



Fachhochschule
Nordwestschweiz

Fachhochschule Nordwestschweiz

Studiengang Angewandte Psychologie

Bachelorthesis 2016



**Beratungsdienste für Ausbildung
und Beruf Aargau**

Einsatz von eignungsdiagnostischen Instrumenten in der Berufs-, Studien- und Laufbahnberatung

Autor: Marco Schraner

Ort, Datum: Olten, Juni 2015

Begleitung: lic. phil. Joannis Avramakis

Praxispartner: ask! Aarau

Abstract

Ziel der vorliegenden Arbeit war es, aufzuzeigen wie in der Berufs-, Studien- und Laufbahnberatung der ask!, eignungsdiagnostische Instrumente zur Urteilsbildung eingesetzt werden. Dabei wurden der Nutzen, aber auch die Grenzen von eignungsdiagnostischen Instrumenten aufgezeigt, sowie die Schwierigkeiten beleuchtet, die im Arbeitsalltag entstehen können. Basierend auf theoretischen Erkenntnissen und langjähriger Praxiserfahrung, wurde in einem Experteninterview ein online-Fragebogen erstellt, der von insgesamt 24 Beraterinnen und Beratern ausgefüllt wurde. Die Auswertung der Befragung wurde qualitativ durch Kategorienbildung vorgenommen und durch quantitative Ergebnisse untermauert. Dieses Vorgehen resultierte in besser abgesicherten Ergebnissen, was wiederum zuverlässigere Aussagen über die Praxissituation zuließ.

Die Bachelorarbeit umfasst 83'898 Zeichen inklusive Leerzeichen und ohne Anhang.

Inhaltsverzeichnis

Abstract	I
Inhaltsverzeichnis	II
Selbständigkeitserklärung	IV
1 Einleitung	1
1.1 Ausgangslage	2
1.2 Zielsetzung	2
2 Theoretische Grundlagen	3
2.1 Die Klassische Testtheorie (KTT)	3
2.2 Qualitätsanforderungen an einen psychologischen Test	5
2.2.1 Objektivität	5
2.2.2. Reliabilität	6
2.2.3 Validität	8
2.2.4 Normierung (Eichung)	11
2.3 Psychologische Diagnostik	12
2.3.1 Definition psychologischer Diagnostik	12
2.3.2 Definition: Eignung und Eignungsdiagnostik	12
2.3.3 Drei Ansätze der Eignungsdiagnostik	13
2.4 Übersicht einiger Testverfahren und deren Validität	14
2.4.1 Intelligenz und Beruf	14
2.4.2 Kritik an der Validität zu eignungsdiagnostischen Verfahren	18
2.4.3 Allgemeine Persönlichkeitskonstrukte und spezifische Persönlichkeitseigenschaften	18
2.4.4 Motivation und Interessen	20
2.4.5 Problemfelder eigenschaftsorientierter Verfahren	23
2.4.6 Interview	24
2.5 Klinische vs. statistische Urteilsbildung	25
3 Methodik	29
3.1 Auswertung der Befragung	31
4 Ergebnisse	32
4.1 Interpretation der Ergebnisse (deduktiv)	32
4.2 Interpretation induktive Kategorien	35
4.3 Zusammenfassungen der aussagekräftigsten und somit wichtigsten Kategorien	37

5 Diskussion	40
5.1 Kritische Reflexion des methodischen Vorgehens	40
5.2 Fazit und Ausblick.....	41
5.3 Handlungsempfehlungen	42
7 Literaturverzeichnis	43
8 Tabellenverzeichnis	45
9 Abbildungsverzeichnis.....	45
10 Anhang.....	46

Selbständigkeitserklärung

Hiermit erkläre ich, dass die vorliegende Bachelor-Thesis selbständig und ausschliesslich mit den angegebenen Quellen, Hilfsmitteln und Hilfeleistungen entstanden ist und Zitate kenntlich gemacht wurden.

Aarau, 7. Juni 2016

Marco Schraner

1 Einleitung

Die ask! – Beratungsdienste für Ausbildung und Beruf Aargau sind ein Kompetenz- und Servicezentrum welche für Menschen von 14 bis 65 Jahren sowie Organisationen, Beratungsdienste unkompliziert anbietet. Zu den Angeboten gehören Berufs-, Studien- und Laufbahnberatung, Jugendpsychologischer Dienst, Lehrpersonenberatung, Dienstleistungen wie Mentoring, Coaching sowie Vermittlung von Wissen an Fachpersonen. <http://www.beratungsdienste-aargau.ch/mehr-zu-ask/portrait.html>

Mit der vorliegenden Arbeit wurde ermittelt, nach welchen Kriterien eignungsdiagnostische Instrumente in der Berufs-, Studien- und Laufbahnberatung verwendet und zur diagnostischen Urteilsbildung eingesetzt werden. Dabei wurden Besonderheiten von Tests erfragt, Themenbereiche ausserhalb der typischen Kernthematik von Intelligenz-, Persönlichkeits-, Leistungs- und Interessentest berücksichtigt und nach thematischen Lücken, für die keine Tests zur Verfügung stehen gefragt. Damit soll aufgezeigt werden, wie mit eignungsdiagnostischen Instrumenten umgegangen wird und welchen Nutzen sie für die Beraterinnen und Berater in der Arbeitspraxis erbringen. Besonderes Augenmerk galt unter anderem, den Auswahlkriterien die zur Verwendung oder nicht-Verwendung der jeweiligen Tests führte.

Die aktuelle Vorgehensweise der Berufs-, Studien- und Laufbahnberaterinnen und Beratern soll aufzeigen, wie und aus welchen Gründen eignungsdiagnostische Instrumente verwendet werden. Grosser Wert wird auf den Vergleich von theoretischen Grundlagen und den in der Praxis angewandten Methoden gelegt. Besondere Aufmerksamkeit erhält die Kontroverse klinische vs. statistische Urteilsbildung. Gemäss Kuncel, Klieger, Connelly und Ones (2013) besteht in der psychologischen Praxis eine deutliche Präferenz für klinisches gegenüber statistischem Vorgehen bei der diagnostischen Urteilsbildung. Damit verbunden ist eine skeptische Haltung gegenüber rechnerischer Testverfahren. Ob dies auch auf die vorliegende Stichprobe zutrifft und was dies für Konsequenzen für die jeweiligen Beratungssituationen haben kann, wird aufgezeigt.

1.1 Ausgangslage

Wenn von psychologischer Diagnostik die Rede ist, stellt sich immer wieder die Frage welche Verfahren am besten eingesetzt und kombiniert werden können, um zu einem adäquaten Ergebnis zu gelangen. Obwohl sich viele Untersuchungen nicht explizit auf die Berufs-, Studien- und Laufbahnberatung fokussieren, so können doch viele Erkenntnisse darauf übertragen werden. So ist zum Beispiel seit langer Zeit bekannt, dass die in der Theorie empfohlenen, oftmals standardisierten, mechanischen Vorgehensweisen, in der Praxis zu Gunsten von kasuistischen oder holistischen vernachlässigt werden. Zusätzlich zu dem Spannungsfeld zwischen theoretischen Erkenntnissen und praktikabler Umsetzung, in welchem sich die Beraterinnen und Berater zwangsläufig befinden, kommt eine Umstrukturierung der ask! hinzu. Seit Januar 2016 sind Beratungen für Personen welche eine Berufsausbildung oder eine Matura abgeschlossen und das fünfundzwanzigste Lebensjahr erreicht haben nicht mehr kostenlos. Dies kann auch Berater und Beraterinnen der ask! vor neue Herausforderungen stellen und ist ein Indikator dafür, dass Ökonomie einen grösseren Stellenwert in der Beratung einnimmt.

1.2 Zielsetzung

Ziel der Arbeit ist es, zu ermitteln, wie in der Berufs-, Studien- und Laufbahnberatung bestehende Verfahren eingesetzt werden, um im Sinne einer klinischen oder statistischen Urteilsbildung Informationen zu gewinnen. Dabei werden Schwierigkeiten, welche sich eventuell erst im praktischen Arbeitsalltag zeigen genauso angesprochen, wie die Vorteile die eignungsdiagnostische Instrumente mit sich bringen. Als Fazit soll eine pro und contra Diskussion im Sinne von klinischer vs. statistischer Urteilsbildung oder praktisches vs. theoretisches Vorgehen geführt werden. Gleichzeitig sollen auch Schwächen von jeweiligen Verfahren beschrieben und somit einen optimalen Einsatz von eignungsdiagnostischen Instrumenten ermöglicht werden.

2 Theoretische Grundlagen

2.1 Die Klassische Testtheorie (KTT)

Abgesehen von wenigen Ausnahmen sind die heute gebräuchlichen Tests nach den Regeln der Klassischen Testtheorie (KTT) konzipiert worden. Die KTT liefert eine theoretische Begründung der Reliabilität (Messgenauigkeit) eines Tests. Dabei ist eine grundlegende Annahme, dass Testwerte aus Persönlichkeitsfragebögen, Intelligenztests, Konzentrationstests usw. fehlerbehaftet sind. Wenn beispielsweise jemand bei einem Intelligenztest einen IQ Wert von 125 erreicht so ist dies nur der beobachtete Wert, der wahre Wert kann höher oder auch niedriger sein. Diese Abweichung kommt durch Messfehler zustande. Wenn der Intelligenztest wiederholt wird, ohne dass Erinnerungs- und Übungseffekte auftreten und der beobachtete IQ ist nun 131, muss der Messfehler jedes Mal unterschiedlich gross gewesen sein, da sich die Intelligenz der Person nicht verändert. Das heisst, ihr wahrer Wert ist gleich geblieben. Schmidt-Atzert & Amelang (2012).

Schmidt-Atzert & Amelang (2012) liefern für die Grundannahmen der KTT folgende Umschreibung.

Die KTT beginnt mit einigen wenigen Grundannahmen (Axiomen). Diese werden a priori angenommen und nicht etwa empirisch durch Untersuchungen begründet.

Sie stellen die Grundlage für mathematische Ableitungen dar, die schliesslich zu Formeln führen, mit denen wir beispielsweise die Messgenauigkeit eines Tests berechnen. (S. 11)

Moosbrugger & Kelava (2012) umschreiben die Axiome der KTT wie folgt:

1. Existenzaxiom

Der Erwartungswert (Mittelwert bei unendlich vielen Messwiederholungen) des Messfehlers ist 0. Folglich entspricht der Erwartungswert des beobachteten Werts dem wahren Wert.

2. Verknüpfungaxiom

Jeder beobachtete Wert X setzt sich aus dem wahren Wert T (truescore) und einem unsystematischen Messfehler E (Error) zusammen.

3. Unabhängigkeitsaxiom

Es gibt keinen Zusammenhang zwischen Messfehlern und wahren Werten oder zwischen den Messfehlern zweier verschiedener Messungen.

Schmidt-Atzert & Amelang (2012) schliessen daraus, dass das Ergebnis in einem Test also der beobachtete Wert nicht als absolut genaue Messung angesehen werden darf. Für jede Person existiert in einem Test ein wahrer Wert, dieser ist zumindest über einen bestimmten Zeitraum konstant. Er könnte theoretisch ermittelt werden, indem man den Test unendlich oft durchführt und dabei sicherstellt, dass keine Erinnerungs- und Übungseffekte auftreten.

Der Mittelwert oder Erwartungswert aller Messergebnisse (also aller beobachteter Werte) wäre dann der wahre Wert.

Weiter leiten Schmidt-Atzert & Amelang (2012) aus den Grundannahmen der KTT die Reliabilität ab. Aus den oben aufgeführten Axiomen lässt sich mathematisch eine Aussage über die Reliabilität eines Tests ableiten. Diese Formel stellt die wichtigste Ableitung aus den Annahmen der klassischen Testtheorie dar:

$$Rel = \frac{Var(T)}{Var(X)}$$

Definition: Die Reliabilität eines Tests ist der Anteil der Varianz der wahren Werte (T) an der Varianz der beobachteten Werte (X).

Ein Reliabilitätskoeffizient von .80 bedeutet demzufolge, dass die beobachtete Varianz der Testwerte zu 80 Prozent auf Unterschiede zwischen den wahren Werten der Testpersonen zurückzuführen ist und zu 20 Prozent auf Fehlervarianz beruht. Die Formel hilft, die inhaltliche Bedeutung der Reliabilität zu begreifen.

Noch können wir die Reliabilität eines Tests nicht empirisch bestimmen. Wir kennen immer nur die Varianz der beobachteten Werte, wissen also, wie stark die vorliegenden Messwerte streuen. Wahre Werte und deren Varianz entziehen sich der Beobachtung (Schmidt-Atzert & Amelang, 2012).

2.2 Qualitätsanforderungen an einen psychologischen Test

Moosbrugger & Kelava (2012) klären auf:

Wenn man mit der Frage konfrontiert wird, worin der eigentliche Unterschied zwischen einem unwissenschaftlichen Test (etwa einer Fragensammlung) und einem wissenschaftlich fundierten, psychologischen Test besteht, so ist die Antwort darin zu sehen, dass sich ein psychologischer Test dadurch unterscheidet, dass er hinsichtlich der Erfüllung der sog. Testgütekriterien empirisch überprüft wurde. (S. 23)

Die Testgütekriterien stellen also ein Instrument zur Qualitätsbeurteilung von psychologischen Tests dar. Das Testmanual eines vorliegenden Tests sollte in geeigneter Weise darüber informieren, welche Testgütekriterien in welcher Weise erfüllt sind. Als Gütekriterien hat sich in den vergangenen Jahren eine Reihe von Aspekten etabliert, die nicht zuletzt auch die Basis zur berufsbezogenen Eignungsbeurteilung bilden. Die wichtigsten werden hier aufgelistet (Moosbrugger & Kelava, 2012).

2.2.1 Objektivität

Moosbrugger & Kelava (2012) definieren Objektivität folgendermassen: „Ein Test ist dann objektiv, wenn er dasjenige Merkmal, das er misst, unabhängig von Testleiter und Testauswerter misst. Ausserdem müssen klare und anwenderunabhängige Regeln für die Ergebnisinterpretation vorliegen“. (S. 23)

Objektivität bedeutet also, dass den Testdurchführenden kein Verhaltensspielraum bei der Durchführung, Auswertung oder Interpretation eingeräumt wird. Somit wäre völlige Objektivität, wenn sowohl jeder beliebige Testleiter, der einen bestimmten Test mit einer bestimmten Testperson durchführt, die Testleistung der Testperson genau gleich auswertet und interpretiert. Sinnvollerweise wird das Gütekriterium der Objektivität in drei Aspekte unterteilt nämlich in die Durchführungs-, Auswertungs- und Interpretationsobjektivität (Moosbrugger & Kelava, 2012).

Schmidt-Atzert & Amelang (2012) zeigen auf, was bei den jeweiligen Differenzierungen beachtet werden sollte.

Durchführungsobjektivität (Standardisierung des Tests, genaue Instruktionen, klare Anweisungen für Testleiter u. a. zum Umgang mit Fragen, Beispielitems)

Auswertungsobjektivität (genaue Anweisungen zum Gebrauch von Schablonen, Angaben zum Umgang mit nicht beantworteten Items, Umgang mit unterschiedlichen Beobachtungsergebnissen oder Beurteilungen, bei nicht völlig standardisierbaren Auswertungen Massnahmen zur bestmöglichen Gewährleistung der Objektivität, bei computerbasierten Tests soll die Auswertung zu kontrollieren sein)

Interpretationsobjektivität (Fallbeschreibungen im Manual, Hinweise zur Berücksichtigung von Testerfahrung etc., Angaben zur benötigten Sachkunde)

2.2.2. Reliabilität

Für die Reliabilität finden Moosbrugger & Kelava (2012) folgende Definition: "Ein Test ist dann reliabel (zuverlässig), wenn er das Merkmal, das er misst, exakt, d.h. ohne Messfehler, misst". (S. 11)

Um das Ausmass der Reliabilität zu bestimmen wurden im Rahmen der KTT mehrere Verfahren entwickelt. Dabei werden vier Vorgehensweisen unterschieden (Moosbrugger & Kelava, 2012).

2.2.2.1 Retest-Reliabilität

Um die Reliabilität nach dem Retest-Verfahren zu bestimmen, wird derselbe Test, unter der idealen Annahme, dass sich das zu messende Merkmal selbst nicht verändert hat, zu zwei verschiedenen Zeitpunkten vorgelegt. Die Reliabilität wird dann als Korrelation zwischen den beiden Testergebnissen ermittelt. Dabei ist zu beachten, dass die ermittelte Korrelation in Abhängigkeit vom Zeitintervall zwischen beiden Testungen variieren kann. Je nach Zeitabstand ist eine Vielzahl von Einflüssen auf die Messungen denkbar, die sich reliabilitätsverändernd auswirken können, insbesondere Übungs- und Erinnerungseffekte oder ein sich tatsächlich veränderndes Persönlichkeitsmerkmal (Moosbrugger & Kelava, 2012).

2.2.2.2 Paralleltest-Reliabilität

Etliche reliabilitätsverändernde Einflüsse wie Übungs- und Erinnerungseffekte, aber auch Merkmalsveränderungen können eliminiert bzw. kontrolliert werden, wenn die Reliabilität nach dem Paralleltest-Verfahren bestimmt wird. Dieses Verfahren wird oftmals als Königsweg der Reliabilitätsbestimmung bezeichnet. Hierfür wird die Korrelation zwischen den beobachteten Testwerten in zwei parallelen Testformen berechnet, die aus inhaltlich möglichst ähnlichen Items sogenannten Item Zwillingen bestehen. Parallel sind zwei Testformen dann, wenn sie trotz nicht identischer Item Stichproben zu gleichen wahren Werten und Varianzen der Testwerte führen (Moosbrugger & Kelava, 2012).

2.2.2.3 Testhalbierungs-Reliabilität

Oftmals ist es nicht möglich, einen Test zu wiederholen oder parallele Testformen herzustellen weil z.B. die Testpersonen zu einem zweiten Termin nicht zur Verfügung stehen, die Verzerrungen durch eine Wiederholung zu hoch waren, oder dass ein Item Pool nicht gross genug ist, um zwei parallele Testformen herzustellen. In solchen Fällen ist es angebracht, den Test in zwei möglichst parallele Testhälften zu teilen und die Testhalbierungs-Reliabilität auch Split-Half-Reliabilität genannt als Korrelation der beiden Testhälften zu bestimmen (Moosbrugger & Kelava, 2012).

2.2.2.4 Innere Konsistenz

Die Konsistenzanalyse stellt eine Verallgemeinerung der Testhalbierungsmethode dar. Jedes Item eines Tests wird als eigenständiger Testteil betrachtet. Je stärker die Testteile untereinander positiv korrelieren, desto höher ist die interne Konsistenz des Verfahrens. Das Mass der internen Konsistenz einer Skala wird als Cronbachs Alpha bezeichnet (Moosbrugger & Kelava, 2012).

2.2.2.5 Interrater-Reliabilität

Die Interrater-Reliabilität kommt dann zum Einsatz, wenn beispielsweise von verschiedenen Personen, die sich in der gleichen Situation befunden, Angaben zur Situationseinschätzung gefordert sind (Rost, 1996).

2.2.3 Validität

Was ist Validität? Gemäss Moosbrugger & Kelava (2012) ist Validität, aus dem Englischen validity also Gültigkeit, der Begriff der das Ausmass zusammenfasst, in dem ein Test misst was er zu messen vorgibt oder anders gesagt, misst was er messen soll. Diese vereinfachenden Zusammenfassungen drücken aus, dass Validität ein umfassendes und sehr wichtiges Gütekriterium zur Beurteilung eines diagnostischen Verfahrens darstellt. Die Validität ist den Gütekriterien Objektivität und Reliabilität übergeordnet. Denn wenn ein Test nicht gültig ist, weil er zum Beispiel etwas anderes erfasst, als er sollte, sind Objektivität oder Reliabilität nicht mehr von Belang. Validität ist jedoch auch das komplexeste und am schwierigsten zu bestimmende Gütekriterium. Die Frage, ob ein Test misst, was er messen soll, klingt zunächst sehr einfach, ist aber auf den zweiten Blick schwierig zu beantworten. Es gibt verschiedene Methoden und verschiedene Kriterien, um diese Frage zu beantworten. Validität wird aus diesem Grund als ein breit definiertes Gütekriterium verstanden, das sich auf verschiedene Qualitätsaspekte eines Tests bezieht. Statt von der Validität eines Tests zu sprechen, ist es daher angemessener, die Validität verschiedener möglicher Interpretationen von Testergebnissen zu betrachten (Moosbrugger & Kelava, 2012).

Unterschieden wird meistens in drei Arten der Validität. Schuler (2014) erwähnt aber, dass auch weitere Validitätsbegriffe Verwendung finden.

2.2.3.1. Kriteriumsvalidität

Unter Kriteriumsvalidität versteht man den Zusammenhang zwischen Testergebnis und konkreten Leistungen oder Verhaltensweisen ausserhalb der Testsituation. Das Kriterium muss für den vorgesehenen Einsatzbereich des Tests relevant sein. so werden beispielsweise Intelligenztests zur Vorhersage von Schul- oder Ausbildungserfolg verwendet. In diesem Fall bezeichnet man den Schul- bzw. Ausbildungserfolg als Kriterium. Ein Kriterium ist immer etwas Konkretes, direkt Messbares, wie Prüfungsleistungen, erzielter Umsatz, oder auch Verhaltensweisen wie Studienabbruch (Schuler, 2014).

2.2.3.2 Inhaltsvalidität

Die Inhaltsvalidität gibt an, wie repräsentativ die Inhalte eines Tests für das zu messende Konstrukt sind. Tests bestehen aus Items, und diese sollen das Merkmal repräsentieren, das mit dem Test erfasst werden soll. Wir stellen uns vor, man könne alle Items finden, die zur Messung von Intelligenz oder eines anderen Merkmals grundsätzlich geeignet sind. Dazu sind genaue Kenntnisse des Merkmals erforderlich. Ein inhaltsvalider Test besteht aus einer repräsentativen Auswahl von Items aus dem Item Universum. Das klingt ziemlich hypothetisch, und das ist es auch. Deshalb wird die Inhaltsvalidität auch nicht bei allen Tests thematisiert. Bei einigen Tests ist es aber sehr sinnvoll, nach deren Inhaltsvalidität zu fragen (Schuler, 2014).

2.2.3.3 Konstruktvalidität

Unter Konstruktvalidität versteht man empirische Belege dafür, dass ein Test das Konstrukt erfasst, welches er erfassen soll und nicht ein anderes.

Ob ein Test Intelligenz oder ein Persönlichkeitsmerkmal misst, ist eine Frage nach dessen Konstruktvalidität. Konstrukte existieren nicht real. Vielmehr handelt es sich um gedankliche Konstruktionen, um abstrakte Konzepte. Für die Wissenschaft haben sie sich aber als nützlich erwiesen, und auch Laien verwenden Konstrukte, auch wenn sie beispielsweise unter Intelligenz oder Angst nicht genau das Gleiche verstehen wie ein Wissenschaftler (Schuler, 2014).

2.2.3.4 konvergente und diskriminante Validität

Für die Konstruktvalidierung besteht eine besondere Herausforderung. Es gilt nachzuweisen, dass der Test das angestrebte Konstrukt erfasst und nicht etwa ein anderes. Dabei wird zum einen in konvergente Validität unterschieden welche den Zusammenhang mit weiteren Indikatoren des Konstrukts, das man erfassen will abbildet und der diskriminanten Validität, welche den Zusammenhang mit Indikatoren anderen Konstrukte aufzeigt (Schuler, 2014).

2.2.3.5 Augenscheinvalidität

Augenscheinvalidität ist dann gegeben, wenn der Laie bzw. Proband glaubt zu erkennen, was der Test misst und daher eher Vertrauen in das Verfahren entwickelt. Unter Marketinggesichtspunkten kann dies nützlich sein, für die Unverfälschbarkeit ist die Nachvollziehbarkeit allerdings abträglich (Schuler, 2014).

Auch Beauducel und Leue (2014) geben zu bedenken, dass ob die Augenscheinvalidität eher zu einer Erhöhung der Verfälschung der Antworten führt oder ob diese Transparenz die Akzeptanz eines Inventars erhöht und damit das Kooperationsverhalten der Probanden erhöht vom jeweiligen Kontext abhängig sind. Moosbrugger & Kelava (2012) schreiben jedoch der Augenscheinvalidität vor dem Hintergrund der Mittelbarkeit der Ergebnisse und der Akzeptanz von Seiten der Testpersonen eine erhebliche Bedeutung zu. Nicht zuletzt auch wegen der Bekanntheit der Intelligenzforschung haben z.B. Intelligenz-Tests eine hohe Augenscheinvalidität, da Laien aufgrund von Inhalt und Gestaltung des Tests es für glaubwürdig halten, dass damit Intelligenz gemessen werden kann. Aus der wissenschaftlichen Perspektive ist die Augenscheinvalidität allerdings nicht immer zufriedenstellend, denn die Validität eines Tests muss auch empirisch durch Kennwerte belegt werden (Moosbrugger & Kelava, 2012).

2.2.3.6 Inkrementelle Validität

„Die inkrementelle Validität bezeichnet das Ausmass, in dem die Vorhersage des praktisch relevanten externen Kriteriums verbessert werden kann, wenn zusätzliche Testaufgaben oder Testskalen zu den bereits existierenden hinzugenommen werden“ (Moosbrugger & Kelava, 2012, S. 166).

Um inkrementelle Validität etwas besser zu veranschaulichen zeigt Schuler (2014) folgendes Beispiel. Intelligenztests korrelieren durchschnittlich $r = .51$ mit Berufserfolg. Demnach kann Berufserfolg zu 26 Prozent durch Intelligenz aufgeklärt werden. Strukturierte Interviews sind ein ebenso guter Prädiktor des Berufserfolgs $r = .51$. Es wäre aber naiv zu glauben, dass sich die Validitäten einfach addieren lassen. Würden Eignungsinterviews nichts anderes als Intelligenz messen, könnten sie die mit Intelligenztests erzielte Vorhersage nicht verbessern. Im Durchschnitt beträgt die Validität einer Kombination von Intelligenztest und strukturiertem Eignungsinterview $R = .63$. Strukturierte Interviews klären dem zu folge durchschnittlich weitere 14 Prozent der Varianz des Berufserfolgs auf (Schuler, 2014).

2.2.4 Normierung (Eichung)

Für die Normierung oder auch Eichung eines Tests wird ein Bezugssystem erstellt, mit dessen Hilfe die Ergebnisse einer Testperson im Vergleich zu den Merkmalsausprägungen anderer Personen eindeutig eingeordnet und interpretiert werden können. Zweck der Normierung ist es, möglichst aussagekräftige Vergleichswerte von solchen Personen zu erhalten, die der Testperson hinsichtlich relevanter Merkmale (z.B. Alter, Geschlecht, Schulbildung) ähnlich sind. Die Testergebnisse der untersuchten Person werden dann bei der normorientierten Beurteilung in Relation zu den Testergebnissen von Personen aus der Eichstichprobe interpretiert (Moosbrugger & Kelava, 2012).

2.3 Psychologische Diagnostik

2.3.1 Definition psychologischer Diagnostik

In der wissenschaftlichen Literatur finden sich unterschiedliche Definitionen von Psychologischer Diagnostik.

Schmidt-Atzert und Amelang liefern folgenden Vorschlag:

Psychologische Diagnostik ist eine Teildisziplin der Psychologie. Sie dient der Beantwortung von Fragestellungen, die sich auf die Beschreibung, Klassifikation, Erklärung oder Vorhersage menschlichen Verhaltens und Erlebens beziehen. Sie schliesst die gezielte Erhebung von Informationen über das Verhalten und Erleben eines oder mehrerer Menschen sowie deren relevanter Bedingungen ein. Die erhobenen Informationen werden für die Beantwortung der Fragestellung interpretiert. Das diagnostische Handeln wird von psychologischem Wissen geleitet. Zur Erhebung von Informationen werden Methoden verwendet, die wissenschaftlichen Standards genügen. (S. 19)

2.3.2 Definition: Eignung und Eignungsdiagnostik

Schuler (2014) definiert Eignung als Erfolgswahrscheinlichkeit einer Person in einem bestimmten Umfeld. Dabei ist zu bedenken, dass Eignung keine generelle Eigenschaft einer Person darstellt, sondern ein Relationsmerkmal, das den Grad der Passung von Person und Aufgabe widerspiegelt.

Unter Eignungsdiagnostik versteht Schuler (2014) die Entwicklung, Prüfung und Anwendung psychologischer Verfahren, zum Zwecke eignungsbezogener Erfolgsprognosen und Entscheidungshilfen, auf Grundlage wissenschaftlicher Theorien, Messmethoden und Klassifikationsmodellen. Im beruflichen Kontext sind die praktischen Einsatzbereiche alle Arten berufsbezogener Einschätzung, Beratung, Selektion und Zuordnung wie sie in der institutionellen Berufsberatung sowie der Auswahl und Entwicklung von Mitarbeitern vorkommen (Schuler, 2014).

2.3.3 Drei Ansätze der Eignungsdiagnostik

Schuler (2014) zeigt auf, dass bei der Klassifikation von eignungsdiagnostischen Instrumenten in drei methodische Ansätze unterschieden wird. In den Eigenschaftsansatz, den Simulationsansatz und biografischen Ansatz. Jeder dieser Ansätze folgt einem eigenen Validitätsprinzip und ist dementsprechend auch mit einer eigenständigen Validierungsstrategie verbunden. Mit dem Eigenschafts- oder Konstruktansatz werden Merkmale erfasst die als relativ stabil angenommen werden. Hierzu zählt zum Beispiel sprachgebundene Intelligenz oder Gewissenhaftigkeit, diese werden typischerweise mit psychologischen Tests gemessen. Das vorrangig verfolgte Validitätsprinzip bei psychologischen Tests ist die Konstruktvalidität. Zielsetzung des Simulationsansatzes ist die Erfassung von Verhalten das in ähnlicher Form am Arbeitsplatz gefordert wird. Das dazu gehörende Validitätsprinzip ist die inhaltliche Repräsentativität oder Inhaltsvalidität, typische Erhebungsform ist die Arbeitsprobe. Typische Erfassungsmethoden bei dem biografischen Ansatz sind biografische Fragen in mündlicher oder schriftlicher Form, die Validitätslogik ist die kriteriumsbezogene Validität (Schuler, 2014).

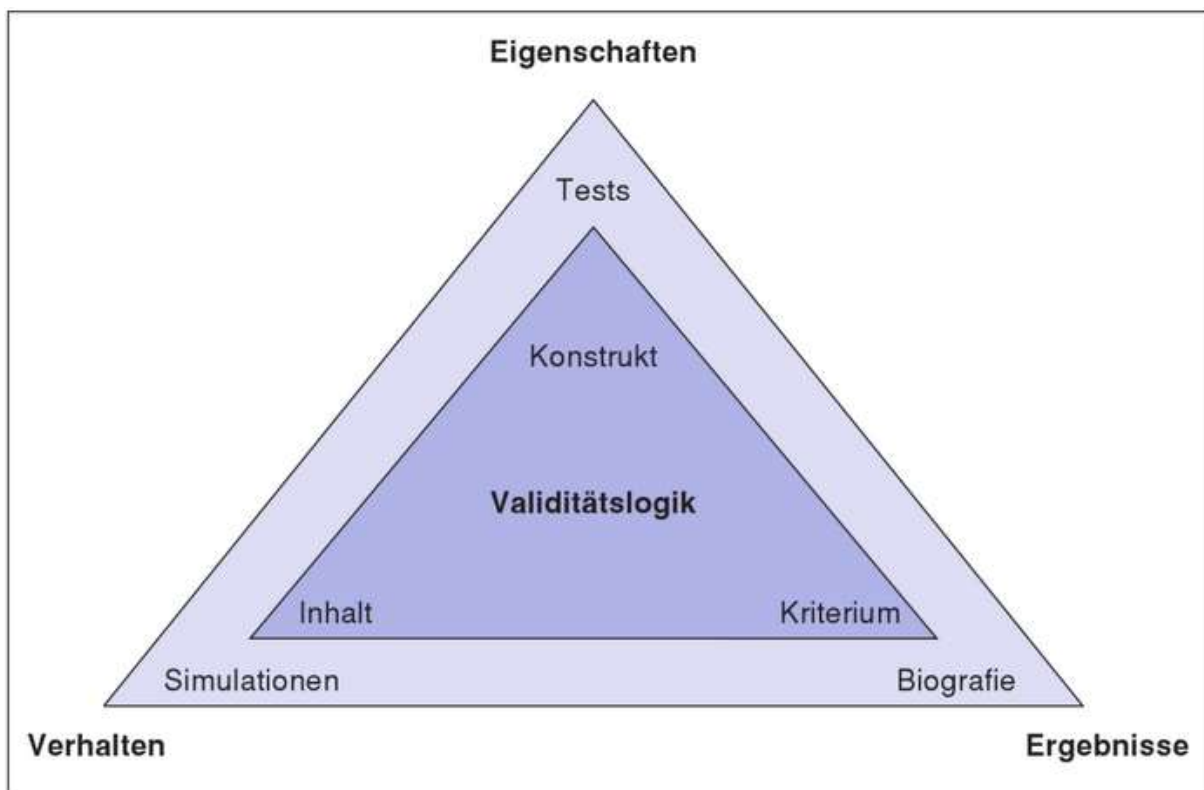


Abbildung 1: Die Ansätze der Eignungsdiagnostik und ihre Validitätslogik. Quelle: Schuler (2014, S. 153).

Schuler (2014) gibt zu bedenken, dass in der praktischen Berufseignungsdiagnostik zumeist komplexe Anforderungen abgebildet werden müssen und sich daher die Berücksichtigung mehrerer methodischer Zugänge empfiehlt. Erst durch die Systematische Kombination unterschiedlicher Verfahren wird sichergestellt, dass alle relevanten Facetten des interessierenden Merkmalsbereichs abgedeckt werden. Durch das multimethodale Verfahrensrepertoire ist eine grosse Spannweite möglicher Erhebungsformen abgedeckt. Beispielsweise vergangenheitsbezogene wie gegenwärtige Verhaltensbeschreibung, Selbsteinschätzung und Fremdeinschätzung, sowie typisches und maximales Verhalten. Ein multimethodales Verfahren gewährleistet eine erhöhte Generalisierbarkeit der eignungsdiagnostischen Befunde auf zukünftiges berufliches Verhalten. Darüber hinaus hat die Messung eines Merkmals mit mehreren Methoden den Vorteil höherer Reliabilität. Eine hohe Diagnosequalität ergibt sich also nicht allein aus der Konstruktion der Einzelverfahren, sondern auch aus der theoriegestützten Systematik des Gesamtverfahrens (Schuler, 2014).

2.4 Übersicht einiger Testverfahren und deren Validität

Schuler (2014) beschreibt, dass die Klassische Methode zur Operationalisierung psychologischer Konstrukte der Test ist. Psychologische Tests sind gewissermassen der Prototyp von konstruktorientierter Diagnoseverfahren. Sie sind in der wissenschaftlich kontrollierten Eignungsdiagnostik die am häufigsten verwendeten Instrumente, ihr Einsatz erstreckt sich auf praktisch alle Berufsbereiche und beruflich relevante Fähigkeiten. Testverfahren welche in der Berufseignungsdiagnostik eingesetzt werden, können in zwei Klassen eingeteilt werden. Kognitive Fähigkeitstests, die maximale Leistung messen und persönlichkeitsorientierte Verfahren, die typische Leistung beziehungsweise typisches Verhalten erfassen. Zu den kognitiven Fähigkeitstests gehören die allgemeinen Intelligenztests, Tests spezifischer kognitiver Fähigkeiten, Aufmerksamkeits- und Konzentrationstests, Tests zur sensorischen und motorischen Leistung und sonstige Leistungstests. Zu den persönlichkeitsorientierten Verfahren gehören die allgemeinen Persönlichkeitstests, spezifische Persönlichkeitstests sowie Einstellungs-, Motivations- und Interessentests (Schuler, 2014).

2.4.1 Intelligenz und Beruf

Kognitive Fähigkeitstests wurden früh als eignungsdiagnostisches Instrument eingesetzt. Bereits während der beiden Weltkriege wurden ein allgemeiner kognitiver Fähigkeitstest und ein sprachfreier Alternativtest bei Millionen von Rekruten eingesetzt (Schuler, 2014).

Dabei zeigte sich eine grosse Abweichung der Intelligenzwerte zwischen den Berufsgruppen. Weiterhin wird deutlich, dass die Streuung der Intelligenzausprägung innerhalb der Berufe abnimmt, je höher die Qualifikationsanforderungen des jeweiligen Berufs sind. Die schwächste Gruppe stellen die Landarbeiter (farm workers), Wirtschaftsprüfer (accountants) erzielten die besten Werte (Ackermann & Humphreys, 1991).

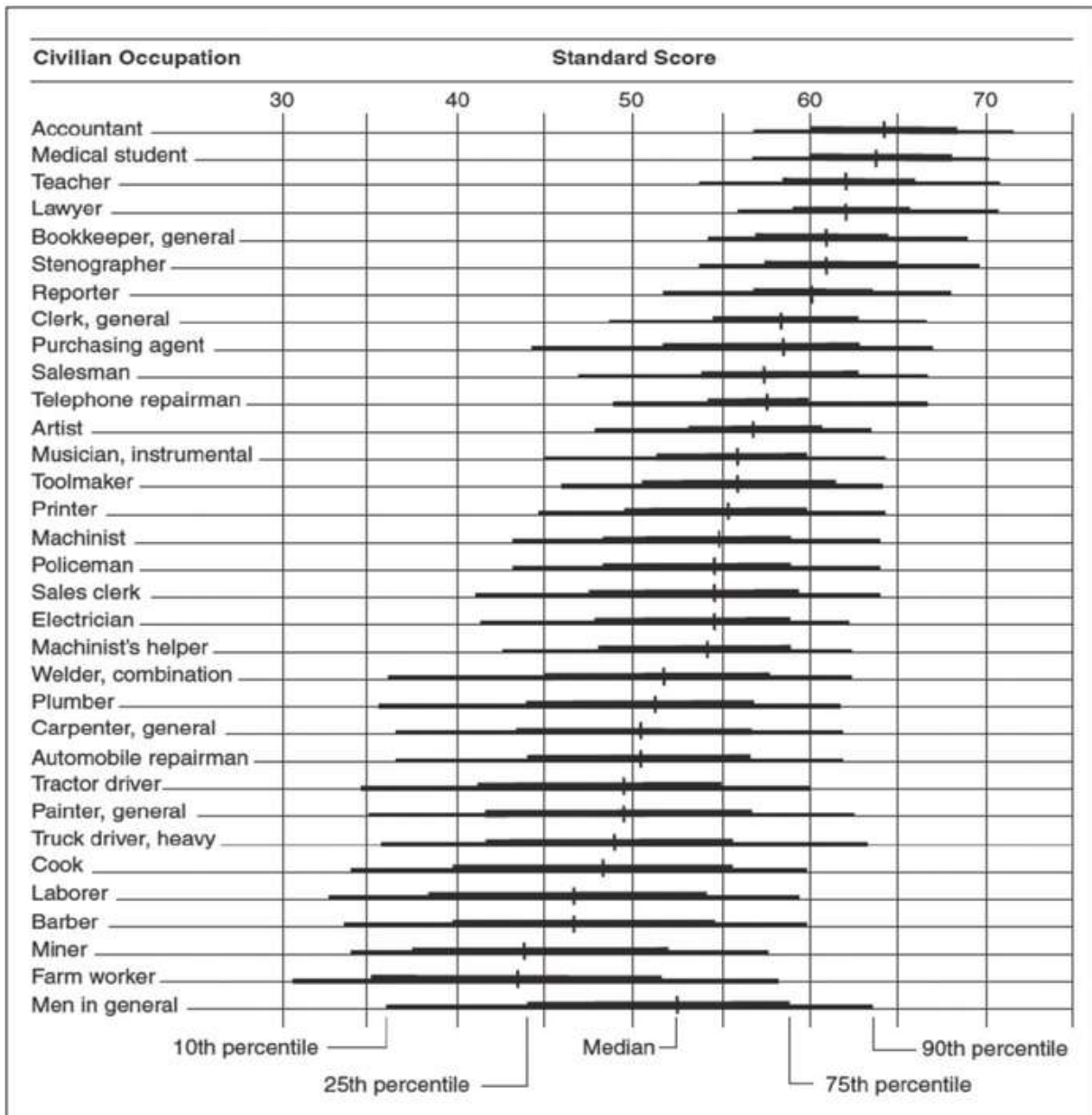


Abbildung 2: Werte des Army General Classification Test. Quelle : Schuler (2014, S. 159).

Ähnliche Ergebnisse finden sich auch im deutschen Raum. Eine Aufstellung von verschiedenen Berufen mit einem grossen Datensatz von (N=30477) kam zum Resultat, dass sich die Stelleninhaber bezüglich ihres Intelligenzniveaus gravierend unterscheiden (Engelbrechts, 1994).

Wie können diese Befunde erklärt werden? Schallberger (2000) geht davon aus, dass eine Passung zwischen den kognitiven Fähigkeiten des Ausführenden und den beruflichen Anforderungen notwendig ist. Weiterhin unterscheiden sich die Berufe in ihren Anforderungen an die kognitive Leistungsfähigkeit. Personen, die diese Anforderungen nicht erfüllen, werden dauerhaft erfolglos sein und der selbst- wie auch Fremdselektion unterliegen. Andererseits dürften Personen, die die Berufsanforderungen deutlich übertreffen, auf Dauer unterfordert sein und werden dazu neigen, die Arbeitstätigkeit zu wechseln (Schallberger, 2000).

Umgekehrt wirkt sich die Arbeitstätigkeit auf die kognitive Leistungsfähigkeit der Berufsinhaber aus. Besonders herausfordernde Tätigkeiten erhalten die kognitiven Fähigkeiten und führen zur Entfaltung bereichsspezifischer kognitiver Fähigkeiten (Häfeli, Kraft & Schallberger, 1988).

Schmidt und Hunter (1998) gehen davon aus, dass es praktisch keinen Beruf gibt für den allgemeine Fähigkeitstests keine kriteriumsbezogene Validität für berufliche Leistungsvariablen aufweisen. Ergänzend zeigen sie, dass kognitive Fähigkeiten auch mit anderen Erfolgskriterien zusammenhängen. So erweist sich die Allgemeine Intelligenz mit .56 im Vergleich zu anderen Merkmalen als stärkster Einzelprädiktor für Trainings- und Ausbildungserfolg.

Die folgende Tabelle zeigt die Validität kognitiver Verfahren für berufseignungsdiagnostische Zwecke auf Meta-Ebene auf (Schmidt & Hunter, 1998).

Ansatz	Prädiktor	Angaben aus Schmidt und Hunter (1998)			Neuere Befunde	
		Val	mR	inkrV	Val	Neuere Metaanalyse
Eigenschaft	Allgemeine kognitive Fähigkeitstests	.51			.62	Salgado et al. (2003)
	„Integrity“-Tests	.41	.65	.14	.18	Van Iddekinge, Roth, Raymark und Odle-Dusseau (2012)
	Gewissenhaftigkeitstests	.31	.60	.09	.31	Barrick, Mount und Judge (2001)
	Fachkenntnistests	.48	.58	.07		
	Interessen	.10	.52	.01	.23	Van Iddekinge, Roth, Putka und Lanivich (2011)
Simulation	Probezeit	.44	.58	.07		
	Arbeitsproben	.54	.63	.12	.39	Roth, Bobko und McFarland (2005)
	Situational Judgement Tests				.34	McDaniel, Morgeson, Finnegan, Campion und Braverman (2001)
	Assessment Center	.37	.53	.02	.28	Hermelin, Lievens und Robertson (2007)
Biografie	Grafologie	.02	.51	.00		
	Biografische Daten	.35	.52	.01	.32	Bliesener (1996)
	Unstrukturiertes Einstellungsgespräch	.36	.55	.04	.20	Huffcutt und Arthur (1994)
	Strukturiertes Einstellungsgespräch	.51	.63	.12	.40	Schmidt und Rader (1999)

Abbildung 3: Zusammenfassung metaanalytischer Befundlage zur Validität eignungsdiagnostischer Verfahren. Quelle: Schuler (2014, zitiert nach Schmidt & Hunter, 1998).

Die Quellen zu aktuelleren Metaanalysen werden direkt auf der rechten Seite in der Tabelle aufgeführt. Folgende Abkürzungen werden verwendet: Val = kriteriumsbezogene Validität (Kriterium = Leistungsbeurteilung) mR = multiple Korrelation mit allgemeiner kognitiver Fähigkeit als erster und mit dem jeweiligen Prädiktor als zweiter Variable und inkrv = inkrementelle Validität, das heisst Validitätszuwachs durch Hinzunahme des zweiten Prädiktors (Schuler, 2014).

2.4.2 Kritik an der Validität zu eignungsdiagnostischen Verfahren

Schuler (2014) gibt zu bedenken, dass wenn Bewerbergruppen in bestimmten Berufssparten z.B. aus spezialisierten Fachakademikern bestehen, so sind diese Gruppen bereits durch ihren Ausbildungsweg stark nach kognitiven Fähigkeiten vorselektiert, was zu einer Varianzeinschränkung führt. Ein Beispiel dafür liefert die Metaanalyse von Funke, Krauss, Schuler & Stapf (1987), bei welcher zur Prognose wissenschaftlich-technischer Leistung von Mitarbeitern im Forschungs- und Entwicklungsbereich von Grossunternehmen eine mittlere Validität von $r=.16$ aufwies.

Eine weitere Herausforderung im Umgang mit kognitiven Fähigkeitstests besteht gemäss Schuler (2014) vor allem bei Massenuntersuchungen wie z.B. Studierfähigkeitstests an denen Personen mehrmals teilnehmen können, oder bei besonders prestigeträchtigen Berufen auf die sich Bewerber sehr intensiv vorbereiten. Denn eine wiederholte Bearbeitung von identischen oder strukturell ähnlichen kognitiven Testverfahren führt in der Regel zu einer bedeutsamen Verbesserung der erzielten Testleistung. Gegenmassnahmen die geeignet sind um unerwünschte Wiederholungseffekten zu begegnen, sind unter anderem, die Verwendung von nicht öffentlich zugänglichen Verfahren, Teilnehmer nur einmalig zu einem Verfahren zu zulassen oder grosse Zeitintervalle von bis zu einem Jahr bis zur wiederholten Teilnahme vorzugeben. Auch der Einsatz von Parallelverfahren eignet sich wie auch die Verwendung von wenig übungsanfälligen eignungsdiagnostischer Verfahren (Schuler, 2014).

2.4.3 Allgemeine Persönlichkeitskonstrukte und spezifische Persönlichkeitseigenschaften

Schuler (2014) beschreibt als nicht kognitives eignungsdiagnostisches Verfahren ein allgemeines Modell das Fünf-Faktoren-Modell der Persönlichkeit (auch kurz „Big Five“ genannt). Auch wenn ein Diskurs darüber besteht wie viele und welche Faktoren die Persönlichkeitsstruktur eines Menschen ausreichend beschreibt, so wird dieses Modell von vielen Forscherinnen und Forschern akzeptiert. Auch in neueren Studien konnten wiederholt fünf abstrakte Faktoren nachgewiesen werden die dieses Modell bestätigen. Bemerkenswert ist dabei die Konvergenz ganz unterschiedlicher Forschungsansätze. Ähnliche Befunde finden sich in lexikalischen Ansätzen, in faktorenanalytischen Reanalysen und in interkulturellen Vergleichsstudien. Eine grobe Charakterisierung der Fünf Faktoren zeigt die folgende Tabelle (Schuler, 2014).

Neurotizismus	<ul style="list-style-type: none"> • Ängstlichkeit • Reizbarkeit • Depression • Soziale Befangenheit • Impulsivität • Verletzlichkeit
Extraversion	<ul style="list-style-type: none"> • Herzlichkeit • Geselligkeit • Durchsetzungsfähigkeit • Aktivität • Erlebnishunger • Frohsinn
Offenheit für Erfahrungen	<ul style="list-style-type: none"> • Offenheit für Phantasie • Offenheit für Ästhetik • Offenheit für Gefühle • Offenheit für Handlungen • Offenheit für Ideen • Offenheit des Normen- und Wertesystems
Verträglichkeit	<ul style="list-style-type: none"> • Vertrauen • Freimütigkeit • Altruismus • Entgegenkommen • Bescheidenheit • Gutherzigkeit
Gewissenhaftigkeit	<ul style="list-style-type: none"> • Kompetenz • Ordnungsliebe • Pflichtbewusstsein • Leistungsstreben • Selbstdisziplin • Besonnenheit

Abbildung 4: Charakterisierung der Big Five. Quelle: Schuler (2014, zitiert nach Ostendorf & Angleitner, 2004).

Obwohl sich das Fünf-Faktoren-Modell als Orientierungsrahmen für Praxis und Forschung sehr bewährt hat, stellt eine Beschreibung anhand von nur fünf Merkmalen eine doch recht undifferenzierte Charakterisierung der Persönlichkeit dar. Beispiele für differenziertere Modelle sind der 16-Persönlichkeits-Faktoren-Test nach Cattell, bei welchem nach 16 Eigenschaften unterschieden werden kann. Das Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung (BIP) mit 14 Skalen zu beruflicher Orientierung, Arbeitsverhalten, sozialen Kompetenzen und psychischer Konstitution. Ein Beispiel dafür, dass selbst innerhalb eines breiter gefassten Merkmals noch weiter differenziert werden kann, ist das Leistungsmotivationsinventar (LMI), mit dem 17 Teilaspekte der Leistungsmotivation unterschieden werden (Schuler, 2014).

2.4.4 Motivation und Interessen

Motivation spielt gemäss Schuler (2014) zur Erklärung von Verhalten in der Persönlichkeitspsychologie wie auch im Alltagsleben eine grosse Rolle. Jedoch gestalten sich Versuche Motive und Motivation zu messen schwierig. Diese treibende Kraft ist als solche nicht beobachtbar und in manchen psychologischen Theorien wird angenommen sie sei auch für das handelnde Individuum nicht erkennbar. Eine Messung erfordert den Nachweis eines kausalen Zusammenhangs zwischen Motivation und dem Verhaltensereignis. Unter den heute eingesetzten Tests zur Erfassung der Leistungsmotivation finden sich nahezu ausschließlich Verfahren vom Fragebogentypus. Beispiele hierfür sind der Leistungsmotivationstest (L-M-T) von Hermans, Petermann und Zielinski (1978), der als Teilaspekte Leistungsstreben, Ausdauer und Fleiß, Leistungshemmende Prüfungsangst und Leistungsfördernde Prüfungsangst unterscheidet, sowie das Leistungsmotivationsinventar (LMI, Schuler & Prochaska, 2001), das Leistungsmotivation als Verhaltensdisposition erfasst, die in mehreren breiten Persönlichkeitsmerkmalen verankert ist (Schuler, 2014).

Es besteht gemäss Schuler (2014) weitgehend Einigkeit darüber, dass die Typologie der Berufsorientierung nach Holland eine sehr ergiebige Heuristik für die Eignungsdiagnostik in der Berufsberatung ist. Dahinter steckt die Theorie, dass Interessen grundlegende Persönlichkeitsorientierungen sind und die Berufswahl durch allgemeine Wesensmerkmale bedingt sind.

1. Realistische Orientierung	<p>Personen dieses Typus sind „männlich“, aktiv, aggressiv; sie sind an physischer Aktivität interessiert und motorisch befähigt; sie bevorzugen konkrete Gegebenheiten gegenüber abstrakten Problemen; sie haben eher konventionelle Werthaltungen – sowohl im politischen wie im ökonomischen Bereich. Bei der Berufswahl tendieren sie am meisten zu handwerklichen und technischen sowie zu land- und forstwirtschaftlichen Berufen.</p>
2. Intellektuelle Orientierung	<p>Dieser Typus ist aufgabenorientiert und in gewissem Sinne „asozial“; er trachtet danach, Probleme intellektuell zu bewältigen, „durch Manipulation von Ideen, Worten und Symbolen“, nicht dagegen durch physische oder soziale Aktivität. Er hat ein starkes Bedürfnis, Zusammenhänge zu verstehen und besitzt eher unkonventionelle Wertvorstellungen und Einstellungen. Menschen dieses Typus werden vor allem in naturwissenschaftlichen und mathematischen Berufen gefunden.</p>
3. Soziale Orientierung	<p>Menschen dieses Typus sind sozial orientiert und von sozialer Verantwortung erfüllt. Sie haben ein starkes Bedürfnis nach Beachtung und sozialer Interaktion und verfügen über gute verbale und soziale Fähigkeiten. Sie tendieren dazu, Probleme eher emotional und durch soziale Aktivität zu bewältigen als intellektuell. Typische Berufe in diesem Bereich sind: pädagogische und sonderpädagogische Berufe, Sozialarbeiter, klinischer Psychologe, Berufsberater.</p>
4. Konventionelle Orientierung	<p>Dieser Typus bevorzugt weitgehend strukturierte verbale und numerische Aktivitäten und Untergebenenrollen; er ist konformistisch („extracptive“) eingestellt und vermeidet unklare Situationen sowie Probleme, die soziale Aktivität oder ausgeprägte physische Fähigkeiten erfordern; er identifiziert sich mit Machtpositionen und schätzt materiellen Besitz und Status. Bei der Berufswahl und der beruflichen Ausbildung werden z.B. folgende Berufe gewählt: Buchhalter, Rechnungsprüfer, Bankangestellter, Statistiker, EDV-Operator.</p>
5. Unternehmerische Orientierung	<p>Menschen dieser Art verstehen sich als starke, männliche Führerpersönlichkeiten. Sie besitzen ausgeprägte verbale Fertigkeiten und fühlen sich wohl, wenn sie anderen etwas verkaufen oder mit ihnen in Konkurrenz treten können. Sie vermeiden jedoch klar definierte verbale Situationen sowie Aufgaben, die einen längeren, angestregten intellektuellen Einsatz erfordern. Die beruflichen Präferenzen betreffen u. a. folgende Berufe: Hotelier, Unternehmer, Industriebereiter, Immobilienhändler, Wahlkampfmanager, Versicherungsvertreter und dergleichen.</p>
6. Künstlerische Orientierung	<p>Dieser Typus ähnelt dem intellektuellen hinsichtlich seiner „int-razeptiven“ und eher „asozialen“ Ausrichtung, er unterscheidet sich jedoch von ihm durch sein Bedürfnis nach Selbst-Ausdruck mit Hilfe künstlerischer Medien. Er meidet hochgradig strukturierte Probleme und Aufgaben, die grobmotorische Fertigkeiten erfordern. Menschen dieser Art haben ferner eine geringere Ichstärke, sind eher feminin und leiden häufiger unter emotionalen Störungen. Sie tendieren natürlich vor allem zu künstlerischen oder mit dem Kultur- und Kunstleben befassten Berufen.</p>

Abbildung 5: Hollands Typenschema. Quelle: Schuler (2014, zitiert nach Seifert, 1977).

Eine Besonderheit des Hollandschen Ansatzes, der ihn für die Berufs- und Studienberatung besonders interessant macht, ist die Einschätzung von Berufen nach den gleichen Kategorien. Die theoretische Annahme ist, dass Menschen nach (beruflichen) Umwelten suchen, die es ihnen ermöglichen, ihre Fähigkeiten bestmöglich anzuwenden und ihre Werthaltungen zu verwirklichen. Dies ist ihnen am besten dort möglich, wo hohe Kongruenz zwischen den Orientierungen der Person und den Anforderungen der Umwelt vorliegt. Das Ausmaß an Kongruenz ergibt sich nicht allein durch die Übereinstimmung oder Nichtübereinstimmung der dominierenden Werthaltungen, sondern durch die Prüfung der Orientierungsmuster von Person und Umwelten auf Übereinstimmung. Mit dem in Abbildung XY wiedergegebenen hexagonalen Modell soll dies möglich sein. Die räumliche Nähe im Modell symbolisiert die Ähnlichkeit der Orientierungen (Schuler, 2014).

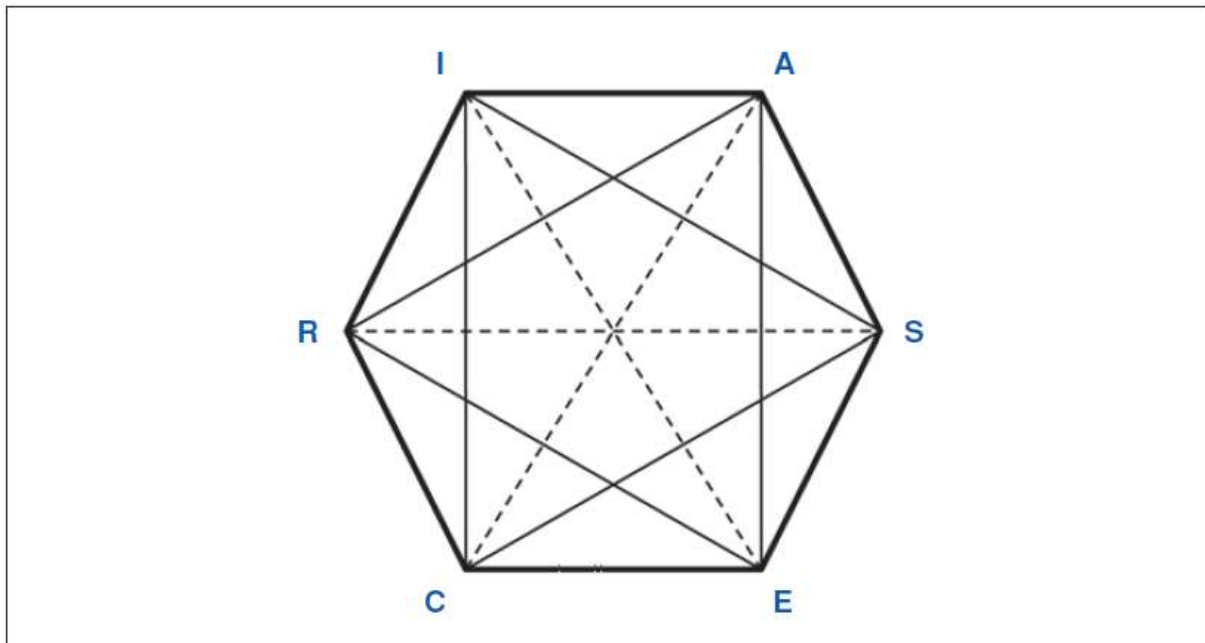


Abbildung 6: Hexagonales Modell zur Bestimmung der Ähnlichkeit zwischen Persönlichkeitstypen, Umwelttypen und deren Beziehungen. Quelle: Schuler (2014, zitiert nach Bergmann & Eder, 2005).

Bergmann und Eder (2005) erläutern, dass es nach dem Modell von Holland vier Abstufungen von Kongruenzen gibt:

Wenn z. B. eine praktisch-technisch orientierte Person (R) einen praktisch-technischen Beruf (R) ergreift, dann liegt maximale Person-Umwelt-Kongruenz vor; ergreift dieselbe Person (R) einen intellektuell-forschenden (I) oder konventionellen (C) Beruf, dann ergibt sich eine mittlere und bei einem künstlerisch-sprachlichen (A) oder unternehmerischen (E) Beruf niedrige Kongruenz. Ergreift ein R-Typ einen sozialen Beruf (S), so handelt es sich um eine inkongruente Wahl. (S. 11)

2.4.5 Problemfelder eigenschaftsorientierter Verfahren

Schuler (2014) macht auf Abwägungsüberlegungen die beim Einsatz von eignungsdiagnostischen Verfahren in der Praxis auftreten aufmerksam. Es stehen eine Vielzahl von Verfahren zur Verfügung um einen möglichst guten Überblick über das Fähigkeits- und Eigenschaftsprofil einer Person zu erhalten. Dieser Überblick geht allerdings auf Kosten der Detailtiefe, da spezifische Einzelinformationen aufgrund des eingeschränkten zeitlichen Rahmens oder wegen anderer Restriktionen nicht erfasst werden können. Wenn andererseits bei einer eignungsdiagnostischen Untersuchung der Schwerpunkt auf einer Detaillierung liegt, geht damit zwangsläufig die Breite des Überblicks über das allgemeine Fähigkeits- und Eigenschaftsprofil verloren. Diese Abwägungsproblematik wurde von Cronbach als „Bandwidth-fidelity“-Dilemma („Bandbreiten-Detailtiefen-Dilemma“) bezeichnet. Ausgangspunkt der „Bandwidth-fidelity-Überlegungen“ stellt immer die Beziehung zwischen dem eignungsdiagnostischen Prädiktor und dem vorherzusagenden Berufserfolgskriterium dar. Wenn ein eng umgrenztes Kriterium vorhergesagt werden soll, so muss auch ein spezifischer Prädiktor eingesetzt werden. Ist hingegen ein allgemeines Globalmaß das Kriterium, so muss auch der Prädiktor ein vergleichbares Abstraktionsniveau aufweisen. Das „Bandwidth-fidelity“-Dilemma wird von dieser Warte aus also durch eine systematische Vorgehensweise aufgelöst. Erst nach einer Konkretisierung des gewünschten Zielkriteriums erfolgt die Festlegung auf einen adäquaten, vom Abstraktionsniveau her zum Kriterium symmetrischen Prädiktor (Schuler, 2014).

Eine Befürchtung bei Anwendung von eignungsdiagnostischen Verfahren, welche auch Schuler (2014) erwähnt, ist die Gefahr der systematischen Verzerrung der Selbstdarstellung. Bei Leistungstests wie auch bei simulationsorientierten Verfahren stellt dies insofern kein Problem dar, als sie darauf ausgelegt sind maximales Verhalten, also Bestleistungen zu erfassen. Bei eigenschaftsorientierten Verfahren sowie bei den Biografie bezogenen Methoden, Fragebogen und Interviews, wird dagegen angenommen, ihre auf typisches Verhalten fokussierende Messtechnik würde zur verzerrten Erfassung der wahren Werte führen (Schuler, 2014).

Dieses sozial erwünschte Antwortverhalten stellt allerdings ein deutlich geringeres Problem dar als häufig angenommen. Ones und Viswesvaran (1998) konnten an einer aussagekräftigen Stichprobe aufzeigen, dass sich eine Auspartialisierung von Sozialer Erwünschtheit aus der Korrelation zwischen Persönlichkeitswerten und Berufserfolg nicht auswirkt. Das heißt, die Validität von Persönlichkeitsmassen wird durchschnittlich nicht durch die Selbstdarstellung beeinträchtigt.

2.4.6 Interview

Das Interview ist ein Prozess der gegenseitigen Steuerung und Urteilsbildung, geprägt durch Absichten und Erwartungen bei welchem sich die Gesprächspartner gegenseitig beeinflussen. Dabei gilt die Grundannahme, dass das vergangene Verhalten der beste Prädiktor für zukünftiges Verhalten ist. Interviews verfügen über eine hohe Augenscheinvalidität und sind gut akzeptiert. Sie sind interaktive, intuitiv, flexibel und bieten vielseitige diagnostische Möglichkeiten, sind aber auch anfällig für Verzerrungen und Verfälschungseffekte (Schuler, 2014).

Diverse Mechanismen in der Informationsverarbeitung tragen dazu bei, dass Interviewer Verfälschungseffekten unterliegen können. Dazu gehören die Übergewichtung des ersten Eindrucks sowie negativer Emotionen, das Unterschätzen oder ganz Vernachlässigen von relevanten positiven Informationen. Bestätigungstendenzen können dazu führen, dass im Sinne der selbsterfüllenden Prophezeiung die eigenen Vorannahmen bestätigt werden. Sympathieeffekte wie auch der Halo-Effekt haben zur Folge, dass emotionale Bewertungen die Leistungsbeurteilung verzerren. Der Massstabeffekt beschreibt die Tendenz vergleichende Beurteilungen vorzunehmen wobei es zu Fehlurteilen durch Vergleich mit ungeeigneten Bezugsgruppen kommt (Schuler, 2014).

Interviews unterscheiden sich in ihrem Grad der Strukturierung. Sie reichen von unstrukturiert über teilstrukturiert bis hin zu vollstrukturiert. Strukturierte Interviews weisen bessere psychometrische Gütekriterien auf und können Verzerrungen und Verfälschungseffekten entgegen wirken (Schuler, 2014).

Kriterienbezogene Validität für berufliche Leistung			
Interviewtyp	k	N	p
Unstrukturiert	39	9'330	.33
Strukturiert	106	12'847	.44

Tabelle 1: Kriterienbezogene Validität für berufliche Leistung. Quelle: McDaniel et al., (1994).

Interrater-Reliabilität			
Strukturierungsgrad	k	N	M IRR
Sehr niedrig	11	242	.36
Niedrig	21	2342	.51
Mittel	28	3122	.65
Hoch	36	23'997	.79
Sehr hoch	29	2'543	.76

Tabelle 2: Interrater-Reliabilität. Quelle: Huffcutt et al., (2013).

2.5 Klinische vs. statistische Urteilsbildung

Beauducel und Leue (2014) schreiben, dass sich bereits in den 1920er Jahren zwei gegenläufige Positionen zur Urteilsbildung herausstellten. Als wesentliche Strategien der Informationsverarbeitung und Urteilsbildung werden die klinische Urteilsbildung und die statistische Urteilsbildung differenziert.

Bei der klinischen Urteilsbildung werden Informationen auf der Basis subjektiver, intuitiver und informeller Einschätzungen zu einem diagnostischen Urteil integriert. Ihr zugrunde liegen nicht psychometrische Daten. Diese bestehen aus Eindrücken oder Interpretationen des Diagnostikers die anhand von Äusserungen des Probanden im Gespräch oder im Interview gewonnen werden. Statistische Urteilsbildung basiert auf festen Regeln, Formeln und Algorithmen die zu einem diagnostischen Urteil integriert werden. Sie basieren auf psychometrischen Daten welche anhand von Persönlichkeitsinventaren oder Leistungstests erhoben werden, bei welchen eindeutige Instruktionen zur Durchführung, Auswertung und Interpretation bestehen (Beauducel & Leue, 2014).

Meehl (1954) äusserte Kritik an der diagnostischen Urteilsbildung, wenn Daten erhoben werden, für die keine hinreichenden Informationen über Objektivität und Reliabilität vorliegen und die nach einem nicht transparenten Vorgehen zu einer diagnostischen Entscheidung zusammen gefügt werden. Es geht bei der Art der Datenerhebung nach Meehl vor allem um die Frage, ob Durchführungs-, Auswertungs- und Interpretationsobjektivität für die erhobenen Daten sichergestellt werden können und ob die erhobenen Prädiktoren hinreichend reliabel sind.

In Studien konnte aufgezeigt werden, dass diagnostische Einschätzungen auf der Basis statistischer Urteilsbildung genauso gut wenn nicht noch besser sind, als Einschätzungen auf Basis klinischer Urteilsbildung (Meehl, 1954).

Die Auswertungen der Metaanalyse von Kuncel et al. (2013) bestätigen die Resultate aus Meehl's Studie. Obwohl in der psychologischen Praxis eine klare Präferenz für klinisches Vorgehen besteht, konnte aufgezeigt werden, dass statistische Urteile in der Personalauswahl aussagekräftiger sind. Wenn es um Aussagen über Arbeits-, Trainings- oder akademische Leistung geht sind statistische Urteile sogar deutlich überlegen. Da in der Praxis allerdings zahlreiche Vorurteile gegenüber dem rechnerischen Ansatz bestehen und dieser unter Umständen spezifische Besonderheiten des Einzelfalls nicht berücksichtigt, ist es unwahrscheinlich, dass sich der klinische Ansatz eins zu eins austauschen lässt. Eine mögliche Alternative wäre ein Kompromiss, bestehend aus einem statistischen Urteil mit klinischem Sachverstand (Kuncel et al., 2013).

Vergleich statistisches versus klinisches Urteil		
Kriterium	statistisches Urteil	klinisches Urteil
Arbeitsleistung	.44	.28
Trainingsleistung	.31	.16
Akademische Leistung	.58	.48

Tabelle 3: Vergleich statistisches versus klinisches Urteil. Quelle: Kuncel et al., (2013).

Benedikt Hell von der Fachhochschule Nordwestschweiz hat in seinem Artikel „Bauchgefühl versus Verstand“ mehrere plausible Antworten die erklären können, wieso intuitive Urteile weniger zuverlässig sind.

Zum einen vertrauen intuitiv Urteilende zu sehr auf einige wenige, meist deutlich herausstechende Informationen. Weiterhin sind solche Urteile oft intransparent, denn es werden fallweise einzelne Bewerberaspekte wie Berufserfahrung, Kenntnisse, Fähigkeiten oder Persönlichkeitsmerkmale beobachtet und individuell bewertet, ohne dass zuvor festgelegt wurde, welche Aspekte auf welche Weise in die Bewertung eingehen sollen. Hierdurch wird gleich mehreren Fehlerquellen Tür und Tor geöffnet. Die Urteile sind stark situationsabhängig, denn sie beruhen auf einzelnen Verhaltensausschnitten, die in der betreffenden Situation gerade beobachtet werden konnten. Zudem ist die Gewichtung von Einzelaspekten subjektiv und fallabhängig. Und nicht zuletzt geht der Bezug zu den Stellenanforderungen verloren. Nicht zuletzt reduzieren Standardisierungsmaßnahmen systematische Fehler: Einer selektiven Wahrnehmung wird entgegengewirkt und die Informationsverarbeitung wird durch unterstützende Bewertungshilfen (zum Beispiel Checklisten für Interviewbeobachtungen) erleichtert (Hell, 2015).

Auch Tversky und Kahneman (1974) weisen auf Verzerrungstendenzen im diagnostischen Prozess hin. Sie beschreiben verschiedene Heuristiken und Verzerrungstendenzen, die bei der Präsentation von Informationen bzw. bei der Verarbeitung von Informationen relevant sein können und zumeist dazu dienen, die Komplexität von Informationen zu reduzieren.

Mögliche Heuristiken der Informationsverarbeitung, die beim diagnostischen Prozess auftreten können.
<ul style="list-style-type: none"> • Repräsentativitätsheuristik • Verfügbarkeitsheuristik • Simulationsheuristik • Ankerheuristik

Abbildung 7: Heuristiken der Informationsverarbeitung. Quelle: Beauducel & Leue (2014, zitiert nach Tversky & Kahneman, 1974).

Unter der Repräsentativitätsheuristik werden Urteilsfehler zusammengefasst, die darauf basieren, dass Menschen Ereignisse mit einer bestimmten Wahrscheinlichkeit einer bestimmten Kategorie zuordnen. So wurde in einer Untersuchung den Versuchspersonen die schriftliche Beschreibung einer Frau vorgelegt, welche sich aktiv für Frauenrechte und Emanzipation einsetzte. Danach wurden die Probanden gefragt, was denn nach dieser Beschreibung wahrscheinlicher sei, dass diese Frau eine Bankangestellte, oder eine Bankangestellte und Feministin sei. Die Mehrzahl der Versuchspersonen schätzte die Wahrscheinlichkeit, dass die Frau Bankangestellte und Feministin sei, wesentlich höher ein. Da die Wahrscheinlichkeit für das gleichzeitige Auftreten beider Ereignisse nicht größer sein kann, als die Wahrscheinlichkeit, dass eines der beiden Ereignisse alleine eintritt, ist diese Einschätzung falsch (Tversky & Kahneman, 1974).

Ähnlichkeit zur Repräsentationsheuristik hat die Verfügbarkeitsheuristik insofern, da auch hier Auftretenswahrscheinlichkeiten falsch eingeschätzt werden. Dabei werden Urteile über Ereignisse dadurch beeinflusst, wie schnell oder wie leicht sie aus dem Gedächtnis abgerufen werden können. Das bedeutet, dass wir Ereignisse als wahrscheinlicher einschätzen an die wir uns leicht erinnern, als solche, an welche wir uns nur schwer erinnern können (Tversky & Kahneman, 1974).

Im Vergleich zu der Verfügbarkeitsheuristik bezieht sich die Simulationsheuristik auf die Leichtigkeit mit der Ereignisse konstruiert werden können. Dabei wird eine Kausalität zwischen dem Verhalten einer Person und dem Auftreten eines Ereignisses geprüft. So ist es möglich, dass Stereotypen den Abruf und die Konstruktion von Ereignissen und die Beurteilung ihrer Auftretenswahrscheinlichkeit beeinflussen können (Beauducel & Leue, 2014).

Ankerheuristik bedeutet, dass bestimmte Informationen das Einschätzen einer Situation oder das Fällen einer Entscheidung massgeblich beeinflussen. Dabei spielt es keine Rolle, ob diese Information für die Situation relevant ist oder nicht. So überschätzten Probanden nach auswendig lernen ihrer letzten beiden Sozialversicherungsnummern die Anzahl der Ärzte in New York deutlich, obwohl die erste Zahl keine logische Verbindung zu der zweiten hat (Tversky & Kahneman, 1974).

Beauducel und Leue (2014) haben einige Vorschläge zur Erhöhung der diagnostischen Genauigkeit zusammen getragen.

Wie kann man die diagnostische Genauigkeit erhöhen?
<ul style="list-style-type: none">• Bei der Urteils- bzw. Entscheidungsfindung sollten Informationen verschiedener methodischer Zugänge integriert werden.• Es sollten Regeln und Hypothesen im diagnostischen Prozess generiert und geprüft werden.• Das Zustandekommen diagnostischer Entscheidungen sollte explizit und transparent gemacht werden.• Verzerrungstendenzen und Fehlerquellen bei der Verarbeitung von Informationen sollten in der Diagnostikausbildung bewusst gemacht werden und Strategien zum Erkennen der Verzerrungstendenzen sowie deren Reduktion geschult werden.

Abbildung 8: wie kann man die diagnostische Genauigkeit erhöhen. Quelle: Beauducel & Leue (2014, S. 61).

3 Methodik

Um in Erfahrung zu bringen, wie eignungsdiagnostische Instrumente in der Berufs-, Studien- und Laufbahnberatung eingesetzt werden, ist es unerlässlich die Beraterinnen und Berater zu ihrem Verhalten zu befragen. In einem Experteninterview mit Herrn Avramakis, welcher als Fachspezialist Stabsstelle Entwicklung bei der ask! in Aarau arbeitet und selbst langjährige Erfahrung als Berufs-, Studien- und Laufbahnberater hat, wurde ein Konzept für die Befragung der Mitarbeiter erstellt. Da aus früheren Projekten bekannt ist, dass die Beraterinnen und Berater eher gewillt sind einen kurzen Fragebogen auszufüllen, als sich für Interviews zur Verfügung zu stellen, fiel die Entscheidung wie Daten gewonnen werden sollen nicht schwer. Die Herausforderung bestand darin, mit wenigen Fragen genügend qualitative und quantitative Daten zu erheben um daraus repräsentative Schlussfolgerungen treffen zu können. Im Gespräch mit Herrn Avramakis wurde offene Frage formuliert, die sich jeweils auf eine spezifische Beratung beziehen. Um genügend Informationen zu erhalten, war geplant einige Mitarbeiter der ask! auszuwählen und zu bitten, jeweils fünf Fälle aus ihrer Akten auszusuchen und für jeden einzelnen Fall den Fragebogen auszufüllen, was zu einem $n = 50$ geführt hätte.

Der Fragebogen wurde im Unipark Tool erstellt und Herrn Avramakis, welcher sich für den Pretest zur Verfügung stellte, für eine erste Beantwortung zugänglich gemacht. Dabei fiel auf, dass mit den wenigen offenen Fragen die Gefahr bestand zu wenige Informationen zu generieren. Deshalb wurde der Fragebogen um eine Matrixfrage erweitert, was den Vorteil hatte, dass nun auch leichter und zuverlässiger quantitative Informationen gewonnen werden konnten. Durch die Matrixfrage musste allerdings auch das Untersuchungsdesign verändert werden, da diese Art Fragen nicht Fallspezifisch formuliert werden konnten. In weiteren Gesprächen wurden zusätzliche offene wie auch Matrixfragen erstellt. Grundlage für die jeweiligen Fragen lieferten zum einen die Berufserfahrung von Herrn Avramakis, die theoretischen Grundlagen, wie auch Empfehlungen von Dr. Katja Pässler, welche an der Fachhochschule Nordwestschweiz als Psychologiedozentin im Bereich angewandte Eignungsdiagnostik tätig ist. Der endgültige Fragebogen enthielt insgesamt sechs offene Fragen, von welchen die letzte den befragten Personen die Möglichkeit gab, wichtige Informationen, welche nicht abgefragt wurden, anzugeben. Zusätzlich sind drei Matrixfragen integriert, welche anhand von einer vierstufigen Likert Skala beantwortet werden können. Es wurde bewusst auf eine Fluchtkategorie in Form von einer fünfstufigen Skala verzichtet, da es sich bei den Befragten um Fachleute handelt, welche eindeutig Stellung zu ihrem Arbeitsalltag und den gestellten Fragen beziehen können sollten. Dichotom konnte nur eine zusätzliche Frage beantwortet werden, welche sich spezifisch auf den Unterschied von kostenpflichtigen und unentgeltlichen Beratungen bezieht,

wobei auch hier ein Textfeld für zusätzliche Aussagen angefügt wurde. Auf Pflichtfragen wurde bewusst verzichtet, lediglich beim Überspringer der Fragen erscheint eine Meldung, die bestätigt werden muss, um eine Beantwortung zu umgehen. Der definitive Fragebogen ist im Anhang A aufgelistet, die Empfehlungen von Frau Pässler im Anhang B.

Die Designveränderungen führten dazu, dass nicht einzelne Mitarbeiter für die Befragung ausgesucht wurden, sondern alle Beraterinnen und Berater der ask! per Email angeschrieben und zur Fragebogen Beantwortung eingeladen wurden. So wurde pro Mitarbeiter ein Fragebogen ausgefüllt. Die Befragung war ursprünglich für drei Wochen geplant. Acht Tage nach der ersten Einladung wurde eine Erinnerung an alle Empfänger gesandt. Da die Umfrage in der Ferienzeit durchgeführt wurde, gingen zahlreichen Abwesenheitsmeldungen ein. Daher wurde beschlossen, die Umfragezeit um eine Woche zu verlängern. Einladungsemail wie auch Erinnerungsemail sind im Anhang C ersichtlich. Insgesamt wurden 48 Personen angeschrieben, von welchen 24 an der Umfrage teilgenommen haben, was einem $n = 24$ entspricht.

3.1 Auswertung der Befragung

Um die Antworten auf die offenen Fragen auszuwerten, wurde eine qualitative Inhaltsanalyse nach Mayring (2010) durchgeführt. Da bereits bei der Erstellung des Fragebogens umfassendes Vorwissen zur Verfügung stand und dadurch Hypothesen bezüglich des Forschungsgegenstandes in die Fragekonstruktion einfließen, konnten zur Auswertung der Antworten auf offene Fragen, deduktive Kategorien gebildet werden.

Um möglichst präzise Formulierungen der Kategorien zu erreichen und unnötige Überschneidungen zu vermeiden, wurde ein Kodierleitfaden erstellt. Das Textmaterial konnte in drei Kategorien unterteilt werden. Textstellen, welche keine zusätzlichen Informationen lieferten, wie z.B. „Ja“ oder „ich verstehe die Frage nicht“ wurden ausgegraut und für die Auswertung nicht beachtet. Die Aussagen, welche sich auf das Untersuchungsdesign bezogen, wurden in der kritischen Reflexion berücksichtigt. Textstellen welche zu den deduktiven Kategorien zugeordnet werden konnten wurden blau eingefärbt. Aus dem restlichen Textmaterial wurden soweit möglich induktive Kategorien gebildet und grün eingefärbt. Bei der Bildung der induktiven Kategorien ist speziell zu erwähnen, dass aufgrund der teilweise sehr kurzen Antworten das Paraphrasieren nicht immer sinnvoll umgesetzt werden konnte. Es kam vereinzelt zu Überschneidungen mit der Generalisierung. Antworten welche z.B. nur aus einem Wort bestanden konnten nicht paraphrasiert werden und wurden deshalb direkt übernommen. Dies war der Zusammenfassung zu Kategorien jedoch nicht abträglich. Das Kategoriensystem ist im Anhang D einsehbar.

Da auch sämtliche Matrixfragen als deduktive Kategorien verwendet wurden, konnten die Aussagen aus zwei Perspektiven ausgewertet werden. Zum einen wurden die Kategorien aufgrund der Häufigkeit des Vorkommens relativ gewichtet (Häufigkeitsanalyse), zum anderen die Zustimmung oder Ablehnung einer Kategorie mittels Matrixfragen ermittelt. Bei der Gegenüberstellung von quantitativen und qualitativen Daten wurden zusätzlich Erklärungen einzelner Textteile durch Kontext und weitergehende Interpretation durchgeführt (Explikation). Daraus ergaben sich aussagekräftigere Interpretationen. Die Auswertung der Matrixfragen ist im Anhang E ersichtlich.

4 Ergebnisse

4.1 Interpretation der Ergebnisse (deduktiv)

In diesem Kapitel werden die quantitativen Ergebnisse aus den Matrixfragen mit den deduktiven Kategorien verglichen und daraus Rückschlüsse gezogen.

Den höchsten Wert aller Matrixfragen erzielten die „Gütekriterien“ mit 84% Zustimmung auf die Frage wie wichtig folgende Aspekte sind. Interessanterweise wurde „Gütekriterien“ zur Testauswahl in den offenen Fragen kein einziges Mal genannt. Das könnte zum einen daran liegen, dass die Testverfahren welche der ask! zur Verfügung stehen, bereits vorselektiert sind. Die Mitarbeiter können sich somit darauf verlassen, dass alle Testverfahren, welche sie verwenden, die Qualitätsanforderungen erfüllen. Andererseits ist es ebenfalls denkbar, dass die Priorität von Gütekriterien selbstverständlich ist und die Mitarbeiter es daher nicht für nötig empfunden haben dies in den offenen Fragen zu thematisieren. Unabhängig davon welche der beiden Vermutungen zutrifft, sind Testgütekriterien als einer der wichtigsten Faktoren erkannt.

Die eigene Erfahrung und Routine mit Testverfahren wurde mit 82% Zustimmung als wichtiger Aspekt bei der Testauswahl bestätigt, bei den offenen Fragen aber nur zwei Mal genannt. Daraus lässt sich ableiten, dass die Tests nicht aufgrund der eigenen Erfahrung ausgesucht werden, aber grundsätzlich Erfahrung und Routine für den Umgang mit Tests und deren Auswertung, wie auch Interpretation derselben wichtig sind. Gestützt wird diese Annahme durch die Aussage: „ je mehr Erfahrung ich mit den von mir hauptsächlich eingesetzten Verfahren gesammelt habe, desto mehr kann ich aus ihnen herausholen“.

Ebenfalls mit 82% Zustimmung wurden die sprachlichen Anforderungen eines Tests bei dessen Auswahl bestätigt. Mit fünf Nennungen bei den offenen Fragen, was einer kleinen bis mittleren Ausprägung entspricht, besteht auch hier eine gewisse Diskrepanz. Eine mögliche Erklärung wäre, dass die prozentuale Zustimmung deshalb so hoch ist, weil das Verstehen von Testfragen und Testauswertungen eine Grundvoraussetzung für die Vorhersagekraft von Tests darstellt. Die wenigen Nennungen könnten ein Hinweis sein, dass sprachliche Schwierigkeiten nicht sehr oft auftreten und deshalb nicht so präsent sind, oder sprachliche Anforderungen von den Beraterinnen und Beratern durch andere Kategorien mitabgedeckt sind, wie zum Beispiel die Voraussetzungen der Probanden.

Die „online-Verfügbarkeit“ erreichte mit 81% und 8 Nennungen in beiden Befragungen den jeweils dritthöchsten Wert. Durch diese Kongruenz wird offensichtlich, dass die online-Verfügbarkeit von Tests als sehr wichtig für den Arbeitsalltag angesehen wird. Es lässt sich jedoch durch einzelne

Aussagen noch mehr ableiten. Bei der Frage, ob bei kostenpflichtigen Beratungen auch Tests eingesetzt werden, war eine Antwort, sie müssen online zur Verfügung stehen. Da online-Tests vom Kunden ausserhalb der Beratung selbständig bearbeitet werden können, ist es gut möglich, dass die „online-Verfügbarkeit“ sich auf Testökonomie bezieht. Untermauert wird dies durch die Tatsache, dass die Kategorie „Länge und Dauer von Tests“ die am häufigsten genannte deduktive Kategorie darstellt. Diese wird später noch ausführlicher beschrieben.

Vertrauen in die klinische Vorgehensweise stellt mit sechs Nennungen eine mittlere Ausgeprägte Kategorie dar, die durch mehrere Matrixfragen bestätigt wurde. Auf die Frage was getan werden müsste, wenn die Beratende Person, der Klient oder die Klientin beim Ergebnis eines Tests stutzig werden, stimmten 74% zu, dass die Testauswertung überprüft werden sollte, um mögliche Fehler auszuschliessen. Bei der Aussage, Tests würden hauptsächlich das bestätigen, was ich im Vorfeld vermutet habe, war die Zustimmung mit 56% zu 44% minimal positiv. Beide Matrixfragen sind skeptisch zu betrachten. Die minimale positive Zustimmung von gerade mal acht Prozent ist im Vergleich zu anderen Kategorien und in Anbetracht der kleinen Stichprobe von 37 Antworten nicht sehr aussagekräftig. Die 74% Zustimmung, dass die Testauswertung überprüft werden sollte lässt isoliert betrachtet vermuten, dass mehr Vertrauen in klinische Vorgehensweisen besteht als in Tests. Das relativiert sich jedoch in Anbetracht, dass ebenfalls 74% zustimmten, die Aussagen des Klienten oder der Klientinnen zu überprüfen. 78% stimmten sogar zu, den Eindruck den die Beratenden über den Klienten oder die Klientin gewonnen haben zu überprüfen. Auch die dritte Matrixfrage: „Tests sind dann sinnvoll, wenn ich aufgrund meiner Erfahrung Klienten/innen nicht gut genug einschätzen kann“ welche darauf abzielte, nur Tests zu verwenden wenn klinisches Vorgehen als unzureichend erachtet wird, wurde mit 58% nicht-Zustimmung nicht bestätigt. Die Letzte Matrixfrage, welche darauf abzielte, dass Erfahrung stärker gewichtet wird als Tests, „Tests unterstützen bei der Einschätzung von Klienten/innen, sind jedoch weniger wichtig als langjährige Erfahrung“, konnte mit 48% Zustimmung den Verdacht aus der Theorie nicht bestätigen.

Die Länge und Dauer der Tests empfanden nur 66% der Befragten als wichtig. Dies ist insofern überraschen, da diese Kategorie mit 15 Nennungen bei den offenen Fragen am meisten vorkam. Unabhängig von der im Verhältnis geringen Zustimmung, kann aufgrund der einzelnen Antworten abgeleitet werden, dass Zeitdruck in den Beratungsgesprächen in den letzten Jahren zugenommen hat und für die meisten Berater und Beraterinnen ein wichtiger Faktor zu sein scheint. Mit dem Hintergrundwissen, dass ab Januar 2016 nicht mehr alle Beratungen kostenlos sind, lässt sich vermuten, dass ökonomische Gedanken zunehmend Einfluss auf die Beratungssituation nehmen werden.

Die zweithäufigste genannte Kategorie ist „Vertrauen in Testverfahren“ mit insgesamt 14 Nennungen. Auch wenn sie mit 48 % zu 51% nicht durch die Zustimmung zur Frage, „Tests liefern Informationen, die ich durch kein anderes Verfahren aufdecken kann“, bestätigt werden kann, zeigen sich keine Anzeichen dafür, dass die Beraterinnen und Beratern den Tests nicht vertrauen würden. Die Aussage, „selbst ein langes Gespräch zeigt meiner Erfahrung nach jedoch nie alle Aspekte auf, die mit einem Test in kurzer Zeit erhoben werden können“, untermauert das Vertrauen in Testverfahren. Das in der Theorie oftmals beschriebene Misstrauen gegenüber statistischen Verfahren konnte in der Stichprobe nicht bestätigt werden. Dies überrascht nicht nur vor dem theoretischen Hintergrund. Eine weitere mit acht Nennungen überdurchschnittlich oft vertretene Kategorie ist „Skepsis gegenüber Testverfahren“, welche eigentlich als Gegenstück zu „Vertrauen in Testverfahren“ in den Fragebogen aufgenommen wurde. Es stellt sich die Frage, wie es möglich ist, dass zwei sich gegenüberstehende Kategorien stark vertreten sind. Bei der Matrixfrage, was zu tun ist wenn ein Testresultat stutzig macht, wurde der Antwort „der Test sollte wiederholt werden“ mit 92% nicht zugestimmt. Das ist die höchste negative Ausprägung im Fragebogen und zeigt deutlich auf, dass Tests als funktionierend wahrgenommen werden. Gemäss Aussagen wird die Validierung von Tests im Gespräch mit den Probanden und Probandinnen durchgeführt und somit die Aussagekraft des Tests für die einzelne Person relativiert.

Die Kategorie „Test als Absicherung“ wurde sechs Mal genannt und verfügt somit über eine mittlere Ausprägung. Sie konnte jedoch durch keine der beiden Matrixfragen deutlich bestätigt werden. „Ich verwende Tests vor allem dann, wenn ich mir bei einem/r Klienten/in unsicher bin“ wurde mit 67% Nicht-Zustimmung nicht bestätigt und „Durch Tests kann ich mich und meine Einschätzung absichern“ erreichte eine Zustimmung von 63%. Eine genauere Betrachtung der Textstellen zeigt einen wesentlichen Unterschied zwischen den Antworten auf offene Fragen und Antwortmöglichkeiten der Matrixfragen auf. Bei den offenen Fragen wurde ein Test als Absicherung für die Probanden genannt, nicht aber für die Beratenden. Bei den Matrixfragen wurde nur die Absicherung für die Beratenden abgefragt. Somit stellt ein Testresultat schon eine Absicherung der eigenen Meinung dar, allerdings nur für die Probanden.

Die beiden Kategorien „Wichtigkeit von Kosten der Testverfahren“ und „Empfehlungen von Arbeitskollegen“ erreichten eine Zustimmung von 60% und von 61%. Die erste der beiden Kategorien wurde in den Antworten einmal genannt die zweite gar nicht. Somit scheinen beide Kategorien keine wichtige Rolle bei der Testauswahl darzustellen.

Die Empfehlung aus der Theorie, um die Kontroverse statistische vs. klinische Urteilsbildung zu umgehen, nämlich ein Kompromiss bestehend aus einem statistischen Urteil mit klinischem Sachverstand zu verfolgen, scheint bei den Mitarbeitern nicht gut anzukommen. Einerseits wurde die Kategorie bei den offenen Fragen nicht genannt, andererseits sind bei der Matrixfrage „Klienten/innen anhand von Tests zu beurteilen wäre grundsätzlich immer sinnvoll“ mit 71% Nicht-Zustimmung als Vorschlag abgelehnt. Auch die zweite Matrixfrage „Ideal zur Informationsgewinnung ist es, zuerst Tests durchzuführen und diese dann durch individuelle Fragen zu ergänzen“ wurde mit nur 48% Zustimmung als Handlungsempfehlung ebenfalls nicht angenommen.

„Komfort bei der Auswertung“ erhielt eine Zustimmung von 72%, obwohl die Kategorie nur zwei Mal genannt wurde. Diese Diskrepanz könnte durch die Kategorie „online-Verfügbarkeit“ erklärt werden. Da online-Tests leicht in der Auswertung sind, ist es denkbar, dass die starke Zustimmung zum „Komfort bei der Auswertung“ durch die „online-Verfügbarkeit“ repräsentiert ist.

„Das Layout als wichtiges Kriterium bei der Testauswahl“ erhielt eine Zustimmung von 60%. Genannt wurde die Kategorie fünf Mal was einem kleinen bis mittleren Wert entspricht. Aufgrund einzelner Aussagen wird auch in dieser Kategorie sichtbar, dass das Layout für die Beraterinnen und Berater weniger wichtig zu sein scheint als für die Probanden. Vor allem die Testauswertung sollte für die Probanden leicht verständlich dargestellt sein.

4.2 Interpretation induktive Kategorien

In diesem zweiten Teil der Ergebnisse werden die aus den verbliebenen Teststücken gebildeten induktiven Kategorien interpretiert.

Insgesamt zweiundzwanzig Textstellen lassen sich zu der Kategorie „Testauswahl aufgrund der Fragestellung“ zusammenfassen. Dadurch wird offensichtlich, dass sich die Beraterinnen und Berater differenziert mit den Tests auseinandersetzen und die Passung von Tests zu der gestellten Situation oberste Priorität hat.

Aus elf Textstellen wurden die Kategorie „Aufzeigen oder Spiegelung eigen- und Fremdbild“ gebildet. Dabei wird thematisiert, dass Tests unabhängig davon ob sie zu einem erwarteten Ergebnis oder zu einem unerwarteten geführt haben gut geeignet sind um entweder durch ein Fremdbild das Eigenbild zu konkretisieren, aber auch Diskrepanzen zwischen Eigen- und Fremdbild zu thematisieren. Diese Kategorie hat eine gewisse Nähe zu der deduktiven Kategorie „Test als Absicherung“, da beide Kategorien die Gegenüberstellung von Eigen- und Fremdbild beschreiben.

Zwölf Textstellen ergeben die Kategorie „Voraussetzungen der Teilnehmer bestimmen die Testauswahl“. Auch in diesem Fall wird eine differenzierte Auseinandersetzung mit den Tests und der Ausgangslage des Klienten oder der Klientin offensichtlich. Hierbei wird noch etwas genauer ausgesagt, welche Aspekte im Zentrum stehen können. Das Alter könnte im Zusammenhang mit den Normstichproben an welchen die jeweiligen Tests erhoben wurden stehen. Leider wurden Voraussetzungen nicht detaillierter beschrieben. Damit bleibt unklar ob sich dies auch auf die Sprachkenntnisse beziehen oder auf andere Aspekte wie Jobsituation, Schulbildung usw.

„Testanwendung in erweitertem Themenbereichen“ besteht aus elf Textstellen. Diese Kategorie wurde aus Antworten auf die Frage „Zu welchen anderen Themen, abgesehen von Interessen, Persönlichkeit und Leistung / Fähigkeit, verwenden Sie Tests“ gebildet. Die Menge an Antworten verdeutlicht, dass auch ausserhalb der in der Theorie beschriebenen und geläufigen Themenbereichen Tests Verwendung finden. Dadurch wird auch die deduktive Kategorie „Vertrauen in Testverfahren“ ein weiteres Mal bestätigt.

Vier Teststellen bilden die Kategorie „Test als Gesprächsstoff, oder Gesprächsöffner“. Wenn auch nur wenig genannt, wird hier ein weiterer Nutzen von Tests aufgezeigt, welcher in der Literatur nicht erwähnt wurde. Dieser Nutzen wurde bereits im Expertengespräch thematisiert, später jedoch nicht als deduktive Kategorie festgelegt. Diese Kategorie als Matrixfrage zu verwenden wäre besonders hilfreich gewesen, um abzuschätzen wie stark alle Befragten dem Testgebrauch als Gesprächsförderung zustimmen würden.

Eine der beiden kleinsten Kategorien mit drei Teststellen ist „Testauswahl aufgrund von Vorgaben der Organisation“. Da die Vorgaben dem Autor zu wenig detailliert bekannt sind, ist es schwierig diese Kategorie einzuschätzen. Besonders die Tatsache, dass Vorgaben sehr wenig genannt wurden ist überraschend. Einige konkrete Aussagen wie „Ich wähle häufig umfassende, mehrschichtige Tests, wie der NST“ oder die Antwort auf die Frage „Setzen Sie Tests auch in kostenpflichtigen Beratungen ein und falls ja, gibt es Unterschiede zu den Tests die in unentgeltlichen Beratungen eingesetzt werden“ die mit „nur selten bei RAV-Beratungen“ beantwortet wurde, lassen eher nicht auf einheitliche Vorgaben schliessen. Denkbar ist natürlich auch, dass die Vorgaben nicht für alle Mitarbeitende gleich sind. Dies erleichtert jedoch die Einstufung dieser Kategorie nicht.

Die andere der beiden kleinsten Kategorien ist „Anwesenheit von Beratenden“. Diese kann Inhaltlich als Gegenstück zu „online-Verfügbarkeit“ angesehen werden. In beiden Fällen wird die Auswahl von Tests davon abhängig gemacht, ob sie online verfügbar sind, also von Klienten und Klientinnen selbst ausgefüllt werden können, oder aber ob der Test nur in Papierform vorliegt und die Beratenden somit für die Durchführung anwesend sein müssen. Es ist offensichtlich, dass auch hier ökonomische Gründe Einfluss nehmen.

Die letzte Kategorie präsentiert sich als Ausnahmefall. Insgesamt zehn noch nicht zugeordnete Antworten wurden als „sonstige“ zusammengefasst. Hierbei handelte es sich um Teststellen, die inhaltlich zusätzliche Informationen lieferten, somit also nicht ausgegraut werden sollten. Gleichzeitig liessen sie sich thematisch weder zu einer bestehenden noch zu neuen Kategorien zusammenfügen. Durch diesen Kompromiss sind die Informationen für Interpretationen präsent, ohne für jede einzelne Teststelle eine eigene Kategorie formulieren zu müssen.

4.3 Zusammenfassungen der aussagekräftigsten und somit wichtigsten Kategorien.

Um in der Vielfalt des Kategoriensystems einen besseren Überblick zu gewährleisten, werden in diesem Kapitel die wichtigsten und somit aussagekräftigsten Kategorien zusammengefasst und graphisch dargestellt.

Abbildung 9 zeigt die relevanten Kategorien welche auf die zunehmende Priorisierung der Testökonomie hinweist.

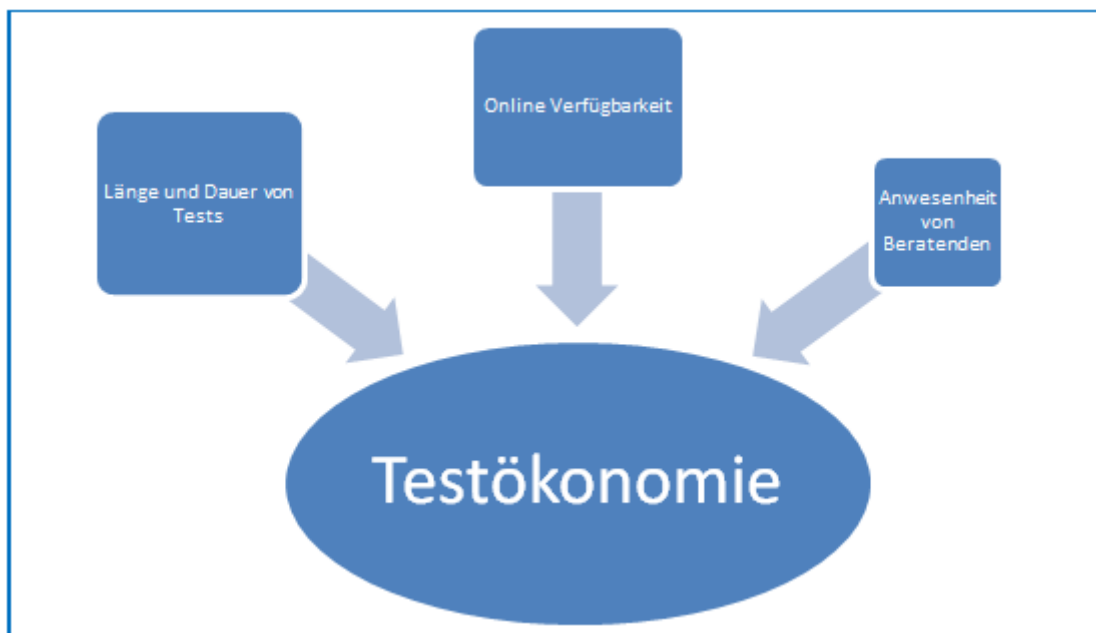


Abbildung 9: Zusammenfassung relevanter Kategorien für die Testökonomie.

In der Abbildung 10 sind die Kategorien die zur Auswahl von Testverfahren führen aufgelistet. Die Grösse der jeweiligen Kategorie stimmt mit der Priorität die sie einnehmen überein. Routine und Erfahrung wie auch Gütekriterien sind speziell markiert, da sie als wichtig empfunden werden jedoch auch als selbstverständlich.

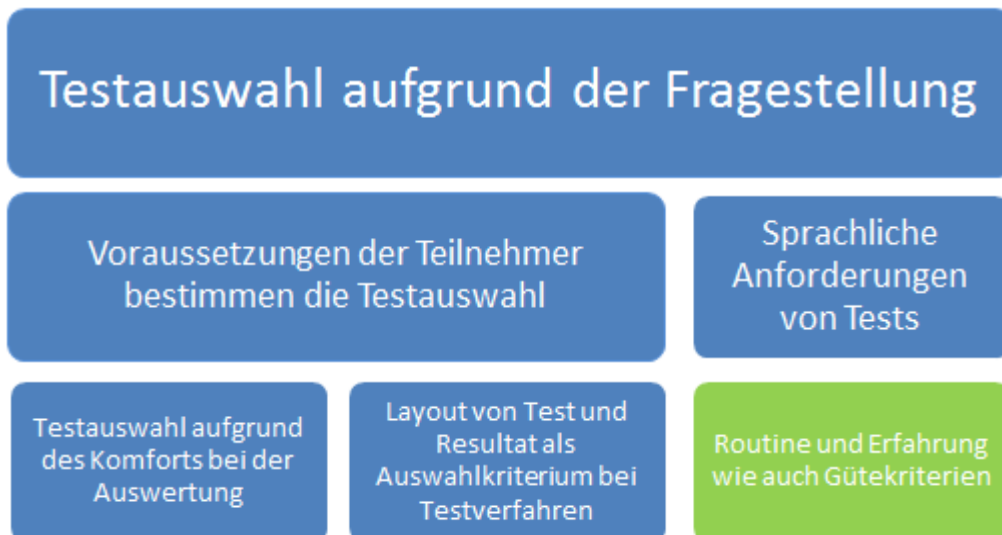


Abbildung 10: Relevante Kategorien für die Testauswahl.

In der Abbildung 11 werden die Pros. und Contras. für klinisches Vorgehen Stichwortartig dargestellt. Das gleichmässige Auftreten von beiden Kategorie typen unterstreicht ein reflektiertes Vertrauen in das klinische Vorgehen.



Abbildung 11: Vertrauen in klinisches Vorgehen.

Abbildung 12 zeigt eine Gegenüberstellung von zwei sich grundsätzlich widersprechenden Kategorien. Die vier-Felder Tafel lässt vermuten, welcher Umgang mit den Testverfahren praktiziert wird.

		Vertrauen in Testverfahren	
		Ja	Nein
Skepsis gegenüber Testverfahren	Ja	Differenzierter Umgang mit Testverfahren.	Tests werden nicht als angebrachtes Diagnoseinstrument wahrgenommen.
	Nein	Eventuell übermässiges Vertrauen in Tests, wenig kritische Auseinandersetzung.	Vermutlich werden Tests gar nicht oder nicht bewusst verwendet.

Abbildung 12: Verbindung von Skepsis und Vertrauen in Testverfahren.

5 Diskussion

5.1 Kritische Reflexion des methodischen Vorgehens

Bei der vorliegenden Arbeit wurde den theoretischen Grundlagen eine grosse Gewichtung beigemessen. Dies wurde als wichtig und angemessen erachtet, da aufgrund aktueller Untersuchungen bekannt ist, dass zwischen theoretischen Erkenntnissen und Empfehlungen und die in der Praxis verwendeten Vorgehensweisen grosse Diskrepanzen bestehen. Aufgrund des ausführlich gestalteten Theorieteils konnte eine Gegenüberstellung von Praxis versus Theorie gewährleistet werden, was massgeblich zu den Handlungsempfehlungen beigetragen hat.

Das Experteninterview und die damit verbundenen Nachbesprechungen haben sich als wegweisend für den Verlauf der Untersuchung herausgestellt. Der gemeinsam erarbeitete online Fragebogen hat sich als an die Situation angepasstes und als äusserst geeignetes Instrument für die Datenerhebung herausgestellt. Auch die Zeit die benötigt wurde um den Fragebogen auszufüllen konnte mit 15 – 20 Minuten sehr gut eingeschätzt werden, denn die durchschnittliche Beantwortung hat die 15 Minuten Grenze nie überschritten. Obwohl sich nicht alle Vorannahmen aus der Theorie wie auch aus Praxisempfehlungen bestätigt haben, so war die Praxiserfahrung des Experten äusserst hilfreich. Die offenen Fragen durch Matrixfragen zu ergänzen und somit Vorannahmen zu überprüfen hat sich für die Auswertung als überaus geeignet erwiesen.

Kritisch zu betrachten sind teilweise Formulierungen bei den offenen Fragen. So wurde vereinzelt bemängelt, dass die Fragen teilweise zu wenig konkret formuliert seien, so dass einige Antworten nicht aussagekräftig waren oder sogar ganz ausblieben. In diesem Zusammenhang zeigt sich eine weitere Schwäche. Den Pretest durch den Experten durchführen zu lassen, war einerseits zeitsparend und für die weitere Verbesserung des Fragebogens sehr hilfreich, andererseits besteht die Möglichkeit, dass eine dritte neutrale Person, welche bei der Fragenkonstruktion nicht beteiligt war, zu offen formulierte Fragen als solche erkannt und moniert hätte. Der Zeitpunkt für die Befragung war nicht ideal gewählt und hätte im Vorfeld besser geplant und vor die Ferienzeit verlegt werden können.

5.2 Fazit und Ausblick

Zusammenfassend kann festgehalten werden, dass die Berater und Beraterinnen der ask! die ihnen zur Verfügung stehenden eignungsdiagnostischen Instrumente differenziert und in Abwägung zur Fragestellung einsetzen. Die in der Theorie oftmals genannte Skepsis gegenüber eignungsdiagnostischen Instrumenten konnte nicht bestätigt werden. Skeptische Aussagen bezogen sich nicht auf Tests und deren Güte, sondern differenziert auf die Aussagekraft für die spezifische Situation. Zahlreiche Aussagen zeigen, entgegen der Annahme, dass den Testverfahren grosses Vertrauen entgegen gebracht wird. So wurden diverse Situationen genannt in denen weitere Tests wünschenswert wären. Dem in der Literatur vorgeschlagene Kompromiss bestehend aus einem statistischen Urteil mit klinischem Sachverstand wird wenig Zustimmung zu teil. Dies könnte zum einen daran liegen, dass die Berater und Beraterinnen nicht genügend über Vorteile des Kompromisses aufgeklärt sind. Andererseits könnte auch der Zeitfaktor entscheidenden Einfluss haben. In Anbetracht, dass Testökonomie einen zunehmend hohen Stellenwert einzunehmen scheint, wird es wohl auch in Zukunft nicht einfacher werden, Tests als Grundlage für Beratungssituationen zu verwenden. Das Vertrauen in klinische Vorgehensweisen scheint in vernünftigen Mass vorhanden zu sein. Zu erkennen ist dies daran, dass Tests in Gesprächen validiert werden, was als Zugeständnis zu Gesprächssituationen gewertet werden kann. Es wird jedoch eingeräumt, dass auch Vorannahmen und Eindrücke aus Interviews verfälscht sein können. Eine reflektierende Haltung gegenüber Interviews wird auch durch deren Absicherung durch Tests erkennbar. Wobei die Absicherung sich auf die Meinung der Probanden bezieht was einer Spiegelung des Eigen-, mit dem Fremdbild gleichkommt.

Aufgrund diverser Kategorien und Einzelaussagen wurde klar, dass die Testauswahl sich nicht alleine auf die Fragestellung fixieren kann. Voraussetzungen von Teilnehmern wie Alter, Sprachkenntnisse und Ausbildung sind starke Indikatoren für die Testauswahl, wie auch ein leicht verständliches Layout und gut interpretierbares Testresultat. Gütekriterien wie auch Erfahrung und Routine scheinen eine Selbstverständlichkeit im Umgang mit Testverfahren zu sein. Von der Theorie nicht berücksichtigt wird, dass Testverfahren auch als Gesprächseröffnung und als Gesprächsstoff Lieferant eingesetzt werden. Inwiefern die Testauswahl durch Vorgaben der Organisation eingeschränkt sind konnte nicht restlos geklärt werden.

Für eine weiterführende Untersuchung, wäre es interessant einigen Beratungsgesprächen beizuwohnen, um die gewonnenen Erkenntnisse überprüfen zu können. Dadurch würde ein eventuell sozial Erwünschtes Antwortverhalten, welches in dem Fragebogen nicht aufgedeckt werden konnte, zu Tage treten. Zudem wäre es sinnvoll, in Einzelinterviews mit den Beratern und Beraterinnen die Thematik vertieft zu besprechen. Dabei könnten unabhängig von dieser Arbeit neue Kategorien ersichtlich werden und zusätzlich das erstellte Kategoriensystem überprüft und falls notwendig angepasst und erweitert werden.

5.3 Handlungsempfehlungen

Da die Erfahrung und Routine im Umgang mit Testverfahren eine sehr hohe Zustimmung erhalten hat und als selbstverständlich klassifiziert wurde, sollte auf jeden Fall gewährleistet werden, dass Neueinsteiger in der Branche die Gelegenheit haben diverse Tests kennen und einsetzen zu lernen.

Falls Testverfahren als Grundlage für Beratungssituationen im Sinne des Kompromisses bestehend aus einem statistischen Urteil mit klinischem Sachverstand aus ökonomischen oder anderen Gründen nicht durchsetzbar sind, so wären standardisierte Interviews ein anerkanntes Verfahren um die Validität der Beratungen zu erhöhen. Durch die Strukturierung würden die Interviews transparenter, besser nachvollziehbar und vor allem objektiver was sich durch eine Messung der Interrater-Reliabilität erkennbar wäre. Möglicherweise könnten Interviews dann auch als Gesprächsöffner hilfreich sein.

Um dem zunehmendem Zeitdruck etwas entgegen zu setzen, wäre es sinnvoll auf Testverfahren zurückgreifen zu können, welche in kurzer Zeit durchgeführt und ausgewertet werden können. In Anbetracht, dass Intelligenz der beste Einzelprädiktor für Trainings- und Ausbildungserfolg darstellt, wäre eine genauere Betrachtung des Wonderlic Personnel Test empfehlenswert. Dieser Intelligenztest steht den Beratern und Beraterinnen der ask! bis anhin nicht zur Verfügung. Obwohl der Wonderlic Personnel Test nur allgemeine Intelligenz misst, und keinerlei detaillierte Rückschlüsse auf verbale-, numerische- oder figurale Intelligenz zulässt, ist er mit einer Testdauer von 12 Minuten sehr schnell durchführbar, wobei die Auswertung innerhalb von ein bis zwei Minuten zu bewerkstelligen sein müsste. Mit einer Retest Reliabilität von .94 Dobrill (1988) und deutlichen Korrelationen mit anderen kognitiven Leistungstests Hicks, Harrison und Engle (2015) sind auch die Gütekriterien zufriedenstellend.

7 Literaturverzeichnis

- Beauducel, A. & Leue, A. (2014). *Psychologische Diagnostik*. Göttingen : Hogrefe Verlag GmbH & Co. KG
- Bergmann, C. & Eder, F. (2005). *Allgemeiner Interessen-Struktur-Test mit Umwelt-Struktur-Test (AIST-R/UST-R)*. Göttingen : Beltz Test
- Dobrill, C. & Warner, H. (1988). Further studies of the Wonderlic Personnel Test as a brief measure of intelligence. *Journal of Consulting and Clinical Psychology*, 56, 145–147.
- Engelbrecht, W. (1994). Computerunterstützte berufsbezogene Testauswertung im Dienst der Berufsberatung. *Zeitschrift für Arbeits- und Organisationspsychologie*, 38, 175 –181.
- Häfeli, K., Kraft, U. & Schallberger, U. (1988). *Berufsausbildung und Persönlichkeitsentwicklung*. Bern : Huber
- Hell, B. (2015). Erfolgreiche Personalauswahl: Bauchgefühl versus Verstand. Verfügbar unter : <http://www.hrtoday.ch/article/erfolgreiche-personalauswahl-bauchgef-hl-versus-verstand> [16.05.2016].
- Huffcutt, A. I., Culbertson, S. S. & Weyhrauch, W. S. (2013). Employment Interview Reliability: New meta-analytic estimates by structure and format. *International Journal of Selection and Assessment*, 21, 264-276.
- Kuncel et al. (2013). Mechanical Versus Clinical Data Combination in Selection and Admissions Decisions: A Meta-Analysis. *Journal of Applied Psychology*, 98, 1060–1072.
- Mayring, P. (2010). *Qualitative Inhaltsanalyse : Grundlagen und Techniken*. Weinheim : Beltz
- McDaniel, M.A., Whetzel, D.L., Schmidt, F.L. & Maurer, S.D. (1994). The validity of employment interviews: A comprehensive review and meta-analysis. *Journal of applied psychology*, 79, 599- 616.
- Moosbrugger, H. & Kelava, A. (2012). *Testtheorie und Fragebogenkonstruktion*. Heidelberg : Springer-Verlag Berlin

Ones, D. S. & Viswesvaran, C. (1998). The effects of social desirability and faking on personality and integrity assessment for personnel selection. *Human Performance*, 11, 245-269.

Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and review of the evidence*. Minneapolis: University of Minnesota.

Rost, J. (1996). *Lehrbuch Testtheorie, Testkonstruktion*. Bern : Huber

Schallberger, U. (2000). Berufliche Tätigkeit als „Determinante“ interindividueller Differenzen. In M. Amelang (Hrsg.), *Enzyklopädie der Psychologie: Differentielle Psychologie. Band 4: Determinanten interindividueller Unterschiede* (S. 407 – 454). Göttingen : Hogrefe.

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, 262-274.

Schmidt-Atzert, L. & Amelang, M. (2012). *Psychologische Diagnostik*. Heidelberg: Springer-Verlag Berlin

Schuler, H. (2014). *Lehrbuch der Personalpsychologie*. Göttingen : Hogrefe

Seifert, K. H. (1977). *Handbuch der Berufspsychologie*. Göttingen : Hogrefe

Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty : Heuristics and biases. *Science*, 185, 1124-1131.

Wonderlic Inc. (1992). Wonderlic Personnel Test: Libertyville : Author. Validity comparison of the General Ability Measure for Adults with the Wonderlic Personnel Test. *North American Journal of Psychology*, 3, 173-182.

8 Tabellenverzeichnis

TABELLE 1: KRITERIENBEZOGENE VALITITÄT FÜR BERUFLICHE LEISTUNG. QUELLE: MCDANIEL ET AL., (1994).	24
TABELLE 2: INTERRATER-RELIABILITÄT. QUELLE: HUFFCUTT ET AL., (2013).	24
TABELLE 3: VERGLEICH STATISTISCHES VERSUS KLINISCHES URTEIL. QUELLE: KUNCEL ET AL., (2013).	26

9 Abbildungsverzeichnis

ABBILDUNG 1: DIE ANSÄTZE DER EIGNUNGSDIAGNOSTIK UND IHRE VALIDITÄTSLOGIK. QUELLE: SCHULER (2014, S. 153).	13
ABBILDUNG 2: WERTE DES ARMY GENERAL CLASSIFICATION TEST. QUELLE : SCHULER (2014, S. 159).	15
ABBILDUNG 3: ZUSAMMENFASSUNG METAANALYTISCHER BEFUNDLAGE ZUR VALIDITÄT EIGNUNGSDIAGNOSTISCHER VERFAHREN. QUELLE: SCHULER (2014, ZITIERT NACH SCHMIDT & HUNTER, 1998).	17
ABBILDUNG 4: CHARAKTERISIERUNG DER BIG FIVE. QUELLE: SCHULER (2014, ZITIERT NACH OSTENDORF & ANGLEITNER, 2004).	19
ABBILDUNG 5: HOLLANDS TYPENSHEMA. QUELLE: SCHULER (2014, ZITIERT NACH SEIFERT, 1977).	21
ABBILDUNG 6: HEXAGONALES MODELL ZUR BESTIMMUNG DER ÄHNLICHKEIT ZWISCHEN PERSÖNLICHKEITSTYPEN, UMWELTTYPEN UND DEREN BEZIEHUNGEN. QUELLE: SCHULDER (2014, ZITIERT NACH BERGMANN & EDER, 2005).	22
ABBILDUNG 7: HEURISTIKEN DER INFORMATIONSVERARBEITUNG. QUELLE: BEAUDUCEL & LEUE (2014, ZITIERT NACH TVERSKY & KAHNEMAN, 1974).	26
ABBILDUNG 8: WIE KANN MAN DIE DIAGNOSTISCHE GENAUIGKEIT ERHÖHEN. QUELLE: BEAUDUCEL & LEUE (2014, S. 61).	28
ABBILDUNG 9: ZUSAMMENFASSUNG RELEVANTER KATEGORIEN FÜR DIE TESTÖKONOMIE.	37
ABBILDUNG 10: RELEVANTE KATEGORIEN FÜR DIE TESTAUSWAHL.	38
ABBILDUNG 11: VERTRAUEN IN KLINISCHES VORGEHEN.	38
ABBILDUNG 12: VERBINDUNG VON SKEPSIS UND VERTRAUEN IN TESTVERFAHREN.	39

10 Anhang

Anhang A: Fragebogen

1 Anfang

Herzlich Willkommen zur Umfrage

Im Rahmen meiner Bachelor-Thesis in Zusammenarbeit mit "ask! - Beratungsdienste für Ausbildung und Beruf Aargau" möchte ich untersuchen, wie in der Berufs-, Studien und Laufbahnberatung eignungsdiagnostische Instrumente zur Urteilsbildung eingesetzt werden.

Um dies zu ermitteln bedarf es Ihrer professionellen Einschätzung.

Bitte nehmen Sie sich zur Beantwortung des Fragebogens 15-20 Minuten Zeit, Sie leisten damit einen entscheidenden Beitrag zum Gelingen dieser Untersuchung. Sie können die Befragung jederzeit unterbrechen und zu einem späteren Zeitpunkt fortfahren.

Ihre Angaben werden anonym ausgewertet und selbstverständlich vertraulich behandelt.

Vielen Dank für Ihre Teilnahme

Marco Schraner

2 Standardseite

Wenn Sie für eine Befragung einen Test (Interessen-, Persönlichkeits-, Leistungs- oder Fähigkeitstest) verwenden, nach welchen Kriterien wählen Sie diesen aus?

Die Frage kann auch stichwortartig beantwortet werden



Gibt es Besonderheiten bei bestimmten Typen von Tests?

3 Standardseite

Zu welchen anderen Themen, abgesehen von Interessen, Persönlichkeit und Leistung / Fähigkeit, verwenden Sie Tests?

4 Standardseite

Einige allgemeine Fragen zu eignungsdiagnostischen Tests. Wie stark stimmen Sie mit den Aussagen überein?

	stimme nicht zu	stimme eher nicht zu	stimme eher zu	stimme zu
Ich verwende Tests vor allem dann, wenn ich mir bei einem/r Klienten/in unsicher bin.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Klienten/innen anhand von Tests zu beurteilen wäre grundsätzlich immer sinnvoll.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tests sind dann sinnvoll, wenn ich aufgrund meiner Erfahrung Klienten/innen nicht gut genug einschätzen kann.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tests bestätigen hauptsächlich das, was ich im Vorfeld vermutet habe.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tests liefern Informationen, die ich durch kein anderes Verfahren aufdecken kann.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tests unterstützen bei der Einschätzung von Klienten/innen, sind jedoch weniger wichtig als langjährige Erfahrung.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Durch Tests kann ich mich und meine Einschätzung absichern.

Ideal zur Informationsgewinnung ist es, zuerst Tests durchzuführen und diese dann durch individuelle Fragen zu ergänzen.

5 Standardseite

Wenn Sie oder Ihr/e Klient/in bei dem Ergebnis eines Tests (Interessen-, Persönlichkeits-, Leistungs- oder Fähigkeitstest) "stutzig" werden, wie gehen Sie weiter vor?



6 Standardseite

Wenn ich oder der/die Klient/in bei dem Ergebnis eines Tests "stutzig" werden...

	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu
sollte die Testauswertung überprüft werden, um mögliche Fehler auszuschliessen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
sollte der Test wiederholt werden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
sollte ich meinen Eindruck über den/die Klienten/in überprüfen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
sollten die Aussagen des/r Klienten/in überprüft werden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

7 Standardseite

Nochmals kurz zurück zur Auswahl der Tests. Bitte geben Sie an, wie wichtig Ihnen folgende Aspekte sind.

	unwichtig	eher unwichtig	eher wichtig	wichtig
Die Länge und Dauer des Tests	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die eigene Erfahrung / Routine mit den Tests	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die sprachlichen Anforderungen der Tests	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Komfort bei der Auswertung	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Online Verfügbarkeit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Gütekriterien	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

- Das Layout
- Die Kosten für die jeweiligen Testverfahren
- Die Empfehlungen von Arbeitskollegen/innen

8 Standardseite

Gibt es Situationen in denen Sie gerne Testen würden aber kein geeigneter Test zur Verfügung steht?



Setzen Sie Tests auch in kostenpflichtigen Beratungen ein und falls ja, gibt es Unterschiede zu den Tests die in unentgeltlichen Beratungen eingesetzt werden?

- nein
- ja

9 Standardseite

Bitte geben Sie ihr Geschlecht an.

- weiblich
- männlich

Wie lange sind Sie als Berater/in tätig?

- weniger als 2 Jahre
- 2 bis 5 Jahre
- 6 bis 10 Jahre
- länger als 10 Jahre

10 Standardseite

Gibt es Informationen, die Sie mitteilen möchten, für die in der Umfrage kein Platz war?



11 Endseite

Vielen Dank, dass Sie an der Befragung teilgenommen haben.

Sie können nun mit anklicken des Buttons "Close window" die Umfrage beenden.

Falls der "Close window" Button nicht funktionieren sollte, können Sie das Fenster auch wie gewöhnlich oben rechts mit x schliessen.

Anhang B: Empfehlungen von Dr. Katja Pässler

Frage Student:

Guten Morgen Frau Pässler

Was meine Bachelorarbeit angeht, so habe ich diese Woche mein erstes Gespräch mit Herrn Avramakis vom ASK (Beratungsdienste für Ausbildung und Beruf Aargau) in Aarau geführt. Er ist der Durchführung meiner Arbeit sehr zugetan und wir haben etwas diskutiert wie wir vorgehen könnten. Meiner Erfahrung nach besteht in der Gesellschaft und selbst unter Fachleuten eine grosse Skepsis gegenüber standardisierten Messverfahren wie wir sie gerade in der Eignungsdiagnostik besprechen. Ich habe meine letzte Semesterarbeit über das Thema Intuition vs. Standardisierung geschrieben und würde dieses Thema gerne mit einflechten. Herr Avramakis kann sich gut vorstellen, dass ich untersuchen würde wie Diagnostische oder Eignungsdiagnostische Instrumente eingesetzt werden, welchen Stellenwert sie für die Mitarbeiter haben und inwiefern Intuition als Entscheidungshilfe genutzt wird. Ebenfalls hat er vorgeschlagen, ich könnte einen Onlinefragebogen an die Mitarbeiter versenden, da diese eher geneigt sind 10-15 Minuten zu investieren als sich einem 30 Minütigen oder eventuell noch längerem Interview zu stellen.

Selbstverständlich sollte meine Arbeit auch dem ASK einen Nutzen bringen, und so könnte ich mir vorstellen, dass meine Fragestellung beantworten soll wie Eignungsdiagnostische Instrumente gewinnbringend eingesetzt werden können. Ich wüsste gerne ob sie meine Idee für sinnvoll und durchführbar halten?

Freundliche Grüsse

Marco Schraner

Feedback Dozentin:

Lieber Herr Schraner

Ich finde Ihre Fragestellung sehr interessant. Es gibt bereits einige Studien, die den Einsatz von eignungsdiagnostischen Instrumenten im Kontext der Personalauswahl und -entwicklung bei Unternehmen untersuchen. Hier wurden eine Reihe von Faktoren identifiziert, nach denen Verfahren für den Einsatz ausgewählt wurden. Die Frage nach der prädiktiven Validität der Verfahren für den

Kontext Personalauswahl spielte dabei eine untergeordnete Rolle. Für Ihre Arbeit wäre es sicherlich interessant zu untersuchen, warum bestimmte Verfahren eingesetzt bzw. nicht eingesetzt werden. Bzw. wie Beratungspersonen ihren Verfahrenskanon auswählen (oftmals spielen dort Empfehlungen von Kollegen bzw. auch die Kosten der Verfahren eine Rolle). Spannend wäre sicher auch die Frage, wie Informationen aus Verfahren genutzt und kombiniert werden, u.a. im Sinne einer klinischen oder diagnostischen Urteilsbildung.

Freundliche Grüße

Katja Pässler

Anhang C: Einladungsemail und Erinnerungsemail für die Umfrage

Sehr geehrte Damen und Herren

Mein Name ist Marco Schraner, ich studiere an der FHNW in Olten Angewandte Psychologie in der Vertiefungsrichtung Arbeits-, Organisations- und Personalpsychologie (AOP).

Im Rahmen meiner Bachelorarbeit, in Zusammenarbeit mit Joannis Avramakis vom „ask! – Beratungsdienste für Ausbildung und Beruf Aargau“ möchte ich untersuchen, wie in der Berufs-, Studien und Laufbahnberatung

eignungsdiagnostische Instrumente eingesetzt werden.

Um dies zu ermitteln bedarf es Ihrer professionellen Einschätzung.

Bitte nehmen Sie sich 15-20 Minuten Zeit zur Beantwortung meiner Umfrage, Sie leisten damit einen entscheidenden Beitrag zum Gelingen dieser Untersuchung.

Sie können die Befragung jederzeit unterbrechen und zu einem späteren Zeitpunkt fortfahren.

Hier geht's zur Umfrage: http://ww3.unipark.de/uc/Umfrage_ask/

Bitte stellen Sie sicher, dass Ihr Browser cookies zulässt, damit Sie die Umfrage jederzeit unterbrechen können.

Wie lasse ich cookies zu im Internet Explorer: <http://windows.microsoft.com/de-ch/windows-vista/block-or-allow-cookies>

Wie lasse ich cookies zu im Firefox: <https://support.mozilla.org/de/kb/cookies-erlauben-und-ablehnen>

Für Fragen dürfen Sie mich gerne unter marco.schraner1@students.fhnw.ch kontaktieren.

Vielen Dank für Ihre Teilnahme

Freundliche Grüsse

Marco Schraner

Liebe Berufs-, Studien und Laufbahnberaterinnen und Berater

Vielen Dank für die bereits zahlreichen eingegangenen beantworteten Umfragen. Falls Sie noch nicht die Gelegenheit gehabt haben die Umfragen auszufüllen, möchte ich Sie daran erinnern, dass Sie noch bis am **Samstag, den 08. Mai** dafür Zeit haben.

Vielen Dank für Ihre Unterstützung.

Freundliche Grüsse

Marco Schraner

Sehr geschätzte Berufs-, Studien und Laufbahnberaterinnen und Berater

Ich bedanke mich ganz herzlich für Ihre zahlreiche Teilnahme an meiner Umfrage. Aufgrund der vielen „out of Office“ Rückmeldungen die ich erhalten habe, befürchte ich, dass einige von Ihnen bis anhin gar nicht die Chance hatten an der Umfrage teilzunehmen. Ich möchte allen Beraterinnen und Beratern die Möglichkeit geben teilzunehmen und freue mich über jede Rückmeldung. Deshalb habe ich die Umfragezeit bis zum **Sonntag den 15. Mai verlängert**.

Vielen Dank für Ihre Unterstützung.

Freundliche Grüsse

Marco Schraner

Anhang D Kategoriensystem

Deduktive Kategorienbildung

Kodierleitfaden

Kategorie	Definition	Ankerbeispiel	Kodierregeln
Skepsis gegenüber Testverfahren	Skeptisch, vorsichtig oder negative Äusserungen über Testverfahren.	Viele Tests überzeugen mich persönlich nicht, weil sie zu durchschaubar sind. Ich persönlich habe mich schon als Jugendlicher über Verfahren geärgert, bei denen ich schon im Vorherein wusste, was ich ankreuzen muss, damit mein angestrebtes Resultat herauskommt	
Test als Absicherung	Textstellen die sich auf Test dienen zur Absicherung der Einschätzung beziehen.	Wenn sich jemand zwischen Alternativen nicht entscheiden kann	Äusserungen wenn Tests zur Bestätigung von Einschätzungen verwendet werden oder wenn ohne Test gar keine Einschätzung zu Stande kommt.
Vertrauen in klinische Vorgehensweise	Textstellen welche die sich positiv auf die klinischen Urteilsbildungen beziehen.	Bilde Hypothesen aus Wissen und Erfahrungen der einzelnen Beratung und diskutiere diese mit dem/der Kunden/in	Dabei sind vor allem Informationen aus den Beratungsgesprächen gemeint, wie auch Validierung von Testergebnissen.
Zustimmung zu statistisches Vorgehen mit klinischem Sachverstand		Kein Ankerbeispiel vorhanden.	
Betonung von Länge und Dauer von Tests (Ökonomie)	Aussagen die die Testökonomie in den Mittelpunkt rücken.	Zeitlicher Aufwand	
Erfahrung und Routine mit Tests	Aussagen die die eigene Erfahrung mit Tests in den Mittelpunkt rücken.	Ja. Fehlende Erfahrung mit div. Tests.	
Sprachliche Anforderungen von Tests	Alle Textstellen die sich auf Sprachliche Anforderungen von Testverfahren beziehen.	Bei Migrantinnen/Migranten, die der deutschen Sprache (noch) nicht mächtig sind	
Online Verfügbarkeit von Tests	Textstellen welche die Online Verfügbarkeit	Online-Verfügbarkeit	

	thematisieren.		
Testauswahl aufgrund des Komforts bei der Auswertung	Textstellen die aufzeigen, dass eine komfortable Auswertung für die Berater oder die Kunden Vorteile bringt.	Auswertungshilfen	
Testauswahl aufgrund von Empfehlungen von Arbeitskolleginnen und Kollegen	Wenn bei Testauswahl die Empfehlungen von Arbeitskollegen erwähnt werden.	Kein Ankerbeispiel vorhanden.	
Gütekriterien als Grundlage der Testauswahl	Wenn Gütekriterien bei der Testauswahl erwähnt werden.	Kein Ankerbeispiel vorhanden.	
Kosten für die jeweiligen Testverfahren als Auswahlkriterium	Die Kosten (Nutzungsgebühren) von Tests werden erwähnt.	Kosten	Nur Äusserungen die sich auf die Kosten für die Tests beziehen. Andere ökologische Zusammenhänge wie Zeitdauer der Tests werden nicht berücksichtigt.
Testlayout als Auswahlkriterium bei Testverfahren	Textstellen die das Layout des Tests wie auch der Testresultate ansprechen.	Darstellung Ergebnis	
Vertrauen in Testverfahren	Textstellen die sich positiv auf Tests beziehen.	Auswertung überprüfen und/oder Beantwortung der Fragen zusammen anschauen (falls möglich) und dadurch Erklärungen fürs Ergebnis finden.	Äusserungen die Tests positiv bewerteten, oder in heiklen Situationen die Tests nicht in Frage stellen.

Deduktives Kategoriensystem

Kategorie	Stelle im Text
Skepsis gegenüber Testverfahren	2.Frage 1.Antwort / 2.Frage 8.Antwort / 2.Frage 15.Antwort / 2.Frage 17.Antwort / 2.Frage 19.Antwort / 5.Frage 15.Antwort / 5.Frage 22.Antwort / 5.Frage 6.Antwort /
Test als Absicherung	1.Frage 13.Antwort / 1.Frage 15.Antwort / 1.Frage 17.Antwort / 10.Frage 6.Antwort / 10.Frage 7.Antwort / 10.Frage 12.Antwort /
Vertrauen in klinische Vorgehensweise	5.Frage 7.Antwort / 5.Frage 8.Antwort / 5.Frage 9.Antwort / 5.Frage 16.Antwort / 5.Frage 18.Antwort / 5.Frage 24.Antwort / 5.Frage 4.Antwort / 5.Frage 17.Antwort / 5.Frage

	20.Antwort / 5.Frage 21.Antwort / 5.Frage 23.Antwort /
Zustimmung zu statistisches Vorgehen mit klinischem Sachverstand	Nicht explizit genannt
Betonung von Länge und Dauer von Tests (Ökonomie)	1.Frage 5.Antwort / 1. Frage 1.Antwort / 1.Frage 9.Antwort / 1.Frage 18.Antwort / 1.Frage 19.Antwort / 1.Frage 21.Antwort / 2.Frage 1.Antwort / 2.Frage 12.Antwort / 2.Frage 13.Frage / 2.Frage 18.Antwort / 8.Frage 7.Antwort / 10.Frage 10.Antwort / 1.Frage 23.Antwort / 1.Frage 4.Antwort / 1.Frage 4.Antwort /
Erfahrung und Routine mit Tests	8.Frage 3.Antwort / 8.Frage 16.Antwort /
Sprachliche Anforderungen von Tests	1.Frage 19.Antwort / 1.Frage 23.Antwort / 8.Frage 4.Antwort / 8.Frage 14.Antwort / 1.Frage 9.Antwort /
Online Verfügbarkeit von Tests	1.Frage 1.Antwort / 1.Frage 18.Antwort / 1.Frage 23.Antwort / 2.Frage 7.Antwort / 2.Frage 23. Antwort / 2. Frage 22.Antwort / 9.Frage 3.Antwort /
Testauswahl aufgrund des Komforts bei der Auswertung	1.Frage 10.Antwort / 1.Frage 23.Antwort /
Testauswahl aufgrund von Empfehlungen von Arbeitskolleginnen und Kollegen	Nicht genannt
Gütekriterien als Grundlage der Testauswahl	Nicht explizit genannt
Kosten für die jeweiligen Testverfahren als Auswahlkriterium	2.Frage 13.Antwort /
Layout von Test und Resultat als Auswahlkriterium bei Testverfahren	1.Frage 4.Antwort / 1.Frage 9.Antwort / 1.Frage 10.Antwort / 2.Frage 10.Antwort / 2.Frage 13.Antwort /
Vertrauen in Testverfahren	5.Frage 1.Antwort / 5.Frage 13.Antwort / 5.Frage 10.Antwort / 8.Frage 4.Antwort / 8.Frage 11.Antwort / 8.Frage 10.Antwort / 8.Frage 14.Antwort / 8.Frage 15.Antwort / 8.Frage 20.Antwort / 8.Frage 23.Antwort / 10.Frage 5.Antwort / 10.Frage 13.Antwort / 5.Frage 19.Antwort / 10.Frage 14.Antwort /

Induktives Kategoriensystem

Textstelle/Antwort	Zitat	Paraphrase	Generalisierung	Kategorie
1.Frage 1.Antwort	Fragestellung	Fragestellung	Tests werden aufgrund der Fragestellung die sich stellt ausgewählt.	Testauswahl aufgrund der Fragestellung
1.Frage 2.Antwort	Je nach Fragestellung der Beratung Beispiele: braucht der Kunde weitere	Je nach Fragestellung der Beratung		

	Berufsideen (Interessen), will er mehr über sich erfahren (Persönlichkeit) oder muss abgeklärt werden, ob jemand z.B. EBA			
1.Frage 3.Antwort	Je nach Fragestellung oder Untersuchungsgegenstand (z.B. geht es darum, mehr über sich selber zu erfahren? Das intellektuelle Potential abzubilden etc....)	Je nach Fragestellung oder Untersuchungsgegenstand		
1.Frage 4.Antwort	Fragestellung	Fragestellung		
1.Frage 5.Antwort	Nach Fragestellung in der Beratung	Nach Fragestellung in der Beratung		
1.Frage 6.Antwort	was ist die Fragestellung? was ist die Ausgangslage? was ist das Ziel?	was ist die Fragestellung? was ist die Ausgangslage?		
1.Frage 7.Antwort	Fragestellung	Fragestellung		
1.Frage 8.Antwort	ganz ausdrücklicher Wunsch d Kunden und Definition wozu es sein soll Was ist die Grundfrage / das Anliegen / das Beratungsziel?	Was ist die Grundfrage / das Anliegen		
1.Frage 9.Antwort	Fragestellung	Fragestellung		
1.Frage 10.Antwort	Fragestellung	Fragestellung		
1.Frage 11.Antwort	Nach der Fragestellung	Nach der Fragestellung		
1.Frage 12.Antwort	Art der Fragestellung	Art der Fragestellung		
1.Frage 14.Antwort	nach der Fragestellung, oft Kombination von zwei Tests	nach der Fragestellung		
1.Frage	Die Auswertung muss	Die Auswertung muss		

19.Antwort	zur Beantwortung der Fragestellung beitragen. Das Ergebnis fließt in die Beratung ein	zur Beantwortung der Fragestellung beitragen		
1.Frage 20.Antwort	Welche Fragestellung im Vordergrund steht	Welche Fragestellung im Vordergrund steht		
1.Frage 21.Antwort	Nach der Fragestellung.	Nach der Fragestellung.		
1.Frage 22.Antwort	Fragestellung	Fragestellung		
1.Frage 23.Antwort	Fragestellung bei unklar oder sehr breit geäußerten Interessen	Fragestellung		
2.Frage 2.Antwort	Ja - auf jeden Fall - auch hier je nach Fragestellung der Beratung	Fragestellung		
1.Frage 20.Antwort	Ziel der Beratung	Ziel der Beratung		
1.Frage 9.Antwort	allgemeine Interessenabklärung oder spezifische,	allgemeine Interessenabklärung oder spezifische,		
Textstelle/Antwort	Zitat	Paraphrase	Generalisierung	Kategorie
2.Frage 14.Antwort	NST gute Gesprächsgrundlage - allerdings Kombination mit anderen Tests erforderlich	Gesprächsgrundlage	Tests liefern Gesprächsstoff und Helfen dabei Diskussion in Gang zu bringen.	Test als Gesprächsstoff, oder Gesprächsöffner
5.Frage 5.Antwort	Tests sind Grundlagen für Gespräche und Diskussionen. Sie helfen oft, Themen auf die Spur zu kommen.	Grundlagen für Gespräche		
10.Frage 9.Antwort	Für mich ist ein Test immer eine Gesprächsgrundlage, ein anderer Zugang als eine Anamnese. Das ermöglicht ein anderes, vertieftes Gespräch	Gesprächsgrundlage		
5.Frage	Von "Entlarfung" halte	Basis für ein Gespräch.		

14.Antwort	ich nichts, sondern sie sind Basis für ein Gespräch.			
Textstelle/Antwort	Zitat	Paraphrase	Generalisierung	Kategorie
1.Frage 6.Antwort	wenn der/die Kund(in) findet: "ich habe keine Ahnung" und das Gespräch von diversen: "Ich weiss nicht" begleitet ist.	das Gespräch von diversen: "Ich weiss nicht" begleitet ist	Tests helfen dabei das Selbstbild aufzuzeigen oder eigen- und Fremdbild miteinander zu vergleichen.	Aufzeigen oder Spiegelung eigen- und Fremdbild
1.Frage 8.Antwort	Kunde hat Mühe differenziert zu verbalisieren und seine Bedürfnisse, Talente, Charaktereigenschaften klar zu reflektieren	Bedürfnisse, Talente, Charaktereigenschaften klar zu reflektieren		
3.Frage 17.Antwort	Spiegelung des IST-Zustandes, wo sind die Energien, Motivationen, Visionen	Spiegelung des IST-Zustandes		
3.Frage 18.Antwort	Befindlichkeit; Finden, Benennen und Eingrenzen von Problematiken; Ergründen der Belastungsfähigkeit und dem Mass der momentanen Belastung	Finden, Benennen und Eingrenzen von Problematiken		
5.Frage 11.Antwort	ich nutze die Chance, das, was stutzig macht zu thematisieren und zu klären	zu thematisieren und zu klären		
5.Frage 2.Antwort	Ich frage beim Kunden nach und bespreche das Resultat	Ich frage beim Kunden nach		
5.Frage 3.Antwort	Ich thematisiere die "Stutzigkeit" im Gespräch.	thematisiere die "Stutzigkeit"		

5.Frage 12.Antwort	Dabei scheint das Fremdbild/das Testresultat nicht mit dem Selbstbild übereinzustimmen. Diese Einsicht kann sehr gut in das kommende Reflexionsgespräch aufgenommen werden.	Dabei scheint das Fremdbild/das Testresultat nicht mit dem Selbstbild übereinzustimmen		
5.Frage 15.Antwort	Dies sind interessante Punkte, um beim Kunden eine Auseinandersetzung mit sich einzuleiten. Manchmal hinterfrage ich das Zustandekommen dieses Ergebnisses, bei Leistungstest kann Prüfungs- und Aufgabenlöse-Verhalten beobachtet und angesprochen werden	um beim Kunden eine Auseinandersetzung mit sich einzuleiten		
10.Frage 12.Antwort	Ausserdem stellt eine Testbesprechung ein hervorragendes Mittel dar, heikle Punkte anzusprechen; im Sinne von: "Im Test meine ich eine Tendenz zu folgender Problematik zu erkennen, was kommt Ihnen dazu in den Sinn?"	Stellt ein hervorragendes Mittel dar, heikle Punkte anzusprechen		
Textstelle/Antwort	Zitat	Paraphrase	Generalisierung	Kategorie
1.Frage 4.Antwort	Alter	Alter	Allgemeine Voraussetzung aber auch Alter, Zielsetzung bestimmen die Testauswahl.	Voraussetzungen der Teilnehmer bestimmen die Testauswahl
1.Frage 5.Antwort	nach den Voraussetzungen der	Voraussetzungen der Kunden		

	Kunden			
1.Frage 7.Antwort	Ausbildung des Kunden, Alter	Ausbildung des Kunden, Alter		
1.Frage 9.Antwort	Alter	Alter		
1.Frage 12.Antwort	Alter des Kunden/der Kundin	Alter		
1.Frage 19.Antwort	Nach der Klientel	Nach der Klientel		
1.Frage 20.Antwort	Alter, Schulbildung	Alter, Schulbildung		
1.Frage 22.Antwort	Alter	Alter		
1.Frage 9.Antwort	Komplexität des Tests	Komplexität des Tests		
2.Frage 13.Antwort	Komplexität, Anspruchsniveau	Komplexität, Anspruchsniveau		
1.Frage 1.Antwort	Zielgruppe	Zielgruppe		
1.Frage 10.Antwort	Zielgruppe	Zielgruppe		
Textstelle/Antwort	Zitat	Paraphrase	Generalisierung	Kategorie
2.Frage 20.Antwort	Es muss zum Kunden passen			Sonstige
1.Frage 18.Antwort	Eichung			
1.Frage 12.Antwort	Zeitpunkt der Beratung			
2.Frage 6.Antwort	manchmal muss man entscheiden: entweder oder, oder aber man gibt der Beantwortung eine Bewertung, hoch, tief, skalierend usf. manchmal wird die Zeit gestoppt, manchmal kann man Fragen frei beantworten oder es wird assoziiert schliesslich sind auch Zeichnungen möglich			
8.Frage 13.Antwort	Wenn sich jemand zwischen Alternativen nicht entscheiden kann			
8.Frage 8.Antwort	Berufswahlreife, "gefühlte" Unterstützung vom Umfeld daheim und in			

	der Schule			
9.Frage 6.Antwort	nur selten bei RAV-Beratungen			
1.Frage 16.Antwort	Interessen oder Persönlichkeitstest			
1.Frage 10.Antwort	Handhabung			
2.Frage 16.Antwort	Ich wähle häufig umfassende, mehrschichtige Tests, wie der NST			
Textstelle/Antwort	Zitat	Paraphrase	Generalisierung	Kategorie
9.Frage 1.Antwort	Unterschiedliche Weisungen / Konzepte		Die Organisation gibt vor wann welche Tests verwendet werden dürfen.	Testauswahl aufgrund von Vorgaben der Organisation
9.Frage 5.Antwort	Persönlichkeitstests dürfen nur in der kostenpflichtigen Beratung eingesetzt werden	nur in der kostenpflichtigen Beratung		
9.Frage 7.Antwort	Wir haben in unserer Organisation Vorgaben, welche Tests in welchen Packages angeboten werden können	Wir haben in unserer Organisation Vorgaben,		
Textstelle/Antwort	Zitat	Paraphrase	Generalisierung	Kategorie
3.Frage 1.Antwort	Ja, Fragebogen wie z.B. der BIP z.B. beinhalten solche Aspekte	Ja, Fragebogen wie z.B. der BIP z.B. beinhalten solche Aspekte	Testeinsatz ausserhalb von den Bereichen Interessen, Persönlichkeit und Leistung / Fähigkeit	Testanwendung in erweitertem Themenbereichen
3.Frage 2.Antwort	Berufswahl, Fach- oder Führungsausbildung, Laufbahnberatung,	Berufswahl, Fach- oder Führungsausbildung, Laufbahnberatung,		

	RAV-Beratung	RAV-Beratung		
3.Frage 5.Antwort	Keine Tests, aber Arbeitsmittel: Bilder, Fragebogen, Listen.... z.B. zu Motivation, Test zum Thema Lernen - sehr selten.	Test zum Thema Lernen		
3.Frage 6.Antwort	gestalterische Fähigkeiten	gestalterische Fähigkeiten		
3.Frage 7.Antwort	Entscheidungsfindung	Entscheidungsfindung		
3.Frage 8.Antwort	Arbeitsmittel: zB. Karriereanker um Bedürfnisse und Werte zu spiegeln	Arbeitsmittel: zB. Karriereanker um Bedürfnisse und Werte zu spiegeln		
3.Frage 10.Antwort	Wartegg: Ängste, Lernproblematiken	Wartegg: Ängste, Lernproblematiken		
3.Frage 15.Antwort	meist Studienwahl oder berufliche Neuorientierung	meist Studienwahl oder berufliche Neuorientierung		
3.Frage 20.Antwort	Motivation: LMI Burn out	Motivation: LMI Burn out		
3.Frage 21.Antwort	Psychische Verfassung, Passung Konzentrationsfähigkei t. Technisches Verständnis	Psychische Verfassung, Passung Konzentrationsfähigkei t. Technisches Verständnis		
Textstelle/Anw ort	Zitat	Paraphrase	Generalisieru ng	Kategorie
1.Frage 18.Antwort	Papier-Version		Antworten die als Gemeinsamk eit die Anwesenheit der Beratenden Person voraussetzen .	Anwesenheit der Beratenden Person erforderlich
2.Frage 13.Antwort	Präsenz des Beratenden			
2.Frage 22.Antwort	Leistungstests mache ich nur im Beratungssetting			

Antworten offene Fragen Bachelor-Thesis

1. Frage	Wenn Sie für eine Befragung einen Test (Interessen-, Persönlichkeits-, Leistungs- oder Fähigkeitstest) verwenden, nach welchen Kriterien wählen Sie diesen aus?
1. Antwort	Fragestellung, Zielgruppe, Zeitbedarf, Online-Verfügbarkeit
2. Antwort	Je nach Fragestellung der Beratung Beispiele: braucht der Kunde weitere Berufsideen (Interessen), will er mehr über sich erfahren (Persönlichkeit) oder muss abgeklärt werden, ob jemand z.B. EBA oder EFZ absolvieren kann (Leistungstest)
3. Antwort	Je nach Fragestellung oder Untersuchungsgegenstand (z.B. geht es darum, mehr über sich selber zu erfahren? Das intellektuelle Potential abzubilden etc....)
4. Antwort	Alter, Hintergrund, Fragestellung, Aufwand für Ratsuchenden, Aufwand für BB, Darstellung Ergebnis
5. Antwort	Nach Fragestellung in der Beratung, nach den Voraussetzungen der Kunden, Zeitaufwand, diagnostischen Grundlagen
6. Antwort	was ist die Fragestellung? was ist die Ausgangslage? was ist das Ziel? ...oder, wenn der/die Kund(in) findet: "ich habe keine Ahnung" und das Gespräch von diversen: "Ich weiss nicht" begleitet ist.
7. Antwort	Fragestellung, Ausbildung des Kunden, Alter;
8. Antwort	ganz ausdrücklicher Wunsch d Kunden und Definition wozu es sein soll Was ist die Grundfrage / das Anliegen / das Beratungsziel? Kunde hat Mühe differenziert zu verbalisieren und seine Bedürfnisse, Talente, Charaktereigenschaften klar zu reflektieren

9. Antwort	Fragestellung, Alter, Komplexität des Tests, Darstellung der Auswertung, Zeitfaktor, deutsche Sprachkenntnisse, allgemeine Interessenabklärung oder spezifische,
10. Antwort	Fragestellung, Zielgruppe, Auswertungsmöglichkeit, Handhabung, Visuelle Darstellung,
11. Antwort	Nach der Fragestellung
12. Antwort	Art der Fragestellung Zeitpunkt der Beratung Alter des Kunden/der Kundin
13. Antwort	Welche Frage möchte der Kunde / die Kundin mittels Test angehen? Bringt das Resultat ihn/sie allenfalls einen Schritt weiter?
14. Antwort	nach der Fragestellung, oft Kombination von zwei Tests
15. Antwort	nach Art der Fragestellung und nach Engagement des Kunden Kunde wünscht Klarheit in einem oder in mehreren dieser Aspekte; er ist vielleicht selber verunsichert oder unbestimmt in seinen Aussagen Ich möchte meine Einschätzung dem Kunden mit einem solchen Verfahren 'sachlich' belegen durch Aussagen, welche der Kunde gemacht hat Prozess in Beratung intensivieren
16. Antwort	Interessen oder Persönlichkeitstest
17. Antwort	Kundennutzen; Förderung meines Verständnisses für die Persönlichkeit und die Bedürfnisse des Kunden;
18. Antwort	Online-Version; Papier-Version; Zeitlicher Aufwand; Eichung;
19. Antwort	Nach dem Ertrag: welcher Test bringt mir

	schneller mehr. Die Auswertung muss zur Beantwortung der Fragestellung beitragen. Das Ergebnis fließt in die Beratung ein. Nach der Klientel: Verständnis der deutschen Sprache
20. Antwort	Welche Fragestellung im Vordergrund steht sowie Alter, Schulbildung, Ziel der Beratung
21. Antwort	Nach der Fragestellung. Nach den zeitlichen Ressourcen.
22. Antwort	Fragestellung, Alter
23. Antwort	Kriterien zur Auswahl von Interessenstests. >> sprachliches Niveau des/der Kundin >> Durchführungszeit >> gerne Onlineversionen >>Auswertungshilfen >>Fragestellung bei unklar oder sehr breit geäußerten Interessen

2. Frage	Gibt es Besonderheiten bei bestimmten Typen von Tests?
1. Antwort	Ja, bei Leistungstests gibt es kaum Online-Versionen. Zudem existieren bei vielen Tests keine Schweiz-spezifischen Normen. Würde in Laufbahnberatung gern auch projektive Tests einsetzen, dies ist bei kostenpflichtigen Beratung allerdings nicht mehr möglich, und auch sonst sind die Zeitfenster so eng, dass projektive Tests kaum mehr Platz haben.
2. Antwort	Ja - auf jeden Fall - auch hier je nach Fragestellung der Beratung
3. Antwort	Was ist mit dieser Frage genau gemeint? Meines Erachtens gibt es schon Besonderheiten, weil sonst würde es die verschiedenen Tests ja gar nicht geben.
4. Antwort	Ja
5. Antwort	Diese Frage verstehe ich nicht ganz. Natürlich gibt es Besonderheiten, jeder Test ist anders und damit besonders...
6. Antwort	manchmal muss man entscheiden: entweder oder, oder aber man gibt der Beantwortung eine Bewertung, hoch, tief, skalierend usf. manchmal wird die Zeit gestoppt, manchmal kann man Fragen frei beantworten oder es wird assoziiert schliesslich sind auch Zeichnungen möglich
7. Antwort	Online Auswertungen vs paper pencil

8. Antwort	nein (ausgeschlossen Leistungs- & Fähigkeistests, weil ich diese praktisch nicht mehr anwende)
9. Antwort	Jeder Test hat seine Besonderheiten
10. Antwort	Darstellung der Ergebnisse bei Interessentest ist für Jugendliche wichtig
11. Antwort	Diese Frage ist mir zu unspezifisch - sorry.
12. Antwort	Zeitaufwand für Instruktion, Durchführung, Auswertung
13. Antwort	Aus Sicht der Kundschaft: Länge, Komplexität, Anspruchsniveau, Verständlichkeit der Auswertungen Aus Sicht der Beratenden: Präsenz des Beratenden, Zeit für die Auswertung, Kosten
14. Antwort	NST gute Gesprächsgrundlage - allerdings Kombination mit anderen Tests erforderlich
15. Antwort	Jugendliche müssen die Gelegenheit erhalten, bei uns mit niveaugerechten Verfahren die Vielfalt der Berufsarbeiten kennenlernen zu können Bilder können bei ihnen häufig vielschichtigere Aussagen erzeugen als textliche Interessenschecks. Die Leistungstests bei uns sind meistens schulfächerunabhängiger als die zur Selektion eingesetzten Eignungstests. (Diese Frage ist m. E. zu offen formuliert - es gibt an jedem Verfahren etwas Besonderes)
16. Antwort	Ich wähle häufig umfassende, mehrschichtige Tests, wie der NST.
17. Antwort	Viele Tests überzeugen mich persönlich nicht, weil sie zu durchschaubar sind. Ich persönlich habe mich schon als Jugendlicher über Verfahren geärgert, bei denen ich schon im Vorherein wusste, was ich ankreuzen muss, damit mein angestrebtes Resultat herauskommt.
18. Antwort	Online-Version bedarf keine zeitliche Investition seitens BSLB für die Durchführung; die Besprechung der Testergebnisse je nach Test erübrigt sich oder man kann es telefonisch erledigen.
19. Antwort	Gewisse Interessentests beinhalten einen

	Persönlichkeitsaspekt (NST) Selbsteinschätzungen wie beim BIP oder Explorix können das Ergebnis verzerren Die Normen: nicht schweizerisch
20. Antwort	Es muss zum Kunden passen.
21. Antwort	Ja. Jeder Test hat diese. Das ist der Grund, ihn auszuwählen. Die Frage ist mir nicht klar.
22. Antwort	Leistungstests mache ich nur im Beratungssetting Interessen- und Persönlichkeitstests können von Kunden auch online bearbeitet werden
23. Antwort	Ja, die meiste digital, wenige noch Papier,...
24. Antwort	bei geringen Sprachkenntnissen Bildverfahren

3. Frage	Zu welchen anderen Themen, abgesehen von Interessen, Persönlichkeit und Leistung / Fähigkeit, verwenden Sie Tests?
1. Antwort	Ja, Fragebogen wie z.B. der BIP z.B. beinhalten solche Aspekte. Mit gesundheitsspezifischen Tests kenne ich mich ansonsten weniger aus.
2. Antwort	Berufswahl, Fach- oder Führungsausbildung, Laufbahnberatung, RAV-Beratung
3. Antwort	Keine
4. Antwort	Verwende wenig Tests.
5. Antwort	Keine Tests, aber Arbeitsmittel: Bilder, Fragebogen, Listen.... z.B. zu Motivation, Test zum Thema Lernen - sehr selten.
6. Antwort	gestalterische Fähigkeiten
7. Antwort	Entscheidungsfindung
8. Antwort	Arbeitsmittel: zB. Karriereanker um Bedürfnisse und Werte zu spiegeln
9. Antwort	bis jetzt keine anderen Themen
10. Antwort	Wartegg: Ängste, Lernproblematiken
11. Antwort	zu keinen anderen
12. Antwort	zu keinen anderen Themen
13. Antwort	keine weiteren

14. Antwort	Zu keinen.
15. Antwort	meist Studienwahl oder berufliche Neuorientierung
16. Antwort	zu keinen andern Themen
17. Antwort	Spiegelung des IST-Zustandes, wo sind die Energien, Motivationen, Visionen
18. Antwort	Befindlichkeit; Finden, Benennen und Eingrenzen von Problematiken; Ergründen der Belastungsfähigkeit und dem Mass der momentanen Belastung
19. Antwort	zu keinen weiteren Themen
20. Antwort	Motivation: LMI Burn out
21. Antwort	Psychische Verfassung, Passung Konzentrationsfähigkeit. Technisches Verständnis (aber das fällt wohl unter Fähigkeitstests)
22. Antwort	keine anderen Themen
23. Antwort	zu keinen anderen
24. Antwort	keine

5. Frage	Wenn Sie oder Ihr/e Klient/in bei dem Ergebnis eines Tests (Interessen-, Persönlichkeits-, Leistungs- oder Fähigkeitstest) "stutzig" werden, wie gehen Sie weiter vor?
1. Antwort	Zusammen mit der Klientin / dem Klienten überprüfen wir, ob die Aufgaben richtig verstanden wurden. Falls es nicht am Aufgabenverständnis liegt, überprüfe ich die Auswertung. Und falls auch diese o.k. ist, dann schauen wir zusammen, ob das Ergebnis nicht doch korrekt sein könnte, aber halt anders ausgefallen ist, als wir erwartet hätten. In seltenen Fällen wiederholen wir den Test.
2. Antwort	Ich frage beim Kunden nach und bespreche das Resultat
3. Antwort	Ich thematisiere die "Stutzigkeit" im Gespräch.
4. Antwort	Wenn es für Fragestellung relevant im Gespräch thematisieren.
5. Antwort	Tests sind Grundlagen für Gespräche und Diskussionen. Sie helfen oft, Themen auf die Spur zu kommen.

6. Antwort	ich bespreche es, mache darauf aufmerksam im Sinne von: was fällt Ihnen auf, das ist mir aufgefallen. Ich finde es interessant es bestätigt mir, dass Tests nicht alles sind es bestätigt mir, dass da noch was anderes sein könnte, oder ist etwas unbewusst, ist etwas ungenau gearbeitet worden, wie war die Tagesform?
7. Antwort	Validierung der Ergebnisse im Beratungsgespräch
8. Antwort	zuerst klären was die Dimension die stutzig macht bedeutet differenzieren, situativ beschreiben lassen unter welchen Bedingungen es für ihn zutrifft, wann nicht
9. Antwort	Es kommt sehr darauf an, wie der Klient reagiert. Interessiert den Klienten die Unsicherheit oder lehnt er das Resultat eher ab. Je nach dem gehe ich mehr in die Tiefe und suche mit dem Klienten nach Gründen.
10. Antwort	ich gehe auf die Itemebene und versuche so, eine Erklärung zu finden
11. Antwort	ich nutze die Chance, das, was stutzig macht zu thematisieren und zu klären
12. Antwort	Dabei scheint das Fremdbild/das Testresultat nicht mit dem Selbstbild übereinzustimmen. Diese Einsicht kann sehr gut in das kommende Reflexionsgespräch aufgenommen werden.
13. Antwort	Auswertung überprüfen und/oder Beantwortung der Fragen zusammen anschauen (falls möglich) und dadurch Erklärungen fürs Ergebnis finden.
14. Antwort	Wertneutrale Transparenz - sage, Sie hätten sich so in der Momentaufnahme entschieden, welche möglichen Erklärungen gebe es? Tests validieren sich immer auch im Gespräch. Von "Entlarfung" halte ich nichts, sondern sie sind Basis für ein Gespräch.
15. Antwort	Dies sind interessante Punkte, um beim Kunden

	<p>eine Auseinandersetzung mit sich einzuleiten. Manchmal hinterfrage ich das Zustandekommen dieses Ergebnisses, bei Leistungstest kann Prüfungs- und Aufgabenlöse-Verhalten beobachtet und angesprochen werden. Je nach momentaner psychischer Verfassung des Kunden (bei Jugendlichen häufiger) kann ein Testergebnis auch zu tendenziös sein. Da gilt: ein Test ist kein Test!</p>
16. Antwort	<p>Ich stelle entsprechende Fragen, allenfalls spiegle ich der Klientin meine Wahrnehmung.</p>
17. Antwort	<p>Offene, aber vorsichtige Konfrontation des Probanden mit den Testergebnissen</p>
18. Antwort	<p>Ich versuche herauszufinden was sagt der Klient, die Klientin dazu. Die Validierung von Klient ist mir sehr wichtig. Eventuell stellt sich ein Aha-Effekt.</p>
19. Antwort	<p>Nachfragen: Die Hypothese vom Klient einholen, warum das Ergebnis "fremd" erscheint? Kennt sich der Klient selber zu wenig? Der Klient könnte eine Fremdeinschätzung holen. Der Klient hat vielleicht die Fragen falsch verstanden. Einen anderen ähnlichen Test zwecks Vergleichs vorschlagen.</p>
20. Antwort	<p>Ich thematisiere diesen Punkt.</p>
21. Antwort	<p>Das Testergebnis wird hinterfragt und mit dem Kunden verifiziert. Im Handbuch nachlesen, überprüfen, z. B. Normen. Das Testergebnis kann z. B. in der Intervention/Supervision besprochen werden.</p>
22. Antwort	<p>Unsicherheit/Unklarheit im Gespräch definieren und nach Möglichkeit klären. Test ist aus meiner Sicht immer eine "Momentaufnahme", die in einer bestimmten Situation, Verfassung, Stimmung gemacht wird und somit nicht absolut</p>

23. Antwort	Situation ansprechen, nachfragen
24. Antwort	Bilde Hypothesen aus Wissen und Erfahrungen der einzelnen Beratung und diskutiere diese mit dem/der Kunden/in

8. Frage	Gibt es Situationen in denen Sie gerne Testen würden aber kein geeigneter Test zur Verfügung steht?
1. Antwort	Nein
2. Antwort	Nein
3. Antwort	Ja. Fehlende Erfahrung mit div. Tests.
4. Antwort	bei Migrantinnen/Migranten, die der deutschen Sprache (noch) nicht mächtig sind, aber bildungsnahe Biografien haben 50+Tests im Sinne von: war jetzt das alles und wohin könnte es (noch) gehen? (interessanterweise gibt es hier va. Soziale Interessentests, KV/ mit Menschen arbeiten- die Berufswelt wäre aber um einiges grösser!
5. Antwort	Nein
6. Antwort	Kaum
7. Antwort	Ja, wir haben in den Waben weniger Zeit für Tests.
8. Antwort	Berufswahlreife, "gefühlte" Unterstützung vom Umfeld daheim und in der Schule
9. Antwort	Nein
10. Antwort	die Möglichkeit "Kreativität" zu testen
11. Antwort	Eignungsabklärung zu konkreten Berufslehren

12. Antwort	Kaum
13. Antwort	Wenn sich jemand zwischen Alternativen nicht entscheiden kann
14. Antwort	Selten bei Personen, welche sich schlecht ausdrücken können oder mit Verbalem und Schriftlichem schwer tun. Um schulungsgewohnte Personen ganzheitlich erfassen zu können
15. Antwort	Motivationsbildertest, wie damals Achtnich. Alternativmöglichkeit zu NST mit moderneren Fragen
16. Antwort	Kaum, je mehr Erfahrung ich mit den von mir hauptsächlich eingesetzten Verfahren gesammelt habe, desto mehr kann ich aus ihnen herausholen.
17. Antwort	Nein
18. Antwort	Wenige
19. Antwort	Nein
20. Antwort	Ja, z. B. wie kreativ ist jemand und wofür reicht seine Begabung/Talent im gestalterischen Berufsfeld?
21. Antwort	eher selten
22. Antwort	nein
23. Antwort	In Situationen, bei denen ich den Eindruck gewinne, dass der/die Kundin eine verfehlte oder fehlende Selbstwahrnehmung (Sicherheit, Vertrauen, Wirksamkeit) zeigt.

9. Frage	Setzen Sie Tests auch in kostenpflichtigen Beratungen ein und falls ja, gibt es Unterschiede zu den Tests die in unentgeltlichen Beratungen eingesetzt werden?
1. Antwort	Unterschiedliche Weisungen / Konzepte
2. Antwort	ich mache keine kostenpflichtigen Test. Nein konnte ich nicht anklicken plus einen Kommentar geben....

3. Antwort	sie müssen online zur Verfügung stehen
4. Antwort	Nein
5. Antwort	Persönlichkeitstests dürfen nur in der kostenpflichtigen Beratung eingesetzt werden.
6. Antwort	nur selten bei RAV-Beratungen
7. Antwort	Wir haben in unserer Organisation Vorgaben, welche Tests in welchen Packages angeboten werden können.
8. Antwort	Nein, bisher keine Unterschiede
9. Antwort	Mutmassung, keine Erfahrung

10. Frage	Gibt es Informationen, die Sie mitteilen möchten, für die in der Umfrage kein Platz war?
1. Antwort	Scheint bestens zu klappen. Viel Erfolg!
2. Antwort	Nein
3. Antwort	Es gab Antwortmöglichkeiten, die mir zu undifferenziert formuliert waren und Fragen, die zu suggestiv formuliert waren.

4. Antwort	das würde mich wunder nehmen: wie wird die Persönlichkeit definiert? wie wird Psychologie definiert? es wird keine Frage gestellt: haben Sie lieber Tests oder ziehen Sie ein Gespräch vor? Klar gibt es einen Abschnitt, wo gefragt wird, welche Bedeutung ich persönlich dem Test beimesse, es wird aber nicht gefragt, was die Alternative ist und warum die Alternative auch eine Alternative wäre-
5. Antwort	der Test gilt hauptsächlich der Überprüfung und Ergänzung des Selbstbildes (des Kunden) und nicht des Fremdbildes (des Beraters) als Berater gebe ich meinen Fremdbildern / Projektionen nicht mehr so viel Raum wie damals. es bewahrt mich und die Kundin davor, dass von mir am Ende ein Urteil gefällt wird und sie mit einem "Rat geschlagen" wird.
6. Antwort	Es wäre hilfreich, eine Art Früherfassungsverfahren zu entwickeln, um Jugendliche zu eruieren, die von ihrem Umfeld wenig Unterstützung in der BW erwarten können. Dies müsste bereits vor oder in der 7. Klasse geschehen.
7. Antwort	Grundsätzlich geht es mir beim Einsatz von Tests weniger um mich, sondern mehr darum dem Kunden die Möglichkeit zu geben anhand von Aufgaben Themen zu reflektieren (Interesse, Persönlichkeit, Potenzial, Leistung) und im gemeinsamen Gespräch entlang von auftauchenden Fragen zu verifizieren.
8. Antwort	Ich führe keine kostenpflichtigen Beratungen durch.
9. Antwort	Es sollte mehr zwischen Interessen -und Persönlichkeitstests unterschieden werden. Für mich ist ein Test immer eine Gesprächsgrundlage, ein anderer Zugang als eine Anamnese. Das ermöglicht ein anderes, vertieftes Gespräch. Dieser Aspekt wurde hier nicht berücksichtigt. M.E. wurde dem Aspekt "stuzig machen" viel zu viel Gewicht beigemessen. Es ist wie es ist und darüber findet

	das Gespräch statt.
10. Antwort	Ich berate seit über 30 Jahren alle Altersgruppen, ausser Gymnasiasten, und möchte am wenigsten auf Interessenstests/-checks verzichten. Sie werden anteilmässig am meisten eingesetzt. In den letzten Jahren ist der ökonomische Aufwand von Tests stets wichtiger geworden, was mein Berater-Verständnis für Aus- und Weiterbildungschancen für alle trübt.
11. Antwort	Leider sind die gestellten Fragen teilweise unklar und dementsprechend nicht präzis beantwortbar.
12. Antwort	Ich sehe Tests als ausgezeichnete Ergänzung zu meiner persönlichen Wahrnehmung und erlebe oft eine Unterstützung meiner Intuition. Ausserdem stellt eine Testbesprechung ein hervorragendes Mittel dar, heikle Punkte anzusprechen; im Sinne von: "Im Test meine ich eine Tendenz zu folgender Problematik zu erkennen, was kommt Ihnen dazu in den Sinn?"
13. Antwort	Tests mit Bildern finde ich für bestimmte Klienten besser geeignet als Fragebogen.
14. Antwort	Bei der Frage: Tests unterstützen bei der Einschätzung von Klienten. Sind jedoch weniger wichtig als langjährige Erfahrung. Meinen Sie damit, dass langjährige Berufserfahrung einer Beratungsperson einen Test unnötig machen kann? Wenn ja. Erfahrung hilft für die Einschätzung. Selbst ein langes Gespräch zeigt meiner Erfahrung her jedoch nie alle Aspekte auf, die mit einem Test in kurzer Zeit erhoben werden können.
15. Antwort	Nein

Anhang E Auswertung der Matrixfragen

4.Frage: Einige allgemeine Fragen zu eignungsdiagnostischen Tests. Wie stark stimmen Sie mit den Aussagen überein?		
Code Antwortoption	n	%
Ich verwende Tests vor allem dann, wenn ich mir bei einem/r Klienten/in unsicher bin		
1 stimme nicht zu	13	35%
2 stimme eher nicht zu	12	32%
3 stimme eher zu	10	27%
4 stimme zu	2	5%
Klienten/innen anhand von Tests zu beurteilen wäre grundsätzlich immer sinnvoll.		
1 stimme nicht zu	11	30%
2 stimme eher nicht zu	15	41%
3 stimme eher zu	9	24%
4 stimme zu	2	5%
Tests sind dann sinnvoll, wenn ich aufgrund meiner Erfahrung Klienten/innen nicht gut genug einschätzen kann.		
1 stimme nicht zu	12	33%
2 stimme eher nicht zu	9	25%
3 stimme eher zu	13	36%
4 stimme zu	2	6%
Tests bestätigen hauptsächlich das, was ich im Vorfeld vermutet habe.		
1 stimme nicht zu	5	14%
2 stimme eher nicht zu	11	30%
3 stimme eher zu	19	51%
4 stimme zu	2	5%
Tests liefern Informationen, die ich durch kein anderes Verfahren aufdecken kann.		
1 stimme nicht zu	2	5%
2 stimme eher nicht zu	17	46%
3 stimme eher zu	12	32%
4 stimme zu	6	16%
Tests unterstützen bei der Einschätzung von Klienten/innen, sind jedoch weniger wichtig als langjährige Erfahrung.		
1 stimme nicht zu	3	8%
2 stimme eher nicht zu	16	44%
3 stimme eher zu	13	36%
4 stimme zu	4	11%
Durch Tests kann ich mich und meine Einschätzung absichern.		
1 stimme nicht zu	3	9%
2 stimme eher nicht zu	10	29%
3 stimme eher zu	15	43%
4 stimme zu	7	20%
Ideal zur Informationsgewinnung ist es, zuerst Tests durchzuführen und diese dann durch individuelle Fragen zu ergänzen.		
1 stimme nicht zu	11	31%
2 stimme eher nicht zu	7	20%
3 stimme eher zu	12	34%
4 stimme zu	5	14%

6. Frage: Wenn ich oder der/die Klient/in bei dem Ergebnis eines Tests „stutzig“ werden...

Code Antwortoption	n	%
Sollte die Testauswertung überprüft werden, um mögliche Fehler auszuschliessen.		
1 Stimme nicht zu	5	19%
2 Stimme eher nicht zu	2	7%
3 Stimme eher zu	14	52%
4 Stimme zu	6	22%
Sollte der Test wiederholt werden.		
1 Stimme nicht zu	13	48%
2 Stimme eher nicht zu	12	44%
3 Stimme eher zu	2	7%
4 Stimme zu	0	0%
Sollte ich meinen Eindruck über den/die Klienten/in überprüfen.		
1 Stimme nicht zu	2	7%
2 Stimme eher nicht zu	4	15%
3 Stimme eher zu	14	52%
4 Stimme zu	7	26%
Sollten die Aussagen des/r Klienten/in überprüft werden.		
1 Stimme nicht zu	3	11%
2 Stimme eher nicht zu	4	15%
3 Stimme eher zu	17	63%
4 Stimme zu	3	11%

7. Frage: Nochmals kurz zurück zur Auswahl der Tests. Bitte geben Sie an, wie wichtig Ihnen folgende Aspekte sind.

Code Antwortoption	n	%
Die Länge und Dauer des Tests		
1 Unwichtig	4	12%
2 eher unwichtig	7	21%
3 eher wichtig	11	33%
4 wichtig	11	33%
Die eigene Erfahrung / Routine mit den Tests		
1 Unwichtig	2	6%
2 eher unwichtig	4	12%
3 eher wichtig	10	30%
4 wichtig	17	52%
Die sprachlichen Anforderungen der Tests		
1 Unwichtig	0	0%
2 eher unwichtig	6	19%
3 eher wichtig	14	44%
4 wichtig	12	38%
Komfort bei der Auswertung		
1 Unwichtig	0	0%
2 eher unwichtig	9	27%
3 eher wichtig	13	39%
4 wichtig	11	33%

Online Verfügbarkeit		
1 Unwichtig	0	0%
2 eher unwichtig	6	19%
3 eher wichtig	18	56%
4 wichtig	8	25%
Die Gütekriterien		
1 Unwichtig	1	3%
2 eher unwichtig	4	12%
3 eher wichtig	15	45%
4 wichtig	13	39%
Das Layout		
1 Unwichtig	4	13%
2 eher unwichtig	9	28%
3 eher wichtig	15	47%
4 wichtig	4	13%
Die Kosten für die jeweiligen Testverfahren		
1 Unwichtig	0	0%
2 eher unwichtig	13	39%
3 eher wichtig	11	33%
4 wichtig	9	27%
Die Empfehlungen von Arbeitskollegen		
1 Unwichtig	1	3%
2 eher unwichtig	12	36%
3 eher wichtig	17	52%
4 wichtig	3	9%