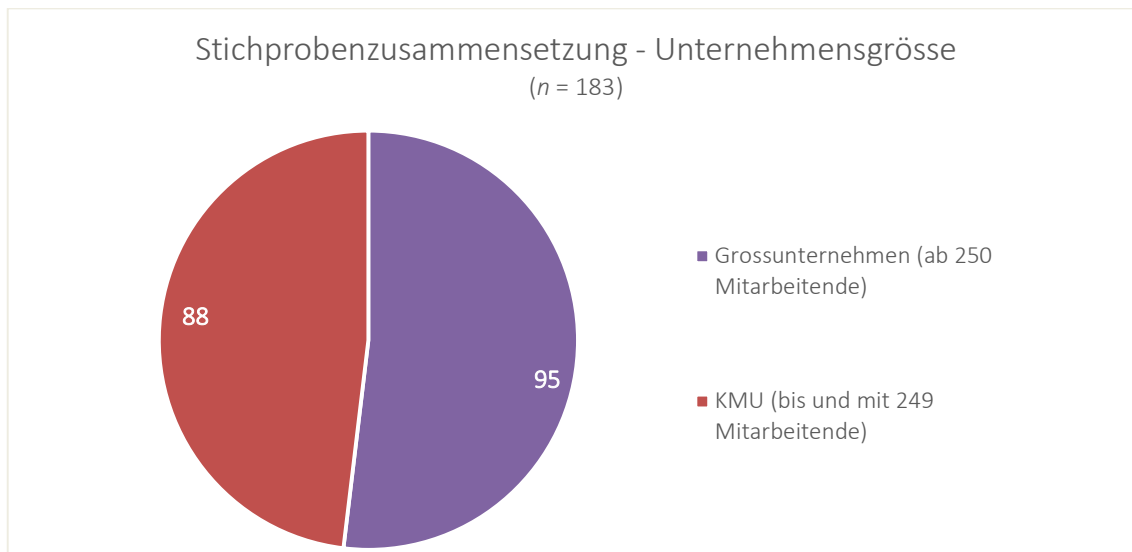


Abbildung 24. Stichprobenszusammensetzung – Unternehmensgrösse – grobe Auswertung.

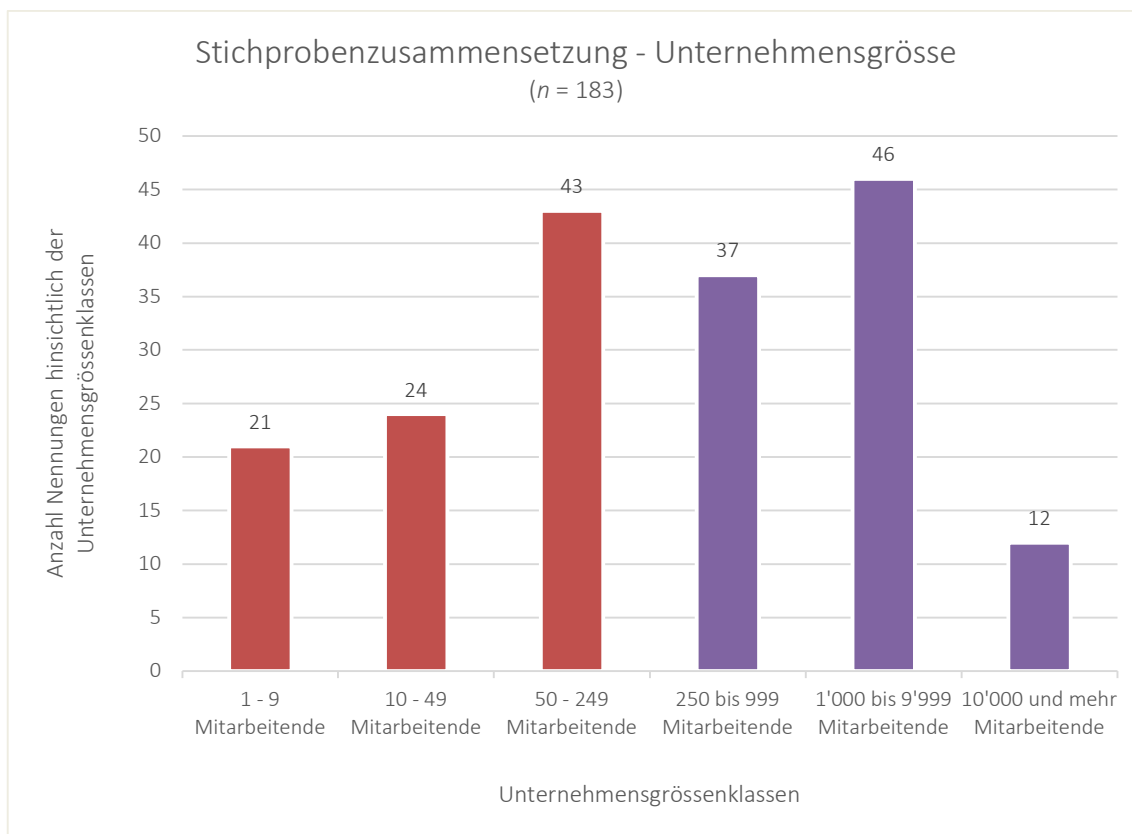


Abbildung 25. Stichprobenszusammensetzung – Unternehmensgrösse – spezifischere Auswertung.

### Psychologische Vertretung in der HR-Abteilung

Die Mehrheit ( $n = 121$ , 66%) der Befragten berichtet, dass in ihrer HR-Abteilung kein Psychologe bzw. keine Psychologin arbeitet. 54 (30%) der Umfrageteilnehmenden gaben hingegen an, dass in ihrer HR-Abteilung eine psychologische Vertretung vorhanden ist. Acht Personen (4%) konnten zu dieser Frage keine Angabe machen. Die Abb. 26 visualisiert die genannten Ergebnisse.

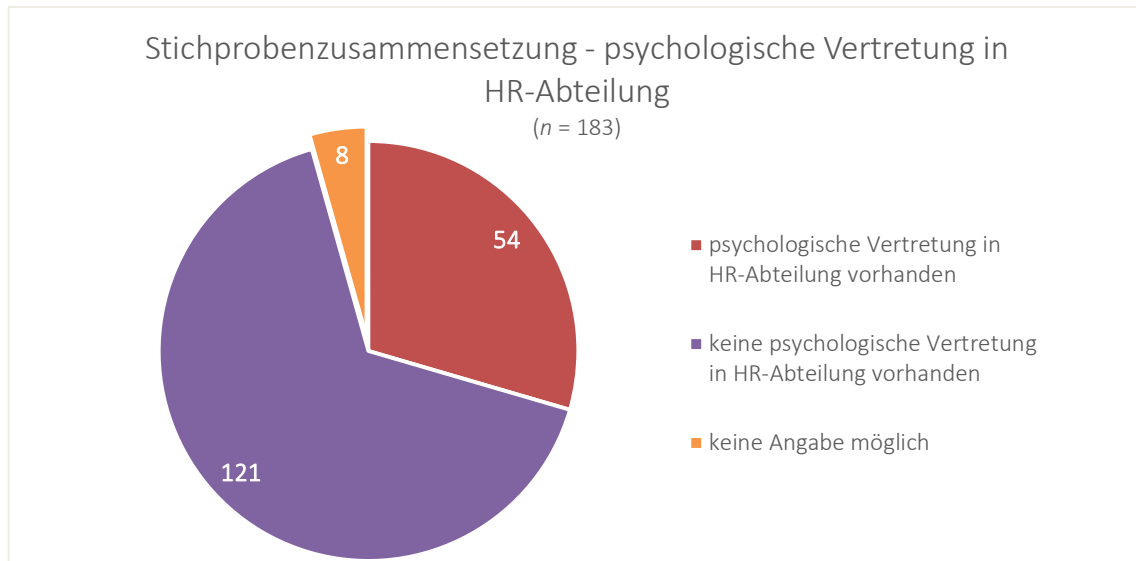


Abbildung 26. Stichprobenszusammensetzung – psychologische Vertretung im HR.

### 4.3 OF1 – Häufigkeit

62% ( $n = 113$ ) der befragten Personen setzen Leistungstests und Persönlichkeitsinventare in ihrem Unternehmen im Rahmen von Personalentscheidungen ein. 38% ( $n = 70$ ) der Befragten verzichten hingegen auf Testeinsatz. Die Abb. 27 visualisiert die genannten Ergebnisse.

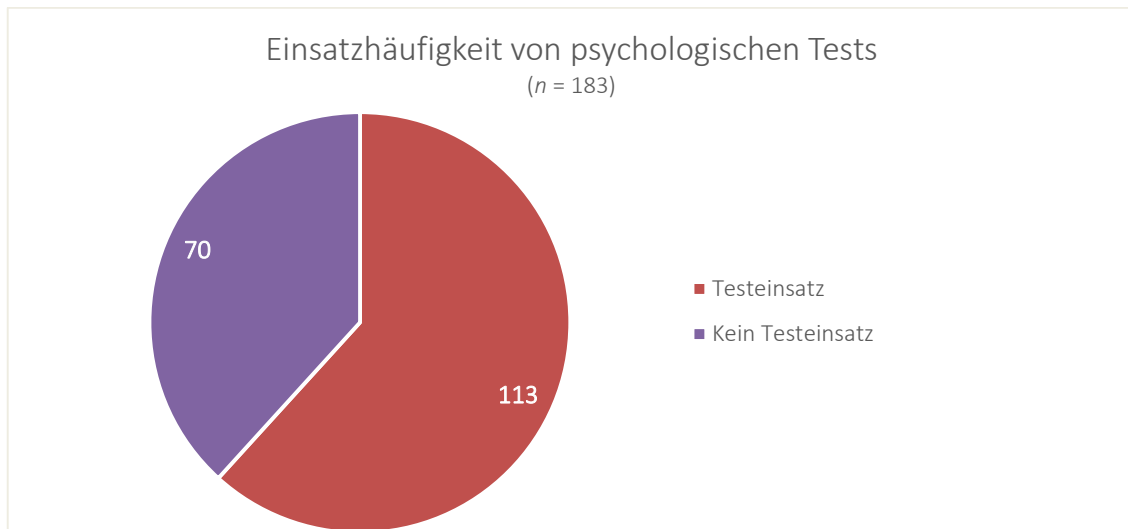


Abbildung 27. Auswertung Fragestellung OF1 zu Häufigkeit.

## 4.4 OF2 – Testart / Testnamen

Von den 113 Teilnehmenden, die psychologische Tests einsetzen in ihrem Unternehmen, wurden 131-mal Testnamen genannt und 16-mal eine Testbatterie genannt. 24-mal wurde aus diversen Gründen keine Angabe zu eingesetzten Tests gemacht. U.a. wurden fehlende Kenntnis oder Vertraulichkeit als Gründe genannt.

Mit 88% wurden mehrheitlich Persönlichkeitstests genannt. 12% der Nennungen entsprechen Leistungstests, wobei neben allgemeinen Intelligenztests auch ein Aufmerksamkeits- und Konzentrationstest genannt wurden. Die Abb. 28 stellt das dar.

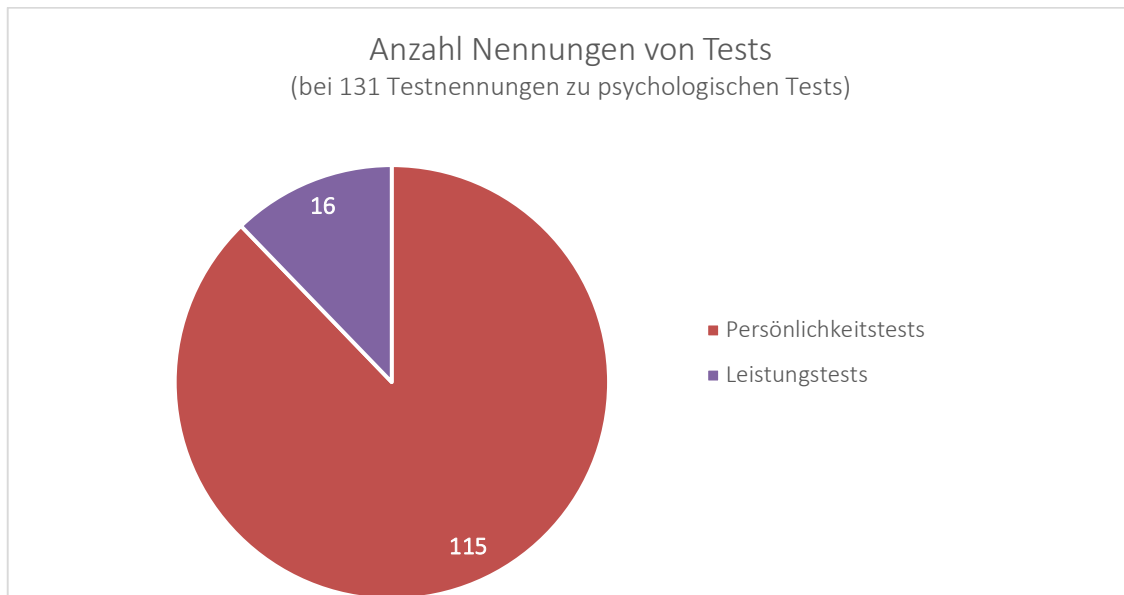


Abbildung 28. Auswertung der Fragestellung OF2 zu eingesetzten Testarten.

Unter den 131 Nennungen betreffend psychologische Tests finden sich 51 verschiedene Testnamen. Davon wurden 23 Tests mehrfach genannt. 28 Tests wurden nur einfach genannt. Die nachfolgende Abb. 29 visualisiert die Testnamen, die mehrfach genannt wurden. Die daran anschließende Tab. 11 enthält die Namen der je einmal genannten Tests. Die Daten mit Informationen über die genannten Test und die ausgeschriebenen Namen sind unter dem Anhang J zu finden.

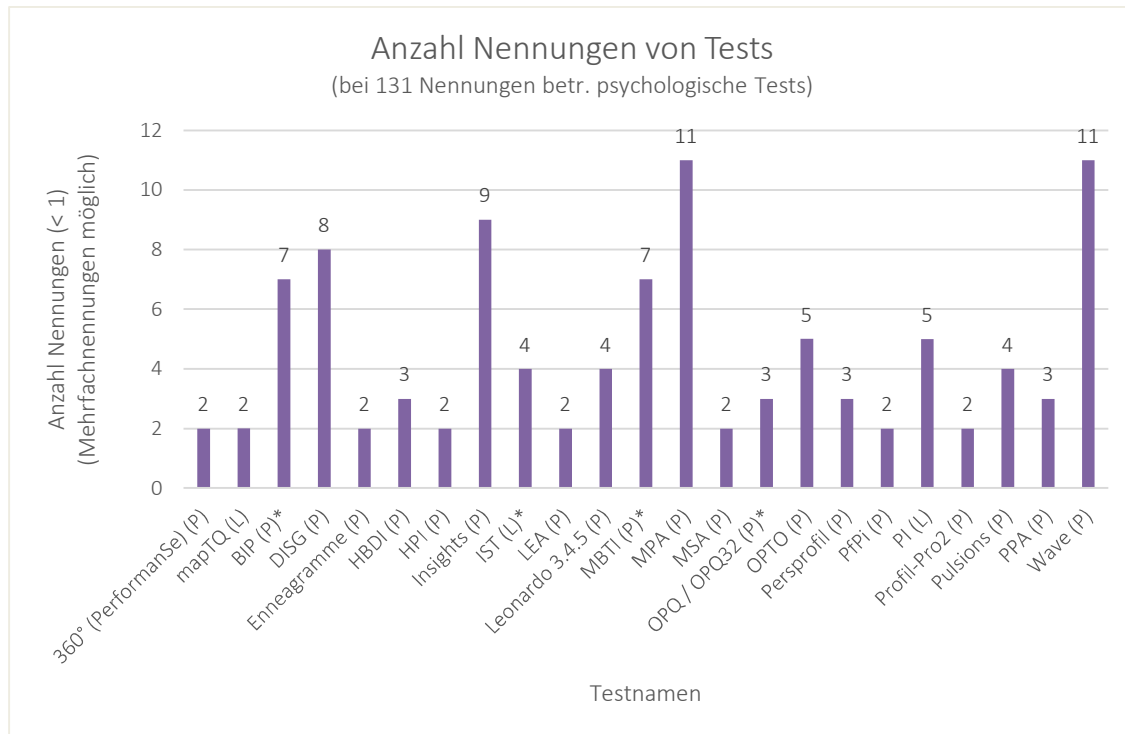


Abbildung 29. Tests, die mehr als einmal genannt wurden. P = Persönlichkeitstest, L = Leistungstest, \* = Testrezension bei SDBB und / oder TBS-DTK vorhanden.

Tabelle 11

Tests, die einmal genannt wurden. P = Persönlichkeitstest, L = Leistungstest, \* = Testrezension bei SDBB und / oder TBS-TK vorhanden.

Fragebogen zu Kompetenzen, Handlungseigenschaften und Persönlichkeit	Leistungstests
<ul style="list-style-type: none"> <li>• ADEPT15</li> <li>• Arc-en-ciel</li> <li>• Belbin</li> <li>• Clifton Strength</li> <li>• COMPER</li> <li>• ECHO</li> <li>• F-JAS</li> <li>• HEXACO*</li> <li>• IDI</li> <li>• ISK (P)*</li> <li>• L.A.B.E.L</li> <li>• LIFO</li> <li>• LMI*</li> <li>• MOTIVA</li> <li>• NEO-FFI / -PIR*</li> <li>• Nova Profil</li> <li>• Positive Intelligence</li> <li>• PMI</li> <li>• Profiler Tonus</li> <li>• Profilsan</li> <li>• QE Pro</li> <li>• QCP</li> <li>• Sosie</li> </ul>	<ul style="list-style-type: none"> <li>• CAPP</li> <li>• d2-R*</li> <li>• GIA</li> <li>• M-KIT*</li> <li>• Swift Aptitude</li> </ul>

Neben einzelnen Testnamen wurden auch ganze Testbatterien bzw. genutzte Testplattformen von den Fragebogenteilnehmenden genannt. 16 solcher Nennungen konnten gezählt werden. Die Abb. 30 zeigt genannte Testbatterien bzw. -plattformen auf.

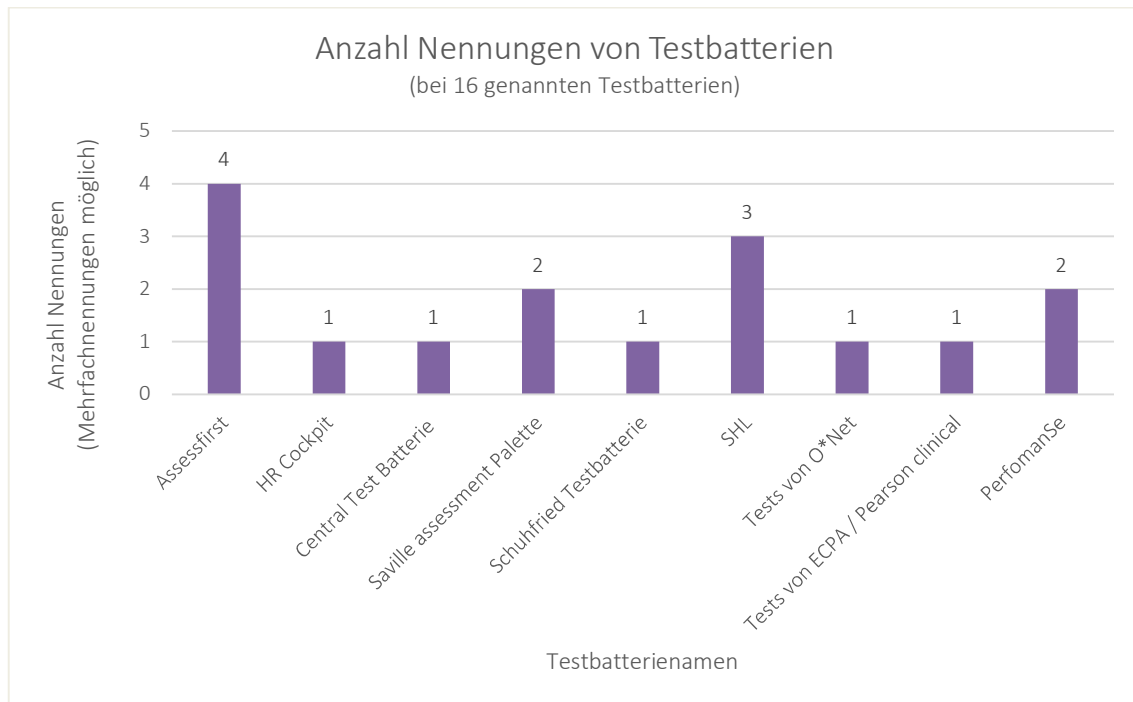


Abbildung 30. Auswertung zu OF2 – genannte Testbatterien.

## 4.5 UF1 – Qualität

Von den 113 Fragebogenteilnehmenden, die psychologische Tests in ihrem Unternehmen zu Personalentscheidungen einsetzen, wurden 131 Nennungen von Testnamen verzeichnet (vgl. Kap. 4.4). Von den 131 Testnennungen entsprechen 27 einem Test, der über eine Rezension vom TBS-DTK oder des SDBB verfügen. Das bedeutet, dass durch die 113 Fragebogenteilnehmenden, die den Testeinsatz bejahten, zu 21% rezensierte Tests eingesetzt werden (siehe Abb. 31 und 32).

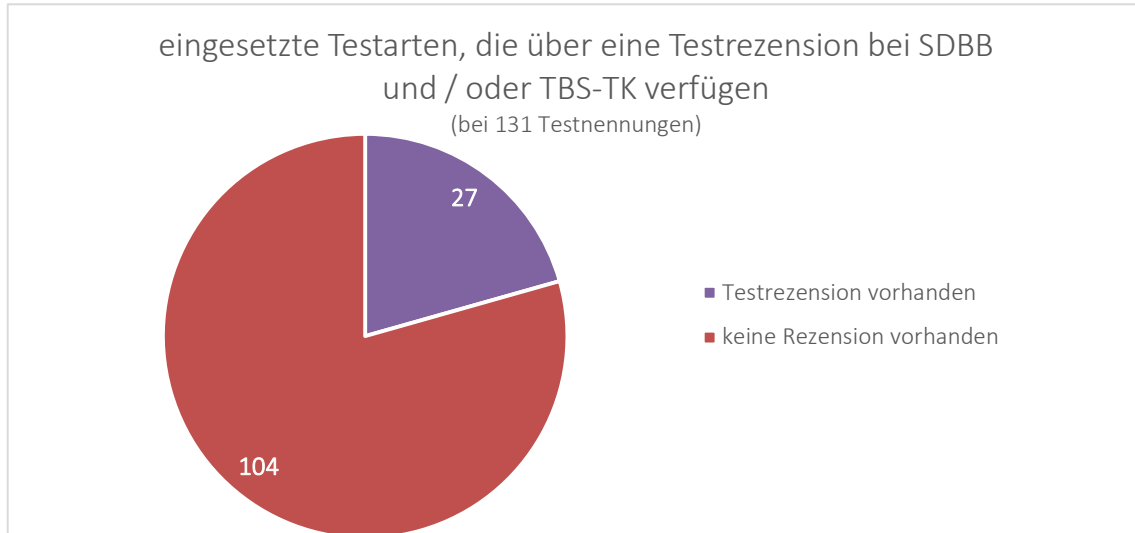


Abbildung 31. Auswertung der Fragestellung UF1 zu Testrezensionen.

Die Abb. 32 zeigt die genannten rezensierten Tests und ihre Verwendungshäufigkeit auf.

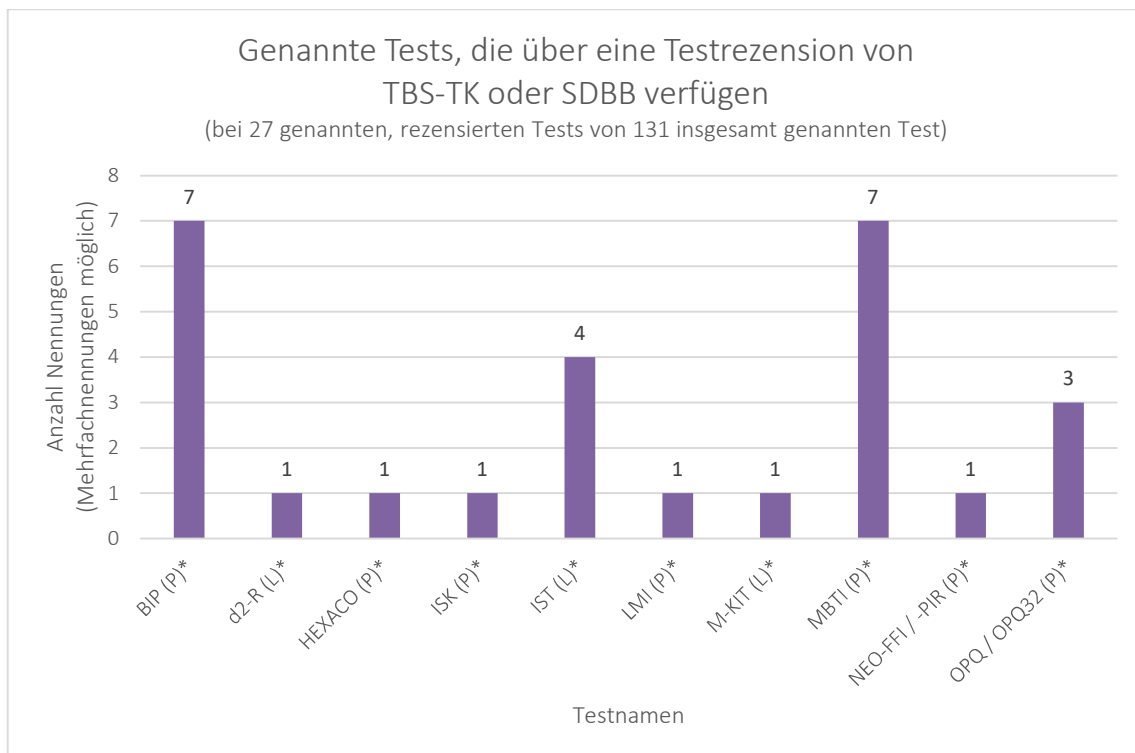


Abbildung 32. Auswertung der Fragestellung UF1 zu Testrezensionen.

## 4.6 UF2 – Zielgruppen

Von den 113 Umfrageteilnehmenden, die psychologische Tests in ihrem Unternehmen einsetzen, nannten die Teilnehmenden insb. Personen in Führungspositionen als Zielgruppe für psychologische Tests. Die Abb. 33 zeigt dies auf.

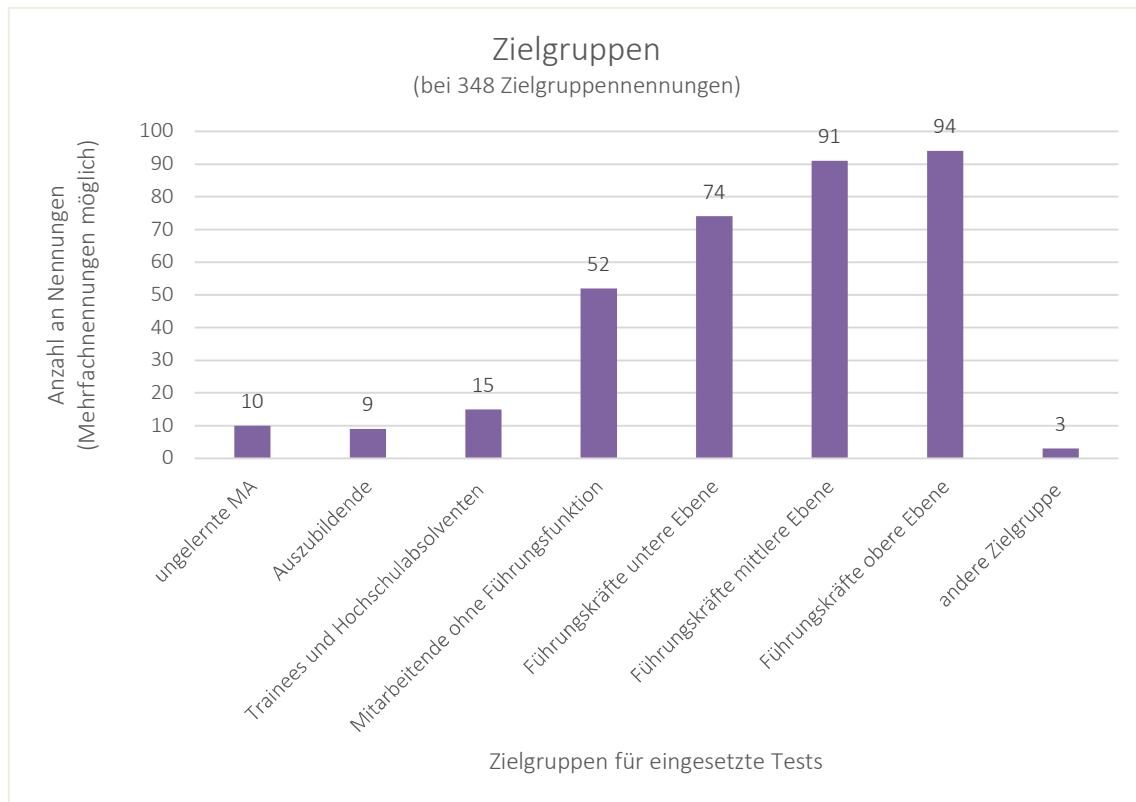


Abbildung 33. Auswertung der Fragestellung UF2 zu den Zielgruppen.

Wie die Abb. 33 zeigt, wurden auch andere Zielgruppen genannt, die nicht in der Fragebogenauswahl vertreten waren. Zweimal wurde die Zielgruppe *stellensuchende Klientinnen und Klienten* genannt und einmal die Zielgruppe *Personal, welches im sozialen oder schulischen Bereich tätig ist, und starken psychosozialen Risiken ausgesetzt ist*. Neun Umfrageteilnehmende verwiesen im offenen Antwortfeld darauf, dass sie Tests einsetzen bei der Auswahl von fachlichen Specialistinnen und Spezialisten. Damit präzisierten sie die Auswahlkategorie *Mitarbeitende ohne Führungsfunktion*.

## 4.7 UF3 – Testzweck

Von den 113 Fragebogenteilnehmenden, die den Testeinsatz bejahten, setzen 59% ( $n = 67$ ) Tests für die Personalauswahl und -entwicklung ein. 27% ( $n = 30$ ) setzten Tests nur zum Zweck der Personalauswahl ein. 14% ( $n = 16$ ) setzten Tests nur zum Zweck der Personalentwicklung ein (siehe Abb. 34).

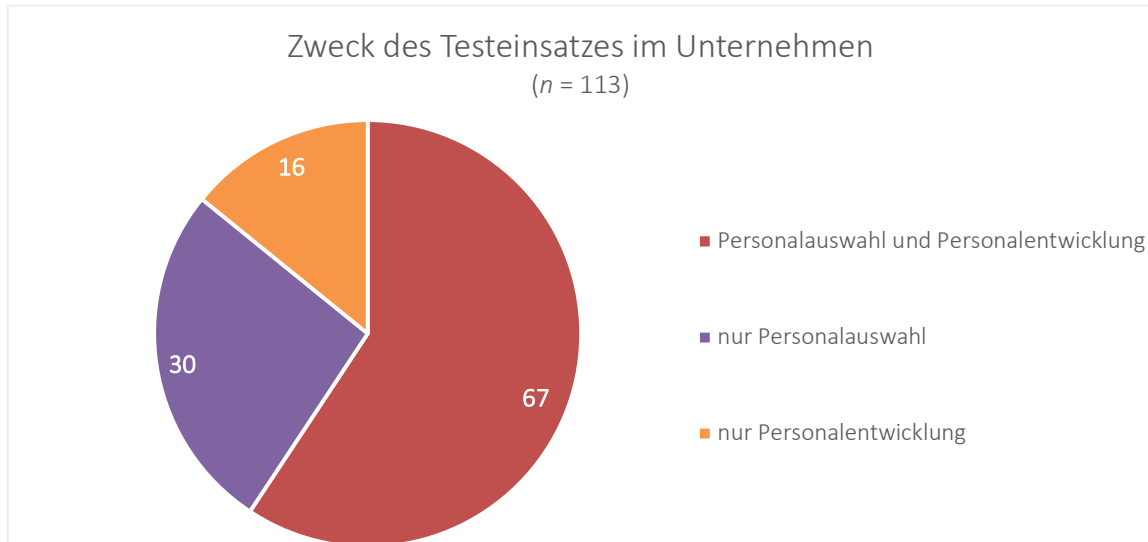


Abbildung 34. Auswertung der Fragestellung UF3 zu den Testzwecken.

Im offenen Textfeld wurden die Testzwecke spezifiziert bzw. wurden weitere Zwecke genannt, die in der Fragebogenauswahl nicht vertreten waren. Achtmal wurde im Zusammenhang mit der Personalentwicklung der Zweck *Teamaufbau*, *-entwicklung* und *Konfliktmanagement* genannt. Zusätzlich wurden je einmal die Zwecke *Nachfolgeplanung*, *Outplacement*, *Studierendenauswahl* und *Stellenkonzeption* genannt.

## 4.8 UF4 – Informationsbasis

Die Fragebogenteilnehmenden nutzen bzw. nutzten vielfältige Informationsquellen zur Entscheidung für den Einsatz psychologischer Tests. Die 113 Umfrageteilnehmenden, die psychologische Tests in ihrem Unternehmen einsetzen, nannten insb. den Austausch mit Fachexpertinnen, -experten und Testanbietenden als Informationsquelle. Die Abb. 35 zeigt dies auf.

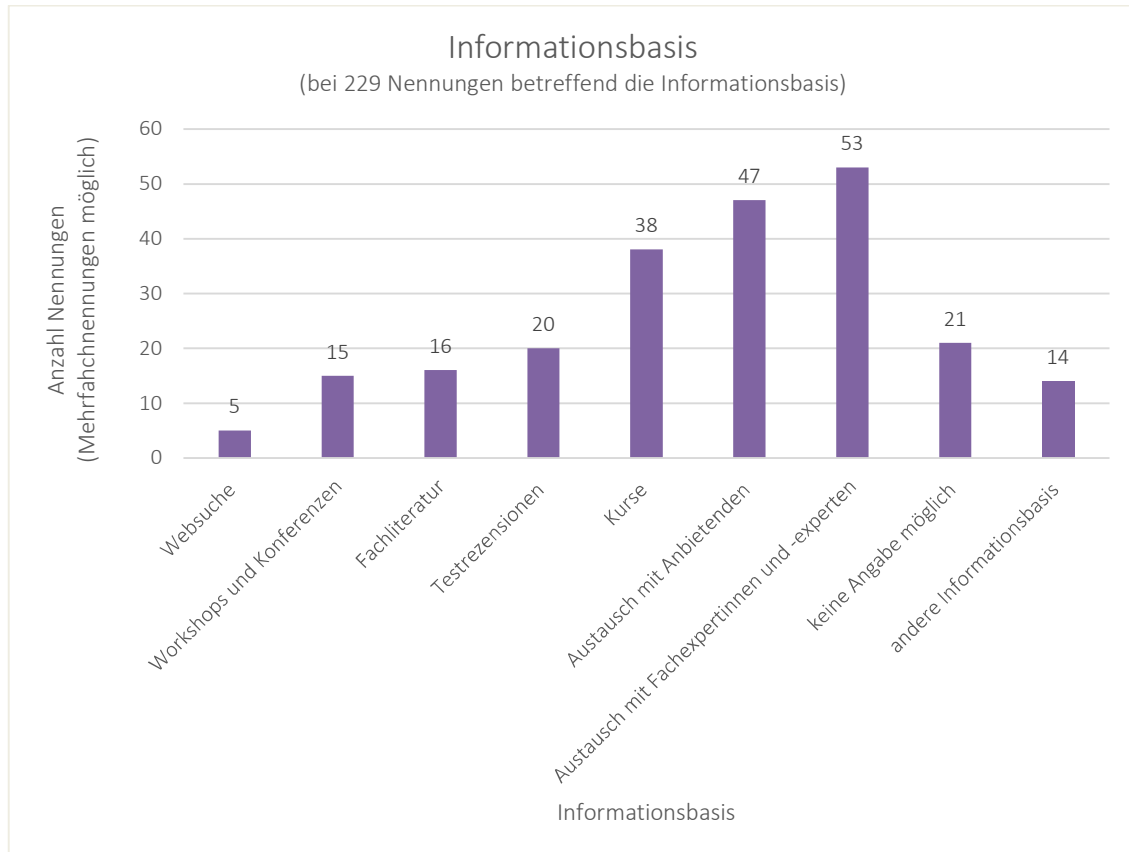


Abbildung 35. Auswertung der Fragestellung UF4 zur Informationsbasis.

Wie die Abb. 35 zeigt, wurden auch andere Informationsquellen genannt, die nicht in der Auswahl vertreten waren. Die Tab. 12 zeigt diese Nennungen.

Tabelle 12

Weitere Nennungen zur Informationsbasis.

Weitere Nennungen zur Informationsbasis
<ul style="list-style-type: none"> <li>• Auswahl externem Dienstleistung überlassen (3x genannt)</li> <li>• Austausch mit Kolleginnen und Kollegen (3x genannt)</li> <li>• Selbstversuche (2x genannt)</li> <li>• Spezifische Test-Zertifizierung und entsprechende Überzeugung zum psychologischen Test (2x genannt)</li> <li>• Persönliche Erfahrung (2x genannt)</li> <li>• Ausschreibung (1x genannt)</li> <li>• Psychologische Ausbildung und entsprechende Beschäftigung mit psychologischen Tests (1x genannt)</li> </ul>

## 4.9 UF5 bis UF7 – Hintergründe

Während die Resultate der OF1, OF2 und UF1 bis UF4 durch einfache Auszählungen erhalten werden konnten, mussten zur Ergebnisberechnung für die UF5 bis UF7 einige Berechnungsschritte unternommen werden. Um diesen Prozess möglichst nachvollziehbar darstellen zu können, dient das nachfolgende Flussdiagramm (siehe Abb. 36). Die darin beschriebenen Punkte werden in den nun folgenden Unterkapiteln detailliert erläutert.



Abbildung 36. Prozess der quantitativen Auswertung, um Ergebnisse für die Fragestellungen UF5 – 7 erhalten zu können.

### 4.9.1 Berechnung der Korrelationen auf Itemebene

An den erhobenen Daten wurden 39 statistische Tests in Form von Spearman-Korrelationen ( $r_s$ ) durchgeführt, was auf 40 aufgerundet wurde. Die Wahrscheinlichkeit ( $p$ ) wird durch die Bonferroni-Korrektur a posteriori berechnet, was zu folgender Anpassung führt: bei Deklaration von Signifikanz bei nachfolgenden Korrelationen auf Itemebene muss ein Wert von  $p^* < 0.001$  unterschritten werden. Die Rechnung zur Bonferroni-Korrektur wird nachfolgend dargestellt:

$$p^* = \frac{\alpha}{n} \rightarrow p^* = \frac{0.05}{40} \rightarrow p^* = 0.00125$$

Von der Unterscheidung von Signifikanz zu hoher und höchster Signifikanz wird dabei abgesehen.

Die berechneten korrelativen Zusammenhänge sind in Anhang K dargestellt und betreffend die ebengenannte, korrigierte Signifikanzgrenze eingeordnet.

## 4.9.2 Faktorenbildung zur Datenreduktion

39 Items im Fragebogen dienten der Aufklärung einflussnehmender Faktoren auf den Einsatz psychologischer Tests in Unternehmen in der Schweiz. Um schliesslich eine Regression rechnen zu können mit einer handlichen Anzahl latenter Variablen, ist zunächst eine Datenreduktion angezeigt. Die nachfolgenden Unterkapitel befassen sich mit dem Prozess der Datenreduktion, beginnend mit einer explorativen Faktorenanalyse (EFA), gefolgt von einer Reliabilitätsanalyse und einer daran anschliessenden konfirmatorischen Faktorenanalyse (CFA), wie dies auch die nachfolgende Abb. 37 visualisiert.

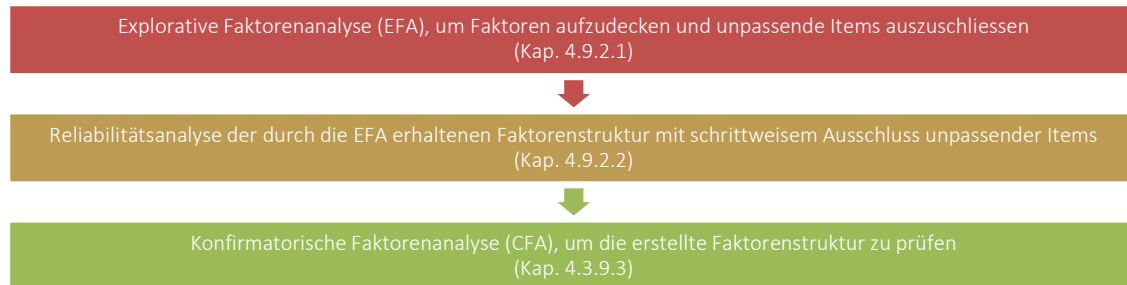


Abbildung 37. Ablauf der Datenreduktion und anschliessende Bestätigung der reduzierten Faktorenstruktur.

### 4.9.2.1 Durchführung der explorativen Faktorenanalyse (EFA)

Die faktorielle Struktur des verwendeten Fragebogens wurde durch eine EFA mittels der Extraktionsmethode *Maximum Likelihood*, der Rotationstechnik *Promax* und dem Abbruchkriterium *Parallelanalyse nach Horn* erhalten. Da SPSS dieses spezifische Abbruchkriterium nicht anbietet, wurde zur Durchführung der EFA die Software *Jamovi* verwendet. Zur Berechnung der Varianzaufklärung diente schliesslich die Software *Excel*.

Zunächst werden die Voraussetzungen zur Durchführung der EFA geprüft, anschliessend folgt die Besprechung der Durchführungsmodalitäten. Dann folgt das Ergebnis der EFA mit entsprechender Angabe der Fit-Indizes.

#### Voraussetzungsprüfung

Die Durchführung einer EFA verlangt das Prüfen bzw. das Begutachten einiger Voraussetzungen. Leonhardt (2017) setzt für die Stichprobengrösse eine mindestens dreimal so grosse Anzahl voraus, wie es Variablen gibt. In die EFA werden 39 unabhängige Variablen miteinbezogen, das Dreifache davon – 120 – ist kleiner als die Stichprobengrösse von  $n = 183$ . Diese Voraussetzung ist damit erfüllt. Ein besonders wichtiges Kriterium, welches zu erfüllen sei, ist das Vorliegen von Korrelationen zwischen den unabhängigen Variablen (Hartzinger & Nagel, 2013). Die Durchsicht der Korrelationsmatrix (siehe Anhang L) offenbart eine hohe Anzahl an Korrelationen zwischen den vorliegenden unabhängigen Variablen, wodurch auch diese Voraussetzung erfüllt ist. Des Weiteren sollen die Daten intervallskaliert sein und einer multivariaten Normalverteilung folgen (Hartzinger & Nagel, 2013; Bühner, 2011). Bereits der Kolmogorov-Smirnov-Test auf univariate Normalverteilung zeigte sich als signifikant, womit die Annahme einer entsprechenden Normalverteilung also verworfen werden muss (*durchschnittliches*  $p < 0.001$ ). Die Sichtprüfung der Histogramme jedoch lässt bei den meisten unabhängigen Variablen eine Normalverteilung erahnen. Da in Jamovi oder SPSS keine Berechnungen zur multivariaten Normalverteilung vorgenommen werden können, erfolgte die Berechnung mittels online Tool *StatistikGuru* (Hemmerich, 2018). Der Henze-Zirkler-Test auf multivariate Normalverteilung zeigte sich ebenfalls signifikant ( $p < .001$ ). Demnach liegt keine multivariate Normalverteilung vor und die entsprechende Voraussetzung ist also nicht erfüllt (siehe Anhang M).

Faktorenanalysen seien relativ robust ggü. dieser Voraussetzungsverletzung (Döring, 2016). Da eine EFA explorativen Charakter besitzt, ist diese Voraussetzung auch nicht als restriktiv zu verstehen (Hartzinger & Nagel, 2013). Besondere Aufmerksamkeit sollen aber nach der Durchführung der EFA den Indizes zur Modellgüte geschenkt werden – dem Bartlett-Test auf Sphärizität und dem Kaiser-Mayer-Olkin-Kriterium (Hartzinger & Nagel, 2013; Döring, 2016). Zur Vollständigkeit vorweggenommen: beide genannten Indizes zeigen Modellgüte an.

### **Durchführungsmodalitäten**

Zur Durchführung der EFA fiel die Wahl auf die Extraktionsmethode *Maximum Likelihood*. Gerade diese Methode führe trotz Verletzung der Voraussetzung betreffend die multivariate Normalverteilung zur besten Reproduktion der Ausgangskorrelationsmatrix (Bühner, 2011). Demnach kann davon ausgegangen werden, dass im Vergleich zur *Hauptkomponentenanalyse* die Maximum Likelihood Extraktionsmethode geeigneter ist zur Durchführung dieser EFA.

Die Rotationsmethode *Promax* entspricht einer Kombination aus orthogonaler und schiefwinkliger Rotation (Bühl, 2016). Bühner (2011) empfiehlt, bei korrelierenden Faktoren ( $r > .10$ ) die Anwendung der Promax-Methode. Dass die Faktoren untereinander korrelieren, konnte durch die Ausgabe der Korrelationsmatrix (Anhang L) bereits festgestellt werden.

Als Abbruchkriterium wurde die *Parallelanalyse nach Horn* gewählt. Vergleichsweise hatte das Abbruchkriterium nach *Kaiser Gutmann* lediglich eine zweifaktorielle Struktur vorgeschlagen, denn nur zwei Faktoren zeigten Eigenwerte deutlich über 1. Dadurch hätten sich die unabhängigen Variablen lediglich in eine Logik positiver und negativer Einflussfaktoren einbetten lassen, womit keine wirkliche Datenreduktion hätte erwirkt werden können. Diese Methode bzw. die dadurch erhaltene Lösung wurde deshalb verworfen.

Die Parallelanalyse nach Horn zeigt im entsprechenden Screeplot, dass fünf Faktoren mit ihren Eigenwerten noch über dem simulierten Eigenwertverlauf liegen. Die Tab. 13 zeigt die vorgeschlagene Faktorenstruktur mit den entsprechenden MSA-Werten, also denjenigen Werten, die die Eignung der Items zur Durchführung der EFA zeigt. Die Tab. 13 zeigt auch, welche Items ausgeschlossen wurden. Die Rohdaten der EFA sind in Anhang N abgelegt. Nach der Tab. 13 werden die jeweiligen Ausschlüsse erläutert.

Tabelle 13

Durch EFA gefundene Faktoren, Erläuterungen betreffend die Itemausschlüsse und Varianz- bzw. Gesamtvarianzaufklärung durch die gebildeten Faktoren.

Faktoren bzw. Items	Faktorladung	MSA-Wert	Massnahmen	quadrierte Ladungen auf jeweiligen Faktor	Summe der quadrierten Faktorladungen in %	quadrierte Summe der quadrierten Faktorladungen in %	Kumulation der quadrierten Summen der quadrierten Faktorladungen in % (Varianzaufklärung)
<b>Optimierung der Entscheidungsgüte</b>							
Einarbeitung	0.393	0.890	-	0.128	37%	13%	13%
Personalentwicklung	0.562	0.941	-	0.293			
Person-Job-Fit	0.734	0.930	-	0.533			
Person-Group Fit	0.696	0.902	-	0.483			
cultural Fit	0.766	0.890	-	0.500			
Reliabilität	0.726	0.941	-	0.411			
Kriteriumsvalidität	0.804	0.926	-	0.572			
wissenschaftliche Fundierung	0.645	0.927	-	0.301			
Entscheidungssicherheit	0.431	0.931	-	0.185			
zusätzliche Puzzlezeile	0.418	0.927	-	0.181			
Objektivierung	0.487	0.938	-	0.193			
Minimierung v. Fehlbesetzungen	0.696	0.943	-	0.475			
Return on Investment	0.745	0.934	-	0.514			
Vergleich zwischen Bewerbenden	0.701	0.924	Item ausgeschlossen aufgrund inhaltlicher Überlegungen				
<b>Praktikabilität und Entlastung</b>							
Benchmarks und Normen	0.615	0.884	-	0.438	39%	15%	28%
Fairness	0.878	0.861	-	0.884			
Handhabbarkeit	0.49	0.913	-	0.278			
Erweiterung der Gesprächsbasis	0.359	0.836	-	0.159			
Vorauswahl treffen	0.369	0.779	-	0.174			
<b>Herausforderungen und Risiken</b>							
hohe Lizenzkosten	0.679	0.865	-	0.549	44%	19%	48%
hohe Zeitintensität	0.67	0.892	-	0.529			
ungewollte Selbstselektion	0.846	0.894	-	0.601			
Arbeitgebendenattraktivität leidet	0.8	0.923	-	0.654			
fehlende Akzeptanz durch Bewerbende	0.896	0.882	-	0.560			
Scheingenauigkeit	0.298	0.889	-	0.142			
mangelnde Zielgruppenangepasstheit	0.371	0.850	-	0.259			
zu generisch	0.323	0.893	-	0.215			
<b>Kontext</b>							
Effizienz d. schwerpunktmässige Selektierung	0.315	0.868	-	0.106	22%	5%	53%
Expertise	0.345	0.733	-	0.056			
Unternehmenskultur	0.547	0.873	-	0.497			
Finanzen	0.561	0.621	-	0.155			
Akzeptanz durch Linienvorgesetzte	0.59	0.678	-	0.411			
Anzahl Bewerbungen / Arbeitsmarkt	0.343	0.757	-	0.108			
Trends und Verbreitung	0	0.378	Item ausgeschlossen da es nicht auf entsprechenden Faktor (oder einen anderen Faktor) lädt und weil der MSA-Wert inakzeptabel ist, der Einbezug in die EFA demnach nicht empfohlen ist.				
<b>Zugänglichkeit zu Tests und Theorie</b>							
Science-Practitioner-Gap	0.462	0.672	Item ausgeschlossen, da MSA nur ausreichend ist und Faktor nur aus zwei zu unähnlichen Items besteht.				
Vielfältiges Angebot erschwert Testauswahl	-0.349	0.587	Item ausgeschlossen, da MSA nur ausreichend ist und Faktor nur aus zwei zu unähnlichen Items besteht.				
<b>Persönliche Einstellung</b>							
persönliche Befürwortung von Testeinsatz	x	0.956	Die Items, laden eigentlich moderat auf Faktor <i>Optimierung der Entscheidungsgüte</i> , werden aber zu einem separaten Faktor zusammengeschlossen, da vermutet wird, dass der resultierende Faktor als Mediator oder als weiterer unabhängiger Faktor eingesetzt werden kann, wobei sich diese Vermutung orientiert an der TRA von Fishbein und Ajzen (1975).				
positive Erfahrung mit Tests	x	0.940					
<b>Einfluss der persönlichen Meinung</b>							
Einfluss der persönlichen Meinung	X	0.608	Dieses Item lädt auf keinen Faktor, was auch inhaltlichen Überlegungen entspricht.				

Der Bartlett-Test auf Sphärizität bestätigte im Allgemeinen die Eignung der Daten zur Durchführung der Faktorenanalyse ( $\chi^2_{(df = 741)} = 3745, p < .001$ ). Das Kaiser-Meyer-Olkin-Kriterium (KMO) bestätigt diese Eignung ebenfalls ( $KMO = .897$ ). Nach Bühl (2016) kann dies als gut bis sehr gut bezeichnet werden.

Die MSA-Werte lagen bei beinahe allen Items in mindestens akzeptablen Bereich. Lediglich ein Item – *Trends und Verbreitung* – zeigte mit  $MSA = 0.378$  einen inakzeptablen Wert und musste folglich ausgeschlossen.

Das Item *Vergleich zwischen Bewerbenden* wurde trotz seines hervorragenden MSA-Wertes ausgeschlossen. Es fehlte die inhaltliche Logik betreffend die Zugehörigkeit zum Faktor *Optimierung der Entscheidungsgüte*. Vielmehr wurde erwartet, dass es sich dem Faktor *Praktikabilität und Entlastung* zuordnen lassen würde.

Aufgrund weiterer inhaltlicher und semantischer Überlegungen wurde der Faktor *Zugang zu Tests und entsprechender Theorie* ausgeschlossen. Der Name des Faktors suggeriert eine gewisse inhaltliche Breite, welche aber durch die kleine Anzahl an Items nicht abgedeckt werden kann. So wurde durch Ausschluss dieses Faktors eine vierfaktorielle Lösung gewählt, welche 53% der Varianz zu erklären vermag. Die vier Faktoren lauten:

1. *Optimierung der Entscheidungsgüte*, also fundiertere Personalentscheidungen durch den Testeinsatz treffen zu können.
2. *Praktikabilität und Entlastung*, also durch Testeinsatz Unterstützung bei Personalentscheidungen erhalten zu können.
3. *Herausforderungen und Risiken*, die mit dem Testeinsatz einhergehen, wie fehlender Nutzen, hoher Aufwand oder die Gefahr, Bewerbende dadurch zu verlieren.
4. *Kontext*, also äussere Einflussfaktoren, die durch die HR-Mitarbeitenden kaum zu beeinflussen sind.

In der nun folgenden Reliabilitätsanalyse wird geprüft, ob die den Faktoren zugeordneten Items ihren jeweiligen Faktor auch wirklich adäquat vertreten.

#### 4.9.2.2 Durchführung der Reliabilitätsanalyse

Die Durchführung der Reliabilitätsanalyse orientiert sich an Bühner (2011), Bühl (2016), Zöfel (2003) und Döring (2016). Die Berechnungen erfolgten mittels SPSS und Excel. Mittels Excel wurden die Indizes *Variationskoeffizient* und *Itemschwierigkeit* berechnet, durch SPSS wurden die Indizes *Trennschärfenkoeffizient* und *Cronbachs Alpha* der Gesamtskalen bzw. *Cronbachs Alpha, wenn Item gelöscht*, berechnet. Zur Bewertung der eben genannten Indizes wurden die nachfolgenden Regeln verwendet, welche in Tab. 14 dargestellt sind. Die Reliabilitätsanalyse mit den berechneten Indizes sind in Tab. 15 abgebildet.

Tabelle 14  
Indizes zur Einordnung der Werte, die im Rahmen der Reliabilitätsanalyse berechnet wurden.

Indizes zur Einordnung der Reliabilitätsanalyse				
	Variationskoeffizient	Itemschwierigkeit (SI)	Trennschärfenkoeffizient (TKS)	Reliabilitätskoeffizient Cronbachs Alpha (Cronbachs $\alpha$ )
<b>Ziel bzw. Regel</b>	mittlere Streuung wird angestrebt	SI von 0 – 20%: wenig informative Items, Itemausschluss empfohlen SI von 80 – 100%: wenig informative Items, Itemausschluss empfohlen 20 – 80%: informative Items, Item sollte beibehalten werden (Döring, 2016; Bühner, 2011)	vergleichsweise niedrige TKS-Werte sollen zum Ausschluss der jeweiligen Items führen (Bühl, 2016)  TKS 0.3 bis 0.5: mittelmässig TKS > 0.5: hoch (Döring, 2016)	> 0.6 akzeptable Reliabilität > 0.7 gute Reliabilität > 0.8 sehr gute Reliabilität > 0.9 sehr gute Reliabilität, wobei möglicherweise redundante Items vorliegen (Bühner, 2011)

Tabelle 15

Durch Reliabilitätsanalyse gefundene Indizes, Erläuterungen betreffend die Itemausschlüsse.

	N	Min-Wert	Max-Wert	Mittelwert	Std.-Abweichung	Variationskoeffizient	Schwierigkeitsindex (SI)	Trennschärfen-koeffizient (TSK)	Cronbachs $\alpha$ , wenn Item weggelassen	Cronbachs $\alpha$ für entwickelte Skala
<b>Optimierung der Entscheidungsgüte</b>										
Einarbeitung	183	1	6	3.49	1.382	40%	50%	0.518	0.927	Cronbachs $\alpha$ = 0.921 > sehr guter Cronbachs Alpha Wert: alle Items werden beibehalten
Personalentwicklung	183	1	6	4.49	1.279	29%	70%	0.632	0.923	
Person-Job-Fit	183	1	6	3.84	1.316	34%	57%	0.717	0.921	
Person-Group-Fit	183	1	6	3.70	1.224	33%	54%	0.638	0.923	
cultural Fit	183	1	6	3.42	1.392	41%	48%	0.566	0.925	
Reliabilität	183	1	6	3.45	1.345	39%	49%	0.717	0.921	
Kriteriumsvalidität	183	1	6	3.11	1.408	45%	42%	0.610	0.924	
wissenschaftliche Fundierung	183	1	6	3.67	1.327	36%	53%	0.700	0.921	
Entscheidungssicherheit	183	1	6	3.98	1.387	35%	60%	0.704	0.921	
zusätzliche Puzzleteile	183	1	6	4.46	1.239	28%	69%	0.585	0.925	
Objektivierung	183	1	6	3.89	1.456	37%	58%	0.709	0.921	
Minimierung v. Fehlbesetzungen	183	1	6	3.55	1.514	43%	51%	0.766	0.919	
Return on Investment	183	1	6	3.30	1.538	47%	46%	0.725	0.920	
<b>Praktikabilität und Entlastung</b>										
Benchmarks und Normen	183	1	6	3.66	1.412	39%	53%	0.642	0.651	Cronbachs $\alpha$ = 0.744 > guter Cronbachs Alpha Wert: alle Items werden beibehalten
Fairness	183	1	6	4.04	1.427	35%	61%	0.676	0.637	
Handhabbarkeit	183	1	6	3.84	1.432	37%	57%	0.482	0.708	
Erweiterung des Gesprächsbedarfs	183	1	6	4.17	1.617	39%	63%	0.383	0.748	
Vorauswahl treffen	183	1	6	3.50	1.617	46%	50%	0.400	0.741	
<b>Herausforderungen und Risiken</b>										
hohe Lizenzkosten	183	1	6	3.34	1.459	44%	47%	0.660	0.855	Cronbachs $\alpha$ = 0.874 > sehr guter Cronbachs Alpha Wert: alle Items werden beibehalten
hohe Zeitintensität	183	1	6	3.09	1.450	47%	42%	0.637	0.858	
ungewollte Selbstselektion	183	1	6	2.86	1.526	53%	37%	0.696	0.851	
Arbeitgebendenattraktivität leidet	183	1	6	2.55	1.382	54%	31%	0.725	0.849	
fehlende Akzeptanz durch Bewerbende	183	1	6	2.78	1.366	49%	36%	0.666	0.855	
Scheingenauigkeit	183	1	6	3.46	1.565	45%	49%	0.552	0.867	
mangelnde Zielgruppenangepasstheit zu generisch	183	1	6	3.83	1.537	40%	57%	0.567	0.865	
	183	1	6	3.28	1.333	41%	46%	0.564	0.865	
<b>Kontext</b>										
Effizienz d. schwerpunktmässige Selektierung	183	1	6	4.04	1.549	38%	61%	0.381	0.556	Cronbachs $\alpha$ = 0.657 > akzeptabler bis guter Cronbachs Alpha Wert: nur fünf der sechs Items werden in diese Skala miteinbezogen.
Expertise	183	1	6	3.69	1.755	48%	54%	0.314	0.583	
Unternehmenskultur	183	1	6	3.49	1.634	47%	50%	0.432	0.533	
Finanzen	183	1	6	3.87	1.611	42%	57%	0.395	0.549	
Akzeptanz durch Linienvorgesetzte	183	1	6	3.90	1.521	39%	58%	0.465	0.524	
Anzahl Bewerbungen / Arbeitsmarkt	183	1	6	3.26	1.797	55%	45%	0.135 > TSK im Vergleich sehr tief	0.657 > Item ausgeschlossen, da sich dadurch $\alpha$ erheblich verbessert.	
<b>Persönliche Einstellung</b>										
persönliche Befürwortung von Testeinsatz	183	1	6	3.94	1.559	40%	59%	0.616	Berechnung nicht möglich, weil nur zwei Items	Cronbachs $\alpha$ = 0.762 > guter Cronbachs Alpha Wert: Faktor wird beibehalten
positive Erfahrung mit Tests	183	1	6	3.94	1.502	38%	59%	0.616		
<b>Einfluss der persönlichen Meinung</b>										
Einfluss der persönlichen Meinung	183	1	6	3.74	1.672	45%	55%	Berechnungen nicht möglich, da Faktor aus nur einer Variablen besteht.		

Die berechneten Indizes zeigten sich bei den meisten Items mindestens im akzeptablen Bereich. Jedoch musste das Item *Auswahlmöglichkeiten aufgrund Bewerbendenanzahl* ausgeschlossen werden. Zum einen verfügt dieses Item über einen vergleichsweise tiefen TSK-Wert ( $TSK = 0.135$ ). Zudem verbessert der Itemausschluss die Reliabilität des Faktors erheblich (von  $Cronbachs \alpha = 0.613$  zu  $Cronbachs \alpha = 0.657$ ).

*Persönliche Einstellung*<sup>14</sup> und *Einfluss der persönlichen Meinung*<sup>15</sup> wurden ebenfalls als Faktoren aufgenommen, wodurch nun eine sechsfaktorielle Lösung verfolgt wird. Weitere Anpassungen wurden nicht vorgenommen. Das Faktorenmodell mit seinen enthaltenen Items wurde anschliessend einer CFA unterzogen, deren Durchführung im nun folgenden Kap. 4.9.2.3 erläutert wird.

#### 4.9.2.3 Durchführung der konfirmatorischen Faktorenanalyse (CFA)

Aufgrund der Reliabilitätsanalyse wurden einige Items und eine Skala ausgeschlossen. Die Faktorenstruktur soll nun durch eine konfirmatorische Faktorenanalyse bestätigt werden.

Dazu wurde wieder die Extraktionsmethode Maximum Likelihood gewählt. Dabei wurden die Items ihren jeweiligen Faktoren zugeordnet. Die folgenden gängigen statistischen Indizes der CFA zur Bewertung der Passung des Faktorenmodells sind herangezogen worden:

Der Comparative Fit Index (CFI) ergab einen Wert von 0.818. Der Tucker-Lewis Index (TLI) lag bei 0.803. Bühner (2011) empfiehlt als Daumenregel, beim CFI und beim TLI den Cut-off-Wert bei 0.95 und höher zu setzen. Die berechneten Werte liegen darunter, womit der Modellfit in Frage gestellt wird.

Der Root Mean Square Error of Approximation (RMSEA) betrug 0.0770. Der RMSEA sollte bei einer Stichprobengrösse von  $n < 250$  unter 0.06, bei einer Stichprobengrösse von  $n \geq 250$  unter 0.08 liegen (Bühner, 2011). Man kann hier sagen, dass das wohl eher knapp erfüllt ist.

Der Standardized RMR (SRMR) beträgt 0.0810. Schreiber (2017) verweist darauf, dass Werte gleich oder unter 0.08 einem akzeptablen Modellfit entsprechen, womit dieses Kriterium sehr knapp nicht erfüllt ist.

Der Test auf exakten Fit durch den Chi-Quadrattest zeigt sich signifikant ( $\chi^2_{(df = 517)} = 1079, p < .001$ ), womit Modellfit angezeigt ist. Die Rohdaten der CFA befinden sich im Anhang O.

Bei den fünf berechneten Indizes zeigten nur zwei Indizes – strenggenommen nur ein Index – einen guten Modellfit an. Schreiber (2017) empfiehlt, mehrere Indizes in die Interpretation miteinzubeziehen. Schliesslich sind einige sensibel betreffend die Stichprobengrösse, andere lehnen bei komplexer Modellstruktur den Modellfit fälschlicherweise eher ab.

Diejenigen Indizes, die dem vorliegenden Faktorenmodell den Fit deutlich absprechen – CFI, TLI und RMSEA – sind nach Schreiber (2017) auch diejenigen, die eben auf komplexe Modellstrukturen eher mit fälschlicher Ablehnung des Fits reagieren. Diejenigen Indizes, die den Modellfit (knapp) implizieren – Chi-Quadrattest und SRMR – sind ggü. komplexen Modellstrukturen unempfindlich (Schreiber, 2017). Demnach lässt sich annehmen, dass das erstellte Faktorenmodell einen angemessenen Fit zu den Daten aufweist und daher für weitere Analysen und Interpretationen geeignet ist. Deshalb steht der Berechnung der Regressionsanalyse nun lediglich noch die entsprechende Voraussetzungsprüfung im Weg. Zunächst werden die Korrelationen der neu gebildeten Faktoren zur abhängigen Variablen berechnet (siehe Kap. 4.9.2.4), ehe die Regressionsanalyse (siehe Kap. 4.9.2.5) durchgeführt wird.

---

<sup>14</sup> *Persönliche Einstellung* vereint Befürwortung und gesammelte positive Erfahrungen bzw. Ablehnung und negative Erfahrungen in Bezug auf Tests und deren Einsatz.

<sup>15</sup> *Einfluss der persönlichen Meinung* bezieht sich auf die Partizipationsmöglichkeit der HR-Mitarbeitenden bei der Entscheidung, ob Tests im Unternehmen eingesetzt werden sollen oder nicht.

#### 4.9.2.4 Berechnung der Korrelationen auf Faktorebene

Mit den aus der CFA erhaltenen Faktoren wurde eine Korrelationsmatrix berechnet. In der Tab. 16 sind die Spearman-Korrelationen ( $r_s$ ) zwischen den erhaltenen Faktoren und der abhängigen Variablen *Testeinsatz* zu sehen. Alle berechneten Korrelationen zeigten sich signifikant.

Tabelle 16  
Korrelationen auf Testeinsatz auf Faktorebene. \* = 2-seitige Signifikanz, wobei von Unterscheidung von Signifikanz zu hoher und höchster Signifikanz abgesehen wird.

Anz. unabhängige Variablen	Faktornamen	Deskriptive Werte		Korrelationen zur abhängigen Variable <i>Testeinsatz</i>
1	Optimierung der Entscheidungsgüte	M	3.72	$r_s = 0.430^*$
		SD	0.983	
		V	26%	
2	Praktikabilität und Entlastung	M	3.84	$r_s = 0.246^*$
		SD	1.057	
		V	28%	
3	Herausforderungen und Risiken	M	3.15	$r_s = -0.488^*$
		SD	1.059	
		V	34%	
4	kontextuelle Einflussfaktoren	M	3.80	$r_s = 0.254^*$
		SD	1.049	
		V	28%	
5	persönliche Einstellung	M	3.94	$r_s = 0.533^*$
		SD	1.375	
		V	35%	
6	Einfluss der persönlichen Meinung	M	3.74	$r_s = 0.411^*$
		SD	1.672	
		V	45%	

#### 4.9.2.5 Berechnung binär logistischer Regressionen

Die Berechnung von binären logistischen Regressionen verlangt eine Voraussetzungsprüfung. Diese wurde entlang der vorgeschlagenen Regeln von IBM (2021) und der UZH (2024) durchgeführt und erfüllt. Die Voraussetzungsprüfung kann unter Anhang Q eingesehen werden.

Die anschliessende Berechnung der binären logistischen Regression wurde in SPSS vorgenommen. Dabei wurde die Methode *vorwärts LR* (vorwärts Likelihood-Ratio-Test) nach Empfehlung von Bühl (2016) gewählt, um dasjenige Modell zu finden, welches die höchste Modellgüte besitzt. So wurden die folgenden unabhängigen Variablen, die aus der CFA (siehe Kap. 4.8.2.3) erhalten wurden, in die Berechnung aufgenommen: *Optimierung der Entscheidungsgüte, Praktikabilität und Entlastung, Herausforderungen und Risiken, kontextuelle Einflussfaktoren, persönliche Einstellung, Einfluss der persönlichen Meinung*. Die höchste Modellgüte besass dasjenige Modell, welches lediglich die unabhängigen Variablen miteinbezug, welche über signifikante Regressionskoeffizienten auf die abhängige Variable *Testeinsatz* zeigten: *Herausforderungen und Risiken, Einfluss der persönlichen Meinung* und *persönliche Einstellung*. Dieses Modell wurde schliesslich weiterverfolgt und wird im Folgenden dokumentiert.

Die binär logistische Regressionsanalyse zeigte hier, dass sowohl das Modell als Ganzes ( $\chi^2_{(df=3)} = 100.31, p < .001$ ), als auch die einzelnen Koeffizienten der Variablen signifikant sind. Betreffend die Einschätzung der Modellgüte kann das *Nagelkerkes R<sup>2</sup>* herangezogen werden, welches die Erklärungskraft des Modells angibt. Dieses liegt bei  $R^2 = .57$ , was nach Backhaus, Erichson, Plinke & Weiber (2008) darauf schliessen lässt, dass das Regressionsmodell eine sehr gute Erklärungskraft besitzt.<sup>16</sup>

<sup>16</sup> Einteilung der *Nagelkerkes R<sup>2</sup>* nach Backhaus et al. (2008):  
 $R^2 > .20$  = akzeptabel |  $R^2 > .40$  = gut |  $R^2 > .50$  = sehr gut

Zudem wurde die Effektstärke  $f^2$  berechnet, um das *Nagelkerkes*  $R^2$  einschätzen zu können. Dazu wurde die nachfolgende Formel verwendet:

$$f^2 = \frac{R^2}{1 - R^2} = \frac{0.57}{1 - 0.57} \cong 1.33$$

Mit  $f^2 \cong 1.33$  handelt es sich, nach Cohen (1988) interpretiert, um einen starken Effekt.<sup>17</sup>

Die Ergebnisse der binären logistischen Regressionsanalyse werden in der nachfolgenden Tab. 17 und in der Abb. 38 dargestellt und daran anschliessend besprochen.

Tabelle 17

Berechnung der binären logistischen Regression mithilfe von SPSS, wobei das OR-1 (in %), orientiert an Morris & Gardner (1988), selbst berechnet und der Tabelle hinzugefügt wurde.

unabhängige Variablen	Regressionskoeffizient (B)	Standardfehler (SE)	Wald	df	Signifikanz	Exp(B) / OR	OR – 1 (in %)
<b>persönliche Einstellung</b>	0.87	0.19	21.64	1	< .001	2.39	139%
<b>Herausforderungen und Risiken</b>	-0.83	0.25	10.97	1	< .001	0.44	-56%
<b>Einfluss der persönlichen Meinung (pers. Meinung)</b>	0.64	0.14	21.28	1	< .001	1.90	90%
<b>Konstante</b>	-2.36	1.28	3.41	1	0.065	0.095	x

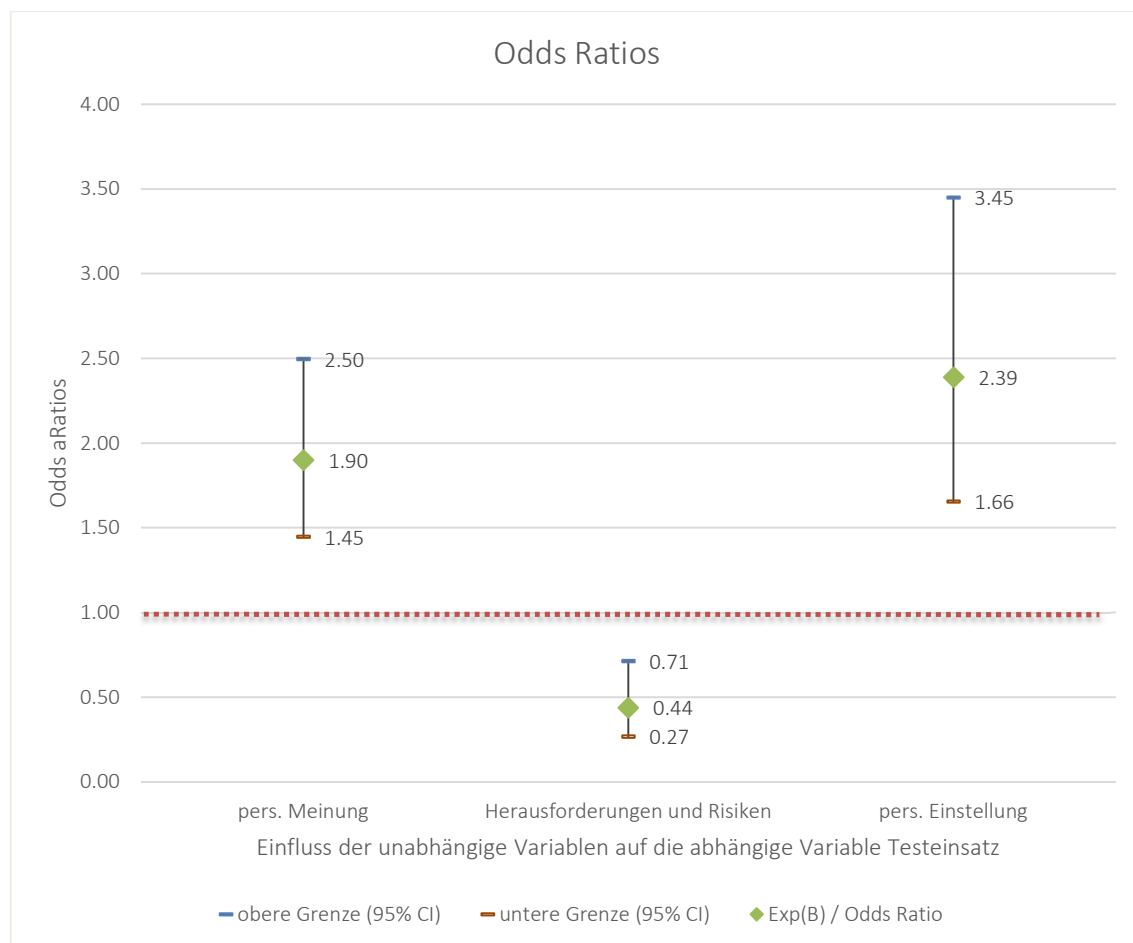


Abbildung 38. Ergebnisse der Regressionsanalyse in grafischer Form. Die Konfidenzintervalle wurden entlang der nachfolgenden Formeln, orientiert an Morris & Gardner (1988), berechnet: obere Grenze: 95% CI =  $e^{B-1.96 \times SE}$ , untere Grenze: 95% CI =  $e^{B-1.96 \times SE}$ .

<sup>17</sup> Einteilung der Effektstärke  $f^2$  nach Cohen (1988):

$f^2 > 0.02$  = schwacher Effekt |  $f^2 > 0.15$  = mittlerer Effekt |  $f^2 > 0.35$  = starker Effekt

Die Tab. 17 zeigt, dass die Regressionskoeffizienten der folgenden unabhängigen Variablen signifikant ausfallen: *persönliche Einstellung* (Wald = 21.61,  $p < .001$ ), *Herausforderungen und Risiken* (Wald = 10.97,  $p < .001$ ), *Einfluss der persönlichen Meinung* (Wald = 21.28,  $p < .001$ ).

Steigt die Bewertung in Bezug auf den *Einfluss der persönlichen Meinung betreffend Testeinsatz im Unternehmen* um eine Einheit, also würde auf der Antwortskala eine Stufe höher gewählt, so steigt die relative Wahrscheinlichkeit für *Testeinsatz im Unternehmen* um 90%, bzw. um den Faktor von 1.90 an. Das OR von 1.90 zeigt eine positive Assoziation zwischen dem *Einfluss der persönlichen Meinung betreffend Testeinsatz im Unternehmen* und dem tatsächlichen *Testeinsatz im Unternehmen*. Einfach gesagt: wird der *Einfluss der eigenen Meinung in Bezug auf den Einsatz psychologischer Tests im Unternehmen* als positiv bewertet, desto wahrscheinlicher ist es, dass im Unternehmen auch Tests eingesetzt werden.

Sinkt die Bewertung betreffend die *Herausforderungen und Risiken durch Testeinsatz* um eine Einheit, also würde auf der Antwortskala eine niedrigere Stufe gewählt, so steigt die relative Wahrscheinlichkeit für den *Testeinsatz im Unternehmen* um 56%, bzw. um den Faktor von 0.44. Das OR von 0.44 liegt unter 1 und zeigt damit einen negativ gerichteten Zusammenhang an zwischen der unabhängigen Variablen *Herausforderungen und Risiken* und der abhängigen Variablen *Testeinsatz im Unternehmen*. Einfach gesagt: werden *Herausforderungen und Risiken durch Testeinsatz* wahrgenommen, ist *Testeinsatz im Unternehmen* weniger wahrscheinlich.

Steigt die *persönliche Einstellung ggü. dem Einsatz psychologischer Tests* um eine Einheit, also würde auf der Antwortskala eine höhere Stufe gewählt, so steigt die relative Wahrscheinlichkeit für den *Testeinsatz im Unternehmen* um 139%, bzw. um den Faktor von 2.40. Das OR von 2.40 zeigt demnach eine positive Assoziation zwischen *persönlicher Einstellung* und dem *Testeinsatz im Unternehmen*. Einfach gesagt: je positiver die *persönliche Einstellung ggü. dem Einsatz psychologischer Tests* ist, desto wahrscheinlicher werden im Unternehmen Tests eingesetzt.

#### 4.9.2.6 Berechnung linearer Regressionen zur Exploration

Die Korrelationsmatrix der zusammengefassten Faktoren (siehe Anhang P) lässt weitere Zusammenhänge erahnen, die durch die binär logistische Regressionsanalyse nicht aufgedeckt werden können. Die Vermutung liegt nahe, dass die *persönliche Einstellung* nicht nur als unabhängige Variable fungiert, sondern auch als Mediator zwischen weiteren unabhängigen Variablen und der abhängigen Variablen. Aufgrund der unterschiedlichen Skalenniveaus ist die Berechnung einer Mediationsanalyse nicht möglich. Die abhängige Variable müsste mindestens Intervallskalenniveau besitzen (Backhaus et al. 2008). Um dennoch der geäußerten Vermutung nachzugehen, werden Pearson-Korrelationen ( $r_p$ ) berechnet (siehe Tab. 18). Danach wird eine weitere explorative, lineare Regression in SPSS gerechnet.

Tabelle 18

Korrelationen auf persönliche Einstellung auf Faktorebene. \* = 2-seitige Signifikanz, wobei von der Unterscheidung von Signifikanz zu hoher und höchster Signifikanz abgesehen wird.

Anz. unabh. abhängige Variablen	Faktornamen	Deskriptive Werte		Korrelationen zur abhängigen Variable persönliche Einstellung
		M	SD	
1	Optimierung der Entscheidungsgüte	M	3.72	$r_p = 0.806^*$
		SD	0.983	
		V	26%	
2	Praktikabilität und Entlastung	M	3.84	$r_p = 0.609^*$
		SD	1.057	
		V	28%	
3	Herausforderungen und Risiken	M	3.15	$r_p = -0.603^*$
		SD	1.059	
		V	34%	
4	kontextuelle Einflussfaktoren	M	3.80	$r_p = 0.415^*$
		SD	1.049	
		V	28%	
5	Einfluss der persönlichen Meinung	M	3.94	$r_p = 0.237^*$
		SD	1.375	
		V	35%	

Die Berechnung linearer Regressionen verlangte ebenfalls eine Voraussetzungsprüfung. Diese orientierte sich an den Vorgaben zur Berechnung linearer Regressionen von Döring (2016), sowie Eid, Gollwitzer und Schmitt (2017). Auch diese Voraussetzungsprüfung wurde erfüllt und kann unter Anhang R eingesehen werden.

Entlang der Empfehlung von Bühl (2016) wurde zur Berechnung der linearen Regressionen die *schrittweise* Methode gewählt, wobei die nachfolgend aufgezählten Faktoren als unabhängige Variable in die Berechnung aufgenommen wurden: *Optimierung der Entscheidungsgüte, Praktikabilität und Entlastung, Herausforderungen und Risiken, kontextuelle Einflussfaktoren* und *Einfluss der persönlichen Meinung*. Als abhängige Variable wurde die *persönliche Einstellung* eingesetzt.

Die höchste Modellgüte besass dasjenige Modell, welches lediglich die unabhängigen Variablen miteinbezog, welche über einen signifikanten Regressionskoeffizienten auf die abhängige Variable *persönliche Einstellung* zeigten: *Optimierung der Entscheidungsgüte* und *Herausforderungen und Risiken* ( $F_{(df = 3, 179)} = 129.21, p < .001$ ). Dieses Modell wurde schliesslich weiterverfolgt. Die Resultate dieser Berechnung sind in der nachfolgenden Tab. 19 und der Abb. 39 eingetragen. Der Tab. 19 bzw. Abb. 39 anschliessend finden sich die Erläuterungen der Resultate.

Tabelle 19

Berechnung der linearen Regression mithilfe von SPSS.

Faktoren	Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten		T	Signifikanz
	Regressionskoeffizient (B)	Standardfehler (SE)	$\beta$			
Optimierung der Entscheidungsgüte	0.94	0.07	0.67		12.89	< .001
Herausforderungen und Risiken	-0.21	0.07	-0.16		-2.99	0.003
Kontext	0.14	0.06	0.10		2.18	0.030
Konstante	0.59	0.51	x		1.15	0.253

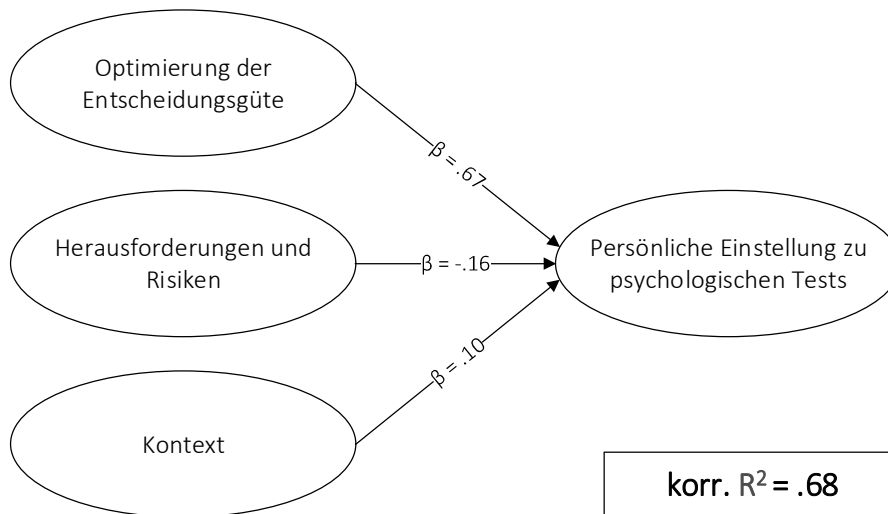


Abbildung 39. Ergebnisse der durchgeführten linearen Regression.

Sinkt die Bewertung betreffend die *Herausforderungen und Risiken* durch Testeinsatz um eine Einheit, also würde auf der Antwortskala eine niedrigere Stufe angewählt, so steigt die Bewertung der *persönlichen Einstellung ggü. Tests* um 0.21 Stufen auf der Antwortskala. Einfach gesagt: werden *Herausforderungen und Risiken* durch Testeinsatz wahrgenommen, leidet die *persönliche Einstellung ggü. dem Testeinsatz*.

Steigt die Bewertung betreffend die *Optimierung der Entscheidungsgüte* durch Testeinsatz um eine Einheit, also würde auf der Antwortskala eine ganze Stufe höher angewählt, so steigt die Einschätzung der *persönlichen Einstellung ggü. Tests* um 0.94 Stufen auf der Antwortskala. Einfach gesagt: sehen Personen die Möglichkeit zur Optimierung der Entscheidungsgüte durch den Einsatz psychologischer Tests, stärkt das ihre Einstellung ggü. dem Testeinsatz.

68% ( $korr. R^2 = .68$ ) der Varianz der persönlichen Einstellung ggü. Tests lässt sich durch die beiden unabhängigen Variablen *Optimierung der Entscheidungsgüte* und *Herausforderungen und Risiken* erklären. Um die Bedeutsamkeit der *korrigierten R<sup>2</sup>* ( $korr. R^2$ ) zu beurteilen, wurde die Effektstärke  $f^2$  nach Cohen berechnet. Dazu wurde die folgende Formel verwendet:

$$f^2 = \frac{korr. R^2}{1 - korr. R^2} = \frac{0.68}{1 - 0.68} \cong 2.13$$

Mit  $f^2 \cong 2.13$  lässt sich der Effekt, orientiert an Cohen (1988), als stark beschreiben.<sup>18</sup>

<sup>18</sup> Einteilung der Effektstärke  $f^2$  nach Cohen (1988):

$f^2 > 0.02$  = schwacher Effekt |  $f^2 > 0.15$  = mittlerer Effekt |  $f^2 > 0.35$  = starker Effekt

## 5 Diskussion

Dieses Kapitel widmet sich der Diskussion der vorangegangenen Kapiteln 1 bis 4. Zunächst werden die in Kap. 4 dargestellten Ergebnisse in Kurzform wiederholt, interpretiert und diskutiert (siehe Kap. 5.1). Mit der Beantwortung der Hypothesen, welche in Kap. 2 formuliert wurden, schliesst dieses Unterkapitel.

Die Ergebnisse machen die Formulierung von Gestaltungsempfehlungen notwendig. Diese werden im Unterkapitel 5.2 besprochen. In Kap. 5.3 werden Limitationen erörtert, die die Ergebnisse hinsichtlich der Repräsentativität möglicherweise einschränken. In Unterkapitel 5.4 folgt ein kurzes Schlusswort, welches den Abschluss der Diskussion und der gesamten Masterthesis bildet.

### 5.1 Interpretation und Beurteilung der quantitativen Ergebnisse

Entlang der in Kap. 1 genannten Fragestellungen werden nun die Ergebnisse besprochen, wobei zuerst die Beschreibung und Einordnung der Stichprobe folgt.

#### 5.1.1 Stichprobe

Etwas mehr als die Hälfte der Umfrageteilnehmenden stammt aus der Romandie. Ein Grossteil der Teilnehmenden verfügt über einen tertiären Ausbildungsabschluss, wobei Absolventinnen und Absolventen mit sehr verschiedenen beruflichen und akademischen Ausbildungen vertreten sind. Auch in Bezug auf die Berufserfahrung herrscht eine grosse Variabilität. Die Stichprobe umfasst Berufseinsteigende sowie Personen, die ihre gesamte berufliche Laufbahn der Personalarbeit gewidmet haben. Die Stichprobe zeigt Berufserfahrungen von null bis 34 Berufsjahren.

Wenn man die Anteile der grössten Wirtschaftsabteilungen der Schweiz betrachtet (vgl. Tab. 6), muss festgestellt werden, dass Wirtschaftsabteilungen, wie Handel, Gastgewerbe und Bauwesen in der Stichprobe unterdurchschnittlich vertreten sind. Hingegen sind die grössten Wirtschaftsabteilungen, namentlich *verarbeitendes Gewerbe* sowie *Gesundheit und Soziales*, adäquat vertreten.

Die Anteile an Teilnehmenden aus Grossunternehmen und Teilnehmenden aus KMUs sind in etwa gleich. Knapp ein Drittel der Befragten dürfen in ihren Unternehmen auf interne Unterstützung durch eine Psychologin bzw. einen Psychologen zurückgreifen.

Zusammenfassend überzeugt die Stichprobe mit einer hohen Variabilität hinsichtlich der demografischen Merkmale. Besonders erfreulich ist, dass die Romandie in der Stichprobe gut vertreten ist, da bisherige Studien kaum Daten zu dieser Region erheben konnten bzw. sich auf die Deutschschweiz oder auf Deutschland konzentriert hatten.

#### 5.1.2 Häufigkeit und Testarten (OF1 und OF2)

Beinahe zwei Drittel der Teilnehmenden berichten, dass sie in ihrem Unternehmen zum Zweck von Personalentscheidungen psychologische Tests einsetzen. Es ist anzunehmen, dass dies den wahren Anteil an Unternehmen in der Schweiz überschätzt, die psychologische Tests einsetzen.

Vergleichsweise berichten die Studien von Hell et al. (2021) oder Véron et al. (2024), dass ca. ein Drittel der Unternehmen Persönlichkeitstests einsetzen und knapp 7% Leistungstests nutzen.

Die Umfrage hat wahrscheinlich insb. HR-Mitarbeitende angezogen, die bereits über Erfahrung mit dem Einsatz psychologischer Tests verfügen. Es ist daher anzunehmen, dass vom erhaltenen Ergebnis betreffend die Anteile an testnutzenden Unternehmen nicht auf die Grundgesamtheit geschlossen werden darf.

Im Vergleich zu Leistungstests werden mehrheitlich Persönlichkeitstests eingesetzt. Grosser Beliebtheit erfreuen sich die typenbildenden Instrumente DISG, Insights und MBTI. Auch der MPA, der Wave und der BIP wurden häufig genannt. Zudem wurden viele Einzelmeldungen zu Testnamen verzeichnet, was die Vielfalt des Testangeboten widerspiegelt (vgl. Hell, 2013; ZIPD, 2024). Ähnliche Ergebnisse zu eingesetzten Tests zeigten sich auch bei den Untersuchungen von Berchtold (2005), Hossiep et al. (2015) und Hossiep et al. (2022).

Die Ergebnisse der Umfrage zeigen, dass Leistungstests kaum eingesetzt werden. 13% der Nennungen entsprechen Leistungstests, wobei insbesondere Tests zur allgemeinen Intelligenz genannt wurden. Andere Leistungstests, wie z.B. Tests zu Fachwissen werden nicht und Tests zu Konzentration nur vereinzelt eingesetzt.

### 5.1.3 Qualität (UF1)

Nur wenige der genannten Tests verfügen über eine Testrezension vom TBS-DTK oder vom SDBB. Angesichts der Tatsache, dass es mehr als 3'000 Persönlichkeitstests und 800 Leistungstests gibt (ZIPD, 2024) und das TBS-DTK lediglich 60, das SDBB lediglich 100 Rezensionen führt, ist das Resultat kaum verwunderlich.

Um die Qualität von Persönlichkeitstests wirklich beurteilen zu können, greift die Frage nach Rezensionen zu kurz. Zur Beurteilung der Qualität der eingesetzten Persönlichkeitstests soll das Schema herangezogen werden, welches bereits in Kap. 2.1.1.1 sinngemäss dargelegt wurde: Tests, die Ergebnisse als Taxonomien präsentieren, sind zur Beschreibung von Persönlichkeit ungeeignet. Wie die Ergebnisse zeigen, sind es jedoch diese Verfahren, die am meisten eingesetzt werden – Verfahren, die sich zumeist auf psychoanalytische Theorien berufen. Aufgrund deren fehlenden wissenschaftlichen Fundierung, müssen diese Tests hinsichtlich ihrer fehlenden Konstruktvalidität als mangelhaft bezeichnet werden.

Der durch das TBS-DTK und das SDBB positiv rezensierte BIP wird laut Ergebnissen ebenfalls gerne genutzt. Wie die Studie von Hossiep et al. (2022) aufzeigt, hat sich dessen Nutzung in den letzten Jahren mehr und mehr etabliert.

Intelligenztests, die moderate bis hohe Kriteriumsvalidität aufweisen (vgl. Tab. 2), werden laut Ergebnissen kaum eingesetzt.

Die qualitativen Ergebnisse (siehe Kap. 3.5.2) bieten entsprechende Erklärungsansätze für den Einsatz genannter Tests. Positive Erfahrungen mit Tests im Allgemeinen, aber auch mit spezifischen Testverfahren, sind wohl ausschlaggebend. Zudem wurde in den qualitativen Interviews genannt, dass die Handhabbarkeit von Tests ausschlaggebend sei für deren Einsatz. Insbesondere bei online Typentests werden Auswertung und Interpretation automatisch ausgeführt und ganze Rappports druckfertig ausgegeben, was von HR-Mitarbeitenden geschätzt wird.

Eine dritte mögliche Erklärung für den Einsatz von Tests mit mangelnder Güte findet sich in der qualitativen Kategorie *Trends und Verbreitung*. Auch König et al. (2010) haben sich dieser Thematik gewidmet. Die Flut an Testangeboten und die damit verbundene Unsicherheit führt zur Nachahmung anderer Unternehmen, was letztlich ihre Verwendung weiter fördert. Die zunehmende Verbreitung eines Tests erhöht demnach die Wahrscheinlichkeit seines Einsatzes.

#### 5.1.4 Zielgruppe (UF2)

Psychologische Tests werden insb. für die Auswahl von Führungskräften eingesetzt. Ein ähnliches Bild hat sich ebenfalls in den Studien von Armoneit et al. (2020) und Hossiep et al. (2015) gezeigt. Auch Mitarbeitende ohne Führungsfunktion werden gerne getestet. Manche der Umfrage-Teilnehmenden präzisierten im offenen Antwortfeld ihre Auswahl betreffend die Mitarbeitenden ohne Führungsfunktion und unterstrichen, dass sie psychologische Tests bei der Auswahl von fachlichen Spezialisten und Spezialistinnen einsetzen.

Kaum eingesetzt werden Tests laut den Ergebnissen bei ungelerten Mitarbeitenden, Auszubildenden oder Hochschulabsolventen und -absolventinnen.

Damit entsprechen die Ergebnisse betreffend die Führungskräfte der Studie von Armoneit et al. (2020) und Hossiep et al. (2015). Die Ergebnisse betreffend die Auszubildenden und die Hochschulabsolventinnen und -absolventen konnten jedoch nicht repliziert werden.

#### 5.1.5 Zweck (UF3)

59% der Befragten setzen psychologische Tests für die Personalauswahl und -entwicklung ein. 27% nutzen Tests ausschliesslich zur Personalauswahl und 14% setzen Tests lediglich für Entwicklungszwecke ein. Damit zeichnen die vorliegenden Ergebnisse ein anderes Bild, als die Ergebnisse von Hossiep et al. (2015). Ihre Ergebnisse zeigten, dass psychologische Tests mehrheitlich für Entwicklungszwecke eingesetzt werden. Es stellt sich nun die Frage, ob das durch regionale Unterschiede (Schweiz – Deutschland) erklärt werden kann oder ob die vorliegenden Daten zu wenig repräsentativ sind, um damit auf die Grundgesamtheit schliessen zu können.

#### 5.1.6 Informationsbasis (UF4)

Die befragten HR-Mitarbeitenden nutzen diverse Informationsquellen, um sich über einzusetzende Tests zu informieren. Am häufigsten genannt wurde der Austausch mit Fachexpertinnen und -experten. Ebenfalls informieren sich die Befragten durch den Austausch mit den Testanbietenden oder durch Kurse. Vergleichsweise werden Testrezensionen oder Fachliteratur kaum zur Information genutzt.

Dass Testrezensionen wenig berücksichtigt werden, lässt sich wohl am ehesten dadurch erklären, dass das TBS-DTK oder das SDBB nur eine geringe Menge des vielfältigen Testangebots abdecken. Hingegen ist die Sammlung von Testrezensionen vom ZPID ausufernd und dadurch unübersichtlich. Der Situation abträglich ist auch die Tatsache, dass zu einem Test verschiedene Rezensionen bestehen.

Dass Fachliteratur nur wenig genutzt wird, widerspiegelt mit grosser Wahrscheinlichkeit den Science-Practitioner-Gap, welcher in Kap. 2.3 thematisiert wurde.

Die Hauptinformationsquellen zu Tests sind der direkte Austausch mit den Testanbietenden und / oder der Austausch mit Fachexpertinnen und -experten. Vor dem Hintergrund, dass vor allem Tests mit fraglicher Qualität eingesetzt werden, ist die Informationsquelle *Austausch mit Testanbietenden* äusserst kritisch zu werten. Die Beziehung zwischen Testanbietenden und Testanwendenden scheint komplex<sup>19</sup>: Expertinnen und Experten, die während einer Weiterbildung der FHNW die Resultate diskutiert hatten, erklärten sich die rege Nutzung mangelhafter Instrumente u.a. dadurch, dass sich Testanbietende stark an ihre Kundschaft und damit an die Unternehmen binden. Dadurch wird nicht nur ein Wechsel zu einem anderen Anbietenden erschwert (N.E., persönl. Mitteilung, 14.06.24). Die

---

<sup>19</sup> Es bestand die Möglichkeit, die quantitativen Ergebnisse der vorliegenden Thesis bereits vor Abgabe durch HR-Expertinnen und -Experten im Rahmen einer Weiterbildung an der FHNW diskutieren zu lassen.

Vermutung liegt nahe, dass durch Anbietende keine ungefilterten, sachlichen Informationen erhalten werden können.

Websuchen werden kaum als Informationsquellen genutzt, was als positiv bewertet werden muss. „16PF“, „DISG-Modell“, „Welches Tier steckt in Dir?“ – das sind die ersten drei Treffer bei Google, wenn man nach Persönlichkeitstests sucht. Der vierte und damit einzige Treffer auf der ersten Suchseite, der sich angemessen mit der Thematik der Persönlichkeitstestung auseinandersetzt, ist eine Seite der Universität Leipzig, die eine kostenlose Persönlichkeitserfassung entlang des Big Five-Modells anbietet.

Bittet man ChatGPT nach Testempfehlungen für die Personalauswahl, folgen nach dem NEO-PI-R sogleich der MBTI und der Hogan Personality Inventory (HPI) (siehe Anhang S). Eine angemessene Informationsbasis stellen Websuchen oder ChatGPT in diesen Belangen demnach nicht dar.

Zusammenfassen lässt sich sagen: Rezensionsdatenbanken bieten entweder zu wenige (TBS-DTK, SDBB) oder zu viele Rezensionen (ZPID) an, um sich einen Überblick über die Testlandschaft verschaffen zu können. Der Austausch mit Testanbietenden muss hinsichtlich des Informationsgehalts als gefärbt und damit als eingeschränkt betrachtet werden. Websuchen bieten keine angemessenen Informationen. Es darf also keinesfalls unterschätzt werden, wie schwer sich die adäquate Informationsbeschaffung zu psychologischen Tests gestaltet.

### 5.1.7 Hintergrund (Einflussfaktoren, UF5 bis 7)

Die Spearman-Korrelationen zwischen der abhängigen Variablen *Testeinsatz* und den unabhängigen Variablen *persönliche Einstellung* (Beweggrund), *Optimierung der Entscheidungsgüte* (Entscheidungskriterium), *Praktikabilität* und *Entlastung* (Entscheidungskriterium) und *kontextuelle Einflussfaktoren* (äusserer Einflussfaktor) zeigen sich signifikant positiv. Die Korrelation zwischen *Testeinsatz* und der Variablen *Risiken, Herausforderungen und Risiken* (Entscheidungskriterium) zeigt sich signifikant negativ.

Berechnungen binär logistischer Regressionen konnten drei signifikante einflussnehmende Faktoren auf die abhängige Variable *Testeinsatz* aufzeigen: persönliche Einstellung, Herausforderungen und Risiken und Einfluss der eigenen Entscheidung.

Je positiver die persönliche Einstellung der HR-Mitarbeitenden ggü. psychologischen Tests ist, desto eher werden sie in ihrem Unternehmen eingesetzt. Die persönliche Einstellung ist demnach von zentraler Bedeutung für den *Testeinsatz*. Erklärt werden kann das durch die TRA nach Fishbein und Ajzen (1975), welche bereits im Kap. 3.2.5 zu den qualitativen Ergebnissen hinzugezogen wurde. Verhalten wird durch Verhaltensintention beeinflusst. Verhaltensintention wird u.a. durch die persönliche Einstellung beeinflusst, wie sich das auch in den vorliegenden Ergebnissen zeigt.

Die Wahrnehmung von Herausforderungen und Risiken in Bezug auf *Testeinsatz* verringert die Wahrscheinlichkeit, dass im Unternehmen *Testeinsatz* erfolgt. Demnach hat eine Auseinandersetzung mit der Thematik stattgefunden, wobei die Entscheidung, auf Tests zu verzichten demnach getroffen wurde, dass einzusetzende Ressourcen und zu befürchtende Risiken höher wiegen, als ein allfälliger Nutzen durch *Testeinsatz*. Insbesondere vor dem Hintergrund des Personalmangels und der Befürchtung von fehlender Akzeptanz ist dieses Resultat nachzuvollziehen.

Die Wahrscheinlichkeit von Testeinsatz hängt auch davon ab, ob die Meinung der HR-Mitarbeitenden in Bezug auf psychologische Tests überhaupt Einfluss auf die Entscheidung zum Testeinsatz hat. Das Resultat ist nicht nur deshalb positiv zu werten, weil es zeigt, dass die befragten HR-Mitarbeitenden partizipieren dürfen. Diese Variable dient auch als Kontrolle, ob die erhobenen unabhängigen Variablen aus rein kontextueller Sicht überhaupt Einfluss auf die personalstrategische Entscheidung für oder gegen den Testeinsatz nehmen können.

Nun kann nach dem TRA-Modell von Fishbein und Ajzen (1975) angenommen werden, dass die Einstellung als Mediator für Verhaltensintention und schliesslich Verhalten. Daher wurde, in explorativem Sinne, ein weiteres Regressionsmodell gerechnet. So zeigten sich die unabhängigen Variablen *Optimierung der Entscheidungsgüte*, *Herausforderungen und Risiken* und der *Kontext* als signifikante Einflussfaktoren auf die *persönliche Einstellung*. Das bedeutet, dass die Zustimmung betreffend die Variablen *Optimierung der Entscheidungsgüte* eher zu einer positiven *Einstellung* führt, währenddessen hohe Zustimmungen betreffend die Variablen *Herausforderungen und Risiken* zu einer negativen *Einstellung* führt. Zudem zeige sich auch der *Kontext* als signifikanter einflussnehmender Faktor – zumindest auf die *Einstellung* betreffend Testeinsatz. König et al. (2010) konnten diesen in ihrer Studie als einflussnehmenden Faktor auf den Testeinsatz darlegen.











Die Ergebnisse offenbaren eine gewisse Widersprüchlichkeit. Einerseits wird der Einsatz psychologischer Tests geschätzt aufgrund ihrer Güte bzw. aufgrund dessen, dass sie die Güte des Entscheidungsprozesses verbessern. Andererseits werden Tests eingesetzt, die gerade wegen ihrer Güte in der Kritik stehen. Dieser Umstand unterstreicht die Notwendigkeit von Gestaltungsempfehlungen. Nachfolgend werden nun die Hypothesen beantwortet (siehe Kap. 5.1.8). Daran anschliessend werden die Gestaltungsempfehlungen formuliert (siehe Kap. 5.2).

## 5.1.8 Beantwortung der Hypothesen

Entlang der besprochenen Ergebnisse können die meisten der in Kap. 2.4 aufgestellten Hypothesen bestätigt werden. Die Tab. 20 zeigt die Beantwortung der Hypothesen und lässt damit die gerade eben besprochenen Ergebnisdarstellungen und -interpretationen noch einmal Revue passieren.

Tabelle 20

Beantwortung der Hypothesen, die in Kap. 2.4 aufgestellt wurden. Häkchen = Annahme, Kreuz = Ablehnung der jeweiligen Hypothese.

Beantwortung der Hypothesen zu den Oberfragestellungen (OF)	
Beantwortung der Hypothesen zur OF 1 - Verwendungshäufigkeit	
	H1.1: 62% der in der Schweiz tätigen Unternehmen setzen zum Zweck der Personalauswahl und -entwicklung psychologische Tests ein.
	H1.2: Vergleichsweise werden mehr Persönlichkeitstests als Leistungstests eingesetzt. 88% aller eingesetzten Tests entsprechen Persönlichkeitstests, lediglich 12% entsprechen Leistungstests.
Beantwortung der Hypothese zur OF2 - Testart	
	H1.3: In der Schweiz tätige Unternehmen setzen verschiedenste psychologische Tests im Rahmen von Personalentscheidungen ein. Befragte nannten 51 verschiedene Tests.
Beantwortung der Hypothesen zu den Unterfragestellungen (UF)	
Beantwortung der Hypothese zur UF1 - Qualität eingesetzter Tests	
	H1.4: Lediglich 13% der genannten Tests verfügen über eine Testrezension durch das TBS-DTK und / oder SDBB. Eine Vielzahl der eingesetzten Tests verfügt über keine Testrezension der genannten Datenbanken und steht betreffend der Qualität zudem in der Kritik.
Beantwortung der Hypothese zur UF2 - Einsatzzweck	
	H1.5: Psychologische Tests werden dort, wo sie eingesetzt werden, bei 59% für Personalauswahl und -entwicklung eingesetzt. In 27% werden sie nur zum Zweck der Personalauswahl und in 14% nur zur Personalentwicklung eingesetzt. Diese Hypothese muss verworfen werden. Tests werden laut vorliegenden Ergebnissen eher zur Personalauswahl als zur Personalentwicklung eingesetzt.
Beantwortung der Hypothese zur UF3 - Zielgruppen	
	H1.6: Je höher die vorgesehene Hierarchiestufe der zu rekrutierenden Person, desto eher werden Tests zur Personalauswahl eingesetzt. Insbesondere Bewerbende für Führungspositionen oder fachliche Schlüsselpositionen werden getestet.
Beantwortung der Hypothese zur UF4 - Informationsbasis	
	H1.7: Durch HR-Mitarbeitende werden verschiedenste Informationsquellen genutzt, um sich über einzusetzende Tests zu informieren, wobei der Austausch mit Fachpersonen und Testanbietenden vergleichsweise am häufigsten genutzt wird.
Beantwortung der Hypothesen zu UF5 bis UF7 - Hintergründe in Bezug auf Testeinsatz	
	H1.8: Persönliche Einstellung der HR-Mitarbeitenden hat Einfluss auf die Wahrscheinlichkeit für Testeinsatz im Unternehmen.
	H1.9: Das Entscheidungskriterium betreffend die wahrgenommenen Herausforderungen und Risiken durch Testeinsatz nimmt Einfluss auf die Wahrscheinlichkeit für Testeinsatz im Unternehmen. Andere Entscheidungskriterien zeigen zwar signifikante Korrelationen, aber keine signifikanten Regressionskoeffizienten.
	H1.10: Äussere Einflussgrössen beeinflussen die persönliche Einstellung der HR-Mitarbeitenden ggü. psychologischen Tests, zeigen sich aber im Regressionsmodell nicht als einflussnehmend auf Testeinsatz. Manche der Variablen zu äusseren Einflussgrössen zeigen signifikante Korrelationskoeffizienten. Hypothese muss verworfen werden. Äussere Einflussgrössen haben laut vorliegenden Ergebnissen keinen Einfluss auf Testeinsatz. Ein indirekter Einfluss lässt sich vermuten, aber nicht statistisch bestätigen.

## 5.2 Gestaltungshinweise

Die nachfolgenden Gestaltungshinweise konzentrieren sich auf die Verbesserung der Qualitätsproblematik. Dabei könnten die Zentralisierung von Testrezensionen und die Sensibilisierung der Testanwendenden hinsichtlich der Testqualität hilfreich sein. Die nachfolgenden Unterkapitel gehen auf diese Punkte ein. Aufgrund des spezifischen Forschungsinteresses sind detaillierte Gestaltungshinweise nicht möglich, jedoch allgemeine Hinweise und Vorschläge.

### 5.2.1 Zentralisieren von Testrezensionen

Aktuell gibt es zu deutschsprachigen Tests Rezensionen, welche in verschiedenen Datenbanken verstreut sind. Wie in Kap. 2.2.3 erwähnt, befinden sich bei SDBB 100 Rezensionen, bei TBS-DTK 60 und beim ZPID Tausende von Testrezensionen. Dies führt zu einem Verlust an Übersicht und Ressourcen, da die Daten nicht geordnet und dadurch schwer zugänglich sind.

#### Harmonisierung der Rezensionssysteme und Teilen von Ressourcen

Obwohl es das Testsystem der *European Federation of Psychologists Associations* (EFPA) gibt, an welchem sich das *British Psychological Society* (BPS), das *Nederlands Instituut van Psychologen* (NIP) und das TBS-DTK orientieren (Kersting, 2014), sieht das SDBB davon ab und orientiert sich nach eigenen Aussagen an den wichtigsten Gütekriterien (persönl. Mitteilung, M.B., 21.05.24). Durch die fehlende Harmonisierung können Synergien u.a. zwischen Deutschland und der Schweiz, bzw. dem TBS-DTK und dem SDBB, nicht genutzt werden. Beide Institute haben den Auftrag, Tests zu rezensieren, um die Gesellschaft vor unseriösen Instrumenten zu schützen. Dennoch führen das TBS-DTK und das SDBB parallel zueinander Rezensionen zu denselben Tests durch. So wurden bspw. der BIP oder der NEO-FFI durch beide Institute rezensiert. Es wäre sinnvoll, würden die beiden Institute ihre Kräfte bündeln und parallele Arbeit vermeiden, um so Ressourcen zu sparen, die in andere Qualitätsverbesserungen einfließen könnten.

Die Herausforderungen betreffend die Harmonisierung besteht wohl im Umstand, dass die Koordination und Kooperation zwischen verschiedenen Instituten und Datenbanken intensive Zusammenarbeit erfordern. Die Vereinbarung gemeinsamer Standards wäre unumgänglich. Es ist denkbar, dass unterschiedliche Interessen der Institute und Datenbankbetreibenden die Zusammenarbeit erschweren oder gar verunmöglichen könnten.

#### Konzept von Meta-Testrezensionen aufnehmen

Die in ZPID enthaltenen Rezensionen sind zumeist unabhängig erstellt worden. Die 3'738 erstellten Testrezensionen gehen in der Masse unter und erhalten nicht die Aufmerksamkeit, die sie verdienen. Zudem wurde auch hier parallele Arbeit geleistet: die 3'738 Rezensionen wurden für 1'663 Testverfahren erstellt (ZPID, 2023). In der Praxis stellt sich dann die Frage, an welcher der verschiedenen Rezensionen man sich für einen bestimmten Test orientieren soll.

Ähnlich dem Konzept von Metaanalysen, könnte die Erstellung von Meta-Testrezensionen in Betracht gezogen werden. Es wäre bei der aktuellen Datenlage des ZPID angezeigt, bereits erstellte Rezensionen durch die Ergebnisse weiterer Rezensionen zu extrapolieren und damit bereits geleistete Arbeit weiter zu verwenden. Eine Zusammenarbeit mit psychologischen Hochschulen wäre denkbar, um die Erstellung solcher Meta-Testrezensionen bspw. als Masterarbeiten zu vergeben. So könnte der Thematik um psychologische Tests und Rezensionen zudem neuer Schwung verliehen werden.

Eine mögliche Herausforderung betreffend die Meta-Testrezensionen ist wohl das Urheberrecht der bereits erstellten Testrezensionen. Rechtliche Rahmenbedingungen müssten demnach vor der Verwendung von Rezensionen geklärt werden.

### **Sprachliche und regionale Barrieren überwinden**

Psychologische Tests werden zumeist in verschiedenen Sprachversionen entwickelt. Da sich die Testrezensionssysteme des BPS, des NIP oder des TBS-DTK sehr ähnlich sind und sich alle nach der ISO 9000ff orientieren, wäre es auch hier sinnvoll, Synergien zu nutzen und erstellte Testrezensionen zu übersetzen. Die Überwindung sprachlicher Barrieren ist die wohl einfachste und effizienteste Methode die Testrezensionsdatenbanken des TBS-DTK oder des SDBB mit qualitativ hochwertigen Testrezensionen zu füllen. Neben Übersetzungsdiensten bieten sich dafür auch maschinelle Übersetzungsdienste als kostengünstige Alternativen an.

## **5.2.2 Informieren und Sensibilisieren der Testanwendenden**

Verschiedene Vereine in der Schweiz und in Deutschland engagieren sich für die Sensibilisierung und Information in Bezug auf eignungsdiagnostische Qualitätsstandards (siehe Kap. 2.2.3, bzw. Fellner, 2019). Durch Vorträge, Tagungen und Informationsveranstaltungen werden HR-Mitarbeitende und Interessierte über aktuelle Erkenntnisse informiert. Das BPS oder das NIP bieten postgraduale Lizenzierungen für testanwendende Psychologinnen und Psychologen an und setzen damit auf den kompetenzorientierten Ansatz (Kersting, 2014). Das DIN-Portal bietet Lizenzierungen an, wenn Personen sich mit dem DIN33430 adäquat auseinandergesetzt haben.

Problematisch ist, dass die meisten Testanwendenden nicht über eine psychologische Ausbildung verfügen, wodurch die Idee der postgradualen Lizenzierung nicht an der Quelle der Problematik ansetzt. Des Weiteren erreichen Vorträge oder Tagungen nur einen Bruchteil der Testanwendenden. Zudem sind solche Veranstaltungen mit Kosten und Zeitaufwand verbunden. So ist denkbar, dass dies bereits eine erste Hürde darstellt für potenzielle Teilnehmende.

Um also das Ziel der Information und Sensibilisierung verfolgen, eine Grosszahl der Testanwendenden erreichen zu können und die Kosten für Vereine und Testanwendende niedrig zu halten, würde sich das Verfassen und regelmässige Teilen von Beiträgen auf Karriereportalen wie z.B. LinkedIn oder in Fachzeitschriften anbieten. Hier könnten aktuelle Erkenntnisse und Best-Practice-Beispiele verbreitet werden, um eine breite Zielgruppe zu erreichen.

## 5.3 Limitationen in der Methodik

Trotz aller Erkenntnisse, die gewonnen werden konnten, müssen auch die Limitationen der Erhebung besprochen werden.

### 5.3.1 Repräsentativität hinsichtlich Stichprobenszusammensetzung

Aufgrund der Ad-Hoc-Stichprobe und des Anspruchs auf Anonymität lässt sich die Zusammensetzung der Stichprobe nicht näher bestimmen, was die Beurteilung ihrer Repräsentativität erschwert. Es ist denkbar, dass mehrere HR-Mitarbeitende desselben Unternehmens an der Umfrage teilgenommen haben, was die Ergebnisse bezüglich der Häufigkeit eingesetzter Tests verzerren könnte.

Da Personalauswahlprozedere von Unternehmen streng gehütet werden, lag die Vermutung nahe, dass eine direkte Erhebung bei den Unternehmen ohne die Möglichkeit zur Anonymität der Datenerhebung abträglich sein könnte. Zudem hätte das direkte Anschreiben von mehr als 11'000 Unternehmen (vgl. Stichprobenschätzung in Tab. 10) aus Zeitgründen kaum bewerkstellig werden können. Daher wurde ein offenes Sampling gewählt, was jedoch die Repräsentativität der Daten zu Nutzungshäufigkeit limitiert.

Eine Überschätzung des Anteils an Testanwendenden ist wahrscheinlich. Dies kann einerseits durch die gerade eben genannten Gründe erklärt werden. Andererseits ist es denkbar, dass sich insb. Testanwendende von der Umfrage angesprochen fühlten.

### 5.3.2 Repräsentativität betreffend die Stichprobengröße

Eine weitere Limitation dieser Thesis betrifft die Stichprobengröße, die trotz sorgfältiger Planung kleiner ausfiel als geplant. Die in Kap. 3.3.2 berechnete Stichprobengröße von  $n = 385$  und die damit angestrebte Fehlermarge, die laut Döring und Bortz mit 5% gerade noch zumutbar ist, konnte nicht unterschritten werden. Die Fehlermarge, die angenommen werden muss, beträgt 7.2%, was die Generalisierbarkeit der Ergebnisse einschränken könnte, die einer Auszählung gefolgt sind (OF1, OF2, UF1 bis UF4). Die entsprechende Berechnung der Fehlermarge wird nachfolgend dargestellt:

- Z-Wert für 95%-Konfidenzintervall ( $z$ ) = 1.96
- Standardabweichung ( $p$ ) = 0.5
- Stichprobengröße ( $n$ ) = 183
- gesucht: Fehlermarge ( $e$ )

$$e = z \times \sqrt{\frac{p \times (1-p)}{n}} \rightarrow 100 \times 1.96 \times \sqrt{\frac{0.5 \times (1-0.5)}{183}} \approx 7.2\%$$

### 5.3.3 Unterschiede in Regressionskoeffizienten: binär vs. linear

Ein weiterer Punkt betrifft die methodische Einschränkung der verwendeten Regressionsmodelle. Es besteht eine signifikante Korrelation ( $r_s = .533$ ) und Regression ( $OR = 2.39$ ) zwischen persönlichen Einstellungen und der abhängigen Variablen Testeinsatz. Das deutet darauf hin, dass persönliche Einstellung möglicherweise ein Mediator und nicht (nur) eine unabhängige Variable ist. Dies steht auch im Einklang mit dem TRA-Modell von Fishbein und Ajzen (1975). Allerdings lassen die vorliegenden Daten keine definitive Überprüfung dieser Vermutung zu, was die Aussagekraft des vermuteten Zusammenhangs natürlich limitiert.

Die Unterschiede zwischen den Regressionsmodellen – eines binär logistisch und eines linear – beeinflussen nicht nur die Interpretation der Zusammenhänge, sondern auch die Ergebnisse selbst. Zudem ist der Common Method Bias (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003) zu berücksichtigen. Backhaus et al. (2008) verweisen darauf, dass das Nagelkerkes  $R^2$  mit dem

Bestimmtheitsmass einer linearen Regression vergleichbar sei. Doch das Bestimmtheitsmass ( $R^2 = .68$ ) ist höher als das Nagelkerkes  $R^2$  ( $R^2_{Nagelkerkes} = .57$ ). Dies könnte auf den Common Method Bias zurückzuführen sein, der natürliche Korrelation bei gleicher Abfrageart beschreibt (Podsakoff et al., 2003).

Die Berechnung des linearen Bestimmtheitsmasses für abhängige und unabhängige Variablen basierte auf Selbsteinschätzungen auf einer sechsstufigen Skala. Es ist anzunehmen, dass das Antwortverhalten der Personen hier bereits zu Korrelationen geführt hat, die von tatsächlichen Assoziationen zwischen den abhängiger und unabhängiger Variablen zu unterscheiden sind.

Im Gegensatz dazu wurde für die Berechnung der Nagelkerkes  $R^2$ , also dem binär logistischen «Bestimmtheitsmass», lediglich die Abfrage der unabhängigen Variable durch Selbsteinschätzung auf einer sechsstufigen Skala verlangt. Die abhängige Variable wurde dichotom abgefragt. Aufgrund ungleicher Methode kann hier von Abwesenheit des Common Method Bias ausgegangen werden. Daher könnten die höheren Werte des linearen Bestimmtheitsmasses  $R^2$  im Vergleich zum binär logistischen Nagelkerkes  $R^2$  auf den Common Method Bias zumindest teilweise zurückzuführen sein.

### 5.3.4 Verzicht auf weitere explorative Auswertungen

Der quantitative Datensatz würde Vergleiche zwischen Gruppen zulassen. Von Interesse wären regionale Vergleiche (Romandie – Deutschschweiz, ähnlich zur Studie von Hossiep et al., 2022) oder Vergleiche zwischen KMUs und Grossunternehmen bezüglich der eingesetzten Tests. Beide genannten Gruppenvergleiche hätten den Vorteil, dass die Gruppen in etwa gleich viele Teilnehmende umfassen. Darauf wurde verzichtet bzw. musste verzichtet werden, weil die vorliegende Thesis bereits an der Grenze der maximal erlaubten Anzahl an Zeichen kratzt. Damit bieten die erhobenen quantitativen Daten die Möglichkeit zur weiteren Forschung.

## 5.4 Schlusswort

Trotz aller genannten Limitationen zeichnen die Daten in Bezug auf eingesetzte Tests ein deutliche Bild ab: v.a. der Einsatz von Persönlichkeitstest ist in der Schweiz verbreitet. Problematisch ist dabei, dass häufig mangelhafte Persönlichkeitstests eingesetzt werden, die Testpersonen typisieren. Dadurch wird mit dem Konzept der Persönlichkeit unseriös umgegangen.

Die Flut an Testangeboten ist Nährboden für Anwendungen von Heuristiken, auf welche Menschen zurückgreifen, wenn sie Entscheidungen unter Unsicherheit treffen müssen. Man kann annehmen, dass sich Tests wie der MBTI oder der Insights verbreiten, weil sie sich auf die Theorien des bekannten Psychoanalytikers C.G. Jung stützen. Dass diese Theorien zwischenzeitlich überholt sind, wird kaum wahrgenommen.

Vielleicht liegt es aber auch an der menschlichen Präferenz, durch Kategorisierung mit hoher Komplexität zurechtzukommen – typisierende Tests einzusetzen, um die hohe Komplexität von Persönlichkeit zu reduzieren.

Gleichwohl – es ist unumgänglich, über psychologische Tests weiter zu informieren und auf ebengenannte Qualitätsproblematik zu sensibilisieren. Die vorliegende Masterthesis soll dazu einen Impuls in diese Richtung geben.

## 6 Literaturverzeichnis

- Armoneit, C., Schuler, H. & Hell, B. (2020). Nutzung, Validität, Praktikabilität und Akzeptanz psychologischer Personalauswahlverfahren in Deutschland 1985, 1993, 2007, 2020: Fortführung einer Trendstudie. *Zeitschrift für Arbeits- und Organisationspsychologie*, 64(2), 67-82. doi: 10.1026/0932-4089/a000311
- Asendorpf, J. (2019). *Persönlichkeitspsychologie für Bachelor* (3. Aufl.). Heidelberg: Springer.
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2008). *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung* (12. Aufl.). Heidelberg: Springer.
- Baumeister, R., Bratslavsky, E., Finkenauer, C. & Vohs, K. (2001). Bad is stronger than good. *Review of general psychology*, 5(4), 323-370. doi: 10.1037/1089-2680.5.4.323
- BAV (2024). *Psychologische Tauglichkeitsuntersuchungen*.  
<https://www.bav.admin.ch/bav/de/home/rechtliches/rechtsgrundlagen-vorschriften/richtlinien/richtlinien-bahn/psychologische-tauglichkeitsuntersuchungen.html>
- Beauducel, A. & Leue, A. (2014). *Psychologische Diagnostik*. Göttingen: Hogrefe.
- Berchtold, M. (2005). Häufigste Auswahlverfahren in der Personalselektion. *HR Today*, 12, 37-39.
- BFS (2023a). *Sprachen*. <https://www.bfs.admin.ch/bfs/de/home/statistiken/bevoelkerung/sprachen-religionen/sprachen.html>
- BFS (2023b). *Beschäftigte nach Vollzeitäquivalenten und Wirtschaftsabteilungen*.  
<https://www.bfs.admin.ch/bfs/de/home/statistiken/industrie-dienstleistungen/unternehmen-beschaeftigte/beschaeftigungsstatistik/beschaeftigte.assetdetail.27165007.html>
- BFS (2023c). *Marktwirtschaftliche Unternehmen nach Wirtschaftsabteilungen und Grössenklasse im Jahr 2021*. <https://www.bfs.admin.ch/bfs/de/home/statistiken/industrie-dienstleistungen/unternehmen-beschaeftigte.assetdetail.26665235.html>
- BFS (2024). *Berufliche Mobilität*. <https://www.bfs.admin.ch/bfs/de/home/statistiken/arbeits-erwerb/erwerbstaetigkeit-arbeitszeit/erwerbsbevoelkerung/berufliche-mobilitaet.html>
- Birkeland, S., Manson, T., Kisamore, J., Brannick, M. & Smith, M. (2006), A Meta-Analytic Investigation of Job Applicant Faking on Personality Measures. *International Journal of Selection and Assessment*, 14, 317-335. doi: 10.1111/j.1468-2389.2006.00354.x
- Bogner, K. & Landrock, U. (2015). *Antworttendenzen in standardisierten Umfragen (Version 1.1). (GESIS Survey Guidelines)*. Mannheim: GESIS - Leibniz-Institut für Sozialwissenschaften. doi: 10.15465/gesis-sg\_016
- Bühl, A. (2016). *SPSS 23. Einführung in die moderne Datenanalyse* (15. Aufl.). Hallbergmoos: Pearson.
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion* (3. Aufl.). Hallbergmoos: Pearson.

- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2. Aufl.). Hillsdale, NJ: Lawrence Erlbaum.
- Diagnostik- und Testkuratorium. (2018). TBS-DTK. Testbeurteilungssystem des Diagnostik- und Testkuratoriums der Föderation Deutscher Psychologinnenvereinigungen. *Psychologische Rundschau*, 69, 109-116. doi: 10.1026/0033-3042/a000401
- Döring, N. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5. Aufl.). Heidelberg: Springer.
- Duden (2023). *Beweggrund*. <https://www.duden.de/rechtschreibung/Beweggrund>
- Eckardt, H. (1973). Psychologische Eignungsprognosen: Ihre Stellung in einem Gesamtsystem arbeitsmarktorientierter und eignungsorientierter Beratungsinformationen. *Mitteilungen aus der Arbeitsmarkt-und Berufsforschung*, 6(2), 177-189.
- Eichhorn, K. (2019). *Item und Skala: Empirische Untersuchungen zur Gültigkeit psychologischer Messungen anhand physikalischer Merkmale*. Unveröffentlichte Dissertation, Ludwig-Maximilians-Universität München.
- Eid, M., Gollwitzer, M. & Schmitt, M. (2017). *Statistik und Forschungsmethoden*. Weinheim: Beltz.
- Fellner, K. (2019). *Moderne Personalauswahl*. Wiesbaden: Springer.
- Fishbein, M. & Ajzen, I. (1975). *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*. Reading: Addison-Wesley.
- Guest, G. Namey, E. & Chen, M. (2020). A simple method to assess and report thematic saturation in qualitative research. *PLoS One*, 15(5). doi: 10.1371/journal.pone.0232076
- Hartig, J., Frey, A. & Jude, N. (2008). Validität. In: H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 135-163). Berlin: Springer.
- Hartzinger, R. & Nagel, H. (2013). *Statistik mit SPSS* (2. Aufl.). München: Pearson.
- Helfferrich, C. (2011). *Die Qualität qualitativer Daten. Manual für die Durchführung qualitativer Interviews* (4. Aufl.). Wiesbaden: Springer.
- Hell, B. (2013). Mehr Licht im Persönlichkeitstest-Dschungel. *HR Today*, 9. <https://www.hrtoday.ch/de/article/mehr-licht-im-persoenlichkeitstest-dschungel>
- Hell, B., Pässler, K. & Nido, M. (2023). Wissenschafts-Praxis-Gap in der Eignungsdiagnostik: Die Validität von Personalauswahlverfahren aus Sicht von HR-Fachpersonen im Vergleich zu metaanalytischen Erkenntnissen. *Wirtschaftspsychologie*, 25(2). <https://irf.fhnw.ch/handle/11654/37760>
- Hell, B., Vögeli, S., & Hermann, M. (2021). Employer Branding in Schweizer Unternehmen. *Personal Schweiz*, 3, 18-20. doi: 10.5281/zenodo.4695935
- Hemmerich, W. (2018). *StatistikGuru: Multivariate Normalverteilung online prüfen*. <https://statistikguru.de/rechner/multivariate-normalverteilung.html>

- Herrmann, A., Goldstein, D., Stadler, R., Landwehr, J., Heitmann, M., Hofstetter, R. & Huber, F. (2011). The effect of default options on choice-Evidence from online product configurators. *Journal of Retailing and Consumer Services*, 18(6), 483-491. doi: 10.1016/j.jretconser.2011.06.005
- HFG. (2023). *Bundesgesetz über die Forschung am Menschen*. <https://www.fedlex.admin.ch/eli/cc/2013/617/de>
- Hogrefe Testzentrale. (2024). *MSCEIT*. <https://www.testzentrale.ch/shop/mayer-salovey-caruso-test-zur-emotionalen-intelligenz-69909.html>
- Hogrefe. (2022, 19. April). Eignungsdiagnostik mit Intelligenztests im HR-Bereich. *Themen*. [https://www.hogrefe.com/ch/thema/eignungsdiagnostik-mit-intelligenztests-im-hr-bereich-1?fbclid=IwZXh0bgNhZW0CMTEAAR01ordSj\\_fPpwt31wb7TPNAdHJC2Jz-E666jhDX1UkVchv8z23G1B4AZ9c\\_aem\\_bPuQdU\\_17--ru9rIBGmJOQ](https://www.hogrefe.com/ch/thema/eignungsdiagnostik-mit-intelligenztests-im-hr-bereich-1?fbclid=IwZXh0bgNhZW0CMTEAAR01ordSj_fPpwt31wb7TPNAdHJC2Jz-E666jhDX1UkVchv8z23G1B4AZ9c_aem_bPuQdU_17--ru9rIBGmJOQ)
- Holtbrügge, D. (2004). *Personalmanagement*. Heidelberg: Springer. doi: 10.1007/978-3-662-09237-8\_1
- Hossiep, R. & Mühlhaus, O. (2005). *Personalauswahl und -entwicklung mit Persönlichkeitstests*. Göttingen: Hogrefe.
- Hossiep, R., Weiss, S., Netzer, M. & Hossiep, R. (2022). Personalauswahl: Typentests noch immer weit verbreitet. *Wirtschaftspsychologie heute*, 12.
- Hossiep, R., Weiss, S., Shecke, J. (2015). *Ist Persönlichkeit gefragt? – Eine Erhebung unter den grössten deutschen Unternehmen*. Forschungsbericht. Bochum: Ruhr-Universität, Projektteam Testentwicklung.
- Hunter J. & Hunter R. (1984). Validity and utility of alternate predictors of job performance. *Psychological Bulletin*, 96(1), 72-98. doi: 10.1037/0033-2909.96.1.72
- IBM (2021). *Binäre logistische Regression – Voraussetzungen*. <https://www.ibm.com/docs/de/spss-statistics/beta?topic=regression-binary-logic>
- Iller, M., Mäder, R. & Schreiber, M. (2021). *Handbuch Fragebogen zur Erfassung der Persönlichkeit (HEXACO-PI-R1)*. [https://www.laufbahndiagnostik.ch/downloads/de/Handbuch\\_HEXACO-PI-R1.pdf](https://www.laufbahndiagnostik.ch/downloads/de/Handbuch_HEXACO-PI-R1.pdf)
- Kahlke, E. & Schmidt, V. (2004). *Handbuch Personalauswahl*. Heidelberg: Economica.
- Kanning, U. (2015) *Personalauswahl zwischen Anspruch und Wirklichkeit. Eine wirtschaftspsychologische Analyse*. Heidelberg: Springer.
- Kauffeld, S. & Grohmann, A. (2011). *Arbeits-, Organisations- und Personalpsychologie für Bachelor*. Heidelberg: Springer.
- Kerstin, U. (2004). Choosing How to Choose: Institutional Pressures Affecting the Adoption of Personnel Selection Procedures. *International Journal of Selection and Assessment*, 12(4), 327-342. doi: 10.1111/j.0965-075X.2004.00288.x

- Kersting, M. (2006). Zur Beurteilung der Qualität von Tests: Resümee und Neubeginn. *Psychologische Rundschau*, 57, 243-253. doi: 10.1026/0033-3042.57.4.243
- Kersting, M. (2008). *Qualität in der Diagnostik und Personalauswahl – der DIN-Ansatz*. Göttingen: Hogrefe.
- Kersting, M. (2014). Qualitätsstandards der Personalauswahl. In H. Schuler & U. Kanning (Hrsg.), *Lehrbuch der Psychologie* (S. 325-356). Göttingen: Hogrefe.
- Khamis H. (2004). Measures of Association: How to Choose? *Journal of Diagnostic Medical Sonography*, 24(3). 155-162. doi: 10.1177/8756479308317006
- König, C., Klehe, U., Berchtold, M. & Kleinmann, M. (2010). Reasons for being selective when choosing personnel selection procedures. *International Journal of Selection and Assessment*, 18(1), 17-27. doi: 10.1111/j.1468-2389.2010.00485.x
- Krings, T. (2017). *Erfolgsfaktoren effektiver Personalauswahl*. Wiesbaden: Springer. doi: 10.1007/978-3-658-16456-0\_5
- Kristof-Brown, A., Zimmerman, R. & Johnson, E. (2005). Consequences of individuals' fit at work: A meta-analysis of person–job, person–organization, person–group, and person–supervisor fit. *Personnel psychology*, 58(2), 281-342. doi: 10.1111/j.1744-6570.2005.00672.x
- Krohne, W. & Hock, M. (2015). *Psychologische Diagnostik: Grundlagen und Anwendungsfelder* (2. Aufl.). Stuttgart: W. Kohlhammer.
- Kuckartz, U. (2014). *Mixed Methods: Methodologie, Forschungsdesigns und Analyseverfahren*. Heidelberg: Springer. doi: 10.1007/978-3-531-93267-5
- Kuckartz, U. (2018) *Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung* (4. Aufl.). Weinheim: Beltz Juventa.
- Kuckartz, U., Dresing, T., Rädiker, S. & Stefer, C. (2007). *Qualitative Evaluation: der Einstieg in die Praxis*. Wiesbaden: Springer VS.
- Latham, G. & Finnegan, B. (1993). Perceived practicality of unstructured, patterned, and situational interviews. In H. Schuler, J. Farr, & M. Smith (Hrsg.), *Personnel selection and assessment: Individual and organizational perspectives* (S. 41-55). Mahwah, NJ: Lawrence Erlbaum.
- Lee, Y., Berry, C. & Gonzalez-Mulé, E. (2019). The importance of being humble: A meta-analysis and incremental validity analysis of the relationship between honesty-humility and job performance. *Journal of Applied Psychology*, 104(12), 1535-1546. doi: 10.1037/apl0000421
- Lenzner, T., Neuert, C. & Otto, W. (2016). *Kognitives Pretesting*. Mannheim: GESIS Leibniz-Institut für Sozialwissenschaften. doi: 10.15465/gesis-sg\_010
- Leonhardt, R. (2017). *Lehrbuch Statistik* (4. Aufl.). Bern: Hogrefe.
- Mayer, D., Salovey, P., Caruso, R. & Sitarenios, G. (2001). Emotional intelligence as a standard intelligence. *Emotion*, 1(3), 232-242. doi: 10.1037/1528-3542.1.3.232

- Mayring, P. & Brunner, E. (2006): Qualitative Textanalyse - Qualitative Inhaltsanalyse. In: T. Schmid & V. Flaker (Hrsg.), *Von der Idee zur Forschungsarbeit. Forschen in Sozialarbeit und Sozialwissenschaft*. Wien: Böhlau.
- Moosbrugger, H. & Höfling, V. (2012). Standards für psychologisches Testen. In: H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 203-224). Berlin: Springer. doi: 10.1007/978-3-642-20072-4\_9
- Moosbrugger, H. & Kelava, A. (2020). *Testtheorie und Fragebogenkonstruktion* (3. Aufl.). Berlin: Springer.
- Morris, J. & Gardner, M. (1988). Calculating confidence intervals for relative risks (odds ratios) and standardized ratios and rates. *British Medical Journal*, 296, 1313-1316. doi: 10.1136/bmj.296.6632.1313
- Mücke, A. & Pekruhl, U. (2006). *HRM in Schweizer Betrieben. Quantitative Studien im Überblick*. Olten: FHNW - Institut für Personalmanagement und Organisation.
- Myers, D. (2014). *Psychologie* (10. Aufl.). Berlin: Springer.
- Neyer, F. & Asendorpf, J. (2018). *Psychologie der Persönlichkeit* (6. Aufl.). Berlin: Springer.
- O'Boyle E., Forsyth D., Banks G., Story P. & White C. (2015). A Meta-Analytic Test of Redundancy and Relative Importance of the Dark Triad and Five-Factor Model of Personality. *Journal of Personality* 83(6), 644-664. doi: 10.1111/jopy.12126.
- O'Boyle, E., Forsyth, D., Banks, G. & McDaniel, M. (2012). A meta-analysis of the Dark Triad and work behavior: A social exchange perspective. *Journal of Applied Psychology*, 97(3), 557-579. doi: 10.1037/a0025679
- Ones, D., Viswesvaran, C. & Schmidt, F. (2012). Integrity tests predict counterproductive work behaviors and job performance well: Comment on Van Iddekinge, Roth, Raymark, and Odle-Dusseau (2012). *Journal of Applied Psychology*, 97(3) 537-542. doi: 10.1037/a0024825
- Petermann, F., & Daseking, M. (2015). *Diagnostische Erhebungsverfahren*. Göttingen: Hogrefe.
- Podsakoff, P., MacKenzie, S., Lee, J. & Podsakoff, N. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *The Journal of applied psychology*, 88(5), 879-903. doi: 10.1037/0021-9010.88.5.879
- Porst, R. (2014). *Fragebogen. Ein Arbeitsbuch* (4. Aufl.). Wiesbaden: Springer. doi:10.1007/978-3-658-02118-4
- Rammstedt, B., Kemper, C., Klein, M., Beierlein, C. & Kovaleva, A. (2014). *Big Five Inventory (BFI-10). Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS)*. doi: 10.6102/zis76
- Robins, R., John, O. & Caspi, A. (1998). The typological approach to studying personality. Methods and Models for Studying the Individual. In: R. Cairns, J. Kagan, & L. Bergman (Hrsg.), *The individual in developmental research: Essays in honor of Marian Radke-Yarrow* (S. 135-160). Beverly Hills, CA: Sage.

- Ruhdorfer-Ritt, D. (2013). Personalauswahl. In: M. Landes & E. Steiner (Hrsg.) *Psychologie der Wirtschaft. Psychologie für die berufliche Praxis* (S. 175-195). Wiesbaden: Springer. doi: 10.1007/978-3-531-18957-4\_8
- Ryan, A. & Ployhart, R. (2014). A century of selection. *Annual Review of Psychology*, 65, 693–717. doi: 10.1146/annurev-psych-010213-115134
- Sackett, P., Zhang, C. Berry, C. & Lievens, F. (2022). Revisiting Meta-Analytic Estimates of Validity in Personnel Selection: Addressing Systematic Overcorrection for Restriction of Range. *Journal of Applied Psychology*, 107, 2040-2068. doi: 10.1037/apl0000994
- Sackett, P., Zhang, C., Berry, C., & Lievens, F. (2023). Revisiting the design of selection systems in light of new findings regarding the validity of widely used predictors. *Industrial and Organizational Psychology*, 1-18. doi:10.1017/iop.2023.24
- Salgado, J. & Tauriz, G. (2014). The Five-Factor Model, forced-choice personality inventories and performance: A comprehensive meta- analysis of academic and occupational validity studies. *European Journal of Work and Organizational Psychology*, 23, 3-30. doi: 10.1080/1359432X.2012.716198
- Salovey, P., Caruso, D. R., & Sitarenios, G. (2001). Emotional intelligence as a standard intelligence. *Emotion*, 1(3), 232-242. doi: 10.1037/1528-3542.1.3.232
- Schmidt, F. & Hunter, J. (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. doi: 10.1037/0033-2909.124.2.262
- Schmidt, F., Oh, I. & Shaffer, J. (2016). The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 100 Years of Research Findings. *Fox School of Business Research Paper*, 1-74. doi: 10.13140/RG.2.2.18843.26400
- Scholz, C. (2000). *Personalmanagement: informationsorientierte und verhaltenstheoretische Grundlagen* (5. Aufl.). München: Franz Vahlen.
- Schreiber, J. (2017). Update to core reporting practices in structural equation modeling. *Research in Social & Administrative Pharmacy*, 13(3), 634–643. doi: 10.1016/j.sapharm.2016.06.006
- Schuler, H. & Kanning, U. (2014). *Lehrbuch der Psychologie* (3.Aufl.). Göttingen: Hogrefe.
- Schuler, H. (2000). *Psychologische Personalauswahl* (3. Aufl.). Göttingen: Hogrefe.
- Schuler, H. (2004). *Organisationspsychologie - Grundlagen und Personalpsychologie*. Göttingen: Hogrefe.
- Schuler, H. (2013). *Psychologische Personalauswahl. Eignungsdiagnostik für Personalentscheidungen und Berufsberatung* (4. Aufl.). Göttingen: Hogrefe.
- Schuler, H., Höft, S. & Hell, B. (2014). Eigenschaftsorientierte Verfahren in der Personalauswahl. In H. Schuler & U. Kanning (Hrsg.), *Lehrbuch der Personalpsychologie* (S. 149-213). Göttingen: Hogrefe.
- Schwarb, T. (1996). *Die wissenschaftliche Konstruktion der Personalauswahl*. München: Rainer Hampp.

- Stemmler, G. & Margraf-Stiksrud, J. (2015). *Lehrbuch Psychologische Diagnostik*. Bern: Hans Huber.
- Urech, R. (2020, 31. Januar). Antreten zum Persönlichkeitstest: Bewerber müssen immer öfter auch ihren Charakter analysieren lassen. *Aargauer Zeitung*.  
<https://www.aargauerzeitung.ch/leben/antreten-zum-personlichkeitstest-bewerber-muessen-immer-oft-er-auch-ihren-charakter-analysieren-lassen-ld.1190272>
- UZH (2024). *Logistische Regressionsanalyse – Voraussetzungen*.  
[https://www.methodenberatung.uzh.ch/de/datenanalyse\\_spss/zusammenhaenge/lreg.html](https://www.methodenberatung.uzh.ch/de/datenanalyse_spss/zusammenhaenge/lreg.html)
- Van der Linden, D., Pekaar, K., Bakker, A., Schermer, J., Vernon, P., Dunkel, C. & Petrides, K. (2017). Overlap between the general factor of personality and emotional intelligence: A meta-analysis. *Psychological Bulletin*, 143(1), 36-52. doi: 10.1037/bul0000078.
- Véron, T., Mitrovic, T. & Hell, B. (2024). Erfolgsfaktoren Employer Branding und agiles Recruiting. *personalsCHWEIZ*, 1.
- Walther, B. (2022). *Kendall-Tau*. <https://bjoernwalther.com/kendall-tau-korrelationskoeffizient-in-spss-berechnen/>
- Weichbrodt, J. (2022). *Mixed Methods Forschungsdesign. Vorlesungsunterlagen vom 25.04.2022*. Olten: FHNW.
- Wilhelmy, A. & Kleinmann, M. (2016). Selektion und Attraktion. In: S. Kauffeld, S. & D. Spurk (Hrsg.) *Handbuch Karriere und Laufbahnmanagement* (S. 1-32). Berlin: Springer. doi: 10.1007/978-3-662-45855-6\_7-1
- Zöfel, P. (2003). *Statistik für Psychologen*. München: Pearson.
- ZPID. (2023). *Verzeichnis Testverfahren. Kurznamen. Langnamen. Autoren. Testrezensionen* (29. Aufl.). doi: 10.23668/psycharchives.12312
- ZPID. (2024). *Verzeichnis zur Suche nach Testverfahren*. <https://zpidu16.zpid.de/pubpsych/>

## Anhang A

### Abkürzungsverzeichnis

---

Abb.	Abbildung
AC	Assessment Center
BAV	Bundesamt für Verkehr
BFS	Bundesamt für Statistik
BPS	British Psychological Society
BV	Bundesverfassung
CFA	Confirmatory Factor Analysis (konfirmatorische Faktorenanalyse)
CFI	Comparative Fit Index (Fit Index bei Faktoranalysen)
CHC-Modell	Cattell-Horn-Carroll-Modell
df	Degrees of Freedom (Freiheitsgrade)
DIN	Deutsche Industrienorm
DISC	Dominance-Influence-Supportiveness-Conscientiousness-Test (Dominanz, Einfluss, Beständigkeit, Gewissenhaftigkeit)
DL	Dienstleistung(en)
DTK	Diagnostik- und Testkuratorium
EFA	Exploratory Factor Analysis, zu Deutsch: explorative Faktorenanalyse
EFPA	European Federation of Psychologists Associations
FK	Führungskraft
HFP	Höhere Fachprüfung
HR	Human Resources
Hrsg.	Herausgebende
Int.	Interview
IQ	Intelligenzquotient
IST	Intelligenz-Struktur-Test
Kap.	Kapitel
KMO	Kaiser-Meyer-Olkin-Kriterium
KMU	Kleines bis mittelgrosses Unternehmen
<i>M</i>	Mittelwert
MAS	Master of advanced Studies
MBTI	Myers-Briggs Type Indicator (Myers-Briggs-Typenindikator)
MPA	Master Person Analysis (Persönlichkeitstest von Master)
MSA	Measurement System Analysis
<i>n</i>	Stichprobengrösse
NEO-PI-R	NEO Personality Inventory-Revised (NEO-Persönlichkeitsinventar revidiert)
NIP	Niederlands Instituut van Psychologen
O*NET	Occupational Information Network (Berufsinformationsnetzwerk)
OF	Oberfragestellungen
<i>p</i>	kleinste Signifikanz
Pos.	Position
<i>r</i>	Korrelationskoeffizient (Zusammenhangsmass)
<i>R</i> <sup>2</sup>	Determinationskoeffizient / Bestimmtheitsmass
RMSEA	Root Mean Square Error of Approximation (Fit Index bei Faktoranalysen)
<i>r<sub>p</sub></i>	Korrelationskoeffizient Pearson
<i>r<sub>s</sub></i>	Korrelationskoeffizient Spearman $\rho$
SD	Standardabweichung (Streuungsmass)
SDBB	Schweizerisches Dienstleistungszentrum Berufsbildung, Berufs-, Studien- und Laufbahnberatung
SI	Itemschwierigkeit
SPSS	statistical Package for the Social Sciences (statistisches Auswertungsprogramm von IBM)
SR	Systematische Rechtssammlung

SRMR	Standardized Root Mean Square Error of Approximation (Fit Index bei Faktoranalysen)
Tab.	Tabelle
TBS	Testbewertungssystem
TBS-DTK	Testbewertungssystem des Diagnostik- und Testkuratoriums
TKS	Trennschärfenkoeffizient
TLI	Tucker-Lewis Index (Fit Index bei Faktoranalysen)
TRA	Theory of Reasoned Action (nach Fishbein und Ajzen, 1975)
UF	Unterfragestellungen
V	Variationskoeffizient
VG	Vorgesetzte
VG	Vorgesetzter bzw. Vorgesetzte
vgl.	vergleiche
ZPID	Leibniz-Institut für Psychologie
$\alpha$	Cronbachs Alpha (interne Konsistenz)
$\beta$	Beta-Koeffizient (Standardisierter Regressionskoeffizient)
$\rho$	wahrer Korrelationskoeffizient in der Population, verwendet in Meta-Analysen
$\chi^2$	Chi-Quadrat-Test

## Anhang B

Leitfaden zur Durchführung der Expertinnen- und Experteninterviews, Erstellung orientiert sich an Helfferich (2004)

Themenblock: Einstieg			
Erzählauffordernde Leitfragen	Check: Wurde das erwähnt? Memo für mögliche Nachfragen	Zusätzliche konkrete Fragen in genau dieser Formulierung	Auswahl an passenden Aufrechterhaltungs- und Steuerungsfragen
Wie sieht Ihr beruflicher Werdegang aus?	Ausbildungen?	Welche Ausbildungen und Weiterbildungen haben Sie absolviert?	> Und dann? Wie ging es weiter?
	Berufserfahrung?	Wie lange arbeiten Sie bereits in diesem Unternehmen?	> Können Sie dazu noch etwas mehr erzählen?
		In welcher Funktion arbeiten Sie in diesem Unternehmen?	> Haben Sie immer in derselben Funktion gearbeitet?
Frühere Arbeitsorte?	Haben Sie bereits in anderen Unternehmen im HR / Personalwesen gearbeitet?	> Und dann? Wie ging es weiter?	
Haben Sie in Ihrem Arbeitsumfeld selbst schon einmal einen psychologischen Test machen müssen?	Art?	Handelte es sich um einen Leistungstest oder ein Persönlichkeitsinventar?	> Wie war das für Sie?
	Zweck?	Zu welchem Zweck mussten Sie den Test ausfüllen?	
Was sind Ihrer Meinung nach Vorteile, die für den Einsatz von psychologischen Tests im Rahmen von Personalentscheidungen sprechen?	Begründung?	-	> Gibt es noch weitere Vorteile, die Sie nennen können?
Was sind Ihrer Meinung nach Nachteile, die gegen den Einsatz von psychologischen Tests im Rahmen von Personalentscheidungen sprechen?	Begründung?	-	> Gibt es noch weitere Nachteile, die Sie nennen könnten?
Themenblock: Oberfragestellungen			
Erzählauffordernde Leitfragen	Check: Wurde das erwähnt? Memo für mögliche Nachfragen	Zusätzliche konkrete Fragen in genau dieser Formulierung	Auswahl an passenden Aufrechterhaltungs- und Steuerungsfragen
Setzen Sie in Ihrem Unternehmen psychologische Tests zur Personalauswahl oder zur Personalentwicklung ein?	Leistungstests?	Wenn ja: welche? (siehe Frageblock 1.1)	-
	Persönlichkeitsinventare?	Wenn nein: weshalb nicht? (siehe Frageblock 1.2)	-
Themenblock: Unterfragestellungen 1.1			
Erzählauffordernde Leitfragen	Check: Wurde das erwähnt? Memo für mögliche Nachfragen	Zusätzliche konkrete Fragen in genau dieser Formulierung	Auswahl an passenden Aufrechterhaltungs- und Steuerungsfragen
Welche Tests setzen Sie ein?	Nennung der eingesetzten Tests?	Wie heisst der Test genau?	> Gibt es noch weitere Leistungstests, die Sie in Ihrem Unternehmen einsetzen?
		Wie heissen die Tests genau?	> Gibt es noch weitere Persönlichkeitsinventare, die Sie in Ihrem Unternehmen einsetzen?
Weshalb setzen Sie genau diese Tests ein?	Beweggründe?	Welche Beweggründe haben Sie zur Auswahl dieses bzw. dieser Tests geführt?	> Haben Sie gewisse Erfahrungen, die Sie geleitet haben? Wenn ja: welche?
	Entscheidungskriterien (Kosten-Nutzen-Kalkül)?	Welche Abwägungen haben Sie bei der Entscheidung gemacht bzw. machen müssen?	> Gab es gewisse Vorgaben, die Sie geleitet haben? Wenn ja: welche?
	Informationsbasis?	Welche Informationsquellen haben Sie genutzt, um sich für den Test bzw. die Tests entscheiden zu können?	> Wo haben Sie sich informiert? > Gab es noch weitere Informationen bzw. Informationsquellen, die Sie geleitet haben? Wenn ja: welche?
Zu welchem Zweck bzw. zu welchen Zwecken setzen Sie die Tests ein?	Personalentscheidungen?	Für welche Personalentscheidungen setzen Sie die Tests ein?	> Welche weiteren Personen hatten Einfluss auf die Entscheidung?
			> Welche äusseren Umstände hatten Einfluss auf die Entscheidung?
			> Setzen Sie die Tests für die Personalauswahl ein? > Setzen Sie die Tests für die Personalentwicklung ein? > Zu welchen weiteren Zwecken setzen Sie die Tests ein?

Für welche Zielgruppe setzen Sie genannte Tests ein?	Zielgruppe?	Für welche Zielgruppe setzen Sie die genannten Tests ein? Weshalb wird diese Zielgruppe berücksichtigt? Weshalb werden diese Zielgruppen berücksichtigt?	> Können Sie dazu noch etwas mehr erzählen? > Wieso werden andere Zielgruppen nicht berücksichtigt?
Ziehen Sie in Erwägung, das momentane Prozedere in Zukunft zu verändern bzw. anzupassen?	Zukünftige Planung?	Wie sähe das Prozedere in Zukunft aus?	> Können Sie dazu noch etwas mehr erzählen? > Wie sollte das Prozedere Ihrer Meinung nach aussehen?
<b>Themenblock: Unterfragestellungen 1.2</b>			
Erzählauffordernde Leitfragen	<b>Check: Wurde das erwähnt? Memo für mögliche Nachfragen</b>	<b>Zusätzliche konkrete Fragen in genau dieser Formulierung</b>	<b>Auswahl an passenden Aufrechterhaltungs- und Steuerungsfragen</b>
Weshalb setzen Sie keine psychologischen Tests zur Personalauswahl oder Personalentwicklung ein?	Beweggründe?	Welche Beweggründe haben Sie dazu geführt?	> Haben Sie gewisse Erfahrungen, die Sie geleitet haben? Wenn ja: welche?
	Entscheidungskriterien (Kosten-Nutzen-Kalkül)?	Welche Abwägungen haben gegen den Einsatz gesprochen?	> Gab es gewisse Vorgaben, die Sie geleitet haben? Wenn ja: welche?
	Informationsbasis?	Auf Basis welcher Informationen bzw. Informationsquellen ist die Entscheidung erfolgt?	> Wo haben Sie sich informiert? > Gab es noch weitere Informationen bzw. Informationsquellen, die Sie geleitet haben? Wenn ja: welche?
	Externale Einflüsse?	Gab es noch weitere Faktoren, die Einfluss auf die Entscheidung hatten?	> Welche weiteren Personen hatten Einfluss auf die Entscheidung? > Welche äusseren Umstände hatten Einfluss auf die Entscheidung?
Ziehen Sie einen Einsatz psychologischer Tests für die Zukunft in Erwägung?	Zukünftige Planung?	Wie sähe das Prozedere in Zukunft aus?	> Können Sie dazu noch etwas mehr erzählen? > Wie sollte das Prozedere Ihrer Meinung nach aussehen?
<b>Themenblock: Unterfragestellungen 2</b>			
Erzählauffordernde Leitfragen	<b>Check: Wurde das erwähnt? Memo für mögliche Nachfragen</b>	<b>Zusätzliche konkrete Fragen in genau dieser Formulierung</b>	<b>Auswahl an passenden Aufrechterhaltungs- und Steuerungsfragen</b>
Haben Sie in der Vergangenheit den Testeinsatz in Ihrem Unternehmen anders gehandhabt?	Handhabung?	Wie wurde der Testeinsatz früher in Ihrem Unternehmen gehandhabt?	> Können Sie dazu noch etwas mehr erzählen?
	Prozedere?	Weshalb hat man das Prozedere verändert?	> Wie kam es dazu? > Welche weiteren Einflussgrößen haben zur Änderung geführt?
Wie war es an anderen Arbeitsorten? Haben Sie dort psychologische Tests zur Personalauswahl oder zur Personalentwicklung eingesetzt?	Handhabung?	Wie wurde der Testeinsatz an anderen Arbeitsorten gehandhabt?	> Können Sie dazu noch etwas mehr erzählen? > Wie kam es dazu?
<b>Themenblock: Abschluss</b>			
Haben Sie noch Anmerkungen?			
Gibt es Unklarheiten?			
E-Mail Adresse erfragen, um Masterarbeit und jeweiliges Transkript zusenden zu können (Spätsommer 2024)			
Danksagung für Teilnahme am Interview			

## Anhang C

Rohdaten zur Planung der geschichteten Stichprobe

06.02.00.01.01a Vollzeitäquivalente nach Wirtschaftsabteilungen (BFS, 2024a)				
Quartalswerte, in 1'000				
Wirtschaftsbranchen			2023	
Abschnitte und Abteilungen			I	II
B-S	5-96	Total	4 206.0	4 234.6
B-F	5-43	SEKTOR II	1 025.4	1 033.4
<b>B</b>	<b>5-9</b>	<b>Bergbau und Gewinnung von Steinen und Erden</b>	<b>4.6</b>	<b>4.6</b>
<b>C</b>	<b>10-33</b>	<b>Verarbeitendes Gewerbe/Herstellung von Waren</b>	<b>633.0</b>	<b>638.3</b>
	10-12	Herstellung von Nahrungsmitteln und Tabakerz.	79.9	80.5
	13-15	Herstellung von Textilien und Bekleidung	11.4	11.4
	16-18	Herstellung von Holzwaren, Papier und Druckerz.	57.2	57.5
	19-20	Kokerei, Mineralölverarb. u. Herstel. von chem. Erz.	30.0	29.8
	21	Herstellung von pharmazeutischen Erzeugnissen	45.3	45.3
	22-23	Herstellung von Gummi- und Kunststoffwaren	37.7	37.9
	24-25	Herstellung von Metallerzeugnissen	90.8	91.5
	26	Herstellung von Datenverarbeitungsgeräten und Uhren	111.4	114.2
	27	Herstellung von elektrischen Ausrüstungen	26.2	26.0
	28	Maschinenbau	74.2	74.8
	29-30	Fahrzeugbau	15.8	16.0
	31-33	Sonstige Herstellung von Waren, Rep. und Inst.	53.2	53.4
<b>D</b>	<b>35</b>	<b>Energieversorgung</b>	<b>28.5</b>	<b>28.1</b>
<b>E</b>	<b>36-39</b>	<b>Wasserversorgung, Beseitigung von Umweltverschmutz.</b>	<b>18.2</b>	<b>18.3</b>
<b>F</b>	<b>41-43</b>	<b>Baugewerbe/Bau</b>	<b>341.1</b>	<b>344.1</b>
	41-42	Hoch- und Tiefbau	108.1	110.7
	43	Sonstiges Ausbaugewerbe	233.0	233.4
G-S	45-96	SEKTOR III	3 180.6	3 201.2
<b>G</b>	<b>45-47</b>	<b>Handel, Instandhaltung und Rep. von Kraftfahrzeugen</b>	<b>517.4</b>	<b>518.0</b>
	45	Handel und Reparatur von Motorfahrzeugen	80.3	80.4
	46	Grosshandel	204.0	206.3
	47	Detailhandel	233.2	231.4
<b>H</b>	<b>49-53</b>	<b>Verkehr und Lagerei</b>	<b>224.1</b>	<b>219.1</b>
	49	Landverkehr und Transport in Rohrfernleitungen	118.0	113.5
	50-51	Schifffahrt und Luftfahrt	15.3	16.0
	52	Lagerei sowie Erbring. v. sonst. Dienstl. für den Verk.	58.4	58.8
	53	Post-, Kurier- und Expressdienste	32.4	30.8
<b>I</b>	<b>55-56</b>	<b>Gastgewerbe/Beherbergung und Gastronomie</b>	<b>197.9</b>	<b>200.2</b>
	55	Beherbergung	64.2	67.6
	56	Gastronomie	133.7	132.7
<b>J</b>	<b>58-63</b>	<b>Information und Kommunikation</b>	<b>168.7</b>	<b>167.8</b>
	58-60	Verlagswesen, audiovisuelle Medien und Rundfunk	24.4	23.9
	61	Telekommunikation	24.7	24.6
	62-63	Informationstechnologische und Informationsdienstl.	119.7	119.3
<b>K</b>	<b>64-66</b>	<b>Erbringung von Finanz- und Versicherungsdienstl.</b>	<b>219.3</b>	<b>220.1</b>
	64	Erbringung von Finanzdienstleistungen	108.1	108.3
	65	Versicherungen	42.1	43.2
	66	Mit Finanz- u. Versicherungsdienstl. verb. Tätigk.	69.0	68.6
<b>L</b>	<b>68</b>	<b>Grundstücks- und Wohnungswesen</b>	<b>43.9</b>	<b>44.9</b>
<b>M</b>	<b>69-75</b>	<b>Erbringung von freiberufl., wissen. u. techn. Dienstl.</b>	<b>397.4</b>	<b>397.9</b>
	69	Rechts- und Steuerberatung, Wirtschaftsprüfung	73.4	72.1
	70	Unternehmensverwaltung u. -führung; Unternehmensber.	119.0	116.8
	71	Architektur- und Ingenieurbüros	120.7	121.7
	72	Forschung und Entwicklung	33.2	33.8
	73-75	Sonstige freiberufl., wissen. und techn. Dienstl.	51.1	53.4
<b>N</b>	<b>77-82</b>	<b>Erbringung von sonstigen wirtschaftlichen Dienstl.*</b>	<b>238.5</b>	<b>256.6</b>
	77+79-82	Erbringung von sonst. wirtschaftlichen Dienstl. (ohne 78)	152.1	156.6
	78	Vermittlung und Überlassung von Arbeitskräften	86.4	100.0
<b>O</b>	<b>84</b>	<b>Öffentliche Verwaltung</b>	<b>171.6</b>	<b>172.5</b>
<b>P</b>	<b>85</b>	<b>Erziehung und Unterricht</b>	<b>253.2</b>	<b>250.4</b>
<b>Q</b>	<b>86-88</b>	<b>Gesundheits- und Sozialwesen</b>	<b>575.5</b>	<b>579.5</b>
	86	Gesundheitswesen	334.3	336.5
	87	Heime (ohne Erholungs- und Ferienheime)	153.8	154.7
	88	Sozialwesen (ohne Heime)	87.4	88.3
<b>R</b>	<b>90-93</b>	<b>Kunst, Unterhaltung und Erholung</b>	<b>61.5</b>	<b>61.5</b>
<b>S</b>	<b>94-96</b>	<b>Erbringung von sonstigen Dienstleistungen</b>	<b>111.6</b>	<b>112.5</b>

\* = zu diverse Gruppe

über ca. 200'000 Mitarbeitende im Vollzeitäquivalent pro Wirtschaftsabteilung; 1 Interview

über ca. 500'000 Mitarbeitende im Vollzeitäquivalent pro Wirtschaftsabteilung; 1-2 Interviews

### Einverständniserklärung zur Teilnahme an der studentischen Masterarbeit «Einsatz psychologischer Tests bei Personalentscheidungen»

---

Studienverantwortung: Tanja Senn

Kontaktinformationen: E-Mail: tanja.senn@students.fhnw.ch, Telefon: [REDACTED]

#### Informationen zur Studie

Psychometrische Testverfahren wie Intelligenztests oder Persönlichkeitsinventare werden zur Personalauswahl und Personalentwicklung eingesetzt (Leinweber, 2013). Das Angebot für solche Tests ist mannigfaltig (Hell, 2013). Rund ein Drittel der Schweizer Unternehmen setzt neben gängigen Auswahlverfahren wie dem Einstellungsinterview oder der Prüfung der Bewerbungsunterlagen auch Persönlichkeitstests zur Personalauswahl ein (Urech, 2020). Im deutschsprachigen Raum gibt es schätzungsweise 3'000 verschiedene Persönlichkeitstests (Hell, 2013). Daten über Art und Einsatzhäufigkeit von Persönlichkeitstests oder Leistungstests zur Personalauswahl und -entwicklung in Schweizer Unternehmen fehlen. Damit fehlen auch Daten, für welche Zielgruppen und zu welchem Zweck gewisse Tests eingesetzt werden, welche Beweggründe, Entscheidungskriterien und Einflussgrössen zur Wahl bestimmter Tests im Unternehmen geführt haben. Diese Masterarbeit widmet sich dieser Forschungslücke.

#### Teilnahme

Ihre Teilnahme an der Studie beinhaltet ein qualitatives Interview von ungefähr einer Stunde, welches auditiv aufgezeichnet wird. Die auditive Aufzeichnung dient der anschliessenden Transkription des Interviewgesprächs, wobei eine Anonymisierung Ihrer personenbezogenen Daten vorgenommen wird. Ihnen werden Fragen zum *Einsatz psychologischer Tests bei Personalentscheidungen* gestellt.

#### Rücktritt von der Teilnahme

Die Teilnahme an dieser Studie ist freiwillig. Sie haben jederzeit das Recht, von der Teilnahme an der Studie zurückzutreten, ohne dafür einen Grund anzugeben.

#### Datenschutz, Vertraulichkeit und zukünftige Nutzung

Die im Rahmen dieser Studie gesammelten Daten werden nur für diese Studie genutzt. Ihr Name oder andere Informationen zu Ihrer Identität werden vertraulich behandelt, nicht publiziert und nicht an Dritte weitergegeben. Ihr Beitrag wird in anonymisierter oder pseudonymisierter Form sicher aufbewahrt und nach Abschluss der Studie gelöscht.

#### Einverständnis

Ich habe die Informationen in dieser Erklärung gelesen und verstanden. Mir wurde die Möglichkeit eingeräumt, Fragen zur Studie zu stellen und diese wurden zu meiner Zufriedenheit beantwortet. Ich gebe mein Einverständnis, an dieser Studie teilzunehmen und gebe meine Erlaubnis für die Aufbewahrung und Verwendung meiner im Rahmen dieser Studie gesammelten Daten.

#### Unterschriften

Name der teilnehmenden Person

Datum

Unterschrift

Einverständnis eingeholt von

Datum

Unterschrift

## Anhang E

Abfrage der abhängigen Variablen und der spezifischen Folgefragen

Abfrage der abhängigen Variable				
Variablenname	Item	Antwortskala	Quellen für Antwortskala	Skalenniveau
<b>Testeinsatz</b>	Setzen Sie in Ihrem Unternehmen psychologische Tests im Rahmen von Personalentscheiden ein?	Mehr- bzw. Einfachauswahlaufgabe: <ul style="list-style-type: none"> <li>Ja, wir führen Tests intern durch</li> <li>Ja, wir lassen Tests extern durchführen</li> <li>Nein</li> </ul>	Orientiert an entsprechender Frage aus Interviewleitfaden	Nominalskaliert
Spezifische Abfrage betreffend die eingesetzten Tests, wenn Testeinsatz bejaht wird				
Variablenname	Item	Antwortskala	Quellen für Antwortskala	Skalenniveau
<b>Testnamen</b>	Welche psychologischen Tests setzen Sie in Ihrem Unternehmen ein? (Bitte nennen Sie die Namen und gegebenenfalls die Kürzel der eingesetzten Testverfahren im untenstehenden Textfeld!)	Offene Antwortmöglichkeit in Form eines Textfeldes	orientiert an entsprechender Frage aus Interviewleitfaden	Nominalskaliert (Häufigkeiten der genannten Tests werden anschliessend berechnet)
<b>Auswahlgrund</b>	Weshalb haben Sie sich für die genannten psychologischen Tests entschieden? (Bitte nennen Sie die Gründe im untenstehenden Textfeld!)	Offene Antwortmöglichkeit in Form eines Textfeldes	orientiert an entsprechender Frage aus Interviewleitfaden	Nominalskaliert (Häufigkeiten der genannten Tests werden anschliessend berechnet)
<b>Zielgruppe</b>	Für welche Zielgruppen werden Tests eingesetzt? (Mehrfachauswahl möglich!)	Mehrfachauswahlaufgabe plus offene Antwortmöglichkeit, falls die betroffene Zielgruppe nicht in der Auswahl genannt wurde. <ul style="list-style-type: none"> <li>ungelernte Mitarbeitende</li> <li>Auszubildende</li> <li>Trainees oder Hochschulabsolventinnen und -absolventen</li> <li>Ausgebildete Mitarbeitende ohne Führungsfunktion</li> <li>Führungskräfte untere Ebene</li> <li>Führungskräfte mittlere Ebene</li> <li>Führungskräfte obere Ebene</li> <li>andere Zielgruppen (bitte beschreiben)</li> </ul>	orientiert an Armoneit et al., 2022 und (Hossiep, Weiss & Schecke, 2015)	Ordinalskaliert (geordnet nach Ausbildungsgrad bzw. Akademisierungsstufe)
<b>Zweck</b>	Zu welchem Zweck setzen Sie die genannten psychologischen Tests ein? (Mehrfachauswahl möglich!)	Mehrfachauswahlaufgabe plus offene Antwortmöglichkeit, falls weitere bzw. andere Zwecke mit dem Testeinsatz verfolgt werden: <ul style="list-style-type: none"> <li>Personalauswahl</li> <li>Personalentwicklung</li> <li>andere Zwecke (bitte beschreiben)</li> </ul>	Formulierung orientiert sich an der Fragestellung der vorliegenden Thesis.	Nominalskaliert
<b>Informationsbasis</b>	Auf Basis welcher Informationen wurde die Wahl für den genannten Test / für die genannten Tests getroffen? (Mehrauswahl möglich!)	Mehrfachauswahlaufgabe plus offene Antwortmöglichkeit, falls verwendete Informationsbasis nicht in der Auswahl genannt wurde: <ul style="list-style-type: none"> <li>einfache Websuche</li> <li>Fachliteratur (Fachzeitschriften und -bücher)</li> <li>Testrezensionen</li> <li>Austausch mit Fachexperten und -expertinnen</li> <li>Workshops und Konferenzen</li> <li>Kurse und Schulungen</li> <li>Austausch mit Testanbieterinnen und -anbietern</li> <li>keine Angabe möglich</li> <li>andere Basis (bitte beschreiben)</li> </ul>	Eigene Formulierung.	Nominalskaliert

## Anhang F

### Abfrage der demografischen Daten

Abfrage der demografischen Daten				
Abfrage der Daten betreffend Unternehmen				
Variablenname	Item	Antwortskala	Quellen für Antwortskala	Skalenniveau
Wirtschaftsabteilung	Welcher der folgenden Wirtschaftsabteilungen gehört das Unternehmen an, in dem Sie arbeiten?	<p>Einfachwahlaufgabe mit offener Antwortmöglichkeit, falls den Teilnehmenden die Zuordnung nicht möglich ist:</p> <ul style="list-style-type: none"> <li>Land- und Forstwirtschaft</li> <li>Verarbeitendes Gewerbe / Herstellung von Waren</li> <li>Energie-, Wasser- und Umweltversorgung</li> <li>Baugewerbe / Bau, Handel / Reparaturgewerbe</li> <li>Verkehr und Lagerei</li> <li>Gastgewerbe</li> <li>Information und Kommunikation</li> <li>Finanz- und Versicherungsdienstleistungen</li> <li>Grundstücks- und Immobilienverwaltung</li> <li>Erbringung von wissenschaftlichen und technischen Dienstleistungen</li> <li>öffentliche Verwaltung</li> <li>Erziehung und Unterricht</li> <li>Gesundheits- und Sozialwesen</li> <li>Kunst</li> <li>Unterhaltung und Erholung</li> <li>andere Wirtschaftsabteilung (bitte beschreiben)</li> </ul>	Auswahl abgeleitet von (BFS, 2023c).	Nominalskaliert
Grösse des Unternehmens	Wie viele Mitarbeitende beschäftigt Ihr Unternehmen?	<p>Einfachwahlaufgabe mit geschlossener Antwortmöglichkeit:</p> <ul style="list-style-type: none"> <li>1 - 9 Mitarbeitende</li> <li>10 - 49 Mitarbeitende</li> <li>50 - 249 Mitarbeitende</li> <li>250 bis 999 Mitarbeitende</li> <li>1'000 bis 9'999 Mitarbeitende</li> <li>10'000 und mehr Mitarbeitende</li> </ul>	Ranges orientiert an (BFS, 2023b)	Ordinalskaliert
Psychologin bzw. Psychologe intern	Arbeiten in Ihrer HR-Abteilung ein Psychologe bzw. eine Psychologin?	<p>Einfachwahlaufgabe mit geschlossener Antwortmöglichkeit:</p> <ul style="list-style-type: none"> <li>Ja</li> <li>Nein</li> <li>Das weiss ich nicht</li> </ul>	Eigene Formulierung.	Nominalskaliert
Abfrage der Daten betreffend Teilnehmenden				
Variablenname	Item	Antwortskala	Quellen für Antwortskala	Skalenniveau
Höchster Ausbildungsabschluss	Was ist Ihr höchster Ausbildungsabschluss?	<p>Einfachwahlaufgabe mit offener Antwortmöglichkeit, falls der Abschluss nicht aufgeführt wurde</p> <ul style="list-style-type: none"> <li>Lehrabschluss EFZ</li> <li>HR-Assistentin / HR-Assistent</li> <li>Personalfachmann / -frau mit eidg. Fachausweis</li> <li>dipl. Leiter / -in Human Resources (HFP)</li> <li>MAS Human Resource Management</li> <li>Bachelor in Psychologie oder Wirtschaftswissenschaften</li> <li>Master in Psychologie oder Wirtschaftswissenschaften</li> <li>anderer Abschluss (bitte beschreiben)</li> </ul>	Auswahl abgeleitet vom offiziellen schweizerischen Informationsportal für Berufs-, Studien- und Laufbahnberatung (2024)	Ordinalskaliert
Berufserfahrung	Über wie viele Jahre Berufserfahrung im Bereich HR verfügen Sie? (Bitte tragen Sie die ungefähre Anzahl als ganze Zahl im Textfeld ein!)	Offene Antwortmöglichkeit für maximal zwei Ziffern	Eigene Formulierung.	Verhältnisskaliert

## Anhang G

Pretest-Seiten in Tivian

### 1 Begrüssungsseite\_PreTester

Liebe Tester

Ich danke Euch für Eure Zeit!

Hier noch einige Hinweise:

- Bitte wählt anschliessend Eure Initialen und den Testdurchgang aus.
- Bitte beachtet, dass die Übersetzungen zum aktuellen Zeitpunkt noch nicht eingepflegt sind. Wählt also bei der Sprache "Deutsch" aus.
- Am Ende des Fragebogens werden Euch Feedbackfragen zum Fragebogen gestellt. Für schriftliches Feedback stehen Textfelder bereit.

Ich bin gespannt auf Euer Feedback!

Liebe Grüsse

Tanja

Bitte wähle Deine Initialen im Dropdown-Menü aus.

Bitte wähle den Testdurchgang im Dropdown-Menü aus.

### 13 Feedback\_I

Bitte bewerte die Testfragen (und formuliere allenfalls ein kurzes Feedback):

Die gestellten Fragen fand ich..

	trifft gar nicht zu					trifft völlig zu
.. verständlich.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
.. relevant.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
.. genau formuliert.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Hast Du Anmerkungen zu den Testfragen?

Wenn ja, bitte nutze das Textfeld! Wenn nein, klicke auf "Weiter"!

## 14 Feedback\_II

Bitte bewerte den Test an sich (und formuliere allenfalls ein kurzes Feedback):

	trifft gar nicht zu					trifft völlig zu
Der Test ist benutzerfreundlich gestaltet.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Das Layout ist ansprechend.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Es gab keine technischen Probleme.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Das Ausfüllen hat 15min gedauert.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Mit welchem Endgerät hast Du den Fragebogen ausgefüllt?

- Handy
- Tablet
- Computer
- anderes Endgerät (bitte beschreiben)

Hast Du Anmerkungen zum Test an sich?

Wenn ja, bitte nutze das Textfeld! Wenn nein, klicke auf "Weiter"!

## 15 Endseite

Lieber Tester, Deine Antworten wurden gespeichert.  
Danke für Deine Zeit und Unterstützung!

## Anhang H

Pretest Rohdaten, mit SPSS berechnet

Berechnung Retest-Reliabilität (hinsichtlich Stabilität), ausgedrückt in Kendalls Tau b Koeffizient

Items	E.S. - 1. DG	E.S. - 2. DG	F.S. - 1. DG	F.S. - 2. DG	L.S. - 1. DG	L.S. - 2. DG
positive Erfahrungen	1	1	6	6	5	5
Einarbeitungsbedarfe	3	4	5	4	2	2
Personalentwicklung	5	4	5	4	5	2
Eignung	4	4	5	5	5	5
Person-Group-Fit	4	2	3	3	2	2
UK-Fit	4	2	5	4	5	2
genaue Resultate erhalten	3	2	6	6	3	4
auf zukünftige Leistung schliessen	2	2	5	4	2	2
wissenschaftliche Fundierung	3	2	6	6	5	5
Vergleichbarkeit unter Bewerbenden	4	4	6	6	6	6
Normen & Benchmarks	4	2	5	4	5	5
gleiche Voraussetzungen schaffen	4	4	6	5	5	5
praktische Durchführung	3	2	4	3	6	5
Entscheidungssicherheit	5	4	4	5	6	4
weitere Informationen	5	4	5	5	5	5
Objektivierung	4	3	5	5	4	5
Fehlbesetzungen	2	2	2	5	3	3
Interviewfragen	5	4	5	5	1	4
Return on Investment	4	2	6	4	4	5
Vorauswahl	4	5	6	3	6	6
bisherige Praxis ausreichend	2	2	1	2	1	2
hohe Lizenzkosten	2	4	5	5	4	2
hoher Zeitaufwand	2	2	4	3	2	3
ungewollte Selbstselektion	4	2	6	4	5	2
Arbeitgebendenattraktivität sinkt	2	2	4	4	4	2
keine Akzeptanz durch Bewerbende	2	2	4	4	3	3
Scheingenauigkeit	4	4	2	2	1	1
adäquat für alle Anforderungsstufen	2	4	6	6	5	5
Verallgemeinerung	4	3	5	5	2	3
persönliche Befürwortung	5	2	6	6	5	5
Meinungseinfluss im Unternehmen	1	1	1	1	6	4
Scientist-practitioner gap	2	5	2	2	2	3
vielfältiges Testangebot	5	5	5	5	6	4
Ineffizienz	1	2	5	4	x	x
wenige Bewerbungen	5	5	5	6	x	x
Konkurrenzorientierung	2	2	1	1	x	x
fehlende interne Expertise	6	6	3	2	x	x
fehlender UK-Fit	2	2	1	1	x	x
fehlende Finanzen	5	2	2	2	x	x
fehlende VG-Akzeptanz	1	2	6	4	x	x
Effizienz	x	x	x	x	5	5
viele Bewerbungen	x	x	x	x	4	4
positiv Konkurrenzorientierung	x	x	x	x	6	5
vorhandene interne Expertise	x	x	x	x	2	4
vorhandener UK-Fit	x	x	x	x	1	2
vorhandene Finanzen	x	x	x	x	3	5
vorhandene VG-Akzeptanz	x	x	x	x	1	5

Kendall $\tau$   Signifikanz	0.492	$p < 0.001$	0.638	$p < 0.001$	0.486	$p < 0.001$
Durchschnittliches Kendall $\tau$	<u>0.561</u> $p < 0.001$					
Spearman $\rho$   Signifikanz	0.551	$p < 0.001$	0.709	$p < 0.001$	0.599	$p < 0.001$
Durchschnittliches Spearman $\rho$	<u>0.649</u> $p < 0.001$					

## Anhang I

Fragebogen in allen Sprachversionen

### Deutsche Version

Sehr geehrte HR-Mitarbeitende

Ziel dieser Befragung ist es, Ihre persönlichen Meinungen und die Handhabung in Ihrem Unternehmen zu erfassen. Durch Ihre Antworten kann ein Überblick über den Testeinsatz in der Schweiz generiert werden, sowie Faktoren, die für oder gegen den Einsatz sprechen.

Zum Aufbau des Fragebogens: zuerst werden Ihnen für- und gegensprechende Aussagen hinsichtlich psychologischer Tests präsentiert. Diese wurden aus einer vorgelagerten qualitativen Interviewstudie und aus Literaturrecherche erhalten. Diese Aussagen sollen durch Ihre Teilnahme an der Umfrage nun auf Relevanz und Vollständigkeit geprüft werden. Anschliessend folgen einige wenige demografische Fragen zu Ihrer Person und zum Unternehmen, in dem Sie tätig sind.

Bitte beachten Sie:

- Das Ausfüllen des Fragebogens dauert ca. 15 Minuten.
- Bitte füllen Sie den Fragebogen bis 05.04.2024 aus.
- Alle Daten werden anonym erhoben. Ein Rückschluss auf Ihre Person ist nicht möglich.
- Allfällige personenbezogenen Daten (z.B. E-Mail-Adresse) werden streng vertraulich behandelt.
- Ihre Angaben werden ausschliesslich für diese Masterthesis verwendet und nach deren Abgabe unwiderruflich gelöscht.
- In meiner Masterarbeit werden zusammengefasste Ergebnisse dieser Umfrage enthalten sein.

Wenn Sie mit der Bearbeitung des Fragebogens weiterfahren, stimmen Sie den oben aufgeführten Punkten zu. Klicken Sie dazu auf «Weiter», um zur nächsten Seite zu gelangen.

Wie sehr treffen die folgenden Aussagen zu?

Bitte wählen Sie eine Antwortmöglichkeit aus.

Antwortskala: trifft überhaupt nicht zu, trifft überwiegend nicht zu, trifft eher nicht zu, trifft eher zu, trifft überwiegend zu, trifft voll und ganz zu

Den Einsatz von psychologischen Tests im Rahmen von Personalentscheidungen empfinde ich als sinnvoll, weil ...

.. ich selbst positive Erfahrungen damit gemacht habe.

.. man dadurch Einarbeitungsbedarfe feststellen kann.

.. man die Testergebnisse auch zu Personalentwicklungszwecken verwenden kann.

.. man die Eignung von Bewerbenden zur besetzenden Stelle feststellen kann.

.. man die Passung zwischen Bewerbenden und bestehendem Team feststellen kann.

.. man die kulturelle Passung zwischen Bewerbenden und dem Unternehmen feststellen kann.

.. deren Resultate genau sind.

.. die gewonnenen Resultate auf zukünftige Leistungen von Bewerbenden schliessen lassen.

.. mich deren wissenschaftliche Fundierung überzeugt.

.. man dadurch zwischen verschiedenen Bewerbenden vergleichen kann.

.. die Normen und Benchmarks eine Einordnung der Testresultate ermöglichen.

.. man für verschiedene Bewerbende gleiche Voraussetzungen schaffen kann.

.. ich deren Durchführung als praktisch empfinde.

.. dadurch die Entscheidungssicherheit erhöht wird.

.. man dadurch weitere Informationen über Bewerbende erhalten kann.

.. es den Prozess der Personalentscheidung objektiviert.

.. dadurch Fehlbesetzungen minimiert werden können.

.. sich die Testergebnisse in einem weiteren Interview aufgreifen lassen.

.. ich darin einen langfristigen finanziellen Nutzen erkenne (Return on Investment).

.. da sie helfen, eine Vorauswahl unter den erhaltenen Bewerbungen zu treffen.

Sehen Sie weitere Vorteile in Bezug auf den Einsatz psychologischer Tests zu Personalentscheiden?  
(Wenn ja: bitte nutzen Sie das untenstehende Textfeld! Wenn nein: klicken Sie auf «Weiter»!)

Den Einsatz von psychologischen Tests im Rahmen von Personalentscheidungen empfinde ich als problematisch, weil ...
.. die Kosten dafür zu hoch sind (Lizenzkosten).
.. der dafür erforderliche Zeitaufwand zu hoch ist.
.. geeignete Bewerbende ihre Bewerbung deshalb zurückziehen.
.. dadurch die Arbeitgebendenattraktivität sinkt.
.. Bewerbende solche Verfahren nicht akzeptieren.
.. sie einem Scheingenauigkeit vorgaukeln.
.. deren Formulierungen nicht für alle Anforderungsstufen verständlich sind.
.. deren Fragen zu allgemein gehalten sind.
Sehen Sie weitere Nachteile in Bezug auf den Einsatz psychologischer Tests zu Personalentscheiden? (Wenn ja: bitte nutzen Sie das untenstehende Textfeld! Wenn nein: klicken Sie auf «Weiter»!)
Im Allgemeinen befürworte ich den Einsatz psychologischer Tests im Rahmen von Personalentscheidungen. Setzen Sie in Ihrem Unternehmen psychologische Tests im Rahmen von Personalentscheiden ein? Ja, wir führen Tests intern durch; Ja, wir lassen Tests extern durchführen; Nein; Anderes Prozedere (bitte beschreiben)
Welche psychologischen Tests setzen Sie in Ihrem Unternehmen ein? (Bitte nennen Sie die Namen und gegebenenfalls die Kürzel der eingesetzten Testverfahren im untenstehenden Textfeld!)
Weshalb haben Sie sich für die genannten psychologischen Tests entschieden? (Bitte nennen Sie die Gründe im untenstehenden Textfeld!)
Für welche Zielgruppen werden Tests eingesetzt? (Mehrfachauswahl möglich!) <ul style="list-style-type: none"> <li>• ungelernete Mitarbeitende</li> <li>• Auszubildende</li> <li>• Trainees oder Hochschulabsolventinnen und -absolventen</li> <li>• Ausgebildete Mitarbeitende ohne Führungsfunktion</li> <li>• Führungskräfte untere Ebene</li> <li>• Führungskräfte mittlere Ebene</li> <li>• Führungskräfte obere Ebene</li> <li>• andere Zielgruppen (bitte beschreiben)</li> </ul>
Zu welchem Zweck setzen Sie die genannten psychologischen Tests ein? (Mehrfachauswahl möglich!) <ul style="list-style-type: none"> <li>• Personalauswahl</li> <li>• Personalentwicklung</li> <li>• andere Zwecke (bitte beschreiben)</li> </ul>
Auf Basis welcher Informationen wurde die Wahl für den genannten Test / für die genannten Tests getroffen? (Mehrauswahl möglich!) <ul style="list-style-type: none"> <li>• einfache Websuche</li> <li>• Fachliteratur (Fachzeitschriften und -bücher)</li> <li>• Testrezensionen</li> <li>• Austausch mit Fachexperten und -expertinnen</li> <li>• Workshops und Konferenzen</li> <li>• Kurse und Schulungen</li> <li>• Austausch mit Testanbieterinnen und -anbietern</li> <li>• keine Angabe möglich; andere Basis (bitte beschreiben)</li> </ul>
Meine persönliche Meinung hat bzw. hatte Einfluss darauf, wie mein Unternehmen den Einsatz psychologischer Tests zu Personalentscheidungen handhabt.

In meinem Unternehmen werden keine psychologischen Tests im Rahmen von Personalentscheidungen eingesetzt, weil...
... sie unseren Selektionsprozess ineffizient machen würden.
... wir nur wenige Bewerbungen erhalten und dadurch kaum Auswahlmöglichkeiten bestehen.
... wir uns da am Vorgehen konkurrierender Unternehmen orientieren.
... intern die Expertise dafür fehlt.
... es nicht zur Unternehmenskultur passt.
... die finanziellen Mittel dafür fehlen.
... die Akzeptanz der Linienvorgesetzten dafür fehlt.
In meinem Unternehmen werden psychologische Tests im Rahmen von Personalentscheidungen eingesetzt, weil...
... sie unseren Selektionsprozess effizient machen.
... wir viele Bewerbungen erhalten und damit grosse Auswahlmöglichkeiten bestehen.
... wir uns da am Vorgehen konkurrierender Unternehmen orientieren.
... intern die Expertise dafür vorhanden ist.
... es zur Unternehmenskultur passt.
... die finanziellen Mittel dafür vorhanden sind.
... das von den Linienvorgesetzten akzeptiert ist.

Ich finde, dass wissenschaftliche Forschungsergebnisse zu Eignungsdiagnostik leicht zugänglich sind.
Das vielfältige Angebot an psychologischen Tests macht die Auswahl schwer.
Welcher der folgenden Wirtschaftsabteilungen gehört das Unternehmen an, in dem Sie arbeiten? Land- und Forstwirtschaft; Verarbeitendes Gewerbe / Herstellung von Waren, Energie-, Wasser- und Umweltversorgung; Baugewerbe / Bau, Handel / Reparaturgewerbe; Verkehr und Lagerei; Gastgewerbe; Information und Kommunikation; Finanz- und Versicherungsdienstleistungen; Grundstücks- und Immobilienverwaltung; Erbringung von wissenschaftlichen und technischen Dienstleistungen, öffentliche Verwaltung; Erziehung und Unterricht; Gesundheits- und Sozialwesen; Kunst, Unterhaltung und Erholung; andere Wirtschaftsabteilung (bitte beschreiben)
Wie viele Mitarbeitende beschäftigt Ihr Unternehmen? <ul style="list-style-type: none"> <li>• 1 - 9 Mitarbeitende</li> <li>• 10 - 49 Mitarbeitende</li> <li>• 50 - 249 Mitarbeitende</li> <li>• 250 bis 999 Mitarbeitende</li> <li>• 1'000 bis 9'999 Mitarbeitende</li> <li>• 10'000 und mehr Mitarbeitende</li> </ul>
Arbeitet in Ihrer HR-Abteilung ein Psychologe bzw. eine Psychologin? <ul style="list-style-type: none"> <li>• Ja</li> <li>• Nein</li> <li>• Das weiss ich nicht</li> </ul>
Was ist Ihr höchster Ausbildungsabschluss? <ul style="list-style-type: none"> <li>• Lehrabschluss EFZ</li> <li>• HR-Assistentin / HR-Assistent</li> <li>• Personalfachmann / -frau mit eidg. Fachausweis</li> <li>• dipl. Leiter / -in Human Resources (HFP)</li> <li>• MAS Human Resource Management</li> <li>• Bachelor in Psychologie oder Wirtschaftswissenschaften</li> <li>• Master in Psychologie oder Wirtschaftswissenschaften</li> <li>• anderer Abschluss (bitte beschreiben)</li> </ul>
Über wie viele Jahre Berufserfahrung im Bereich HR verfügen Sie? (Bitte tragen Sie die ungefähre Anzahl als ganze Zahl im Textfeld ein!)
Haben Sie Fragen oder Anmerkungen? Für anonyme Anmerkungen nutzen Sie gerne das nachfolgende Textfeld. Bei Fragen wenden Sie sich gerne via E-Mail an mich: <a href="mailto:tanja.senn@students.fhnw.ch">tanja.senn@students.fhnw.ch</a>
Herzlichen Dank für Ihre Teilnahme! Sind Sie an den Ergebnissen interessiert? Ab September 2024 können Sie meine Masterarbeit unter der Website <a href="https://www.researchgate.net/profile/Tanja-S">https://www.researchgate.net/profile/Tanja-S</a> einsehen. Möchten Sie eine pdf-Datei der Masterarbeit per E-Mail erhalten? Dann hinterlassen Sie gerne im nachfolgenden Textfeld Ihre E-Mail-Adresse oder schicken Sie mir Ihre E-Mail-Adresse mit dem Betreff «Masterarbeit» an: <a href="mailto:tanja.senn@students.fhnw.ch">tanja.senn@students.fhnw.ch</a>
Vielen Dank für Ihre Teilnahme! Ihre Antworten wurden gespeichert. Sie können das Fenster nun schliessen.

Cher·e·s professionnel·le·s des ressources humaines,  
 Cette enquête vise à recueillir vos avis personnels et les méthodes pratiquées dans votre entreprise. Vos réponses serviront à générer une vue d'ensemble de l'utilisation des tests en Suisse, ainsi que des facteurs favorables ou défavorables à cet égard.

Pour la structure du questionnaire : Tout d'abord, vous seront présentées des déclarations favorables et défavorables concernant les tests psychologiques. Ces déclarations ont été obtenues à partir d'une étude qualitative précédente et d'une recherche documentaire. Votre participation à l'enquête permettra désormais d'évaluer la pertinence et l'exhaustivité de ces déclarations. Ensuite, vous serez invité·e·s à répondre à quelques questions démographiques vous concernant, ainsi qu'à votre entreprise.

Veillez noter:

- Cette enquête dure 15 minutes environ.
- Veuillez répondre à cette enquête avant le 05.04.2024.
- Toutes les données sont collectées de manière anonyme, empêchant ainsi toute possibilité d'identification de votre personne.
- Les éventuelles données à caractère personnel (p. ex. adresse e-mail) sont traitées de manière strictement confidentielle.
- Vos données seront exclusivement utilisées pour la réalisation de ce mémoire et seront définitivement supprimées après la soumission de ce dernier.
- Une synthèse des résultats de cette enquête sera incluse dans mon mémoire.

En répondant à cette enquête, vous acceptez les points mentionnés ci-dessus. Cliquez sur «Suivant» pour passer à la page suivante.

Dans quelle mesure les affirmations suivantes sont-elles justes?

Veillez choisir une réponse.

Échelle de réponse: pas du tout d'accord, pas d'accord, pas vraiment d'accord, plutôt d'accord, d'accord, entièrement d'accord

Selon moi, il est utile d'utiliser des tests psychologiques dans le cadre de décisions en matière de personnel car ...

... j'en tire moi-même des expériences positives.

... les résultats de ces tests permettent de déterminer les besoins en matière d'intégration.

... les résultats de ces tests peuvent également servir au développement du personnel.

... les résultats de ces tests permettent de déterminer si la personne correspond au poste.

... les résultats de ces tests permettent de déterminer si la personne s'intégrera bien dans l'équipe existante.

... les résultats de ces tests permettent de déterminer si la personne s'intégrera bien dans l'entreprise du point de vue culturel.

... les résultats de ces tests sont précis.

... les résultats de ces tests permettent de déduire les prestations futures de la personne.

... le fondement scientifique de ces tests me convainc.

... les résultats de ces tests permettent de comparer différents profils.

... les normes et références permettent de classer les résultats de ces tests.

... les mêmes conditions préalables peuvent être appliquées à différents profils.

... ces tests sont pratiques à mettre en œuvre.

... les résultats de ces tests nous confortent dans nos décisions.

... ces tests permettent de recueillir de plus amples informations sur la personne.

... ces tests objectivent le processus décisionnel en matière de personnel.

... ces tests permettent de limiter les erreurs de recrutement.

... les résultats de ces tests peuvent être repris dans le cadre d'un autre entretien.

... j'y perçois un avantage financier à long terme.

... les résultats de ces tests aident à faire une présélection parmi les candidatures reçues.

Voyez-vous d'autres avantages à utiliser des tests psychologiques pour les décisions en matière de personnel? (Si oui, veuillez remplir la zone de texte ci-dessous! Si non, cliquez sur «Suivant»!)

<p>Selon moi, il est problématique d'utiliser des tests psychologiques dans le cadre de décisions en matière de personnel car ...</p>
<p>... les coûts encourus dans ce cadre sont trop élevés (coûts de licence).</p>
<p>... la réalisation de ces tests prend trop de temps.</p>
<p>... la réalisation de ces tests peut inciter des personnes au profil adéquat à retirer leur candidature.</p>
<p>... la réalisation de ces tests nuit à l'attractivité de l'entreprise en tant qu'employeur.</p>
<p>... les personnes ne tolèrent pas de telles méthodes.</p>
<p>... ces tests donnent l'illusion d'une précision des résultats qui n'existe pas.</p>
<p>... la formulation de ces tests n'est pas adaptée à tous les niveaux d'exigence.</p>
<p>... les questions de ces tests sont de nature très générale.</p>
<p>Voyez-vous d'autres désavantages à utiliser des tests psychologiques pour les décisions en matière de personnel? (Si oui, veuillez remplir la zone de texte ci-dessous! Si non, cliquez sur «Suivant»!)</p>
<p>De manière générale, je suis pour l'utilisation de tests psychologiques dans le cadre de décisions en matière de personnel.</p>
<p>Avez-vous recours à des tests psychologiques dans le cadre de décisions en matière de personnel au sein de votre entreprise? Oui, nous réalisons des tests en interne; Oui, nous faisons réaliser des tests en externe; Non, autre procédure (veuillez préciser)</p>
<p>À quels tests psychologiques avez-vous recours au sein de votre entreprise? (Veuillez indiquer les noms et, le cas échéant, les abréviations des méthodes de test utilisées dans la zone de texte ci-dessous!)</p>
<p>Pour quelles raisons avez-vous opté pour les tests psychologiques mentionnés? (Veuillez indiquer les raisons dans la zone de texte ci-dessous!)</p>
<p>Pour quels groupes cibles des tests sont-ils utilisés? (Plusieurs réponses possibles!)</p> <ul style="list-style-type: none"> <li>• personnel non qualifié</li> <li>• personnes en apprentissage</li> <li>• stagiaires ou personnes diplômées d'une haute école</li> <li>• personnel formé sans fonction de direction</li> <li>• cadres de niveau inférieur</li> <li>• cadres de niveau intermédiaire</li> <li>• cadres de niveau supérieur</li> <li>• autres groupes cibles (veuillez préciser)</li> </ul>
<p>Dans quel but utilisez-vous les tests psychologiques mentionnés? (Plusieurs réponses possibles!) sélection du personnel; développement du personnel; autre (veuillez préciser)</p>
<p>Sur la base de quelles informations avez-vous choisi le test mentionné / les tests mentionnés? (Plusieurs réponses possibles!)</p> <ul style="list-style-type: none"> <li>• recherche simple sur le Web</li> <li>• littérature spécialisée (magazines et livres spécialisés)</li> <li>• critiques de tests</li> <li>• échange avec des spécialistes en la matière</li> <li>• ateliers et conférences</li> <li>• cours et formations</li> <li>• échange avec des prestataires spécialisés dans l'offre de tests</li> <li>• aucune indication possible</li> <li>• autre base (veuillez préciser)</li> </ul>
<p>Mon avis personnel a ou a eu une influence sur la manière dont mon entreprise aborde l'utilisation de tests psychologiques pour les décisions en matière de personnel.</p>

Mon entreprise n'a pas recours à des tests psychologiques dans le cadre des décisions en matière de personnel car ...
... ces tests rendraient notre processus de sélection inefficace.
... nous ne recevons que peu de candidatures et n'avons donc guère de possibilités de sélection.
... nous nous inspirons de la manière dont nos entreprises concurrentes procèdent.
... nous ne disposons pas en interne de l'expertise nécessaire à cette fin.
... la pratique consistant à utiliser des tests psychologiques est contraire à notre culture d'entreprise.
... nous ne disposons pas des ressources financières nécessaires à cette fin.
... les responsables hiérarchiques sont contre.
Mon entreprise a recours à des tests psychologiques dans le cadre des décisions en matière de personnel car ...
... ces tests améliorent l'efficacité de notre processus de sélection.
... nous recevons beaucoup de candidatures et avons donc un grand choix de possibilités de sélection.
... nous nous inspirons de la manière dont nos entreprises concurrentes procèdent.
... nous disposons en interne de l'expertise nécessaire à cette fin.
... la pratique consistant à utiliser des tests psychologiques entre dans le cadre de notre culture d'entreprise.
... nous disposons des ressources financières nécessaires à cette fin.
... les responsables hiérarchiques sont pour.

<p>Selon moi, les résultats des recherches scientifiques sur le diagnostic d'aptitude sont facilement accessibles.</p> <p>L'offre diversifiée de tests psychologiques rend la sélection difficile.</p> <p>Auquel des secteurs d'activité suivants l'entreprise dans laquelle vous travaillez appartient-elle?</p> <ul style="list-style-type: none"> <li>• agriculture et sylviculture</li> <li>• industrie manufacturière</li> <li>• production et distribution d'électricité et d'eau</li> <li>• construction</li> <li>• commerce / réparation</li> <li>• transport et entreposage</li> <li>• hébergement et restauration</li> <li>• information et communication</li> <li>• activités financières et d'assurance</li> <li>• activités immobilières</li> <li>• activités scientifiques et techniques</li> <li>• administration publique</li> <li>• enseignement</li> <li>• santé humaine et action sociale</li> <li>• arts</li> <li>• spectacles et activités récréatives</li> <li>• autre secteur d'activité (veuillez préciser)</li> </ul>
<p>Combien de personnes votre entreprise emploie-t-elle?</p> <ul style="list-style-type: none"> <li>• 1 à 9 personnes</li> <li>• 10 à 49 personnes</li> <li>• 50 à 249 personnes</li> <li>• 250 à 999 personnes</li> <li>• 1'000 à 9'999 personnes</li> <li>• 10'000 personnes et plus</li> </ul>
<p>Votre service RH compte-t-il un ou une psychologue?</p> <ul style="list-style-type: none"> <li>• Oui</li> <li>• Non</li> <li>• Je ne sais pas</li> </ul>
<p>Quel est votre plus haut diplôme de fin d'études?</p> <ul style="list-style-type: none"> <li>• Certificat de fin d'apprentissage CFC</li> <li>• Assistante RH / Assistant RH</li> <li>• Spécialiste en ressources humaines avec brevet fédéral</li> <li>• Directrice / Directeur des ressources humaines dipl. (EPS)</li> <li>• MAS Gestion des ressources humaines</li> <li>• Bachelor en psychologie ou sciences économiques</li> <li>• Master en psychologie ou sciences économiques</li> <li>• autre diplôme (veuillez préciser)</li> </ul>
<p>Combien d'années d'expérience professionnelle avez-vous dans le domaine des RH? (Veuillez indiquer un nombre approximatif sous forme de nombre entier dans la zone de texte!)</p>
<p>Avez-vous des questions ou remarques? N'hésitez pas à utiliser la zone de texte suivante pour formuler vos remarques anonymes. De même, n'hésitez pas à me contacter par e-mail en cas de questions: <a href="mailto:tanja.senn@students.fhnw.ch">tanja.senn@students.fhnw.ch</a></p>
<p>Merci beaucoup pour votre participation! Vous souhaitez connaître les résultats? À compter du mois de septembre 2024, vous pourrez consulter mon mémoire sur le site Web suivant: <a href="https://www.researchgate.net/profile/Tanja-S">https://www.researchgate.net/profile/Tanja-S</a>. Vous souhaitez recevoir mon mémoire au format PDF par e-mail? Dans ce cas, n'hésitez pas à me laisser votre adresse e-mail dans la zone de texte suivante ou à me l'envoyer par e-mail en indiquant l'objet «Mémoire»: <a href="mailto:tanja.senn@students.fhnw.ch">tanja.senn@students.fhnw.ch</a></p>
<p>Un grand merci pour votre participation. Vos réponses ont été enregistrées. Vous pouvez maintenant fermer la fenêtre.</p>

Englische Version

Dear HR Personnel

The purpose of this survey is to record your personal opinions and how things are done in your company. Your answers will be used to generate an overview of testing use in Switzerland, as well as factors weighing in favour of or against the use of testing.

For the questionnaire structure: Initially, you will encounter statements supporting and opposing psychological tests. These statements have been derived from a preceding qualitative interview study and literature review. Your participation in the survey will now serve to assess the relevance and comprehensiveness of these statements. Subsequently, you will be asked a few demographic questions about yourself and your company.

Please note:

- The questionnaire takes around 15 minutes to complete.
- Please complete the questionnaire by 05/04/2024.
- All data is gathered anonymously. It cannot be used to identify you.
- All personal data (e.g., e-mail addresses) will be kept strictly confidential.
- Your answers will be used solely for this Master's thesis and will be permanently deleted after they have been submitted.
- My Master's thesis will contain a summary of the results of this survey.

By proceeding further with this questionnaire, you agree to the points set out above. Click 'Continue' to go to the next page.

How much do the following statements apply?

Please select an answer option.

Answer scale: completely disagree; predominantly disagree; slightly disagree; predominantly agree; completely agree

I think that using psychological tests as part of personnel decisions makes sense, because...

...I have personally had good experiences with it.

...they can be used to determine training requirements.

...you can also use the test results for employee development purposes.

...they can determine the suitability of the applicant for the post to be filled.

...they show whether the applicant is a good fit with the existing team.

...they show whether the applicant is a good cultural fit with the company.

...their results are accurate.

...the results obtained indicate how applicants will perform in the future.

...I find the science behind them convincing.

...you can use them to compare between multiple applicants.

...the standards and benchmarks allow the test results to be categorised.

...they allow you to apply the same requirements to multiple applicants.

...I find them practical to carry out.

...they allow you to make more reliable decisions.

...they can provide additional information about applicants.

...they objectify the personnel selection process.

...they can minimise incorrect hires.

...the test results can be brought up in a subsequent interview.

...I recognise a long-term financial benefit in them (return on investment).

...they help make a preliminary selection from the applications received.

Do you see any further benefits in using psychological tests as part of personnel decisions?

(If so, please use the text box below! If not, click 'Continue'.)

I think that using psychological tests as part of personnel decisions is problematic because...
...the costs involved are too high (licence costs).
...the time required is too great.
...suitable applicants will withdraw their application as a result of them.
...they make the employer appear less attractive.
...applicants do not agree to such procedures.
...they create a misleading impression of accuracy.
...their wording cannot be understood at all performance levels.
...their questions are phrased too generally.
Do you see any further disadvantages in using psychological tests as part of personnel decisions? (If so, please use the text box below! If not, click 'Continue'.)
In general, I approve of using psychological tests as part of making personnel decisions.
Do you use psychological tests as part of making personnel decisions in your company? <ul style="list-style-type: none"> <li>• Yes, we conduct tests internally</li> <li>• Yes, we have tests conducted externally</li> <li>• No</li> <li>• Other procedure (please give details)</li> </ul>
Which psychological tests do you use as part of making personnel decisions in your company? (Please give the name and, where applicable, the abbreviation of the test procedures used in the text box below.)
What made you decide to use the abovementioned psychological tests? (Please give your reasons in the text box below.)
For which target groups are tests used? (Multiple answers are possible!) <ul style="list-style-type: none"> <li>• untrained employees</li> <li>• apprentices</li> <li>• trainees or university graduates</li> <li>• trained non-managerial employees</li> <li>• junior managers</li> <li>• middle managers</li> <li>• senior managers</li> <li>• other target groups (please give details)</li> </ul>
For what purpose do you use the abovementioned psychological tests? (Multiple answers are possible!) Personnel selection; personnel development; other purposes (please give details)
Based on which information was the choice for the abovementioned test(s) made? (Multiple answers are possible!) <ul style="list-style-type: none"> <li>• basic internet search</li> <li>• professional literature (journals and books)</li> <li>• test reviews</li> <li>• exchanges with experts</li> <li>• workshops and conferences</li> <li>• courses and training programmes</li> <li>• exchanges with test providers</li> <li>• no information possible</li> <li>• other basis (please give details)</li> </ul>
My personal opinion affects/affected how my company approaches the use of psychological tests for personnel decisions.

Psychological tests are not used in my company as part of making personnel decisions, because...
...they would render our selection process inefficient.
...we receive only a small number of applications and rarely need to choose between applicants.
...we base our approach on the approach of competing companies.
...we lack the expertise internally to use them.
...it does not suit our company culture.
...we do not have the financial resources for them.
...our line managers do not accept it.
Psychological tests are used in my company as part of making personnel decisions, because...
...they make our selection process efficient.
...we receive a lot of applications and have to choose between a large number of applicants.
...we base our approach on the approach of competing companies.
...the expertise is available internally.
...it suits our company culture.
...we have the financial resources for them.
...our line managers accept it.

I find that scientific research results on aptitude testing are easily accessible.
The extensive number of psychological tests makes choosing one difficult.
<p>To which of the following economic sectors does the company at which you work belong?</p> <ul style="list-style-type: none"> <li>• agriculture and forestry</li> <li>• manufacturing industries / goods manufacturing</li> <li>• energy, water, and environmental supply</li> <li>• building and construction industry</li> <li>• trade / repair industry</li> <li>• transport and storage</li> <li>• hospitality industry</li> <li>• information and communication</li> <li>• finance and insurance services</li> <li>• land and property administration</li> <li>• provision of scientific and technical services</li> <li>• public administration</li> <li>• education and teaching</li> <li>• health and social services</li> <li>• art</li> <li>• entertainment and relaxation</li> <li>• other economic sector (please give details)</li> </ul>
<p>How many people work at your company?</p> <ul style="list-style-type: none"> <li>• 1-9 employees</li> <li>• 10-49 employees</li> <li>• 50-249 employees</li> <li>• 250-999 employees</li> <li>• 1,000-9,999 employees</li> <li>• 10,000 employees or over</li> </ul>
<p>Is there a psychologist working in your HR department?</p> <ul style="list-style-type: none"> <li>• Yes</li> <li>• No</li> <li>• I do not know</li> </ul>
<p>What is your highest academic qualification?</p> <ul style="list-style-type: none"> <li>• EFZ apprentice certificate</li> <li>• HR assistant</li> <li>• federally certified personnel specialist</li> <li>• qualified Human Resources director</li> <li>• MAS Human Resource Management</li> <li>• Bachelor's in psychology or economics</li> <li>• Master's in psychology or economics</li> <li>• other qualification (please give details)</li> </ul>
<p>How many years of job experience in HR do you have? (Please enter the approximate number as a whole number in the text box)</p>
<p>Do you have any questions or comments? If you have any anonymous comments, please feel free to use the text box below. You can also send me any questions by e-mail to: <a href="mailto:tanja.senn@students.fhnw.ch">tanja.senn@students.fhnw.ch</a></p>
<p>Thank you very much for your time! Are you interested in the results? From September 2024, you will be able to view my Master's thesis on the website <a href="https://www.researchgate.net/profile/Tanja-S.">https://www.researchgate.net/profile/Tanja-S.</a> Would you like to receive a PDF version of my Master's thesis by e-mail? Please leave your e-mail address in the text box below or send me an e-mail with the subject "Master's Thesis" to: <a href="mailto:tanja.senn@students.fhnw.ch">tanja.senn@students.fhnw.ch</a>.</p>
<p>Thank you for your participation! Your responses have been saved. You may now close the window.</p>

## Anhang J

Übersicht über Testnamen, entsprechende Testabkürzungen und Information über Anbietende, operationalisierte Konstrukte und theoretische Basis

Testname	Testname Kurzform	Einfach-nennung 1-0	Mehrfach-nennung 1-0	Anzahl	P 1-0	L 1-0	Test-rezension 1-0	P-Tests	L-Tests	Rez.	Website	Abfrage von	Modell
360° (PerformanSe)	360° (PerformanSe) (P)	0	1	2	1	0	0	2	0	0	<a href="https://performanse.com/en/360-by-performanse">https://performanse.com/en/360-by-performanse</a>	Perspektiven (z.B. MA und VG) vergleichen zu den acht bis 14 verschiedenen Fähigkeitsdimensionen, die individuell zusammengestellt werden können.	
AON ADEPT 15 (Adaptive Employee Personality Test)	ADEPT15 (P)	1	0	1	1	0	0	1	0	0	<a href="https://assessment.aon.com/getmedia/c8b72fd4-d352-4537-af63-7aeae0489663/adept-15_Infolyer_DE.pdf">https://assessment.aon.com/getmedia/c8b72fd4-d352-4537-af63-7aeae0489663/adept-15_Infolyer_DE.pdf</a>	Persönlichkeitsmodell entlang der sechs Dimensionen Aufgaben-, Adaption-, Leistungs-, Teamarbeits-, Interaktions- und emotionalem Stil	x
AON mapTQ (Map Talent Questionnaire)	mapTQ (L)	0	1	2	0	1	0	0	2	0	<a href="https://assessment.aon.com/de-de/loesungen/was-ist-online-assessment">https://assessment.aon.com/de-de/loesungen/was-ist-online-assessment</a>	Kognitive Fähigkeiten werden entlang der Dimensionen deduktiv-logisches, numerisches, induktiv-logisches Denken und Arbeitsgedächtnis abgefragt.	x
Arc-en-ciel	Arc-en-ciel (P)	1	0	1	1	0	0	1	0	0	<a href="https://www.aec-disc.de/">https://www.aec-disc.de/</a>	Potenzialanalyse entlang der Abfrage von Motivation und Persönlichkeit	x
Belbin	Belbin (P)	1	0	1	1	0	0	1	0	0	<a href="https://www.belbin.com/about/belbin-team-roles">https://www.belbin.com/about/belbin-team-roles</a>	Persönlichkeitsabfrage zur Zuteilung zu einer von neun Teamrollen, die in drei Dimensionen gegliedert sind: kommunikationsorientierte, wissensorientierte und handlungsorientierte Teamrolle.	Belbin-Modell zu Teamrollen
BIP (Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung)	BIP (P)*	0	1	7	1	0	1	7	0	7	<a href="https://www.testzentrale.ch/shop/bochumer-inventar-zur-berufsbezogenen-persoeneichkeitsbeschreibung.html">https://www.testzentrale.ch/shop/bochumer-inventar-zur-berufsbezogenen-persoeneichkeitsbeschreibung.html</a>	14 Dimensionen werden entlang der persönlichen Eignungsvoraussetzungen wie Arbeitsverhalten, Berufliche Orientierung, soziale Kompetenzen psychische Konstitution abgefragt	x
CAPP (L)	CAPP (L)	1	0	1	0	1	0	0	1	0	<a href="https://assessments.careers/capp-assessment/">https://assessments.careers/capp-assessment/</a>	logisches Denken	x
Clifton Strength	Clifton Strength (P)	1	0	1	1	0	0	1	0	0	<a href="https://www.gallup.com/cliftonstrengths/de/253595/Home.aspx">https://www.gallup.com/cliftonstrengths/de/253595/Home.aspx</a>	Persönlichkeit wird entlang der Dimensionen Durchführung, Einflussnahme, Beziehungsaufbau und strategisches Denken abgefragt.	x
COMPER	COMPER (P)	1	0	1	1	0	0	1	0	0	<a href="https://www.mpsolutions.ch/site/tests.php">https://www.mpsolutions.ch/site/tests.php</a>	Selbstbeurteilung von Fähigkeiten und Selbstbild	x
d2-R	d2-R (L)*	1	0	1	0	1	1	0	1	1	<a href="https://www.testzentrale.ch/shop/test-d2-revision.html">https://www.testzentrale.ch/shop/test-d2-revision.html</a>	Aufmerksamkeit und Konzentration	x
DISC	DISG (P)	0	1	8	1	0	0	8	0	0	<a href="https://www.disg-modell.de/ueber-disg/einfuehrung/">https://www.disg-modell.de/ueber-disg/einfuehrung/</a>	Persönlichkeitstest, welcher acht Dimensionen der Persönlichkeit misst, welche am Typenmodell des Psychoanalytikers C.G. Jung orientiert sind. Durch den DISG-Test soll das eigene Verhalten kennengelernt und das Verhalten anderer verstanden werden. Der Test soll insbesondere bei Personalentwicklung (Kommunikationsverbesserung) eingesetzt werden, so die Anbieter. DISG-Abkürzung bzw. -Konstrukte: D: Dominanz, I: Initiative, S: Stetigkeit (wohl eher Hilfsbereitschaft), G: Gewissenhaftigkeit	DISC-Modell
ECHO (PerformanSe)	ECHO (P)	1	0	1	1	0	0	1	0	0	<a href="https://performanse.com/en/personality-test-echo">https://performanse.com/en/personality-test-echo</a>	Persönlichkeit des MA kennenlernen durch vier Dimensionen: Beweggründe, Soft Skills, Wissen über sich selbst, Beziehung zum Job	IDEAC-Modell (Influence, Decision, Efficacy, Agility, Collaboration)
Enneagramme	Enneagramme (P)	0	1	2	1	0	0	2	0	0	<a href="https://enneagram-personality.com/de">https://enneagram-personality.com/de</a>	drei Intelligenzzentren - mental, emotional und instinktiv - werden über neun verschiedene Funktionstypen der menschlichen Psyche erfragt.	x
F-JAS (Fleishman – Job Analyse System für eigenschaftsbezogene Anforderungsanalysen)	F-JAS (P)	1	0	1	1	0	0	1	0	0	<a href="https://www.testzentrale.ch/shop/fleishman-job-analyse-system-fuer-eigenschaftsbezogene-anforderungsanalysen.html">https://www.testzentrale.ch/shop/fleishman-job-analyse-system-fuer-eigenschaftsbezogene-anforderungsanalysen.html</a>	Tätigkeitsbeschreibung entlang der Dimensionen Kognition, Psychomotorik, physischen Merkmalen, Sensorik und sozialen Fähigkeiten und Fertigkeiten.	x

Testname	Testname Kurzform	Einfach-nennung 1-0	Mehrfach-nennung 1-0	Anzahl	P 1-0	L 1-0	Test-rezension 1-0	P-Tests	L-Tests	Rez.	Website	Abfrage von	Modell
GIA (General Intelligence Assessment, Thomas Int.)	GIA (L)	1	0	1	0	1	0	0	1	0	<a href="https://www.thomasinternational.ch/wp-content/uploads/2021/05/GIA-Informationen.pdf">https://www.thomasinternational.ch/wp-content/uploads/2021/05/GIA-Informationen.pdf</a>	Auffassungsgabe wird entlang der fünf Dimensionen logisches Denken, Auffassungsgabe, numerische Fähigkeiten, Wortbedeutung und räumliche Vorstellungskraft getestet.	x
HBDI (Herrmann Brain Dominance Instrument)	HBDI (P)	0	1	3	1	0	0	3	0	0	<a href="https://hbdi.de/hbdi-einzelprofil/">https://hbdi.de/hbdi-einzelprofil/</a>	Denkweise wird durch Selbstbeurteilung eher dem rationalen, organisationalen, experimentellen oder führenden Denken zugeordnet.	Vier-Quadranten-Modell des Gehirns
HEXACO	HEXACO (P)*	1	0	1	1	0	1	1	0	1	<a href="https://signa-assessment.com/tests-psychometriques/hexapro/">https://signa-assessment.com/tests-psychometriques/hexapro/</a>	bewertete Bereiche und Facetten Integrität vs. Emanzipation, Emotionalität vs. Distanziertheit, Extravertiertheit vs. Introvertiertheit, Gelassenheit vs. Sicherheit, Gewissenhaftigkeit vs. Unbekümmertheit, Offenheit vs. Ernsthaftigkeit	Big Six-Modell der Persönlichkeit von Ashton und Lee
Hogan Personality Inventory (HPI)	HPI (P)	0	1	2	1	0	0	2	0	0	<a href="https://www.hoganassessments.com/assessment/hogan-personality-inventory/">https://www.hoganassessments.com/assessment/hogan-personality-inventory/</a>	Persönlichkeit wird entlang der sieben Skalen Ausgeglichenheit, Ambition, soziale Umgänglichkeit, Einfühlungsvermögen, Besonnenheit, Wissbegierde und Lernansatz erfragt.	OCEAN
Intercultural Development Inventory (IDI)	IDI (P)	1	0	1	1	0	0	1	0	0	<a href="https://www.idiinventory.com/assessment-and-reports">https://www.idiinventory.com/assessment-and-reports</a>	Abfrage, ob monoculturales oder interculturales Mindset besteht, Einteilung in fünf Typen.	intercultural development Continuum
Insights	Insights (P)	0	1	9	1	0	0	9	0	0	<a href="https://insights.ch/was-ist-insights-mdi/">https://insights.ch/was-ist-insights-mdi/</a>	Dieser Test ist laut Anbieterwebsite eine Weiterentwicklung des DISG-Modells. Neben den im DISG erfassten Verhalten werden hier aber zusätzlich noch die folgenden drei Konstrukte zu erfassen versucht: Motive, emotionale Intelligenz, Verkaufskompetenz	x
ISK	ISK (P)*	1	0	1	1	0	1	1	0	1	<a href="https://www.testzentrale.ch/shop/inventar-sozialer-kompetenzen.html">https://www.testzentrale.ch/shop/inventar-sozialer-kompetenzen.html</a>	vier Subdimensionen zu soziale Orientierung, Offensivität, Selbststeuerung, Flexibilität	x
IST-2000 R / IST-Screening	IST (L)*	0	1	4	0	1	1	0	4	4	<a href="https://www.testzentrale.ch/shop/intelligenz-struktur-test.html">https://www.testzentrale.ch/shop/intelligenz-struktur-test.html</a>	Analogien, Zahlenreihen, Matrizen	verbale, numerische und figurale Intelligenz
L.A.B.E.L (Liste d'Adjectifs Bipolaires et en Echelle de Likert)	L.A.B.E.L (P)	1	0	1	1	0	0	1	0	0	<a href="https://www.mpsolutions.ch/site/tests.php">https://www.mpsolutions.ch/site/tests.php</a>	Selbstbewertung der eigenen Persönlichkeit entlang Adjektive.	x
LEA - Leadership Effectiveness analysis	LEA (P)	0	1	2	1	0	0	2	0	0	<a href="https://www.mrg.com/assessments/leadership/">https://www.mrg.com/assessments/leadership/</a>	22 Führungsverhaltensweisen, gruppiert in sechs Kernfunktionen, werden gemessen.	
Leonardo 3.4.5	Leonardo 3.4.5 (P)	0	1	4	1	0	0	4	0	0	<a href="https://leonardo345.com/de/willkommen/">https://leonardo345.com/de/willkommen/</a>	Persönlichkeit entlang vier psychologischer Dimensionen: zwischenmenschliche Beziehungen (introvertiert oder extravertiert), Umgang mit Informationen (praktisch oder konzeptuell), das Treffen von Entscheidungen (analytisch oder intuitiv) und Arbeitsmanagement (strukturiert oder offen). Zudem wird die Grundeinstellung in Bezug auf fünf strategische Polaritäten des Unternehmens: Zeit, Organisation, Wertachse (Konsens oder Fokus), Achse der Prozess- bzw. Produktorientierung sowie die Steuerungsachse (Stabilität oder Instabilität) gemessen.	x
LIFO (Life Orientations)	LIFO (P)	1	0	1	1	0	0	1	0	0	<a href="https://lifoproducts.de/methode">https://lifoproducts.de/methode</a>	Persönlichkeit wird entlang zweier Kontinuen bestimmt- Nähe zu Aufgaben vs. Nähe zu Menschen und langer Zeithorizont vs. Kurzer Zeithorizont - in vier Kategorien unterteilt: Unterstützend / Hergebend, Bewahrend / Festhaltend, Anpassend / Harmonisierend, Bestimmend / Übernehmend.	humanistischen Theorien von Erich Fromm, Carl Rogers und Peter Drucker.

Testname	Testname Kurzform	Einfach-nennung 1-0	Mehrfach-nennung 1-0	Anzahl	P 1-0	L 1-0	Test-rezension 1-0	P-Tests	L-Tests	Rez.	Website	Abfrage von	Modell
LMI (Leistungsmotivationsinventar)	LMI (P)*	1	0	1	1	0	1	1	0	1	<a href="https://www.testzentrale.ch/shop/leistungsmotivationsinventar.html">https://www.testzentrale.ch/shop/leistungsmotivationsinventar.html</a>	Berufliche Leistungsorientierung bzw. -motivation wird mit den zehn Items Beharrlichkeit, Dominanz, Engagement, Erfolgszuversicht, Flexibilität, Flow, Furchtlosigkeit, Internalität, Kompensatorische Anstrengung, Leistungsstolz, Lernbereitschaft, Schwierigkeitspräferenz, Selbstständigkeit, Selbstkontrolle, Statusorientierung, Wettbewerbsorientierung und Zielsetzung gemessen	x
M-KIT (modularer Kurz-Intelligenz-Test)	M-KIT (L)*	1	0	1	0	1	1	0	1	1	<a href="https://www.testzentrale.ch/shop/modularer-kurzintelligenztest.html">https://www.testzentrale.ch/shop/modularer-kurzintelligenztest.html</a>	fluide Intelligenz wird entlang der Dimensionen verbal, figural-bildhaft und numerisch getestet.	x
MBTI	MBTI (P)*	0	1	7	1	0	1	7	0	7	<a href="https://eu.themyersbriggs.com/de-DE/tools/MBTI/MBTI-personality-Types">https://eu.themyersbriggs.com/de-DE/tools/MBTI/MBTI-personality-Types</a>	Auf Grundlage von vier Dimensionen kann eine Person einem von 16 Persönlichkeitstypen zugeordnet werden. Die vier Dimensionen lauten: Extraversion (E) – Introversion (I), Empfindung (S) – Intuition (N), Denken (T) – Fühlen (F), Urteilen (J) – Wahrnehmen (P)	x
MOTIVA	MOTIVA (P)	1	0	1	1	0	0	1	0	0	<a href="https://www.motiva-solutions.com/motiva-individual">https://www.motiva-solutions.com/motiva-individual</a>	Persönlichkeit in den Bereichen berufliche Interessen, Motivatoren, Zufriedenheit, Kompetenzen und Stärken, Soft Skills, aktueller Stand der beruflichen Entwicklung und das Mindset werden gemessen.	x
MPA	MPA (P)	0	1	11	1	0	0	11	0	0	<a href="https://master.ch/wp-content/uploads/2020/10/Master-Persoenlichkeitsanalyse-MPA.pdf">https://master.ch/wp-content/uploads/2020/10/Master-Persoenlichkeitsanalyse-MPA.pdf</a>	Abgefragt werden hier innere Antriebskräfte, Sozial- und Arbeitsverhalten anhand von neun Persönlichkeitsdimensionen, 28 Eigenschaften und acht Teamrollen. Weitere Informationen gibt das Unternehmen über dieses Instrument nicht Preis. Einsetzen könne man das Instrument u.a. für Personalauswahl, Personal- und Teamentwicklung	x
MSA - Motivstrukturanalyse	MSA (P)	0	1	2	1	0	0	2	0	0	<a href="https://msaprofil.com/">https://msaprofil.com/</a>	18 Grundmotive in Form von Kontinuen zu Wissen, Prinzipientreue, Macht, Status, Ordnung, materielle Sicherheit, Freiheit, Beziehung, Fürsorge, Familie, Idealismus, Anerkennung, Wettkampf, Risiko, Essen, körperliche Aktivität, Sinnlichkeit und Spiritualität werden erhoben.	x
NEO-FFI / PIR	NEO-FFI / -PIR (P)*	1	0	1	1	0	1	1	0	1	<a href="https://www.testzentrale.ch/shop/neo-fuenf-faktoren-inventar-nach-costa-und-mc-crae.html">https://www.testzentrale.ch/shop/neo-fuenf-faktoren-inventar-nach-costa-und-mc-crae.html</a>	Die Persönlichkeit wird entlang der Dimensionen Offenheit, Neurotizismus, Extraversion, Verträglichkeit und Gewissenhaftigkeit erfasst.	x
Nova Profil	Nova Profil (P)	1	0	1	1	0	0	1	0	0	<a href="https://www.swissnova.swiss/nova-profil-geneve-vaud-neuchatel-zurich/">https://www.swissnova.swiss/nova-profil-geneve-vaud-neuchatel-zurich/</a>	Entlang des DISC-Modells werden die vier Faktoren D: Dominanz, I: Initiative, S: Stetigkeit (wohl eher Hilfsbereitschaft) und G: Gewissenhaftigkeit erhoben.	DISC-Modell
OPQ / OPQ32 (SHL)	OPQ / OPQ32 (P)*	0	1	3	1	0	1	3	0	3	<a href="https://online.shl.com/de/de-de/products/opq32r">https://online.shl.com/de/de-de/products/opq32r</a>	Persönlichkeit entlang der Big5	Big5
OPTO	OPTO (P)	0	1	5	1	0	0	5	0	0	<a href="https://master.ch/wp-content/uploads/2020/10/Performance-Optimierung-OPTO.pdf">https://master.ch/wp-content/uploads/2020/10/Performance-Optimierung-OPTO.pdf</a>	Persönlichkeitstest, welcher acht Dimensionen der Persönlichkeit misst, welche u.a. für das Verhalten und die Arbeitsleistung von Mitarbeitenden und Führungskräften entscheidend sind: Einfluss, Resilienz, Kooperation, Effizienz, Ausföhrung, Compliance, Agilität, Innovation. Die Grundlage für den OPTO-Test sei das Big Five Persönlichkeitsmodell.	Big5
Positive Intelligence	Positive Intelligence (P)	1	0	1	1	0	0	1	0	0	<a href="https://www.positiveintelligence.com/saboteurs/">https://www.positiveintelligence.com/saboteurs/</a>	Durch den PQ (Positive Intelligence Quotient), auch geistige Fitness, wird die relative Stärke des positiven mentalen Muskels (Sage) zum negativen mentalen Muskel (Saboteur) gemessen.	x

Testname	Testname Kurzform	Einfach-nennung 1-0	Mehrfach-nennung 1-0	Anzahl	P 1-0	L 1-0	Test-rezension 1-0	P-Tests	L-Tests	Rez.	Website	Abfrage von	Modell
Persprofile	Persprofil (P)	0	1	3	1	0	0	3	0	0	<a href="https://versusprofile.com/our-solutions/persprofile/">https://versusprofile.com/our-solutions/persprofile/</a>	Misst Verhaltenstendenzen, Motivatoren, emotionale Intelligenz und soziale Fähigkeiten.	Psychoanalytische Lehre von Carl Jung und William Marston, Forschung zu emotionaler Intelligenz durch Peter Salovey und Daniel Goleman.
PFPI (OCEAN)	PFPI (P)	0	1	2	1	0	0	2	0	0	<a href="https://www.pearsonclinical.fr/pfpi">https://www.pearsonclinical.fr/pfpi</a>	Persönlichkeit nach Big5-Modell	OCEAN
PI - Predictive index	PI (L)	0	1	5	0	1	0	0	5	0	<a href="https://de.predictiveindex.com/assessments/cognitive-assessment/">https://de.predictiveindex.com/assessments/cognitive-assessment/</a>	Lerntempo wird entlang der drei kognitiven Fähigkeitskategorien verbales, numerisches und abstraktes Denken getestet.	x
PMI (Positive Mindset Index)	PMI (P)	1	0	1	1	0	0	1	0	0	<a href="https://www.psytoolkit.org/survey-library/pmi.html">https://www.psytoolkit.org/survey-library/pmi.html</a>	Skala aus den sechs Faktoren Glück, Selbstvertrauen, Kontrolle, emotionale Stabilität, Motivation und Optimismus	x
Professional Profile 2 (Profil-Pro2)	Profil-Pro2 (P)	0	1	2	1	0	0	2	0	0	<a href="https://www.centraltest.com/solutions/professional-profile-2">https://www.centraltest.com/solutions/professional-profile-2</a>	Persönlichkeit und Motivation	x
PROFILER TONUS	Profiler Tonus (P)	1	0	1	1	0	0	1	0	0	<a href="http://www.iris-forco.com">www.iris-forco.com</a>	Persönlichkeit wird entlang von acht Dimensionen erfragt: Energieform, Aktionsstil, Zusammenarbeit, Ausdrucksweise, Problemlösung, Organisation, Motivationen, Geistform	x
ProfilsScan	ProfilsScan (P)	1	0	1	1	0	0	1	0	0	<a href="https://www.profilscan.com/science/">https://www.profilscan.com/science/</a>	Persönlichkeit, soft Skills, emotionale Intelligenz und Kognition werden erfragt.	x
Pulsions	Pulsions (P)	0	1	4	1	0	0	4	0	0	<a href="https://www.pulsions.com/en/">https://www.pulsions.com/en/</a>	Verhalten (keine weiteren Informationen!)	x
QE Pro	QE Pro (P)	1	0	1	1	0	0	1	0	0	<a href="https://www.generation-qe.com/test-qe-pro/">https://www.generation-qe.com/test-qe-pro/</a>	emotionale Intelligenz	x
QCP (Questionnaire de Comportements Professionnel)	QCP (P)	1	0	1	1	0	0	1	0	0	<a href="https://www.hogrefe.fr/produit/questionnaire-de-comportements-professionnels-qcp/#resume">https://www.hogrefe.fr/produit/questionnaire-de-comportements-professionnels-qcp/#resume</a>	13 Dimensionen, die berufliche Persönlichkeit erfassen: Einfluss haben, soziales Handeln, Verhandeln, Ehrgeiz, neue Erfahrungen suchen, neue Informationen suchen, Selbstvertrauen, Resilienz, Ausgeglichenheit, Nachdenklichkeit und vorsichtiges Handeln, Regeln respektieren, Extraversion, Freundlichkeit.	x
Sosie	Sosie (P)	1	0	1	1	0	0	1	0	0	<a href="https://www.talentlens.com/recruitment/assessments/sosie-2nd-generation-html">https://www.talentlens.com/recruitment/assessments/sosie-2nd-generation-html</a>	Persönlichkeit und Werte	x
Swift Aptitude (Saville)	Swift Aptitude (L)	1	0	1	0	1	0	0	1	0	<a href="https://www.savilleassessment.com/aptitude-tests/">https://www.savilleassessment.com/aptitude-tests/</a>	verbale, bildhafte und numerische Intelligenz werden getestet	x
PPA (Verhaltens-Profil-Analyse, Thomas Int.)	PPA (P)	0	1	3	1	0	0	3	0	0	<a href="https://www.thomas.co/de/ppa">https://www.thomas.co/de/ppa</a>	Persönlichkeit wird erhoben durch die vier Dimensionen Dominanz, Einfluss, Beständigkeit und Compliance, wodurch man schliessen kann auf Befürchtungen, Motivatoren, Werte und Verhaltenspräferenzen einer Person.	x
Wave (Saville)	Wave (P)	0	1	11	1	0	0	11	0	0	<a href="https://www.savilleassessment.com/wave/">https://www.savilleassessment.com/wave/</a>	Persönlichkeit (keine weiteren Informationen)	x
Total		28	23	131	41	10	10	115	16	17	x	x	x
Keine Angabe / Test in online-Recherche nicht gefunden			0	24	0	0	0			0	x	x	x

## Testbatterien

Namen der Testbatterie / des Anbietenden	Einfach-nennung 1-0	Mehrfach-nennung 1-0	Anzahl	P 1-0	L 1-0	Test-rezension 1-0	P-Tests	L-Tests	Rez.	Website	Abfrage von
Assessfirst	1	1	4	x	x	x				<a href="https://www.assessfirst.com/de/">https://www.assessfirst.com/de/</a>	Testbatterie mit drei Untertests zu Motivation, Persönlichkeit und Kognition.
HR Cockpit	0		1	x	x	x				<a href="https://www.constant-dialog.ch/hr/cockpit.html">https://www.constant-dialog.ch/hr/cockpit.html</a>	Enthält neun thematische Analyse-Module, die nach Bedarf zusammengestellt werden können. Die Analyse-Module befassen sich mit Mitarbeitendenbefragungen, Persönlichkeit, Kompetenzen, Führung, Werte, Gesundheit und Ergonomie.
Central Test Batterie	0		1	x	x	x					
Saville assessment Palette	1	0	2	x	x	x					
Schuhfried Testbatterie	1		1	x	x	x				<a href="https://www.schuhfried.com/de/wiener-testsystem/tests-und-test-sets/">https://www.schuhfried.com/de/wiener-testsystem/tests-und-test-sets/</a>	Das Wiener Testsystem ist eine Software, welches auf seiner Plattform verschiedenste psychologische Tests zusammenfasst. Der Kundschaft steht es frei, verschiedene Tests zu kombinieren. Das Testsystem umfasst Tests zu Aufmerksamkeit, Gedächtnis, exekutive Funktionen, Intelligenz, Psychomotorik und Persönlichkeit.
SHL	0	1	3	x	x	x				<a href="https://www.shl.com/shldirect/de/practice-tests/">https://www.shl.com/shldirect/de/practice-tests/</a>	SHL bietet verschiedene Test zu Kognition, Persönlichkeit, Simulationen, berufsbezogenen Tests an.
Tests von O*Net	0		1	x	x	x					
Tests des éditions ECPA Paris	0		1	x	x	x				<a href="https://www.pearsonclinical.fr/">https://www.pearsonclinical.fr/</a>	
PerfomanSe	1	1	2	x	x	x				<a href="https://performanse.com/en">https://performanse.com/en</a>	Enthält ca. ein Dutzend verschiedene Persönlichkeitstests, kognitiven Tests und kollektive Entwicklungstools.

## Anhang K

Korrelationen auf Testeinsatz auf Itemebene.

\* = 2-seitige Signifikanz,  $p < 0.00125$ .

Anz. unabhängige Variablen	Itemnamen	Fragebogenitems	Deskriptive Werte		Korrelationen zur abhängigen Variablen Testeinsatz
Abfrage begünstigender Beweggründe und Entscheidungskriterien					
Den Einsatz von psychologischen Tests im Rahmen von Personalentscheidungen empfinde ich als sinnvoll, weil ...					
1	pers. Erfahrung	.. ich selbst positive Erfahrungen damit gemacht habe.	M	3.94	$r_s = .498^*$
			SD	1.502	
			V	38%	
2	Einarbeitung	.. man dadurch Einarbeitungsbedarfe feststellen kann.	M	3.49	$r_s = 0.219$
			SD	1.382	
			V	40%	
3	Personalentwicklung	.. man die Testergebnisse auch zu Personalentwicklungszwecken verwenden kann.	M	4.49	$r_s = 0.359^*$
			SD	1.279	
			V	29%	
4	Person-Job-Fit	.. man die Eignung von Bewerbenden zur besetzenden Stelle feststellen kann.	M	3.84	$r_s = 0.288^*$
			SD	1.316	
			V	34%	
5	Person-Group-Fit	.. man die Passung zwischen Bewerbenden und bestehendem Team feststellen kann.	M	3.70	$r_s = 0.310$
			SD	1.224	
			V	33%	
6	cultural Fit	.. man die kulturelle Passung zwischen Bewerbenden und dem Unternehmen feststellen kann.	M	3.42	$r_s = 0.277^*$
			SD	1.392	
			V	41%	
7	Reliabilität	.. deren Resultate genau sind.	M	3.45	$r_s = 0.330^*$
			SD	1.345	
			V	39%	
8	Kriteriumsvalidität	.. die gewonnenen Resultate auf zukünftige Leistungen von Bewerbenden schliessen lassen.	M	3.11	$r_s = 0.223$
			SD	1.408	
			V	45%	
9	wissenschaftliche Fundierung	.. mich deren wissenschaftliche Fundierung überzeugt.	M	3.67	$r_s = 0.329^*$
			SD	1.327	
			V	36%	
10	Vergleich zwischen Bewerbenden	.. man dadurch zwischen verschiedenen Bewerbenden vergleichen kann.	M	4.45	$r_s = 0.288^*$
			SD	1.320	
			V	30%	
11	Benchmarks und Normen	.. die Normen und Benchmarks eine Einordnung der Testresultate ermöglichen.	M	3.66	$r_s = 0.183$
			SD	1.412	
			V	39%	
12	Fairness	.. man für verschiedene Bewerbende gleiche Voraussetzungen schaffen kann.	M	4.04	$r_s = 0.176$
			SD	1.427	
			V	35%	
13	Handhabbarkeit	.. ich deren Durchführung als praktisch empfinde.	M	3.84	$r_s = 0.394^*$
			SD	1.432	
			V	37%	
14	Entscheidungssicherheit	.. dadurch die Entscheidungssicherheit erhöht wird.	M	3.98	$r_s = 0.353^*$
			SD	1.387	
			V	35%	
15	inkrementelle Validität	.. man dadurch weitere Informationen über Bewerbende erhalten kann.	M	4.46	$r_s = 0.386^*$
			SD	1.239	
			V	28%	
16	Objektivierung	.. es den Prozess der Personalentscheidung objektiviert.	M	3.89	$r_s = 0.313^*$
			SD	1.456	
			V	37%	
17	Minimierung von Fehlbesetzungen	.. dadurch Fehlbesetzungen minimiert werden können.	M	3.55	$r_s = 0.375^*$
			SD	1.514	
			V	43%	
18	Ableitung von Interviewfragen	.. sich die Testergebnisse in einem weiteren Interview aufgreifen lassen.	M	4.17	$r_s = 0.283^*$
			SD	1.617	
			V	39%	
19	Return on Investment	.. ich darin einen langfristigen finanziellen Nutzen erkenne.	M	3.30	$r_s = 0.361^*$
			SD	1.538	
			V	47%	
20	Vorauswahl	.. sie helfen, eine Vorauswahl unter den erhaltenen Bewerbungen zu treffen.	M	3.50	$r_s = -0.058$
			SD	1.617	
			V	46%	

Anz. unab- hängige Variablen	Itemnamen	Fragebogenitems	Deskriptive Werte		Korrelationen zur abhängigen Variablen Testeinsatz	
Abfrage hemmender Entscheidungskriterien						
		Den Einsatz von psychologischen Tests im Rahmen von Personalentscheidungen empfinde ich als problematisch, weil ...				
21	Kosten	.. die Kosten dafür zu hoch sind.	M 3.34 SD 1.459 V 44%	$r_5 = -0.343^*$		
22	Zeitaufwand	.. der dafür erforderliche Zeitaufwand zu hoch ist.	M 3.09 SD 1.450 V 47%			
23	Ungewollte Selbstselektion	.. geeignete Bewerbende ihre Bewerbung deshalb zurückziehen.	M 2.86 SD 1.526 V 53%			
24	Sinkende Arbeitgebendenattraktivität	.. dadurch die Arbeitgebendenattraktivität sinkt.	M 2.55 SD 1.382 V 54%	$r_5 = -0.381^*$		
25	fehlende Bewerbendenakzeptanz	.. Bewerbende solche Verfahren nicht akzeptieren.	M 2.78 SD 1.366 V 49%			
26	Scheingenauigkeit	.. sie einem Scheingenauigkeit vorgaukeln.	M 3.46 SD 1.565 V 45%			
27	mangelnde Zielgruppenanpasstheit	.. deren Formulierungen nicht für alle Anforderungsstufen verständlich sind.	M 3.83 SD 1.537 V 40%	$r_5 = -0.248^*$		
28	generische Fragen	.. deren Fragen zu allgemein gehalten sind.	M 3.28 SD 1.333 V 41%			
Abfrage theoriegeleiteter Faktoren						
29	pers. Befürwortung	Im Allgemeinen befürworte ich den Einsatz psychologischer Tests im Rahmen von Personalentscheidungen.	M 3.94 SD 1.559 V 40%	$r_5 = 0.452^*$		
30	Einfluss pers. Meinung	Meine persönliche Meinung hat bzw. hatte Einfluss darauf, wie mein Unternehmen den Einsatz psychologischer Tests zu Personalentscheidungen handhabt.	M 3.74 SD 1.672 V 45%			
31	Scientist-practitioner gap	Ich finde, dass wissenschaftliche Forschungsergebnisse zu Eignungsdiagnostik leicht zugänglich sind.	M 3.15 SD 1.256 V 40%			
32	vielfältiges Angebot	Das vielfältige Angebot an psychologischen Tests macht die Auswahl schwer.	M 4.17 SD 1.366 V 33%	$r_5 = -0.023$		
Abfrage kontextueller Faktoren						
		positive Abfrage In meinem Unternehmen werden psychologische Tests im Rahmen von Personalentscheidungen eingesetzt, weil...	negative Abfrage. In meinem Unternehmen werden keine psychologischen Tests im Rahmen von Personalentscheidungen eingesetzt, weil...			
33	Effizienz des Selektionsprozesses	... sie unseren Selektionsprozess effizient machen.	... sie unseren Selektionsprozess ineffizient machen würden.	M 4.04 SD 1.549 V 38%	$r_5 = 0.068$	
34	Anzahl Bewerbungen / Arbeitsmarkt	... wir viele Bewerbungen erhalten und damit grosse Auswahlmöglichkeiten bestehen.	... wir nur wenige Bewerbungen erhalten und dadurch kaum Auswahlmöglichkeiten bestehen.	M 3.26 SD 1.797 V 55%		
35	Expertise	... intern die Expertise dafür vorhanden ist.	... intern die Expertise dafür fehlt.	M 3.69 SD 1.755 V 48%		
36	Unternehmenskultur	... es zur Unternehmenskultur passt.	... es nicht zur Unternehmenskultur passt.	M 3.49 SD 1.634 V 47%	$r_5 = 0.183$	
37	Finanzen	... die finanziellen Mittel dafür vorhanden sind.	... die finanziellen Mittel dafür fehlen.	M 3.87 SD 1.611 V 42%		
38	Akzeptanz durch Linienvorgesetzte	... das von den Linienvorgesetzten akzeptiert ist.	... die Akzeptanz der Linienvorgesetzten dafür fehlt.	M 3.90 SD 1.521 V 39%		
39	Trends und Verbreitung	... wir uns da am Vorgehen konkurrierender Unternehmen orientieren.	... wir uns da am Vorgehen konkurrierender Unternehmen orientieren.	M 2.25 SD 1.406 V 63%	$r_5 = 0.013$	



## Anhang M

Test auf univariate (mit SPSS berechnet) und multivariate Normalverteilung (mit online Tool *Statistik Guru* berechnet) aller intervallskalierten Items

### Prüfung auf univariate Normalverteilung

	Kolmogorov-Smirnov-Test bei einer Stichprobe										
	Parameter der Normalverteilung <sup>a,b</sup>			Extremste Differenzen			Test-statistik	Asymp. Sig. (2-s.)	Monte-Carlo-Signifikanz (2-seitig) <sup>d</sup>		
	N	M	SD	Absolut	Positiv	Negativ			Sig.	Konfidenzintervall Unter-grenze	Ober-grenze
... ich selbst positive Erfahrungen damit gemacht habe.	183	3.94	1.502	0.181	0.098	-0.181	0.181	0.000	0.000	0.000	0.000
... man dadurch Einarbeitungsbedarfe feststellen kann.	183	3.49	1.382	0.184	0.155	-0.184	0.184	0.000	0.000	0.000	0.000
... man die Testergebnisse auch zu Personalentwicklungszwecken verwenden kann.	183	4.49	1.279	0.263	0.153	-0.263	0.263	0.000	0.000	0.000	0.000
... man die Eignung von Bewerbenden zur besetzenden Stelle feststellen kann.	183	3.84	1.316	0.216	0.123	-0.216	0.216	0.000	0.000	0.000	0.000
... man die Passung zwischen Bewerbenden und bestehendem Team feststellen kann.	183	3.70	1.224	0.220	0.130	-0.220	0.220	0.000	0.000	0.000	0.000
... man die kulturelle Passung zwischen Bewerbenden und dem Unternehmen feststellen kann.	183	3.42	1.392	0.168	0.168	-0.166	0.168	0.000	0.000	0.000	0.000
... deren Resultate genau sind.	183	3.45	1.345	0.156	0.133	-0.156	0.156	0.000	0.000	0.000	0.000
... die gewonnenen Resultate auf zukünftige Leistungen von Bewerbenden schliessen lassen.	183	3.11	1.408	0.183	0.179	-0.183	0.183	0.000	0.000	0.000	0.000
... mich deren wissenschaftliche Fundierung überzeugt.	183	3.67	1.327	0.171	0.129	-0.171	0.171	0.000	0.000	0.000	0.000
... man dadurch zwischen verschiedenen Bewerbenden vergleichen kann.	183	4.45	1.320	0.229	0.121	-0.229	0.229	0.000	0.000	0.000	0.000
... die Normen und Benchmarks eine Einordnung der Testresultate ermöglichen.	183	3.66	1.412	0.196	0.110	-0.196	0.196	0.000	0.000	0.000	0.000
... man für verschiedene Bewerbende gleiche Voraussetzungen schaffen kann.	183	4.04	1.427	0.198	0.119	-0.198	0.198	0.000	0.000	0.000	0.000
... ich deren Durchführung als praktisch empfinde.	183	3.84	1.432	0.163	0.108	-0.163	0.163	0.000	0.000	0.000	0.000
... dadurch die Entscheidungssicherheit erhöht wird.	183	3.98	1.387	0.195	0.103	-0.195	0.195	0.000	0.000	0.000	0.000
... man dadurch weitere Informationen über Bewerbende erhalten kann.	183	4.46	1.239	0.270	0.162	-0.270	0.270	0.000	0.000	0.000	0.000
... es den Prozess der Personalentscheidung objektiviert.	183	3.89	1.456	0.169	0.090	-0.169	0.169	0.000	0.000	0.000	0.000
... dadurch Fehlbesetzungen minimiert werden können.	183	3.55	1.514	0.153	0.126	-0.153	0.153	0.000	0.000	0.000	0.000
... sich die Testergebnisse in einem weiteren Interview aufgreifen lassen.	183	4.17	1.617	0.232	0.129	-0.232	0.232	0.000	0.000	0.000	0.000
... ich darin einen langfristigen finanziellen Nutzen erkenne.	183	3.30	1.538	0.162	0.156	-0.162	0.162	0.000	0.000	0.000	0.000
... sie helfen, eine Vorauswahl unter den erhaltenen Bewerbungen zu treffen.	183	3.50	1.617	0.158	0.140	-0.158	0.158	0.000	0.000	0.000	0.000
... die Kosten dafür zu hoch sind.	183	3.34	1.459	0.160	0.160	-0.149	0.160	0.000	0.000	0.000	0.000
... der dafür erforderliche Zeitaufwand zu hoch ist.	183	3.09	1.450	0.189	0.189	-0.145	0.189	0.000	0.000	0.000	0.000
... geeignete Bewerbende ihre Bewerbung deshalb zurückziehen.	183	2.86	1.526	0.228	0.228	-0.111	0.228	0.000	0.000	0.000	0.000
... dadurch die Arbeitgebendenattraktivität sinkt.	183	2.55	1.382	0.222	0.222	-0.131	0.222	0.000	0.000	0.000	0.000
... Bewerbende solche Verfahren nicht akzeptieren.	183	2.78	1.366	0.218	0.218	-0.115	0.218	0.000	0.000	0.000	0.000
... sie einem Scheingenauigkeit vorgaukeln.	183	3.46	1.565	0.153	0.153	-0.132	0.153	0.000	0.000	0.000	0.000
... deren Formulierungen nicht für alle Anforderungsstufen verständlich sind.	183	3.83	1.537	0.192	0.126	-0.192	0.192	0.000	0.000	0.000	0.000
... deren Fragen zu allgemein gehalten sind.	183	3.28	1.333	0.164	0.164	-0.137	0.164	0.000	0.000	0.000	0.000
Im Allgemeinen befürworte ich den Einsatz psychologischer Tests im Rahmen von Personalentscheidungen.	183	3.94	1.559	0.194	0.117	-0.194	0.194	0.000	0.000	0.000	0.000
Einfluss persönlicher Meinung auf Testeinsatz	183	3.74	1.672	0.189	0.119	-0.189	0.189	0.000	0.000	0.000	0.000
Ich finde, dass wissenschaftliche Forschungsergebnisse zu Eignungsdiagnostik leicht zugänglich sind.	183	3.15	1.256	0.175	0.175	-0.134	0.175	0.000	0.000	0.000	0.000
Das vielfältige Angebot an psychologischen Tests macht die Auswahl schwer.	183	4.17	1.366	0.209	0.115	-0.209	0.209	0.000	0.000	0.000	0.000
neut_äEF_Effizienz_Selektionsprozess	183	4.04	1.549	0.170	0.103	-0.170	0.170	0.000	0.000	0.000	0.000
neut_äEF_Bewerbungsanzahl_Auswahlmög	183	3.26	1.797	0.196	0.196	-0.128	0.196	0.000	0.000	0.000	0.000
neut_äF_Vorhandensein_Expertise	183	3.69	1.755	0.181	0.161	-0.181	0.181	0.000	0.000	0.000	0.000
neut_äF_UKPassung	183	3.49	1.634	0.142	0.142	-0.136	0.142	0.000	0.000	0.000	0.000
neut_äF_Vorhandensein_Finanzen	183	3.87	1.611	0.173	0.140	-0.173	0.173	0.000	0.000	0.000	0.000
neut_äF_Vorhandensein_LV_Akzeptanz	183	3.90	1.521	0.188	0.107	-0.188	0.188	0.000	0.000	0.000	0.000
äEF_ZF_Orientierung_Konkurrenz	183	2.25	1.406	0.252	0.252	-0.188	0.252	0.000	0.000	0.000	0.000

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

c. Signifikanzkorrektur nach Lilliefors.

d. Lilliefors-Methode auf der Basis von 10000 Monte-Carlo-Stichproben mit Startwert 2000000.

## Tests auf multivariate Normalverteilung

179	5	4	4	5	4	3	5	3	3	4	2
180	5	5	2	5	5	1	5	5	6	6	1
181	5	6	2	2	6	1	5	5	5	5	2
182	4	5	2	5	2	2	3	2	2	4	1
183	6	5	5	5	4	2	5	4	3	6	3

Methode zur bestimmung der multivariaten Normalverteilung

Henze-Zirkler

### Multivariate Normalverteilung prüfen

#### Ergebnis

Test	HZ Teststatistik	p-Wert	Multivar. NV
Henze-Zirkler	1.0002	< .001	Multivar. NV ist nicht gegeben

## Anhang N

EFA Rohdaten, mit Jamovi berechnet

Faktorladungen

	Faktor					1 - Kommunalität
	1	2	3	4	5	
positive_Erfahrungen	0.501					0.431
Einarbeitungsbedarfe	0.393					0.703
Personalentwicklung	0.562					0.490
Eignung	0.734					0.388
PG_Fit	0.696					0.565
cultural_Fit	0.766					0.619
genaue_Resultate	0.726					0.446
prädiktive_Validität	0.804					0.543
wissenschaftliche_Fundierung	0.645					0.447
Vergleich_zwischen_Bewerbenden	0.701					0.424
Normen_Benchmarks			0.615			0.505
Fairness			0.878			0.260
Praktikabilität		-0.366	0.490			0.475
Entscheidungssicherheit	0.431		0.376			0.369
inkrementelle_Validität	0.418					0.602
Objektivierung	0.487		0.370			0.367
Minimierung_Fehlbesetzungen	0.696					0.349
Interviewfragen			0.359			0.704
Return_on_Investment	0.745					0.382
Vorauswahl			0.369			0.721
zu_hohe_Kosten		0.679				0.491
zu_hoher_Zeitaufwand		0.670				0.521
Bewerbendenrückzug		0.846				0.367
sinkende_Arbeitgebendenattraktivität		0.800				0.307
fehlende_Akzeptanz_Bewerbende		0.896			0.325	0.290
Scheingenauigkeit	-0.439					0.531
allgemeine_Adressatengerechtigkeit		0.371				0.590
allgemeine_Befürwortung	0.628					0.380
Einfluss_persönlicher_Meinung						0.934
S_P_Gap					0.462	0.787
vielfältiges_Testangebot					-0.349	0.776
neut_äEF_Effizienz_Selektionsprozess	0.485			0.315		0.651
neut_äEF_Bewerbungsanzahl_Auswahlmög				0.343		0.811
neut_äF_Vorhandensein_Expertise				0.345		0.742
neut_äF_UKPassung	0.453			0.547		0.399
neut_äF_Vorhandensein_Financen				0.561		0.635
neut_äF_Vorhandensein_LV_Akzeptanz				0.590		0.592
äEF_ZF_Orientierung_Konkurrenz						0.966
zu_allgemeine_Fragen	-0.423	0.323				0.557

Anmerkung. „Maximum likelihood“-Extraktion wurde zusammen mit einer „Promax“-Rotation verwendet

Bartlett's Test auf Sphärizität

$\chi^2$	df	p
3745	741	<.001

KMO-Kriterium für die Angemessenheit der Probennahme

	MSA
Insgesamt	0.897
positive_Erfahrungen	0.940
Einarbeitungsbedarfe	0.890
Personalentwicklung	0.941
Eignung	0.930
PG_Fit	0.902
cultural_Fit	0.890
genaue_Resultate	0.941
prädiktive_Validität	0.926
wissenschaftliche_Fundierung	0.927
Vergleich_zwischen_Bewerbenden	0.924
Normen_Benchmarks	0.884
Fairness	0.861
Praktikabilität	0.913
Entscheidungssicherheit	0.931
inkrementelle_Validität	0.927
Objektivierung	0.938
Minimierung_Fehlbesetzungen	0.943
Interviewfragen	0.836
Return_on_Investment	0.934
Vorauswahl	0.779
zu_hohe_Kosten	0.865
zu_hoher_Zeitaufwand	0.892
Bewerbendenrückzug	0.894
sinkende_Arbeitgebendenattraktivität	0.923
fehlende_Akzeptanz_Bewerbende	0.882
Scheingenauigkeit	0.899
allgemeine_Adressatengerechtigkeit	0.850
allgemeine_Befürwortung	0.956
Einfluss_persönlicher_Meinung	0.608
S_P_Gap	0.672
vielfältiges_Testangebot	0.587
neut_äEF_Effizienz_Selektionsprozess	0.868
neut_äEF_Bewerbungsanzahl_Auswahlmög	0.757
neut_äF_Vorhandensein_Expertise	0.733
neut_äF_UKPassung	0.873
neut_äF_Vorhandensein_Financen	0.621
neut_äF_Vorhandensein_LV_Akzeptanz	0.678
äEF_ZF_Orientierung_Konkurrenz	0.378
zu_allgemeine_Fragen	0.893

## Anhang O

CFA Rohdaten, mit Jamovi berechnet

Faktorladungen					
Faktor	Indikator	Schätzung	Std.-fehler	Z	p
persönliche Einstellung	allgemeine_Befürwortung	1.2432	0.1005	12.371	<.001
	positive_Erfahrungen	1.1531	0.0979	11.783	<.001
Optimierung der Entscheidungsgüte	prädiktive_Validität	0.8654	0.0958	9.033	<.001
	Return_on_Investment	1.1658	0.0977	11.938	<.001
	Minimierung_Fehlbesetzungen	1.2149	0.0937	12.970	<.001
	Objektivierung	1.1040	0.0924	11.942	<.001
	inkrementelle_Validität	0.7510	0.0846	8.878	<.001
	Entscheidungssicherheit	1.0354	0.0887	11.677	<.001
	wissenschaftliche_Fundierung	0.9854	0.0851	11.581	<.001
	genaue_Resultate	0.9966	0.0863	11.543	<.001
	Personalentwicklung	0.8293	0.0858	9.666	<.001
	cultural_Fit	0.7795	0.0967	8.063	<.001
	PG_Fit	0.7761	0.0827	9.389	<.001
	Eignung	0.9680	0.0846	11.447	<.001
	Einarbeitungsbedarfe	0.7426	0.0968	7.673	<.001
Praktikabilität	Normen_Benchmarks	1.0257	0.0973	10.537	<.001
	Vorauswahl	0.7785	0.1224	6.360	<.001
	Interviewfragen	0.7990	0.1217	6.567	<.001
	Praktikabilität	0.9322	0.1026	9.086	<.001
	Fairness	1.1033	0.0964	11.439	<.001
Herausforderungen und Risiken	Bewerbendenrückzug	1.1829	0.0981	12.064	<.001
	fehlende_Akzeptanz_Bewerbende	1.0367	0.0887	11.685	<.001
	allgemeine_Adressatengerechtigkeit	0.8550	0.1101	7.767	<.001
	zu_allgemeine_Fragen	0.7544	0.0950	7.943	<.001
	Scheingenaugigkeit	0.9295	0.1102	8.436	<.001
	sinkende_Arbeitgebendenattraktivität	1.1377	0.0863	13.187	<.001
	zu_hoher_Zeitaufwand	0.9605	0.0989	9.707	<.001
zu_hohe_Kosten	1.0068	0.0983	10.244	<.001	
Kontextuelle Faktoren	neut_äEF_Effizienz_Selektionsprozess	0.8799	0.1224	7.188	<.001
	neut_äEF_Bewerbungsanzahl_Auswahlmö	-0.0654	0.1535	-0.426	0.670
	neut_äF_Vorhandensein_Expertise	0.6335	0.1470	4.310	<.001
	neut_äF_UKPassung	1.3152	0.1235	10.647	<.001
	neut_äF_Vorhandensein_Financen	0.4833	0.1332	3.628	<.001
	neut_äF_Vorhandensein_LV_Akzeptanz	0.8365	0.1190	7.032	<.001

Faktorkovarianzen

		Schätzung	Std- fehler	Z	p
persönliche Einstellung	persönliche Einstellung	1.000 <sup>a</sup>			
	Optimierung der Entscheidungsgüte	0.962	0.0288	33.41	<.001
	Praktikabilität	0.771	0.0566	13.62	<.001
	Herausforderungen und Risiken	-0.721	0.0542	-13.29	<.001
	Kontextuelle Faktoren	0.632	0.0727	8.69	<.001
Optimierung der Entscheidungsgüte	Optimierung der Entscheidungsgüte	1.000 <sup>a</sup>			
	Praktikabilität	0.789	0.0428	18.42	<.001
	Herausforderungen und Risiken	-0.630	0.0526	-11.97	<.001
	Kontextuelle Faktoren	0.547	0.0702	7.79	<.001
Praktikabilität	Praktikabilität	1.000 <sup>a</sup>			
	Herausforderungen und Risiken	-0.500	0.0717	-6.97	<.001
	Kontextuelle Faktoren	0.424	0.0840	5.05	<.001
Herausforderungen und Risiken	Herausforderungen und Risiken	1.000 <sup>a</sup>			
	Kontextuelle Faktoren	-0.577	0.0706	-8.18	<.001
Kontextuelle Faktoren	Kontextuelle Faktoren	1.000 <sup>a</sup>			

<sup>a</sup> fester Parameter

Test auf exakten Fit

$\chi^2$	df	p
1079	517	<.001

Maße der Modellgüte

CFI	TLI	SRMR	RMSEA	RMSEA 90% CI	
				Untere	Obere
0.818	0.803	0.0810	0.0770	0.0706	0.0835

## Anhang P


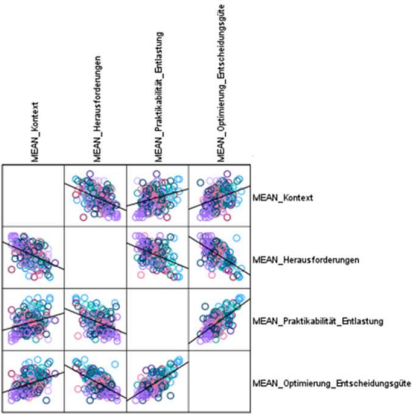
Korrelationsmatrix – Korrelationen unter allen gebildeten Faktoren und der abhängigen Variable Testeinsatz, mit SPSS berechnet

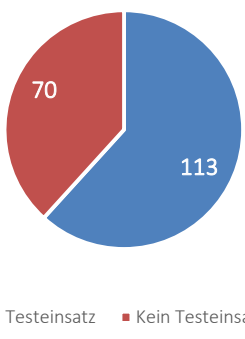
		Spearman-Korrelationen							
		Testeinsatz	MEAN_Optimierung_Entscheidungs-güte	MEAN_Praktikabilität_Entlastung	MEAN_Herausforderungen	MEAN_Kontext	MEAN_persEinstellung	Einfluss persönlicher Meinung auf Testeinsatz	
Testeinsatz	Korrelationskoeffizient	1.000	.430**	.246**	-.488**	.254**	.533**	.411**	
	Sig. (2-seitig)		0.000	0.001	0.000	0.001	0.000	0.000	
MEAN_Optimierung_Entscheidungs-güte	Korrelationskoeffizient	.430**	1.000	.626**	-.548**	.329**	.799**	.205**	
	Sig. (2-seitig)	0.000		0.000	0.000	0.000	0.000	0.005	
MEAN_Praktikabilität_Entlastung	Korrelationskoeffizient	.246**	.626**	1.000	-.361**	.233**	.576**	0.092	
	Sig. (2-seitig)	0.001	0.000		0.000	0.001	0.000	0.213	
MEAN_Herausforderungen	Korrelationskoeffizient	-.488**	-.548**	-.361**	1.000	-.394**	-.596**	-.210**	
	Sig. (2-seitig)	0.000	0.000	0.000		0.000	0.000	0.004	
MEAN_Kontext	Korrelationskoeffizient	.254**	.329**	.233**	-.394**	1.000	.406**	.197**	
	Sig. (2-seitig)	0.001	0.000	0.001	0.000		0.000	0.007	
MEAN_persEinstellung	Korrelationskoeffizient	.533**	.799**	.576**	-.596**	.406**	1.000	.295**	
	Sig. (2-seitig)	0.000	0.000	0.000	0.000	0.000		0.000	
Einfluss persönlicher Meinung auf Testeinsatz	Korrelationskoeffizient	.411**	.205**	0.092	-.210**	.197**	.295**	1.000	
	Sig. (2-seitig)	0.000	0.005	0.213	0.004	0.007	0.000		

\*\* . Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

## Anhang Q

Voraussetzungsprüfung zur Berechnung einer binären logistischen Regression, wobei «Testeinsatz» abhängige Variable verwendet wird.

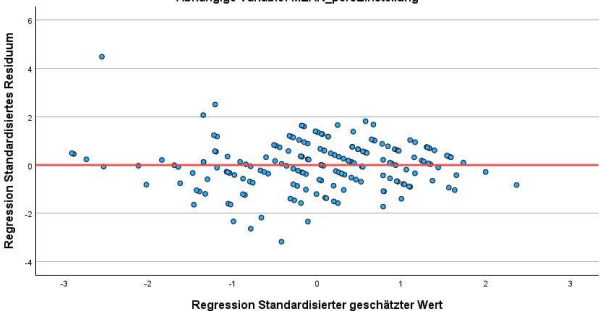
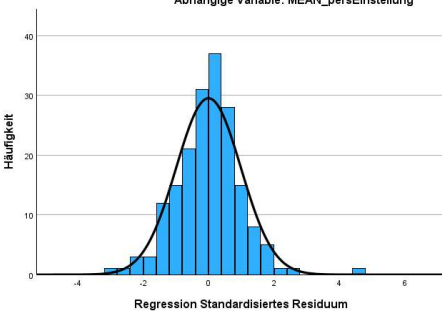
Voraussetzungen nach IBM (2021) Quelle: <a href="https://www.ibm.com/docs/de/spss-statistics/beta?topic=regression-binary-logistic">https://www.ibm.com/docs/de/spss-statistics/beta?topic=regression-binary-logistic</a>																																													
Voraussetzungen zur Berechnung einer binären logistischen Regression	erfüllt?	Ergebnis																																											
Multivariate Normalverteilung unter den unabhängigen Variablen (eher als Empfehlung zu verstehen)	ja	 <p>Wieder im online-Rechner von <i>StatistikGuru</i> berechnet, wobei bei den zusammengefassten Variablen also multivariate Normalverteilung gegeben ist – der Henze-Zirkler-Test ist signifikant geworden.</p>																																											
Abhängige Variable als echte kategoriale Variable-Gruppenzugehörigkeit soll nicht auf Werten einer stetigen Variable beruhen	ja	<p>Die abhängige Variable liegt als kategoriale Variable vor. Die Fragebogenteilnehmenden konnten lediglich zwischen drei verschiedenen Kategorien auswählen: Ja, interner Testeinsatz; Ja, externer Testeinsatz und Nein</p> <p>Es folgte schliesslich eine Variablenzusammenfassung von «Ja, interner Testeinsatz» und «Ja, externer Testeinsatz» zu «Ja, Testeinsatz», mit welcher schliesslich auch die Regression durchgeführt wird.</p>																																											
Voraussetzungen nach Methodenberatung UZH (2024) Quelle: <a href="https://www.methodenberatung.uzh.ch/de/datenanalyse_spss/zusammenhaenge/lreg.html">https://www.methodenberatung.uzh.ch/de/datenanalyse_spss/zusammenhaenge/lreg.html</a>																																													
Keine Kolinearität ( $r < 0.7$ / $VIF < 10$ ) zwischen den unabhängigen Variablen	ja	<table border="1"> <caption>Korrelationen</caption> <thead> <tr> <th></th> <th>mean_Güte_von_Test_Lund_Selektionsverfahren</th> <th>mean_Risiken_Kosten_Auswände</th> <th>mean_Praktikabilität_Durchführung_Auswertung</th> <th>mean_Bessere_Einflussfaktoren</th> </tr> </thead> <tbody> <tr> <td>mean_Güte_von_Test_Lund_Selektionsverfahren</td> <td>1</td> <td>-.583*</td> <td>.682*</td> <td>.260*</td> </tr> <tr> <td>mean_Risiken_Kosten_Auswände</td> <td></td> <td>1</td> <td>-.414*</td> <td>-.345*</td> </tr> <tr> <td>mean_Praktikabilität_Durchführung_Auswertung</td> <td></td> <td></td> <td>1</td> <td>.213*</td> </tr> <tr> <td>mean_Bessere_Einflussfaktoren</td> <td></td> <td></td> <td></td> <td>1</td> </tr> </tbody> </table> <table border="1"> <caption>Kollinearitätsstatistik</caption> <thead> <tr> <th>Modell</th> <th>Toleranz</th> <th>VIF</th> </tr> </thead> <tbody> <tr> <td>1 (Konstante)</td> <td></td> <td></td> </tr> <tr> <td>mean_Güte_von_Test_Lund_Selektionsverfahren</td> <td>.415</td> <td>2.409</td> </tr> <tr> <td>mean_Risiken_Kosten_Auswände</td> <td>.636</td> <td>1.572</td> </tr> <tr> <td>mean_Praktikabilität_Durchführung_Auswertung</td> <td>.535</td> <td>1.870</td> </tr> <tr> <td>mean_Zugänglichkeit_zu_Testen_u_Theorien</td> <td>.955</td> <td>1.047</td> </tr> </tbody> </table> <p>** Die Korrelation ist auf dem Niveau von 0.01 (2-seitig) signifikant.</p> <p>a. Abhängige Variable: Testeinsatz</p>  <p>Es besteht keine Kolinearität unter den unabhängigen Variablen, die <math>r &lt; .7</math> beträgt. Zudem ist der VIF-Wert deutlich unter 10.</p>		mean_Güte_von_Test_Lund_Selektionsverfahren	mean_Risiken_Kosten_Auswände	mean_Praktikabilität_Durchführung_Auswertung	mean_Bessere_Einflussfaktoren	mean_Güte_von_Test_Lund_Selektionsverfahren	1	-.583*	.682*	.260*	mean_Risiken_Kosten_Auswände		1	-.414*	-.345*	mean_Praktikabilität_Durchführung_Auswertung			1	.213*	mean_Bessere_Einflussfaktoren				1	Modell	Toleranz	VIF	1 (Konstante)			mean_Güte_von_Test_Lund_Selektionsverfahren	.415	2.409	mean_Risiken_Kosten_Auswände	.636	1.572	mean_Praktikabilität_Durchführung_Auswertung	.535	1.870	mean_Zugänglichkeit_zu_Testen_u_Theorien	.955	1.047
	mean_Güte_von_Test_Lund_Selektionsverfahren	mean_Risiken_Kosten_Auswände	mean_Praktikabilität_Durchführung_Auswertung	mean_Bessere_Einflussfaktoren																																									
mean_Güte_von_Test_Lund_Selektionsverfahren	1	-.583*	.682*	.260*																																									
mean_Risiken_Kosten_Auswände		1	-.414*	-.345*																																									
mean_Praktikabilität_Durchführung_Auswertung			1	.213*																																									
mean_Bessere_Einflussfaktoren				1																																									
Modell	Toleranz	VIF																																											
1 (Konstante)																																													
mean_Güte_von_Test_Lund_Selektionsverfahren	.415	2.409																																											
mean_Risiken_Kosten_Auswände	.636	1.572																																											
mean_Praktikabilität_Durchführung_Auswertung	.535	1.870																																											
mean_Zugänglichkeit_zu_Testen_u_Theorien	.955	1.047																																											
Intervallskalenniveau bei unabhängigen Variablen	ja	Die unabhängigen Variablen sind intervallskaliert und konnten auf Likert-Antwortskalen bewertet werden.																																											
Binäre Kodierung (0-1-Kodierung) bei der abhängigen Variable	ja	Die abhängige Variable ist nominal-dichotom, wobei «kein Testeinsatz» mit 0 und «Testeinsatz» mit 1 kodiert wurde.																																											

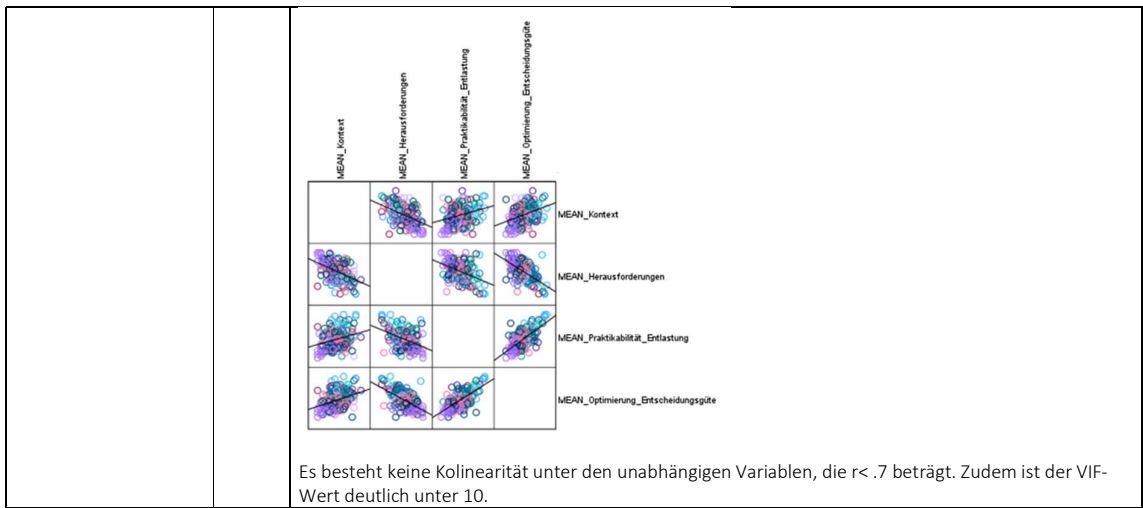
<p>Für jede Gruppe, die durch kategoriale Prädiktoren gebildet wird, gilt <math>n \geq 25</math>.</p>	<p>ja</p>	<p style="text-align: center;">Testeinsatz im Unternehmen (n = 183)</p> 	<table border="1" data-bbox="938 441 1365 592"> <thead> <tr> <th></th> <th>Anzahl Nennungen</th> <th>Prozent</th> </tr> </thead> <tbody> <tr> <td><b>Testeinsatz</b></td> <td><b>113</b></td> <td>62%</td> </tr> <tr> <td><b>Kein Testeinsatz</b></td> <td><b>70</b></td> <td>38%</td> </tr> <tr> <td>Total</td> <td>183</td> <td>100%</td> </tr> </tbody> </table>		Anzahl Nennungen	Prozent	<b>Testeinsatz</b>	<b>113</b>	62%	<b>Kein Testeinsatz</b>	<b>70</b>	38%	Total	183	100%
	Anzahl Nennungen	Prozent													
<b>Testeinsatz</b>	<b>113</b>	62%													
<b>Kein Testeinsatz</b>	<b>70</b>	38%													
Total	183	100%													
<p>Beide, durch Kategorien getrennte Datensätze, verfügen über eine Stichprobengröße weit über je 25.</p>															

## Anhang R

Voraussetzungsprüfung zur Berechnung einer linearen Regression, wobei «persönliche Einstellung» als abhängige Variable verwendet wird.

Voraussetzungen nach Gauss-Marvos Theorem Quelle: Döring (2016) und Eid, Gollwitzer & Schmitt (2017)																																																																																																											
Voraussetzungen zur Berechnung einer linearen Regression	erfüllt?	Ergebnis																																																																																																									
Annahme 1: Linearer Zusammenhang zwischen UV und AV	ja	<p style="text-align: center;"><b>Korrelationen</b></p> <table border="1"> <thead> <tr> <th></th> <th></th> <th>MEAN_persEinstellung</th> <th>MEAN_Optimierung_Entscheidungsgüte</th> <th>MEAN_Praktikabilität_Entlastung</th> <th>MEAN_Herausforderungen</th> <th>MEAN_Kontext</th> </tr> </thead> <tbody> <tr> <td>MEAN_persEinstellung</td> <td>Pearson-Korrelation</td> <td>--</td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td></td> <td>N</td> <td>183</td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>MEAN_Optimierung_Entscheidungsgüte</td> <td>Pearson-Korrelation</td> <td>.806**</td> <td>--</td> <td></td> <td></td> <td></td> </tr> <tr> <td></td> <td>Sig. (2-seitig)</td> <td>&lt;.001</td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td></td> <td>N</td> <td>183</td> <td>183</td> <td></td> <td></td> <td></td> </tr> <tr> <td>MEAN_Praktikabilität_Entlastung</td> <td>Pearson-Korrelation</td> <td>.609**</td> <td>.680**</td> <td>--</td> <td></td> <td></td> </tr> <tr> <td></td> <td>Sig. (2-seitig)</td> <td>&lt;.001</td> <td>&lt;.001</td> <td></td> <td></td> <td></td> </tr> <tr> <td></td> <td>N</td> <td>183</td> <td>183</td> <td>183</td> <td></td> <td></td> </tr> <tr> <td>MEAN_Herausforderungen</td> <td>Pearson-Korrelation</td> <td>-.603**</td> <td>-.585**</td> <td>-.414**</td> <td>--</td> <td></td> </tr> <tr> <td></td> <td>Sig. (2-seitig)</td> <td>&lt;.001</td> <td>&lt;.001</td> <td>&lt;.001</td> <td></td> <td></td> </tr> <tr> <td></td> <td>N</td> <td>183</td> <td>183</td> <td>183</td> <td>183</td> <td></td> </tr> <tr> <td>MEAN_Kontext</td> <td>Pearson-Korrelation</td> <td>.415**</td> <td>.355**</td> <td>.279**</td> <td>-.446**</td> <td>--</td> </tr> <tr> <td></td> <td>Sig. (2-seitig)</td> <td>&lt;.001</td> <td>&lt;.001</td> <td>&lt;.001</td> <td>&lt;.001</td> <td></td> </tr> <tr> <td></td> <td>N</td> <td>183</td> <td>183</td> <td>183</td> <td>183</td> <td>183</td> </tr> </tbody> </table> <p>** Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.</p> <p>Zwischen der abhängigen Variable und den latenten Faktoren bestehen erwähnenswerte Korrelationen.</p>			MEAN_persEinstellung	MEAN_Optimierung_Entscheidungsgüte	MEAN_Praktikabilität_Entlastung	MEAN_Herausforderungen	MEAN_Kontext	MEAN_persEinstellung	Pearson-Korrelation	--						N	183					MEAN_Optimierung_Entscheidungsgüte	Pearson-Korrelation	.806**	--					Sig. (2-seitig)	<.001						N	183	183				MEAN_Praktikabilität_Entlastung	Pearson-Korrelation	.609**	.680**	--				Sig. (2-seitig)	<.001	<.001					N	183	183	183			MEAN_Herausforderungen	Pearson-Korrelation	-.603**	-.585**	-.414**	--			Sig. (2-seitig)	<.001	<.001	<.001				N	183	183	183	183		MEAN_Kontext	Pearson-Korrelation	.415**	.355**	.279**	-.446**	--		Sig. (2-seitig)	<.001	<.001	<.001	<.001			N	183	183	183	183	183
		MEAN_persEinstellung	MEAN_Optimierung_Entscheidungsgüte	MEAN_Praktikabilität_Entlastung	MEAN_Herausforderungen	MEAN_Kontext																																																																																																					
MEAN_persEinstellung	Pearson-Korrelation	--																																																																																																									
	N	183																																																																																																									
MEAN_Optimierung_Entscheidungsgüte	Pearson-Korrelation	.806**	--																																																																																																								
	Sig. (2-seitig)	<.001																																																																																																									
	N	183	183																																																																																																								
MEAN_Praktikabilität_Entlastung	Pearson-Korrelation	.609**	.680**	--																																																																																																							
	Sig. (2-seitig)	<.001	<.001																																																																																																								
	N	183	183	183																																																																																																							
MEAN_Herausforderungen	Pearson-Korrelation	-.603**	-.585**	-.414**	--																																																																																																						
	Sig. (2-seitig)	<.001	<.001	<.001																																																																																																							
	N	183	183	183	183																																																																																																						
MEAN_Kontext	Pearson-Korrelation	.415**	.355**	.279**	-.446**	--																																																																																																					
	Sig. (2-seitig)	<.001	<.001	<.001	<.001																																																																																																						
	N	183	183	183	183	183																																																																																																					
Annahme 2: Werte wurden durch eine Zufallsstichprobe erhalten	ja	Die 183 Fragebogenteilnehmenden wurden durch eine Adhoc-Stichprobe erhalten. Man kann demnach von Zufall sprechen.																																																																																																									
Annahme 3: Bedingter Erwartungswert für jeden Fehlerwert der unabhängigen Variable liegt im Mittel bei 0 (Unabhängigkeit des Fehlerwerts)	ja	<p style="text-align: center;"><b>Streudiagramm</b> Abhängige Variable: MEAN_persEinstellung</p> <p>Fehlerwert sollte für jeden Prädiktor 0 annehmen, damit also die Residuen um die Nullhorizontale streuen. Das ist hier offensichtlich der Fall, womit auch diese Voraussetzung als erfüllt angeschaut werden kann.</p>																																																																																																									
Annahme 4: Exogenität der unabhängigen Variablen	ja	<p>Die Exogenität der unabhängigen Variablen wurden mittels Streudiagramm-Matrix überprüft, wobei keine offensichtliche Verletzung dieser Annahme festgestellt wurde.</p>																																																																																																									

<p>Annahme 5: Homoskedastizität (Konstanz der Varianz) der Fehlerwerte der unabhängigen Variablen</p>	<p>ja</p>	<p style="text-align: center;"><b>Streudiagramm</b> Abhängige Variable: MEAN_persEinstellung</p>  <p>Die visuelle Prüfung hat ergeben, dass kein Muster erkennbar ist; dass über den gesamten Wertebereich der geschätzten Werte die Fehlerwerte etwa die gleiche Varianz aufweisen. Homoskedastizität mit grosser Wahrscheinlichkeit also nicht vor.</p>																																																																																			
<p>Annahme 5: Normalverteilung der Fehlerwerte</p>	<p>ja</p>	<p style="text-align: center;"><b>Histogramm</b> Abhängige Variable: MEAN_persEinstellung</p>  <p><b>Kolmogorov-Smirnov-Test bei einer Stichprobe</b></p> <table border="1" data-bbox="544 703 836 871"> <thead> <tr> <th></th> <th></th> <th>Standardized Residual</th> </tr> </thead> <tbody> <tr> <td>N</td> <td></td> <td>183</td> </tr> <tr> <td>Parameter der Normalverteilung<sup>a,b</sup></td> <td>Mittelwert</td> <td>.0000000</td> </tr> <tr> <td></td> <td>Std.-Abweichung</td> <td>.98894994</td> </tr> <tr> <td rowspan="3">Extremste Differenzen</td> <td>Absolut</td> <td>.049</td> </tr> <tr> <td>Positiv</td> <td>.042</td> </tr> <tr> <td>Negativ</td> <td>-.049</td> </tr> <tr> <td>Teststatistik</td> <td></td> <td>.049</td> </tr> <tr> <td>Asymp. Sig. (2-seitig)<sup>c</sup></td> <td></td> <td>.200<sup>d</sup></td> </tr> </tbody> </table> <p>a. Die zu testende Verteilung ist eine Normalverteilung. b. Aus den Daten berechnet. c. Signifikanzkorrektur nach Lilliefors. d. Dies ist eine untere Grenze der echten Signifikanz.</p> <p>Der Test auf Normalverteilung nach Kolmogorov-Smirnov zeigt sich als nicht signifikant (<math>p &lt; 0.05</math>). Demnach sind die standardisierten Residuen normalverteilt. Die Sichtprüfung des Histogramms bestätigt dies.</p>			Standardized Residual	N		183	Parameter der Normalverteilung <sup>a,b</sup>	Mittelwert	.0000000		Std.-Abweichung	.98894994	Extremste Differenzen	Absolut	.049	Positiv	.042	Negativ	-.049	Teststatistik		.049	Asymp. Sig. (2-seitig) <sup>c</sup>		.200 <sup>d</sup>																																																										
		Standardized Residual																																																																																			
N		183																																																																																			
Parameter der Normalverteilung <sup>a,b</sup>	Mittelwert	.0000000																																																																																			
	Std.-Abweichung	.98894994																																																																																			
Extremste Differenzen	Absolut	.049																																																																																			
	Positiv	.042																																																																																			
	Negativ	-.049																																																																																			
Teststatistik		.049																																																																																			
Asymp. Sig. (2-seitig) <sup>c</sup>		.200 <sup>d</sup>																																																																																			
<p>Weitere Voraussetzungen nach Methodenberatung UZH (2024) Quelle: <a href="https://www.methodenberatung.uzh.ch/de/datenanalyse_sps/zusammenhaenge/mreg.html">https://www.methodenberatung.uzh.ch/de/datenanalyse_sps/zusammenhaenge/mreg.html</a></p>																																																																																					
<p>Keine Autokorrelation</p>	<p>ja</p>	<p><b>Modellzusammenfassung<sup>b</sup></b></p> <table border="1" data-bbox="544 1207 755 1312"> <thead> <tr> <th>Modell</th> <th>Durbin-Watson-Statistik</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>1.967<sup>a</sup></td> </tr> </tbody> </table> <p>a. Einflussvariablen: (Konstante), MEAN_Kontext, MEAN_Praktikabilität_Erntastung, MEAN_Herausforderungen, MEAN_Optimierung_Entscheidungsgüte</p> <p>b. Abhängige Variable: MEAN_persEinstellung</p> <p>Der Durbin-Watson-Wert soll zwischen 1 und 3 liegen, damit man anhand dessen davon ausgehen kann, dass keine Autokorrelation besteht. Der Wert liegt bei 1.967, womit Autokorrelation ausgeschlossen werden kann.</p>	Modell	Durbin-Watson-Statistik	1	1.967 <sup>a</sup>																																																																															
Modell	Durbin-Watson-Statistik																																																																																				
1	1.967 <sup>a</sup>																																																																																				
<p>Intervallskalenniveau bei unabhängigen Variablen</p>	<p>ja</p>	<p>Die unabhängigen Variablen sind intervallskaliert und konnten auf Likert-Antwortskalen bewertet werden.</p>																																																																																			
<p>Intervallskalenniveau bei der abhängigen Variablen</p>	<p>ja</p>	<p>Die abhängige Variable ist ebenfalls intervallskaliert und konnte auf einer Likert-Antwortskala von 1 bis 6 bewertet werden.</p>																																																																																			
<p>Keine Kollinearität (<math>r &lt; 0.7</math> / <math>VIF &lt; 10</math>) zwischen den unabhängigen Variablen</p>	<p>ja</p>	<p><b>Korrelationen</b></p> <table border="1" data-bbox="544 1585 1031 1816"> <thead> <tr> <th></th> <th>mean_Oble_von_n_Test_und_Selektionsverfahren</th> <th>mean_Risiken_Kosten_Auswände</th> <th>mean_Praktikabilität_betr_Durchführung_Auswertung</th> <th>mean_büssere_Einflussfaktoren</th> </tr> </thead> <tbody> <tr> <td>mean_Oble_von_n_Test_und_Selektionsverfahren</td> <td>1</td> <td>-.583**</td> <td>.692**</td> <td>-.260**</td> </tr> <tr> <td>Sig. (2-seitig)</td> <td></td> <td>&lt;.001</td> <td>&lt;.001</td> <td>&lt;.001</td> </tr> <tr> <td>N</td> <td>183</td> <td>183</td> <td>183</td> <td>183</td> </tr> <tr> <td>mean_Risiken_Kosten_Auswände</td> <td>Pearson-Korrelation</td> <td>1</td> <td>-.414**</td> <td>-.345**</td> </tr> <tr> <td>Sig. (2-seitig)</td> <td></td> <td></td> <td>&lt;.001</td> <td>&lt;.001</td> </tr> <tr> <td>N</td> <td>183</td> <td>183</td> <td>183</td> <td>183</td> </tr> <tr> <td>mean_Praktikabilität_betr_Durchführung_Auswertung</td> <td>Pearson-Korrelation</td> <td>.692**</td> <td>1</td> <td>.213**</td> </tr> <tr> <td>Sig. (2-seitig)</td> <td></td> <td>&lt;.001</td> <td></td> <td>&lt;.001</td> </tr> <tr> <td>N</td> <td>183</td> <td>183</td> <td>183</td> <td>183</td> </tr> <tr> <td>mean_büssere_Einflussfaktoren</td> <td>Pearson-Korrelation</td> <td>-.260**</td> <td>-.345**</td> <td>1</td> </tr> <tr> <td>Sig. (2-seitig)</td> <td></td> <td>&lt;.001</td> <td>&lt;.001</td> <td></td> </tr> <tr> <td>N</td> <td>183</td> <td>183</td> <td>183</td> <td>183</td> </tr> </tbody> </table> <p>** Die Korrelation ist auf dem Niveau von 0.01 (2-seitig) signifikant.</p> <table border="1" data-bbox="1047 1669 1299 1816"> <thead> <tr> <th>Modell</th> <th>Toleranz</th> <th>VIF</th> </tr> </thead> <tbody> <tr> <td>1 (Konstante)</td> <td></td> <td></td> </tr> <tr> <td>mean_Oble_von_n_Test_und_Selektionsverfahren</td> <td>.415</td> <td>2.409</td> </tr> <tr> <td>mean_Risiken_Kosten_Auswände</td> <td>.636</td> <td>1.572</td> </tr> <tr> <td>mean_Praktikabilität_betr_Durchführung_Auswertung</td> <td>.535</td> <td>1.870</td> </tr> <tr> <td>mean_Zugänglichkeit_zu_Testis_u_Theorien</td> <td>.955</td> <td>1.047</td> </tr> </tbody> </table> <p>a. Abhängige Variable: Testensatz</p>		mean_Oble_von_n_Test_und_Selektionsverfahren	mean_Risiken_Kosten_Auswände	mean_Praktikabilität_betr_Durchführung_Auswertung	mean_büssere_Einflussfaktoren	mean_Oble_von_n_Test_und_Selektionsverfahren	1	-.583**	.692**	-.260**	Sig. (2-seitig)		<.001	<.001	<.001	N	183	183	183	183	mean_Risiken_Kosten_Auswände	Pearson-Korrelation	1	-.414**	-.345**	Sig. (2-seitig)			<.001	<.001	N	183	183	183	183	mean_Praktikabilität_betr_Durchführung_Auswertung	Pearson-Korrelation	.692**	1	.213**	Sig. (2-seitig)		<.001		<.001	N	183	183	183	183	mean_büssere_Einflussfaktoren	Pearson-Korrelation	-.260**	-.345**	1	Sig. (2-seitig)		<.001	<.001		N	183	183	183	183	Modell	Toleranz	VIF	1 (Konstante)			mean_Oble_von_n_Test_und_Selektionsverfahren	.415	2.409	mean_Risiken_Kosten_Auswände	.636	1.572	mean_Praktikabilität_betr_Durchführung_Auswertung	.535	1.870	mean_Zugänglichkeit_zu_Testis_u_Theorien	.955	1.047
	mean_Oble_von_n_Test_und_Selektionsverfahren	mean_Risiken_Kosten_Auswände	mean_Praktikabilität_betr_Durchführung_Auswertung	mean_büssere_Einflussfaktoren																																																																																	
mean_Oble_von_n_Test_und_Selektionsverfahren	1	-.583**	.692**	-.260**																																																																																	
Sig. (2-seitig)		<.001	<.001	<.001																																																																																	
N	183	183	183	183																																																																																	
mean_Risiken_Kosten_Auswände	Pearson-Korrelation	1	-.414**	-.345**																																																																																	
Sig. (2-seitig)			<.001	<.001																																																																																	
N	183	183	183	183																																																																																	
mean_Praktikabilität_betr_Durchführung_Auswertung	Pearson-Korrelation	.692**	1	.213**																																																																																	
Sig. (2-seitig)		<.001		<.001																																																																																	
N	183	183	183	183																																																																																	
mean_büssere_Einflussfaktoren	Pearson-Korrelation	-.260**	-.345**	1																																																																																	
Sig. (2-seitig)		<.001	<.001																																																																																		
N	183	183	183	183																																																																																	
Modell	Toleranz	VIF																																																																																			
1 (Konstante)																																																																																					
mean_Oble_von_n_Test_und_Selektionsverfahren	.415	2.409																																																																																			
mean_Risiken_Kosten_Auswände	.636	1.572																																																																																			
mean_Praktikabilität_betr_Durchführung_Auswertung	.535	1.870																																																																																			
mean_Zugänglichkeit_zu_Testis_u_Theorien	.955	1.047																																																																																			



Ich führe ein kleines Unternehmen und möchte zur Personalauswahl einen psychologischen Test einsetzen. Was kannst Du mir empfehlen?



Bei der Auswahl eines psychologischen Tests für die Personalauswahl in einem kleinen Unternehmen ist es wichtig, einen Test zu wählen, der wissenschaftlich fundiert, zuverlässig und validiert ist. Hier sind einige empfehlenswerte psychologische Tests:

### 1. Big Five Persönlichkeitstest (NEO-PI-R)

- **Beschreibung:** Misst die fünf Hauptdimensionen der Persönlichkeit: Neurotizismus, Extraversion, Offenheit, Verträglichkeit und Gewissenhaftigkeit.
- **Vorteile:** Sehr gut erforscht und breit anerkannt. Bietet umfassende Einblicke in die Persönlichkeit eines Kandidaten, was für viele verschiedene Rollen nützlich sein kann.
- **Kosten:** Variiert je nach Anbieter, aber in der Regel erschwinglich für kleine Unternehmen.

### 2. Myers-Briggs-Typenindikator (MBTI)

- **Beschreibung:** Klassifiziert Personen in 16 verschiedene Persönlichkeitstypen basierend auf vier dichotomen Merkmalen.
- **Vorteile:** Weit verbreitet und leicht verständlich. Hilft bei der Teamentwicklung und beim Verständnis der Arbeitsstile.
- **Kosten:** Relativ kostengünstig.

### 3. Hogan Personality Inventory (HPI)

- **Beschreibung:** Misst normale Persönlichkeitsmerkmale, die das berufliche Verhalten vorhersagen.
- **Vorteile:** Speziell für den beruflichen Kontext entwickelt. Konzentriert sich auf Eigenschaften, die im Arbeitsumfeld wichtig sind.
- **Kosten:** Kann etwas teurer sein, aber es gibt Optionen für kleine Unternehmen.

### 4. Situational Judgement Tests (SJT)

- **Beschreibung:** Bewertet die Fähigkeit ein  Kandidaten, mit beruflichen Situationen umzugehen.
- **Vorteile:** Präzise und relevant für spezifische Jobs. Kann maßgeschneidert werden, um



Sende eine Nachricht an ChatGPT



ChatGPT kann Fehler machen. Überprüfe wichtige Informationen.