Fachhochschule Nordwestschweiz
Hochschule für Angewandte Psychologie

Verbreitung psychometrischer Online-Tests in nordwestschweizer Unternehmen

BACHELOR THESIS

Juni 2024

Autorin
Perrin Marilou Sophie

Betreuungsperson
Dr. Gemmiti Marco

Praxispartner*in
Jörg Lienert AG
Dr. Winkler Silvan

Abstract

Die vorliegende Bachelorarbeit untersucht die Forschungsfrage, wie weit verbreitet psychometrische Online-Tests in nordwestschweizer Unternehmen sind und präsentiert wichtige Ergebnisse. Die Methodik umfasst eine Kombination aus Desk Research und einer Umfrage, die sowohl quantitative als auch qualitative Aspekte enthält. Die Umfrageergebnisse zeigen, dass die Anwendung von psychometrischen Tests in der Praxis noch nicht weit verbreitet ist. Weniger als die Hälfte aller befragten Unternehmen (N=53) nutzen Online-Tests und davon erfüllen nicht alle die psychometrischen Gütekriterien. Die vielschichtigen Gründe für die geringe Verbreitung werden in der Diskussion kritisch hinterfragt und umfassen mangelndes Wissen über die neuesten wissenschaftlichen Erkenntnisse, Vorurteile sowie knappe Ressourcen. Viele Personalverantwortliche zweifeln an der Fähigkeit der Tests, komplexe individuelle Unterschiede adäquat zu erfassen und überschätzen die Validität weniger zuverlässiger Methoden. Die Ergebnisse verdeutlichen die Notwendigkeit einer besseren Sensibilisierung für den Wert psychometrischer Tests. Diese bieten einen erkenntlichen Mehrwert, da sie objektiv und systematisch die Personalrekrutierung unterstützen und somit indirekt die Unternehmensleistung steigern können. Die Arbeit gibt Handlungsempfehlungen, die den Praxispartner, die Jörg Lienert AG, einbeziehen und auf eine multimethodale Herangehensweise abzielen. Dies betont die Relevanz der Kombination verschiedener Verfahren zur Erzielung einer umfassenderen und valideren Einschätzung der Kandidat:innen.

Anzahl Zeichen (inkl. Leerzeichen): 122’860

Keywords: Psychometrische Tests, Personalauswahl, Gütekriterien, Eignungsdiagnostik

Inhaltsverzeichnis

ABSTRACT **2**

1 EINLEITUNG **7**

1.1 VORSTELLUNG DES THEMAS UND RELEVANZ **7**

1.2 FORSCHUNGSSTAND **7**

1.3 ABLEITUNG DER FRAGESTELLUNG **8**

1.4 VORSTELLUNG DES PRAXISPARTNERS **9**

2 DEFINITIONEN UND BEGRIFFSERKLÄRUNGEN **10**

2.1 EIGNUNGSDIAGNOSTIK **10**

2.2 PERSONALAUSWAHL **11**

2.3 PSYCHOMETRISCHER TEST **11**

2.3.1 DEFINITION KOGNITIVE LEISTUNGSTESTS 12

2.3.2 DEFINITION PERSÖNLICHKEITSTESTS 12

2.4 BEGRIFFSDEFINITION ITEM **13**

2.5 KORRELATION UND KORRELATIONSKOEFFIZIENT **13**

3 THEORIETEIL **14**

3.1 PSYCHOMETRIK: GÜTEKRITERIEN IN DER TESTKONSTRUKTION **14**

3.2 OBJEKTIVITÄT **15**

3.2.1 DURCHFÜHRUNGSOBJEKTIVITÄT 16

3.2.2 AUSWERTUNGSOBJEKTIVITÄT 16

3.2.3 INTERPRETATIONSOBJEKTIVITÄT 16

3.3 RELIABILITÄT ... 16

3.3.1 INNERE KONSISTENZ ... 17

3.3.2 TEST-RETEST-RELIABILITÄT ... 17

3.4 VALIDITÄT ... 17

3.4.1 INHALTSVALIDITÄT ... 18

3.4.2 KONSTRUKTVALIDITÄT ... 18

3.4.3 KRITERIUMSVALIDITÄT ... 19

3.4.4 PROGNOSTISCHE VALIDITÄT ... 19

3.4.5 AUGENSCHEINVALIDITÄT ... 19

3.5 RELEVANTE MODELLE UND THEORIEN DER PERSONALAUSWAHL ... 20

3.5.1 RETURN ON INVESTMENT (ROI) ... 20

3.5.2 TRIMODALER ANSATZ DER EIGNUNGSDIAGNOSTIK ... 20

3.5.3 PERSON-ENVIRONMENT-FIT ... 21

3.5.4 DIGITALISIERUNG IN DER PERSONALREKRUTIERUNG ... 22

3.6 RELEVANZ VON PSYCHOMETRISCHEN ONLINE-TESTS ... 23

3.7 VERBREITUNG VON PSYCHOMETRISCHEN ONLINE-TESTS IN DER PERSONALAUSWAHL ... 25

3.8 HERAUSFORDERUNGEN UND KRITIKPUNKTE BEI DER VERWENDUNG VON ONLINE-TESTS ... 26

4 METHODIK ... 30

4.1 FORSCHUNGSDESIGN ... 30

4.2 SAMPLING ... 32

4.3 DATENERHEBUNG IN FORM EINES FRAGEBOGENS (UMFRAGE) ... 32

4.4 DATENAUFBEREITUNG UND –AUSWERTUNG ... 37

4.5 EINORDNUNG DER GÜTEKRITERIEN DER GENUTZTEN TESTS MITTELS LITERATURRECHERCHE ... 38

5 ERGEBNISSE ... 39

5.1 Darstellung der Ergebnisse 39

5.2 Anwendung der Online-Tests 40

5.3 Verbreitung von psychometrischen Online-Tests in Unternehmen der Nordwestschweiz 43

5.3.1 Auswahl Intelligenztests und Persönlichkeitstests 43

5.4 Kategorisierung der Gründe für möglichen Verzicht auf Persönlichkeitstests 45

5.5 Bewertung der Qualität der vorhandenen Tests im Hinblick auf die Gütekriterien 50

6 DISKUSSION 53

6.1 Verbreitung der Online-Tests 53

6.2 Übereinstimmung der Tests mit Berufserfolg 56

6.3 Bewertung der Qualität der vorhandenen Tests im Hinblick auf die Gütekriterien 57

7 IMPLIKATIONEN UND MÖGLICHE HANDLUNGSEMPFEHLUNGEN FÜR DEN PRAXISPARTNER 60

7.1 Multimethodialität und die Rolle der Jörg Lienert AG 60

8 FAZIT 62

8 LIMITATIONEN UND AUSBLICK 64

9 LITERATUR- UND ABBILDUNGSVERZEICHNIS 66

9.1 Literaturverzeichnis 66

9.2 Abbildungsverzeichnis 73

10 ANHANG 74

10.1 Fragebogen 74

10.2 LISTE DER BEFRAGTEN UNTERNEHMEN 74

10.3 LITERATURVERZEICHNIS BEWERTUNG GÜTEKRITERIEN ANGEWANDTE TESTS 74

1 Einleitung

1.1 Vorstellung des Themas und Relevanz

Jedes Unternehmen steht regelmässig vor der Herausforderung wie sich qualifiziertes Personal gewinnen lässt. Aufgrund der fortschreitenden Automatisierung und Digitalisierung haben die Herausforderungen bei der Personalgewinnung zugenommen, da sich zum einen die Anforderungen an die Arbeitnehmenden und zum anderen der Bedarf an Fachkräften in verschiedenen Berufsgruppen erhöht hat (Krüger, 2018). Infolgedessen kommen ausführlichen Vorüberlegungen bei der Personalrekrutierung eine hohe Bedeutung zu. Voreilige Entscheidungen bei der Personalrekrutierung können Betriebe finanziell erheblich belasten und Kleinunternehmen sogar in die Insolvenz treiben (Bartscher et al., 2012). Nicht nur eine schlechte Planung, sondern auch die Wahl des falschen Mitarbeitenden kann nennenswerte Kosten, Zeitressourcen und weitere Unannehmlichkeiten für Unternehmen verursachen. Daher empfiehlt sich eine zielgerichtete Rekrutierung aus finanzieller Hinsicht zur Begrenzung des Risikos für Fehl- und Mehrinvestitionen. Eine gezielte Planung zur richtigen Personalrekrutierung steigert die Leistungsressourcen der Unternehmen, was sich nachhaltig positiv auf den Unternehmenserfolg auswirkt (Wöhe & Döring, 2013). Die Studie von Ekwoaba et al. (2015) bestätigt ebenfalls, dass eine qualitativ hochwertige Personalrekrutierung zu verbesserten organisatorischen Leistungen führen kann. Dabei spielen die Etablierung von objektiven Rekrutierungs- und Auswahlkriterien eine wichtige Rolle. Darüber hinaus konnte die Studie belegen, dass eine sorgfältige und systematische Auswahlpraxis nicht nur hilft die richtigen Mitarbeitenden zu finden, sondern auch signifikant zur Unternehmensleistung beiträgt. Dies wird durch statistische Analyse der Daten bestätigt, die zeigen, dass objektivere Rekrutierungs- und Auswahlkriterien mit einer besseren Organisationsleistung korrelieren. Zusammengefasst bestätigt dies, dass Investitionen in eine gute Selektion während der Rekrutierung zu einer optimierten Organisationsleistung führen könnte, da qualifizierte Mitarbeitende zur Steigerung der Organisationsleistung beitragen. Dies verdeutlicht die Relevanz der Implementierung von Verfahren, die eine sorgfältige, systematische und objektive Selektion ermöglichen.

1.2 Forschungsstand

Die Forschung und Praxis ist sich einig, dass die Rekrutierung, die Personalentwicklung sowie die langfristige Bindung von passenden Mitarbeitenden erfolgsentscheidende Kriterien für Unternehmen darstellen. Darüber hinaus ist auch bekannt, dass rapid wachsende Technologien und weltwirtschaftliche Veränderungen auch

den Arbeitsmarkt beeinflussen (Steininger, 2020). So werden an Arbeitsplätzen überfachliche Qualifikationen stets relevanter. Wieso werden dann Testverfahren, welche über eine hohe Qualität verfügen im Vergleich zu anderen Verfahren selten eingesetzt, wenn bekannt ist, dass diese einen bemerkbaren Mehrwert zur Rekrutierung geeigneter Mitarbeitenden leisten können (Steininger, 2020)?

Es ist festzustellen, dass bei Mitarbeitenden, Führungskräften und Betriebsräten mehrheitlich Vorurteile oder Bedenken gegenüber dem Einsatz von Testverfahren bestehen (Steininger, 2020). Dies aus ungerechtfertigten Gründen – denn die Intelligenz, welche sich mit psychologischen kognitiven Tests messen lässt, ist die aussagekräftigste Vorhersage für Berufserfolg und Trainingsfähigkeit. Dies wurde durch mehrere Studien bestätigt (Salgado, 2017, zitiert nach Ackerman & Heggestad, 1997; Guion, 1998; Murphy, 2002; Ones et al., 2010; Reeve & Hackel, 2002; Schmidt & Hunter, 1998; Schmitt, 2014; Vinchur & Koppes, 2011).

Der Einsatz von psychologischen Tests bringt zweifellos einen Mehrwert für die Eignungsdiagnostik und somit für die Unternehmen. Infolgedessen werben Eignungsdiagnostiker schon seit Jahren für eine vermehrte Anwendung von psychometrischen Tests (Weinert & Stulle, 2015). Dabei eingeschlossen sind auch Persönlichkeitstests, welche einen erheblichen Einfluss auf die Bestimmung von Verhalten und Leistungen in Arbeitsbedingungen haben (Hough et al., 2015). Die Forschung bestätigt, dass wie bei der Vorhersage des Berufserfolges durch kognitive Tests, auch ein starker Zusammenhang zwischen Berufserfolg und den fünf grossen Persönlichkeitsfaktoren («Big Five») vorhanden ist (Hossiep et al., 2000).

1.3 Ableitung der Fragestellung

Der Forschungsstand zeigt, dass die Investition und Auswahl von passenden Verfahren in der Personalrekrutierung omnipräsent ist und im Interesse der Unternehmen sowie der Mitarbeitenden stehen. Desweiteren bestätigt die Forschung, einen starken Zusammenhang zwischen kognitiven Tests und Persönlichkeitstests sowie der Vorhersage des Berufserfolgs, was wiederum für die Implementierung dieser Tests spricht. Es bestehen, wie im vorhergehenden Kapitel beschrieben, bei Mitarbeitenden, Führungskräften und Betriebsräten mehrheitlich Vorurteile oder Bedenken gegenüber dem Einsatz solcher Testverfahren (Steininger, 2020). Die Bachelorthesis schliesst an den dargelegten Forschungsstand an und stellt die folgende Frage:

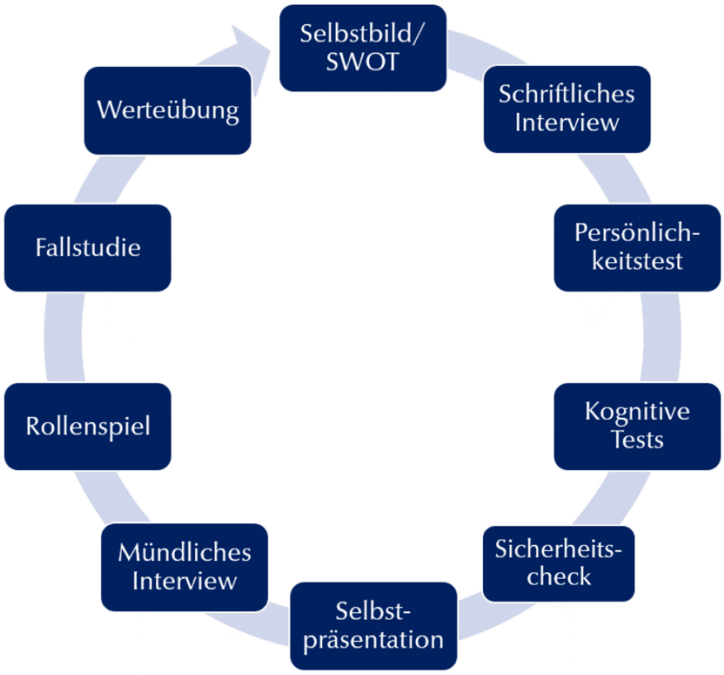
> “Wie weit verbreitet sind psychometrische Online-Tests in der Personalrekrutierung und Personalauswahl von Unternehmen in der Nordwestschweiz?”

Die Forschungsfrage wird im Auftrag des Praxispartners Jörg Lienert AG erarbeitet. Neben der Beantwortung der Fragestellung zielt die Arbeit darauf ab, Erkenntnisse und Handlungsempfehlungen für den Praxispartner zu liefern, welche nachfolgend vorgestellt werden.

1.4 Vorstellung des Praxispartners

Die Jörg Lienert AG ist seit über 40 Jahren als Unternehmen in der Personalrekrutierung und -entwicklung tätig. Dabei beschäftigt sie sich mit der Suche und Auswahl von Fach- und Führungskräften. Das Hauptziel der Jörg Lienert AG ist die Rekrutierung geeigneter Kandidat:innen für diverse Schlüsselpositionen in unterschiedlichen Unternehmen. Dieser Person-Environment-Fit (Abkürzung: P-E-Fit) wird durch Entscheidungen im Bereich Diagnostik und Assessments abgesichert (Jörg Lienert AG, o. D.). Als renommiertes Personalberatungsunternehmen verfügt die Jörg Lienert AG schweizweit über sechs Standorte. Dazu gehören Aarau, Basel, Bern, Zug, Zürich und der Hauptsitz Luzern (Wirtschaftsförderung, o. D.). Die Jörg Lienert AG ist Mitglied von Swiss Assessment und bietet eine breite Palette an Dienstleistungen im Bereich der Personalrekrutierung und -entwicklung an. Darunter gehören beispielsweise Bereiche wie Active Sourcing, klassische Direktansprache und crossmediale öffentliche Ausschreibungen. Die soeben genannten Dienstleistungen werden als Mandate bezeichnet. Folglich werden alle Mandate, von der ersten Kontaktaufnahme bis zur erfolgreichen Besetzung der Stelle, persönlich betreut. Das Besetzen einer Führungsposition ist ein komplexer Prozess, wobei es notwendig ist die verschiedenen Facetten einer Persönlichkeit ganzheitlich zu analysieren. Dies geschieht im Assessment objektiv durch verschiedene Methodenansätze (Jörg Lienert AG, o. D.). Die Jörg Lienert AG hat dafür ein eigen konzipiertes Instrument entwickelt, welches Lienert-Loop genannt wird und bei verschiedensten Rekrutierungsszenarien einen messbaren Mehrwert und eine gewinnbringende Assistenz gewährleistet. Konkret kommt dafür die Lienert-Loop Tool-Box (Abbildung 1) zum Einsatz, welche aus verschiedenen Instrumenten und Verfahren besteht.

Abbildung 1
Lienert-Loop Tool-Box der Jörg Lienert AG



Die Lienert-Loop Tool-Box zeichnet sich durch eine Kombination aus den oben dargestellten Verfahren und Übungen aus. Abhängig vom Kompetenzprofil werden die dafür geeigneten Instrumente ausgewählt. Es ist essenziell, dass die getesteten Kompetenzen vorerst mit der Kundschaft definiert und dann durch mehrere Instrumente präzise registriert werden. Das Einsetzen von vielfältigen Instrumenten erzielt eine objektive und differenzierte Evaluierung. Abschliessend werden die gewonnenen Erkenntnisse zusammengefasst und der Kundschaft präsentiert (Jörg Lienert AG, o.D.).

2 Definitionen und Begriffserklärungen

Im nachfolgenden Teil wird zuerst auf die theoretischen Grundlagen eingegangen und somit auch auf die wichtigsten Begriffe und Definitionen. Dieses Vorgehen ermöglicht eine adäquate Verständlichkeit beim Lesen dieser Arbeit.

2.1 Eignungsdiagnostik

Die Eignungsdiagnostik beschreibt eine Teildisziplin der Psychologie. Sie befasst sich mit der Zuordnung von Individuen zu Berufspositionen bzw. Arbeitstätigkeiten (Maier, o. D.). Dies erfolgt auf der Grundlage von Informationen hinsichtlich Kompetenzen der Person und einer Anforderungsanalyse des Arbeitsplatzes. Das Ziel ist eine bestmögliche Passung (P-E-Fit) zu erreichen. Die Informationen über die jeweiligen Kandidat:innen werden durch eignungsdiagnostische Instrumente, wie z.B. psychologische Tests

(Intelligenztests, Persönlichkeitstests) oder biografische Fragebögen erfasst. Weitere Erhebungsmöglichkeiten stellen ausserdem klassische Vorstellungsgespräche, in Form von unstrukturierten oder strukturierten Interviews dar oder es werden Simulationsverfahren (z.B. Assessment Center, Arbeitsprobe) verwendet (Maier, o.D.). In der vorliegenden Bachelorarbeit wird bei den psychometrischen Tests speziell auf Persönlichkeitstests und kognitive Leistungstests eingegangen.

2.2 Personalauswahl

Personalauswahl beschreibt im Allgemeinen ein Zuordnungsproblem zwischen Arbeitsplätzen und Individuen. Somit steht die Selektion von Individuen im Vordergrund. Ausserdem bestehen weitere Formen der Zuordnung wie Beratung und Platzierung. Das Ziel der Personalauswahl ist eine hohe Passung zwischen Individuum und der bestimmten Berufsposition als auch dem Unternehmen zu erreichen (Nerdinger et al., 2019). In der vorliegenden Bachelorarbeit wird Personalauswahl synonym zu Personalrekrutierung verwendet, obwohl diese sich korrekt definiert abgrenzen (Knorr, 2021). Denn genau betrachtet, ist die Personalrekrutierung der Prozess, der vor der eigentlichen Personalauswahl steht, wo es darum geht, so viele passende Kandidat:innen zu einer Bewerbung zu motivieren. Anschliessend beginnt die Phase der Personalauswahl, wo die Person ausgewählt wird, die mit ihren Kompetenzen am besten zu den Anforderungen der Arbeitsstelle passt (Knorr, 2021).

2.3 Psychometrischer Test

Nach Lienert und Raatz (1998) ist ein psychometrischer Test, ein standardisiertes wissenschaftliches Instrument, das dazu dient, ein oder mehrere bestimmte Persönlichkeitsmerkmale zu untersuchen (z.B. Intelligenz, Gewissenhaftigkeit). Das Hauptziel besteht darin, durch objektive Messungen eine quantitative Einschätzung über das Ausmass der individuellen Ausprägung dieser Merkmale zu erhalten.

Rost (2004) ergänzt diese Definition mit der Bemerkung, dass es nicht immer um quantitative Ergebnisse geht, sondern psychometrische Tests auch qualitative Aussagen ermöglichen können (z.B. die Kategorisierung von Personen entsprechend bestimmter Merkmale). Nachfolgend werden die zwei, für die Bachelorarbeit wichtigsten, psychometrischen Testarten vorgestellt.

2.3.1 Definition Kognitive Leistungstests

Kognitive Leistungstests sind individuell durchgeführte Tests, die verwendet werden, um das Intelligenzniveau einer Person zu bestimmen. Sie werden deshalb auch Intelligenztests genannt (American Psychological Association, 2023) und in der vorliegenden Bachelorarbeit auch synonym als IQ-Tests bezeichnet. Darin werden Fähigkeiten wie die Problemlösung, Konzeptbildung, rationale Schlussfolgerung, Erfassung von Details und die Bewältigung von anderen intellektuellen Aufgaben gemessen. Kognitive Leistungstests umfassen mentale, verbale und Leistungstests mit abgestuften Schwierigkeitsgraden, die vorgängig anhand einer repräsentativen Stichprobe der Bevölkerung standardisiert wurden (American Psychological Association, 2023). Es gibt eine Vielzahl von unterschiedlichen Intelligenztests und individuellen Anbieter auf dem Markt (Kanning, 2022). Laut Hogrefe (2022) wird in der Intelligenzdiagnostik grundsätzlich eine Unterscheidung zwischen Matrizentests und Intelligenzstruktur-Tests gemacht. Intelligenzstruktur-Tests erheben verschiedene Teilgebiete der Intelligenz, wie beispielsweise die figurale, numerische und verbale Intelligenz. Sie bestehen somit aus verschiedenen Bausteinen, die separat betrachtet werden können, zusammen aber eine grössere Einheit bilden (Hogrefe, 2022). Wohingegen Matrizentests die allgemeine kognitive Leistungsfähigkeit messen und die Fähigkeit der Schlussfolgerung und rationalem Denken in einem einzelnen Testwert angeben. Ausserdem sind Matrizentests überwiegend sprachfrei und zeiteffizient, während Intelligenzstruktur-Tests auch spezifischere Fähigkeiten bewerten können (Hogrefe, 2022).

Intelligenztests basieren auf theoretischen Modellen. Das Konstrukt «Intelligenz» wird dementsprechend in Intelligenzfaktoren unterteilt, um so operationalisieren zu können, dass ein psychometrischer Test auch Intelligenz misst. Hierbei gibt es viele verschiedene Modelle in der Forschung (Rost & Häcker, 2022).

2.3.2 Definition Persönlichkeitstests

Persönlichkeitstests können auch als Persönlichkeitsfragebögen beschrieben werden (Schmidt-Atzert & Amelang, 2012). Sie dienen dazu, die verschiedenen Persönlichkeitsmerkmale des Menschen zu messen. Dies wird meist mit vorformulierten Fragen oder Aussagen in schriftlicher Art erhoben. In einer zuvor festgelegten Instruktion wird die Bearbeitung des Persönlichkeitstests erläutert. Das Antwortformat ist definiert und die häufigste verwendete Art sind vorgelegte Antworten, zu denen man durch selektives Ankreuzen Zustimmung bzw. Ablehnung zum Ausdruck bringen kann. Sie stellen folglich ein Messinstrument der Personalauswahl dar, das auf dem Prinzip der Selbstauskunft beruht. So fungiert das Individuum selbst als Informationsquelle über seine eigene

Ausprägung der Persönlichkeit. Den jeweiligen Persönlichkeitstests liegt immer ein Modell zugrunde. So ist das wohl etablierteste Modell zur Beschreibung von Persönlichkeit das «Big Five», welches fünf globale Persönlichkeitsdimensionen umfasst: die Extraversion, Verträglichkeit, Gewissenhaftigkeit, Neurotizismus und Offenheit für Erfahrung (Schmidt-Atzert & Amelang, 2012). Diese grundlegenden Dimensionen können noch in dazugehörige Facetten unterteilt werden. Die «Big Five» sind für die Eignungsdiagnostik hilfreich, denn es liegen viele bestätigende Studien dazu vor. Folglich hat dieses Modell die Forschung zur psychologischen Diagnostik massgeblich beeinflusst und vorangetrieben. Ausserdem ist es auch möglich andere Verfahren zur Erhebung der Persönlichkeit, welche nicht auf dem Big Five-Modell basieren, dennoch damit in einen theoretischen Rahmen einzuordnen (Schmidt-Atzert & Amelang, 2012). So kann der Name eines Testes zwar variieren, aber inhaltlich lässt sich eine Übereinstimmung zu den Dimensionen oder Facetten der Big Five finden. Zusammenfassend kann das Big Five als ein hilfreiches Referenzsystem für diverse Persönlichkeitstests dienen (Schmidt-Atzert & Amelang, 2012).

2.4 Begriffsdefinition Item

Ein Test oder Fragebogen besteht aus verschiedenen Fragen, Aufgaben oder Aussagen (Antwortoptionen), die in der Testkonstruktion «Items» genannt werden. Somit stellen die gegebenen Antworten auf die Items Indikatoren für die latenten Merkmale (z.B. Intelligenz) dar (Witte, 2019).

Demzufolge bilden Items beobachtbare Variablen, weil sie als Indikatoren der psychologischen (verborgenen) Konstrukte fungieren. Aufgrund dessen lässt sich das Antwortverhalten der Teilnehmenden beobachten und messen (Bühner, 2021).

2.5 Korrelation und Korrelationskoeffizient

Es erfolgt eine Definition des Korrelationsbegriffs, da in der vorliegenden Arbeit der Zusammenhang zwischen den psychometrischen Tests und dem späteren Berufserfolg bzw. Arbeitsleistung betrachtet wird. Dieser wird meist in Korrelationen angegeben. Ein Zusammenhang zwischen (mindestens) zwei Variablen lässt sich durch einen Korrelationskoeffizienten messen. Dieser Zusammenhang wird auch Korrelation genannt und deutet auf eine Wechselbeziehung hin (Kuckartz et al., 2010). Bei Korrelationen stehen Variablen miteinander in Beziehung und die Vorzeichen (+ oder -) geben an, ob es sich um positive oder negative Zusammenhänge handelt (Heimsch et al., 2018).

3 Theorieteil

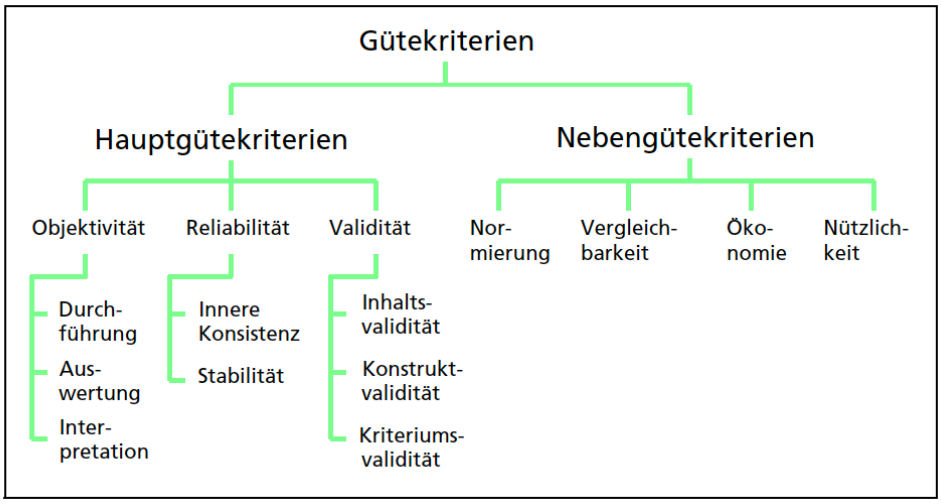
Im folgenden Abschnitt der Arbeit wird auf die wichtigsten theoretischen Grundlagen eingegangen. Dazu werden verschiedene Konstrukte der Testkonstruktion vorgestellt und relevante Modelle und Theorien der Personalauswahl erläutert.

3.1 Psychometrik: Gütekriterien in der Testkonstruktion

Wie bereits erläutert, dienen psychometrische Tests dazu, quantitative wie auch qualitative Aussagen über den relativen Grad einer individuellen Merkmalsausprägung zu treffen (Bühner, 2021).

Ausserdem verfügen psychometrische Tests über weitere Kennzeichen. Sie müssen nach der probalistischen oder der klassischen Testtheorie konstruiert sein (Bühner, 2021). Die wesentlichsten Kennzeichen eines psychometrischen Tests sind jedoch die Hauptgütekriterien: Objektivität, Reliabilität und Validität (Hock et al., 2023). Weitere Kennzeichen wie Strukturiertheit und Zugänglichkeit sind auch zu beachten (Bühner, 2021). Daneben sollten auch die Nebengütekriterien in Abbildung 2 wie Normierung, Vergleichbarkeit, Ökonomie und Nützlichkeit gewährleistet werden (Bühner, 2021). Da nach Hock et al. (2023) und in diverser Literatur die Hauptgütekriterien mit ihren Unterformen (s. Abbildung 2), als die wichtigsten Kriterien bezeichnet werden, wird in der folgenden Bachelorarbeit ausschliesslich auf diese eingegangen.

Abbildung 2
Übersicht der Haupt- und Nebengütekriterien. Adaptiert von Einführung in die Test- und Fragebogenkonstruktion (s.44) von Bühner, 2021, Pearson.



3.2 Objektivität

Objektivität besitzt den Anspruch auf Vergleichbarkeit der Testresultate von unterschiedlichen Testpersonen. Wenn die Ergebnisse des Testes und die daraus folgenden Interpretationen unabhängig von den wechselnden Bedingungen der Durchführung und Auswertung sind, gilt ein Test als objektiv. Dabei inkludiert sind die Personen, die den Test durchführen, auswerten und interpretieren. Es darf somit keinen Einfluss haben, wer als Testleiter:in fungiert. Ausserdem muss die Durchführung, Datenerhebung und Interpretation auch stets gleich und unabhängig von den Versuchspersonen stattfinden. Vorrausetzung dafür ist auch, dass ein Manual (Testanleitung) für den konzipierten Test besteht, welches so detailliert wie möglich ist (Witte, 2019). Nach Moosbrugger und Kelava (2020) muss Objektivität erfüllt sein, wenn man Resultate zwischen unterschiedlichen Testpersonen vergleichen will. Dafür werden Korrelationskoeffizienten zwischen den Ergebnissen von Tests verglichen, die mindestens von zwei unterschiedlichen versuchsleitenden Personen durchgeführt, ausgewertet und berechnet wurden. Die gemessenen Werte sollten bei der Beurteilung grösser als 0,9 sein, damit Objektivität gewährleistet werden kann (Witte, 2019). Eine hohe Objektivität ist dadurch definiert, dass jede beliebige/r Testleiter:in den Test mit allen Testpersonen in der exakt gleichen Art durchführt und gleichzeitig auch alle Testauswerter:innen die individuellen Leistungen der Testpersonen auf dieselbe Weise auswerten und interpretieren (Moosbrugger & Kelava, 2020). Es wird zwischen Durchführungs-, Auswertungs- und

Interpretationsobjektivität unterschieden. Damit sind die verschiedenen Phasen eines Testdurchlaufs von Beginn bis zum Ende gemeint.

3.2.1 Durchführungsobjektivität

Die Durchführungsobjektivität wird durch eine Standardisierung der Testdurchführung sichergestellt. Diese beinhaltet klare Anweisungen, vorgegebene Materialien, eine fixierte Art der Testvorgabe und andere relevante Merkmale (Hock et al., 2023). Eine weitere Voraussetzung dafür ist auch, dass ein Manual besteht (Witte, 2019.)

3.2.2 Auswertungsobjektivität

Auswertungsobjektivität bezeichnet die Erfassung der Daten, die während eines Tests anfallen, sowie deren Zusammenführung und Berechnung zu Testergebnissen. Dazu sind klare Vorschriften und Schablonen notwendig, welche wichtige Informationen zur Auswertung der Daten enthalten (Bühner, 2021).

3.2.3 Interpretationsobjektivität

Interpretationsobjektivität bezieht sich auf zusätzliche Schlussfolgerungen, die basierend auf den Testergebnissen gezogen werden, wie z.B. die Erkenntnis, dass die Intelligenz einer bestimmten Person im Durchschnitt liegt oder über- bzw. unterdurchschnittlich ist (Hock et al., 2023).

3.3 Reliabilität

Ein weiteres Gütekriterium ist die Reliabilität (Präzision, Zuverlässigkeit), welche die Messgenauigkeit beschreibt (Schmidt-Atzert & Amelang, 2012). Eine reliable Messung heisst, dass sie beinahe frei von (unsystematischen) Messfehlern ist. Auch hier wird die Angabe der Reliabilität in Reliabilitätskoeffizienten ausgedrückt, welche zwischen 0 und 1 variieren. Je höher der Wert ist, desto höher fällt die Reliabilität aus. Allerdings bedeutet eine hohe Reliabilität nicht automatisch, dass ein Test besser ist, denn auch hier kommt es auf weitere Faktoren an. Allgemein sagt die Reliabilität aus, wie exakt ein Merkmal durch den Test erfasst wird und ob es sich gleichzeitig auch um die angestrebte Merkmalsdimension handelt, die der Test zu erheben bestrebt (Schmidt-Atzert & Amelang, 2012). Internationalen Richtlinien zufolge muss die Reliabilität einen Wert von mindestens 0,7 besitzen (Kanning, 2022). Auch bei der Reliabilität gibt es verschiedene Unterformen, die nachfolgend erläutert werden.

3.3.1 Innere Konsistenz

Eine Unterform stellt die innere Konsistenz (interne Konsistenz) dar, welche eine sehr genaue Schätzung der Reliabilität ermöglicht. Sie wird über den Cronbachs Alpha-Koeffizienten bestimmt und ist ein Indikator für Eindimensionalität sowie für eine hohe Reliabilität. Die innere Konsistenz ist eine Erweiterung der Methode der Testhalbierung. So wird bei der internen Konsistenz der Test nicht willkürlich in zwei Hälften unterteilt, sondern die im Test verwendeten Items werden aufgeteilt und alle bivariaten Korrelationen zwischen den einzelnen Items werden ausgerechnet und gemittelt. Es ist die am meisten genutzte Unterform der Reliabilität (Döring & Bortz, 2016).

3.3.2 Test-Retest-Reliabilität

Eine weitere Unterform ist die Stabilität, welche auch Test-Retest-Reliabilität oder Testwiederholungsreliabilität genannt wird. Diese sollte nur eingesetzt werden, wenn Testverfahren sehr stabile Merkmale erheben wollen. Denn die Stabilität überprüft die Konsistenz der Ergebnisse innerhalb eines Tests wie auch die zeitliche Stabilität. Folglich wird die gleiche Untersuchung nach einem bestimmten Zeitraum (meistens 2-3 Wochen Zeitabstand zwischen den verschiedenen Messungen) wiederholt und somit die Übereinstimmung der Ergebnisse gemessen (Döring & Bortz, 2016). Weitere Unterformen der Reliabilität umfassen: Paralleltest-Reliabilität und Testhalbierungs-Reliabilität, auf welche aufgrund der geringen Relevanz für die vorliegende Arbeit nicht näher eingegangen wird (Döring & Bortz, 2016).

3.4 Validität

Das letzte Hauptgütekriterium ist die Validität und dieses gilt nach Moosbrugger und Kelava (2020) als das wichtigste in Bezug auf die praktische Anwendbarkeit von Testverfahren und den theoretischen Grundlagen von Merkmalszusammenhängen.

Denn hohe Objektivität und hohe Reliabilität sind zwar essenzielle Bedingungen für psychometrische Verfahren, aber sie stellen keine ausreichenden Voraussetzungen für eine hohe Validität dar. Grundsätzlich liegt Validität vor, wenn der Test auch wirklich das zu beabsichtigende Merkmal misst und nicht willkürlich ein anderes. Die Validität beschäftigt sich mit der inhaltlichen Übereinstimmung mit dem, was der Test wirklich erfasst und mit dem Merkmal, was man mit dem Test erheben möchte. Ausserdem befasst sich die Validität auch mit den Schlussfolgerungen, die aus den Testergebnissen entstehen und den Interpretationen, die auf der Grundlage eines gewissen Aussenkriteriums ausgelegt werden

(Moosbrugger & Kelava, 2020). Auch bei der Validität gibt es weiter differenzierte Unterformen.

3.4.1 Inhaltsvalidität

Eine Form davon ist die Inhaltsvalidität, welche besteht, wenn die Testitems während der Operationalisierung so entwickelt werden, dass sie das zu untersuchende Merkmal angemessen repräsentieren. Die Inhaltsvalidität wird nicht durch numerische Werte bestimmt, sondern erfolgt durch vorhergehende theoretische Grundlagen und rationale Überlegungen. Die Inhaltsvalidität lässt sich am besten durch Expert:innen bewerten. So schlagen Moosbrugger und Kelava (2020) vor, dass die Items ein direktes Feedback zu den untersuchenden Merkmalen erlauben, sodass eine Aussage über die Repräsentativität gemacht werden kann. Inhaltsvalidität kann beispielsweise am Beispiel der Autofahrprüfung illustriert werden. Dort wird die Eignung der Autofahrenden anhand einer Prüfung von 45 Minuten inhaltsvalide getestet, da die Autofahrenden mehrere Manöver und Hindernisse bestreiten müssen und damit auch über eine längere Dauer geprüft werden, als wenn sie nur kurz wenige Meter fahren und dann bereits wieder aus dem Auto aussteigen.

3.4.2 Konstruktvalidität

Eine weitere Unterform ist die Konstruktvalidität, welche gegeben ist, wenn die Korrelation zwischen den Items im Test und dem zu untersuchenden Merkmal wissenschaftlich begründet ist (Moosbrugger & Kelava, 2020). Dies erfordert eine klare Definition von den theoretischen Konstrukten anhand des Forschungstandes. Dadurch wird nachvollziehbar, über welche beobachtbaren Merkmale das spezifische Konstrukt mit welchen Messmethoden erfasst wird (Döring & Bortz, 2016). Wenn aufgrund von dem gezeigten Verhalten der Testperson (z.B. durch gewählte Items) auf das zugrundeliegende psychologische Konstrukt geschlossen werden kann, weist ein Test Konstruktvalidität auf (Moosbrugger & Kelava, 2020). Von vielen Autor:innen wird der Begriff Konstruktvalidität als Synonym für alle Validitätsarten verwendet. Dies aus dem Grund, dass die Konstruktvalidität aussagend dafür ist, wie exakt ein Test erfasst, was er misst (Bühner, 2021). Die Konstruktvalidität lässt sich noch weiter differenzieren. Zu ihr gehören auch die diskriminante, konvergente und faktorielle Validität (Bühner, 2021). Auf deren Definition wird jedoch verzichtet, da es nicht notwendig ist, um die nachfolgenden Kapitel zu verstehen.

3.4.3 Kriteriumsvalidität

Eine hohe Kriteriumsvalidität ist gegeben, wenn von den Ergebnissen im Test auf ein Kriterium bzw. auf mehrere Kriterien ausserhalb der Testsituation geschlossen werden kann. Je enger der Testwert mit einem aussenstehenden Kriterium in Beziehung steht, desto höher das Ausmass der Kriteriumsvalidität (Moosbrugger & Kelava, 2020). In folgende Arten lässt sich die Kriteriumsvalidität weiter differenzieren: Vorhersagevalidität (prognostische oder prädiktive Validität), Übereinstimmungsvalidität, retrospektive Validität und inkrementelle Validität (Bühner, 2021).

3.4.4 Prognostische Validität

Die Vorhersagevalidität ist für die vorliegende Arbeit von hoher Relevanz und wird deshalb näher erläutert. Es handelt sich dabei um Korrelationen zwischen dem im Test erhobenen Merkmal und zeitlich versetzten Aussenkriterien. Als klassisches Beispiel wird oft von Zusammenhang der Leistung in einem Intelligenztest und der Abschlussnote in der Ausbildung berichtet (Bühner, 2021). In dieser Bachelorarbeit wird als prognostische Validität der Zusammenhang von psychometrischen Tests und dem zukünftigen Berufserfolg beleuchtet.

3.4.5 Augenscheinvalidität

Als letzte Unterform der Validität wird die Augenscheinvalidität definiert. Diese kennzeichnet sich dadurch aus, dass auch Laien einen Test durch «blossen Augenschein» als gültig anerkennen. Diese Form der Validität ist besonders wichtig im Hinblick auf die Testpersonen (Moosbrugger & Kelava, 2020). Hohe Augenscheinvalidität begünstigt die Akzeptanz von Laien, gleichzeitig besteht aber auch ein grösseres Risiko von Verfälschbarkeit (Klug, o. D.). Die Einschätzung von Laien (Testpersonen wie auch Personalverantwortliche) bezüglich Glaubwürdigkeit eines Testverfahrens hat einen erheblichen Einfluss auf die Gesamtauswertung. Somit können die Testergebnisse viel besser vermittelt und auch nachvollzogen werden (Moosbrugger & Kelava, 2020). Die Testpersonen entwickeln eher Vertrauen in den Test, da sie glauben zu erkennen, was das Verfahren misst (Schmidt-Atzert & Amelang, 2012). Es ist jedoch wichtig anzumerken, dass Augenscheinvalidität niemals mit Inhaltsvalidität gleichgesetzt werden darf. Oftmals wird von Laien angenommen, wenn ein Test die nach ihnen erkennbaren Merkmalen misst, müsse er automatisch auch inhaltsvalide sein. Die Verwechslungsgefahr ist demnach gross, die Bedeutung jedoch nicht dieselbe (Moosbrugger & Kelava, 2020).

3.5 Relevante Modelle und Theorien der Personalauswahl

3.5.1 Return on Investment (ROI)

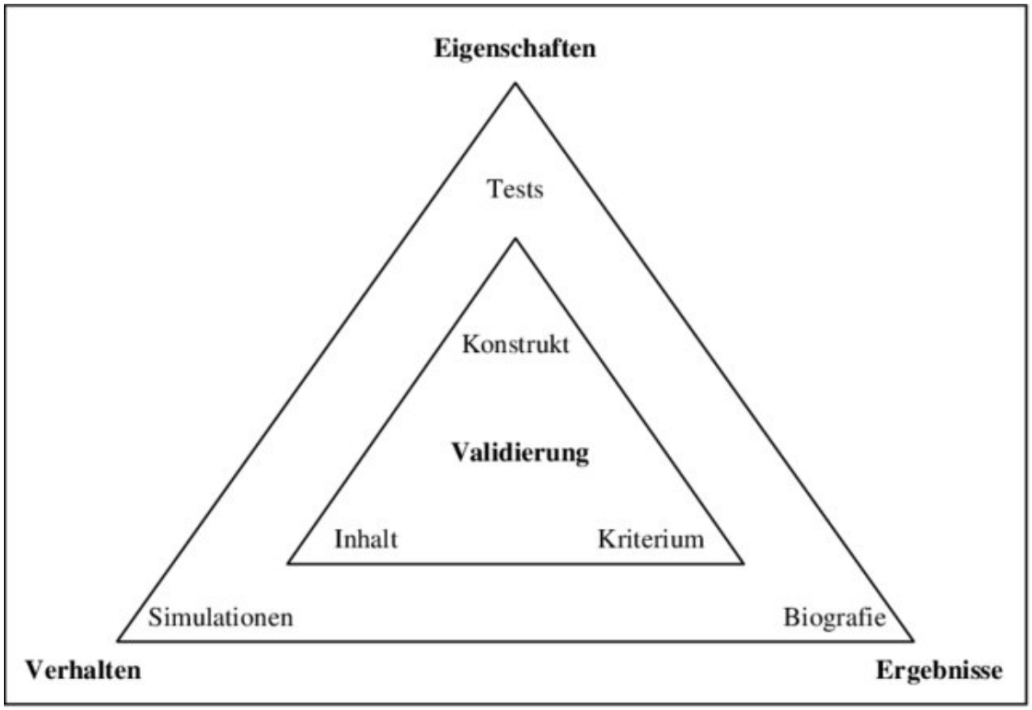
Die Bedeutung der Personalabteilung (Human Resources) hat in jedem Unternehmen zugenommen (Sangeetha, 2010). Es wird stets intensiver nach Personal mit fachlichen Qualifikationen gesucht und wie in der Einleitung erwähnt, herrscht Fachkräftemangel (Krüger, 2018). Es wird von einer Notwendigkeit neuer, innovativer Rekrutierungsstrategien berichtet. Dies aus gutem Grunde – die Personalabteilung hat einen starken Einfluss auf die Kapitalrendite (Return on Investment = ROI) und den gesamten Erfolg des Unternehmens im Markt. Dies betont, wie wichtig es ist, effektive Rekrutierungsstrategien zu entwickeln und auch langfristig engagierte Mitarbeitende im Unternehmen zu entwickeln und zu binden (Sangeetha, 2010). Denn die Kompetenzen und Talente der einzelnen Mitarbeitenden schaffen somit Wettbewerbsvorteile gegenüber anderer Marktkonkurrenz und führen zu nachhaltigem Geschäftserfolg (Sangeetha, 2010).

3.5.2 Trimodaler Ansatz der Eignungsdiagnostik

Nach Höft und Schuler (2014) sind die Persönlichkeitstests und Intelligenztests auch im trimodalen Ansatz der Eignungsdiagnostik zu finden. Der trimodale Ansatz beabsichtigt verschiedene Verfahrensarten bei der Personalauswahl anzuwenden. Er lässt sich in drei elementare Verfahrenstypen einteilen, welche auf verschiedenen Vorgehensweisen basieren. Es bestehen folglich drei Auswahlmöglichkeiten, was den Prozess der Personalauswahl betrifft – deshalb der Name trimodal. Diese drei verschiedenen Konzepte lassen sich nicht nur einzeln anwenden, sondern auch miteinander systematisch kombinieren. Dies macht den trimodalen Ansatz auch zu einem multimodalen Ansatz (Höft & Schuler, 2014). Dieses Vorgehen soll die Treffquote, dass man den richtigen Bewerbenden ausgewählt hat, erhöhen und den Fehler einen unpassenden Bewerbenden zu rekrutieren minimieren oder gar eliminieren (Höft & Schuler, 2014). Die Abbildung 3 verdeutlicht den Aufbau des trimodalen Ansatzes, welche in drei Betrachtungsebenen unterteilt wird (Schuler, 2016). Dadurch kann auf drei unterschiedlichen Ebenen interveniert oder etwas erfasst werden: auf der Ebene Aufgaben/Ergebnisse, der Ebene des Verhaltes und der Eigenschaftsebene. Der trimodale Ansatz ist für die Personalauswahl von grosser Relevanz, da mehrfach bestätigt wurde, dass ein multimodales (multimethodales) Vorgehen gewinnbringend ist und die Validität des Personalauswahlprozesses dadurch erhöht werden kann (Schuler, 2016). Nebst der Eignungsdiagnostik und der Personalauswahl wird er auch

in anderen Gebieten der Personalpsychologie eingesetzt, z.B. in der Leistungsbeurteilung und Personalentwicklung (Schuler, 2016).

Abbildung 3
Darstellung des Trimodalen Ansatzes nach Schuler. Adaptiert von Der trimodale Ansatz der Eignungsdiagnostik (S. 103) von Schuler und Höft (2006), Hogrefe.



3.5.3 Person-Environment-Fit

Das Person-Environment Fit-Modell (abgekürzt: P-E-Fit) basiert auf der interaktionalen Psychologie und besagt, dass das Zusammenspiel zwischen Persönlichkeitsmerkmalen und Merkmalen der Umwelt der Hauptantreiber des menschlichen Verhaltens ist (De Cooman & Vleugels, 2022). Nach Nerdinger et al. (2019) betrachtet das P-E-Fit-Modell die Passung zwischen den Ressourcen und Fähigkeiten, die eine Person mitbringt und den Anforderungen, die eine Aufgabe oder Arbeitsstelle voraussetzt. Falls dieses Gleichgewicht nicht gegeben ist, kann Stress entstehen. Unter Ressourcen und Fähigkeiten sind Qualifikationen, Wissen und Persönlichkeitsmerkmale und Verhaltensarten gemeint. So ist es wichtig, dass diese zur Organisation passen, umgekehrt soll auch die Arbeitstätigkeit und die Organisation zur Person passen. Falls diese Passung nicht vorhanden ist, kann eine Diskrepanz entstehen, welche negative Auswirkungen auf das Individuum und die Organisation haben kann (Nerdinger et al., 2019). So konnte in Studien bestätigt werden, dass Führungspersonen deren Führungsstil den Anforderungen der Organisation entsprach, weniger gesundheitliche Beschwerden und

Ausfälle erlebten, als solche, bei denen die Passung nicht so positiv ausfiel. Ein hoher P-E-Fit kann nicht nur für das Individuum positive Folgen haben (Nerdinger et al., 2019) sondern auch Unternehmen profitieren von einer hohen Passung. Denn diese korreliert mit Mitarbeiterzufriedenheit und Mitarbeiterbindung, welche sich auch in einer höheren Arbeitsleistung und geringeren Fluktuation widerspiegelt (De Cooman & Vleugels, 2022). Daraus kann interpretiert werden, dass der vorangehenden Personalrekrutierung und -auswahl einen hohen Stellenwert beigemessen werden sollte. Durch eine sorgfältige Planung, gut überdachte Entscheidung mithilfe validierter Verfahren kann das Unternehmen davon profitieren und wiederholende, unnötige Kosten für die Rekrutierung umgehen. Damit wird auch der bereits erwähnte ROI angesprochen. Auch dieser profitiert von zufriedenen Mitarbeitenden und hoher Arbeitsleistung, welche beide mit einem guten P-E-Fit gefördert werden.

3.6 Digitalisierung in der Personalrekrutierung

Die Digitalisierung spielt eine stets relevantere Rolle – auch in der Personalauswahl (Schellinger et al., 2020). Da Unternehmen organische Elemente sind, unterliegen sie ständig Veränderungen in ihren Strukturen, Managementstilen bis hin zu den Technologien. Im Kontext der Globalisierung und der rapiden Entwicklung von Informationstechnologien ist das Personalmanagement und dessen Effizienz, was Rekrutierungen betrifft, besonders gefragt (Molotkova et al., 2019). Das Personalmanagement ist das dynamischste Subsystem eines Unternehmens (Molotkova et al., 2019) und eine Neuausrichtung der Methoden und Technologien des Personalmanagements beeinflusst die Gesamteffizienz eines Unternehmens. Der eher geringe Einsatz von Online-Tests könnte sich auch stückweise dadurch erklären, dass das Schweizer Personalmanagement aktuell mehr denn je gefordert ist und mit den wandelnden und herausfordernden Einflüssen, wie beispielsweise der vergangenen Corona-Pandemie und den aktuellen Kriegen mit grösseren Unsicherheiten in der Wirtschaft rechnen muss (Universität St. Gallen, 2023). Laut dem Trend Barometer People Management 2030 der Universität St. Gallen (2023) bestehen Trends wie New Work und digitale Transformation, denen man versucht gerecht zu werden. Gleichzeitig wird auch festgestellt, dass Fachkräftemangel herrscht und HR-Expert:innen berichten, dass sie nicht hinreichend auf die Ansprüche der nahen Zukunft vorbereitet sind. Dabei sind sich Expertinnen einig, dass künftig vor allem Gebiete wie die Rekrutierung noch besser ausgebaut werden müssen (Universität St. Gallen, 2023). Da laut der Universität St. Gallen (2023) das Wissen besteht, dass digitale Kompetenzen stets an Wichtigkeit gewinnen, könnte man daraus interpretieren, dass mit dem fortlaufenden Zuwachs an Digitalisierung zunehmend auch Online-Tests im Bereich der Personalauswahl

zum Einsatz kommen. Nach Steiner (2009) bietet die Digitalisierung im Hinblick auf Testverfahren auch viele Vorteile – es können mit Online-Tests viele Ressourcen eingespart werden. Das bedeutet eine effiziente Auswertung, ortsunabhängige Zugriffsmöglichkeit und das Personal muss nicht vor Ort die Durchführung übernehmen (Steiner, 2009). Gleichzeitig bietet das Internet eine Vielzahl an Online-Testverfahren und dazu unzählige Anbieter. Dies stellt mit der voranschreitenden Digitalisierung eine Herausforderung für Personalverantwortliche dar, da die Durchschaubarkeit schwer einzuschätzen ist und oft auch unseriöse Angebote im Internet kursieren (Steiner, 2009).

Es liegt auf der Hand, dass die Technologie einen grossen Einfluss auf den gesamten Auswahlzyklus hat. Die Forschung ist dabei nicht im gleichen Tempo wie die Digitalisierung. Unternehmen wird geraten mehr in Technologie und Digitalisierung zu investieren, damit sie qualifizierte Talente für sich gewinnen können. Insbesondere den Sozialen Medien und digitalen Vorstellungsgesprächen kommen eine wichtige Rolle zu, wenn sich Unternehmen fortschrittlich präsentieren wollen. Digitale Vorstellungsgespräche stellen eine Möglichkeit dar, die Ressourcen der Personalbeschaffung gezielter einzusetzen. Die Rekrutierung von Talenten wird immer technikaffiner und die traditionelle Rekrutierung wird wohl bald eher eine Ausnahme darstellen (Adeosun & Ohiani, 2020). Nikolaou (2021) berichtet von zwei neuen Trends: asynchrone Befragungen und Gamification. Bei asynchronen Interviews geben Kandidat:innen Antworten auf vorab aufgezeichnete Fragen. Gamification nutzt existierende Tools, wie Persönlichkeitstests oder Tests zur Situationsbeurteilung, in einer spielerischen Form. Zudem fordert die Digitalisierung in der Personalrekrutierung die Offenlegung des Datenschutzes und den Umgang mit personenbezogenen Daten (Adeosun & Ohiani, 2020).

3.6 Relevanz von psychometrischen Online-Tests

Nach Afshari et al. (2014) ist das Identifizieren, Abwägen und Evaluieren eines Kandidaten hinsichtlich den Berufsanforderungen eine wesentliche Aufgabe der Personalauswahl und während des Rekrutierungsprozesses. Denn die Fähigkeiten des Personals, wie z.B. Wissen, Fertigkeiten und Erfahrung spielen eine wichtige Rolle für den Erfolg von Organisationen (Afshari et al., 2014). Der Hauptfokus eines guten Personalwesens (HR) liegt in zwei Bereichen: einerseits müssen sie die bisher erworbenen Fähigkeiten beurteilen und andererseits müssen sie die Kandidat:innen auch bezüglich ihren Potenziale zukünftiger Entwicklungen einschätzen. Aufgrund dessen treffen HR-Fachpersonen die Entscheidung, ob eine bestimmte Person für die ausgeschriebene Berufsposition am besten geeignet ist (Jansen et al., 2023).

Die bestmögliche Passung ist deshalb so wichtig, da man davon ausgeht, dass sie mit zukünftigem Berufserfolg (Job Performance) einhergeht. Steiner berichtet (2009), dass psychologische Tests die Qualität der Auswahlentscheidungen positiv beeinflussen. In der Vergangenheit wurden solche Tests in Unternehmen gelegentlich als Papier-Bleistift-Tests (Paper-Pencil-Tests) verwendet. Die Entwicklung der Online-Tests erhöhte jedoch den Einsatz von Testverfahren in der Personalauswahl. Kanning (2019) behauptet jedoch zehn Jahre später, dass der Einsatz in Unternehmen im Hinblick auf das Erfolgsversprechen und die aussagekräftige Validität gering sei. Zu Unrecht – denn wie aus vielen Studien hervorgeht, gehören die kognitive Intelligenz und die emotionale Stabilität (Persönlichkeitseigenschaft; Gegenpol des Neurotizismus) zu den aussagekräftigsten Prädiktoren für den Berufserfolg (Dhliwayo & Coetzee, 2020).

Eine der bekanntesten und relevantesten Studien, auf der heute noch viele Forschungen aufbauen, ist die Metaanalyse von Schmidt und Hunter (1998), welche sich mit den verschiedenen Verfahren in der Eignungsdiagnostik und den dazugehörigen Gütekriterien auseinandersetzte. In der erstmals 1998 publizierten Studie von Hunter und Schmidt ist ersichtlich, dass kognitive Intelligenz mit Abstand der beste Prädiktor für Berufserfolg darlegt. Dennoch wird 24 Jahre später in Deutschland berichtet, dass der Intelligenztest das seltenste eingesetzte Verfahren in der Praxis darstellt (Hogrefe, 2022). Die Implementierung von Intelligenztests in den Rekrutierungsprozess könnte für viele Unternehmen einen Mehrwert bieten, denn wie Kanning (2019) behauptet, können Führungskräfte auf viele essentielle Aufgaben geprüft werden, die sie als Manager:innen im Berufsalltag bewältigen müssen. Viele Fähigkeiten, die im Intelligenztest erhoben werden, können anschliessend als Massstab dafür dienen, ob Führungskräfte ein logisches Denkvermögen besitzen und wie sie unter Zeitdruck rationale Entscheidungen treffen. Dies sind alles Fähigkeiten, die wichtig im Berufsalltag sind und zum Unternehmenserfolg beitragen, weshalb es einleuchtet, dass die prognostische Validität für kognitive Leistungstests besonders hoch ist.

Auch den Persönlichkeitstests wird eine Bedeutung im Zusammenhang mit Arbeitsleistung und Berufserfolg zugeschrieben (Neyer & Asendorpf, 2024). So berichten Neyer und Asendorpf (2024) von mehreren Studien bei denen die «Big Five» im Zusammenhang mit Eigenschaften einer Führungsperson wichtige Erkenntnisse brachten. Dabei stellte sich heraus, dass besonders die Persönlichkeitsmerkmale Gewissenhaftigkeit, Extraversion und Offenheit mit einer qualitativ hohen Führung korrelieren. Diese drei Dimensionen lassen sich mittels psychometrischen Persönlichkeitsfragebögen gut ermitteln und eignen sich als Prädiktoren für Berufserfolg (Neyer & Asendorpf, 2024). Sackett et al.

(2022) definieren Gewissenhaftigkeit nach wie vor als den aussagekräftigsten Prädiktor aller Persönlichkeitsdimensionen im Hinblick auf die Arbeitsleistung.

3.7 Verbreitung von psychometrischen Online-Tests in der Personalauswahl

Ständig entstehen in der Forschung neue psychologische Verfahren für die Verwendung in der Personalauswahl und es werden kontinuierlich Befunde zu deren Gütekriterien publiziert (Hogrefe, 2020). Doch welche Auswirkungen hat der wissenschaftliche Fortschritt auf die praktische Anwendung der Tests? Und welche Kriterien bestimmen, ob ein gewisses Verfahren letztendlich auch angewendet wird (Hogrefe, 2020)? Es existieren zahlreiche psychometrische Testverfahren, die sich für die Verwendung eignen. Solche Testverfahren können sicherstellen, dass individuelle Aspekte (wie z.B. Intelligenz oder Persönlichkeit) objektiv erfasst werden (Hogrefe, 2020). Dabei haben Forschende wie auch Anbieter der Tests hohe Ansprüche an die Verfahren. Die psychometrischen Tests unterlaufen bereits vor der Veröffentlichung verschiedene Prüfungen, um nachzuweisen, dass die wissenschaftlichen Gütekriterien eingehalten worden sind. Wie bereits in Kapitel 3.4 erwähnt, spielt die Validität dabei eine wichtige Rolle (Hogrefe, 2020). Laut Funk et al. (2015) fällt hierbei auf, dass weniger valide Verfahren überschätzt werden, während sehr valide Verfahren unterschätzt werden. Obwohl Intelligenztests zu ungefähr 45 Prozent die berufliche Leistung von Führungskräften hervorsagen können, ist es erstaunlich, dass diese immer noch sehr selten in Unternehmen eingesetzt werden (Kanning, 2019). Kanning (2019) geht davon aus, dass im deutschsprachigen Raum noch sehr viele Vorurteile gegenüber dem Einsatz von kognitiven Leistungstests bestehen und dies die sparsame Anwendung erklären könnte. Die USA hingegen prüft bei der Besetzung einer Managementposition die Bewerbenden zu etwa 50 Prozent auf die Intelligenz (Kanning, 2019).

Die grosse Anzahl an Testverfahren und unterschiedlichen Anbieter auf dem Markt machen es Personalverantwortlichen nicht einfach eine Auswahl zu treffen. Auf welcher Grundlage sollen sie sich entscheiden? Es existieren hunderte von Verfahren und alle werben damit, die besten zu sein. So ist es nicht erstaunlich, dass schnell der Überblick verloren geht und Unwissen besteht, welche Tests eingesetzt werden sollten, denn ein gutes Marketing ist nicht mit einem qualitativ hochwertigen Test gleichzusetzen (Kanning, 2022). Es fällt auf, dass Beratungsfirmen, welche mögliche Testverfahren an die Kundschaft verkaufen, die relevanten Kennwerte zu den Gütekriterien oftmals nicht preisgeben, was die Einschätzung darüber hinaus erschwert. Somit kann dies als ein Indiz betrachtet werden, an welchem wissenschaftlich geprüfte Testverfahren erkennbar sind. Diese veröffentlichen meist alle wichtigen Werte zu den Gütekriterien und halten wichtige

Informationen nicht zurück. Es ist auffällig, dass sich viele Unternehmen nicht an den vorgestellten Gütekriterien orientieren, sondern anhand der Beliebt- und Bekanntheit eines Testes. Dies erklärt, dass längst überholte Tests immer noch eingesetzt werden. Weiter besteht zudem der Irrtum, dass ein Test eine hohe Qualität aufweist, wenn die Ergebnisse Sinn machen. Beispielsweise füllt eine Personalverantwortliche selbst einen Fragebogen aus und wenn die Resultate ihrem Selbstbild entsprechen, hält sie diesen Test für qualitativ gut. Dies ist jedoch ein fatales Denken, denn wie bereits erwähnt, können Selbstbilder verzerrt sein und Menschen haben die Tendenz das zu glauben, was sie wollen (Bestätigungsfehler) (Kanning, 2022). Vielen Persönlichkeitstests liegt auch der Barnum-Effekt (vlg. Forrer-Effekt) zugrunde: die Auswertung des Testes gibt so allgemeingültige Aussagen und Feedbacks von sich, die praktisch auf jeden Menschen zutreffen können. Deshalb ordnen die meisten Menschen, diese dann auch als valide und aussagekräftig ein (Hartmann, 2023). So haben sich nebst überholten Tests auch Verfahren etabliert, dessen Gütekriterien gar nie wissenschaftlich belegt worden sind (Kanning, 2022).

3.8 Herausforderungen und Kritikpunkte bei der Verwendung von Online-Tests

Die Testverfahren haben von Natur aus Einschränkungen. In Bezug auf die Menge der gesammelten Informationen ist das Interview dem Testverfahren überlegen. Verhaltensübungen können das gegenwärtige Sprachvermögen besser beurteilen, wichtige Aspekte der sozialen Kompetenz sowie die Sympathie eines Bewerbers werden auch deutlich besser erhoben. Oftmals lassen sich die Fachkenntnisse aus dem Lebenslauf besser abschätzen als durch «Prüfungen». Der für die jeweilige Fragestellung passende Methodenmix („Multimethodalität“) ist der optimale Ansatz in der Personalarbeit, anstatt starr an einer Methode festzuhalten (Sarges et al., 2009). An dieser Stelle ist anzumerken, dass die Strukturierung des Interviewprozesses eine massgebende Auswirkung auf die Validität hat. Dasselbe gilt auch für die einheitliche Anwendung (Standardisierung) der verschiedenen Interviewarten. Somit kann ein Interview durchaus in der Form einer Videokonferenz durchgeführt werden, dies muss jedoch mit allen Bewerbenden eingehalten werden. Dasselbe gilt bei Face-to-Face (Kanning & Ohlms, 2024). Dementsprechend kann ein gut strukturiertes Interview viele Vorteile mit sich bringen und eine hohe Validität aufweisen. Laut Funk et al. (2015) werden aber in der Praxis häufiger die unstrukturierten Interviews angewendet, welche eine tiefe Validität besitzen. Dies kann die Auswahl und nachfolgende Stellenbesetzung im Unternehmen langfristig negativ beeinträchtigen. Der Persönlichkeitstest wird laut Funk et al. (2015) am wenigsten häufig eingesetzt, gefolgt von dem Intelligenztest. Mögliche Gründe erklären sie sich dadurch, dass viele Personalverantwortliche gar nicht über die aktuelle Aussagekraft von psychometrischen

Tests informiert sind, weil sie keine wissenschaftlichen Zeitschriften lesen und sich somit zu wenig mit derzeitigen Erkenntnissen auseinandersetzen (Funk et al., 2015). Nebst der unzureichenden Informiertheit über die Validität der psychologischen Verfahren, zeigt sich bei vielen HR-Fachpersonen auch ein mangelndes Vertrauen in die Tests zur Erfassung der Persönlichkeit oder der Intelligenz. Zudem wird auch oft die Relevanz oder die Praktikabilität in Frage gestellt (Funk et al., 2015). Auch im Jahre 2022 werden die sehr validen Messverfahren stets unterschätzt, wie Hogrefe (2022) berichtet. Ein weiterer Grund für die seltene Anwendung könnte auch daran liegen, dass Personalverantwortliche es auch schlichtweg nicht für möglich halten, dass derart komplexe Konstrukte wie Persönlichkeit oder Intelligenz in einem einzelnen Test erfasst werden können (Hogrefe, 2022). Weiter besteht die Vermutung, dass besonders kognitive Tests von potenziellen Kandidat:innen nicht viel Akzeptanz erfahren und auch vermehrte Kosten durch deren Einsatz entstehen können (Hogrefe, 2022).

Nach Moosbrugger und Kelava (2020) sind Kritikpunkte bei psychometrischen Tests wie Intelligenz- und Persönlichkeitstests häufig auch die Anfälligkeit auf Verfälschungen. Besonders Persönlichkeitstests sind sehr empfänglich für ergebnisverfälschendes Verhalten, sogenanntes Faking. Das Risiko ist hoch, wenn der Test eine hohe Augenscheinvalidität besitzt und die Testpersonen somit gut abschätzen können, welches Persönlichkeitsmerkmal gerade gemessen wird und welche Antwort somit sozial erwünscht wäre. Daher sollten die Skalen eines Persönlichkeitsfragebogens faking-resistent gestaltet sein (Moosbrugger & Kelava, 2020). Bei Online-Tests kann ausserdem der Nachteil entstehen, nicht genau zu wissen, wer den Test wirklich ausfüllt. Diese Art von Verfälschung ist abzugrenzen von Faking und häufiger bei kognitiven Leistungstests anzutreffen (Steiner, 2009). Hierbei geht es darum, dass Personen trotz des Verbots in der Testinstruktion, den Taschenrechner trotzdem verwenden. Es können auch andere Personen beigezogen werden, um so das Testergebnis durch die Teamkompetenz zu verbessern. Diese Risiken werden durch die fehlende Kontrolle, die Ortsunabhängigkeit und die Anonymität begünstigt (Steiner, 2009). Müllerschön (2006) hält die Papier-Bleistift-Tests für überholt und befürwortet Online-Tests. Diese computergestützten Verfahren sind nicht nur moderner, sondern liefern eine schnellere Auswertung und weniger Fehler, als die altmodischere Schablonenauswertung (Müllerschön, 2006).

Oft existieren auch Vorurteile, dass Bewerber:innen auf Führungspositionen schon eine hohe Grundintelligenz mitbringen und es deshalb unnötig ist, diese einen Intelligenztest durchlaufen zu lassen. Was nicht zutrifft, denn Studien zeigen, dass mit aufsteigender Hierarchieebene die Durchschnittsintelligenz zunimmt, aber die Unterschiede

zwischen den Führungskräften auf einer Ebene fast genauso gross sind wie in der Gesamtbevölkerung (Kanning, 2019).

Ausserdem berichten Sackett et al. (2022) über neuere Validitätsschätzungen, bei denen diverse Reduktionen der Validität vorgenommen werden mussten, der bisher angenommenen Prognose für Arbeitsleistung und Berufserfolg. Sackett et al. (2022) untersuchten die bisherigen meta-analytischen Schätzungen der Kriteriumsvalidität von Personalauswahlverfahren und stützten sich hierbei auf die bekannte Studie von Schmidt und Hunter (1998). Dabei wurde kritisiert, dass gewisse Überschätzungen stattfanden, da bei der Bewertung nur ein eingeschränkter Wertebereich betrachtet wird. Somit zeigen die neu kalibrierten Schätzungen der Validität, dass viele Auswahlverfahren weiterhin nützlich sind, aber die prädiktiven Beziehungen schwächer sind, als wie bisher angenommen. Neuere Schätzungen ergaben somit, dass unterdessen bei allen Verfahren, eine Überkorrektur stattfand, was nun eine Reduktion der Validität zur Folge hat. Laut Sackett et al. (2022) werden strukturierten Interviews die höchsten Werte der prognostischen Validität zugesprochen. Gleichzeitig wurde bei den Persönlichkeitstests und kognitiven Leistungstests eine Abnahme hinsichtlich der prognostischen Validität beobachtet und der Intelligenztest wird nicht mehr als der beste Prädiktor angesehen (Sackett et al., 2022).

So berichtet auch Salgado (2017), dass aufgrund der unterschiedlichen Validität der verschiedenen Auswahlverfahren in der Personalrekrutierung die grösste Korrelation mit Berufserfolg durch eine Kombination von multiplen Verfahren entsteht. Für die Gesamtarbeitsleistung ist die kognitive Intelligenz der beste Prädiktor. Dies wird wie soeben kurz beschrieben, durch die Methode der Selbsteinschätzung, welche in Persönlichkeitstests genutzt wird, erklärt. Leistungstests werden in der Regel nicht absichtlich verfälscht, bei Fragebögen ist die Möglichkeit schon grösser, ein sozial erwünschtes Verhalten zu zeigen und sich als disziplinierter zu beschreiben, als man in Wirklichkeit ist (Kanning, 2022). Zudem werden die Angaben der befragten Person aufgrund ihres Selbstbildes geäussert, welches auch nicht der Realität entsprechen muss und verzerrt sein kann. Sofern eine Person nicht eine gute Selbstreflexion besitzt, läuft sie in Persönlichkeitstests Gefahr, sich zu über- oder unterschätzen. Dies erklärt, weshalb dem Intelligenztest eine höhere Aussagekraft zugeschrieben wird und dementsprechend auch die prognostische Validität bezüglich Berufserfolgs und Arbeitsleistung höher ausfällt (Kanning, 2022). Wenn es jedoch ein Persönlichkeitsmerkmal gibt, dass die Arbeitsleistung am zuverlässigsten hervorsagt, dann ist es die Gewissenhaftigkeit. Trotz neuerer Schätzungen bleibt Gewissenhaftigkeit der beste Prädiktor für Arbeitsleistung unter den Persönlichkeitsdimensionen (Sackett et al., 2022). Ein weiterer wichtiger Faktor ist die Qualität der Testverfahren. Im Alltag achten viele auf Nährwerte, Herkunft und

Qualitätsrichtlinien von Produkten, nicht nur auf den Preis. Da es für Nicht-Expert:innen schwierig ist, die Qualität und Kriterien von Tests zu beurteilen, wird für die Entscheidungsfindung häufig auf Testberichte zurückgegriffen. Es ist zu erwähnen, dass nicht nur beim Einkauf von Produkten Qualitätsunterschiede bestehen, sondern auch bei Testverfahren. Wie wird dort die Qualität gemessen und beurteilt (Schmidt-Atzert & Amelang, 2012)? Dabei wird noch einmal auf die Gütekriterien in Kapitel 3.1 verwiesen. Im Ergebnisteil dieser Bachelorarbeit wird eine selbst erarbeitete Tabelle präsentiert, die auf die oben diskutierten Qualitätskriterien eingeht und eine Beurteilung vornimmt.

4 Methodik

Die Fragestellung der Bachelorthesis wird durch den theoretischen Teil, eine Problemanalyse und eine Umfrage beantwortet. Im folgenden Abschnitt wird das methodische Vorgehen erläutert und die notwendigen Schritte zur Beantwortung der Fragestellung vorgestellt.

4.1 Forschungsdesign

Im ersten Teil des methodischen Vorgehens dieser Arbeit wurde eine Desk Research durchgeführt. Eine Desk Research ist eine Schreibtischrecherche und gehört der Sekundärforschung an. Sekundärforschung bedeutet, auf bereits gewonnene Erkenntnisse zurückzugreifen, die zuvor schon einmal zusammengefasst worden sind (Thommen et al. (2017). Dies wurde für die Persönlichkeitstests und kognitiven Leistungstests so durchgeführt. Diese Herangehensweise dient der Informationsbeschaffung von wesentlichen Daten, die im Hinblick auf die Fragestellung wertvolle Fakten liefern. Sie unterscheidet sich von der Primärforschung, da keine neuen Daten erhoben werden. Stattdessen werden vorhandene Informationen systematisch recherchiert, analysiert und ausgewertet (marktforschung.de, o. D.). Die Desk Research hilft festzustellen, ob eine empirische Datenerhebung notwendig ist und zeigt, welche Fragen offenbleiben und welche beantwortet werden können. Spezifischere Forschungsfragen, welche nicht durch die alleinige Verwendung von Primärstudien beantwortet werden, können im Anschluss in einer darauf zugeschnittenen Erhebung erfolgen. Sie liefert einen breiten Überblick über die Datenbasis, unterstützt die Planung und den Aufbau des Projektes (marktforschung.de, o.D.). Ausgehend von der Desk Research sowie in Absprache mit dem Praxispartner wird auf eine Problemanalyse abgestützt, die mittels deskriptiver Statistik dargelegt wird. Im Gegensatz zur analytischen Statistik steht hier die Beschreibung der Daten durch Häufigkeitstabellen, Kennwerte oder Grafiken im Fokus. Die deskriptive Statistik zeichnet sich durch die Messung von Variablen sowie die Zuordnung von Zahlen und Variablen aus. Es ist von hoher Wichtigkeit das richtige statistische Messverfahren und Skalenniveau für die zu erfassenden Variablen auszuwählen. Die unterschiedlichen Variablen können dabei nach vier verschiedenen Skalenniveaus eingeteilt werden (Heimsch, 2018).

Für die Erstellung des Fragebogens eignen sich besonders die Nominalskala und die Ordinalskala, welche im Kapitel 4.3 Datenerhebung vorgestellt werden. Zudem wurden Multiple-Choice, Single-Choice, offene Fragen und die Likertskala verwendet.

Für die Befragung der Unternehmen bot sich ein Fragebogen an, welcher quantitative wie auch qualitative Aspekte enthält. Die Datenerhebung erfolgte über das Sampling, welches in Kapitel 4.2 näher beleuchtet wird. Der Fragebogen wurde durch eine Umfrage mithilfe des Programms Tivian erstellt. Die Fragebogenitems wurden zum einen durch die Desk Research und zum anderen in Absprache mit dem Praxispartner erstellt. Der Fragebogen wurde nach Erstellung für einen Pretest an die Personalabteilung des Geschäfts der Autorin sowie an diverse Teamkolleg:innen zum Verständnis zugestellt und aufgrund deren Rückmeldungen erneut angepasst. In Absprache mit der Betreuungsperson und dem Praxispartner wurde der Fragebogen vor der Versendung ein zweites Mal überarbeitet und anschliessend noch ein technischer Pretest durchgeführt, bevor er final versendet wurde. Vor der Versendung des Fragebogens an die Unternehmen wurde jedes telefonisch kontaktiert und über die bevorstehende Umfrage informiert. Dabei wurde darauf geachtet, direkt mit der/dem Personalverantwortlichen zu sprechen, ihre Emailadresse zu erfassen und ansonsten erneut anzurufen. Nachdem alle Unternehmen telefonisch informiert wurden, konnte der Fragebogen versendet werden. Es wurde darauf geachtet nicht mehr als 50 E-Mails täglich zu versenden, um die Wahrscheinlichkeit zu reduzieren im Spamordner zu landen. Anschliessend wurden die Ergebnisse von Tivian in ein Excel exportiert, die Daten bereinigt und die Variablen benannt. Das Excel wird zur Datenauswertung ins Programm Jamovi importiert. Zur Veranschaulichung des Forschungsdesign und Ablauf des Forschungsprozesses wurde Abbildung 4 erstellt:

Abbildung 4
Beschreibung und Abfolge des Forschungsprozesses (eigene Darstellung)

4.2 Sampling

Das Sampling konzentriert sich auf die vier Schweizer Kantone in der Nordwestschweiz Aargau, Basel-Land, Basel-Stadt und Solothurn und den darin liegenden Unternehmen. Es wurde darauf geachtet, dass die Unternehmen eine grössere Mitarbeitendenanzahl verfügen (ab 100 Mitarbeitende). Um herauszufinden wie weit verbreitet psychometrische Online-Tests in der Schweiz sind, wurden 219 Unternehmen kontaktiert. Die Unternehmen wurden mittels einer im Internet erworbenen und nach Kriterien gefilterten Liste ausgewählt. Dabei wurde darauf geachtet nur Unternehmen einzubinden, welche über mehr als 100 Mitarbeitende (MA) verfügen. Dies wurde mit dem Praxispartner so vereinbart, um ein grosses Sampling zu erzielen und nicht zu kleine Unternehmen (<50 MA) in die Stichprobe einzubeziehen. Ziel war es die grössten Unternehmen (nach MA-Zahl) zu befragen. Da bei >250 MA nur ca. 100 Unternehmen herausgefiltert wurden, wurde die Grenze auf 100 Mitarbeitende gesenkt, da davon ausgegangen werden muss, dass die Rücklaufquote nicht sehr hoch ausfallen könnte, um eine Stichprobe von mindestens 50 zu gewährleisten. Aufgrund der mangelnden Aktualität der Liste, enthielt sie jedoch auch Unternehmen mit weniger als 100 oder 50 MA, was erst bei der Auswertung auffiel. Die nicht aktuelle Liste stellte eine zusätzliche Herausforderung dar, sodass die aktuellen HR-Verantwortlichen und ihre Kontaktdaten telefonisch erfragt und manuell erfasst werden mussten.
Der Praxispartner stellte 26 weitere Firmenadressen zur Verfügung. Die erworbene Liste enthielt doppelte, fusionierte oder geschlossene Firmen. Um auf Reminder zu verzichten, rekrutierte die Autorin weitere Firmen. Da die Umfrage anonym war, konnte nicht nachverfolgt werden, wer sie bereits ausgefüllt hatte, und Reminder hätten an alle Unternehmen gesendet werden müssen. Die Anonymisierung ermöglichte ehrliche Antworten, erschwerte jedoch separate Reminder. Nach Erreichen von 53 Antworten wurde darauf verzichtet.

4.3 Datenerhebung in Form eines Fragebogens (Umfrage)

Wie im Forschungsdesign beschrieben, eigneten sich für die Datenerhebung mittels Fragebogens insbesondere die Nominal- und Ordinalskala. Die beiden verwendeten Skalen sowie ihre Besonderheiten werden nachfolgend vorgestellt:

Nominalskala: Eine Nominalskala ordnet Variablenwerte ohne quantitative Ordnung zu. Ein Beispiel ist das Geschlecht, das zufällig mit 1 und 2 versehen wird. Die Zuordnung erfolgt willkürlich, ohne empirische Bedeutung. Antwortmöglichkeiten können in beliebiger Reihenfolge dargestellt werden, da die Kodierung sich in keine Rangordnung bringen lässt und die Auswertung ist auf Häufigkeitszählung beschränkt. Dichotome Variablen wie «1 ja/

2 nein» ermöglichen eine Ordnungsregulation (Heimsch, 2018). Im Fragebogen wurden folgende Items mithilfe einer Nominalskala erfasst:

- Frage 1: Testanwendung und unterschiedliche Positionen mit jeweils vier unterschiedlichen Antwortoptionen
- Frage 2: Auswahl der Persönlichkeitstests mit insgesamt 14 unterschiedlichen Antwortoptionen
- Frage 3: Auswahl der IQ-Tests mit insgesamt 13 unterschiedlichen Antwortoptionen
- Frage 5: Kosten der Tests mit jeweils vier unterschiedlichen Antwortoptionen
- Frage 6: Anmerkungen und Feedbacks zu Tests mit offenem Textfeld
- Frage 7: Branche mit offenem Textfeld
- Frage 9: Angabe E-Mail mit offenem Textfeld

  Anmerkung: Die Items und konkreten Antwortmöglichkeiten können im Anhang unter Fragebogen eingesehen.

Ordinalskala: Variablen können aufsteigend nach Wertigkeit geordnet werden. Ordinalskalierte Variablen haben eine empirische bedeutende Kodezahl, z.B. Schulstufen oder Einkommen. Die Kodierung bezieht sich nicht auf die Differenz zwischen den Zahlen (Heimsch, 2018). Im Fragebogen wurden folgende Items ordinalskaliert erfasst:

Seite 4: Erleben und Empfinden der Tests

- Frage 4a: Übereinstimmung Persönlichkeitstests und Berufserfolg mit fünf Abstufungen in Form einer Likert-Skala
- Frage 4b: Übereinstimmung kognitive Tests und Berufserfolg mit fünf Abstufungen in Form einer Likert-Skala

Seite 5: Demographische Angaben

- Frage 8: Anzahl Mitarbeitende mit sechs aufsteigenden, unterschiedlichen Antwortoptionen

Für die Gestaltung des Fragebogens wurden einerseits Fragen gewählt, bei denen die Personen nur eine Auswahl treffen konnten (sogenannte Single-Choice), mehrere Auswahlmöglichkeiten hatten (sogenannte Multiple-Choice), offene Fragen und Textfelder zur Verfügung standen und letztendlich Fragen, bei denen sie eine Zustimmung oder Ablehnung in verschiedenen Stufen vornehmen mussten sogenannte Likert-Skalen, welche nachfolgend erläutert werden (Fowler, 2013):

Likertskala

Die Likertskalen nach Likert (1932) werden häufig in psychologischen Verfahren verwendet, um Zustimmung oder Ablehnung zu messen. Testpersonen bewerten Aussagen (Items) auf einer Skala, die als Indikatoren für ihre Meinung oder Ausprägung zu einem

Merkmal fungieren (Moosbrugger & Kelava, 2020). Im Fragebogen dieser Bachelorarbeit gaben Personalverantwortliche an, wie die Aussagen auf ihre Unternehmen zutreffen. Eine fünfstufige Likertskala wurde gewählt, da sie sich hinsichtlich Reliabilität und Validität als optimal erwiesen hat (Schmidt-Atzert & Amelang, 2012; Moosbrugger & Kelava, 2020). Die Likertskala fungiert als Ratingskala und war daher besonders geeignet. Diskret gestufte, unipolar beschriftete Ratingskalen wurden verwendet (Moosbrugger & Kelava, 2020). Offene Fragen erforderten keine vorgegebenen Antworten, da das Interesse an qualitativen, neuen Erkenntnissen im Vordergrund stand.

Folgende Fragen wurden mittels Likertskala erfasst:

Seite 4: **Erleben und Empfinden der Tests**

- Frage 4a Übereinstimmung Persönlichkeitstests und Berufserfolg
- Frage 4b Übereinstimmung kognitive Tests und Berufserfolg

Folgende Fragen wurden mittels offener Fragen erhoben:

- Frage 6 Anmerkungen und Feedbacks zu den Tests
- Frage 7 Branchentätigkeit

Abbildung 5
Datenbereinigungsprozess (eigene Darstellung)

Importierte Datenvariable aus Tivian	Umbenannte Datenvariable in Jamovi	Skalenniveau	Stufen
V_1	Testanwendung unterschiedliche Positionen	Nominalskala	4 unterschiedliche Stufen (Optionen)
V_56	Hexaco (K. Lee, M. C. Ashton)	Nominalskala	Ja/nein
V_57	Master Persönlichkeitsanalyse = MPA	Nominalskala	Ja/nein
V_92	FPI-R (Hogrefe)	Nominalskala	Ja/nein
V_58	BIP (Hogrefe)	Nominalskala	Ja/nein
V_59	Myers-Briggs-Typenindikator = MBTI	Nominalskala	Ja/nein
V_60	16 Personalities (kostenlose, Online-Version des MBTI)	Nominalskala	Ja/nein
V_67	Dunkle Triade (Hogrefe)	Nominalskala	Ja/nein
V_65	NEO-FFI (Hogrefe oder Pearson)	Nominalskala	Ja/nein
V_63	DISG oder DISC (Personality, Insides)	Nominalskala	Ja/nein
V_96	ADEPT-15 (AON)	Nominalskala	Ja/nein
V_94	Golden Profiler of Personality = GPOP (Hogrefe)	Nominalskala	Ja/nein
V_66	Eigenentwicklung (selbst konzipierte Tests oder Verfahren)	Nominalskala	Ja/nein
V_68	Andere nämlich... (bitte Namen des Tests angeben)	Nominalskala	Ja/nein

V_69	Andere nämlich... (bitte Namen des Tests angeben) (2)	Nominalskala	Freitexteingabe
V_97	Keine, weil... (bitte Grund angeben)	Nominalskala	Ja/nein
V_98	Keine, weil... (bitte Grund angeben) (Persönlichkeit)	Nominalskala	Freitexteingabe
V_70	IST (Hogrefe)	Nominalskala	Ja/nein
V_93	M-KIT (Hogrefe)	Nominalskala	Ja/nein
V_71	smartPredict (AON)	Nominalskala	Ja/nein
V_72	DESIGMA (Hogrefe)	Nominalskala	Ja/nein
V_73	SHL (CEB SHL Talent Measurement Solutions)	Nominalskala	Ja/nein
V_74	SMART (Schuhfried)	Nominalskala	Ja/nein
V_75	HUCAMA (Hucama AG)	Nominalskala	Ja/nein
V_95	BOMAT (Hogrefe)	Nominalskala	Ja/nein
V_103	WAIS-IV (Pearson)	Nominalskala	Ja/nein
V_102	Raven's 2 (Pearson)	Nominalskala	Ja/nein
V_76	Eigenentwicklung (selbst konzipierte Tests oder Verfahren) (IQ)	Nominalskala	Ja/nein
V_81	Andere nämlich... (bitte Namen des Tests angeben) (2)	Nominalskala	Ja/nein
V_82	Andere nämlich... (bitte Namen des Tests angeben) (IQ) (2)	Nominalskala	Freitexteingabe
V_79	Keine, weil... (bitte Grund angeben) (IQ)	Nominalskala	Ja/nein
V_80	Keine, weil... (bitte Grund angeben) (IQ) (2)	Nominalskala	Freitexteingabe
V_35	Übereinstimmung der Persönlichkeitstests mit Berufserfolg	Ordinalskala	1-5 Likertskala
V_99	Übereinstimmung der kognitiven Tests mit Berufserfolg	Ordinalskala	1-5 Likertskala

V_104	Kosten der Tests	Nominalskala	4 unterschiedliche Stufen (Optionen)
V_40	Allgemeines Feedback zu den Tests	Nominalskala	Freitexteingabe
V_44	Branche	Nominalskala	Freitexteingabe
V_89	Anzahl Mitarbeitende	Ordinalskala	1-6 aufsteigende Rangordnung
V_105	Email-Adresse	Nominalskala	Freitexteingabe

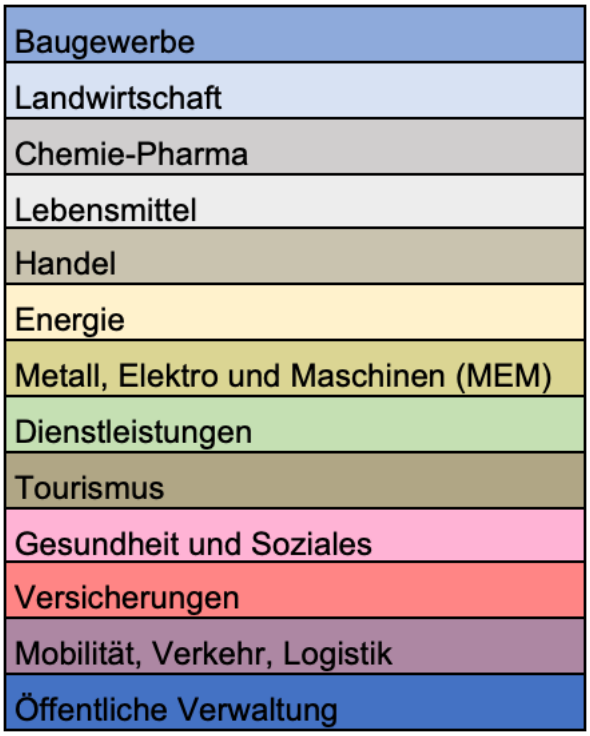
4.4 Datenaufbereitung und –auswertung

Für die Datenaufbereitung wurden die Ergebnisse in ein Excel exportiert, die Daten bereinigt und die Variablen benannt. Die einzelnen Variablen mussten dabei auf das jeweilige Skalenniveau (s. Kapitel 4.3) angepasst werden. Das Excel wurde zur Datenauswertung ins Programm Jamovi importiert, wo die jeweiligen Variablen umbenannt wurden. Für die Auswertung der nominal- und ordinalskalierten Variablen werden:

- Häufigkeitstabellen
- Balken- und Kreisdiagramme
- beobachtete prozentuale Häufigkeiten verwendet.

In der Literatur (Heimsch, 2018) wird ebenfalls diese Vorgehensweise empfohlen. Die Antworten auf die offene Frage bezüglich Branche waren nicht sehr trennscharf. Aufgrund dessen wurde angelehnt an berufsberatung.ch (o. D.) ein eigenes Kategorisierungssystem der Branchen gewählt. Dieses Vorgehen wurde gewählt, da bereits existierende Kategorisierungssysteme zu umfassend waren und teilweise aus bis zu 22 Berufsgruppen bestanden. Anhand der gegebenen Antworten wurden Kategorien gebildet, die nicht zu detailliert, aber gleichzeitig auch nicht zu grob definiert sind. Es wurden 47 «unterschiedliche» Branchen von den 53 Unternehmen angegeben, welche in der untenstehenden Tabelle in insgesamt 13 Branchen zusammengefasst wurden. Folgende Branchen haben sich in Abbildung 6 herauskristallisiert:

Abbildung 6
Eigene Einteilung der 13 Branchen



4.5 Einordnung der Gütekriterien der genutzten Tests mittels Literaturrecherche

Um die psychometrische Qualität der eingesetzten Tests zu bewerten, wurde eine Tabelle erstellt, die im Ergebnisteil präsentiert wird. Diese Tabelle, von der Autorin entworfen, basiert auf Kennwerten aus der Literatur und zielt darauf ab, die Gütekriterien der eingesetzten Tests zu bewerten und eine Übersicht über psychometrische Online-Tests in der Nordwestschweiz zu bieten. Die verwendeten Literaturquellen sind dem Anhang zu entnehmen. Die Tabelle berücksichtigt Kennwerte zur Validität und Reliabilität und deren Unterformen. Aufgrund von Praktikabilitätsgründen konnte keine umfassende Übersicht über alle Gütekriterien gegeben werden und es wurde in Absprache mit dem Praxispartner und der Betreuungsperson die wichtigsten Kriterien ausgewählt. Diese Tabelle ist ein wesentlicher Bestandteil des Methodikteils weil sie in den nächsten zwei Kapiteln wieder aufgegriffen wird und sie die methodische Vorgehensweise und die psychometrische Qualität der Daten verdeutlicht.

5 Ergebnisse

In diesem Kapitel werden die wichtigen Ergebnisse vorgestellt, die für die Beantwortung der Fragestellung relevant sind. Dazu werden Grafiken wie Diagramme und Tabellen verwendet, welche die wichtigsten Kernaussagen veranschaulichen und hervorheben sollen.

5.1 Darstellung der Ergebnisse

Die Umfrage wurde an 219 Unternehmen versendet und von 53 vollständig ausgefüllt und beendet. Dies stellt eine Rücklaufquote von 24.2% dar. Da das Ziel darin bestand mindestens 50 Rückmeldungen zu erhalten, konnte dieses erreicht werden.

Stichprobenbeschreibung nach Branchen

Abbildung 7
Häufigkeiten der Teilnahme der Unternehmen nach Branche

Häufigkeiten

Häufigkeit von Branchen_Cluster

Branchen_Cluster	Anzahl	% von Gesamt	kumulierte %
Baugewerbe	2	3.8%	3.8%
Landwirtschaft	1	1.9%	5.7%
Chemie-Pharma	10	18.9%	24.5%
Lebensmittel	2	3.8%	28.3%
Handel	4	7.5%	35.8%
Energie	1	1.9%	37.7%
MEM	16	30.2%	67.9%
Dienstleistungen	3	5.7%	73.6%
Tourismus	1	1.9%	75.5%
Gesundheit und Soziales	2	3.8%	79.2%
Versicherungen	5	9.4%	88.7%
Mobilität/Verkehr/Logistik	5	9.4%	98.1%
Öffentliche Verwaltung	1	1.9%	100.0%

Die meisten Unternehmen kamen aus der MEM-Branche (16), der Chemie-Pharma-Branche (10) sowie der Versicherungsbranche und der Mobilität/Verkehr/Logistik-Branche (jeweils 5). Die restlichen Häufigkeiten der Teilnehmenden können in der oberen Abbildung 7 entnommen werden.

Stichprobenbeschreibung im Hinblick auf die Mitarbeitendenanzahl

Abbildung 8
Stichprobenbeschreibung anhand Mitarbeitendenanzahl

Häufigkeit von Anzahl Mitarbeitende

Anzahl Mitarbeitende	Anzahl	% von Gesamt	kumulierte %
Weniger als 50	1	1.9%	1.9%
50 - 99	3	5.7%	7.5%
100 - 249	18	34.0%	41.5%
250 - 499	10	18.9%	60.4%
500 - 999	5	9.4%	69.8%
1000 und mehr	16	30.2%	100.0%

Abbildung 8 zeigt einen Auszug aus Jamovi, welcher die Stichprobe im Hinblick auf die Anzahl MA (Unternehmensgrössen) beschreibt. Die meisten Unternehmen hatten 100 bis 249 MA (18), gefolgt von Grossunternehmen mit 1000 und mehr MA (16). Danach folgen Unternehmen mit 250 bis 499 MA (10), 500 bis 999 MA (5) und schliesslich Unternehmen mit weniger als 50 bis 99 MA (4).

5.2 Anwendung der Online-Tests

Abbildung 9
Testanwendung unterschiedliche Positionen

Deskriptivstatistik

	Testanwendung unterschiedliche Positionen	Branchen_Cluster
N	Anwendung Online-Tests best. Positionen	23
	Anwendung Online-Tests alle Positionen	2
	Andere Verfahren / gar keine	26
	Anwendung Tests Papierform	2

Die erste Frage des Fragebogens betrifft die generelle Anwendung von Online-Tests in Unternehmen, wie in Abbildung 9 zu sehen ist. Die meisten Unternehmen (26) nutzen keine Online-Tests, sondern andere Verfahren zur Personalauswahl. 23 Unternehmen nutzen Online-Tests nur bei bestimmten Positionen. Zwei Unternehmen setzen Online-Tests bei allen Positionen (auch Lernende) ein und zwei verwenden Papier-Bleistift-Tests. Anwendung von Online-Testverfahren, aber nur bei bestimmten Positionen. Wenn man die unterschiedlichen Antworten nur im Hinblick auf die Anwendung von Testverfahren (online und Papierform) betrachtet, nutzen mehr als die Hälfte der befragten Unternehmen diese im Rekrutierungsprozess (27).

Das Balkendiagramm (Abbildung 10) zeigt die Testanwendungen nach Branchen. Hierbei fällt auf, dass innerhalb der MEM-Branche der grösste Anteil an Unternehmen und auch drei unterschiedliche Testanwendungen vorhanden sind.

Abbildung 10
Diagramm Testanwendungen nach Branche und Form

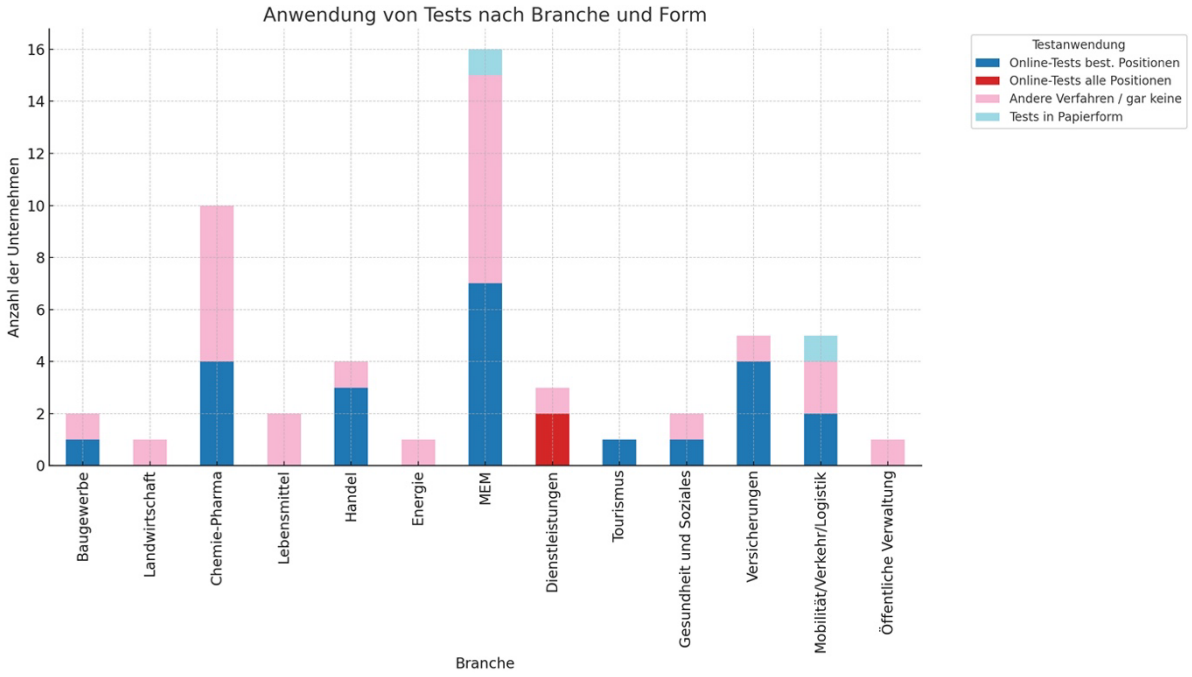
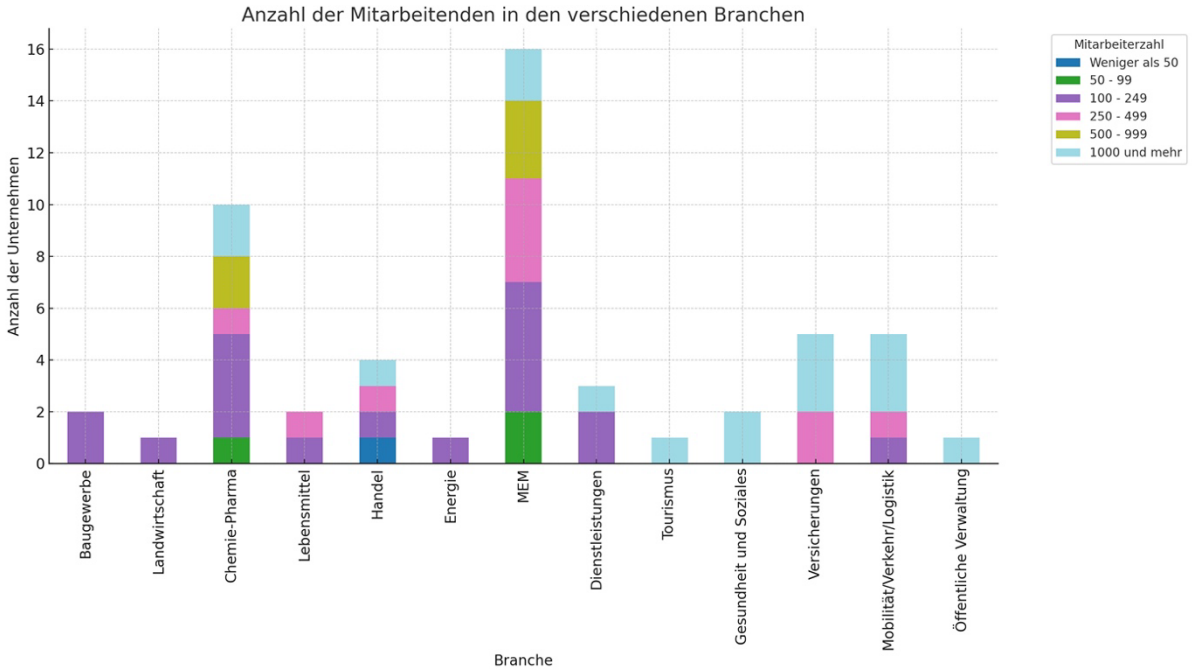


Abbildung 11
Diagramm Anzahl Mitarbeitende in verschiedenen Branchen



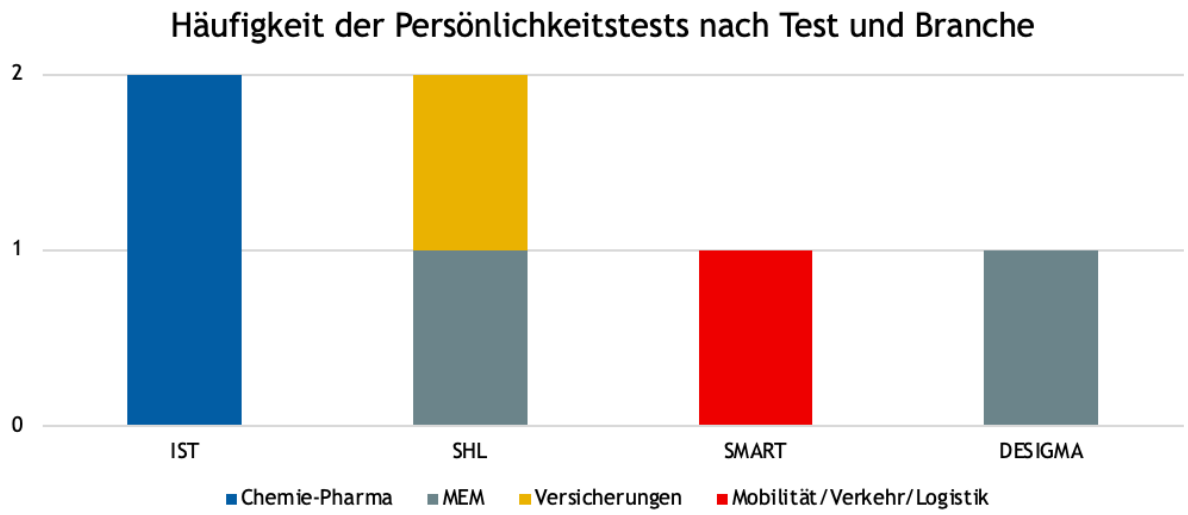
Das Diagramm (Abbildung 11) zeigt die Anzahl der Unternehmen in verschiedenen Branchen, gruppiert nach Anzahl Mitarbeitenden. In der Chemie-Pharma gibt es Unternehmen in allen Kategorien, meist mit 100-249 MA. Die MEM-Branche hat viele Unternehmen in den Kategorien 100-249, 250-499 und 500-999 MA. Baugewerbe, Landwirtschaft, Energie und Tourismus haben insgesamt weniger Unternehmen und beschränken sich auf bestimmte Kategorien.

5.3 Verbreitung von psychometrischen Online-Tests in Unternehmen der Nordwestschweiz

5.3.1 Auswahl Intelligenztests und Persönlichkeitstests

Folgende IQ-Tests werden von **sechs** der befragten **Unternehmen** eingesetzt:

Abbildung 12
Diagramm Angewendete Intelligenztests mit zugehörigen Branchen



In Abbildung 12 sind die angewendeten Intelligenztests nach Branchen in einem Balkendiagramm dargestellt. Von den befragten 53 Unternehmen nutzen sechs insgesamt vier der zur Auswahl stehenden Intelligenztests. Der IST und SHL werden jeweils von zwei Unternehmen genutzt und in den Branchen Chemie-Pharma, MEM und Versicherungen angewendet. Der DESIGMA wird von einem Unternehmen in der MEM-Branche und der SMART von einem in der Mobilität/Verkehrs/Logistik-Branche benutzt. Abbildung 13 listet Intelligenztests auf, die nicht genutzt werden.

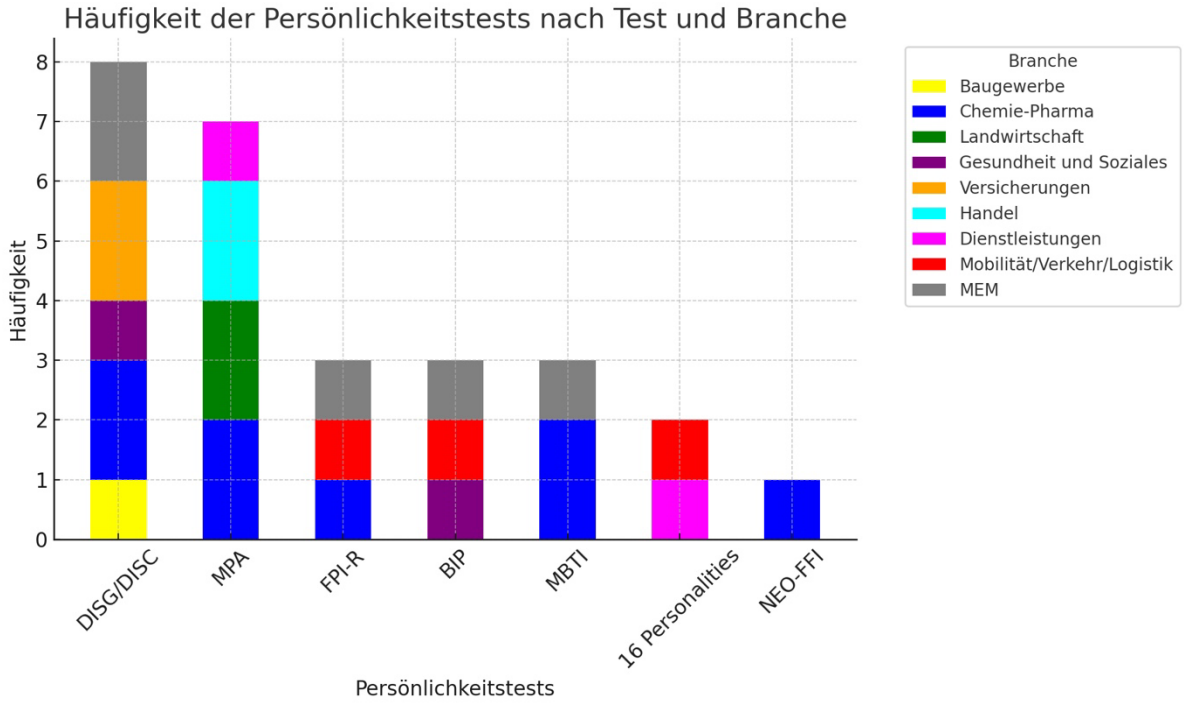
Folgende IQ-Tests werden von **keinen** der befragten **Unternehmen** eingesetzt:

Abbildung 13
Nicht angewendete Intelligenztests

M-KIT (Hogrefe)
smartPredict (AON)
HUCAMA (Hucama AG)
BOMAT (Hogrefe)
WAIS-IV (Pearson)
Raven's 2 (Pearson)

Folgende Persönlichkeitstests werden von **27** der befragten **Unternehmen** eingesetzt:

Abbildung 14
Angewendete Persönlichkeitstests mit zugehörigen Branchen



In Abbildung 14 hat die Autorin die angewendeten Persönlichkeitstests mit den zugehörigen Branchen in einem Balkendiagramm nach Häufigkeit dargestellt. Es ist ersichtlich, dass bei den befragten Unternehmen (N=53) insgesamt sieben von elf der zur Auswahl gestandenen Persönlichkeitstests eingesetzt werden. Insgesamt wenden 27 von den 53 befragten Unternehmen obenstehende Persönlichkeitstests an. Der meisteingesetzte Persönlichkeitstest ist der DISG/DISC mit rund acht Angaben. Kurz danach folgt der MPA mit sieben Angaben. Der FPI-R und der BIP wurden dreimal ausgewählt. Wohingegen der MBTI und der 16 Personalities zweimal ausgewählt wurden und am wenigsten häufig wurde der NEO-FFI mit nur einer Angabe. In der Abbildung 15 werden diejenigen Persönlichkeitstests tabellarisch aufgelistet, welche nicht angewendet werden.

Folgende Persönlichkeitstests werden von **keinen** der befragten **Unternehmen** eingesetzt:

Abbildung 15
Nicht angewendete Persönlichkeitstests

Hexaco (K.Lee, M.C. Ashton)
Dunkle Triade (Hogrefe)
ADEPT-15 (AON)
Golden Profiler of Personality = GPOP (Hogrefe)

Kategorisierung der Gründe für möglichen Verzicht auf IQ-Tests

Abbildung 16
Kategorisierung der Gründe für Verzicht auf IQ-Tests

Deskriptivstatistik

	Gründe für Verzicht auf IQ-Tests	Branchen_Cluster
N	keine Angabe	29
	andere Verfahren	7
	als nicht notwendig erachtet	14
	mangelnde Ressourcen um geeigneten Anbieter zu finden	1
	Tests nur in Kombination mit externen Beratungsfirmen (Expert:innen)	2

Bei Abbildung 16 ist ersichtlich, dass bei der Frage nach kognitiven Leistungstests die Möglichkeit bestand, das Item «Keine, weil (bitte Grund angeben..)» auszuwählen. Hierbei war das Ziel zu ermitteln, was mögliche Kriterien sind, dass Unternehmen keine Intelligenztests anwenden. Die Autorin kategorisierte die freien Antworten in vier Gruppen. 29 Teilnehmende machten keine Angabe, da sie die Tests einsetzen. Häufigster Grund für den Verzicht war der fehlende Mehrwert der Tests (Überflüssigkeit) (14). Weitere Gründe waren die Bevorzugung anderer Verfahren (7), die Nutzung der Tests nur mit geschulten Expert:innen (2) und fehlende Ressourcen bzw. Unsicherheit bei der Auswahl eines Testanbieters (1).

5.4 Kategorisierung der Gründe für möglichen Verzicht auf Persönlichkeitstests

Abbildung 17
Kategorisierung Gründe Verzicht auf Persönlichkeitstests

Häufigkeit von Gründe für Verzicht auf Persönlichkeitstests

Gründe für Verzicht auf Persönlichkeitstests	Anzahl	% von Gesamt	kumulierte %
andere Verfahren	7	38.9 %	38.9 %
als nicht notwendig erachtet	7	38.9 %	77.8 %
Tests nur in Kombination mit externen Beratungsfirmen (Expert:innen)	3	16.7 %	94.4 %
mangelnde Ressourcen um geeigneten Anbieter zu finden	1	5.6 %	100.0 %

Die Kategorisierung für den möglichen Verzicht auf Persönlichkeitstests konnte identisch zum Verzicht der Intelligenztests übernommen werden. Das Ziel war dasselbe wie bei den Intelligenztests und die vier Kategorien bleiben gleich. Wie aus Abbildung 17 ersichtlich ist, sind die Hauptgründe die Nutzung anderer Verfahren und die Einschätzung der Tests, als nicht notwendig (jeweils 7).

Häufigkeit von eigenentwickelten Verfahren (Intelligenztests)

Abbildung 18
Häufigkeit von Eigenentwicklung (IQ-Tests)

Häufigkeit von Eigenentwicklung (selbst konzipierte Tests oder Verfahren) (IQ)

Eigenentwicklung (selbst konzipierte Tests oder Verfahren) (IQ)	Anzahl	% von Gesamt	kumulierte %
nein	43	81.1%	81.1%
ja	10	18.9%	100.0%

Eine weitere Möglichkeit, warum Unternehmen keine Intelligenztests anwenden, kann die Nutzung selbst entwickelter Verfahren sein. Bei der Umfrage verwendet ein kleinerer Anteil eigenentwickelte Tests, was 10 Unternehmen entspricht. Der Grossteil (43) nutzt keine eigenentwickelten Verfahren. Diese Angaben sind in Abbildung 18 auch in Prozentzahlen angegeben.

Häufigkeit von eigenentwickelten Verfahren (Persönlichkeitstests)

Abbildung 19
Häufigkeit von Eigenentwicklung (Persönlichkeitstests)

Häufigkeit von Eigenentwicklung (selbst konzipierte Tests oder Verfahren) (P)

Eigenentwicklung (selbst konzipierte Tests oder Verfahren) (P)	Anzahl	% von Gesamt	kumulierte %
nein	44	83.0%	83.0%
ja	9	17.0%	100.0%

In Abbildung 19 bestand dieselbe Antwortmöglichkeit der Eigenentwicklung bei den Persönlichkeitstests, wo die Anzahl um ein Unternehmen abnimmt und somit 17% auf eigenentwickelte Verfahren einsetzen.

Anwendung von anderen Testverfahren

Abbildung 20
Anwendung von anderen Testverfahren (Persönlichkeit und Intelligenz)

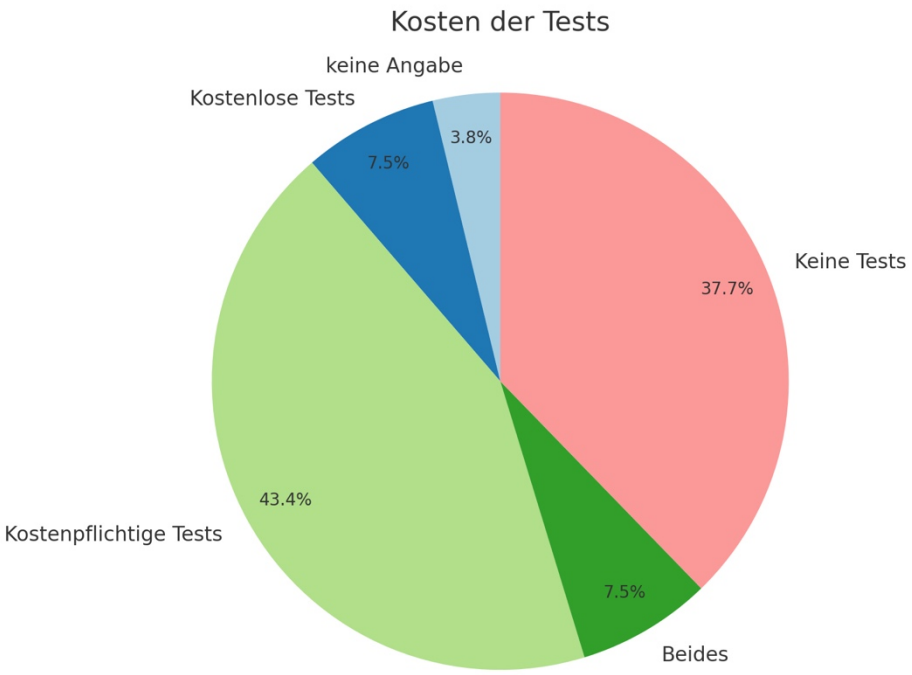
Persönlichkeitstests	Intelligenztests
LPE	Cyquest
Versus Talent Profile	ACE (www.master.ch)
Persönlichkeitsstärken der Uni Zürich	AEC DISC
AEC Disc	DNLA
DNLA	Persolog
Persolog	Thomas International
Sales Master	Cyquest (Lernende), HR Analytics (Kundenberatung)
Insight	Kognitiver Test ist auch im Online Assessment der Firma HR Diagnostics Bestandteil
VPA Persönlichkeits Analyse nach Thomas International	PXT
Cyquest für Lernende und AC Online für Kundenberatung	
Online Assessments von der Firma HR Diagnostics	
PXT	

Ein weiterer Grund für die Nichtanwendung der angebotenen Tests ist der Einsatz anderer Testverfahren. Dafür konnten die befragten Unternehmen das konstruierte Item «Andere, nämlich… (bitte Namen des Testes eingeben» nutzen. Diese freien Antworten sind in der Abbildung 20 festgehalten. Links wurden zwölf Persönlichkeitstests und rechts neun Intelligenztests angegeben. Die Bewertung dieser Tests wird in der Bachelorarbeit nicht vorgenommen, die Auflistung dient nur als Übersicht. Die Untersuchung der

Gütekriterien dieser Tests würde den Rahmen dieser Arbeit sprengen, könnte aber in zukünftiger Forschung erfolgen.

Kosten der Tests

Abbildung 21
Diagramm Kosten der Tests



Eine Frage des Fragebogens bezieht sich auf die Kosten der angewendeten Tests. 23 Unternehmen nutzen kostenpflichtige Tests, vier nutzen kostenlose Tests und vier setzen sowohl kostenpflichtige als auch kostenlose Tests ein. Zwei Unternehmen machten keine Angaben und 20 setzen keine Tests ein. Diese Informationen sind im Kreisdiagramm (Abbildung 21) in Prozentangaben dargestellt.

Kategorisierung der allgemeinen Feedbacks und Rückmeldungen zu den genutzten Tests

Die Autorin versuchte Rückmeldungen zu den Testverfahren zu kodieren und zu kategorisieren. Dies gelang jedoch nicht in einem gleich übersichtlichen Schema, wie bei den Gründen auf den Verzicht von IQ-Tests und Persönlichkeitstests. Hier gaben Unternehmen unterschiedliche Rückmeldungen, die zu individuell waren, um eine gröbere Kategorisierung zu erstellen.

Abbildung 22
Allgemeines Feedback zu den Tests

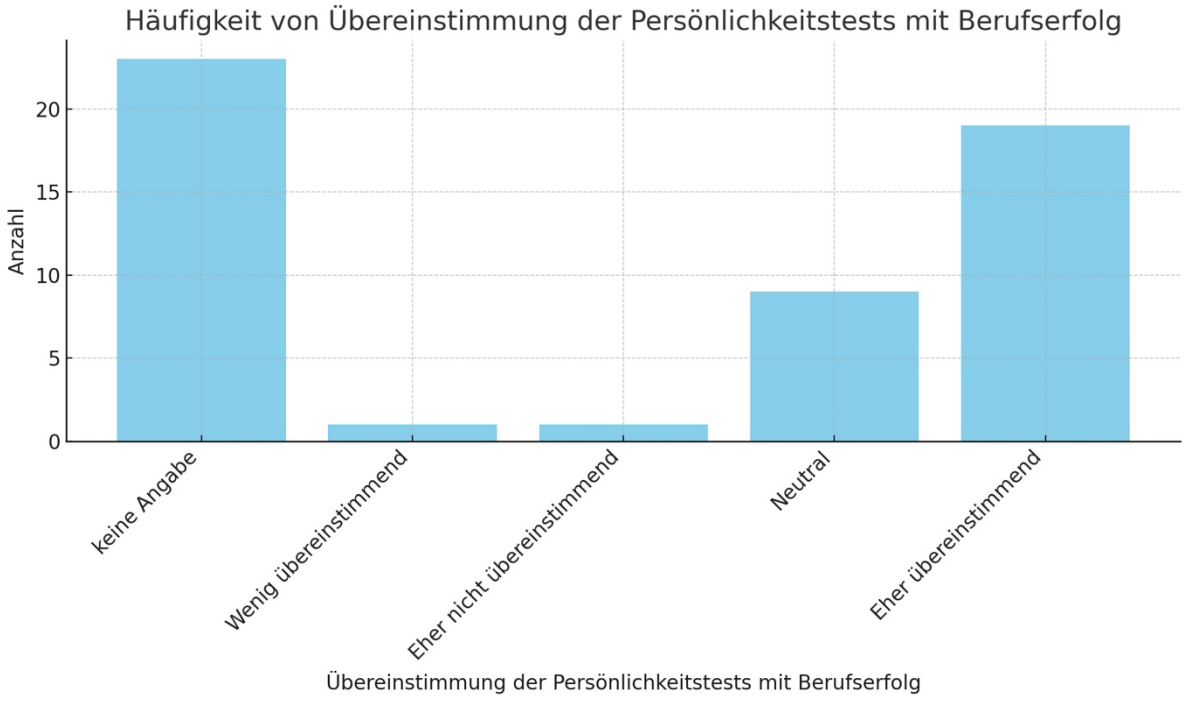
Häufigkeit von Strukturierung_Feedback_allgemein

Strukturierung_Feedback_allgemein	Anzahl	% von Gesamt	kumulierte %
keine Angabe	43	81.1 %	81.1 %
Potenzialanalyse, Unternehmen fördert das Potenzial nicht (fehlende Personalentwicklung)	1	1.9 %	83.0 %
Regelmässige Anwendung wichtig, um Sicherheit für gewisse Berufsgruppen sicherzustellen	1	1.9 %	84.9 %
Kandidat:innen stimmen Testresultaten zu / legen mehr Wert auf Team-Fit / Testresultate werden nicht weiter verglichen	1	1.9 %	86.8 %
Nutzen Tests als Ergänzung zu anderen Verfahren	3	5.7 %	92.5 %
Empfinden IQ-Tests in Rekrutierungen nicht immer sinnvoll	1	1.9 %	94.3 %
Nutzen Online-Assessment mit diversen Tests	1	1.9 %	96.2 %
Empfindet Kosten sehr hoch / wünscht qualitativ hohe, kostenlose Tests	1	1.9 %	98.1 %
Wenden nur Tests mit Analyse durch Expert:innen an (Fehlinterpretationen)	1	1.9 %	100.0 %

Schliesslich wurde von rund zehn Unternehmen ein (freiwilliges) Feedback zu den Testverfahren angegeben, welches zu strukturieren versucht wurde. Es wurde jedoch nur eine Übereinstimmung in der Kategorisierung gefunden, und zwar «nutzen Tests als Ergänzung zu anderen Verfahren» diese Aussage wurde von jeweils drei Unternehmen rückgemeldet. Die restlichen sieben Aussagen sind zu individuell und können deshalb nicht zusammengefasst werden. Für eine bessere Übersicht wurden die diversen Rückmeldungen auf die wesentlichste Kernaussage reduziert und somit kann aus der Abbildung 22 weiteres Feedback entnommen werden.

Übereinstimmung der Persönlichkeitstests mit Berufserfolg

Abbildung 23
Übereinstimmung der Persönlichkeitstests mit Berufserfolg



Bei der Frage nach der Übereinstimmung der Persönlichkeitstests mit Berufserfolg zeigte sich Folgendes: Die häufigste Antwort war «keine Angabe» (23 Nennungen), dies

sind Angaben der Unternehmen, die diese nicht einsetzen oder sich nicht äussern wollten. «Eher übereinstimmend» war somit die grösste Kategorie mit 19 Nennungen. «Wenig übereinstimmend und «eher nicht übereinstimmend» hatten jeweils nur eine Nennung, wie in Abbildung 23 zu sehen ist. Es ist anzumerken, dass es sich hierbei um eine qualitative Einschätzung seitens Personalverantwortlichen handelt.

Übereinstimmung der Intelligenztests mit Berufserfolg

Abbildung 24
Übereinstimmung der kognitiven Leistungstests mit Berufserfolg

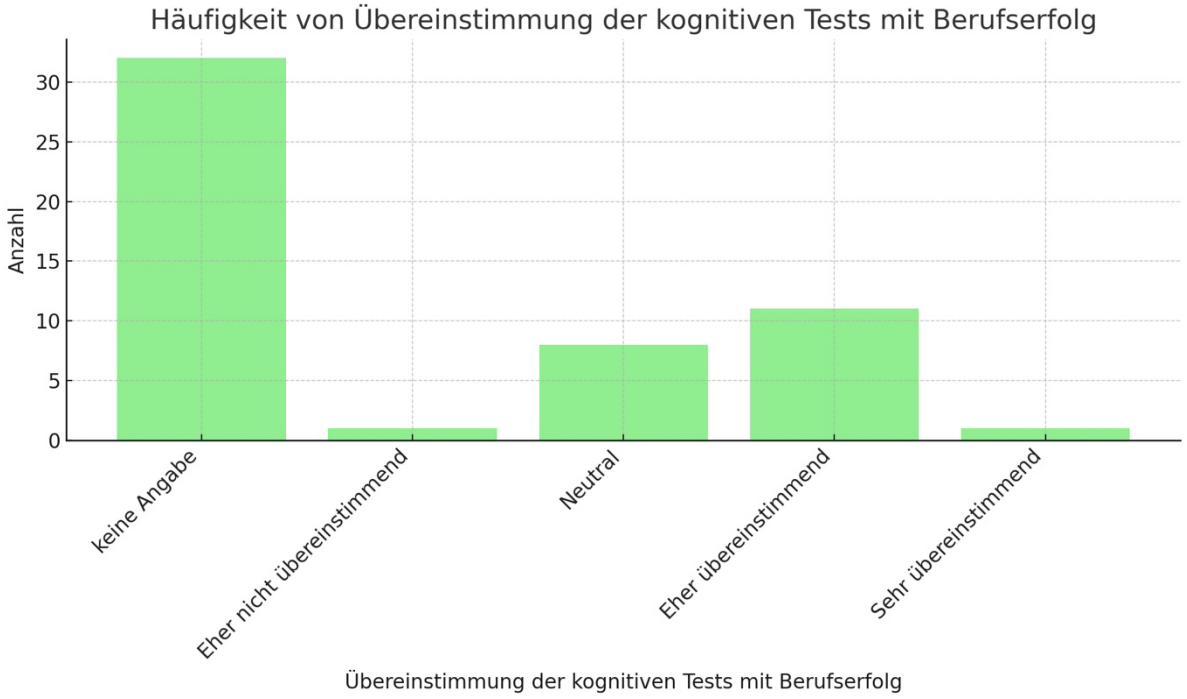


Abbildung 24 zeigt die Übereinstimmung der kognitiven Leistungstests mit dem Berufserfolg. Die häufigste Antwort war «keine Angabe» (32 Nennungen), dies sind Angaben der Unternehmen, die diese nicht einsetzen oder sich nicht äussern wollten. «Eher übereinstimmend» hatte elf Nennungen und war somit die grösste Kategorie. «Neutral» hatte acht Nennungen und danach folgten «eher nicht übereinstimmend» und «sehr übereinstimmend» (jeweils eine Nennung).

5.5 Bewertung der Qualität der vorhandenen Tests im Hinblick auf die Gütekriterien

Wie im theoretischen Teil erläutert, ist die Qualität eines Tests entscheidend für die Aussagekraft und prognostische Validität. Für Nicht-Expert:innen ist dies jedoch oft schwer diese einzuschätzen (Schmidt-Atzert & Amelang, 2012). Nachfolgend wird die, bereits im Methodikteil erwähnte, von der Autorin erstellte, Tabelle (Abbildung 25) mit dazugehöriger

Legende (Abbildung 26) präsentiert. Dabei wird die Testqualität anhand der in Kapitel 2 besprochenen Gütekriterien bewertet.

Abbildung 25
Bewertung der Qualität der eingesetzten Testverfahren

	Gütekriterien				
	Reliabilität		Validität		
Test	**Interne Konsistenz**	**Test-Retest-Reliabilität**	**Inhaltsvalidität**	**Konstruktvalidität**	**Kriteriumsvalidität**
NEO-FFI	α = .72 - .87	r = .71 - .82		r = .54 - .8	
DISG/DISC					
MPA	α = .84	r = .6 – .8		r = .5 – .7	r = .5 – .7
FPI-R	α = .72 - .82	r = .42 – .72			
BIP	α = .74 - .91	r = .71 - .79		r = .54 - .84	r = .41–.49
MBTI					
16 Personalities					
IST	α = .85 - .97	r = .72 - .90		r = .63, .69 - .49	
SHL					
DESIGMA	α = .87 -.96			r = .67	r = .24 - (.44)
SMART	α = .7 - .9				

Abbildung 26
Legende der Kennwerte (Beurteilung der Gütekriterien)

Gütekriterium	Niedrige Werte	Mittelmässige Werte	Hohe Werte	Nicht vorhandene Werte oder widersprüchliche Befunde	Vorhanden, jedoch kein exakter Wert (Einordnung nicht möglich)
Validität					
Konstruktvalidität (bezogen auf die faktorielle Validität)	Kleiner als 0.3	0.3 bis 0.5	Grösser als 0.5	Diese Farbe wird gewählt, wenn die Kennwerte für die einzelnen Gütekriterien nicht gefunden werden können (es liegen keine genaue Angaben vor). Dies kann zum Beispiel der Fall sein, wenn Anbieter diese Werte nicht publiziert haben (Manual) oder wenn widersprüchliche Befunde oder unzureichende, kritische Aussagen zu den einzelnen Tests gefunden werden. Wie im Theorieteil erwähnt, kann dies ein Indiz dafür sein, dass ein Test die Gütekriterien nicht erfüllt.	
Kriteriumsvalidität	Kleiner als 0.2	0.2 bis 0.4	Grösser als 0.4		
Inhaltsvalidität* *keine numerischen Schwellenwerte, Einschätzung basiert auf Gutachten von Expert:innen	Unzuverlässig, führen zu verzerrten Ergebnissen, Konstrukt wird in den Items nicht abgebildet	Teilweise Abdeckung des Konstrukts, können durch Hinzufügen oder Überarbeiten von Items verbessert werden	Vollständige und repräsentative Abdeckung des Konstrukts		
Reliabilität					
Interne Konsistenz	Kleiner als 0.6	0.6 bis 0.69	0.7 bis 0.9		
Test-Retest-Reliabilität	Kleiner als 0.5	0.5 bis 0.69	0.7 bis 0.9		

Anmerkung: In Abbildung 26 wurden die Kennwerte und Kriterien für die Beurteilung der Gütekriterien festgehalten. Die Farben dienen als Orientierungshilfe und ermöglichen eine schnelle Einordnung, ohne dass die genauen Zahlen betrachtet werden müssen. Die Kennwerte und Hinweise wurden aus verschiedenen Literaturquellen zusammengetragen und sind im Anhang zu finden.

6 Diskussion

Im folgenden Kapitel werden die Umfrageergebnisse mit der bestehenden Theorie verknüpft, um die Forschungsfrage zu beantworten und Implikationen für Theorie und Praxis abzuleiten.

6.1 Verbreitung der Online-Tests

Die vorliegende Bachelorarbeit untersucht die Verbreitung und Nutzung von psychometrischen Online-Tests bei der Personalrekrutierung und -auswahl von Unternehmen in der Nordwestschweiz. Psychometrische Tests bieten eine objektive Möglichkeit, die Eignung von Kandidat:innen für bestimmte Positionen zu beurteilen, indem sie Intelligenz und Persönlichkeitsmerkmale messen. Die Relevanz dieser Tests liegt in ihrer Fähigkeit, die Passung zwischen Bewerbenden und den Anforderungen der Positionen sowie des spezifischen Unternehmens zu verbessern. Trotz der erwiesenen Vorteile, die im Theorieteil der Arbeit zusammengefasst wurden, zeigen die Ergebnisse der Umfrage, dass psychometrische Online-Tests in der Nordwestschweiz selten eingesetzt werden. Die Ergebnisse zeigen, dass von den 53 befragten Unternehmen 25 Online-Tests einsetzen, während 26 Unternehmen keine Tests und somit andere Verfahren nutzen. Zwei Unternehmen nutzen Tests, aber in Papierform. Diese Ergebnisse lassen sich mit den theoretischen Erkenntnissen nach Müllerschön (2006) erklären, welcher Online-Tests und deren Vorteile wie z.B. die Ortsunabhängigkeit, gegenüber Papier-Bleistift-Tests befürwortet. Das bedeutet, dass fast die Hälfte aller Befragten zwar Online-Tests einsetzen, diese jedoch nicht den psychometrischen Anforderungen entsprechen bzw. die Gütekriterien erfüllen. Die Ergebnisse deuten auf eine teilweise Nutzung von Online-Tests hin. Dies steht im Gegensatz zu den erwarteten Vorteilen dieser Tests. Insbesondere bei den Intelligenztests ist eine seltene Anwendung zu beobachten, obwohl sie als einer der besten Prädiktoren für beruflichen Erfolg gelten (Schmidt & Hunter, 1998). Von den aufgelisteten Intelligenztests haben insgesamt sechs Unternehmen deren Anwendung bestätigt, weitere neun Intelligenztests wurden angegeben, die sich nicht namentlich in der Auflistung befanden. So kann man davon ausgehen, dass von 53 Unternehmen insgesamt 15 Intelligenztests verwenden, was 28,30% entspricht und somit weniger als ein Drittel darstellt. Dementsprechend decken sich diese Ergebnisse der Umfrage auch mit der Theorie von Kanning (2019), welcher argumentiert, dass Intelligenztests noch immer wenig verbreitet sind. Die Gründe dafür sind vielseitig – die Ergebnisse zeigen, dass andere Verfahren bevorzugt werden und dass die Testverfahren allein nicht als hinreichend empfunden werden. Dasselbe gilt auch für die Persönlichkeitstests. Ausserdem wird von

vielen eine Überflüssigkeit gegenüber der Anwendung von Persönlichkeitstests und Intelligenztests geäussert, was dementsprechend so interpretiert werden kann, dass die Wirkung und Validität der Testverfahren unterschätzt wird, während andere Verfahren überschätzt werden, was auch in der Forschung von Funk et al. (2015) und sieben Jahre später auch von Hogrefe (2022) diskutiert wird. Zudem äussern viele Personalverantwortliche Vorurteile oder Bedenken, was sich auch in der Theorie von Steininger (2020) widerspiegelt.

Von den aufgelisteten Persönlichkeitstests haben insgesamt 27 Unternehmen deren Anwendung bestätigt, weitere zwölf Persönlichkeitstests wurden angegeben, die sich nicht namentlich in der Auflistung befanden. Dementsprechend setzen insgesamt 39 von 53 Unternehmen auf Persönlichkeitstests – deren Psychometrik wird dabei nicht diskutiert. Dies sind 73,58% und entspricht beinahe drei Viertel und mehr als der Hälfte. Das zeigt, dass die Nutzung von Persönlichkeitstests verbreiteter ist, als die von Intelligenztests - unabhängig von den psychometrischen Kriterien. So kann man davon ausgehen, dass die Relevanz der Persönlichkeitstests im Arbeitskontext schon von mehr Unternehmen erkannt wurde. Dies könnte daran liegen, dass die Persönlichkeit für viele Unternehmen in dieser Umfrage wichtiger ist als die Intelligenz und die gewonnenen Erkenntnisse dazu besser verbreitet sind. Wie Neyer und Asendorpf (2024) berichten stehen gewisse Persönlichkeitsmerkmale eng im Zusammenhang mit Arbeitsleistung und einer guten Führung, wobei man annehmen kann, dass diese Unternehmen sich dessen bewusst sind. Die Ergebnisse zeigen, dass Unternehmen aus der MEM-Branche sowie der Chemie-Pharma-Branche häufiger Online-Tests nutzen, als andere Branchen. Dies könnte darauf hinweisen, dass in technischen und naturwissenschaftlichen Branchen ein höherer Wert auf objektive Auswahlkriterien (in Form von Testverfahren) gelegt wird. Dies steht im Einklang mit der Theorie, dass Unternehmen in diesen Sektoren vermehrt komplexere und spezialisierte Kompetenzen suchen, welche durch objektive Testverfahren besser beurteilt werden können (Steiner, 2020). Die häufigsten Gründe für den Verzicht auf Intelligenztests sind die Annahme, dass solche Tests keinen Mehrwert bieten oder dass andere Verfahren (wie z.B. ein klassisches Einstellungsinterview) bevorzugt werden. Bei Persönlichkeitstests stehen ähnliche Gründe im Vordergrund. Eine Übereinstimmung zwischen den Ergebnissen der Umfrage und der Theorie von Hogrefe (2020) lässt sich dadurch finden, indem dass Vorurteile und Unsicherheiten bezüglich der Validität und Praktikabilität psychometrischer Tests vorliegen, was letztendlich die Anwendung hemmen kann. Viele Personalverantwortliche nennen Vorurteile gegenüber der Anwendung von Online-Tests. Oftmals sind die Beweggründe dafür mangelndes Vertrauen und Unwissen über die neuesten wissenschaftlichen Erkenntnisse zur Aussagekraft psychometrischer Online-

Tests. Dies unterstreicht ebenfalls die Forschung von Funk et al. (2015), welche aussagt, dass Personalverantwortliche unzureichend über die neuesten Forschungsergebnisse informiert sind und daher auch die Prognosekraft von psychometrischen Tests unterschätzen und weniger valide Verfahren überschätzen.

Was die Kosten der Tests betrifft, ist zu beobachten, dass die Mehrheit (43.4%) kostenpflichtige Tests einsetzt und nur ein kleiner Teil (7.5%) kostenlose und 7.5% beide verwendet. Es gilt zu erwähnen, dass 37.7% keine Tests verwenden und 3.8% keine Angaben dazu machen. Ein Unternehmen, welches kostenpflichtige Tests verwendet, meldet als allgemeines Feedback zurück, dass die Lizenzen für solche Testverfahren oftmals sehr teuer sind, und wünscht sich valide und reliable Testverfahren zu tieferen Preisen. Damit könnte auch erklärt werden, wieso die Anwendung von psychometrischen Verfahren noch nicht weit verbreitet ist. So ist auch dem Theorieteil zu entnehmen, dass Personalverantwortliche die zusätzlichen Entstehungskosten durch Online-Tests scheuen (Hogrefe, 2022). Wenn man jedoch bedenkt, wie hohe Kosten man durch eine Fehlbesetzung verursachen kann, lohnt sich die nun etwas kostspieligere Investition in die Personalrekrutierung- und Auswahl mit Online-Tests wie auch Wöhe und Döring (2013) berichten. Dazu gehören gut durchdachte und sorgfältig geplante Rekrutierungsstrategien und bewusste Entscheidungen mithilfe objektiver Auswahlkriterien wie psychometrischen Testverfahren (Ekwoaba et al., 2015). Somit sollte nach Sangeetha (2010) der Fokus auf den Mehrwert und den positiven Einfluss einer sauber geplanten Personalauswahl auf den ROI und letztendlich auf den Erfolg eines Unternehmens gelegt werden. Bei der Analyse der Umfrageergebnisse könnte auch interpretiert werden, dass die Kosten für die Durchführung der Tests als Hindernis wahrgenommen werden. Dies betrifft nicht nur die direkten Kosten bspw. für den Einkauf der Lizenz, sondern auch die indirekten Kosten für die Schulung des Personals oder dem Beizug von externen Experten um die Qualität der Durchführung, Auswertung und Interpretation zu gewährleisten. Denn von einigen Unternehmen wird berichtet, dass sie keine Testverfahren anwenden, weil ihnen dazu die geschulten Expert:innen für Durchführung und Interpretation fehlen. Unternehmen müssten möglicherweise die gewinnbringenden Vorteile der Investitionen mehr hervorheben und ein Umdenken sollte geschehen, denn wie Bartscher et al. (2012) berichtet, können voreilige Entscheidungen Unternehmen finanziell enorm belasten.

Ein weiterer Grund, wieso Unternehmen keine Online-Tests anwenden, liegt darin, dass ein Teil der befragten Unternehmen eigenentwickelte Verfahren nutzen. Der Vorteil könnte darin liegen, dass die Anforderungen eines Unternehmens mit einer Eigenentwicklung individuell zugeschnitten werden und damit besser auf die spezifischen Bedürfnisse eingegangen werden kann. Allerdings besteht gleichzeitig auch das Risiko,

dass die selbst entwickelten Verfahren nicht dieselbe wissenschaftliche Basis und Validität aufweisen, wie psychometrische Tests. Dies bestätigt die Beobachtungen von Kanning (2022), dass viele Unternehmen auf eigene, nicht validierte Verfahren zurückgreifen, was auch zur Folge haben kann, dass die Ergebnisse nicht vergleicht werden können und die Aussagekraft somit darunter leidet.

6.2 Übereinstimmung der Tests mit Berufserfolg

Die Übereinstimmung der Persönlichkeitstests mit dem Berufserfolg variiert stark. Viele Unternehmen machten keine Angaben, da diese Tests entweder nicht eingesetzt wurden oder keine ausreichenden Daten zur Einschätzung vorlagen. Neyer und Asendorpf (2024) unterstützen diese Interpretation und berichten, dass Persönlichkeitstests trotz ihres Mehrwerts nicht ausreichend angewendet oder präzise interpretiert werden. Ähnliche Ergebnisse zeigten sich bei den kognitiven Leistungstests: Die meisten Unternehmen machten auch hier keine Angaben, was auf eine geringe Nutzung hinweist. Einige Unternehmen gaben jedoch eine moderate Übereinstimmung an, was die Validität dieser Tests als Prädiktoren für Berufserfolg unterstreicht. Diese Ergebnisse unterstützen die Theorie von Schmidt und Hunter (1998), die die starke Prognosekraft von Intelligenztests hervorheben. Eine genauere Analyse zeigt, dass Persönlichkeitstests eine höhere Übereinstimmung mit dem Berufserfolg aufweisen als Intelligenztests. Dies könnte an der höheren Augenscheinvalidität der Persönlichkeitstests liegen, die ihre Akzeptanz und das Vertrauen in die Ergebnisse erhöht (Schmidt-Atzert & Amelang, 2012; Klug, o. D.). Die Frage der Übereinstimmung war subjektiv und basierte auf dem Bauchgefühl der Personalverantwortlichen, ohne klare Bewertungskriterien. Persönlichkeitstests scheinen leichter verständlich und direkter interpretierbar zu sein als Intelligenztests, die oft komplexere und abstraktere Items enthalten (Moosbrugger & Kelava, 2020). Im Hinblick auf die Möglichkeiten des Fakings weisen Moosbrugger und Kelava (2020) darauf hin, dass Persönlichkeitstests ein höheres Risiko der Verfälschung aufweisen. Entgegen den Erwartungen zeigte die Umfrage jedoch keinen verstärkten Einfluss sozial erwünschten Verhaltens. Es wurde angenommen, dass Persönlichkeitstests häufiger unnatürliches Verhalten fördern, was eine niedrigere Übereinstimmung mit dem späteren Berufserfolg zur Folge haben könnte. Diese theoretischen Überlegungen bestätigten sich jedoch nicht in den empirischen Befunden.

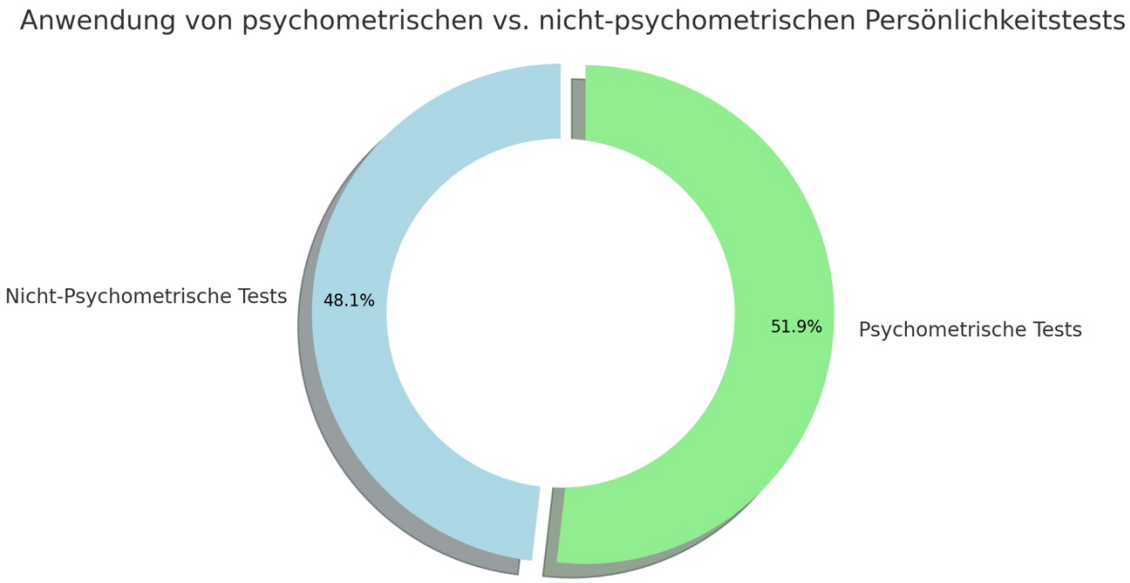
6.3 Bewertung der Qualität der vorhandenen Tests im Hinblick auf die Gütekriterien

Wie im Ergebnisteil beschrieben, wurden die genutzten Tests nach den Gütekriterien beurteilt. Dazu wurde die Abbildung 25 mit ihrer zugehörigen Legende und nach Farben von der Autorin selbst erstellt. Es ist auffällig, dass der DISC/DISG am häufigsten eingesetzt wurde, welcher jedoch über unzureichende und fehlende Hinweise auf solide Reliabilität und Validität verfügt. Dieser Test basiert nicht auf einem fundierten theoretischen Modell und erfüllt die psychometrischen Anforderungen nicht. Obwohl er den wissenschaftlichen Ansprüchen nicht genügt, könnte die häufige Anwendung mit dem Barnum-Effekt zusammenhängen, was bei Persönlichkeitstests nicht selten vorkommt (Hartmann, 2023). Ausserdem stimmt diese Beobachtung auch mit den theoretischen Erkenntnissen von Kanning (2022) überein, welcher aussagt, dass sich viele Unternehmen nicht an den Gütekriterien orientieren, sondern der Bekannt- und Beliebtheit von diversen Testverfahren mehr Gewicht geben und dementsprechend längst überholte Verfahren stets Anwendung finden. Folglich erklärt sich die Autorin den hohen Anwendungseinsatz des DISG/DISC mit den soeben genannten theoretischen Befunden. Wohingegen der zweite meistgenutzte Test der MPA darstellt, welcher mittel bis hohe Werte in den bewerteten Gütekriterien vorweist. Er bietet somit eine wissenschaftlich fundierte Basis und kann als psychometrischer Test bezeichnet werden. Diese Ergebnisse verhalten sich widersprüchlich zu den soeben diskutierten Befunden. Es ist jedoch gut möglich, dass in den Unternehmen, die den MPA einsetzen, Personalverantwortliche besser über die neuesten wissenschaftlichen Erkenntnisse informiert sind und deshalb validere Instrumente einsetzen. Letztendlich kann auch die Verfügbarkeit der einzelnen Verfahren und das verfügbare Budget Einfluss auf die Wahl der Tests haben und somit kommt es zu unterschiedlichen Ergebnissen. Bei den restlichen Testverfahren konnte beobachtet werden, dass der MBTI insgesamt fünfmal angewendet wird (zusammen mit der kostenlosen Variante (16 Personalities), welche auch auf dem MBTI basiert. Der MBTI weist unzureichende bis fehlende Angaben zur Validität und Reliabilität auf und die theoretische Grundlage ist umstritten. Trotz unerfüllter psychometrischer Standards finden diese Tests trotzdem Anwendung. Dies kann wieder mit den bereits oben diskutierten theoretischen Aussagen erklärt werden (Kanning, 2022; Hartmann, 2023). Der BIP und der FPI-R wurden beide je dreimal eingesetzt und der NEO-FFI nur einmal, wobei diese Tests deutlich höhere Werte in Reliabilität und Validität aufweisen. Besonders der geringe Einsatz vom NEO-FFI, welcher auf dem theoretisch fundierten Big-Five-Modell basiert und hohe Werte über alle Gütekriterien hinweg aufweist, bestätigt, dass valide und wissenschaftlich fundierte Tests weniger verbreitet sind. Wenn man die gesamten 27 Persönlichkeitstests anhand ihrer Psychometrik betrachtet, werden insgesamt 14 wissenschaftlich,

psychometrisch fundierte Testverfahren angewendet, während die restlichen 13 davon die Gütekriterien nur unzureichend bis gar nicht erfüllen. Dies wird im folgenden Kreisdiagramm (Abbildung 27) mit Prozentangaben veranschaulicht.

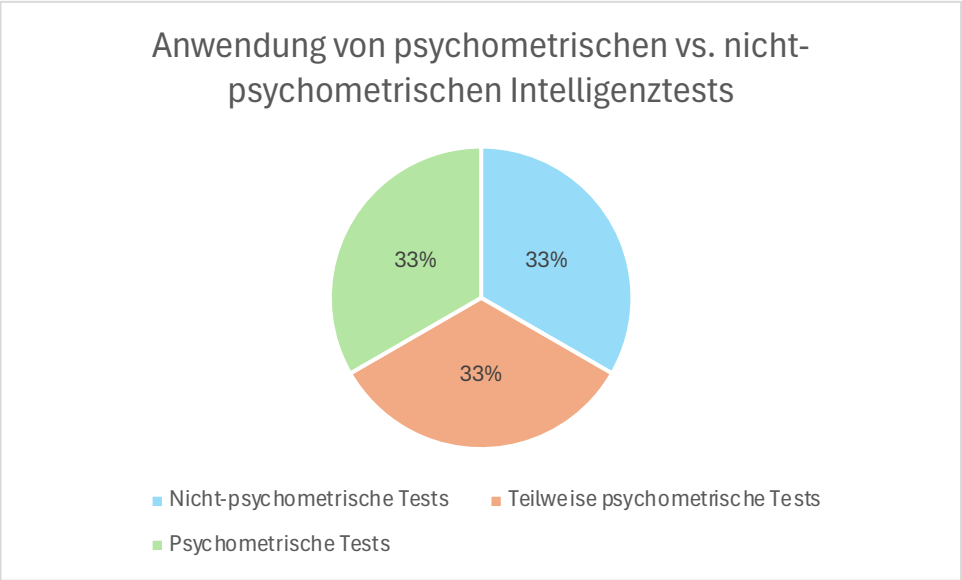
Abbildung 27

Anwendung von psychometrischen vs. nicht-psychometrischen Persönlichkeitstests



Dabei fällt auf, dass nur minim häufiger psychometrische Tests eingesetzt werden. Auch bei den Intelligenztests zeigt sich ein gemischtes Bild hinsichtlich der Gütekriterien. In der untenstehenden Abbildung 28 ist die Einordnung der psychometrischen kognitiven Leistungstests ersichtlich. Bei den Intelligenztests fiel die Bewertung etwas anspruchsvoller aus, da sich die einzelnen Intelligenztests nicht eindeutig in eine psychometrische oder klar nicht psychometrische Kategorie einteilen liessen. Deshalb wurde eine zusätzliche Kategorie «teilweise psychometrisch» gebildet, wobei gewisse Gütekriterien der Psychometrik erfüllt waren und einige nicht. Hierbei fällt auf, dass noch weniger psychometrische Intelligenztests eingesetzt werden als Persönlichkeitstests.

Abbildung 28
Anwendung von psychometrischen vs. nicht-psychometrischen Intelligenztests



Eine mögliche Erklärung dafür, könnte ebenfalls in der Augenscheinvalidität liegen, da es umgekehrt für Laien bei Intelligenztests schwerer abzuschätzen ist, inwiefern der Test nun die Intelligenz misst. Auch die Vermutung, dass Intelligenztests seitens Bewerbenden nicht viel Akzeptanz erfahren und dadurch zusätzliche Kosten entstehen, könnte ein hemmender Faktor für die geringe Nutzung sein (Hogrefe, 2022).

Abschliessend lässt sich in den Ergebnissen der Umfrage eine Übereinstimmung zu den theoretischen Erkenntnissen finden, wie beispielsweise, dass überholte Tests stets Anwendung finden oder sogar solche, deren Gütekriterien nie bestätigt worden sind (Kanning, 2022). Ausserdem bestätigen die Ergebnisse auch Befunde der Theorie, dass die die generelle Anwendung von Testverfahren nicht weit verbreitet ist (Funk et al., 2015) und auch weiterhin viele Vorurteile gegenüber psychometrischen Tests bestehen (Steininger, 2020). Dies signalisiert, dass noch erheblicher Sensibilisierungsbedarf für den Nutzen von Online-Tests und insbesondere deren Gütekriterien besteht.

7 Implikationen und mögliche Handlungsempfehlungen für den Praxispartner

7.1 Multimethodialität und die Rolle der Jörg Lienert AG

Die Jörg Lienert AG könnte eine zentrale Rolle bei der Sensibilisierung von Unternehmen für den Mehrwert psychometrischer Tests spielen. Da der geringe Einsatz der Tests durch die fehlende Expertise oder das Misstrauen der HR-Verantwortlichen erklärt werden kann, sollte dort angesetzt werden. Die Jörg Lienert AG könnte durch gezielte Workshops und Schulungen aufklären und die prognostische Validität dieser Tests erläutern. Auch Kampagnen, die auf eine mögliche Zusammenarbeit mit der Jörg Lienert AG abzielen und die Rolle der Multimethodialität kombiniert mit der Expertise der Expert:innen hervorhebt, könnte gewinnbringend sein. Insbesondere muss vermittelt werden, dass nicht Tests allein bevorzugt werden sollen, sondern weiterhin der trimodale Ansatz nach Schuler im Vordergrund stehen soll. Denn, obwohl Intelligenz von vielen Studien als bester Prädiktor angesehen wird, stellen mehrere Methoden zusammen kombiniert, die treffsicherste Wahl dar. Dabei sollte sichergestellt werden, dass diese von geschulten HR-Fachkräften durchgeführt, ausgewertet und interpretiert werden soll – in diesem Fall brilliert die Jörg Lienert AG. Die Befürwortung der Anwendung verschiedener Verfahren könnte somit propagiert werden, da auch theoretische Erkenntnisse dafür sprechen (Höft & Schuler, 2012). Viele Befunde aus der Literatur unterstützen die Multimethodialität auch, da verschiedene Verfahren zusammen kombiniert eine umfassendere und validere Einschätzung der Kandidat:innen ermöglichen (Sarges et al., 2009).

Obwohl die Kosten für die Implementierung und Nutzung psychometrischer Tests als Hürde wahrgenommen werden, sollte die langfristige Kosten-Nutzen-Analyse im Zusammenhang mit dem ROI in den Vordergrund gestellt werden. Fehlbesetzungen können beachtenswerte Kosten herbeiführen, welche durch die Anwendung validierter Testverfahren vermieden werden können. Es ist empfehlenswert, dass Unternehmen in eine gründlich und wissenschaftlich fundierte Personalauswahl investieren, anstatt nachfolgend hohe finanzielle Auslagen durch falsche Besetzungen zu tragen. Eine sorgfältige Kosten-Nutzen-Analyse sollte im Vorfeld stets erfolgen. Die Jörg Lienert AG könnte die Unternehmen auch dabei unterstützen, die Tests und die anderen Verfahren nicht nur durchzuführen, sondern auch richtig zu interpretieren. Dies ist ausgesprochen relevant, da viele der befragten Unternehmen ihre fehlende Expertise eingestanden. Da eine fehlerhafte Interpretation der Testresultate in falschen Entscheidungen resultieren kann, könnte die Durchführung und Bewertung bei Bedarf auch ausgelagert werden und im

Bereich der Testverfahren ein Outsourcing stattfinden. Damit könnte sichergestellt werden, dass wenn psychometrische Verfahren angewendet werden, solche auch richtig durchgeführt und eine korrekte Interpretation schlussgefolgert wird.

8 Fazit

Im folgenden Kapitel wird die zu Beginn gestellte Fragestellung erneut aufgegriffen und im Hinblick auf die Ergebnisse, wie auch den theoretischen Erkenntnissen beantwortet. Die vorliegende Bachelorarbeit untersucht die Verbreitung und Anwendung psychometrischer Online-Tests bei der Personalrekrutierung und -auswahl von Unternehmen in der Nordwestschweiz. Die Fragestellung lautete: «Wie weit verbreitet sind psychometrische Online-Tests in der Personalrekrutierung und Personalauswahl von Unternehmen in der Nordwestschweiz?»

Die Ergebnisse zeigen, dass diese Tests theoretisch zwar einen grossen Mehrwert bieten, in der Praxis jedoch selten eingesetzt werden. Von den 53 befragten Unternehmen nutzen 25 Online-Tests, während 26 Unternehmen keine Tests verwenden und zwei Unternehmen Tests in Papierform anwenden. Bei einer näheren Analyse der Psychometrik dieser Tests fällt auf, dass insgesamt 14 von den 27 Unternehmen (Papier-Bleistift-Tests wurden auch inkludiert) die Persönlichkeitstests anwenden, die psychometrischen Gütekriterien erfüllen. Die restlichen 13 Unternehmen nutzen Tests, welche unzureichende psychometrische Gütekriterien aufweisen. Bei den Intelligenztests fällt die Anwendung noch spärlicher aus. Nur sechs Unternehmen wandten die (zur Auswahl gestandenen) IQ-Tests an, wobei es zu bemerken gilt, dass nur ein Test von vier die psychometrischen Gütekriterien ganz erfüllt. Weiter wurde ein Test zweimal ausgewählt, welcher die psychometrischen Gütekriterien gar nicht erfüllt und zwei weitere Tests je einmal ausgewählt, welche die psychometrischen Gütekriterien teilweise erfüllen. Dementsprechend werden bei den Intelligenztests nur ein Drittel psychometrische Verfahren angewendet. Eine Limitation stellt jedoch das Auslassen der Bewertung von den übrigen Tests dar, weswegen übergreifend auch keine präzise Aussage zu allen in der Nordwestschweiz angewandten Tests bezüglich Psychometrik gemacht werden kann. Aus den vorliegenden Ergebnissen lässt sich jedoch ableiten, dass Sensibilisierungsbedarf bezüglich Testverfahren in der Personalauswahl vorhanden ist. Viele Personalverantwortliche äussern Bedenken bezüglich des Nutzens und sind nicht über die neuesten wissenschaftlichen Erkenntnisse informiert. Valide Verfahren wie psychometrische Tests werden unterschätzt, während weniger valide Verfahren wie unstrukturierte Interviews überschätzt werden. Aussagen aus der Umfrage bestätigen, dass Vorurteile gegenüber psychometrischen Online-Tests bestehen und andere Verfahren als ausreichend angesehen werden. Ausserdem wird angezweifelt, inwiefern solche Tests komplexe und individuelle Unterschiede adäquat erfassen können. Kosten könnten auch ein Grund des Verzichts auf Tests darstellen, da einige Personalverantwortliche äussern, dass sie geschulte Expert:innen benötigten, um Tests korrekt durchzuführen, auszuwerten

und zu interpretieren. Dieses Bewusstsein und der gegenwärtige Verzicht deuten auf mangelnde Ressourcen und Ausgaben hin, die Unternehmen zu sparen versuchen. Dies hat jedoch gravierende Konsequenzen, da Fehlbesetzungen, wie bereits theoretisch ausgeführt, erhebliche Kosten verursachen. Somit wird empfohlen, eine gründliche Kosten-Nutzen-Analyse durchzuführen und in eine Schulung der HR-Verantwortlichen und Lizenzen für psychometrische Testverfahren zu investieren. Dabei sollte der Fokus jedoch nicht auf einer einzelnen Methode liegen, sondern psychometrische Testverfahren sollten künftig zu bereits anderen Methoden hinzugezogen werden, um eine umfassendere und validere Einschätzung der verschiedenen Kandidat:innen zu erreichen. Denn viele theoretische Erkenntnisse stützen die Verwendung von kombinierten Verfahren. Dabei könnte die Jörg Lienert AG eine zentrale Rolle spielen, die die Unternehmen in diesen Belangen mit ihrem Expertenwissen beraten und unterstützen kann.

8 Limitationen und Ausblick

Im folgenden Kapitel werden die wichtigsten Limitationen der vorliegenden Arbeit und zukünftige Forschungsmöglichkeiten zusammengefasst.

Die Limitationen dieser Studie sind vielfältig und betreffen mehrere Aspekte. Die Stichprobengrösse von N=53 ist eher gering und die Rücklaufquote der befragten Unternehmen auch nicht gross. Dies kann teilweise auch darauf zurückgeführt werden, dass zeitliche Ressourcen zu knapp waren und daher auch auf einen Reminder verzichtet wurde. Die Anonymisierung ermöglichte transparente Antworten, jedoch erschwerte sie das gezielte Versenden einzelner Reminder. Weiter besteht eine Überrepräsentation von Unternehmen aus der MEM- und Chemie-Pharma-Branche. Dies könnte auf eine gewisse Selbstselektion zurückführen, da die Unternehmen letztendlich selber entscheiden konnten, ob sie an der Umfrage teilnehmen wollen. Die Aktualität der Liste stellte eine zusätzliche Herausforderung dar, da viele Informationen veraltet waren, vermehrt nachgefragt werden musste und somit auch nicht die ursprünglich geplante Mitarbeitendenanzahl (>100 MA) eingehalten werden konnte. Diese Unternehmen aus der Stichprobe zu streichen, hätte jedoch zur Folge gehabt, dass der Datensatz noch kleiner wäre und hätte daher aussagekräftige und spannende Erkenntnisse vorenthalten, weshalb alle Unternehmensgrössen beibehalten wurden. Die ungleiche Verteilung der Unternehmensgrössen im Datensatz könnte jedoch eine Beeinflussung der Ergebnisse zur Folge haben. Eine umfassendere und ausgewogenere Stichprobe (gleiche Repräsentation aller Branchen und Unternehmensgrössen) könnte zu präziseren Ergebnissen führen. Der Einfluss auf die Testanwendung durch Faktoren wie Branche, spezifische Unternehmensanforderungen, Unternehmensgrösse wie auch die Verfügbarkeit von Ressourcen wäre spannend in zukünftigen Forschungen näher zu beleuchten. Ausserdem wurde Hogrefe als Literaturquelle häufig genutzt, was die Objektivität der theoretischen Erkenntnisse beeinflussen könnte, da Hogrefe selbst auch ein Anbieter ist. Zusätzlich enthielt der Fragebogen qualitative wie auch quantitative Aspekte, was grundsätzlich positiv ist. Allerdings basierten Items auf subjektiven Einschätzungen, wie beispielsweise die Übereinstimmung der Tests mit dem Berufserfolg. Diese wurden nicht sauber konstruiert, da sie rein aus dem Bauchgefühl der Personalverantwortlichen beantwortet werden mussten und ohne messbare, vorliegende Daten eine derartige Einschätzung wissenschaftlich nicht präzise ist. Es wurde ausserdem auch nicht festgelegt, nach welchen Kriterien die Personalverantwortlichen die Übereinstimmung beurteilen sollen, was ebenfalls ungenau ist. Für eine präzisere Aussagen müssten die individuellen Testresultate der Mitarbeitenden mit dem nachfolgenden Berufserfolg (z.B. Verkaufszahlen oder Dienstjahren) verglichen werden. Dies war im Rahmen dieser Umfrage jedoch nicht

möglich, weshalb eine subjektive Einschätzung eingeholt wurde. Ausserdem wurden die zusätzlich genannten Tests der Unternehmen nicht auf ihre Gütekriterien untersucht, was die Genauigkeit zur exakten Verbreitung psychometrischer Tests in der Nordwestschweiz beeinträchtigen könnte. Aufgrund des begrenzten Rahmens der Arbeit musste eine Auswahl getroffen werden und der Fokus wurde auf die bereits recherchierten Testverfahren gelegt. Ausserdem wurden generell mehr Persönlichkeitstests als Intelligenztests angewendet, was die Aussagekraft der Übereinstimmung und die Übertragbarkeit der Ergebnisse insgesamt erschwert hat. Die unterschiedliche Anwendungshäufigkeit könnte zu einer Verzerrung der Ergebnisse führen und die Vergleichbarkeit und Generalisierung der Ergebnisse auf andere Kontexte beeinträchtigen. Auch wenn kein Test zur exakten Bestimmung der Signifikanz durchgeführt wurde, könnte die Tendenz zur höheren Testnutzung in grösseren Unternehmen in einer weiteren Forschung näher untersucht werden. Die Unternehmensgrösse könnte bedeutend sein und grössere Unternehmen könnten aufgrund eher vorhandener Ressourcen von der Einführung bestimmter Tests profitieren, die ihre umfangreicheren Auswahlverfahren unterstützen.

9 Literatur- und Abbildungsverzeichnis

9.1 Literaturverzeichnis

Adeosun, O. T. & Ohiani, A. S. (2020). Attracting and recruiting quality talent: firm perspectives. *Rajagiri Management Journal*, *14*(2), S. 107–120. https://doi.org/10.1108/ramj-05-2020-0016

Afshari, A. R., Nikolic, M. & Cockalo, D. (2014). Applications of fuzzy decision making for personnel selection problem: A review. *Journal Of Engineering Management And Competitiveness*, *4*(2), S. 68–77. http://www.tfzr.uns.ac.rs/jemc/files/Vol4No2/V4N22014-02.pdf

American Psychological Association. (2023). *Intelligence test*. https://dictionary.apa.org/intelligence-test

Bartscher, T., Stöckl, J. & Träger, T. (2012. *Personalmanagement. Grundlagen, Handlungsfelder*, *Praxis*. Pearson.

berufsberatung.ch. (o. D.). *Übersicht - Berufe*. SDBB/CSFO. https://www.berufsberatung.ch/dyn/show/1470

Bühner, M. (2021). *Einführung in die Test- und Fragebogenkonstruktion*. Pearson.

D., Urhahne, M., Dresel & F., Fischer (Hrsg.), *Psychologie für den Lehrberuf* (S. 471–491*).* Springer eBooks. https://doi.org/10.1007/978-3-662-55754-9_24

De Cooman, R. & Vleugels, W. (2022). *Person–Environment Fit: Theoretical Perspectives, Conceptualizations, and Outcomes*. Oxford Research Encyclopedia. https://doi.org/10.1093/acrefore/9780190224851.013.377

Dhliwayo, P. & Coetzee, M. (2020). Cognitive intelligence, emotional intelligence and personality types as predictors of job performance: Exploring a model for personnel selection. *SA Journal Of Human Resource Management*, *18(0), S. 1-13*. https://doi.org/10.4102/ sajhrm. v18i0.1348

Döring, N. & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5. Auflage). Springer-Lehrbuch. https://doi.org/10.1007/978-3-642-41089-5

Dorsch Lexikon der Psychologie. (o. D.). *Freiburger Persönlichkeitsinventar*. https://dorsch.hogrefe.com/stichwort/freiburger-persoenlichkeitsinventar-fpi-r

Ekwoaba, O., Ikeije, U. U. & Ufoma, N. (2015). *The Impact of Recruitment and selection criteria on organizational performance*. *Global Journal Of Human Resource Management, 3*(2), S. 22-33. https://www.eajournals.org/wp-content/uploads/The-Impact-of-Recruitment-and-Selection-Criteria-on-Organizational-Performance.pdf

Fowler, F. J., Jr. (2013). *Survey research methods*. SAGE Publications.

Funk, L., Nachtwei, J. & Melchers, K. (2015). Die Kluft zwische Wissenschaft und Praxis in der Personalauswahl. *PERSONALquarterly, 03(15),* S. 26-38. https://www.researchgate.net/profile/Jens-Nachtwei/publication/288642308_Die_Kluft_zwischen_Wissenschaft_und_Praxis_in_der_Personalauswahl/links/5688ca8308ae051f9af71006/Die-Kluft-zwischen-Wissenschaft-und-Praxis-in-der-Personalauswahl.pdf

Hartmann, C. (2023, 8. März). *Persönlichkeitstypen*. Psychologie Heute. https://www.psychologie-heute.de/gesellschaft/artikel-detailansicht/42520-persoenlichkeitstypen.html

Heimsch, F. M., Niederer, R. & Zöfel, P. (2018). *Statistik im Klartext: Für Psychologen, Wirtschafts- und Sozialwissenschaftler* (2. Auflage). Pearson.

Hock, M., Peters, J., Renner, K. & Krohne, H. W. (2023). Merkmale und Gütekriterien psychologischer Tests. In M. Hasselhorn, W. Kunde & S. Schneider (Hrsg.), *Psychologische Diagnostik: Grundlagen und Anwendungsfelder* (3. Auflage, S. 31-104). Pearson. https://doi.org/10.17433/978-3-17-042007-6

Höft, S. & Schuler, H. (2014). Personalmarketing und Personalauswahl. In H. Schuler & K. Moser (Hrsg.), *Lehrbuch Organisationspsychologie* (5. Auflage, S. 55-126). Hogrefe. https://doi.org/10.1026/0932-4089/a000162

Hogrefe. (2022, 19. April). Eignungsdiagnostik mit Intelligenztests im HR-Bereich. *Hogrefe*. https://www.hogrefe.com/ch/thema/eignungsdiagnostik-mit-intelligenztests-im-hr-bereich-1

Hogrefe. (2020, 13. Oktober). Einsatz psychologischer Testverfahren in der Personalauswahl. *Hogrefe.* https://www.hogrefe.com/ch/thema/einsatz-psychologischer-testverfahren-in-der-personalauswahl

Hossiep, R., Paschen, M., & Mühlhaus, O. (2000). *Persönlichkeitstests im Personalmanagement: Grundlagen, Instrumente und Anwendungen*. Verlag für Angewandte Psychologie.

Hough, L. M., Oswald, F. L., & Ock, J. (2015). Beyond the Big Five: New Directions for Personal-ity Research and Practice in Organizations. *Annual Review of Organizational Psychology and Organizational Behavior, 2*(1), S. 183–209. https://doi.org/10.1146/annurev-orgpsych-032414-111441

Jansen, L. J., Diercks, J. & Kupka, K. (2023). *Recrutainment*. Springer eBooks. https://doi.org/10.1007/ 978-3-658-38749-5_1

Jörg Lienert AG. (o. D.). *Human Values Search & Diagnostik*. *https://joerg-lienert.ch/*

Kanning, U. P. & Ohlms, M. L. (2024). *Digitale Personalauswahl und Eignungsdiagnostik*. Springer.

Kanning, U. P. (2019, 3. April). *Warum Manager intelligent sein sollten*. Haufe.de News Und Fachwissen. https://www.haufe.de/personal/hr-management/auswahl-von-managern-mit-intelligenztest_80_487472.html?page=25

Kanning, U. P. (2022, 11. Februar). *Woran erkennt man einen guten Test für die Personaldiagnostik?* Haufe.de News und Fachwissen. https://www.haufe.de/personal/hr-management/psychologische-testverfahren-in-der-personalauswahl/test-fuer-personaldiagnostik-auswaehlen_80_494818.html

Klug, A. (o. D.). *Augenscheinvalidität*. https://www.klugmd.de/Wissen/Augenscheinvaliditaet.htm

Knorr, A. (2021, 26. April). *Alles was du über Personalbeschaffung und Personalauswahl wissen musst*. Human Resources and more. https://blog.kenjo.io/de/personalbeschaffung-personalauswahl

Krüger, K. (2018). Herausforderung Fachkräftemangel. https://doi.org/10.1007/978-3-658-20421-1

Kuckartz, U., Rädiker, S., Ebert, T. & Schehl, J. (2010). *Statistik: Eine verständliche Einführung*. VS Verlag und Springer Fachmedien. https://doi.org/10.1007/978-3-531-92033-7_9

Lienert, G. A., & Raatz, U. (1998). *Testaufbau und Testanalyse* (6. Auflage). Beltz Verlagsgruppe. https://doi.org/10.9783621278454

Maier, G. W. (o.D.). *Eignungsdiagnostik*. Gabler Wirtschaftslexikon. https://wirtschaftslexikon.gabler.de/definition/eignungsdiagnostik-35877

marktforschung.de. (o. D.). *Desk Research*. https://www.marktforschung.de/wiki-lexikon/marktforschung/Desk%20Research/

Molotkova, N., Makeeva, M. & Khazanova, D. (2019). *Digitalized personnel management*: *The European Proceedings of Social & Behavioural Sciences*. https://doi.org/10.15405/epsbs.2019.03.75

Moosbrugger, H. & Kelava, A. (2020). *Testtheorie und Fragebogenkonstruktion*. Springer.

Müllerschön, A. (2006). Persönlichkeitstests auf dem Prüfstand. *HR-Today*, November 2006. https://www.muellerschoen-beratung.de/fileadmin/pdf/PersoenlichkeitstestsPruefstand.pdf

Nerdinger, F. W., Blickle, G. & Schaper, N. (2019). *Arbeits- und Organisationspsychologie* (4. Auflage). Springer-Lehrbuch. https://doi.org/10.1007/978-3-662-56666-4

Neyer, F. J. & Asendorpf, J. B. (2024). *Psychologie der Persönlichkeit* (7. Auflage). Springer eBooks. https://doi.org/10.1007/978-3-662-67385-0

Nikolaou, I. (2021). What is the Role of Technology in Recruitment and Selection? *˜The Spanish Journal Of Psychology*, *24*(2), S. 1-6. https://doi.org/10.1017/sjp.2021.6

Rost, D. H. & Häcker, H. O. (2022). Intelligenzfaktoren. Hogrefe AG - Dorsch - Lexikon der Psychologie. https://dorsch. hogrefe.com/stichwort/intelligenzfaktoren

Rost, J. (2004). *Lehrbuch Testtheorie – Testkonstruktion* (2. Auflage). Verlag Hans Huber.

Sackett, P. R., Zhang, C., Berry, C. M. & Lievens, F. (2022). Revisiting meta-analytic estimates of validity in personnel selection: Addressing systematic overcorrection for restriction of range. *Journal Of Applied Psychology*, *107*(11), S. 2040–2068. https://doi.org/10.1037/apl0000994

Salgado, J. F. (2017a). Using ability tests in selection. In H. W. Goldstein, E. D. Pulakos, J. Passmore, & C. Semedo (Hrsg.), *The Wiley Blackwell handbook of the psychology of recruitment, selection and employee retention* (S. 115–150). Wiley Blackwell. https://doi.org/10.1002/9781118972472.ch7

Salgado, J. F. (2017b). *Personnel selection*. Oxford Research Encyclopedia of Psychology. https://doi.org/10.1093/acrefore/9780190236557.013.8

Sangeetha, K. (2010). Effective Recruitment: A Framework. *IUP Journal of Business Strategy 7*(1&2) S. 93-107. https://openurl.ebsco.com/EPDB%3Agcd%3A16%3A

1120080/detailv2?sid=ebsco%3Aplink%3Ascholar&id=ebsco%3Agcd%3A51301171&crl=c

Sarges, W., Wottawa, H. & Roos, C. (2009). *Handbuch wirtschaftspsychologischer Testverfahren, Band II: Organisationspsychologische Instrumente* (2. Auflage). Pabst.

Schellinger, J., Goedermans, M., Kolb, L. M. & Sebai Y. (2020) Digitale Transformation und Human Resource Management. In J. Schellinger, K. O. Tokarski & I. Kissling-Näf (Hrsg.), *Digitale Transformation und Unternehmensführung: Trends und Perspektiven für die Praxis*. (S. 183-222). Springer Gabler. https://doi.org/10.1007/978-3-658-26960-9

Schmidt, F. L. & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*(2), S. 262–274. https://doi.org/10.1037/0033-2909.124.2.262

Schmidt-Atzert, L. & Amelang, M. (2012). *Psychologische Diagnostik (Lehrbuch mit Online-Materialien)*. Springer Science & Business Media. https://doi.org/10.1007/978-3-642-17001-0

Schuler, H. (2016). *Personalpsychologie, trimodaler Ansatz*. Hogrefe AG - Dorsch - Lexikon der Psychologie. https://dorsch.hogrefe.com/stichwort/personalpsychologie-trimodaler-ansatz

Schwaighofer, M., Heene, M. & Bühner, M. (2019). Grundlagen und Kriterien der Diagnostik. In D. Urhahne, M. Dresel & F. Fischer (Hrsg.), *Psychologie für den Lehrberuf*. (S. 471–491). Springer. https://doi.org/10.1007/978-3-662-55754-9_24

St. Gallen, Universität. (2023, 1. März). *«Trend Barometer: People Management 2030»*. Universität St.Gallen. https://www.unisg.ch/de/newsdetail/news/trend-barometer-people-management-2030/

Steiner, H. (2009). *Online-Assessment: Grundlagen und Anwendung von Online-Tests in der Unternehmenspraxis*. Springer Science & Business Media.

Steininger, T. (2020). *Der Einsatz psychologischer Testverfahren in Unternehmen*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-28462-6

Thommen, J., Achleitner, A., Gilbert, D. U., Hachmeister, D. & Kaiser, G. (2017). *Allgemeine Betriebswirtschaftslehre*. Springer eBooks. https://doi.org/10.1007/978-3-658-07768-6

Weinert, S. & Stulle, K. P. (2015). *Executive assessment: Instrumente, Trends, Herausforderungen*. Springer Gabler.

Wirtschaftsförderung Luzern. (o. D.). *Jörg Lienert AG*. https://www.lucerne-business.com/de/unternehmen/joerg-lienert-ag-1304

Witte, K. (2019). *Angewandte Statistik in der Bewegungswissenschaft* (3. Auflage). Springer eBooks. https://doi.org/10.1007/978-3-662-58360-9

Wöhe, G. & Döring, U. (2013). *Einführung in die Allgemeine Betriebswirtschaftslehre* (25. Auflage). Vahlen.

9.2 Abbildungsverzeichnis

Abbildung 1 Lienert-Loop Tool-Box der Jörg Lienert AG 10

Abbildung 2 Übersicht der Haupt- und Nebengütekriterien. Adaptiert von Einführung in die Test- und Fragebogenkonstruktion (s.44) von Bühner, 2021, Pearson. 15

Abbildung 3 Darstellung des Trimodalen Ansatzes nach Schuler. Adaptiert von Der trimodale Ansatz der Eignungsdiagnostik (S. 103) von Schuler und Höft (2006), Hogrefe. 21

Abbildung 4 Beschreibung und Abfolge des Forschungsprozesses (eigene Darstellung) 31

Abbildung 5 Datenbereinigungsprozess (eigene Darstellung) 35

Abbildung 6 Eigene Einteilung der 13 Branchen 38

Abbildung 7 Häufigkeiten der Teilnahme der Unternehmen nach Branche 39

Abbildung 8 Stichprobenbeschreibung anhand Mitarbeitendenanzahl 40

Abbildung 9 Testanwendung unterschiedliche Positionen 40

Abbildung 10 Diagramm Testanwendungen nach Branche und Form 41

Abbildung 11 Diagramm Anzahl Mitarbeitende in verschiedenen Branchen 42

Abbildung 12 Diagramm Angewendete Intelligenztests mit zugehörigen Branchen 43

Abbildung 13 Nicht angewendete Intelligenztests 43

Abbildung 14 Angewendete Persönlichkeitstests mit zugehörigen Branchen 44

Abbildung 15 Nicht angewendete Persönlichkeitstests 45

Abbildung 16 Kategorisierung der Gründe für Verzicht auf IQ-Tests 45

Abbildung 17 Kategorisierung Gründe Verzicht auf Persönlichkeitstests 45

Abbildung 18 Häufigkeit von Eigenentwicklung (IQ-Tests) 46

Abbildung 19 Häufigkeit von Eigenentwicklung (Persönlichkeitstests) 46

Abbildung 20 Anwendung von anderen Testverfahren (Persönlichkeit und Intelligenz) 47

Abbildung 21 Diagramm Kosten der Tests 48

Abbildung 22 Allgemeines Feedback zu den Tests 49

Abbildung 23 Übereinstimmung der Persönlichkeitstests mit Berufserfolg 49

Abbildung 24 Übereinstimmung der kognitiven Leistungstests mit Berufserfolg 50

Abbildung 25 Bewertung der Qualität der eingesetzten Testverfahren 51

Abbildung 26 Legende der Kennwerte (Beurteilung der Gütekriterien) 52

Abbildung 27 Anwendung von psychometrischen vs. nicht-psychometrischen Persönlichkeitstests 58

Abbildung 28 Anwendung von psychometrischen vs. nicht-psychometrischen Intelligenztests 59