



universität
wien

Diplomarbeit

Titel der Arbeit

Zwei neue Subtests für das „Self-Assessment“ Psychologie
durch regelgeleitete Konstruktion

Verfasserin

Nicole Christine Undeutsch

Angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag. rer. nat.)

Wien, im März 2010

Studienkennzahl: 298

Studienrichtung: Psychologie

Betreuer: Univ.-Prof. Dr. Mag. Klaus D. Kubinger

Vorwort und Danksagung

Die vorliegende Arbeit beschäftigt sich mit der Entwicklung zweier diagnostischer Verfahren, deren Items dem Rasch-Modell entsprechen. Diese Verfahren sollen im Rahmen des Self-Assessment Psychologie an der Universität Wien am Institut für Entwicklungspsychologie und Psychologische Diagnostik zum Einsatz kommen.

Schon zu Beginn meines Studiums interessierte ich mich sehr für den Einsatz und die Entwicklung psychologisch diagnostischer Verfahren. Bereits damals verspürte ich großes Interesse irgendwann selbst ein Verfahren zu entwickeln. Daher freute es mich umso mehr, als Prof. Dr. Mag. Klaus D. Kubinger mir dieses Thema anvertraute. Ich möchte mich hiermit einerseits für das Vertrauen, welches er mir durch die Übergabe dieses Themas entgegenbrachte und andererseits für seine Unterstützung, seine Geduld sowie für die kritische Auseinandersetzung mit meinen Ideen bedanken.

Diese Arbeit entstand über einen längeren Zeitraum hinweg, bei dem ich immer wieder Hilfeleistung von mehreren Personen erhielt, denen ebenfalls noch ein besonderer Dank gilt.

Ausdrücklicher Dank gebührt meinem Subbetreuer Mag. Philipp Sonnleitner für die fachliche Unterstützung in der Konstruktionsphase dieser Arbeit.

Ebenso möchte ich mich bei Mag. Lisbeth Weitensfelder für ihre Beiträge bei der Auswahl meiner Items sowie für die kritische Auseinandersetzung mit diesen bedanken.

Auch an Herrn Mag. Poinstingl möchte ich ein Dankeschön richten, da er mir in der Phase der Auswertung bei Problemen hilfreich zur Seite stand.

Ein großes Lob und spezieller Dank gilt Wolfgang Steiner, der mich durch seine Programmierkenntnisse aus einer misereren Lage gerettet hat und somit meine Diplomarbeit in ihrer jetzigen Form möglich machte.

Für ihren fachlichen Rat und für das Korrekturlesen meiner Arbeit möchte ich mich sehr bei Mag. Mario Karolyi und bei Robin Gleeson bedanken.

Zudem bedanke ich mich bei meinen Kollegen und Kolleginnen aus dem Forschungsseminar, die meine Items bzw. meine Verfahren als Testpersonen kritisch bearbeitet haben; sowie bei allen freiwilligen Psychologieinteressenten, die sich bereit erklärt haben meine Verfahren durchzuführen.

Ein besonders herzlicher Dank gilt meiner Studienkollegin Petra Schütz, welche mich durch das ganze Studium begleitete hat, mich motivierte, wenn es sein musste und zum lachen brachte, wenn es nötig war und mir die Zeit wie im Flug vergehen lies. Nicht zuletzt möchte ich ihr natürlich auch für das Korrekturlesen meiner Arbeit Danken.

Natürlich gilt besonderer Dank meiner ganzen Familie, die stets an mich glaubte und mir immer vermittelte wie stolz sie auf mich ist. Von ganzem Herzen möchte ich mich vor allem bei meinen Eltern, Ulrike Moemen und Ludwig Undeutsch bedanken, da sie mir dieses Studium ermöglicht haben, mich immer bedingungslos unterstützen, in schweren Zeiten für mich da sind und an meinen Erfolg glauben.

Auch Alexander Havlik möchte ich für Verständnis, Geduld, Unterstützung und Wertschätzung in den letzten Jahren meinen Dank ausdrücken.

Vielen Dank!

Wien, im März 2010

Abstract

Die vorliegende Arbeit beschreibt die Entwicklung zweier psychologisch-diagnostischer Verfahren für das Self-Assessment Psychologie Wien. Auf Basis der regelgeleiteten Itemkonstruktion wurden der Färbige Matrizentest – FMT und der spezielle Fragebogen zum Interesse am Psychologiestudium (Wien) – FIP entwickelt. Neben der Abhandlung der Themen Self-Assessment und regelgeleitete Itemkonstruktion kommt es im Weiteren zu einer Auseinandersetzung mit Theorien zum Thema Matrizentest und Theorien zum Thema Interessen(fragebogen). Im empirischen Teil dieser Arbeit wurde auf der Literatur aufbauend ein Konstruktionsrational für den FMT entworfen. Dieses sollte für die Vorgabe eines Matrizentests, bei welchem das Fragezeichen in der Mitte der Matrix ist, geeignet sein. Anschließend wurden 28 Items und zwei Instruktionsitems entwickelt. Für den FIP hingegen wurden, mit Bezug zum Interessenkonstrukt von Brickenkamp (1990), neue Regeln zur Itemkonstruktion erstellt. Neben der Skala Psychologie Wien mit 35 Items wurden die Skalen Psychologie Klagenfurt mit 5 Items und typische Irrtümer/falsche Erwartungen mit 5 Items berücksichtigt. Die computerisierten Verfahren wurden einer Stichprobe von 125 Psychologieinteressenten im Alter von 16 bis 17 Jahren vorgegeben. Anhand dieser ließ sich zeigen, dass 24 Items des FMT post-hoc Modellgültigkeit des Rasch-Modells erzielten. Ebenfalls konnte für alle drei Skalen des FIP getrennt Rasch-Modell-Konformität nachgewiesen werden. Im Weiteren wurden schließlich noch die Konstruktionsprinzipien des FMT mittels LLTM analysiert. Jedoch konnten diese die Itemschwierigkeiten nicht besser erklären, als dies im Rasch-Modell der Fall war. Die Arbeit stellt somit zwei neue eindimensionale psychologisch-diagnostische Verfahren für die Praxis zur Verfügung, welche im Weiteren noch hinsichtlich ihrer Gütekriterien untersucht werden sollten.

Schlüsselwörter: regelgeleitete Itemkonstruktion, Matrizentest, Reasoning, spezieller Interessenfragebogen, Interesse, Self-Assessment, RM, LLTM

This thesis describes the development of two psychological diagnostic tests for the Self-Assessment Psychology Vienna. The “Färbige Matrizentest – FMT” and the special “Fragebogen zum Interesse am Psychologiestudium (Wien) – FIP” have been developed by rule-based Item Construction. Apart from issues like self-assessment and rule-based Item Construction, theories about matrices tests and interest questionnaires are discussed. In the empirical part of this work, a construction principle for the FMT was designed with reference to the literature, which would be suitable for the

implementation of a matrices test in which the question mark is at the center of the matrix. Subsequently, 28 items and two instruction items were developed. Whereas for the FIP, which is related to the construct of interest from Brickenkamp (1990), new rules were created. In addition to the scale «psychology Vienna» with 35 items, the scale «psychology Klagenfurt» with five items and the scale «typical mistakes/false expectations» with five items were designed. The tests that were computerized were given to a sample of 125 psychology interested parties, aged 16 to 37 years. Based on this, 24 items of the FMT obtained post-hoc Rasch-Modell-conformity. This could also be demonstrated for the three scales of the FIP, as each of them fits the Rasch-Modell. Finally the construction principles of the FMT were analyzed using the LLTM. However, the LLTM could not explain the item difficulties any better when compared with the Rasch-Modell. The thesis thus provides two new one-dimensional psycho-diagnostic tests for practice, which should be further analyzed concerning quality criteria.

Keywords: rule-based item construction, matrices tests, reasoning, special interest questionnaires, interest, self-assessment, RM, LLTM

Inhaltsverzeichnis

Einleitung.....	9
1 Self-Assessment.....	11
1.1 Self-Assessment zur Studienwahlberatung: Der Wiener Ansatz.....	13
2 Regelgeleitete bzw. rationale Itemkonstruktion.....	16
3 Methodik.....	18
3.1 Das dichotom logistische Modell von Rasch (1 PL-Modell).....	18
3.2 Das Lineare Logistische Test-Modell von Fischer (LLTM).....	21
4 Matrizentest.....	23
4.1 Schlussfolgerndes Denken.....	23
4.1.1 Deduktiver Schluss.....	24
4.1.2 Induktiver Schluss.....	24
4.1.3 Analoger Schluss.....	25
4.2 Schlussfolgerndes Denken bei Matrizentests.....	27
4.3 Bestehende Matrizentests.....	29
4.3.1 Wiener Matrizentest – WMT (Formann & Piswanger, 1979).....	29
4.3.2 Adaptiver Matrizentest – AMT (Hornke, Etzel & Rettig, 1999).....	32
4.4 Färbiger Matrizentest – FMT.....	35
4.4.1 Das Itemuniversum.....	36
4.4.2 Das Regelsystem – Schwierigkeitsgenerierende Operationen.....	37
4.4.3 Materialeigenschaften.....	41
4.4.4 Verteilung der Regeln.....	42
4.4.5 Distraktoren.....	45
5 Interessenfragebogen.....	47
5.1 Interesse.....	47
5.1.1 Erhebung von Interesse.....	49
5.2 Die Theorie von John L. Holland.....	53
5.2.1 Allgemeiner Interessen-Struktur Test/Umwelt-Struktur-Test - Revision (AIST-R/UST-R).....	57
5.2.2 Explorix®.....	58
5.2.3 Multimethodische Objektive Interessentestbatterie (MOI).....	59
5.3 Das Interessenkonzept sensu Brickenkamp (1990).....	61
5.3.1 Generelle Interessen-Skala (GIS).....	63
5.4 Entwicklung des Fragebogens zum Interesse am Psychologiestudium (FIP).....	64
5.4.1 Definition des Konstrukts bzw. Festlegung des Sachverhaltes.....	65
5.4.2 Wahl der Aufgabenform.....	68

5.4.3 Erstellung von Regeln.....	69
6 Datenerhebung.....	73
6.1 Online Testbatterie.....	73
6.1.1 Teil1: Demographische Variablen und Fragen zur Studienwahlreife.....	74
6.1.2 Teil 2: FMT.....	75
6.1.3 Teil 3: FIP.....	77
6.2 Stichprobe.....	78
6.2.1 Demographische Beschreibung der Gesamtstichprobe.....	80
6.2.2 Beschreibung der Gesamtstichprobe in Bezug zur Studienwahlreife.....	81
7 Darstellung der Ergebnisse.....	86
7.1 Analysen zur Rasch-Modell Konformität des FMT.....	86
7.1.1 Kriterienbildung.....	86
7.1.2 LRT nach Andersen - FMT.....	87
7.2 Analysen zur Rasch-Modell Konformität des FIP.....	93
7.2.1 LRT nach Andersen – FIP: Items Psychologie Wien.....	93
7.2.2 LRT nach Andersen – FIP: Items Inhalte aus Klagenfurt.....	94
7.2.3 LRT nach Andersen – FIP: Items typische Irrtümer/falsche Erwartungen.....	96
7.3 Erstellung einer vorläufigen Eich Tabelle für Psychologieinteressenten.....	98
7.4 Analysen des FMT nach dem LLTM.....	98
7.4.1 Überprüfung des Konstruktionsrationalis.....	98
7.4.1 Überprüfung des Konstruktionsrationalis mit Bezug zu Formann (1973).....	101
8 Interpretation und Diskussion.....	103
9 Zusammenfassung.....	107
Literaturverzeichnis.....	109
Anhang.....	119
Anhang 1: Definitionen der ausgewählten Schwerpunkte.....	120
Anhang 2: Eich Tabelle FMT.....	121
Anhang 3: Eich Tabellen FIP.....	122
Lebenslauf.....	127

Einleitung

In den letzten Jahren entwickelte sich im Bereich des Hochschulzugangs und das nicht nur aufgrund von Zugangsbeschränkungen der Trend von fachspezifischen Self-Assessments. Egal für welche Fachrichtung ein Self-Assessment entwickelt wurde, sie dienen alle dem gleichen Zweck; nämlich jene Studieninteressenten zum Studium anzuregen, welche für dieses aufgrund von Fähigkeiten und Neigungen geeignet erscheinen.

Auch am Arbeitsbereich Psychologische Diagnostik, am Institut für Entwicklungspsychologie und Psychologische Diagnostik der Universität Wien, zeigt sich diese Entwicklung. Bereits seit 2005 wird hier ein Self-Assessment für Psychologie angeboten. Seit 2009 gibt es zudem zwischen der Test- und Beratungsstelle des Arbeitsbereichs Psychologische Diagnostik und der Technischen Universität Wien eine Kooperation bei der neue fachspezifische Self-Assessment-Tools entwickelt werden sollen, zum Beispiel für die Studienrichtungen Architektur und Maschinenbau.

Durch diesen Trend steigt der Bedarf an neuen psychologisch diagnostischen Verfahren, die spezielle bzw. allgemeine Studierfähigkeiten oder Interessen erheben. Demzufolge ist es das Ziel dieser Arbeit, zwei neue diagnostische Verfahren für das Self-Assessment Psychologie zu konstruieren. Explizit sollte ein Interessenfragebogen sowie ein Matrizentest entwickelt werden, welche im oben erwähnten Self-Assessment Psychologie zum Einsatz kommen sollen. Dabei sollte eine regelgeleitete Konstruktion erfolgen um die Validität der Verfahren im Vorhinein zu sichern. Außerdem sollte dieses Vorgehen anderen Verfahrenskonstrukteuren ermöglichen, ähnliche Verfahren auf die gleiche Weise für andere Studienrichtungen zu entwickeln. Wesentlich ist, dass die entwickelten Verfahren dem Rasch-Modell entsprechen und somit nachweislich eine Dimension erheben.

Zu Beginn dieser Arbeit wird daher geklärt was man unter einem Self-Assessment versteht. Anschließend wird das Wiener Self-Assessment für Psychologie vorgestellt. Nach einer kurzen Erläuterung der Regelgeleiteten Itemkonstruktion werden die in dieser Arbeit zur Anwendung kommenden Auswertungsverfahren kurz vorgestellt. Darauf folgt die Auseinandersetzung mit dem Thema Matrizentest anhand von folgenden Fragen: „Was versteht man unter „reasoning?“,

„Was misst ein Matrizentest?“ und „Welche bekannten Matrizentests gibt es?“ Abschließen wird dieses Thema mit der Beschreibung des neuen Matrizentests, vor allem in Bezug auf dessen Konstruktionsprinzipien. Einen weiteren Schwerpunkt dieser Arbeit stellt die Auseinandersetzung mit dem Thema Interessen(-fragebogen) dar. Auch hier werden Fragen wie: „Was versteht man unter Interesse?“, „Wie werden Interessen erhoben?“ und „Welche bekannten Interessenfragebögen gibt es?“ geklärt. Nach der Klärung dieser Fragen wird der neu entwickelte Interessenfragebogen vorgestellt, ebenfalls mit besonderem Hinblick auf dessen Konstruktionsprinzipien.

Nach diesen thematischen Abhandlungen wird auf die Datenerhebung dieser Arbeit eingegangen. Zuerst wird die Online-Testbatterie vorgestellt, gefolgt von den Ausführungen zur Stichprobenakquirierung und anschließender Beschreibung dieser. Weiters werden erste testtheoretische Überprüfungen der beiden neu entwickelten Verfahren vorgenommen und dargestellt.

Anmerkung: Aus Gründen der Vereinfachung und zur besseren Lesbarkeit der Arbeit wird auf eine geschlechterspezifische Formulierung verzichtet. Es sei denn das Geschlecht wird explizit genannt.

1 Self-Assessment

Unter Self-Assessment (engl.: die Selbstbeurteilung bzw. die Selbsteinschätzung) versteht man ein psychologisch-diagnostisches Verfahren, welches zumeist aus mehreren einzelnen diagnostischen Verfahren besteht, die spezielle Leistungen, Interessen und Einstellungen erheben sollen. Self-Assessments sind vorwiegend computerisiert und ermöglichen dem Anwender durch sofortige Rückmeldung eine Reflexion über die erhobenen Fähigkeiten und Einstellungen, in Bezug auf zuvor definierte Anforderungen. Explizit geht es um die Aufdeckung des Grades der Übereinstimmung bzw. Passung der Leistungen, Einstellungen und Interessen des Anwenders mit den nötigen Anforderungen einer Ausbildung, einer Studienrichtung, etc.

Nach Kubinger, Moosbrugger, Frebort, Jonkisz und Reiß (2007, S. 323) haben Self-Assessments „das diagnostische Ziel der Feststellung, inwieweit die betreffende Testperson hinsichtlich bestimmter beruflicher oder berufsausbildender Anforderungen geeignet ist“. Man könnte also sagen, dass Self-Assessments als Eignungsbeurteilung fungieren.

Frebort und Kubinger (2008) kommen weiteres zu dem Schluss „„Self-Assessments“, [...], können bei der Studienwahl als Entscheidungsunterstützung einen wichtigen Beitrag leisten.“

Welchen Beitrag und welchen Nutzen Self-Assessments liefern, kann von zwei unterschiedlichen Standpunkten betrachtet werden. Es kann der Nutzen für angehende Studierende, also Studieninteressenten, von dem Nutzen für die Universitäten bzw. die Studienfachrichtungen selbst unterschieden werden.

Dilger, Gerholz, Klieber und Sloane (2008) weisen auf drei grundlegende Funktionen des Self-Assessments für Studieninteressenten hin: diese seien eine Informationsfunktion, eine Reflexionsfunktion und eine Entscheidungsfunktion. Unter Informationsfunktion versteht man die Möglichkeit von angehenden Studierenden, durch das Self-Assessment Informationen über die interessierende Studienrichtung zu erhalten, wie zum Beispiel über Inhalte des Studiums, über Anforderungen des Studiums oder auch über Schwerpunkte an diesem Studienort. Die Reflexionsfunktion ergibt sich daraus, dass man sowohl durch die Einsicht in die Anforderungen und durch die Durchführung des Self-Assessments als auch durch die daraus resultierende Rückmeldung seine eigenen Fähigkeiten und Neigungen genauer abwägen kann. Es ergibt sich zudem die Möglichkeit, durch Einsicht oder durch Rückmeldung Stärken bzw. Schwächen zu erkennen. So hat man die Möglichkeit diesen aufgedeckten Schwächen durch gezielte Förderungen entgegenzuwirken. Die Entscheidungsfunktion ergibt sich schließlich daraus, dass

der Studieninteressierte die Möglichkeit hat, durch seine gewonnenen Informationen in Bezug auf seine Studienwahlentscheidung einzuordnen, zu beurteilen und zu gewichten.

Auch für die einzelnen Studiengänge ergeben sich unterschiedliche Funktionen. Unter anderem zählen Dilger et al. (2008) die Darstellungsfunktion und die Beratungsfunktion zum Nutzen des Self-Assessments für einzelne Studiengänge. Die Darstellungsfunktion ergibt sich daraus, dass Informationen über den Studiengang, die Inhalte und Schwerpunkte mittels eines Self-Assessment übermittelt werden können. Zusätzlich erlaubt die Beratungsfunktion Anforderungen des Studiums darzustellen.

Betrachtet man die angegebenen Funktionen aus der Sicht beider Positionen, scheinen sich diese zu decken. Die gemeinsamen Interessen zeigen sich auch, wenn man die Ziele eines Self-Assessments betrachtet. Die Ziele der Studiengänge sind einerseits mehr geeignete und motivierte Studierende aufzunehmen, und andererseits diese vom Abbrechen des Studiums abzuhalten. Das Ziel der Studieninteressenten ist, dass sie schon anfänglich das richtige Studium wählen, welches ihren Vorstellungen, Neigungen und Fähigkeiten am besten entspricht, ohne dieses abbrechen zu „müssen“ und ein neues Studium anzufangen.

Es geht also bei Self-Assessments (im universitären Bereich) auch um die Erhöhung der Transparenz der Studienanforderungen gegenüber Interessenten. Es sollen am besten diejenigen Studieninteressenten „ausgewählt“ werden, die das Studium erfolgreich abschließen und die später auch erfolgreich in ihrem jeweiligen Berufsfeld arbeiten werden.

Dass sowohl Studiengänge als auch Studieninteressierte von Self-Assessment Systemen profitieren und dies als eine positive Hochschulentwicklung ansehen, zeigt sich durch den Trend, dass immer mehr Self-Assessments angeboten werden.

Als Beispiele seien zu nennen: der „Studienkompass Psychologie“ (Universität Hamburg, 2009), das „SelfAssessment für zukünftige Studierende“ (Verbund Norddeutscher Universitäten, 2009) und die Self-Assessments Maschinenbau, Informatik / Elektrotechnik, Psychologie und Physik der RWTH¹ Aachen (Lehrstuhl für Betriebs- und Organisationspsychologie der RWTH Aachen, 2008). Außerdem das Self-Assessment Psychologie der Test- und Beratungsstelle der Universität Wien (2005), Arbeitsbereich Psychologische Diagnostik, auf das im Folgenden genauer eingegangen wird.

¹ Rheinisch-Westfälische Technische Hochschule Aachen

1.1 Self-Assessment zur Studienwahlberatung: Der Wiener Ansatz

Das Self-Assessment für den Studiengang Psychologie ist ein Forschungsprojekt der Test- und Beratungsstelle des Arbeitsbereichs Psychologische Diagnostik, Fakultät für Psychologie, Universität Wien.

Der erste Schritt bei der Entwicklung des Self-Assessments bezog sich auf die Suche nach einem Anforderungsprofil (Frebort & Kubinger, 2006), das heißt den Voraussetzungen, die für ein Psychologiestudium notwendig sind, um dieses zu bewältigen.

Die DIN 33430² schlägt für die Aufdeckung der Anforderungen eine *Anforderungsanalyse* vor, das heißt „die Ermittlung von personenrelevanten psychischen und psycho-physischen Voraussetzungen für den zu besetzenden Arbeitsplatz, das Aufgabenfeld, die Ausbildung bzw. das Studium oder den Beruf“ (Reimann, 2004, S. 109).

Somit wurde vor der Konstruktion des Self-Assessments mittels „critical incident technique“³ (Flanagan, 1954, zitiert nach Reimann, 2004, S. 112) ein Anforderungsprofil für Psychologiestudenten ermittelt. Es wurden einerseits 131 Psychologiestudierende befragt, die ihre Antworten schriftlich wiedergaben, und andererseits wurden sieben Professoren bzw. Lehrveranstaltungsleiter interviewt. Nach einer Kategorisierung in Situationsklassen wurden die einzelnen Verhaltensnennungen zu Verhaltensklassen über die Situation hinweg zusammengefasst (Frebort & Kubinger, 2006, S. 410).

Das *Anforderungsprofil* setzt sich wie folgt zusammen:

- Hohe Leistungsmotivation und Zielorientierung
- Realistisches Konzept eigener Fähigkeiten
- Planerisches Vorgehen
- Effizienter Lernstil
- Fähigkeit zum folgerichtigen Denken
- Hohe Frustrationstoleranz, Durchhaltevermögen
- Angemessenes Anspruchsniveau, selbstständige Informationssuche
- Leistungsförderlicher Attributionsstil

2 Die DIN 33430 ist die praxisorientierte Prozessnorm, die Qualitätskriterien für die Auswahl, Planung, Durchführung und Auswertung von berufsbezogenen Eignungsbeurteilungen formuliert (Deutsche Psychologen Akademie GmbH des BDP, 2002).

3 Die „critical incident technique“ (Flanagan, 1954, zitiert nach Reimann, 2005) ist eine Arbeits- und Anforderungsanalysemethode. Grundannahme ist, dass sich gute und schlechte Mitarbeiter auf der Verhaltensebene in kritischen Situationen unterscheiden. Mittels Fragebogen, Beobachtungen, Interviews, etc. sollen die für Erfolg und Misserfolg kritischen Ereignisse erfasst und gewichtet werden; diese werden dann zusammengefasst und kategorisiert um abschließend als Anforderungen an das Verhalten beschrieben zu werden.

Darauf aufbauend wurde das Self-Assessment aus psychologisch-diagnostischen Verfahren zusammengestellt, die teilweise neu entwickelt wurden, bereits publiziert waren oder eigens dafür adaptiert wurden (Kubinger et al., 2007).

Im Folgenden werden die Verfahren, welche beim Self-Assessment bisher zum Einsatz kamen, mit ihrer zu messenden Eigenschaft aufgelistet (Frebort & Kubinger, 2008):

- Merkfähigkeit, Lernstrategie und Belastbarkeit werden mit dem LAMBDA – Lernen auswendig, Merken, Belastbarkeit, Denken-analytisch von Kubinger und Maryschka (nicht publiziert) erhoben.
- Belastbarkeit wird mit dem Rechnen in Symbolen von Schmotzer, Kubinger und Maryschka (1994) erhoben. Der RIS ist diesem Fall Bestandteil des LAMBDA und kommt in der Stressphase zum Einsatz.
- Das Anspruchsniveau an die eigenen Leistungen wird mit dem Kodiertest (Subtest 2) der Arbeitshaltungen von Kubinger und Ebenhöf (1996) erhoben.
- Studienrelevantes Vorwissen und Selbsteinschätzung werden mittels LEWITE Psychologie von Frebort (nicht publiziert) erhoben. Dieser beruht auf dem LEWITE – Lexikon Wissenstest von Wagner-Menghin (2004).
- Die Passung der Interessen wird mittels des Allgemeinen Interessen-Struktur-Test revidiert – AIST-R von Bergmann und Eder (2005) erhoben.
- Die Fähigkeit zum logischen (schlussfolgernden) Denken wird mittels Matrizenaufgaben erhoben, welche von Zimmerhofer und Hornke 2007 für das Self-Assessment adaptiert wurden (nicht publiziert).
- Planvolles Vorgehen und die Freude am Lernen und der Umgang mit Misserfolg werden mittels eines Fragebogens zum Arbeitsverhalten (nicht publiziert) erhoben. Dieser beruht auf dem HAKEMP 90 – dem Fragebogen zur Handlungskontrolle von Kuhl (2005/2006).

Nach der Durchführung der genannten Verfahren bekommt die Testperson eine Rückmeldung über ihre Ergebnisse. Für jedes Verfahren ist ein eigener Abschnitt vorgesehen, in dem das Verfahren nochmals kurz beschrieben wird, um der Testperson die Zuordnung der Verfahren zu den Aufgaben zu erleichtern. Es folgt eine Erklärung, was das Verfahren misst und inwieweit diese Fähigkeit für das Psychologiestudium von Bedeutung ist. Abschließend wird darauf hingewiesen, in welchem Ausmaß die jeweiligen Fähigkeiten ausgeprägt sein sollten, bevor der Testperson mitgeteilt wird, wie sie im Vergleich zu anderen Testpersonen abgeschnitten hat.

Es erfolgt jedoch „keine Rückmeldung der Art »geeignet« vs. »nicht geeignet« [...], sondern ein

Stärken-Schwächen-Profil mit Tipps, wie sie [die Testpersonen, Anm. der Verfasserin] am besten mit konkreten Defiziten umgehen können“ (Kubinger et al., 2007, S. 330).

Bezüglich jedes Verfahrens wird folglich rückgemeldet, ob die erhobene Fähigkeit bzw. Einstellung als Stärke oder als Schwäche in Bezug auf das Psychologiestudium zu betrachten ist.

Das Self-Assessment für das Psychologiestudium an der Universität Wien ist bereits seit Oktober 2005 im Einsatz (Frebort & Kubinger, 2006).

Fritsch (2008) prüfte das Self-Assessment erstmals auf prognostische Validität hinsichtlich der späteren Studienleistungen der Teilnehmer. Diese ersten Validierungsergebnisse lieferten jedoch kaum zufrieden stellende Ergebnisse, da „[...]nur eine sehr geringe Anzahl an Prüfungsnoten zur Bildung des Validierungskriteriums vorhanden waren“ (S. 49). Fritsch vermutet: „Der Grund für die geringe Anzahl verfügbarer Prüfungsnoten dürfte also darin liegen, dass eine Vielzahl der TestteilnehmerInnen am Self-Assessment das Psychologiestudium gar nicht erst beginnt oder nur kurze Zeit betreibt“ (S. 66).

Da eine hohe prognostische Validität in Bezug auf den Ausbildungserfolg erwünscht wäre, solche Belege jedoch noch nicht vorliegen, meinen Kubinger et al. (2007): „Wenn aber, wie zumeist, eine belegte Konstruktvalidität als Argument für den Einsatz eines bestimmten psychologisch-diagnostischen Verfahrens vorgebracht wird, dann muss dieses Argument auch für das Self-Assessment gelten“ (S.323).

Demzufolge versuche ich in dieser Arbeit durch regelgeleitete Itemkonstruktion bereits in der Konstruktionsphase die Validität (Konstruktvalidität und Kontentvalidität) so hoch wie möglich zu halten.

2 Regelgeleitete bzw. rationale Itemkonstruktion

Seit einigen Jahren liegt ein besonderes Augenmerk darauf, dass bereits bei der Entwicklung eines diagnostischen Verfahrens zu Gunsten der Validität die Itemkonstruktion rational bzw. regelgeleitet erfolgt.

„Rationale Itemgenerierung“ bedeutet, daß für die Erstellung von Items zu einem bestimmten Bereich so weitgehende Vorschriften entwickelt werden, daß voneinander unabhängig arbeitende Testautoren zu gleichartigen, d.h. weitgehend übereinstimmenden Items gelangen, und zwar geschieht dies in einem ersten Schritt und nicht erst durch eine Auswahl im Anschluß an irgendwelche Analysen. (Feger, 1984, S. 24)

Schon in einigen früheren Arbeiten wird kritisiert, dass die Testkonstruktion bisher eher intuitiv, das heißt erfahrungsgeleitet stattfand, und erst im nachhinein Items durch empirische Aufgabenanalysen abgesichert wurden (vgl. Klauer, 1978, 1984; Schott & Wieberg, 1984; Hornke & Habon, 1984b, 1986). Daher versuchen immer mehr Autoren durch regelgeleitete bzw. rationale Itemkonstruktion ihr Vorgehen zu rechtfertigen bzw. abzusichern. Bereits 1984 vertraten Schott und Wieberg die Ansicht, dass durch regelgeleitete Itemkonstruktion die inhaltliche Validität⁴ (Kontentvalidität) sichergestellt werden könne. Ebenso meint Embretson (1999), dass die Annahme eines kognitiven Designs bei der Itemgeneration den Vorteil hat, dass die Konstruktvalidität bereits auf Itemebene gegeben ist.

Vor der Itemkonstruktion sollte folglich genau durchdacht werden, wie man am besten Items für den interessierenden Fähigkeitsbereich bzw. Persönlichkeitsbereich erstellen könnte. Hornke und Habon (1984b) meinen, es „müßte vor der Aufgabenkonstruktion eine Konstruktionstheorie entwickelt und validiert werden, die garantiert/wahrscheinlich macht, dass jedes konstruierte Item zu der durch den Regelsatz inhaltlich definierten homogenen Grundmenge (Itemuniversum) gehört“ (S. 204).

4 Inhaltliche Validität liegt vor, „wenn das eignungsdiagnostische Verfahren selbst, quasi definitionsgemäß, das optimale Kriterium der interessierenden Eigenschaft darstellt. D.h., die ausgewählten Aufgaben sollen repräsentativ für alle prinzipiell zur Messung der Eigenschaft geeigneten Aufgaben sein.“ (Westhoff, Hellfritsch, Hornke, Kubinger, Lang, Moosbrugger, Püschel & Reimann, 2004, S. 188).

Nach Klauer (1978) sollte die Itemkonstruktion in drei Schritten ablaufen: zuerst sollte der Inhalt bzw. der Sachverhalt, der sich später in den Aufgaben wiederfinden soll, vollständig dargestellt werden. Danach muss eine geeignete Aufgabenform gewählt werden, das heißt ein geeignetes Itemformat. Zuletzt sollten Regeln formuliert werden, welche die Transformation des Inhalts in die Aufgabenform ermöglichen, wodurch eine Grundmenge von Items entsteht.

Holling, Blank, Kuchenbäcker und Kuhn (2008, S. 367) sind zudem der Ansicht, dass Itemeigenschaften, welche die Itemschwierigkeit⁵ bedeutsam beeinflussen, ebenfalls die kognitiven Prozesse, die während des Lösungsprozesses auftreten, bestimmen müssen. Im Weiteren meinen Hornke und Rettig (1989, S. 141), dass „Regeln unter Bezug auf substantielle psychologische Ergebnisse über kognitive Prozesse beim Problemlösen auch Vorhersagen der Itemschwierigkeiten erlauben“ können.

Aus diesem Grund sollen nach Hornke, Küppers und Etzel (2000) schwierigkeitsgenerierende Faktoren für die Entwicklung von Items festgelegt werden. Hat man nun die Absicht, Itemschwierigkeiten bestimmen zu wollen, sollten die Regeln schwierigkeitsgenerierende Faktoren sein, die einen Bezug zu den kognitiven Anforderungen haben, welche für das Lösen der Items benötigt werden. Diese Faktoren bilden schließlich ein Itemuniversum und können durch Kombination miteinander unterschiedliche (bestimmbare) Schwierigkeitsgrade von Items erzeugen (Hornke, Küppers & Etzel, 2000).

Abschließend sei jedoch darauf hingewiesen, dass einer empirischen Aufgabenanalyse immer eine rationale Aufgabenanalyse vorausgehen sollte, dass jedoch auf empirische Validierungen nicht verzichtet werden kann (Schott & Wieberg, 1984).

5 Bei Persönlichkeitsfragebögen ist es nach Kubinger (2009b) treffender „vom „Grad der Herausforderung“ zu sprechen, mit dem ein Item die Tp konfrontiert, in bestimmter Weise zu reagieren“ (S. 83).

3 Methodik

Im folgenden Kapitel sollen die im Rahmen der statischen Analysen zur Anwendung kommenden Verfahren kurz beschrieben werden, um das Verständnis des Empirischen Teils dieser Arbeit für den Leser zu erhöhen.

Ziel dieser Arbeit ist es, Verfahren zu entwickeln, für die das Rasch-Modell gilt. Zusätzlich soll mittels des Linear Logistischen Test-Modells überprüft werden, ob die theoretisch angenommenen kognitiven Operationen geeignet sind, die Schwierigkeiten der Items abzuschätzen (vgl. Formann, 1973; Kubinger, 2008, 2009a).

3.1 Das dichotom logistische Modell von Rasch (1 PL-Modell)

Das dichotom logistische Modell von Rasch (1980) findet seine Anwendung bei dichotomen Daten. Diese Daten ergeben sich, wenn ein Verfahren nur zwei Antwortkategorien zulässt, zum Beispiel richtig und falsch bzw. ja und nein. Für andere Daten gibt es Erweiterungen des Modells, daher wird im folgenden Abschnitt nur noch allgemein vom Rasch-Modell gesprochen. Das Rasch-Modell beruht auf einer probabilistischen itemcharakteristischen (IC) Funktion, welche die Wahrscheinlichkeit für eine Reaktion einer Person bei einem Item festlegt (Kubinger, 1989).

Das Modell „beschreibt die Wahrscheinlichkeit, dass Testperson (Tp) v Item í löst („+“), in Abhängigkeit eines Personenparameters ξ_v , d.i. die (wahre) Fähigkeit von v, und eines Itemparameters σ_i , d.i. die (wahre) Schwierigkeit von í“ (Kubinger, 2003, S. 416).

Die Wahrscheinlichkeit ein Item zu lösen $P(+)$ bzw. nicht zu lösen $P(-)$ wird durch folgende Formeln dargestellt (Kubinger, 1989):

$$P(+|\xi_v, \sigma_i) = \frac{e^{\xi_v - \sigma_i}}{1 + e^{\xi_v - \sigma_i}} \quad (1.1)$$

$$P(-|\xi_v, \sigma_i) = \frac{1}{1 + e^{\xi_v - \sigma_i}} \quad (1.2)$$

Die Wahrscheinlichkeit ein Item zu lösen $P(+)$ wird, wie aus (1.1) ersichtlich, umso höher, je höher die Fähigkeit einer Person (also ξ_v) ist bzw. je niedriger die Itemschwierigkeit (also σ_i) ist. Die Wahrscheinlichkeit ein Item nicht zu lösen $P(-)$, vgl. (1.2), ist demnach umso größer, je niedriger ξ_v und je höher σ_i ist. Zusätzlich kann postuliert werden, dass $P(+)$ dann 0,5 ist, wenn das Item genauso schwierig bzw. leicht ist, wie die Person fähig ist, das heißt $\xi_v = \sigma_i$ (Kubinger, 2003).

Explizit „hängt die Wahrscheinlichkeit einer richtigen Antwort nur von der Differenz zwischen dem Personen- und dem Aufgabenparameter ab“ (Fischer, 1974, S. 209).

Als Grundannahme des Rasch-Modells gilt die *lokal stochastische Unabhängigkeit*, das heißt, dass die Lösung einer Aufgabe nur von der Fähigkeit einer Person und der Schwierigkeit einer Aufgabe abhängt und nicht davon, welche Aufgaben die Person bereits bearbeitet hat oder noch bearbeiten wird (Kubinger, 2003).

Das Rasch-Modell geht zudem von einer *erschöpfenden Statistik* aus. Damit ist gemeint, dass die Anzahl der gelösten Items die gesamte Information über eine Person hinsichtlich der gemessenen Fähigkeit beinhaltet, egal welche Items die Person gelöst hat (Kubinger, 1989).

Ferner gilt die Annahme der *Eindimensionalität*, dass also alle Items bzw. Aufgaben dieselbe latente Fähigkeit bzw. Eigenschaft messen, sprich homogen sind (Fischer, 1974).

Folgender Nutzen ergibt sich, wenn Tests bzw. Items dem Rasch-Modell folgen:

1. Das Rasch-Modell ermöglicht „*spezifisch objektive Vergleiche*“. Das bedeutet, dass der Vergleich zweier Personen unabhängig davon ist, welche und wie viele Items die Personen bearbeitet haben. Ebenso sollte es beim Vergleich zweier Items keine Rolle spielen, welche Personen diese Items bearbeitet haben (Kubinger, 1989, S. 23). Dadurch kann dem Modell *Stichprobenunabhängigkeit* im statistischen Sinne zugeschrieben werden (Kubinger, 2003). „Das heißt, die Schätzungen der Parameter sind insofern stichprobenunabhängig, als die Wahl der Stichprobe aus einer bestimmten Population für die statistische Inferenz dieser Parameter keine Rolle spielt“ (Kubinger, 2009b, S.89). Diese Eigenschaft des Rasch-Modells bildet die Grundlage für die Prüfbarkeit des Modells. Denn lässt sich feststellen, dass die Differenz zwischen den Parametern σ_i und σ_j von Stichprobe zu Stichprobe unterschiedlich ist, würde das den Annahmen des Modells widersprechen und das Rasch-Modell könnte nicht gelten (Kubinger, 2009b).
2. Durch die Geltung des Rasch-Modells ist zudem die Grundlage für adaptives Testen

gegeben, da ein fairer Vergleich, der Leistungen von zwei Testpersonen, welche bei einem psychologisch diagnostischen Verfahren unterschiedliche Items bearbeitet haben, dann erst möglich ist (Kubinger, 2000).

3. Weiters ist das Rasch-Modell nach Kubinger (2000) unabdingbar, um die Anzahl der gelösten Aufgaben eines psychologisch-diagnostischen Verfahrens als Gesamtscore einer Person heranzuziehen. Gemeint ist hiermit, wie bei vielen psychologisch-diagnostischen Verfahren praktiziert, dass Items als gelöst oder nicht gelöst codiert werden und schließlich die Anzahl der gelösten Items als Wert für die Fähigkeit einer Person angesehen wird. Bereits Fischer (1974) erbrachte den Beweis, dass dieser Verrechnungsmodus nur dann fair ist, wenn für die Items eines psychologisch-diagnostischen Verfahrens das Rasch-Modell gilt. Kommt ein solcher Verrechnungsmodus zur Anwendung, dann müsste davon ausgegangen werden, dass, egal welche Items eine Person bearbeitet hat, die Anzahl der gelösten Aufgaben die gesamte Information (welche erhoben wurde) über eine Person beinhaltet. Womit eigentlich die Annahme der erschöpfenden Statistik zu Grunde liegen sollte. Ebenso müsste für diese Verrechnung die Annahme der Eindimensionalität vorausgesetzt sein.

Ob ein Verfahren bzw. Items dem Modell von Rasch folgen und somit die oben genannten Eigenschaften aufweisen, kann mittels Modelltests überprüft werden. Wie bereits erwähnt, bildet die Annahme der Stichprobenunabhängigkeit eine Möglichkeit um die Geltung des Modells zu überprüfen. Dies ist möglich indem man die Stichprobenunabhängigkeit hinterfragt. Dafür teilt man die Stichprobe in Teilstichproben und berechnet getrennte Itemparameter. Die Teilstichproben können sich dadurch ergeben, indem man die Stichprobe anhand eines externen Kriteriums (Geschlecht) oder internen Kriteriums (Score) in zwei Teile teilt. Neben der Möglichkeit eines graphischen Modelltests zur Feststellung, welche Items sich als auffällig erweisen, kann zum Beispiel der Likelihood-Quotienten-Test von Andersen eingesetzt werden, um die Stichprobenunabhängigkeit zu hinterfragen. Dieser LQT erprobt schließlich, ob die erhobenen Daten durch die in den Teilstichproben geschätzten Parameter besser beschrieben werden können als durch die Parameter in der Gesamtstichprobe. Unterscheiden sich die Itemparameter nicht bzw. nur zufällig, dann kann Modell-Konformität für die Daten angenommen werden (Kubinger, 1989).

3.2 Das Lineare Logistische Test-Modell von Fischer (LLTM)

Das LLTM ist eine Erweiterung bzw. ein Spezialfall des Rasch-Modells. Das Modell entstand aus der Idee heraus, dass die Itemschwierigkeit eine Funktion der Schwierigkeit der beteiligten kognitiven Operationen ist (Fischer, 1995).

Inhaltlich bedeutet das LLTM, dass sich die Schwierigkeit jeder Aufgabe aus der je nach Art und Anzahl gewichteten Summe der Schwierigkeiten einzelner Teiloperationen zusammensetzt. Diese Teiloperationen repräsentieren dabei die Hypothese, welche kognitiven Lösungsschritte es zur Lösung der Aufgabe als Ganzes bedarf. (Kubinger, 2003, S. 421)

Formell ist das LLTM definiert als ein Rasch-Modell mit der folgenden Zusatzannahme (Fischer, 1995):

$$\sigma_i = \sum_{j=1}^m q_{ij} \eta_j \quad (1.3)^6$$

Der Itemparameter σ_i kann somit in eine gewichtete Summe von Basisparametern $\eta_1, \eta_2, \dots, \eta_n$ zerlegt werden. Unterschiede bezüglich der Schwierigkeit zwischen Items lassen sich beim LLTM darauf zurückführen, dass bei einem leichteren Item weniger kognitive Operationen beteiligt sind als bei einem schwierigeren (Fischer, 1995).

Die Schätzung der Basisparameter η_j erfolgt wie beim Rasch-Modell vorzugsweise mit der CML-Methode (Fischer, 1995).

Die Feststellung der Gültigkeit des LLTM benötigt nach Fischer (1995) zwei unabhängige Schritte:

1. die Feststellung der Gültigkeit des Rasch-Modells
2. die Feststellung der Zusatzannahme (im Vergleich zum Rasch-Modell) des LLTM.

Ist die Gültigkeit des Rasch-Modells gegeben, wird mit einem weiteren Likelihood-Ratio-Test überprüft, ob die Daten durch m statt durch k bzw. durch $k-1$ Parameter genauso gut zu erklären

6 Zum besseren Verständnis wurde die Notation in Anlehnung an Kubinger (1989) beibehalten.

sind (z.B. nach Kubinger, 1989, S. 34):

$$Z = -2 \ln \frac{L(\text{Daten der } S | \hat{\eta})}{L(\text{Daten der } S | \hat{\sigma})} \quad (1.4)$$

Fischer (1974) weist darauf hin, dass der Likelihood-Ratio-Test in diesem Fall prüft, ob die Annahme der Basisparameter beibehalten werden kann oder verworfen werden muss.

Hat man schließlich für einen Test die Gültigkeit des LLTM festgestellt, kann man für jedes neu konstruierte Item die Gültigkeit des Rasch-Modells annehmen, vorausgesetzt man verwendet hierfür die festgestellten Teiloperationen zur Konstruktion. Zudem lassen sich mit der Kombination der Teiloperationen Items mit gewünschter Schwierigkeit erzeugen (Kubinger, 2003).

4 Matrizentest

Wie in Kapitel 1.2 beschrieben, beinhaltet das Anforderungsprofil des Psychologiestudiums an der Universität Wien die Fähigkeit zum logischen (folgerichtigen) Denken bzw. schlussfolgernden Denken (engl. reasoning). Da nach Kubinger (2009b) Matrizentests (zum Beispiel Wiener Matrizen-Test, vgl. 4.3.1) zu den typischen Repräsentanten der Leistungstests gehören, welche die Fähigkeit „reasoning“ erheben, sollte daher im Rahmen dieser Arbeit für das Self-Assessment Psychologie ein Matrizentest entwickelt werden.

Das folgende Kapitel gliedert sich in drei Abschnitte. Im ersten Abschnitt wird die Fähigkeit des schlussfolgernden Denkens skizziert und im zweiten Abschnitt werden bereits bestehende Matrizentests vorgestellt, insbesondere jene, welche für diese Arbeit von besonderer Bedeutung waren. Im dritten und letzten Abschnitt werden der für diese Arbeit neu entwickelte Färbige Matrizentest - FMT sowie seine Konstruktionsprinzipien vorgestellt.

4.1 Schlussfolgerndes Denken

Schlussfolgerndes Denken gehört zu unserem Alltag und ist eine grundlegende Fähigkeit, mit deren Hilfe bereits der Säugling aus Konsequenzen, die sein Verhalten hat, auf zukünftige schließt und so versucht, Regelmäßigkeiten zu finden und damit Ordnung in seine Umwelt zu bringen. (Oerter & Dreher, 2002, S. 487)

Explizit definiert Kubinger (2009b) „Reasoning ist die Fähigkeit, Gesetzmäßigkeiten oder logisch zwingende Zusammenhänge erkennen und zweckentsprechend verwerten zu können“ (S. 206). Schlussfolgerndes Denken ist folglich eine Fähigkeit, die sich schon früh im Leben zu entwickeln beginnt und uns das adäquate Zurechtfinden in unserer Umwelt ermöglicht.

Da nicht alle Menschen die gleichen Fähigkeiten und Fertigkeiten aufweisen, zumal sich das schlussfolgernde Denken individuell entwickelt, ist es der Psychologie ein besonderes Anliegen, diese Fähigkeit zu messen. Es wird darauf hingewiesen, dass beim schlussfolgernden Denken Unterscheidungen getroffen werden können. In diesem Zusammenhang unterscheiden zum Beispiel Beckmann und Guthke (1999) sowie Oerter und Dreher (2002) drei Arten des

schlussfolgernden Denkens: den deduktiven Schluss, den induktiven Schluss und den Analogieschluss. Im Folgenden wird auf diese genauer eingegangen.

4.1.1 Deduktiver Schluss

Deduktives Schließen erfolgt vom Allgemeinen auf das Spezielle. Der Schluss ergibt sich zwingend und ist somit gesichert (Oerter & Dreher, 2002).

Ein Beispiel für einen deduktiven Schluss anhand eines Syllogismus:

- (a) Alle Blumen sind violett.
- (b) Der Löwenzahn ist eine Blume.
- (c) Der Löwenzahn ist violett.

Die Konklusion (c) basiert auf einer Deduktion. Es geht folglich um das Erkennen von Sachverhalten, die nicht explizit dargestellt, jedoch durch das Dargestellte beinhaltet sind. Aus dem Beispiel lässt sich erkennen, dass solche Aussagen nicht inhaltlich richtig sein müssen, um eine Deduktion zuzulassen. Nach Beckmann und Guthke (1999) ist es bedeutsam, dass es bei der Konklusion im Falle eines deduktiven Schlusses zu keiner Erweiterung der Prämisseninformation kommt. Die Autoren postulieren zudem, dass deduktives Schließen benötigt wird, um generelle Prinzipien und Regeln anzuwenden.

4.1.2 Induktiver Schluss

Dem deduktiven Schließen wird das induktive Schließen gegenübergestellt. Hier fungiert der Schluss vom Speziellen auf das Allgemeine. Durch das Erkennen von Regelmäßigkeiten oder Zusammenhängen sowie durch Erfahrungen und Wahrscheinlichkeiten schließen wir von mehreren speziellen Fällen auf das Allgemeine (Oerter & Dreher, 2002). Wenn man das obige Beispiel etwas umformt, könnte ein Beispiel für einen induktiven Schluss anhand eines Syllogismus lauten:

- (a) Der Löwenzahn ist violett.
- (b) Der Löwenzahn ist eine Blume.
- (c) Alle Blumen sind violett.

Die Konklusion (c) wäre in diesem Fall ein induktiver Schluss, der nicht zwingend, aber möglich ist. Man folgert sozusagen von allen einem selbst bekannten Fällen auf einen allgemeinen Schluss. Auf diese Weise können sich neue Erkenntnisse ergeben, da das Wissen durch die Prämissen (a und b) in der Konklusion erweitert wird. Konklusionen bei induktiven Schlüssen sind allerdings mit Unsicherheit verbunden, da nicht genügend Information in den Prämissen steckt, um diese Konklusionen eindeutig behaupten zu können (Beckmann & Guthke, 1999). Der induktive Schluss könnte folglich leicht widerlegt werden, indem man nur eine Beobachtung macht, die dieser Aussage nicht entspricht (man sieht beispielsweise eine gelbe Blume). Induktive Schlüsse finden sich im Alltag zum Beispiel in Form von Vorurteilen wieder. Aber auch die Wissenschaft bedient sich der Methode des induktiven Schließens (Oerter & Dreher, 2002).

Beckmann und Guthke (1999) betonen, dass die Unterscheidung zwischen dem induktiven und dem deduktiven Schluss bezüglich der Schlussrichtung unzureichend sei, und eine Unterscheidung hinsichtlich der Informationserweiterung erforderlich ist.

4.1.3 Analoger Schluss

Das analoge Schließen ist eine Art induktiver Schluss, der sich jedoch auf die Gleichwertigkeit von Verhältnissen bezieht (Oerter & Dreher, 2002). Man erkennt zum Beispiel eine Ähnlichkeit zwischen dem Verhältnis A:B und dem Verhältnis C:D. Ebenso versteht man darunter, dass ein Objekt bzw. Gegenstand durchleuchtet wird und auf dieser Grundlage eine Schlussfolgerung bezüglich eines anderen Objekts bzw. Gegenstands getätigt wird. Diese Schlüsse sind ebenfalls nur wahrscheinlich (Beckmann & Guthke, 1999).

Ein Beispiel, wie analoges Schließen erhoben werden könnte:

Fohlen:Stute = Kücken: ?

a) Hahn, b) Ente c) Hengst d) Henne e) Vogel

In obigem Beispiel geht es um das Erkennen der Relation zwischen den Begriffen Fohlen und Stute. Anschließend soll diese Relation mit dem Begriff Kücken und einem weiteren hergestellt werden. Die Lösung wäre somit d) Henne, da die Stute der weibliche Elternteil des Fohlen ist und die Henne der weibliche Elternteil des Kücken.

Goswami (2001) meint, dass wir mit Hilfe des analogen Schließens einen sehr großen Teil unserer Alltagsprobleme lösen können: „Wir benutzen Analogien jedes Mal, wenn wir uns an eine vergangene Situation erinnern, um mit einer neuen Situation umzugehen“ (S. 294).

Bei der Feststellung der Fähigkeit des schlussfolgernden Denkens sollte daher beachtet werden, welche Art bzw. welche Arten von Schlüssen vollzogen werden.

Aufgaben des schlussfolgernden Denken in Intelligenztests geben in Worten oder Symbolen eine Auswahl möglicher Lösungen vor, aus denen diejenige herauszufinden ist, welche dem vorher bekanntgegebenen Lösungsprinzip entspricht, oder sie bedienen sich eingekleideter Aufgaben welche durch Anwendungen eines Prinzips zu lösen sind. (Fröhlich, 2002, S. 391)

Wie in der Definition von Fröhlich ersichtlich, gibt es unterschiedliche Methoden, mit denen man versucht schlussfolgerndes Denken zu messen. Beckmann und Guthke (1999) nennen zum Beispiel Klassifikationsaufgaben, Analogien, Vervollständigung von Reihen (zum Beispiel Zahlenreihen), Syllogismen und Matrizenaufgaben. Je nach Aufgabenform kann es sein, dass eine unterschiedliche Art des schlussfolgernden Denkens notwendig ist, um die Aufgaben zu lösen.

Da nun thematisiert wurde, welche Arten des schlussfolgernden Denkens es gibt und welche Aufgabenformen, um schlussfolgerndes Denken zu erheben, interessiert gegenwärtig, welche Art des schlussfolgernden Denkens bei Matrizentests gemessen wird bzw. gemessen werden soll.

4.2 Schlussfolgerndes Denken bei Matrizentests

Der Definitionssatz zum induktiven Denken von Klauer (1993, S. 143) besagt:

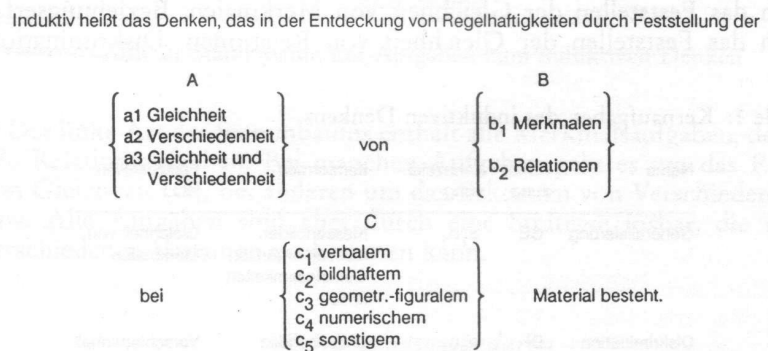


Abb. 1: Definitionssatz zum induktiven Denken

Abb. 1: Definitionssatz zum induktiven Denken von Klauer (1993, S. 143)

Durch die unterschiedlichen Kombinationen der Variablen A, B und C aus Abb. 1 mit ihren Ausprägungen berücksichtigt Klauer 30 verschiedene Variationen des induktiven Denkens. Er betrachtet die Facetten A und B als Kernfacetten, werden diese miteinander kombiniert, dann ergeben sich sechs Kernaufgaben des induktiven Denkens (siehe Tab. 1) (Klauer, 1993).

Tabelle 1: Kernaufgaben des induktiven Denkens nach Klauer (1993, S. 144)

Name	Kürzel	Kennzeichnung	Itemformen	Festzustellen ist...
Generalisierung	GE	a1b1	Klassen bilden, Klassen ergänzen, Gemeinsamkeiten finden	Gleichheit von Merkmalen
Diskrimination	DI	a2b1	Unpassendes streichen	Verschiedenheit von Merkmalen
Kreuzklassifikation	KK	a2b1	Vierfelderschema, Sechsfelderschema, Neunfelderschema	Gleichheit und Verschiedenheit von Merkmalen
Beziehungserfassung	BE	a1b2	Folgen ordnen, Folgen ergänzen, Einfache Analogie	Gleichheit von Relationen
Beziehungsunterscheidung	BU	a2b2	Gestörte Folge	Verschiedenheit von Relationen
Systembildung	SB	a3b2	Matrize, vollständige Analogie	Gleichheit und Verschiedenheit von Relationen

Demzufolge gehören Matrizentestitems zur Kernaufgabe der Systembildung, bei welcher die Gleichheit und Verschiedenheit von Relationen festgestellt wird.

Dem soll hier nicht widersprochen werden. Doch betrachtet man das Lösen einer

Matrizenaufgabe genauer, dann geht es neben dem Suchen und Erkennen von Merkmalen (bzw. Regeln) auch um die Anwendung dieser (vgl. Formann, 1973; Putz-Osterloh, 1981). Falls eine Testperson ein Item zu lösen versucht, dessen Regelstruktur der Person noch nicht bekannt ist, da zum Beispiel im Beispielitem die Regel nicht vorgekommen ist, dann erzeugt sie bei der Suche nach Merkmalen (bzw. Regeln) selbst eine Regel bzw. ein System, welches die gleichen bzw. verschiedenen Merkmale oder Relationen in einem Item in Zusammenhang bringt. Die Testperson generiert sozusagen durch das Erkennen von Zusammenhängen für sie „neues Wissen“. Sie kann zum Beispiel schlussfolgern, dass sich die spezielle Beobachtung in Spalte eins einer Matrix auf die anderen beiden Spalten verallgemeinern lässt. Dies wäre der Definition von Kapitel 4.1.2 zu Folge ein induktiver Schluss und würde sich mit Klauers Annahme decken. Hat eine Testperson nun eine Regel aufgestellt bzw. erfasst oder ist ihr diese aus vorherigen Items bereits bekannt, dann geht es in einem weiteren Schritt darum, diese Regel auch richtig anzuwenden (zum Beispiel in Spalte 2 oder beim nächsten Item). Die Anwendung von Regeln unterliegt der Definition nach dem deduktiven Denken (vgl. 4.1.1). Auch die Tatsache, dass sich eine Lösung zwingend durch die Anwendung von zuvor festgelegten Regeln ergibt, spricht ebenfalls für einen deduktiven Schluss beim Lösen einer Matrizenaufgabe. Beckmann und Guthke (1999) argumentieren ebenfalls:

Die Trennung von Induktion und Deduktion ist eine metaphysische, eigentlich ist jede Deduktion das Resultat einer vorausgegangenen Induktion; das Resultat induktiven Denkens manifestiert sich letztlich über einen deduktiven Denkakt. Die erfolgreiche Anwendung (Deduktion) der induktiv gewonnenen Regel auf den Einzelfall (=nächstes Item) soll dem Diagnostiker Auskunft darüber geben, ob die Regel erfolgreich induziert wurde. (S. 6)

Zusammenfassend kann gesagt werden, dass Klauers Annahme des induktiven Denkprozesses bei Matrizenaufgaben zutrifft, dieser jedoch alleine nicht ausreichend ist, um eine Matrizenaufgabe zu lösen. Es wird somit angenommen, dass für das Lösen eines Matrizentests sowohl induktives als auch deduktives Denken notwendig ist, und daher schlussfolgerndes Denken als Ganzes und nicht einzelne Komponenten davon erhoben werden. Demnach ist ein Matrizentest eine gute Wahl, um die Fähigkeit „reasoning“ zu erheben.

4.3 Bestehende Matrizentests

Zu den bekanntesten Matrizentests zählen Ravens Progressive Matrices, welche seit etwa 70 Jahren in Verwendung sind. John C. Raven veröffentlichte 1938 die Standardform des Progressiven Matrizen-Tests – SPM (Heller, Kratzmeier & Lengfelder, 1998a), welcher für alle Anwender geeignet sein sollte. Nachdem Raven erkannt hatte, dass es einen Bedarf für spezielle Zielgruppen gab, entwickelte er auch die Verfahren Coloured Progressive Matrices – CPM (J. C. Raven, Court & J. Raven, 2001) und Advanced Progressive Matrices – APM (Heller, Kratzmeier & Lengfelder, 1998b). Die CPM finden ihre Anwendung bei Kindern, minderbegabten Erwachsenen und im klinisch psychologischen Bereich zum Beispiel in der Gerontopsychologie (J. C. Raven & J. Raven, 2001). Der APM wurde hingegen mit schwierigeren Aufgaben entwickelt, um dem Deckeneffekt, den man im SPM finden konnte, entgegenzuwirken und um folglich bei Personen mit überdurchschnittlichen Fähigkeiten besser differenzieren zu können (Heller, et al. 1998b).

Neben den Matrizentests von Raven gibt es unter anderem den Wiener Matrizen-Test – WMT (Formann & Piswanger, 1979) sowie den Adaptiven Matrizentest - AMT (Hornke, Etzel & Rettig, 1999). Diese beiden Verfahren sind aufgrund ihrer Entwicklung von besonderer Bedeutung für meine Testentwicklung und werden daher im folgenden Abschnitt genauer beleuchtet.

4.3.1 Wiener Matrizentest – WMT (Formann & Piswanger, 1979)

Der WMT misst logisch schlussfolgerndes Denken im Sinne der sprachfreien Intelligenz. Sowohl das Konzept als auch die Intention des WMT sind den Matrizentests von Raven sehr ähnlich. Im Gegensatz zu Raven wurde jedoch beim WMT ein besonderes Augenmerk auf die Entwicklung und Auswahl der Items gelegt. Formann definierte bereits in seiner Dissertation (1973) ein Itemuniversum, welches dem WMT zugrunde liegt. Dieses Itemuniversum bezieht sich auf jene Faktoren, welche eine Matrizenaufgabe charakterisieren und somit deren Schwierigkeit maßgeblich beeinflussen. Mit vorab konstruierten Regeln entwickelte er (1973) 42 Items, von denen sich 24 Rasch-Modell konforme Items im WMT wiederfinden.

Konzept des WMT

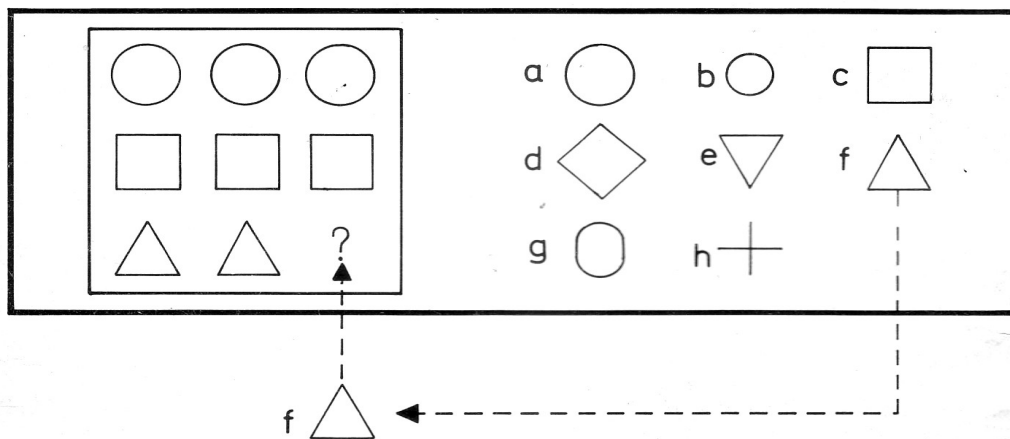


Abb. 2: Einführungsbeispiel aus dem Aufgabenheft des WMT (Formann & Piswanger, 1979, S. 2)

Die Aufgaben des WMT bestehen aus einer Matrix mit 9 Figuren (vgl. linke Seite der Abb. 2). Diese 9 Figuren sind nach bestimmten Regeln angeordnet. Die neunte Figur (rechts unten) ist stets ein Fragezeichen. An die Stelle des Fragezeichens soll eine der acht Antwortmöglichkeiten (a-h), welche auf der rechten Seite der Aufgabe abgebildet sind, gesetzt werden. Die Aufgabe einer Testperson besteht darin, aus den vorgegebenen acht Alternativen diejenige auszuwählen, welche logisch sinnvoll an die Stelle des Fragezeichens gesetzt werden kann. Das Auswählen der richtigen Alternative gelingt der Testperson, indem sie die Regel/n, welche für die Figuren auf der linken Seite gelten, ausfindig macht.

Itemuniversum

Laut Formann (1973, S. 18) charakterisieren folgende Faktoren eine Aufgabe:

Faktor 1: die Anzahl der Komponenten, aus welchen jede Figur einer Matrizenaufgabe besteht.

Faktor 2: die für die Lösung relevanten Materialeigenschaften jeder Komponente einer Figur (das heißt: Form, Muster, Anzahl, räumliche Anordnung).

Faktor 3: die Regeln, welche zur gegebenen Anordnung führen, wobei auf jede relevante Materialeigenschaft einer Komponente nur eine Regel anzuwenden ist.

Faktor 4: die Richtung der Regelgeltung (waagrecht, senkrecht, waag- und senkrecht).

Formann (1973) meint, dass durch die Variation der Anzahl der Faktoren - eins bis vier - Items mit unterschiedlichem Schwierigkeitsgrad konstruiert werden können. Durch die Steigerung der Komponenten, durch die Steigerung der für die Lösung relevanten Materialeigenschaften und

durch die Anwendung mehrerer Regeln in unterschiedliche Richtungen können folglich schwierigere Items erzeugt werden. „Zusammenfassend kann gesagt werden, dass F1 bis F4 zusammen mit der Menge aller Repräsentanten der Materialeigenschaften ein Itemuniversum definieren“ (S. 18).

In Bezug auf *Faktor 1* und *Faktor 2* beschreibt Formann (1973) die grundlegenden Merkmale der Figur einer Matrizenaufgabe genauer. Er ging davon aus, dass jede Figur eines Matrizentestitems aus mindestens einer *Komponente* besteht, wobei Komponenten jene Einheiten sind, aus denen eine Figur aufgebaut ist. Zudem könne man diese Komponenten einer Figur durch vier *Materialeigenschaften* (Form, Muster, Anzahl und räumliche Anordnung) beschreiben. Jede dieser Materialeigenschaften besitze unterschiedliche Repräsentanten. Die Materialeigenschaft Muster habe zum Beispiel die Repräsentanten: weiß, schwarz, gestreift, punktiert, kariert etc.

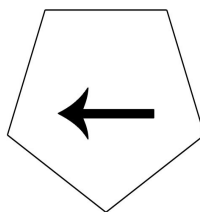


Abb. 3: Beispiel einer Figur mit zwei Komponenten

Zur Erläuterung: die Figur in Abb. 3 besteht aus zwei Komponenten. Die erste Komponente ist: ein (Anzahl) auf der Spitze (räumliche Anordnung) stehendes weißes (Muster) Fünfeck (Form). Die zweite Komponente ist: ein (Anzahl) nach links zeigender (räumliche Anordnung) schwarzer (Muster) Pfeil (Form).

Um nun einen Zusammenhang zwischen den einzelnen Figuren herzustellen, der schließlich auch die fehlende Figur determiniert, spricht Formann (1973) von Regeln (*Faktor 3*), die für jede einzelne Materialeigenschaft jeder Komponente gelten können, aber nicht müssen. Er unterscheidet folgende Regeln (welche sich nur auf relevante Materialeigenschaften einer Komponente beziehen):

Fortsetzen: Diese Regel besagt, dass pro Zeile und/oder Spalte dreimal derselbe Repräsentant einer Materialeigenschaft einer Komponente (zum Beispiel Form, Muster, Anzahl oder räumliche Anordnung) angeordnet ist.

Variieren: Diese Regel besagt, dass pro Zeile und/oder Spalte dieselben drei Repräsentanten

einer Materialeigenschaft angeordnet sind. Ihre Anordnung variiert dabei zufällig.

Überlagern: Diese Regel besagt, dass ein Repräsentant einer Materialeigenschaft einer Komponente pro Zeile und/oder Spalte dadurch entsteht, dass die beiden anderen Repräsentanten einer Materialeigenschaft in derselben Zeile und/oder Spalte übereinander gelegt werden.

Schließlich bezieht sich *Faktor 4* darauf, dass die Richtung der Regelgeltung je Materialeigenschaft einer Komponente unterschiedlich sein kann. Regeln können entweder waagrecht (je Zeile), senkrecht (je Spalte) oder waagrecht und senkrecht gelten.

Die 24 Items des WMT wurden als eine Stichprobe dieses Itemuniversums konstruiert. Sie messen nachweislich eindimensional und zwar bei allen Probanden. Der WMT ist ein speed- und power Test, er hat eine Zeitbegrenzung von 25 Minuten. Der Testwert einer Person ergibt sich durch die Anzahl der gelösten Aufgaben, mit Hilfe der Normtabellen kann ihr eine Position innerhalb ihrer Referenzstichprobe zugeschrieben werden (Formann & Piswanger, 1979).

4.3.2 Adaptiver Matrizentest – AMT (Hornke, Etzel & Rettig, 1999)

Der AMT misst die Fähigkeit zum schlussfolgernden Denken und ist ab dem 14. Lebensjahr einsetzbar. Das Konzept des Adaptiven Matrizentests entspricht im Großen und Ganzen dem des Wiener Matrizen-Tests (vgl. Kapitel 4.3.1).

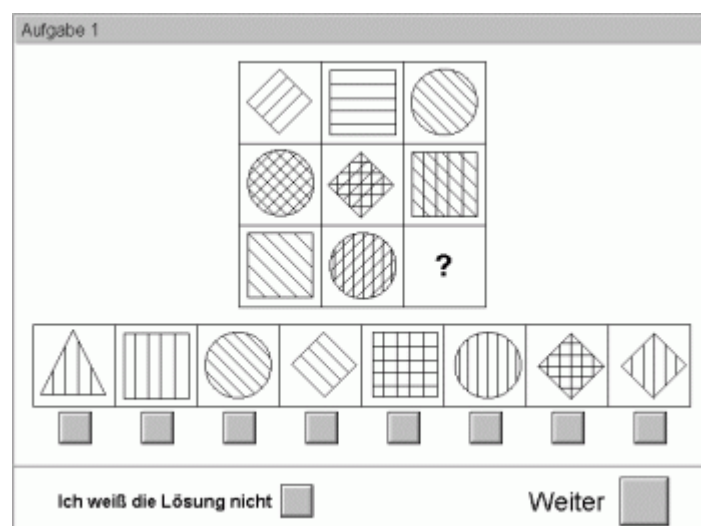


Abb. 4: Aufgabe 1 aus dem AMT; Quelle: <http://www.schuhfried.at/index.php?id=377&L=4>

Anstelle der Antwortmöglichkeiten auf der rechten Seite werden im AMT die Antwortmöglichkeiten unterhalb der Matrix dargeboten (siehe Abb. 4). Der Testperson stehen

außerdem neun Antwortmöglichkeiten zur Verfügung. Zusätzlich zu acht möglichen grafischen Antworten (im Sinne einer möglichen Figur) hat die Testperson die Möglichkeit, ein Kreuz bei der Antwortmöglichkeit „Ich weiß die Lösung nicht“ zu machen. Die Vorgabe der Items des AMT erfolgt zudem adaptiv. Ebenso können neben der Standardtestform (ca. 23 Items) noch drei weitere Formen vorgegeben werden: S1 - Screening: dient als kurzer Überblick mit ca. 13 Items, S3 - Precision: genauere Erfassung um Unterschiede besser feststellen zu können mit ca. 30 Items, S11 - Verkehrspsychologische Kurzform: ähnlich wie S1, mit leichteren Einstiegsitems. Die Durchführungsdauer beträgt inklusive Instruktions- und Übungsphase je nach Testform zwischen 20 und 60 Minuten. (Hornke, Etzel & Rettig, 1999).

Itemuniversum

Als Grundlage für die Entwicklung des AMT dienten die Itemkonstruktionsregeln von Hornke und Habon (1984a). Die Autoren wollten eine hinreichend große Menge homogener Items erstellen, mit denen adaptives Testen möglich ist, um folglich für jedes Fähigkeitsniveau genügend Items zu besitzen. Bei der Erstellung der Itemkonstruktionsregeln lehnten sie sich unter anderem an die Arbeit von Formann (1973) an. Sie gingen davon aus, dass die Schwierigkeit eines Items dadurch bestimmt ist, wie viele und welche kognitiven Operationen für die Lösung eines Items benötigt werden. Verschiedene Studien zeigten allerdings, dass die Items, welche mit den Regeln von Formann (1973) konstruiert wurden, nicht den Annahmen des LLTM (vgl. Kapitel 3.2) folgen (Hornke & Rettig, 1989). Außerdem kritisierten Hornke und Habon (1984a), dass diese Items für überdurchschnittliche Problemlöser nicht angemessen erscheinen. Sie entwickelten daher auf Basis der Analysen von bestehenden Matrizentests ein neues Regelsystem, welches die Konstruktion schwierigerer Items ermöglichen sollte. Diese mit dem neuen Regelsystem entwickelten Items sollten dem Rasch-Modell entsprechen und dem LLTM folgen (vgl. Hornke & Habon, 1984a; Hornke & Rettig, 1989; Hornke, Küppers & Etzel, 2000).

Hornke und Habon (1984a) unterscheiden drei schwierigkeitsgenerierende Komponenten. Erstens die Lösungsoperationen (die Regeln), welche entdeckt und angewandt werden sollen. Zweitens der Realisierungsmodus, der sich darauf bezieht, in welcher Beziehung die betroffenen Komponenten zueinander stehen. Und Drittens die Lösungsrichtung, welche sich auf den Bereich der Regelgeltung bezieht (vgl. auch Hornke & Rettig, 1989; Hornke, Küppers & Etzel, 2000). Im Folgenden werden die Regeln, wie sie in den Werken von Hornke und Habon (1984a), von Hornke und Rettig (1989) und von Hornke, Küppers und Etzel (2000) dargestellt wurden, kurz beschrieben.

1) Lösungsoperationen (Regeln)

Je Zeile und/oder Spalte des Schemas

- ...ist dreimal derselbe Repräsentant einer Komponente angeordnet ► *Identität*
- ...entsteht ein Repräsentant einer Komponente, indem die beiden restlichen, in derselben Zeile und/oder Spalte angeordneten Repräsentanten einer Komponente übereinander gelegt werden ► *Addition*
- ...entsteht ein Repräsentant einer Komponente, indem von einer vollständigen Figur eine Teilfigur ausgeblendet wird ► *Subtraktion*
- ...entsteht ein Repräsentant einer Komponente, indem von den beiden anderen in der Zeile und/oder Spalte angeordneten Repräsentanten einer Komponente nur jene Elemente übernommen werden, die sowohl in der einen als auch in der anderen Figur vorkommen ► *Disjunktion*
- ...entsteht ein Repräsentant einer Komponente, indem von den beiden anderen in der Zeile und/oder Spalte angeordneten Repräsentanten einer Komponente nur jene Elemente übernommen werden, die jeweils nur in der einen, nicht aber in der anderen Figur vorkommen (wobei es in beiden Figuren auch gemeinsame Elemente geben kann) ► *Einzelelementaddition*
- ...entsteht ein Repräsentant einer Komponente, indem die Veränderungsregel, welche aus der ersten Figur die zweite Figur hervorbrachte, auf die veränderte zweite Figur angewendet wird, um die dritte Figur zu erzeugen ► *Seriation*
- ...variiert die Reihenfolge von drei eindimensionalen Repräsentationen („offene Gestalten“) beliebig ► *Variation offener Gestalten*
- ...variiert die Reihenfolge von drei zweidimensionalen Repräsentationen („geschlossene Gestalten“) beliebig ► *Variation geschlossener Gestalten*

2) Realisierungsmodus

- Wenn sich die Regeln, welche in den Items enthalten sind, auf verschiedene, deutlich unterscheidbare Komponenten der dargestellten Figur beziehen, spricht man von *Separierten Komponenten (SC)*. Die Unterscheidbarkeit kann zum Beispiel durch folgende Merkmale gefördert werden: groß vs. klein, innerhalb vs. außerhalb, etc.
- Wenn sich die Regeln auf dasselbe Figurelement beziehen, aber auf unterschiedliche Merkmale desselben (zum Beispiel Farbe und Muster eines Kreises), spricht man von *integrierten Komponenten (IC)*.

- Besteht eine „geschlossene Gestalt“ aus Repräsentanten, für die unterschiedliche Regeln gelten, und sind diese Repräsentanten nur durch die für sie geltenden Regeln erkennbar, dann spricht man von *integrierten Repräsentanten (IR)*. Hier soll die Testperson die geschlossene Gestalt auflösen.

3) Lösungsrichtung

Hier wird unterschieden, ob sich die Lösungsoperationen

- ...auf die Spalten der Matrix beziehen ► *senkrecht*
- ...auf die Zeilen der Matrix beziehen ► *waagerecht* oder
- ...auf die Zeilen und Spalten beziehen ► *senkrecht/waagerecht*

Hornke, Küppers & Etzel (2000) haben für den AMT nach den oben genannten Regeln 270 neue Items entwickelt, welche Parallelversionen zu zuvor konstruierten Items waren. Das heißt, dass nur die grafischen Elemente verändert, die Regeln jedoch übernommen wurden. Das bedeutet, dass „[...]der Genotyp der jeweiligen Aufgabe - beibehalten bleibt und lediglich die phänotypische Gestalt verändert wird“ (S. 184).

4.4 Färbiger Matrizentest – FMT

Auch der FMT sollte in Anlehnung an die Arbeiten von Formann (1973) und Hornke und Habon (1984a) regelgeleitet konstruiert werden. Es wird angenommen, dass sich aufgrund der vorgeschlagenen Regelsysteme der Autoren neue Matrizenitems entwickeln lassen, die dem Rasch-Modell entsprechen und deren angenommene kognitive Operationen sich mit dem LLTM abbilden lassen. Wie der AMT (vgl. 4.3.2) sollte der FMT ein Computermatrizentest werden, um für die Vorgabe im Wiener Self-Assessment für Psychologie geeignet zu sein. Im Gegensatz zu den bereits genannten Matrizentests sollte der FMT so konzipiert werden, dass die gesuchte Figur diejenige ist, die sich in der Mitte der Matrix befindet und nicht wie bisher gehabt rechts unten. Das bedeutet, dass das Fragezeichen sich nun in der Mitte der Matrix befindet. Folglich sind nur jene Regeln denkbar, die dies zulassen. Zusätzlich sollte der FMT, wie der Name schon besagt, Farben beinhalten.

Zu Beginn der Testentwicklung steht die Definition des zu messenden *Konstrukts* (vgl. Kapitel 2)

im Vordergrund. Der FMT soll die Fähigkeit zum logisch schlussfolgernden Denken erheben, wobei eine Definition sowie eine Beschreibung dessen, wie Matrizen test diese Fähigkeiten erheben, bereits in Kapitel 4.1 gegeben wurde. Der Zugang zur Messung des Konstrukts durch regelgeleitete Itemkonstruktion, welche auf kognitiven Lösungsprozessen beruhen (vgl. Formann, 1973; Hornke & Habon, 1984a; Hornke, Küppers & Etzel, 2000) scheint sich bis jetzt bewährt zu haben und wird daher auch für den FMT gewählt.

Als *Antwortformat* für den FMT sollte das Multiple-Choice Format „eins aus acht“ verwendet werden. Dabei werden der Testperson acht Matrizenfiguren (die Lösung und sieben Distraktoren) als Antwortmöglichkeiten zur Verfügung gestellt. Aus diesen soll diejenige Figur ausgewählt werden, die logisch sinnvoll in der Mitte der Matrix platziert werden kann.

4.4.1 Das Itemuniversum

Wie für den AMT (Hornke, Küppers & Etzel, 2000) werden für den FMT drei schwierigkeitsgenerierende Komponenten angenommen.

1) Das **Regelsystem** (welches in Kapitel 4.4.2 ausführlich dargestellt wird)

2) Der **Realisierungsmodus**

Hier wird zwischen den Realisierungsmodi Separierte Komponenten (SC) und Integrierte Komponenten (IC) unterschieden. Beziehen sich Regeln einer Matrizenaufgabe auf einander deutlich unterscheidbare Komponenten einer Figur (vgl. Abb. 5 Variation o und Variation g, S. 39), dann wird dieser Realisierungsmodus als Separierte Komponente (SC) deklariert. Kommt es bei Matrizenaufgaben jedoch vor, dass Regeln sich auf unterschiedliche Merkmale derselben Komponente beziehen (zum Beispiel Muster und Form bzw. Anzahl und Form, vgl. Abb. 7, S. 41, Identifikation Addition und Identität), dann spricht man von integrierten Komponenten (IC).

3) Die **Richtung der Regelgeltung**

Die Richtung der Regelgeltung bezieht sich, wie der Ausdruck besagt, auf den Bereich für den eine Regel gilt. Es wird zwischen waagrecht, senkrecht und waagrecht/senkrecht unterschieden. Je nachdem ob sich eine Regel auf eine Zeile, Spalte oder Zeile und Spalte bezieht.

4.4.2 Das Regelsystem – Schwierigkeitsgenerierende Operationen

Das Regelsystem, welches dem FMT zugrunde liegt, besteht aus Regeln, die gleichfalls kognitiven Operationen entsprechen. Es wird angenommen, dass diese Operationen schwierigkeitsgenerierend sind, im weiteren bedeutet dies, dass je mehr Regeln ein Item beinhaltet, dieses umso schwieriger wird.

Zusätzlich zu den aus der Literatur bekannten Regeln, welche den neuen Bedingungen angepasst wurden, sollte der Versuch einer neuen Regel „Identifikation einer Addition“ erprobt werden.

Regel 1 – Identität

In jeder Zelle einer Zeile und/oder einer Spalte befindet sich derselbe Repräsentant einer Materialeigenschaft. (vgl. Abb. 7, S. 41)

Regel 2 – Addition

Je Zeile und/oder Spalte der Matrix entsteht der Repräsentant einer Materialeigenschaft in der letzten Zelle (der dritten Zelle) durch das Übereinanderlegen des ersten und zweiten Repräsentanten einer Materialeigenschaft in derselben Zeile bzw. Spalte. (vgl. Abb. 5, S. 39)

Regel 3 – Subtraktion

Je Zeile und/oder Spalte der Matrix entsteht der Repräsentant einer Materialeigenschaft in der letzten Zelle (der dritten Zelle), indem man den zweiten vom ersten Repräsentanten einer Materialeigenschaft in derselben Zeile bzw. Spalte abzieht. (vgl. Abb. 5, S. 39)

Regel 4 – Seriation

Je Zeile und/oder Spalte der Matrix entsteht ein Repräsentant einer Materialeigenschaft durch die Veränderungsregel, welche auf den Repräsentanten in der ersten Zelle angewandt wurde, um den Repräsentanten in der zweiten Zelle zu erzeugen. Im weiteren muss diese Veränderungsregel auf den Repräsentanten in der zweiten Zelle angewandt werden, um den Repräsentanten in der dritten Zelle zu erzeugen (zum Beispiel Verschiebung der Position). (vgl. Abb. 6, S. 40)

Regel 5 – Variation offene Gestalten (Variation o)

Je Zeile und/oder Spalte variiert die Position dreier eindimensionaler Repräsentanten einer Materialeigenschaft beliebig. (vgl. Abb. 5, S. 39 und Abb. 7, S. 41)

Regel 6 – Variation geschlossene Gestalten (Variation g)

Je Zeile und/oder Spalte variiert die Position dreier zweidimensionaler Repräsentanten einer Materialeigenschaft beliebig. (vgl. Abb. 5, S. 39)

Regel 7 – Identifikation Addition

Repräsentanten einer Materialeigenschaft sollen durch die Lage im Raum und ihre Anzahl als positive bzw. negative Zahlen erkannt werden. Dann sollen je Zeile und/oder Spalte die erste und die zweite Zelle miteinander addiert werden, um die Lage und Anzahl der Repräsentanten dieser Materialeigenschaft in der dritten Zelle bestimmen zu können. (vgl. Abb. 7, S. 41)

Zur Veranschaulichung der Regeln werden im Folgenden drei Beispiele dargestellt, in denen alle Regeln enthalten sind:



Abb. 5: Aufgabe 5 des FMT

Abb. 5 stellt die Aufgabe 5 des FMT dar. Bei diesem Item kommen vier Regeln zur Anwendung. Die erste Regel ist die *Variation o*, welche sich auf den eindimensionalen Pfeil bezieht. In jeder Zeile variieren drei Repräsentanten dieses Pfeils: einer der schräg nach oben zeigt, einer der nach rechts zeigt und einer der schräg nach unten zeigt. Die zweite Regel ist die *Variation g*, welche sich auf die Materialeigenschaft Muster der Ellipse bezieht. Je Zeile kommen die drei Repräsentanten rot, grün und blau einmal vor, die Reihenfolge kann dabei beliebig variieren. Die dritte Regel ist die *Addition*, welche sich auf die Zeilen, genauer auf die Materialeigenschaft Muster des oberen Dreiecks der Figur bezieht. Addiert man je Zeile das obere Dreieck aus der ersten mit dem aus der zweiten Zelle, erhält man den Repräsentanten für die dritte Zelle. Die vierte Regel ist die *Subtraktion*, welche sich auf die Spalten der Matrix bezieht. Betroffen ist die Materialeigenschaft Muster der unteren Dreiecke. In jeder Spalte entsteht in der dritten Zelle der Repräsentant einer Materialeigenschaft, indem man das Muster des Dreiecks in der zweiten Zelle vom Muster des Dreiecks in der ersten Zelle abzieht.

Folglich ist die Lösung eine Figur mit einem Pfeil der schräg hinauf zeigt, mit einer blauen Ellipse, einem oberen Dreieck, das senkrecht gestreift ist und einem unteren Dreieck, welches kein Muster beinhaltet.

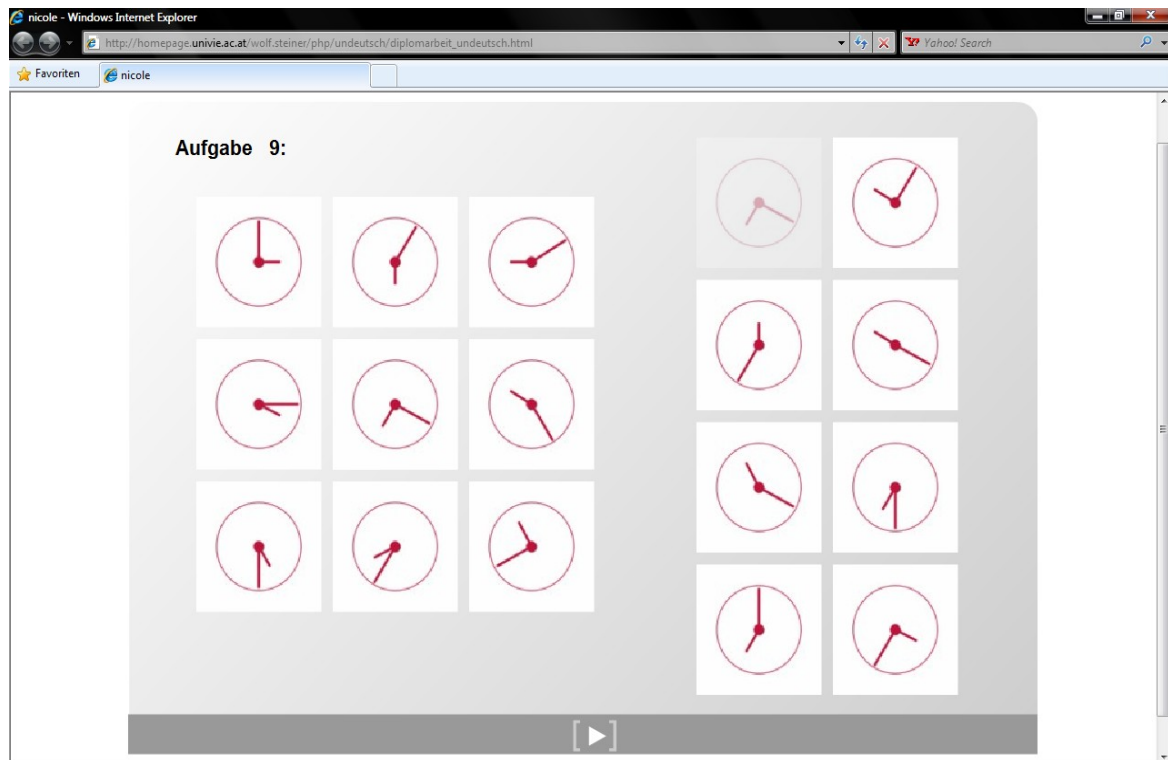


Abb. 6: Aufgabe 9 des FMT

Abb. 6 zeigt die Aufgabe 9 des Färbigen Matrizen Tests. Dieses Item beinhaltet zwei Regeln. Die erste Regel ist eine *Seriation* waagrecht/senkrecht für eine integrierte Komponente, welche sich auf den Minutenzeiger der Uhr bezieht. Dieser verändert seine Position sowohl waagrecht als auch senkrecht um eine Einheit. Zeilenweise gesehen verändert sich der Zeiger jeweils um fünf Minuten und spaltenweise verändert er sich jeweils um 15 Minuten. Die zweite Regel ist ebenfalls eine *Seriation* waagrecht/senkrecht IC, welche sich auf den Stundenzeiger bezieht. Der Stundenzeiger verändert seine Position waagrecht gesehen um 15 Minuten und senkrecht um fünf Minuten von einer Figur zur anderen. Somit ist die Lösung eine Uhr, welche die Uhrzeit 7:20 anzeigt (diese befindet sich, zur Veranschaulichung bereits in der Mitte der Matrix).

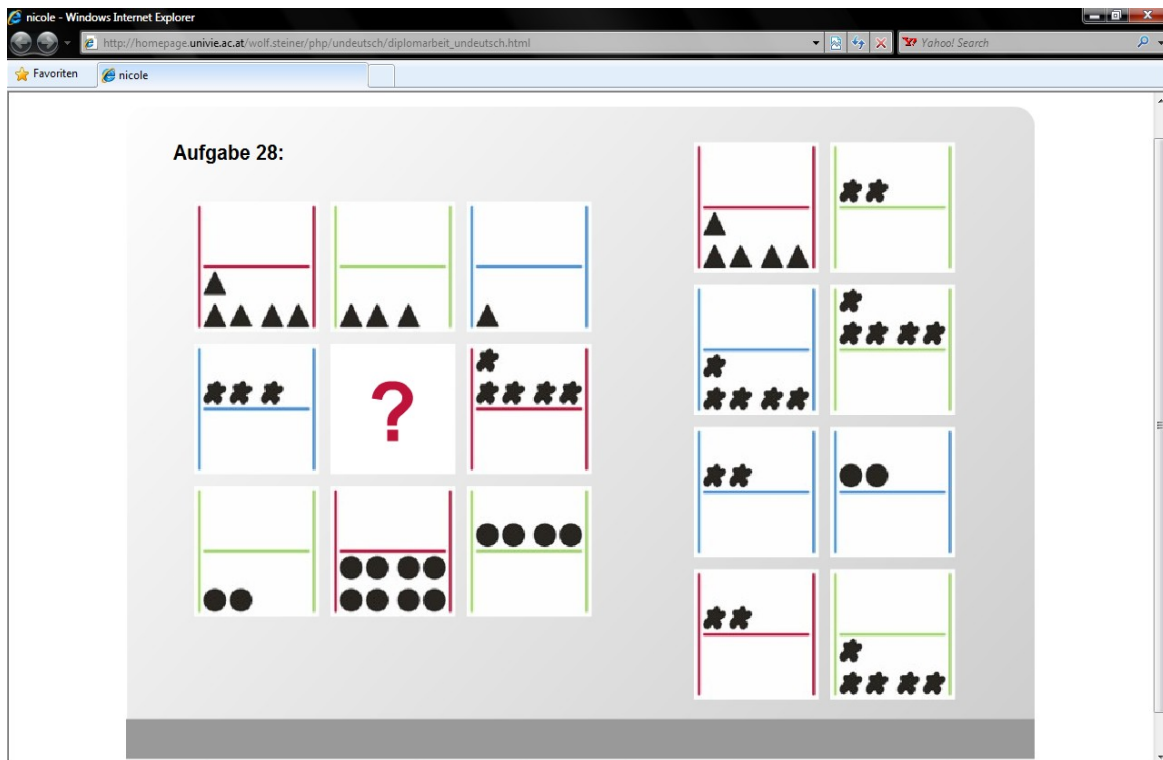


Abb. 7: Aufgabe 28 des FMT

Die Aufgabe 28 des FMT beinhaltet drei Regeln (siehe Abb. 7). Die erste Regel, *Identifikation Addition* bezieht sich senkrecht (IC) auf die Anzahl der Figuren, die sich entweder oberhalb oder unterhalb der Mittellinie befinden. Die Testperson soll hier erkennen, dass Figuren oberhalb der Mittellinie positiven Zahlen entsprechen und Figuren unterhalb der Mittellinie negativen Zahlen entsprechen. Gleichzeitig soll sie die erste Zelle mit der Zweiten Zelle addieren. Die zweite Regel, die *Variation o* betrifft die Farbe des großen eindimensionalen „H“ und variiert senkrecht. Rot, Grün und Blau kommen folglich je Spalte einmal vor, egal in welcher Reihenfolge. Die dritte Regel, die *Identität*, betrifft die Form der Figuren, die ober- und unterhalb der Mittellinie des „H“ stehen. Die Form der Figuren hat waagrecht gesehen (IC) immer die gleiche Merkmalsausprägung (Gestalt). Somit wäre die Lösung der Matrix ein blaues „H“, da dieses in der Spalte noch nicht vorkommt. Die Figuren der Lösung haben die Form von „Flecken“, da in der zweiten Zeile nur Flecken vorhanden sind. Schließlich ist die Anzahl der Figuren gleich fünf, und sie befinden sich unterhalb der Mittellinie (da $(-3) + (-5) = (-8)$).

4.4.3 Materialeigenschaften

Damit der Prozess der Itemkonstruktion weiter objektiviert wird, wurden von vornherein Repräsentanten für die Materialeigenschaften *Muster*, *Farbe* und *Form* festgelegt (vgl. Hornke &

Habon, 1984b; Hornke & Rettig, 1989). Im Folgenden werden die festgelegten Materialeigenschaften für die drei Repräsentanten dargestellt:

- *Farben*: hellgrün, mittelblau, dunkelrot, schwarz
- *Form*: Dreiecke, Fünfecke, Ellipsen bzw. Kreise, Striche, Kleckse, Pfeile, Vierecke
- *Muster*: gestreift, kariert, gepunktet, leer, gefüllt (anzumerken ist, dass diese Formen in ihrer Größe und ihrem Winkel variieren können).

4.4.4 Verteilung der Regeln

Auf der Grundlage der beschriebenen Konstruktionsregeln wurden 30 Items für den Färbigen Matrizentest entwickelt. Da Items, welche nur eine Regel beinhalteten, zu einfach erschienen und dadurch wohl nur redundante Information erhoben werden würde (in dem Sinne, dass alle Testpersonen die Items lösen könnten), wurde auf die Konstruktion von Items mit nur einer Regel verzichtet. Es wurden daher Items mit zwei, drei und vier Regeln entwickelt (vgl. Tabelle 2).

Tabelle 2: Verteilung der Regelanzahl je Item

Anzahl der Regeln	Anzahl der Items	Anteil an der Gesamtanzahl
2 Regeln	11	36,60%
3 Regeln	14	46,60%
4 Regeln	5	16,60%

Da außerdem angenommen wurde, dass Items mit vier Regeln für Studieninteressenten bereits sehr schwierig erscheinen, wurden davon nur wenige Items konstruiert. Die meiste Information könne, so wurde angenommen, aus Items mit zwei und drei Regeln gewonnen werden, daher wurden von diesen mehr konstruiert.

Tabelle 3 (S. 44) präsentiert eine Übersicht über alle 30 Items und ihre Konstruktionsregeln. Pro Item wird aufgelistet, aus wie vielen und welchen Regeln jedes Item besteht. Für die Konstruktion der Items des FMT wurden 84 Regeln herangezogen. Diese 84 Regeln ergeben sich wie folgt: 12 mal *Identität* (14,3%), 12 mal *Addition* (14,3%), 10 mal *Subtraktion* (11,9%), sechs mal *Seriation* (7,1%), 13 mal *Variation o* (15,5%), 28 mal *Variation g* (33,3%) und drei mal *Identifikation Addition* (3,6%). Die Regeln scheinen annähernd gleichverteilt, auffällig ist jedoch der Überhang der *Variation g*, der sich daraus ergibt, dass Form und Farbe leichter durch diese

Regel umzusetzen waren. Die Regel *Identifikation einer Addition* hingegen kam als Versuch seltener zum Einsatz, da nur drei Items eliminiert werden mussten, falls sich herausstellen sollte, dass diese Items nicht dem Modell von Rasch entsprechen. Die Regelrichtung *waagrecht* kommt 43 mal (51,2%) zum Einsatz und die Regelrichtung *senkrecht* 37 mal (44%). Die Regelrichtung *waagrecht und senkrecht* kommt allerdings nur vier mal (4,8%) zum Einsatz, da unter Bezugnahme auf Hornke und Rettig (1989) angenommen wurde, dass diese Regelrichtung leichter sei als die beiden anderen. Zudem kam der Realisierungsmodus *separierte Komponenten* (SC) 45 mal (53,6%) zum Einsatz und der Realisierungsmodus *integrierte Komponenten* (IC) 39 mal (46,4%).

Tabelle 3: Verteilung der Regeln je Item

Itema	Identität	Addition	Subtraktion	Seriation	Variation o	Variation g	Id. Addition	w	s	w/s	SC	IC
Item 1	1	1	0	0	0	1	0	1	2	0	3	0
Item 2	0	0	1	0	1	0	0	1	1	0	0	2
Item 3	0	1	0	0	0	2	0	1	2	0	3	0
Item 4	1	0	1	0	0	2	0	2	1	1	1	3
Item 5	0	1	1	0	1	1	0	3	1	0	2	2
Item 6	3	0	0	0	0	0	0	2	1	0	1	2
Item 7	0	0	0	1	0	0	1	0	1	1	2	0
Item 8	0	0	1	0	0	2	0	2	1	0	3	0
Item 9	0	0	0	1	0	1	0	1	1	0	0	2
Item 10	0	1	0	0	0	2	0	2	1	0	1	2
Item 11	0	0	0	2	0	0	0	0	0	2	0	2
Item 12	0	1	0	0	0	1	0	1	1	0	2	0
Item 13	1	0	1	0	0	2	0	2	2	0	2	2
Item 14	0	1	1	0	0	0	0	1	1	0	0	2
Item 15	0	0	0	0	1	1	0	1	1	0	2	0
Item 16	0	1	0	1	0	1	0	2	1	0	1	2
Item 17	0	1	0	0	0	2	0	2	1	0	0	3
Item 18	1	0	0	0	0	1	0	2	0	0	0	2
Item 19	0	0	1	0	0	1	0	1	1	0	2	0
Item 20	0	0	0	0	0	1	1	1	1	0	2	0
Item 21	0	1	1	0	0	0	0	1	1	0	2	0
Item 22	1	0	0	0	0	2	0	3	0	0	1	2
Item 23	1	0	1	0	1	1	0	2	2	0	4	0
Item 24	1	1	0	0	2	0	0	2	2	0	2	2
Item 25	0	0	0	1	2	0	0	1	2	0	2	1
Item 26	0	0	0	0	2	1	0	1	2	0	2	1
Item 27	0	1	0	0	1	1	0	1	2	0	3	0
Item 28	1	0	1	0	0	1	0	2	1	0	1	2
Item 29	0	1	0	0	1	1	0	1	2	0	0	3

4.4.5 Distraktoren

Hornke und Habon (1984a) weisen darauf hin, dass auch die Erstellung der Distraktoren regelgeleitet erfolgen sollte – in dem Sinne, dass Distraktoren durch Teillösungen zu Stande kommen. Geht man zum Beispiel davon aus, dass ein Item durch das Erkennen und Anwenden von zwei Regeln zu lösen ist, dann könnte eine Teillösung so aussehen, dass nur eine dieser Regeln richtig erkannt wurde und die andere nicht. Die Autoren schlagen hierfür einen Schlüssel zur Konstruktion der Distraktoren vor (vgl. Tab. 4).

Tabelle 4: Schlüssel zur Konstruktion der Distraktoren (Hornke & Habon, 1984a, S. 40)

Distraktoren- Nummer	richtige		Falsche gleiche Repräsentanten Paare		Falsche einzelne Repräsentanten	
	1. Komp.	2. Komp.	1. Komp.	2. Komp.	1. Komp.	2. Komp.
1	2	3	2	2	1	0
2	2	2	2	2	1	1
3	1	2	2	2	2	1

Der Schlüssel zur Konstruktion der Distraktoren soll anhand einer Beispielaufgabe (Hornke & Habon, 1984a) kurz erläutert werden.

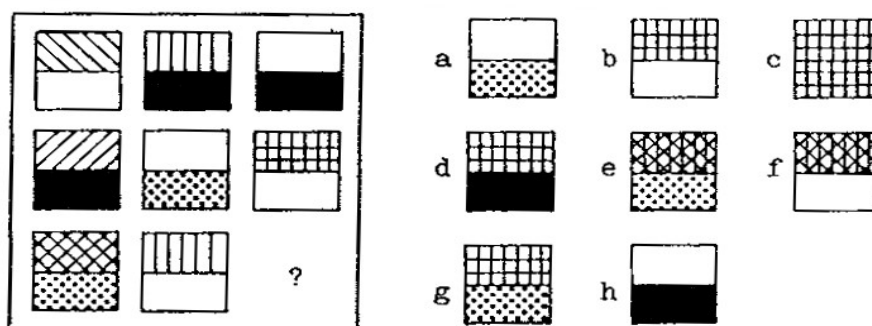


Abb. 8: Beispielaufgabe (Hornke & Habon, 1984a, S. 38)

In Abb. 8 wurde die Distraktorenverteilung Nummer 1 angewandt. Die Lösung wurde nach dem Zufallsprinzip der Position g zugewiesen. Nach der Distraktorenverteilung wurden nun zwei richtige Lösungen der 1. Komponente (= untere Hälfte der Figur) erzeugt und zufällig auf Positionen a und e zugeteilt. Falsche gleiche Repräsentantenpaare der 1. Komponente sollte es zwei geben, daher wurden zwei unterschiedliche Teillösungen konstruiert und zufällig auf Positionen d und h sowie b und f verteilt. Außerdem wurde ein falscher einzelner Repräsentant der 1. Komponente erzeugt und der übrig bleibenden letzten Position c zugeschrieben (Hornke &

Habon, 1984a).

Leider ist in der Arbeit von Hornke und Habon (1984a) die Vorgehensweise hinsichtlich der Konstruktion der 2. Komponente (obere Hälfte der Figur) nicht genau ersichtlich. Die Autoren weisen lediglich darauf hin, ebenso vorgegangen zu sein wie bei der 1. Komponente. Diese Vorgehensweise scheint jedoch nicht immer durchführbar zu sein, da ein zufälliges Zuteilen der 2. Komponente zu den Positionen (a-h) zur Entstehung identer Distraktoren führen könnte. Daher scheint hier eine willkürliche Streuung der Verteilung der 2. Komponente hinsichtlich ihrer Position unabdingbar.

Zur Distraktorenerstellung wurde der in Tabelle 5 dargestellte Verteilungsschlüssel für Komponenten verwendet. Dadurch wurde die Subjektivität der Konstruktion von Distraktoren zum Teil eingeschränkt. Zum Teil deshalb, weil bei der Anwendung der Verteilungsschlüssel darauf geachtet wurde, dass keine identen Distraktoren erzeugt werden. Außerdem wurde nicht wie bei Hornke und Habon (1984a) eine fixe Kombination von Verteilungsschlüsseln festgelegt, die pro Item angewandt werden, sondern Schlüssel zur Konstruktion erstellt, die pro Regelgeltung angewandt werden konnten und daher (bei der Anwendung mehrerer Regeln pro Item) unterschiedlich miteinander kombiniert werden konnten.

Tabelle 5: Schlüssel zur Konstruktion der Distraktoren für einzelne Regeln

Dis. Nr.	richtige (K.)	falsche gleiche Repräsentanten Vierlinge (K.)	falsche gleiche Repräsentanten Drillinge (K.)	falsche gleiche Repräsentanten Paare (K.)	falsche einzelne Repräsentanten (K.)
1	1	0	0	0	6
2	1	0	0	1	4
3	1	1	0	1	0
4	2	0	0	0	5
5	2	0	0	1	3
6	2	0	0	2	1
7	2	0	1	0	2
8	2	0	1	1	0
9	3	0	0	0	4
10	3	0	0	1	2
11	3	0	0	2	0
12	3	0	1	0	1
13	3	1	0	0	0
14	4	0	0	1	1
15	5	0	0	0	2
16	5	0	0	1	0

5 Interessenfragebogen

Auch wenn das Konstrukt Interesse im Anforderungsprofil des Wiener *Self-Assessments für Psychologie* (vgl. Kapitel 1.2) nicht explizit angeführt wird, scheint es offensichtlich zu sein, dass dieses bei der Entscheidung für eine Studienrichtung eine wichtige Rolle spielt. Hierfür lassen sich in der Literatur einige Nachweise finden. So zeigen Schiefele, Streblow und Brinkmann (2007) zum Beispiel, dass sich Studienabbrecher und Weiterstudierende zum Abbruchzeitpunkt neben ihrer Demotivation vor allem in ihrem Studieninteresse unterscheiden. Bergmann (1998) zeigt, dass Studenten bei einer Übereinstimmung der Interessen mit dem gewählten Studium (=kongruenter Studienwahl) zufriedener sind, sich erfolgreicher fühlen und geringere Tendenzen zum Studienwechsel und Studienabbruch zeigen. Müller (2001) belegt zudem einen Zusammenhang zwischen Studieninteresse und Lernbereitschaft.

Die genannten Beispiele dürften verdeutlichen, warum die Erhebung von Interesse im Wiener Self-Assessment von großer Bedeutung ist.

Im diesem Kapitel folgt daher zuerst eine kurze Darstellung des Konstruktes Interesse sowie ein Überblick über etwaige Probleme bei der Konstruktion von Interesseninventaren. Im Weiteren werden bestehende Interessenfragebögen sowie deren theoretische Fundierung vorgestellt. Im letzten Abschnitt wird schließlich der für diese Arbeit neu entwickelte Fragebogen zum Interesse am Psychologiestudium Wien (FIP) präsentiert, insbesondere soll hier ein Überblick über die Konstruktionsprinzipien geschaffen werden.

5.1 Interesse

Durch eine etymologische und morphologische Analyse des Begriffs „Interesse“ zeigte Lägeler (2005), dass diesem in den letzten Jahrhunderten mehrere Bedeutungen zukamen (näheres siehe Lägeler, 2005). Auch heute noch wird unter dem Begriff Interesse nicht nur eine Bedeutung subsumiert. Im Duden-Fremdwörterbuch findet man unter Interesse folgende Begriffsbestimmungen (Dudenredaktion, 2000, S. 449):

1. (ohne Plural) geistige Anteilnahme, Aufmerksamkeit; Ggs. ↑ Desinteresse.
2. a) (meist Plural) Vorliebe, Neigung; b) Neigung zum Kauf.

3. a) (meist Plural) Bestrebung, Absicht; b) das, woran jmdn. od. etwas wichtig od. nützlich ist; Vorteil, Nutzen.
4. (nur Plural; veraltet) Zinsen

Im Falle der Studienwahl scheint die unter Punkt 2. a) angeführte Bedeutung *Vorliebe* bzw. *Neigung* die zutreffendste zu sein. Für eine wissenschaftliche Theorie ist hingegen eine präzise Definition des Konstruktes erforderlich. In der Psychologie variieren diese Definitionen je nach wissenschaftlicher Positionierung des Autors (Todt, 1978).

Eines der wesentlichsten Unterscheidungsmerkmale innerhalb der verschiedenen Interessenkonzepte betrifft die Stabilität von Interessen. Hier wird zwischen Interesse als Zustand (state) und Interesse als Disposition (trait) unterschieden (Bergmann & Eder, 2005).

- *Interesse als Zustand (state)*: Im State-Ansatz, welcher auf den Situationismus zurückgeht, versteht man unter Interesse einen Zustand, welcher von Kontextvariablen wie Zeit und Situation beeinflusst wird (Bergmann, 2003). Folglich entsteht das Interesse an einer Situation, einem Gegenstand oder einem Objekt in der Situation selbst, aufgrund von äußeren Einflüssen.
- *Interesse als Disposition (trait)*: Im Trait-Ansatz, welcher auf den Dispositionismus zurückgeht, werden Interessen als „relativ stabile, situationsunabhängige individuumspezifische Merkmale angesehen, die integrale Bestandteile der Persönlichkeit darstellen“ (Bergmann, 2003, S. 225). „Here, interest is interpreted as a relatively stable tendency to occupy oneself with an object of interest“ (Krapp, 2002, S. 388). Das Interesse wird hier wie eine Persönlichkeitseigenschaft als Eigenschaft einer Person aufgefasst, welche in ihr verankert ist und welche sie dazu bewegt mit Objekten etc. zu interagieren.

Eine relativ allgemeine Definition von Interesse lässt sich bei Fröhlich (2002) finden:

„Interesse (interest). Allgemeine und umfassende Bezeichnung für Einstellungen oder Erwartungen, eine innerlich begründete Tendenz, auf bestimmte Gegenstände oder Gegebenheiten der Umwelt besonders zu achten (Selektivität der Aufmerksamkeit), gesteigerte (emotionale) Anteilnahme an bestimmten für bedeutungsvoll erachteten

Gegenständen oder Ereignissen, Ideen u. ä. zu zeigen, sich für bestimmte Ereignisse oder Sachverhalte besonders zu interessieren und ihren Wirkungen und Ursachen nachzugehen usw.“ (Graumann 1965a, zitiert nach Fröhlich, 2002, S. 246)

Nachdem nun feststeht, was man unter Interesse im Allgemeinen versteht, und dass sich viele Autoren auf unterschiedliche Konstrukte bzw. Definitionen von Interesse beziehen, sollen nun weitere Unbestimmtheiten bei der Konstruktion von Interesseninventaren beleuchtet werden.

5.1.1 Erhebung von Interesse

So unterschiedlich wie die Auffassungen von Interesse sind, so unterschiedlich sind auch die Methoden zur Erfassung von diesem. Super (1957) beschreibt in dem Zusammenhang vier Methoden zur Feststellung von Interessen:

- „Expressed interests are expressions or professions of specific interest; they are preferences“ (Super, 1957, S. 218). Mit «ausgedrückten Interessen» sind jene Interessen gemeint, die eine Person selbst anführt, zum Beispiel in einem Gespräch: „Ich interessiere mich für Tanzstile.“ oder „Ich würde gerne Kunstkritiker werden“.
- „Manifest interests are expressed not in words, but in action, through participation in activities“ (Super, 1957, S. 218). Mit «manifesten Interessen» sind jene gemeint, die sich dadurch feststellen lassen, indem man zum Beispiel Freizeitbeschäftigungen einer Person (Tanzkurse, Mathematikclub etc.) oder die Wahl eines Ausbildungszweiges (HTL, Tourismusschule) betrachtet. Sie zeigen sich folglich durch das Verhalten einer Person.
- „Inventoried interests are estimates of interests based on responses to a large number of questions concerning likes and dislikes, or concerning the order or preference for groups of activities“ (Super, 1957, S. 219). «Aufgenommene Interessen» sind jene Interessen, die mit Hilfe von Fragen bzw. Fragebögen erhoben werden. Personen sollen hier zu einer Vielzahl von Aussagen ihre Interessen äußern. Diese Methode wird als subjektive Methode bezeichnet.
- „Tested interests are manifest interests but interests manifested under controlled rather than in life situations“ (Super, 1957, S. 219). «Getestete Interessen» werden festgestellt, indem über bestimmte Faktoren in einem Test Rückschlüsse auf die Interessen einer Person gezogen werden. Würde man einer Person einen Text über mehrere

Gegenstandsbereiche vorlegen und anschließend ihr Wissen zu den einzelnen Bereichen erheben, dann könnte das Wissen als Faktor dienen, der Rückschlüsse über die Interessen einer Person ermöglicht. Diese Methode der Erhebung von Interessen kann auch als objektive Methode bezeichnet werden.

Die einzelnen Methoden liefern ihre Vor- und Nachteile. Sie sollten daher in Hinblick auf die theoretische Positionierung, die Zielsetzung und die Anwendungssituation, welche ein Autor intendiert, berücksichtigt werden (Bergmann & Eder, 2005).

Todt (1978) liefert in seiner Arbeit einen Überblick über die „Probleme“ bei der Entwicklung seines Interesseninstrumentars (vgl. Abb. 9). Er weist darauf hin, dass bei der Konstruktion in erster Linie entschieden werden sollte, welche **Erhebungsmethode** man verwenden möchte. Jedoch führt er nicht wie Super (1957) vier verschiedene Methoden an; er geht von der Wahl zwischen objektiven Methoden („tested interests“) und subjektiven Methoden („inventoried interests“) aus.

„Expressed interests“ und „manifest interests“ scheinen sich für psychologisch-diagnostische Gespräche besser zu eignen als für psychologisch-diagnostische Verfahren. Bei beiden kann angenommen werden, dass vielleicht ein, zwei oder drei Aussagen zu bevorzugten Tätigkeiten oder Berufen gemacht werden. Interesseninventare („inventoried interests“ & „tested interests“) jedoch versuchen durch die Abdeckung eines möglichst breiten Spektrums an Interessengebieten ein ganzes Interessenprofil einer Person zu erfassen. Somit scheinen diese eher geeignet um Berufswahl- und Schullaufbahnentscheidungen zu unterstützen.

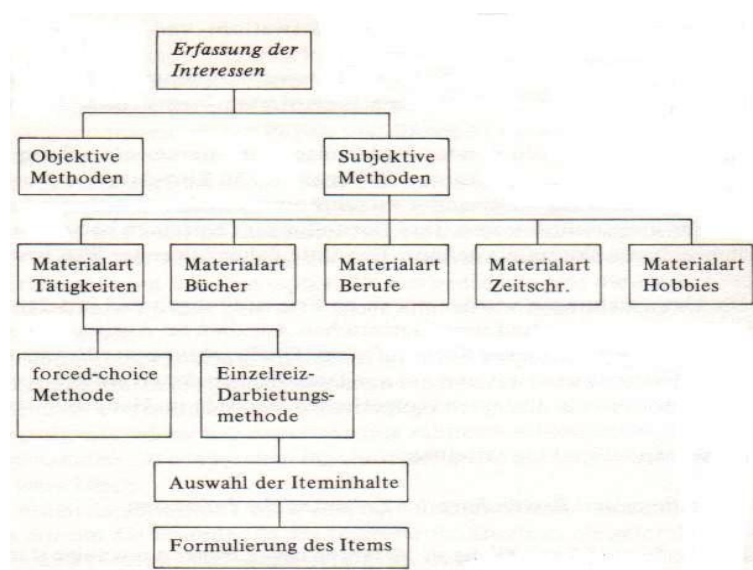


Abb. 9: Probleme von Todt bei der Entwicklung eines Interessenfragebogens (1978, S. 28)

Neben der Frage nach der Methodik stellt sich auch die Frage nach der **Materialart**. Die meisten Fragebögen fordern Testpersonen auf, ihr Interesse in Bezug auf eine Vielzahl von Tätigkeiten zu bekunden (zum Beispiel AIST-R/UST-R, B-I-T. II, DIT, GIS; die genaue Bezeichnung der Verfahren befindet sich am Ende dieses Kapitels). Jedoch soll auch in manchen Verfahren zum Beispiel zu Büchertiteln (DIT) oder zu bestimmten Berufen (DIT) Stellung genommen werden. Auch Zeitschriftentitel (DIT) werden herangezogen, um Interessen zu erheben. Es stellt sich somit dem Testautor die Frage, welche Art von Aufgabenmaterial für seine Zielgruppe geeignet ist und welche sich für den angestrebten Interessenbereich bzw. die Interessenbereiche am ehesten eignet.

Todt (1978) zeigt, dass Jugendliche relativ gleich auf unterschiedliche Materialarten reagieren können. Jedoch ist er der Meinung, dass „die „Materialart“ Tätigkeiten die flexibelste Methodenkonstruktion zulässt“ (S. 207).

Weiters weist Todt (Abb. 9) auf die Wichtigkeit der Wahl des **Antwortformats** hin. Als Beispiele nennt er das forced-choice Format oder die Einzelreizdarbietungsmethode. Auch Bergmann und Eder (2005) sprechen in diesem Zusammenhang von der Entscheidung zwischen forced choice und free choice Antwortformat.

- Das *forced-choice Antwortformat* verlangt von der Testperson, dass sie sich für eine (oder zwei) von zwei oder mehr vorgegebenen Alternativen entscheidet. Es soll dabei die bevorzugte der angegebenen Alternativen gewählt werden bzw. teilweise wird auch gewünscht, dass die am meisten abgelehnte zusätzlich gekennzeichnet wird. Ein Verfahren, welches im deutschsprachigen Raum dieses Antwortformat anwendet, ist der Berufs-Interessen-test II (B-I-T. II) von Irle und Allehoff (1988). Zwei der vier Formen des B-I-T. II verlangen bei einer Gegenüberstellung von vier Tätigkeiten die Markierung jener Tätigkeit, die der Testperson am besten gefällt und die Markierung der Tätigkeit, die ihr am wenigsten gefällt.

Die Kritik am forced-choice Format besteht vor allem darin, dass die Stärke bzw. Ausprägung eines Interesses für eine bestimmte Tätigkeit oder einen bestimmten Bereich bei allen Personen gleich hoch ist und hier im weiteren nicht differenziert werden kann (vgl. Bergmann & Eder, 2005; Todt, 1978).

- Das *free-choice Antwortformat* hingegen bietet der Testperson die Möglichkeit bezüglich einer einzelnen Äußerung Stellung zu nehmen. Beim free-choice Format differenziert man anhand einer Skala, mittels welcher Testpersonen ihre Interessen bezüglich einer Äußerung bekunden sollen. Neben dichotomem Antwortformat (MOI) kommt hier des

öfteren eine mehrkategoriale Ratingskala (AIST-R/UST-R, GIS, B-I-T. II Form BA und BB, DIT) zum Einsatz, zum Beispiel interessiert mich sehr, interessiert mich etwas, interessiert mich kaum, interessiert mich überhaupt nicht.

Der Vorteil von Ratingskalen liegt darin, dass differenzierte Antworten möglich sind und somit das Interesse genauer erfasst werden kann.

Als Kritik äußert Todt (1972), dass bei diesem Antwortformat ein individuelles Bezugssystem zum Tragen kommt. Er fügt jedoch hinzu, dass dieses Problem allen psychologischen Fragebogentechniken zugrunde liegt.

Schließlich geht es noch darum, festzulegen welche **Iteminhalte** erfasst werden sollen. Diese Inhalte sollten bestmöglich eingegrenzt werden, um bei der **Itemformulierung** einen präzisen Rahmen zur Verfügung zu haben, aus dem die Inhalte stammen können (Itemuniversum). Todt (1978) meint, „dass die Auswahl der Inhaltsbereiche (Interessenrichtungen) durchaus ein Problem darstellt und bisher meistens durch informelle Übereinkunft gelöst wurde“.

Nachdem einzelne Problematiken der Interessenerfassung gesondert betrachtet wurden, werden nun bestehende Interessenfragebögen vorgestellt. Im deutschsprachigen Raum gibt es eine Vielzahl an Interessenfragebögen bzw. „-tests“, unter anderen zu nennen sind in der Anordnung ihres Erscheinungsjahres: der Differentielle Interessen-Test (DIT) (Todt, 1972), der Berufs-Interessen-test II (B-I-T. II) (Irle & Allehoff, 1988), die Generelle Interessen-Skala (GIS) (Brickenkamp, 1990), der Fragebogen zum Studieninteresse (FSI) (Schiefele, Krapp, Wild & Winteler, 1992), Explorix® - das Werkzeug zur Berufswahl und Laufbahnplanung (Jörin, Stoll, Bergmann & Eder, 2004), der Allgemeine Interessen-Struktur-Test mit Umwelt-Struktur-Test Revision (AIST-R/UST-R) (Bergmann & Eder, 2005) sowie die unlängst erschienene Multimethodische Objektive Interessentestbatterie (MOI) (Proyer & Häusler, 2008). In den folgenden Abschnitten soll jedoch keine Abhandlung dieser Verfahren dargestellt werden, vielmehr liegt das Hauptaugenmerk auf jenen Verfahren, die auf Basis einer Theorie bzw. eines Modells entwickelt wurden. Da nicht besonders viele Theorien in Bezug auf Interessenfragebögen bestehen, werden im Folgenden die Theorien vorgestellt und in Bezug darauf die Interessenfragebögen, welche sich auf diese Theorie beziehen.

5.2 Die Theorie von John L. Holland

Die Theorie von Holland ist eine der bekanntesten Theorien, welche Interessenfragebögen zugrunde liegt. Dies zeigt sich alleine durch die Existenz unzähliger Studien, sowie durch Artikel über die Beeinflussung der Wissenschaft auf diesem Gebiet (vgl. Campbell & Borgen, 1999; Rayman & Atanasoff, 1999) und durch die Tatsache, dass einige Interessenfragebögen auf dieser Theorie beruhen (siehe dazu Kapitel 5.2.1, 5.2.2, 5.2.3). Holland sieht in Interessen eine Äußerungsform der Persönlichkeit (Bergmann & Eder, 2005). Seine Theorie ist strukturell-interaktiv, da sie von einer Persönlichkeitsstruktur ausgeht und eine wechselseitige Beziehung zwischen Person und Umwelt („Passung“) annimmt (Jörin, Stoll, Bergmann & Eder, 2004).

Die Theorie von Holland (1997) basiert auf **vier Grundannahmen**:

- 1. In our culture, most persons can be categorized as one of six personality types: Realistic, Investigative, Artistic, Social, Enterprising, or Conventional....*
- 2. There are six model environments: Realistic, Investigative, Artistic, Social, Enterprising, and Conventional....*
- 3. People search for environments that will let them exercise their skills and abilities, express their attitudes and values, and take on agreeable problems and roles....*
- 4. Behavior is determined by an interaction between personality and environment. (Holland, 1997, S. 2 ff)*

Die erste Annahme geht davon aus, dass sich Menschen zu sechs „idealen“ Persönlichkeitstypen (realistic, investigative, artistic, social, enterprising, conventional, vgl. Tab. 6., S. 55) zuordnen lassen. Da es wahrscheinlich ist, dass ein Mensch Ähnlichkeiten mit mehreren Typen aufweist, wird er mit allen verglichen und demjenigen zugeordnet, bei dem die größte Ähnlichkeit besteht. Dieser Typ gilt nach Holland als dominanter Persönlichkeitstyp. Holland (1997) berücksichtigt jedoch alle sechs Persönlichkeitstypen bei der Typisierung einer Person, indem diese nach ihrer Ähnlichkeit gerangiert werden, „that allows for the complexity of personality and avoids some of the problems inherent in categorizing a person as a single type“ (S. 3). Dies ermöglicht eine Unterscheidung zwischen 720 verschiedenen Persönlichkeitsmustern und entspricht eher der Komplexität der menschlichen Natur.

Die zweite Annahme postuliert, dass sich nicht nur die Persönlichkeit, sondern auch

verschiedene Umwelten (Berufe, Schulzweige etc.) zu den sechs Dimensionen zuordnen lassen. Dabei wird angenommen, dass sich in Umwelten der Dimension «R» überwiegend Persönlichkeitstypen der Dimension «R» befinden. Holland (1997) meint, wenn sich Personen zusammenfinden, was meist aufgrund von gemeinsamen Interessen, Einstellungen, Werten, Zielen etc. geschieht, dass diese eine Umwelt gestalten, welche denjenigen Typ reflektiert, dem diese Personen am ehesten ähnlich sind.

Die dritte Annahme besagt, dass jeder Persönlichkeitstyp jene (beruflichen) Umwelten aufsucht, die es ihm ermöglichen, seine Fähigkeiten, Fertigkeiten, Werte und Einstellungen auszuleben bzw. einzusetzen oder zu verfolgen. Das heißt, Personen suchen jene Umweltdimension auf, die ihrem Personentyp entspricht oder Umwelten suchen nach passenden Persönlichkeitstypen zum Beispiel im Sinne von Personalauswahlprozessen.

Die vierte Annahme geht davon aus, dass die Interaktion zwischen dem Umwelttyp, in der sich eine Person befindet, und dem Personentyp das Verhalten einer Person (Leistung, Berufswechsel etc.) bestimmt. Holland (1997) meint, dass diese Kombination eine Voraussage über Jobchancen, Berufswahl etc. ermöglicht.

Tabelle 6: Auszug aus der Charakterisierung der Dimensionen R, I, A, S, E und C (Jörin et al., 2004, S. 12-13)

Rubrik	R (Realistic)	I (Investigative)	A (Artistic)	S (Social)	E (Enterprising)	C (Conventional)
Übersetzung	handwerklich-technisch	untersuchend-forschend	künstlerisch, sprachlich	sozial, erziehend-pflegend	unternehmerisch	ordnend-verwaltend
Interessen	Arbeiten mit Händen und Körperkraft, mit Maschinen	Forschen, analysieren, intellektuelle Probleme lösen	Mit der Sprache arbeiten, künstlerisch(e) (Darstellung)	Sozial, helfen, lehren, erziehen, um Probleme kümmern	Managen, organisieren, verkaufen, führen (wirtschaftlich)	Büro, Regeln, Genauigkeit, Ordnung, Routine
Fähigkeiten	Handgeschick	Mathematisch-wissenschaftl. Neugier, Forschung	Kreativität, Ausdrucks-fähigkeit, Ästhetik, Musisch	Einfühlungs-vermögen, Geduld, pädagog., Hilfsbereitschaft	Organisationstalent, Überzeugung, wirtschaftl. Denken, Führungskompetenz	Ordnung, Zahlen, Regelakzeptanz, Monotonie-resistenz
Persönlichkeit	naturverbunden, körperlich robust, handwerklich geschickt, technisch begabt, geradlinig	scharfsinnig, analytisch, geistreich, wissenschaftlich interessiert, intellektuell	fantasievoll, schöpferisch, intuitiv, künstlerisch, gefühlsbetont, originell, eigenwillig, ausdrucksstark	hilfsbereit, warmherzig, gesellig, tolerant, gesprächig, selbstlos, fürsorglich, mitteilsam, treu	führungsfähig, draufgängerisch, aktiv tonangebend, dynamisch, kontaktfreudig, zielstrebig, verantwortungsbereit,	sorgfältig, genau, ordentlich, angepasst, pflichtbewusst, gewissenhaft, diszipliniert, vorsichtig
Werte/Ziele	Greifbares, Gesunder Menschenverstand	Wissen	Künstlerischer Ausdruck, Unkonventionalität, Kultur	Helfen, Beziehungen	Erfolg, Verantwortung	Anpassung, Normen
Berufs-Beispiele	Mechaniker/in, Landwirt/in, Elektriker/in	Biologe/in, Forscher/in, Physiker/in	Schriftsteller/in, Schauspieler/in, Designer/in	Lehrer/in, Sozialarbeiter/in, Krankenpfleger/in	Manager/in, Finanzberater/in, Politiker/in	Sekretär/in, Korrespondent/in, Bankangestellte/r
Berufliches Umfeld	Handwerk, Technik, Landwirtschaft	Forschung, Labor, Wissenschaft	Kunst/Kultur, Theater, Film, Mode, Musik, Journalismus, Design	Bildungswesen, Beratung, Gesundheitswesen	Management, Verkauf, Werbung, PR, Unternehmen/Wirtschaft	Dienstleistungsbetrieb, Verwaltung, Büro, Buchhaltung
Zusammenfassung	Praktisch-technisch, „von Hand“	Denkarbeit, erforschen, analysieren	Kreativ, künstlerisch, musisch	Helfen, pflegen, umsorgen, erziehen, bilden	Managen, initiieren, verkaufen, Finanzen	Genau, pflichtbewusst, systematisch, repetitiv

Im Weiteren werden die zuvor beschriebenen vier Grundannahmen durch zusätzliche sekundäre Konzepte ergänzt. Diese sollen die Grundannahmen untermauern und Vorhersagen oder Erklärungen über weiteres Verhalten ermöglichen. Im Folgenden werden die fünf **sekundären Annahmen** von Holland (1997) näher beschrieben:

1. Als Calculus bezeichnet Holland (1997) seine Annahme, dass die Beziehungen zwischen den sechs Dimensionen in einem hexagonalen Modell dargestellt werden können (vgl. Abb. 10).

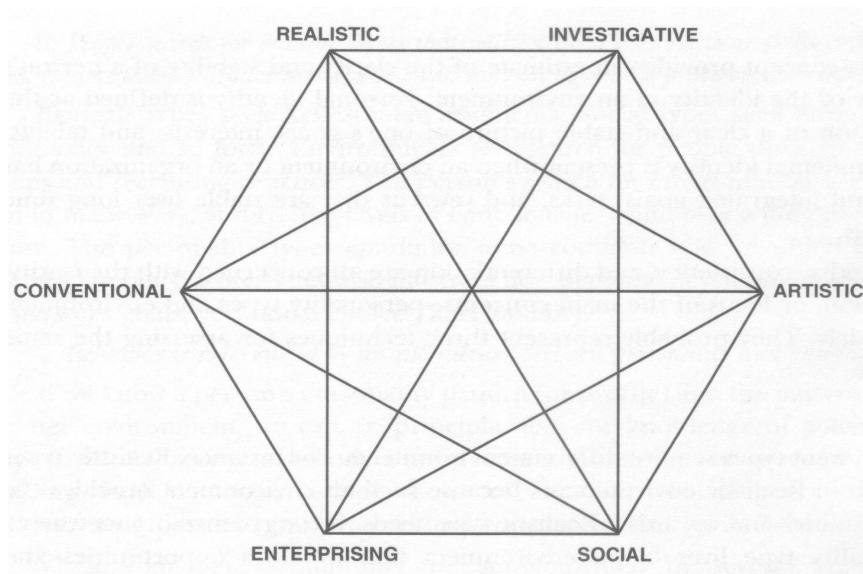


Abb. 10: Hexagonales Modell von Holland (1997, S. 6, S. 35)

2. Konsistenz bezeichnet den Grad der Verbundenheit (Übereinstimmung) zwischen zwei Persönlichkeits- oder Umwelttypen. Holland meint, dass manche Persönlichkeitstypen eher miteinander verbunden sind als andere und dass sich dementsprechend die beruflichen Präferenzen einer Person, deren zwei stärksten Persönlichkeitsausprägungen sehr miteinander verbunden sind, genauer vorhersagen lassen (Holland, 1997). Diese Beurteilung basiert auf der Annahme des hexagonalen Modells der sechs Persönlichkeits- bzw. Umweltdimensionen (vgl. Abb. 10). Je näher die zwei dominantesten Dimensionen einer Person in dem Modell sind, umso konsistenter sind sie. Unterschieden wird zwischen Konsistenz (zum Beispiel S und E oder I und A), mittlerer Konsistenz (zum Beispiel S und C oder E und A) und Inkonsistenz (zum Beispiel S und R oder C und A) (Bergmann & Eder, 2005).
3. Die Differenziertheit bezieht sich auf das Profil einer Person, das sich durch die Ausprägungen in den sechs Dimensionen darstellen lässt. Zeichnet sich im Profil einer

Person eine deutliche Spitze ab, dann spricht man von großer Differenziertheit, somit von einem großen Interesse in einem Bereich. Sind jedoch die Ausprägung auf allen sechs Dimensionen in etwa gleich bzw. mit nur geringen Schwankungen, dann würde man diese Person als eher wenig differenziert bezeichnen. Neben der Breite des Interesses lässt die Differenziertheit auch Aussagen über die Stabilität des Interesses zu (Bergmann & Eder, 2005).

4. Die Identität ermöglicht eine Schätzung der Klarheit und Stabilität einer Persönlichkeit bzw. Umwelt. Eine Persönlichkeit bzw. eine Umwelt gilt dann als stabil, wenn Ziele, Interessen, Talente bzw. Aufgaben klar und über eine längere Zeit stabil sind (Holland, 1997).
5. Kongruenz bezeichnet die Übereinstimmung („Passung“) des Persönlichkeitstyps mit dem Umwelttyp. Dieser Wert basiert ebenfalls auf der Annahme des hexagonalen Modells der sechs Dimensionen (vgl. Abb. 10, S. 56). Je weiter zwei Dimensionen im hexagonalen Modell voneinander entfernt sind, desto weniger haben sie gemeinsam. Demzufolge werden vier Ausprägungen der Kongruenz unterschieden: maximale, mittlere, niedrige Kongruenz und Inkongruenz. Mittels Kongruenz können Aussagen über zum Beispiel die Zufriedenheit mit einem Beruf getätigt werden (Bergmann & Eder, 2005).

Anschließend sollen nun deutschsprachige diagnostische Verfahren vorgestellt werden, die auf Basis des Modells von Holland entwickelt wurden.

5.2.1 Allgemeiner Interessen-Struktur Test/Umwelt-Struktur-Test - Revision (AIST-R/UST-R)

Der AIST-R/UST-R von Bergmann und Eder (2005) ist ein Interessenfragebogen, welcher basierend auf dem Modell von J. L. Holland (1997) konzipiert ist, um schulische bzw. berufliche Interessen zu erheben. Er kann ab dem Alter von 14 Jahren eingesetzt werden und liegt als Papier-Bleistiftverfahren vor. Je nach Fragestellung können beide Verfahren (AIST-R und UST-R) einzeln oder gemeinsam vorgegeben werden.

Der AIST-R besteht aus 60 Items, wobei jeweils zehn eine der sechs Interessendimensionen, nämlich realistic, investigative, artistic, social, enterprising und conventional (vgl. Tab. 6)

erfassen sollen. Die Items des AIST-R sind so formuliert, dass sie kurz und prägnant verschiedene Tätigkeiten aus diesen Interessendimensionen darstellen. Die Testperson hat dabei die Aufgabe, ihr Interesse für jede Tätigkeit auf einer fünfstufigen unipolaren Skala (interessiert mich sehr, interessiert mich ziemlich, interessiert mich etwas, interessiert mich wenig, interessiert mich gar nicht) anzugeben. Bei der Auswertung werden für jede Interessendimension die Werte summiert und anschließend mit der Referenzpopulation verglichen. Außerdem erfolgt eine Typenzuordnung anhand der drei ausgeprägtesten Interessen. Durch die Typenzuordnung werden der Testperson Berufe, deren Umwelttyp sich mit dem Personentyp der Testperson decken, vorgeschlagen. Zusätzlich können Konsistenz und Differenziertheit der Interessen einer Person bestimmt werden.

Der Testteil UST-R besteht analog zum AIST-R aus den gleichen 60 Items. Diese sollen jedoch diesmal in Bezug auf einen bestimmten (zum Beispiel von der Testperson bevorzugten) Beruf nach ihrer Wichtigkeit ebenfalls auf einer fünfstufigen unipolaren Skala beurteilt werden (sehr wichtig, ziemlich wichtig, mittelmäßig wichtig, wenig wichtig, nicht wichtig). Anschließend kann eine Auswertung hinsichtlich der Übereinstimmung (Kongruenz) von Personen- und Umweltstruktur vorgenommen werden. Der UST-R wird häufig dann vorgegeben, wenn der Beruf im Berufsregister nicht aufscheint oder wenn insbesondere der subjektive Umwelttyp eines Berufes durch eine Testperson bestimmt werden soll. (Bermann & Eder, 2005)

5.2.2 Explorix®

Explorix® - das Werkzeug zur Berufswahl und Laufbahnplanung von Jörin, Stoll, Bergmann und Eder (2004) beruht ebenfalls auf dem Holland Modell und ist eine Weiterentwicklung des Self-Directed Search (SDS) nach Holland. Das Verfahren kann ab dem Alter von 15 Jahren durchgeführt werden (Jörin, Stoll, Bergmann & Eder, 2003) und ist sowohl schriftlich als auch am Computer durchführbar. Im Gegensatz zum AIST-R, der sich hauptsächlich auf berufliche Interessen bezieht und damit eher für die erste Berufswahl geeignet ist, erhebt Explorix® den Anspruch, „[...]ein breiteres Spektrum der Persönlichkeit“ (Jörin et al, 2004, S. 8) zu erheben, da er neben den Einschätzungen von Interessen auch Fähigkeitseinschätzungen berücksichtigt. Die Autoren sehen das Verfahren daher als besonders für Erwachsene und Schüler aus weiterführenden Schulen (Laufbahngestaltung) geeignet. Der Aufbau gliedert sich in fünf Abschnitte. Bei den Abschnitten zwei bis vier wird von der Testperson verlangt, mittels des dichotomen Antwortformats Selbsteinschätzungen bezüglich der eigenen Personen durchzuführen:

Im ersten Abschnitt soll die Testperson Angaben zu ihrer Person und ihren Berufswünschen machen.

Im zweiten Abschnitt soll die Testperson eine Liste von Tätigkeiten dahingehend bewerten, ob sie diese Tätigkeiten gerne oder ungerne ausführen würde, wobei die Testperson darauf hingewiesen wird, dass diese Bewertung unabhängig von ihren Fähigkeiten sein soll. Pro Interessendimensionen (realistic, investigative, artistic, social, enterprising und conventional, vgl. Tab. 6) werden hierbei 11 Tätigkeiten vorgegeben.

Es folgt im dritten Abschnitt eine Liste mit Fähigkeiten, bei der die Testperson angeben soll, ob sie diese gut bzw. kompetent ausführen kann (ja) oder nicht gut kann bzw. noch nie ausgeführt hat (nein). Wieder gibt es elf Aussagen pro Interessendimension.

Im vierten Teil folgt eine Liste mit Berufen, bei denen angegeben werden soll, ob die Berufe interessant bzw. ansprechend sind (ja) oder eher uninteressant sind bzw. ablehnend wirken (nein). Hier werden 14 Berufe pro Interessendimension vorgegeben.

Im fünften und letzten Abschnitt soll die Testperson eine Selbsteinschätzung im Vergleich zu Gleichaltrigen bezüglich zweier genereller Fähigkeiten pro Interessendimension auf einer siebenstufigen (1-7) unipolaren Skala (gering = 1, mittel = 4, hoch = 7) vornehmen.

Anschließend werden bei der Auswertung pro Interessendimension die Anzahl der gern- und ja-Antworten mit den Selbsteinschätzungen addiert, um zu einem Gesamtwert pro Dimension zu gelangen. Es wird der Holland-Code ermittelt und Berufe mit allen möglichen Kombinationen der drei am höchsten ausgeprägtesten Interessen betrachtet (Jörin et al., 2003). Die Ermittlung von Kongruenz, Differenziertheit und Kohärenz ist ebenfalls möglich. Es liegen jeweils Formen für Deutschland, Österreich und die Schweiz vor, da es länderspezifische Berufsregister gibt. Neben der Möglichkeit der selbstständigen Durchführung und Auswertung (was aus diagnostischer Perspektive kritisch zu betrachten ist) bieten die Autoren die Möglichkeit einer Onlinedurchführung mit Ergebnisbericht an.

5.2.3 Multimethodische Objektive Interessentestbatterie (MOI)

Die Multimethodische Objektive Interessentestbatterie (MOI) von Proyer und Häusler (2008) beruht ebenfalls auf der Theorie von Holland (1997). Sie erfasst berufliche Interessen auf Basis von expliziten Fragebogenverfahren und objektiven (impliziten) Verfahren. Die MOI ist ein Computerverfahren, welches ab dem Alter von 14 Jahren eingesetzt werden kann. Den Einsatz objektiver Verfahren wählten die Autoren, damit „Informationen zugänglich gemacht werden, die der Testperson in der Selbstbeschreibung nicht zugänglich sind[...]“ (S. 4). Die MOI setzt sich

aus folgenden sechs Subtests zusammen:

1. Der Subtest *Berufliche Identität* versucht anhand von 20 Fragen die berufliche Identität einer Person zu erfassen.
2. Der nonverbale Subtest *Bilderlisten* erhebt mit dichotomem Antwortformat, ob die auf den 60 Bildern vermittelten Inhalte „interessant“ oder „uninteressant“ sind.
3. Der verbale Subtest *Wortlisten* besteht aus 96 beruflichen Tätigkeiten, die ebenfalls mit Hilfe des dichotomen Antwortformats als „interessant“ oder „uninteressant“ eingestuft werden sollen.
4. Der objektive Subtest *Ablenkbarkeit* ist ein Speed-Test, bei dem die Testperson die Aufgabe hat, Geschichten zu lesen und unpassende (erfundene oder nicht passende sinnvolle) Wörter durchzustreichen. Pro Interessendimension gibt es die Geschichte einer Person, die in einem der sechs RIASEC-Bereiche arbeitet. Es handelt sich hierbei um einen Handwerker (R), einen Wissenschaftler (I), einen Schauspieler (A), einen Arzt (S), einen Ladenbesitzer (E) und einen Beamten (C). Folglich liest jede Testperson sechs verschiedene Geschichten. Pro Geschichte steht der Testperson eine Minute Zeit zur Verfügung.
5. Beim objektiven Subtest *Aufteilung* hat die Testperson die Aufgabe, 100.000 Euro auf unterschiedliche Vereine, Projekte oder Organisationen zu verteilen, welche wiederum den sechs Interessendimensionen von Holland zugeteilt werden können. Die Menge des zugeteilten Geldes dient als Indikator für das Interesse.
6. Beim semi-projektiven Subtest *Tachistoskop* werden der Testperson 20 Bilder gezeigt, die tachistoskopisch mehrdeutig sind und nur für ganz kurze Zeit am Bildschirm erscheinen. Die Testperson hat nun die Aufgabe, mittels multiple-choice Antwortformat jene Antwortmöglichkeit auszuwählen, von der sie vermutet, sie gerade am Bildschirm bildlich dargestellt gesehen zu haben. Wobei die Antwortmöglichkeiten in die sechs Interessendimensionen einteilbar sind.

Bei der Durchführung werden zwei Testformen zur Verfügung gestellt, die Langform (S1), welche alle Subtests beinhaltet und 40-60 Minuten beansprucht, sowie die Kurzform (S2), welche ausschließlich die zwei „expliziten“ Subtests beinhaltet und um die 15-20 Minuten dauert.

Bei der Auswertung gibt es eine Variable für die berufliche Identität, welche sich aus dem Summenwert der Fragen des Subtests *berufliche Identität* ergibt. Dieser wird dann mit der Referenzstichprobe verglichen. Außerdem werden die zwei Hauptvariablen *Gesamtergebnis (explizite Interessen)* und *Gesamtergebnis (implizite Interessen)* wiedergegeben. Für jede der

sechs Interessendimensionen wird pro Hauptvariable ein Mittelwert der normierten Testwerte der dazugehörigen Subtests berechnet. Ebenso wird die Differenziertheit und ein Holland-Code sowohl für implizite als auch für explizite Interessen ausgegeben.

Bereits genannte Verfahren (vgl. Kapitel 5.2.1, 5.2.2) nutzen den dreistelligen Holland-Code dazu, um aus dem Berufsregister entsprechende und passende Berufe vorzuschlagen bzw. aufzuzeigen. Im Gegensatz dazu entwickelten die Autoren des MOI ein *Verzeichnis mit beruflichen Tätigkeiten*, die entsprechend dem Holland-Code von Experten kodiert wurden. Es besteht dadurch die Möglichkeit, anhand von einzelnen Tätigkeit ein für die Testperson passendes Berufsprofil zu erstellen.

5.3 Das Interessenkonzept sensu Brickenkamp (1990)

Brickenkamp (1990) definiert „Interessen vereinfacht als emotional-kognitive Verhaltenspräferenzen [...], die sich hinsichtlich verschiedener Merkmale (Interessenrichtungen, Verhaltensmodalitäten, Entstehungsgeschichte, Generalität, Stabilität, Intensität, Vielseitigkeit) voneinander unterscheiden.“ (S. 7). Mit «emotional» meint Brickenkamp die Vorliebe für Tätigkeiten, deren Ausführung mit der Erwartung, bestimmte Gefühle zu erfüllen, verbunden ist. Dies deutet gleichzeitig auf einen motivationalen Faktor von Interessen hin. Mit dem Begriff „kognitiv“ weist Brickenkamp darauf hin, dass es möglich ist, Interessen für eine Beschäftigung zu haben, obwohl diese noch nicht ausgeführt wurde bzw. ausgeführt werden konnte.

Brickenkamp stellte Überlegungen dazu an, dass neben den Interessengebieten (Inhaltsklassen), welche bei den meisten Interessentests erhoben werden, auch noch eine Kategorisierung in Verhaltensklassen möglich sei. Er postuliert hierzu, dass die Entwicklung von Interessen für ein Interessensgebiet in drei Schritten bzw. Phasen abläuft, welche aufeinander aufbauen (vgl. Abb. 11).

1. Phase: Rezeptivität

Bei der Entwicklung von Interessen für ein Interessensgebiet kommt es zuerst zu rezeptivem Interesse bzw. Verhalten. Das bedeutet, dass eine Person eine „erhöhte Aufnahmebereitschaft für Reize und Informationen aus bestimmten Interessengebieten“ (S. 10) hat. In dieser Phase steigt sozusagen die Empfänglichkeit für Inhalte aus einem interessierenden Gebiet. Es kommt zu bevorzugten Tätigkeiten wie „die Suche nach Informationen, Lesen, Zuhören, Beobachten“ (S. 11).

2. Phase: Reproduktivität

In der zweiten Phase entwickelt sich bei einer Person das Bedürfnis, Tätigkeiten aus dem interessierenden Interessengebiet auszuüben bzw. nachzumachen. Es besteht der Wunsch zu reproduktivem Verhalten, womit aktive Beteiligung bei vorgegebenen Verhaltensmustern in diesem Bereich gemeint ist. Bevorzugte Tätigkeiten in dieser Phase sind unter anderem „Mitmachen“, „Nachmachen“ und „Nachahmen“, jedoch ohne negative Akzentuierung dieser Begriffe“ (S. 11).

3. Phase: Kreativität⁷

Schließlich ist die dritte Phase jene Phase, bei der die Person das Bedürfnis bekommt schöpferisch tätig zu werden. Personen in dieser Phase geht es darum, „über vorgefundene Verhaltensmuster hinaus eigene Ideen, Konzepte, Vorstellungen einzubringen, die einen mehr oder weniger großen Neuheitswert besitzen“ (S. 11).

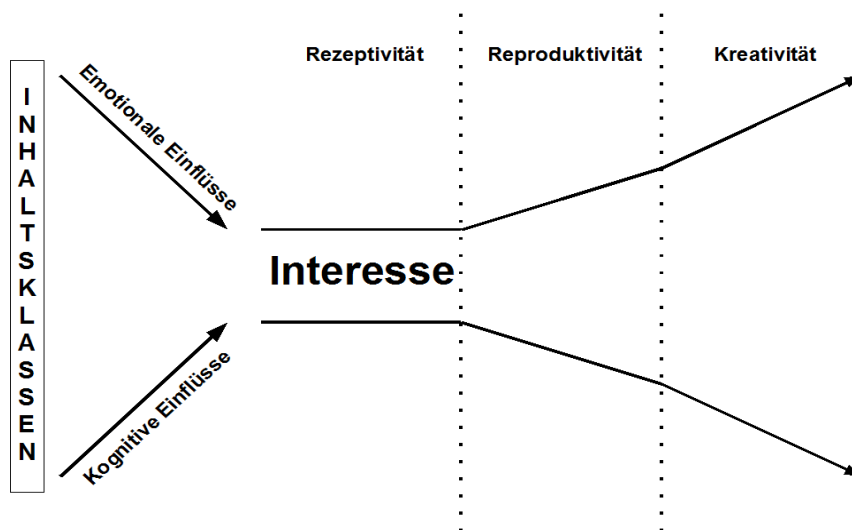


Abb. 11: Eigenkonstruktion: grafische Darstellung des Interessenkonzept sensu Brickenkamp (1990)

Abb. 11 zeigt die grafische Darstellung der Interessenkonzeption sensu Brickenkamp (1990). Kurz zusammengefasst bestimmen emotionale und kognitive Einflüsse die Entwicklung des Interesses für bestimmte Interessengebiete. Die Entwicklung kann dabei in die drei Verhaltensmodalitäten Rezeptivität, Reproduktivität und Kreativität unterteilt werden, welche, wie bereits erwähnt, aufeinander aufbauen.

⁷ Brickenkamp (1990) ist sich durchaus bewusst, dass „eine absolute Trennung kreativen und nicht-kreativen Verhaltens unrealistisch wäre“ (S. 11). Jedoch weißt er darauf hin, dass manche Tätigkeiten mehr kreative Fähigkeiten erfordern als andere.

Aus seiner Definition heraus weist Brickenkamp weiters darauf hin, dass Interessen nicht alleine durch realisiertes Verhalten gezeigt werden und dementsprechend (diagnostische) Verfahren notwendig sind, um emotional-kognitive Verhaltenspräferenzen zu erschließen.

Auf Basis dieser Grundannahmen konstruierte Brickenkamp im Jahre 1990 einen Interessenfragebogen, der im Folgenden beschrieben wird.

5.3.1 Generelle Interessen-Skala (GIS)

Die Generelle Interessen-Skala ist ein von Brickenkamp (1990) konstruierter Interessenfragebogen, welcher versucht, so ökonomisch wie möglich ein breites Spektrum an Interessen zu erfassen und ab dem Alter von 13 Jahren als Papier-Bleistift Verfahren eingesetzt werden kann. Wichtig für Brickenkamp war die Unterscheidung in Interessenklassen und Verhaltensmodalitäten. In Bezug auf die Interessenklassen führte Brickenkamp Befragungen durch und konzipierte auf Basis dieser und seiner Erfahrungen die GIS mit 16 Interessengebieten. Explizit werden folgende Bereiche erfasst:

- | | |
|----------------|------------------------------|
| 1. Musik | 9. Kommunikationstechnologie |
| 2. Kunst | 10. Naturwissenschaft |
| 3. Architektur | 11. Biologie |
| 4. Literatur | 12. Natur/Landwirtschaft |
| 5. Politik | 13. Ernährung |
| 6. Handel | 14. Mode |
| 7. Erziehung | 15. Sport |
| 8. Medizin | 16. Unterhaltung |

In Bezug auf die drei Verhaltensmodalitäten unterscheidet Brickenkamp drei Klassen von Tätigkeiten. Tätigkeiten die rezeptiv sind, welche die reproduktiv sind und Tätigkeiten die als kreativ gelten.

Das Testmaterial der GIS ist ein Fragebogen, der in vier verschiedenen Formen vorliegt. Jede Form enthält identische Items, jedoch entweder Normprofile für männliche oder weibliche Erwachsene oder für männliche oder weibliche Kinder. Insgesamt besteht der Fragebogen aus 48 Items. Von den 48 Items entsprechen jeweils drei Items einem der 16 Interessenbereiche. Von diesen drei Items wiederum entspricht jedes einer der drei Verhaltensmodalitäten (siehe Beispiel, S. 64).

Beispiel (Brickenkamp, 1990, S. 15):

„Kommunikationstechnologische Interessen

Rezeptiv: Sich über neue Computer informieren

Reproduktiv: Mit Computern spielen oder arbeiten

Kreativ: Computerprogramme erstellen“

Für jedes dieser drei Items aus dem oben genannten Beispiel soll die Testperson mittels einer unipolaren sechs-stufigen Ratingskala die Höhe ihres Interesse für diese Tätigkeit bekannt geben (0= kein Interesse bis 5= hohes Interesse). Für die Durchführung werden in etwa 15 bis 20 Minuten benötigt.

Die Auswertung der GIS erfolgt sehr rasch und ist zudem simpel. Für jede der 16 Interessendimensionen müssen lediglich die drei Werte je Zeile (entspricht einer Interessendimension) addiert und im Profil eingetragen werden. Für die Verhaltensmodalitäten muss jede der drei Spalten summiert werden und somit ergibt sich pro Verhaltensklasse ein Wert, der ebenfalls in das Profilblatt eingetragen wird. Zusätzlich besteht die Möglichkeit durch die Summierung aller Werte einen Gesamtwert zu bestimmen.

5.4 Entwicklung des Fragebogens zum Interesse am Psychologiestudium (FIP)

Die Notwendigkeit und der Bedarf eines Interesseninstrumentariums für das Wiener *Self-Assessment für Psychologie* wurde bereits zu Beginn dieses Kapitels geklärt. Hier werden nun im Anschluss die einzelnen Entwicklungsschritte des Fragebogens zum Interesse im Psychologiestudium – FIP skizziert.

Die ersten Überlegungen zielten darauf ab, Klarheit darüber zu bekommen, welche Zielpopulation angesprochen werden soll, welche Methode der Feststellung von Interessen angewendet werden soll und welche Interessengebiete mit dem neuen Verfahren erhoben werden sollen.

Die Frage nach der **Zielpopulation** musste eigentlich nicht mehr gestellt werden, da von Beginn dieser Arbeit an klar war, dass ein Verfahren für die Anwender des Self-Assessment Psychologie konstruiert werden sollte.

Bezüglich der **Methodik** musste eine Entscheidung zwischen objektiven und subjektiven Methoden getroffen werden (vgl. Kapitel 5.1.1). Es stand fest, dass es sich beim Self-Assessment um kein Auswahlverfahren von Seiten der Universität handelt (vgl. Kapitel 1) und somit keine Selektionssituation dadurch geschaffen wird, in der ein Verfälschen von Antworten zu erwarten wäre. Zudem wurde angenommen, dass die zukünftigen Testpersonen die Verfahren ordnungsgemäß durchführen, da es sich um eine Entscheidungshilfe bezüglich ihrer akademischen Entwicklung handelt. Demzufolge schienen objektive Erhebungsmethoden nicht notwendig zu sein und es sollte somit ein subjektiver Interessenfragebogen entwickelt werden.

Zur Frage nach den **Interessengebieten**, welche im Fragebogen integriert sein sollten, wurde vor allem berücksichtigt, dass die Verfahren im Self-Assessment insbesondere von Personen durchgeführt werden, die sich für ein Psychologiestudium interessieren (schließlich wollen jene wissen, ob sie dafür auch geeignet sind). Nun schien die Vorgabe eines allgemeinen Interessenfragebogens, welcher allgemeine berufliche bzw. schulische (bildungsbezogene) Interessen erhebt, größtenteils überflüssig zu sein. Denn die Anwender des Self-Assessments vermuten bereits, dass sie sich für Psychologie interessieren, womit implizit „expressed interests“ (nach Super, 1957; vgl. 5.1.1) gegeben wären. Vielmehr bedarf es einer Abklärung, ob das vermutete Interesse für die Inhalte des Studiums tatsächlich gegeben ist und ob eine realistische Vorstellung über die Inhalte im Studium besteht. In diesem Sinne sollte ein Fragebogen basierend auf einer „speziellen Interessentheorie“ entwickelt werden. Spezielle Interessentheorien verwenden laut Schiefele, Prenzel, Krapp, Heiland und Kasten (1983) „allgemeine Interessentheorien“ als Ausgangspunkt und Grundlage für sich selbst.

Der FIP setzte sich folglich zum Ziel, das Interesse für die Inhalte des Psychologiestudiums zu erheben und gleichzeitig Testpersonen einen ersten Einblick in die Schwerpunkte des Studiums zu verschaffen.

Die Konstruktion sollte wie in Kapitel 2 erwähnt regelgeleitet erfolgen, um die Validität bereits zu Beginn zu sichern und um die Konstruktion eines gleichwertigen Fragebogens für eine andere Studienrichtung zu ermöglichen.

5.4.1 Definition des Konstrukts bzw. Festlegung des Sachverhaltes

Wie in Kapitel 2 erwähnt besteht der erste Schritt der regelgeleiteten Konstruktion darin, das zu messende Konstrukt zu definieren bzw. den genauen Inhalt, welcher sich in den Aufgaben wiederfinden soll, abzugrenzen.

Als grundlegende Theorie für den FIP wurde die Interessenkonzeption sensu Brickenkamp (1990) herangezogen (vgl. Kapitel 5.3). Folglich wurde davon ausgegangen, dass sich Interesse in drei Schritten entwickelt und als emotional-kognitive Verhaltenspräferenzen definiert werden kann. Im Weiteren sollte eine Methode gefunden werden, welche möglichst objektiv und exakt den Bereich absteckt, welcher für die Inhalte der Items zur Verfügung stehen sollte. Wie bereits erwähnt sollten sich die Items auf die Inhalte des Psychologiestudiums beziehen, daher wurde zur Festlegung der Inhalte zuerst der Studienplan des Psychologiestudiums in Wien herangezogen. Bei der Durchsicht des Studienplans wurden die Inhaltsbereiche anhand der Semesterstundenanzahl, welche sie im Studium einnehmen, aufgegliedert (vgl. Tabelle 7).

Tabelle 7: Auflistung der Inhaltsbereiche im Psychologiestudium nach Sst.

Inhaltsbereich	Sst.	Wahlfach Sst.	Gesamt (möglich) Sst.
Psychologische Diagnostik	14	Spezielle Psychologische Diagnostik: 6, 10, 16	14 (20, 24, 30)
Klinische Psychologie	12	Klinische und Gesundheitspsychologie: 6, 10, 16	12 (18, 22, 28)
Entwicklungspsychologie	8	Angewandte Kinder- und Jugendpsychologie: 6, 10, 16	8 (14, 18, 24)
Sozialpsychologie	8	Angewandte Sozialpsychologie: 6, 10, 16	8 (14, 18, 24)
Wirtschaftspsychologie	8	Wirtschaftspsychologie: 6, 10, 16	8 (14, 18, 24)
Bildungspsychologie	8	Bildung / Evaluation / Training: 6, 10, 16	8 (14, 18, 24)
Forschungsmethoden und Evaluation	8		8 (14, 18, 24)
Methodenlehre	14		14
Allgemeine Psychologie	10		10
Differentielle Psychologie	8		8
Biologische Psychologie	8		8
Einführung in die Psychologie	4		4

Da auf die Ökonomie des Fragebogens geachtet werden sollte und dieser daher nicht zu viele Skalen beinhalten sollte, wurden jene Bereiche als Skalen aufgenommen, in denen es die Möglichkeit einer Spezialisierung im Sinne eines Wahlfaches gab. Zusätzlich wurde der Bereich Methodenlehre berücksichtigt, da dieser grundlegend für die meisten anderen Bereiche ist, so erfordert zum Beispiel alleine das Lesen eines Fachartikels grundlegendes methodisches Wissen. Ferner wurden aufgrund der biologischen Aspekte im Wahlfach Klinische und

Gesundheitspsychologie diese zwei Bereiche miteinander verbunden. Es ergaben sich dadurch folgende Skalen, welche im FIP berücksichtigt wurden:

1. Psychologische Diagnostik
2. Methodenlehre
3. Klinische Psychologie (mit Biologischen Aspekten)
4. Entwicklungspsychologie
5. Sozialpsychologie
6. Wirtschaftspsychologie
7. Bildung und Evaluation

Um die Inhalte nun weiter einzugrenzen, schien es erforderlich, sich mit den Hauptinhalten dieser sieben Schwerpunkte genauer auseinanderzusetzen. Auch wenn aufgrund der Erfahrung eines Testautors bereits auf Inhalte geschlossen werden konnte, sollte ein Weg gefunden werden, der so objektiv wie möglich die Inhalte für die Schwerpunkte eingrenzt. Es gäbe die Möglichkeit sich anhand von Skripten oder dem Vorlesungsverzeichnis genauer mit den Inhalten der einzelnen Bereiche zu beschäftigen, jedoch um diesen Prozess objektiver zu gestalten, schien es eine gute Lösung zu sein, Definitionen der einzelnen Bereiche als Inhaltsfestlegung zu verwenden. Demnach wurde eine Liste mit Definitionen aus allen Schwerpunkten zusammengestellt (ein Auszug dieser befindet sich unter Anhang 1).

In einer Befragung von Studiengangwechslern deckten Steckel und Peters (2008) auf, dass vier von zehn Befragten falsche Erwartungen an das Studium gehabt hatten. Es scheint daher in der Bevölkerung teilweise eine falsche Vorstellung über Inhalte in manchen Studien zu kursieren. Daraus ergab sich die Idee, eine Skala in den Fragebogen aufzunehmen, die typische Irrtümer bzw. falsche Erwartungen erfasst, um dadurch solche Studienmotive so früh wie möglich aufzudecken. Um fundierte Irrtümer zu identifizieren, wurden Gutachten aus dem Proseminar „Praktikum zum Psychologischen Diagnostizieren: Eignungsdiagnostik zur Studierendenauswahl“, aus dem WS 2006/07 und dem SS 2007 unter der Leitung von Frau Mag. Martina Frebort durchgearbeitet. Dies geschah mit der Einverständnis von Herrn Mag. Sonnleitner unter der Voraussetzung der Einhaltung der Anonymität und der Verschwiegenheit meinerseits über weitere Inhalte, welche über die „Typischen Irrtümer“ hinausgehen. Das Hauptaugenmerk bei der Sichtung dieser Gutachten lag auf den Vorstellungen der Psychologieinteressenten über die Inhalte des Psychologiestudiums sowie in manchen Fällen auch auf den beruflichen Vorstellungen nach dem Studium. Im Weiteren wurde in einer Bibliothek gesichtet, welche Themen bzw. Inhalte unter dem Begriff Psychologie eingeordnet

waren, jedoch nichts mit dem Psychologiestudium an sich zu tun hatten. Folgende Inhalte ergaben sich dadurch:

- Forensik (Forensische Psychologie), Kriminalpsychologie
- Psychotherapie
- Psychoanalyse
- Erkennen wie Menschen denken, Persönlichkeiten analysieren
- Viel Praxis
- Mediator
- Graphologie
- Verschreibung von Medikamenten
- Langzeitprognosen
- Traumdeutung

Außerdem kam in dieser Phase der Konstruktion des FIP von Mitarbeitern aus dem Arbeitsbereich Psychologische Diagnostik die Anregung, Inhalte des Psychologiestudiums in Wien mit den Inhalten der Psychologiestudien in anderen Bundesländern zu vergleichen. Bei diesem Vergleich fiel auf, dass besonders das Psychologiestudium in Klagenfurt Schwerpunkte beinhaltet, die in Wien nicht behandelt werden. Daraus entstand die Idee, eine Skala für Inhalte zu entwickeln, die in Klagenfurt geboten werden, jedoch in Wien nicht integriert sind. Gegebenenfalls könnten Studenten darauf hingewiesen werden, dass sich ihre Interessen eher mit den Inhalten des Psychologiestudiums in Klagenfurt decken. Bei der Durchsicht der Schwerpunkte in Klagenfurt wurden folgende Bereiche identifiziert, die sich nicht im Studienplan des Wiener Psychologiestudiums wiederfinden:

- Klinisch-psychologische Intervention und Psychotherapie
- Gruppendynamik
- Kulturpsychologie

5.4.2 Wahl der Aufgabenform

Nachdem nun der Rahmen für die Inhalte der Aufgaben festgesetzt wurde, kommt es im zweiten Schritt zur Wahl der Aufgabenform. Welches Antwortformat soll gewählt und welche Materialart soll verwendet werden?

In Kapitel 5.1.1 wurde bereits ein Überblick über mögliche Antwortformate gegeben. Obgleich dabei das forced-choice Antwortformat im Gegensatz zum free-choice Antwortformat nicht besonders gut abschnitt, scheint es nach gründlichen Überlegungen besonders gut für den Fragebogen zum Interesse am Psychologiestudium geeignet zu sein. Hinsichtlich der theoretischen Annahme, dass sich Interessen in drei Schritten entwickeln, scheint ein forced-choice Antwortformat sehr gut geeignet zu sein, um festzustellen auf welcher „Entwicklungsstufe“ sich die jeweilige Testperson bezüglich eines Interessenbereiches befindet. Dies kann zudem im Sinne einer Ratingskala verstanden werden.

Da der FIP mit Bezug auf die Interessenkonzeption von Brickenkamp (1990) entwickelt werden sollte, war es sinnvoll die Items der GIS (1990) näher zu betrachten. Wie unterschied er zwischen den drei Verhaltensmodalitäten und welche Materialart verwendete er hierfür?

Es zeigte sich (Beispiel S. 64), dass die Unterscheidung zwischen den drei Verhaltensmodalitäten aufgrund von Tätigkeiten vorgenommen wurde. Brickenkamp subsumierte einen Teil von Tätigkeiten zu rezeptivem Verhalten, einen Teil zu reproduktivem und einen Teil zu kreativem Verhalten. Demzufolge wurde auch im FIP anhand von Tätigkeiten zwischen den Verhaltensklassen unterschieden. Da aber die Vorgabe von einem Item wie zum Beispiel „Vorlesung besuchen“ alleine noch nicht ausreicht um das Interesse für Inhalte aus den einzelnen Schwerpunkten zu erheben, sollte das Antwortformat noch erweitert werden. Es wurde die Komponente Inhalt noch hinzugefügt; dies war durch verschiedene Materialarten denkbar wie zum Beispiel Vorlesungstitel, Seminartitel, Referatsthemen, Hypothesen etc. Schließlich zeigte Todt (1978), dass Jugendliche relativ gleich auf unterschiedliche Materialarten reagieren (vgl. Kapitel 5.1.1).

5.4.3 Erstellung von Regeln

Nach der Festlegung des Antwortformates kommt es nun zur Bestimmung der Regeln, welche dafür zuständig sind die Inhalte mit dem Antwortformat zu verbinden.

Vorgehensweise in Schritten (Regeln) zusammengefasst

1. Studienplan der Studienrichtung Psychologie besorgen.
2. Den Studienplan durchsehen, Schwerpunkte identifizieren und als Basis für das weitere Vorgehen festlegen. Gegebenfalls können ähnliche oder zusammenhängende Schwerpunkte zu einem zusammengefasst werden. Schwerpunkte können zum Beispiel durch die Semesterstundenanzahl in den einzelnen Bereichen festgelegt werden (wie es

hier im Weiteren gemacht wurde). Anzumerken ist, dass die Semesterstundenanzahl jene Zeitangabe ist, die wöchentlich für ein bestimmtes Fach in einem Semester aufgewandt wird.

3. Nachdem man die Schwerpunkte festgelegt hat, kommt es zur Sammlung von Hauptinhalten in den einzelnen Schwerpunkten. Dazu dient gegebenenfalls eine Definition der Inhalte der einzelnen Schwerpunkte. Falls diese Information nicht vorhanden ist, können auch Skripten zur Hand genommen und auf ihre Hauptthemen hin durchsucht werden.
4. Gegebenenfalls können Schwerpunkte aus anderen Bundesländern des Psychologiestudiums zusammengetragen werden, die sich nicht im Wiener Psychologiestudium befinden, um eine Skala anzufertigen, die diese Inhalte misst.
5. Durch Befragung von Studierenden des Psychologiestudiums herausfinden, welche Tätigkeiten in den einzelnen Schwerpunkten vorkommen, vorkommen könnten oder verlangt werden. Teilweise besteht die Möglichkeit, geforderte Tätigkeiten durch die Durchsicht des Vorlesungsverzeichnisses ausfindig zu machen.
6. In einem weiteren Schritt wird nun eine Liste mit Tätigkeiten angefertigt, die in Schritt 5 zusammengetragen wurden. Diese Liste wird anschließend in die drei Verhaltensmodalitäten unterteilt. Somit sollten drei Listen mit Tätigkeiten zur Verfügung stehen, eine mit rezeptiven Tätigkeiten, eine mit reproduktiven und eine mit kreativen Tätigkeiten. Falls es für eine Verhaltensmodalität zu wenige Tätigkeiten gibt, können noch einzelne durch eigene Überlegungen hinzugefügt werden.
7. Ausgehend von der Definition des Schwerpunkts und den Tätigkeiten werden nun Items erstellt. Ein Item sollte sich dabei auf einen Inhaltsbereich eines Schwerpunktes beziehen. Zudem sollte das Item aus drei Sätzen konstituiert sein, welche jeweils eine Verhaltensmodalität abdecken. Dabei werden so viele Items erstellt, wie es Inhaltsschwerpunkte in einem Bereich gibt. Zur besseren Verständlichkeit sei auf das Instruktionsbeispiel in Abb. 12 S. 72 verwiesen.
8. In einem nächsten Schritt werden nun die vorhandenen Items auf ihr sprachliche Verständlichkeit und ihr generelles Verständnis für die Zielpopulation überprüft. Gleichzeitig wird reflektiert, ob diese Items zu den gegebenen Schwerpunkten passen. Falls Items nicht dem Verständnis für die Zielpopulation entsprechen oder aus anderen Gründen nicht geeignet scheinen, dann können diese mit Hilfe von Experten überarbeitet werden.
9. Zuletzt wird wiederum mit Experten diskutiert, welche Items aus diesem vorhandenen

Pool, die geeigneten für das Self-Assessment sind - mit dem Ziel, vier oder fünf Items (je nach Wunsch) für jeden Schwerpunkt auszuwählen und abschließend den Fragebogen fertig zu stellen.

Die Schritte eins bis vier wurden bereits in diesem Kapitel besprochen. Ein Auszug der Liste der Definitionen befindet sich in Anhang 1. In Tabelle 8 werden die in Schritt fünf und sechs erhobenen Tätigkeiten der einzelnen Verhaltensmodalitäten dargeboten.

Tabelle 8: Aufgliederung der Tätigkeiten zu den Verhaltensmodalitäten

Rezeptive Tätigkeiten	Reproduktive Tätigkeiten	Kreative Tätigkeiten
lesen	vorgeben, durchführen, anwendend	(weiter)entwickeln, optimieren, ableiten
Seminar (Vorlesung) besuchen	teilnehmen, Übung machen	Gutachten interpretieren (Auswertung erklären)
etwas lernen	Dokumentieren, (beobachten, darstellen)	Studie (Projekt, Maßnahme) planen (durchführen, erstellen, ausarbeiten)
(an)hören	Referat (Arbeit, Handout) schreiben (verfassen, halten, erstellen, anfertigen)	Konzept ausarbeiten (Bedingungen überlegen, Merkmale zusammentragen)
(an)sehen	auswerten, untersuchen	analysieren, reflektieren (überprüfen)
recherchieren	Störvariablen erkennen	Neues finden, erforschen, erfinden
		anfertigen, erstellen, entwerfen

Die Schritte sieben bis neun bedingten einen langwierigen Prozess, während dem es zu mehreren Besprechungen mit Experten kam. Die Items wurden dabei bezüglich ihrer Verständlichkeit für die Zielpopulation, der Formulierung sowie zu einer optimalen Darstellung der Inhalte in den einzelnen Schwerpunkten stetig modifiziert.

Schließlich wurde entschieden, für jeden Schwerpunkt sowie für die Skala „Klagenfurt“ und für die Skala „typische Irrtümer/falsche Erwartungen“ jeweils fünf Items auszuwählen. Somit besteht der Fragebogen zum Interesse im Psychologiestudium – FIP nun aus neun Skalen mit je fünf Items; was einer Gesamtanzahl von 45 Items entspricht. Die Items sollten jedoch nicht skalenweise vorgegeben werden, um ein genaues Bearbeiten zu fördern. Sie wurden daher wie die Items im AIST-R/UST-R (Bergmann & Eder, 2005), so vermengt, dass von jeder Skala ein Item vorgegeben wird und sich diese Reihenfolge stets wiederholt.

Ein 46. Item wurde ausgewählt und sollte zum Verständnis der Bearbeitung als Instruktionsbeispiel dienen (vgl. Abb. 12).

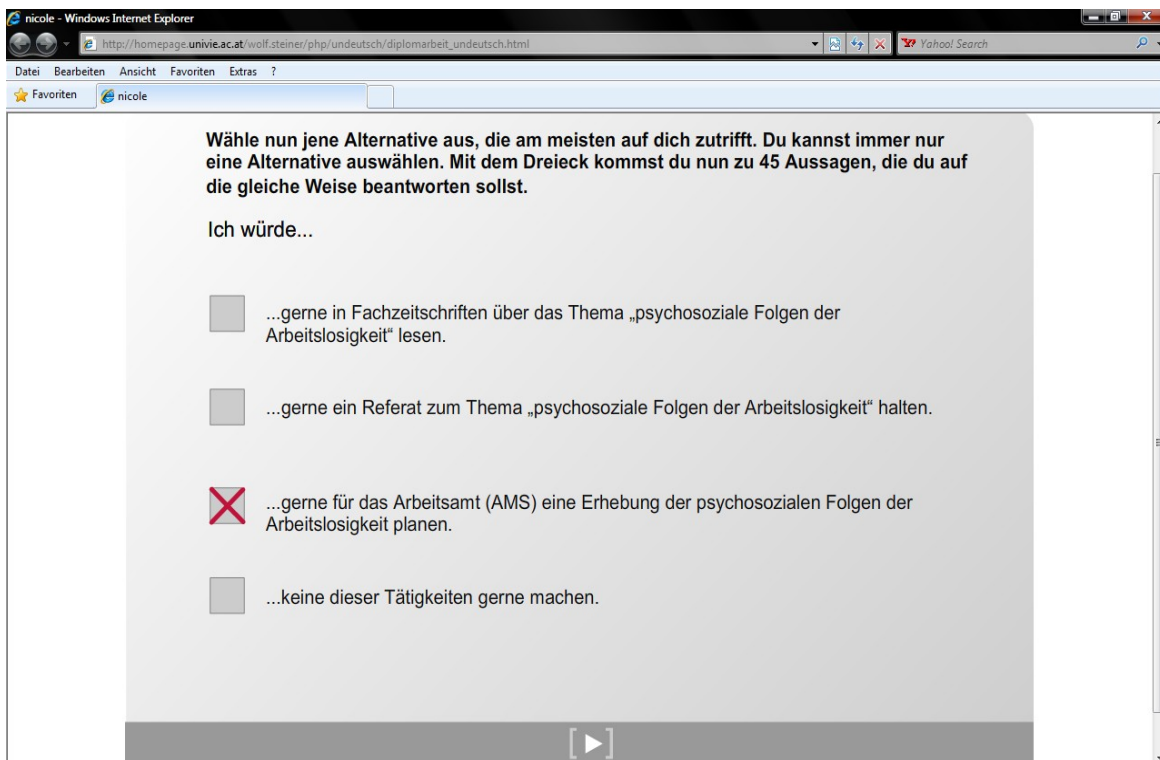


Abb. 12: Instruktionsbeispiel des FIP

6 Datenerhebung

Die in Kapitel 4 und Kapitel 5 entwickelten Verfahren wurden eigens für die Datenerhebung dieser Diplomarbeit programmiert. Die programmtechnische Umsetzung wurde von Herrn Wolfgang Steiner übernommen.

Anschließend wurde die Testbatterie für testtheoretische Analysen einer Stichprobe von 131 Personen vorgegeben.

6.1 Online Testbatterie

Da die Verfahren FMT und FIP für das Wiener Self-Assessment für Psychologie konzipiert wurden und somit später online vorgegeben werden sollten, wurden die beiden als Computerverfahren programmiert und via Internet zugänglich gemacht. Auf diese Weise sollte für die Stichprobe eine vergleichbare Bedingung, wie für zukünftige Benutzer des Self-Assessments, geschaffen werden.

Mittels Passwort konnten sich Psychologieinteressenten von zu Hause aus einloggen und die Testbatterie starten. Zu Beginn wurde den Testpersonen der Ablauf der Testung mit einem kurzen Text erläutert. Die Testpersonen wurden darauf hingewiesen, sich in etwa zwei Stunden für die Testung Zeit zu nehmen. Im Weiteren sollten sie in einem ersten Teil Fragen zu ihrer Person sowie Fragen zur Studienwahlreife beantworten. In einem zweiten Teil sollten sie beim FMT ihr logisch-schlussfolgerndes Denken unter Beweis stellen. Das Prinzip der Aufgaben wurde den Testpersonen zu Beginn mit Hilfe von zwei Beispielaufgaben (vgl. Abb. 14 & Abb. 15) vermittelt. Abschließend sollten die Testpersonen beim FIP noch Fragen hinsichtlich ihrer Interessen für das Psychologiestudium beantworten. Das Prinzip der Aufgabenstellung wurde hier ebenfalls mit einer Beispielaufgabe (vgl. Abb. 16) erläutert.

Zu den Durchführungsbedingungen zählte, dass die Testpersonen während der Vorgabe nicht die Möglichkeit hatten ein Feld unausgefüllt zu lassen bzw. eine Aufgabe nicht zu bearbeiten. Erst nach der Bearbeitung einer Aufgabe bzw. nach dem Ausfüllen einer Seite konnte die nächste Seite mit Hilfe des «Weiter-Button» geöffnet werden. Zudem konnten sie nach dem Öffnen der nächsten Seite nicht mehr auf die vorherige Seite zurückkehren.

Im Folgenden wird ein kurzer Überblick über die drei Teile der Testung gebogen.

6.1.1 Teil1: Demographische Variablen und Fragen zur Studienwahlreife

Abb. 13 zeigt die zweite Seite der Testbatterie. Der erste Teil besteht aus insgesamt drei Seiten mit 13 (bzw. 15, je nach Beantwortung der gestellten Fragen) auszufüllenden Punkten. Es wird hier nach demographischen Variablen wie Geschlecht, Alter, Schulzweig und letzte absolvierte Klasse gefragt (vgl. Abb. 13), diese sollten später der statistischen Auswertung dienen. Die Variable «letzte absolvierte Klasse» wurde deshalb so formuliert, da es aufgrund des längeren Erhebungszeitraums zu unterschiedlichen Angaben bei derselben Personengruppe hätte kommen können.

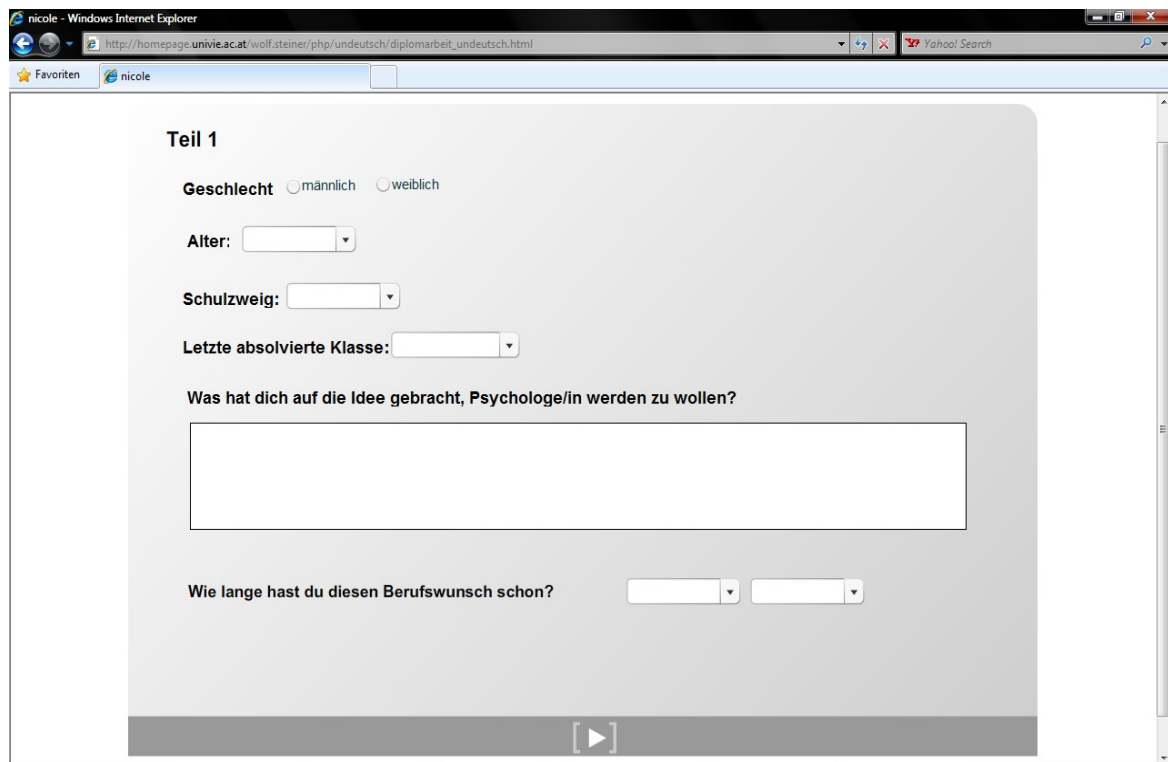


Abb. 13: Seite 2 der Testbatterie dieser Diplomarbeit

Im Weiteren wurden sieben Fragen (teilweise noch eine Zusatzfrage) zur Studienwahlreife gestellt. Diese Fragen stammten aus einem Gesprächsleitfaden zur Berufswahlreife, welcher eigens für Beratungszwecke an der Fakultät für Psychologie der Universität Wien, genauer an der Test- und Beratungsstelle des Arbeitsbereichs Psychologische Diagnostik, entwickelt wurde. Obwohl es sich hierbei um Fragen zur Berufswahlreife handelt, schienen diese vor allem in Bezug zum Self-Assessment Psychologie ebenso als Fragen zur Studienwahlreife für das Psychologiestudium geeignet zu sein.

Die Fragen des Leitfadens wurden von Kubinger (2009b, S. 256) zusammengefasst:

- Was hat Sie auf die Idee gebracht, Psychologe/in werden zu wollen?
- Wie lange haben Sie diesen Berufswunsch schon?
- Kennen Sie jemanden, der Psychologe/in ist?
- Was sind Ihrer Ansicht nach ihre Stärken und Schwächen, die Sie für die Tätigkeit als Psychologe/in mitbringen?
- Welche berufliche Tätigkeit könnten Sie sich noch vorstellen, sollten Sie den Berufswunsch Psychologe/in nicht realisieren können?
- Wissen Sie, wie die Studienbedingungen an der Universität sind?
- Welche Dauer planen Sie für das Studium ein?

Diese Fragen wurden größtenteils übernommen; Änderungen betrafen lediglich die Anredeform. Bei Frage drei wurde die Zusatzfrage „Wen kennst du, der Psychologe/in ist (z.B. Onkel, Freundin der Mutter)?“ gestellt. Außerdem wurde Frage sechs folgendermaßen umformuliert: „Hast du dich bereits mit den Inhalten des Psychologiestudiums vertraut gemacht?“

Die Fragen zur Studienwahlreife wurden deshalb bei der Datenerhebung berücksichtigt, da den Psychologieinteressenten bei der Teilnahme an dieser Diplomarbeit eine Rückmeldung über ihre Ergebnisse in Bezug auf das Psychologiestudium Wien angeboten wurde.

Seite drei des ersten Teils beinhaltet lediglich einen Teilnehmercode und die Frage, ob eine Rückmeldung bezüglich der Ergebnisse erwünscht ist bzw. an welche E-Mailadresse diese erfolgen sollte.

6.1.2 Teil 2: FMT

Der zweite Teil begann mit einer kurzen Instruktion: *Die folgenden 28 Aufgaben bestehen jeweils aus einer Matrix mit einem Fragezeichen in der Mitte. Deine Aufgabe ist es, diejenige Figur aus den acht möglichen Figuren auf der rechten Seite auszuwählen, die logisch sinnvoll an die Stelle des Fragezeichens gesetzt werden kann. Es gibt für jedes Beispiel nur eine Lösung. Es folgen nun zwei Übungsbeispiele.*

Bei den Übungsbeispielen (vgl. Abb. 14 und Abb. 15) wurde den Testperson die Herangehensweise an die Bearbeitung der folgenden Aufgaben erläutert. Hier bekamen die

Testpersonen noch eine Rückmeldung darüber, ob die angegebene Antwort richtige ist oder nicht. Abb. 14 zeigt Übungsaufgabe 1, bei welcher eine falsche Antwortmöglichkeit, zur Veranschaulichung, ausgewählt wurde. Die Testperson bekam hier die Rückmeldung: *Falsche Auswahl! Der Pfeil zeigt die Lösung an, wähle diese durch Anklicken aus...* Hingegen wurde in Abb. 15 die richtige Antwortmöglichkeit angegeben. Somit bekam die Testperson die Rückmeldung *Richtige Auswahl! Klicke auf das Dreieck um zur nächsten Aufgabe zu gelangen...*

Anschließend sollte die Testperson die 28 Aufgaben des FMT bearbeiten. Genauer zum FMT siehe Kapitel 4.4.

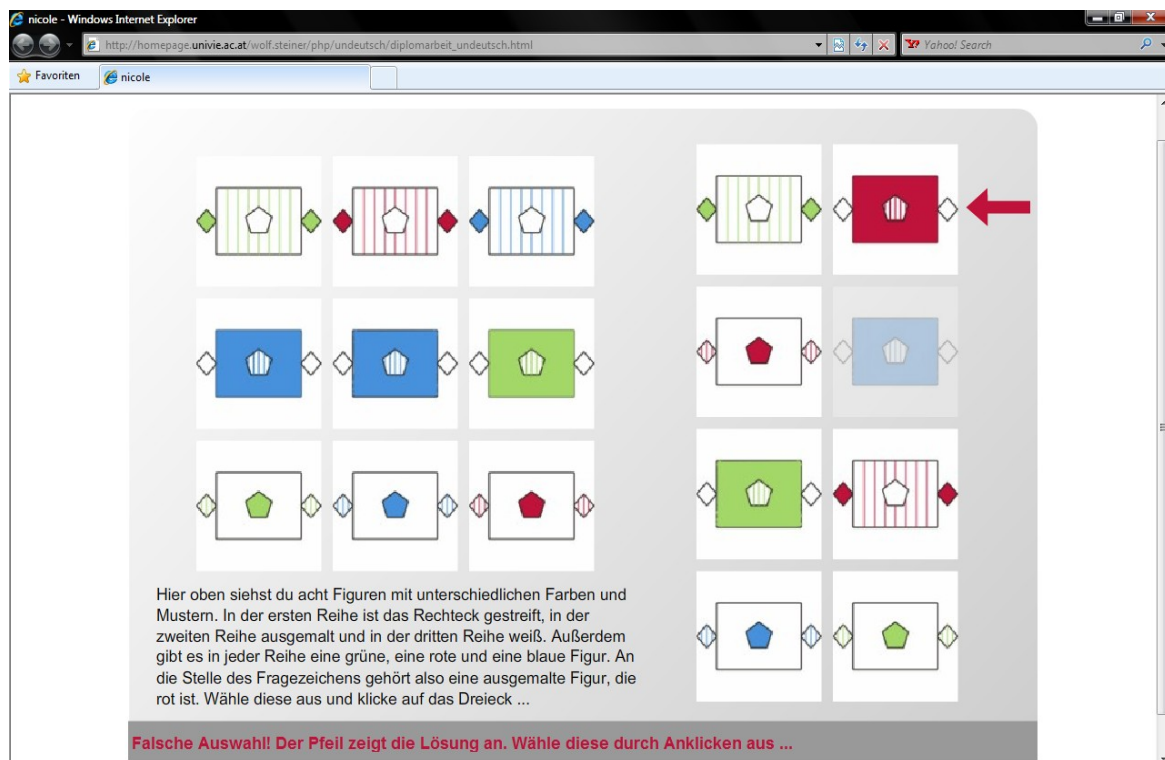


Abb. 14: Übungsbeispiel 1 des FMT



Abb. 15: Übungsbeispiel 2 des FMT

6.1.3 Teil 3: FIP

Der dritte Teil begann ebenfalls mit einer kurzen Instruktion: *Die folgenden Beispiele bestehen jeweils aus vier Aussagen. Deine Aufgabe ist es diejenige Aussage aus den vieren auszuwählen, die am besten für dich passt. Klicke auf das Dreieck, um zu einem Übungsbeispiel zu gelangen...*

Beim Übungsbeispiel des FIP (vgl. Abb. 16) wurde der Testperson ebenfalls die Herangehensweise an die Bearbeitung der folgenden Aussagen erläutert: *Wähle nun jene Alternative aus, die am meisten auf dich zutrifft. Du kannst immer nur eine Alternative auswählen. Mit dem Dreieck kommst du nun zu 45 Aussagen, die du auf die gleiche Weise beantworten sollst. Da es hier keine richtige bzw. keine falsche Antwort gibt wird der Testperson mit einem roten Kreuz in einem der vier grauen Kästchen auf der linken Seite lediglich signalisiert, welche Antwortmöglichkeit sie gewählt hat.*

Im Anschluss an die Übungsaufgabe soll die Testperson die 45 Aufgaben des FIP bearbeiten. Genauer zum FIP siehe Kapitel 5.4.

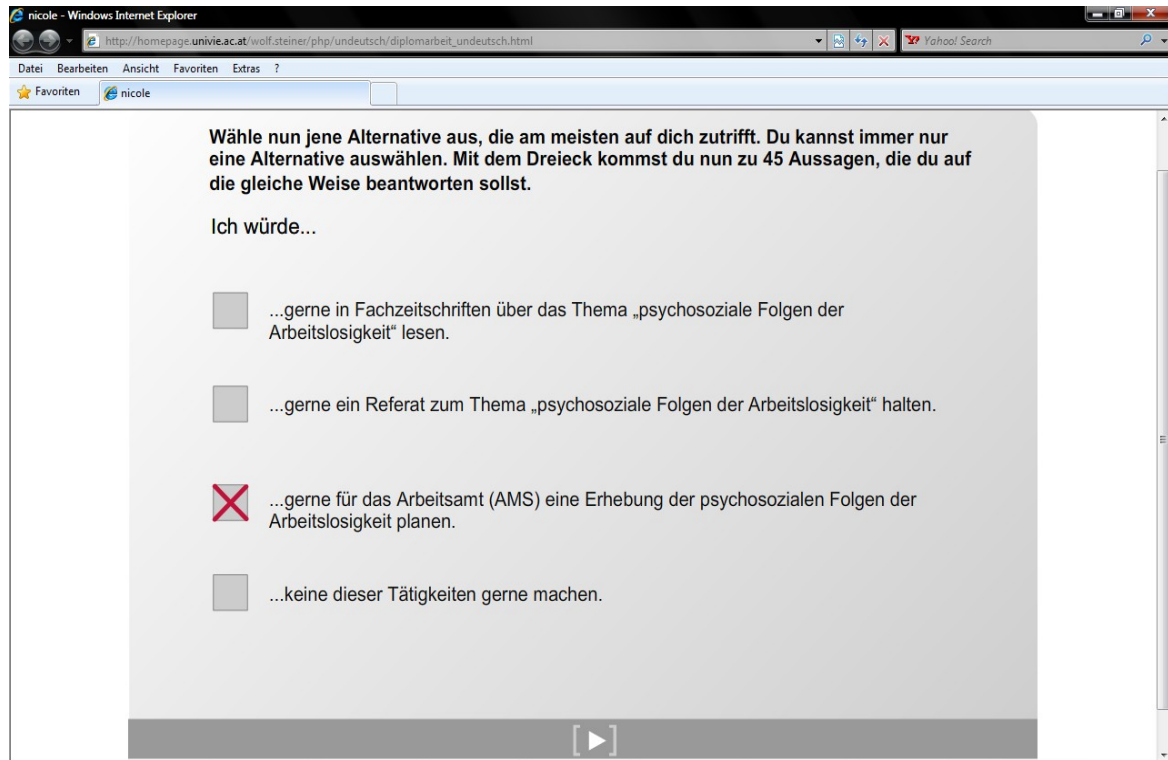


Abb. 16: Übungsbeispiel des FIP

6.2 Stichprobe

Die Datenerhebung fand im Zeitraum vom 29. Juli 2009 bis zum 7. November 2009 statt. Bei der Stichprobe handelt es sich um Interessenten des Psychologiestudiums Wien.

Als Stichprobe wurden deshalb Psychologieinteressenten ausgewählt, da Personen, die zukünftig das Self-Assessment Psychologie durchführen werden, ebenfalls dieser Population angehören. Diese nehmen freiwillig am Self-Assessment Psychologie Wien teil, da sie über ihre „Eignung“ für das Studium bescheid wissen wollen. Somit scheint eine Stichprobe von Psychologieinteressenten, welche ebenfalls freiwillig an der Datenerhebung teilnehmen, eine adäquate Referenzpopulation für zukünftige Self-Assessment Anwender zu sein.

Der erste Teil der Stichprobe bestand aus Interessenten, die auf der Berufsinformationsmesse – BeSt³ im März 2009 über meine Diplomarbeit und die Möglichkeit der Teilnahme an meinen Verfahren informiert wurden. Dies geschah entweder mündlich, durch einen der anwesenden Berater am Psychologiestand, oder sie bekamen einen Flugzettel mit allen Informationen und meiner E-Mail-Adresse. Da ich mich zu diesem Zeitpunkt noch in der Konstruktionsphase

befand, sammelte ich auf der BeSt³ E-Mail-Adressen, um die Interessenten nach der Programmierung und Fertigstellung der Verfahren kontaktieren zu können. Neben Listen, die auf dem Stand auflagen, in welche sich Interessenten eintragen konnten, gab es zudem die Möglichkeit mit einer E-Mail, das Interesse an der Teilnahme zu bekunden. Insgesamt zeigten 207 Personen, im März 2009, durch Bekanntmachung ihrer E-Mail Adressen, ihr Interesse an der Teilnahme meiner Erhebung.

Nach der Fertigstellung der diagnostischen Instrumentarien im Juli 2009 wurden alle 207 Interessenten per E-Mail informiert, dass die Verfahren unter dem angeführten link nun von zu Hause aus durchgeführt werden konnten. Nach etwa einem Monat hatten 29 der 207 Personen an der Erhebung teilgenommen. Am 23. August 2009 wurde ein erstes Erinnerungsmail an alle Personen gesandt. Da es sich um die Haupturlaubszeit handelte, wurde angenommen, dass noch einige Personen an der Erhebung teilnehmen würden.

Schließlich hatten bis zum 10. September 2009 jedoch nur 52 Interessenten an der Datenerhebung teilgenommen. Da es aber Ziel dieser Untersuchung war in etwa 200 Personen für die statistischen Analysen zu erfassen, musste auf weitere Methoden zur Stichprobenakquisierung zurückgegriffen werden. Ab diesem Zeitpunkt wurden in einem Onlineforum Personen angeschrieben, die geplant hatten im Wintersemester 09/10 das Psychologiestudium an der Universität Wien zu beginnen. Jedem Einzelnen in dieser Gruppe wurde eine persönliche Nachricht mit Informationen über die Teilnahme an meiner Diplomarbeit gesendet und die Option auf eine Rückmeldung über ihre Ergebnisse nach den statistischen Analysen angeboten. Auf diese Weise wurden im Zeitraum von 10. September 2009 bis zum 1. November 2009 134 Personen kontaktiert. Der Zeitraum von drei Wochen war deshalb so lange, da pro Tag nur maximal 20 Personen mittels persönlicher Nachricht kontaktiert werden konnten und da immer wieder neue Personen dem Forum beigetreten sind.

Da bis 22. September 2009 noch immer erst 67 Personen an der Datenerhebung teilgenommen hatten, wurden noch weitere Maßnahmen getroffen. Zuerst wurde eine dritte Erinnerungsmail an alle 207 Interessenten der BeSt³ gesendet. Im Weiteren wurden am 7. Oktober 2009 29 Psychologielehrer aus Wien via Mail kontaktiert und über meine Diplomarbeit informiert. Diese wurden gebeten Interessenten am Psychologiestudium aus ihren Klassen auf meine Diplomarbeit hinzuweisen und über die Möglichkeit einer Rückmeldung zu informieren.

Schließlich bestand die Stichprobe am 15. Oktober 2009 noch immer erst aus 87 Personen und war somit noch weit von der angestrebten Anzahl von 200 Personen entfernt, jedoch sollte mit Ende Oktober die Datenerhebung abgeschlossen werden. Als letzte Maßnahme wurden Flugzetteln an die Teilnehmer der Ringvorlesungsprüfung des Psychologiestudiums Wien, am

17. Oktober und am 24. Oktober 2009, verteilt. Zu diesem Zeitpunkt waren die Psychologieinteressenten zwar schon zum Psychologiestudium Wien inskribiert, jedoch mussten sie erst noch das Auswahlverfahren bestehen, um das Studium offiziell beginnen zu dürfen. Insgesamt wurden an diesen beiden Tagen mit Hilfe von Kollegen in etwa 700 Zettel ausgeteilt, auf welchen die Deadline von Ende Oktober bereits bekannt gegeben wurde. Es nahmen in diesem Zeitraum noch weitere 44 Interessenten an der Erhebung teil. Tatsächlich beteiligten sich 131 Psychologieinteressenten an der Erhebung meiner Diplomarbeit.

6.2.1 Demographische Beschreibung der Gesamtstichprobe

Von den 131 Testpersonen bearbeiteten nicht alle die Verfahren vollständig, da einzelne gegebenenfalls früher abbrachen. Der FMT, welcher zuerst vorgegeben wurde, wurde von 125 Personen bearbeitet, was einem Drop-out von sechs Personen (4,8%) entspricht. Von einer weiteren Person fehlte die Hälfte der Daten des FMT, aufgrund von technischen Problemen, in der Datenbank. Die Person wurde aber in der Stichprobe belassen, da sie trotzdem in die Analysen miteinbezogen werden konnte. Nach der Durchführung des FMT kam es zu weiteren Drop-outs, der FIP wurde schließlich nur von 118 Personen durchgeführt. Folglich bearbeiteten sieben der 125 Personen nur den FMT.

Für die demographische Beschreibung der Daten wurden die 125 Personen herangezogen, die zumindest ein Verfahren vollständig bearbeitet haben.

Hinsichtlich des **Geschlechts** setzte sich die Stichprobe aus 20 Männer und 105 Frauen zusammen (vgl. Tabelle 9).

Tabelle 9: Geschlechterverteilung

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig Männer	20	16,0	16,0	16,0
Frauen	105	84,0	84,0	100,0
Gesamt	125	100,0	100,0	

Das **Alter** der getesteten Personen lag zum Zeitpunkt der Erhebung zwischen 16 und 37 Jahren (vgl. Abb. 17). Der Median lag bei 19 Jahren mit einem Quartilabstand von 2 Jahren. Fünf Personen (4%) waren 16 Jahre alt, 12 Personen (9,6%) 17 Jahre, 29 Personen (23,2%) 18 Jahre, 32 Personen (25,6%) 19 Jahre, 18 Personen (14,4%) 20 Jahre, acht Personen (6,4%) 21 Jahre, zehn Personen (8%) 22 Jahre, zwei Personen (1,6%) 23 Jahre, eine Person (0,8%) 24 Jahre, drei

Personen (2,4%) 25 Jahre, eine Person (0,8%) 26 Jahre, zwei Personen (1,6%) 27 Jahre, eine Person (0,8%) 32 Jahre und eine Person (0,8%) war 37 Jahre alt.

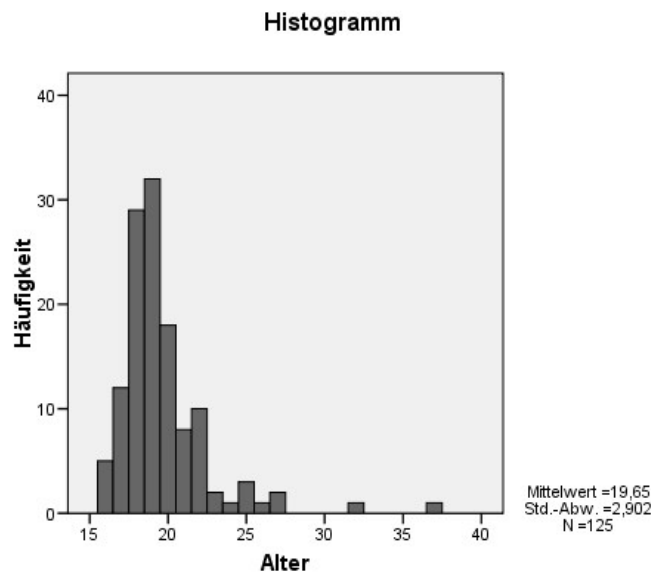


Abb. 17: Histogramm der Altersverteilung

Bezüglich der **Schulbildung** setzte sich die Stichprobe wie folgt zusammen: 73 Personen (58,4%) besuchten eine Allgemein bildende höhere Schule (AHS), 12 Personen (9,6%) eine Handelsakademie (HAK), 13 Personen (10,4%) eine Höhere Bundeslehranstalt (HBLA/HLTW) und sieben Personen (5,6%) besuchten eine Höhere technische Lehranstalt (HTL). 20 Personen (16%) gaben Sonstiges bei ihrer Schulbildung an. Hierzu zählten unter anderem Angaben wie: Berufsreifeprüfung, Bildungsanstalt für Kindergartenpädagogik (BAKIP), Fachabitur etc.

6.2.2 Beschreibung der Gesamtstichprobe in Bezug zur Studienwahlreife

Die Frage „**Was hat dich auf die Idee gebracht, Psychologe/in werden zu wollen?**“ wurde aufgrund des freien Antwortformates in acht Antwortkategorien zusammengefasst (vgl. Tab. 10). Die Kategorie «Erfahrungen» beinhaltet jene Antworten, aus denen hervorging, dass Testpersonen Erfahrungen mit psychisch auffälligen Personen gesammelt haben, dass sie selbst psychologische Hilfe in Anspruch genommen haben oder im psychosozialen Bereich tätig gewesen sind.

Die Kategorie «mit Inhalten beschäftigt» ergibt sich aus Antworten, bei denen ersichtlich wurde, dass Testpersonen in der Schule, durch Bücher oder durch den Kontakt mit Experten Informationen erhielten und sich dadurch mit psychologischen Themen auseinander gesetzt haben.

Die Kategorie «soziale Orientierung» beinhaltet Antworten, welche sich darauf beziehen, dass Testpersonen gerne anderen Menschen helfen, sich mit Problemen von anderen beschäftigen möchten oder generell gerne mit Menschen zusammen arbeiten möchten.

Die Kategorie «Interesse vorhanden» ergibt sich durch Antworten aus denen hervorgeht, dass bestehendes Interesse sie auf die Idee gebracht habe Psychologe zu werden.

In die Kategorie «Stärken vorhanden» wurden Antworten eingegliedert, deren Aussage darin bestand, dass Testpersonen aufgrund von Eigenschaften, wie Empathie oder die Fähigkeit gut mit Menschen umgehen zu können, auf die Idee kamen Psychologe zu werden.

Die Kategorie «angestrebter Beruf» beinhaltet jene Antworten, aus denen hervorging, dass der Beruf Psychologe angestrebt wird, weil dieser Voraussetzung für einen weiteren Beruf ist oder weil die Ausbildung von Vorteil für eine weitere Karriere wäre.

Die Kategorie «Wissen und Verständnis als Ziel» ergab sich aus jenen Antworten, welche sich darauf bezogen, dass der Antrieb für den Beruf daher kommt mehr über Menschen zu erfahren, ein Verständnis für sie zu bekommen und das Wissen auf diesem Gebiet zu erweitern.

Die Kategorie «Rest» ergibt sich schließlich aus allen übrigen Antworten, welche nicht zu einer der anderen Kategorien zugeordnet werden konnten wie zum Beispiel: *ich, mein Freund, alles* oder es wurde nur ein Bindestrich eingefügt etc.

Am häufigsten wurden die Kategorien «mit Inhalten beschäftigt» (29,6%), «Interesse vorhanden» (21,6%), «soziale Orientierung» (14,4%) und «Wissen und Verständnis als Ziel» (12,8%) genannt.

Tabelle 10: Häufigkeitsverteilung der kategorisierten Variable Idee

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Erfahrungen	6	4,8	4,8	4,8
	mit Inhalten beschäftigt	37	29,6	29,8	34,7
	soziale Orientierung	18	14,4	14,5	49,2
	Interesse vorhanden	27	21,6	21,8	71,0
	Stärken vorhanden	5	4,0	4,0	75,0
	angestrebter Beruf	5	4,0	4,0	79,0
	Wissen & Verständnis als Ziel	16	12,8	12,9	91,9
	Rest	10	8,0	8,1	100,0
	Gesamt	124	99,2	100,0	
Fehlend	99	1	,8		
Gesamt		125	100,0		

Die Frage „**Wie lange hast du diesen Berufswunsch schon?**“ wurde einheitlich in Tagen berechnet. Die mittlere Wunschkdauer entspricht 987,3 Tagen (ca. 2,7 Jahre) mit einer Standardabweichung von 1045,518 (ca. 2,8 Jahre) einem Minimum von einem Tag und einem Maximum von 7300 Tagen (entspricht 20 Jahren). Die Variable Wunschkdauer unterliegt einer sehr großen Streuung. Nach dem Median liegt die mittlere Wunschkdauer bei 730 Tagen, was einem Zeitraum von genau 2 Jahren entspricht.

In dem Zusammenhang interessierte, ob es Unterschiede zwischen den Geschlechtern in Bezug zur Wunschkdauer gibt. Es zeigte sich beim t -Test(123)=0,845, $p=0,400$, dass es keine signifikanten Unterschiede zwischen den beiden Geschlechtern in Bezug zur Wunschkdauer gibt.

Bei der Frage „**Kennst du jemanden, der Psychologe/in ist?**“ gaben 57 Personen (45,6%) an, dass sie jemanden kennen der Psychologe ist. Allen Personen, die die Frage mit ja beantworteten, wurde im Anschluss die Frage gestellt „Wen kennst du, der Psychologe/in ist (z.B. Onkel, Freundin der Mutter etc.)?“ Diese Variable wurde aufgrund des freien Antwortformates in fünf Kategorien unterteilt (vgl. Tab. 11). Mit 25 Nennungen (43,9%) wurde die Kategorie Bekannte/r am häufigsten genannt.

Tabelle 11: Häufigkeitsverteilung der kategorisierten Variable Wen

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozent
Gültig Familienmitglied	14	24,6	24,6	24,6
Freund/in	11	19,3	19,3	43,9
Bekannte/r	25	43,9	43,9	87,7
Lehrer/in	4	7,0	7,0	94,7
eigener Psychologe/in	3	5,3	5,3	100,0
Gesamt	57	100,0	100,0	

a psychologe = ja

„**Was sind deiner Ansicht nach deine Stärken und Schwächen, die du für die Tätigkeit als Psychologe/in mitbringst? Welche berufliche Tätigkeit könntest du dir noch vorstellen, solltest du den Berufswunsch Psychologe/in nicht realisieren können?**“ Die Fragen sind für die Studienwahlreife in diesem Sinne wichtig, da Testpersonen einerseits prinzipiell in der Lage sein sollten über ihre eigene Person zu reflektieren und andererseits sich Gedanken über alternative Ausbildungen zu machen.

Vier der 125 Personen (3,2%) machten keine Angaben bezüglich ihrer Stärken und Schwächen.

12 der 125 Personen (9,6%) machten zudem keine Angaben über alternative Berufswünsche oder nennen Berufe, die ebenfalls ein Psychologiestudium als Voraussetzung haben (z.B. Klinische

Psychologin). Die übrigen 113 Personen (90,4%) hatten entweder eine konkrete Vorstellung von einem alternativen Beruf oder sie gaben zumindest den Bereich an, in dem sie gegebenenfalls arbeiten wollen würden.

Die Frage „**Hast du dich bereits mit den Inhalten des Psychologiestudiums vertraut gemacht?**“ beantworteten 93 Personen (74,4%) mit ja und 32 Personen (25,6%) mit nein.

In dem Zusammenhang interessiert, ob es einen Unterschied zwischen Männern und Frauen bei der Beantwortung dieser Frage gibt. Mittels eines χ^2 Tests nach Pearson wurde festgestellt, dass es keinen signifikanten Unterschied zwischen den Geschlechtern gibt ($\chi^2=0,392$, $p=0,531$).

Die Frage „**Welche Dauer planst du für das Studium ein?**“ konnte in der Einheit von Semestern beantwortet werden. Bei der Analyse der Häufigkeiten bzw. der Betrachtung des Histogramms wurde deutlich, dass es einen Ausreißer bei der Beantwortung dieser Frage gab. Die Testperson gab an, dass sie 99 Semester für ihr Studium einplane, deshalb wird diese Person aus der deskriptiv Statistik dieser Frage ausgeschlossen und ihre Angabe als fehlender Wert interpretiert.

Demnach entspricht der Median der Dauer die für das Studium eingeplant wird 10 Semester mit einem Quartilabstand von 4 Semestern, einem Minimum von einem Semester und einem Maximum von 19 Semestern (vgl. Abb. 18).

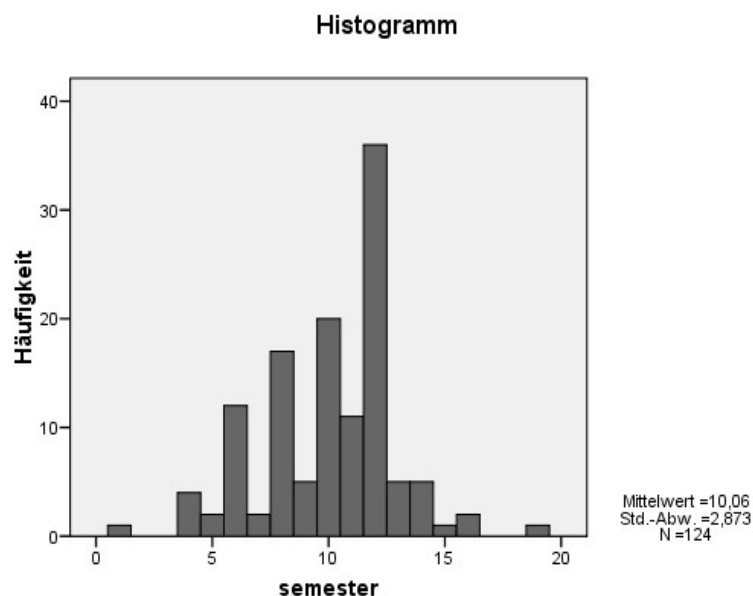


Abb. 18: Histogramm der geplanten Semesteranzahl

Interessant scheint in diesem Zusammenhang zu sein, ob es Unterschiede zwischen den Angaben derjenigen Testpersonen gibt, die zuvor angegeben haben sich über die Inhalte des Studiums informiert zu haben und jenen, die sich nicht informiert haben.

Ein t -Test für unabhängige Stichproben ergab dass es signifikante Unterschiede hinsichtlich der geplanten Semesteranzahl zwischen den Testpersonen gab, die angegeben haben sich nicht mit den Inhalten des Studiums auseinander gesetzt zu haben und jenen, die die Angabe gemacht haben, sich mit den Inhalten des Studiums auseinander gesetzt zu haben ($df=122$, $t=-3,886$, $p\leq 0,001$). Betrachtet man die Mittelwerte in Tab. 12, so scheinen „nicht Informierte“ tendenziell weniger Semester für ihr Studium einzuplanen als „Informierte“. Zudem liegt der Mittelwert der „nicht Informierten“ unter der Mindeststudienzeit.

Tabelle 12: Gruppenstatistik der Variable Semesteranzahl

Inhalte		N	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes
Semester	nein	31	8,42	3,394	,610
	ja	93	10,61	2,463	,255

7 Darstellung der Ergebnisse

Die folgenden Analysen wurden mittels der Software R (The R foundation for statistical computing, 2009) und dem Paket eRm (Mair & Hatzinger, 2007; Poinstingl, Mair & Hatzinger, 2007) durchgeführt. Weiters wurde bei den Analysen durchgängig ein Signifikanzniveau von $\alpha = .01$ angenommen.

7.1 Analysen zur Rasch-Modell Konformität des FMT

Zur Prüfung der Modellgeltung wurde der Likelihood-Ratio-Test nach Andersen (LRT) (vgl. Kapitel 3.1) herangezogen. Mittels Teilungskriterien sollte überprüft werden, ob die Itemparameter in einzelnen Teilstichproben besser geschätzt werden können, als in anderen. Gleich zu Beginn musste das Item a22 aufgrund eines festgestellten Konstruktionsfehlers von der Datenanalyse ausgeschlossen werden. Somit wurden 27 Items des FMT auf Modell-Konformität überprüft.

7.1.1 Kriterienbildung

Als Teilungskriterien fungierten die Variablen Score, Geschlecht, Schultyp und Alter.

- Score

Die Anzahl der gelösten Aufgaben (=Score) diene als internes Teilungskriterium. Die gültigen Fälle wurden hier anhand des Medians des Scores in eine Gruppe mit niedrigem Rohscores und in eine Gruppe mit hohem Rohscore geteilt. Das bedeutet, es wurden 62 Personen die wenige Items gelöst haben = «niedriger Median bzw. gleich dem Median», mit 62 Personen verglichen, die viele Items gelöst haben = «höher Median».

- Geschlecht

Das Geschlecht diene als externes Kriterium und sollte aufzeigen, ob es Unterschiede zwischen Frauen und Männern bei der Bearbeitung der Items des FMT gibt. Das Teilungskriterium Geschlecht bestand aus einer Gruppe mit 105 Frauen und einer Gruppe mit 20 Männern.

- Schultyp

Der besuchte Schultypus diente ebenfalls als externes Teilungskriterium. Die Variable Schulzweig bestand aus den Kategorien AHS, HAK, HBLA/HTLW, HTL und Sonstige. Um als Teilungskriterium zu fungieren wurden die Schultypen in die Kategorien «allgemeiner» und «spezifischer» (berufsbildender) Schultyp unterteilt. Da in der Kategorie Sonstige von sieben Personen angegeben wurde, dass diese ein Gymnasium oder ein deutsches Gymnasium bzw. Abitur absolvierten, wurden diese sieben Personen der allgemeinen Gruppe zugeteilt. Die restliche Kategorie Sonstige wurde der spezifischen Gruppe zugeteilt. Somit bestand das Teilungskriterium Schultyp aus einer Gruppe mit 73 Personen die eine „allgemein bildende höhere Schule“ besuchten und einer Gruppe mit 52 Personen, die angaben eine „berufsbildende höhere Schule“ zu besuchen.

- Alter

Auch das Alter wurde als externes Teilungskriterium herangezogen. Die Variable Alter wurde aufgrund des Medians in die Gruppe «bis 19 Jährige» und «ab 20 Jährige» eingeteilt. Folglich bestand das Teilungskriterium Alter aus einer Gruppe mit 78 Personen die angaben nicht älter als 19 Jahre zu sein und einer Gruppe mit 47 Personen, die angaben 20 Jahre oder älter zu sein.

7.1.2 LRT nach Andersen - FMT

Tabelle 13 zeigt die Ergebnisse der Überprüfung der Modell-Konformität der 27 Items des FMT. Bei den Analysen mit den Teilungskriterien Score und Alter wurde das Item a6 aufgrund von unangemessenem Antwortverhalten nicht in die Berechnungen miteinbezogen. Das bedeutet, dass von allen Personen einer Gruppe die gleiche Antwort gegeben wurde. Genauer betrachtet lies sich feststellen, dass beim Teilungskriterium Alter in der Gruppen «ab 20 Jährige» von allen 47 Personen das Item a6 gelöst wurde. Jedoch lösten auch 75 der 78 Personen in der Gruppe «bis 19 Jährige» das Item. Auch beim Teilungskriterium Score lösten alle 62 Personen der Gruppe «höher Median» das Item. Demzufolge schien das Item a6 zu leicht für die Stichprobe zu sein.

Von den vier Likelihood Ratio Tests nach Andersen erwiesen sich die externen Teilungskriterien als nicht signifikant (vgl. Tab. 13). Dies bedeutet, dass sich die Itemparameter in den Subpopulationen je externem Teilungskriterium nicht signifikant voneinander unterscheiden. Jedoch lieferte der LRT mit dem internen Teilungskriterium Median des Scores ein signifikantes

Ergebnis $\chi^2=60,07$ ($df=25$, $p=0,00$). Es wurde nun davon ausgegangen, dass eventuell Items in die Analysen miteinbezogen wurden, die einen niedrigen Item-fit⁸ besitzen. Um post-hoc die Gültigkeit des Rasch-Modells zu erzielen, sollten im Folgenden auffällige Items aus den Analysen ausgeschieden werden. Welche Items sich als auffällig erwiesen, wurde mittels z-Test nach Fischer und Scheiblechner und mittels graphischer Modellkontrolle (vgl. Abb. 19 & Abb. 20) analysiert. Beim z-Test nach Fischer und Scheiblechner zeigten sich signifikante Werte bei Item a5 ($z=2,021$, $p=0,043$), Item a7 ($z=3,623$, $p\leq 0,001$), Item a14 ($z=-3,094$, $p=0,002$), Item a21 ($z=-2,063$, $p=0,039$), Item a24 ($z=-1,965$, $p=0,049$) und bei Item a26 ($z=2,200$, $p=0,028$). Der grafische Modelltest wurde zur besseren Übersicht einmal für die Items a1 bis a15 (vgl. Abb. 19) und einmal für die Items a16 bis a27 (vgl. Abb. 20) durchgeführt. Hier können diese sechs Items ebenfalls als auffällig bezeichnet werden, da sie auch unter Einbeziehung des Schätzfehlers die 45° Gerade nicht berühren.

Tabelle 13: Ergebnisse LRT für FMT – kein Item ausgeschlossen

Kriterium	χ^2	$\chi^2_{\text{krit } \alpha=0,01}$	df	p	
Score	60,068	44,314	25	0,000	sign.
Geschlecht	34,239	45,642	26	0,129	nicht sign.
Schultyp	25,257	45,642	26	0,504	nicht sign.
Alter	35,821	44,314	25	0,074	nicht sign.

⁸ Der Item-fit spiegelt das Maß an unerwartetem Antwortverhalten wieder.

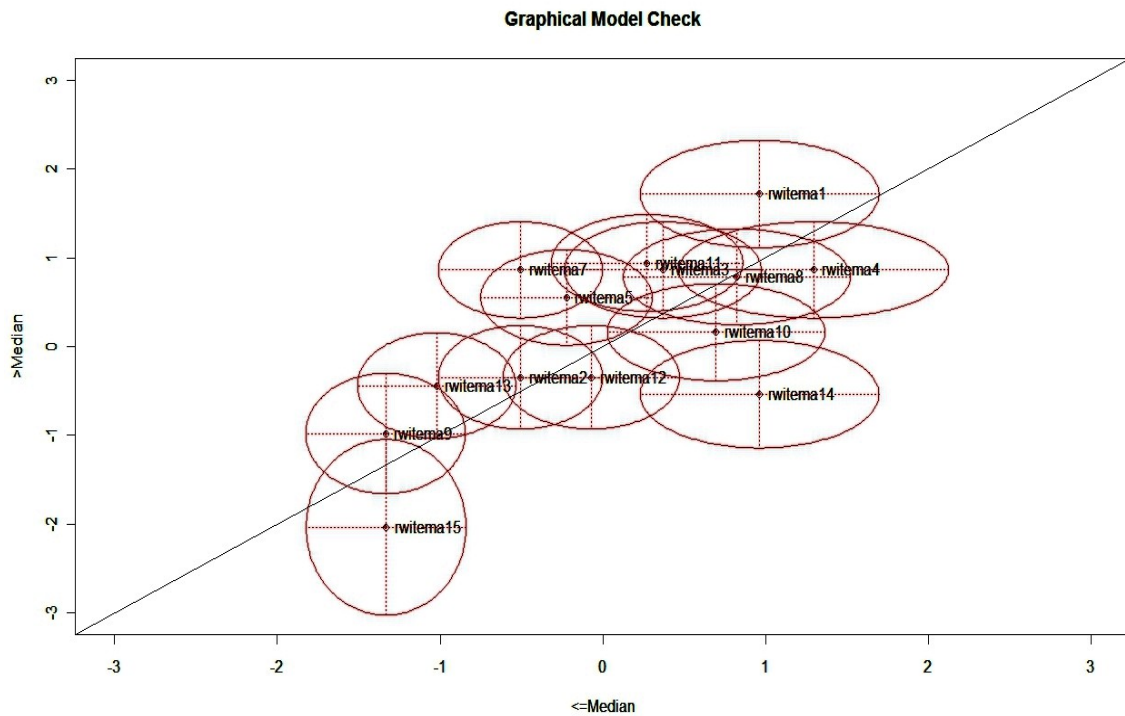


Abb. 19: Graphischer Modelltest FMT – Teilungskriterium Median, Items 1-15

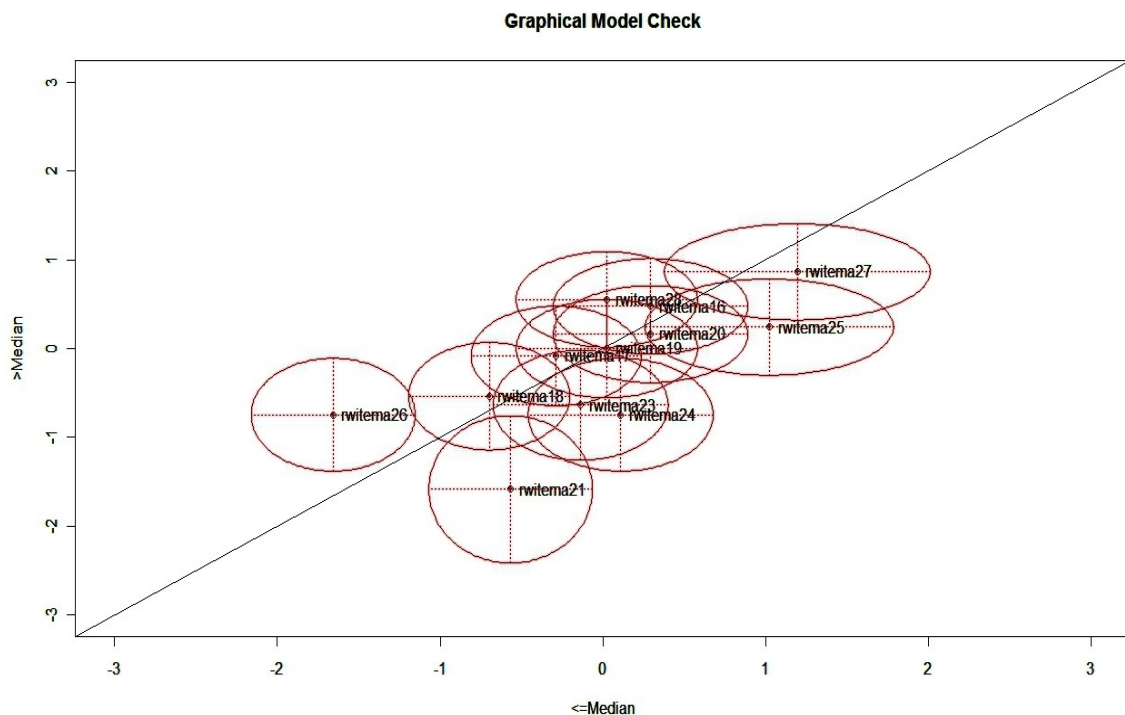


Abb. 20: Graphischer Modelltest FMT – Teilungskriterium Median, Items 16-27

Im Weiteren sollten nun die auffälligen Items näher analysiert werden und nach inhaltlichen Gesichtspunkten, wenn nötig, eliminiert werden, um post-hoc Modell-Konformität zu erlangen. Zuerst wurde analysiert, mit welchen Regeln diese auffälligen Items konstruiert wurden (vgl.

Tab. 14). Bei näherer Betrachtung der Konstruktionsregeln dieser Items fiel auf, dass alle der konstruierten Items, welche sowohl die Regel *Addition*⁹ als auch die Regel *Subtraktion* beinhalteten, sich unter den auffälligen Items befanden (vgl. Tab. 14). Genauer handelt es sich um die Items a5, a14 und a21. Folglich könnte es problematisch sein, diese zwei gegensätzlichen Operationen Addition und Subtraktion gleichzeitig anzuwenden. Im Weiteren zeigte sich jedoch ein Unterschied zwischen diesen drei Items. Bei Item a21 bezogen sich diese Regeln auf eindeutig unterscheidbare Elemente der Figur. Die Figur hatte hier die Form eines Käfers mit Punkten auf dem Rücken. Für die Punkte sollte die Regel *Subtraktion* senkrecht angewandt werden und für die Anzahl der Beine eine waagrechte *Addition*. Auch bei Item a5 kann man von unterscheidbaren Elementen der Figur sprechen. Diese Figur besteht aus einem Körper (Ellipse) mit Flügeln (zwei Dreiecke) oben und unten und einem Pfeil am Ende des Körpers. Für den oberen Flügel der Figur ist eine waagrechte *Addition* anzuwenden und für den unteren Flügel eine senkrechte *Subtraktion*. Zwar beziehen sich beide Regeln auf gleich aussehende Elemente, diese sind aber räumlich eindeutig durch den Körper (die Ellipse) getrennt. Item a14 hingegen scheint keine offensichtliche Trennung der beiden Elemente der Figur, auf welche die beiden Regeln angewandt werden sollen, zu beinhalten. Die Figur von Item a14 sieht aus wie ein Dominostein, der aus einer oberen und einer unteren Hälfte besteht. Für die obere Hälfte sollte eine senkrechte *Addition* und für die untere Hälfte eine waagrechte *Subtraktion* vorgenommen werden. Da aber die Trennung zwischen diesen beiden Teilen nur durch einen Strich erfolgt, könnte es Probleme bei der Differenzierung der beiden Teile gegeben haben.

Somit wurde entschieden von diesen drei Items vorerst nur **Item a14** auszuschneiden, da es im Gegensatz zu den anderen beiden Items die zwei gegensätzlichen Operationen *Addition* und *Subtraktion* wahrscheinlich nicht klar voneinander trennt.

Tabelle 14: Konstruktionsregeln der im z-Test auffälligen Items

Item	Identität	Addition	Subtraktion	Seriation	Variation o	Variation g	Id. Addition	w	s	w/s	SC	IC
Item a5		1	1		1	1		1	1		1	1
Item a7			1				1		1	1	2	
Item a14		1	1					1	1		2	
Item a21		1	1					1	1		2	
Item a24	1	1			2			2	2		2	2
Item a26					2	1		1	2		2	1

Bezüglich Item a7, Item a24 und Item a26 ließen sich keine Gemeinsamkeiten bezüglich der

⁹ Die Bedeutung der einzelnen Regeln kann in Kapitel 4.4.2 (S. 36) nochmals nachgelesen werden.

Konstruktion erkennen, welche sich als auffällig erweisen würde. Somit wurden alle drei Items nochmals einzeln auf Konstruktionsfehler oder andere Unstimmigkeiten bezüglich der Distraktoren oder der Lösung analysiert. **Item a7** (vgl. Abb. 21) war eines von drei Items die mit der Regel *Identifikation einer Addition* erstellt wurden. Jetzt läge die Vermutung nahe, dass eben diese Regel aufgrund ihrer Komplexität oder aufgrund des Einflusses von zusätzlichen kognitiven Faktoren nicht nur die Fähigkeit «reasoning» erhebt. Die Items a19 und a28 wurden ebenfalls mit dieser Regel konstruiert. Betrachtet man allerdings deren Werte im zuvor durchgeführten z-Test nach Fischer und Scheiblechner, so erwiesen sich weder Item a19 ($z=-0,062$, $p=0,950$) noch Item a28 ($z=1,336$, $p=0,182$) als auffällig (vgl. auch Abb. 20). Demzufolge schien die Problematik nicht an dieser Konstruktionsregel zu liegen. Es stellte sich jedoch bei weiterer eingehender Betrachtung des Items a7 heraus, dass einer der Distraktoren ebenfalls als logisch sinnvolle Lösung angegeben werden könnte.

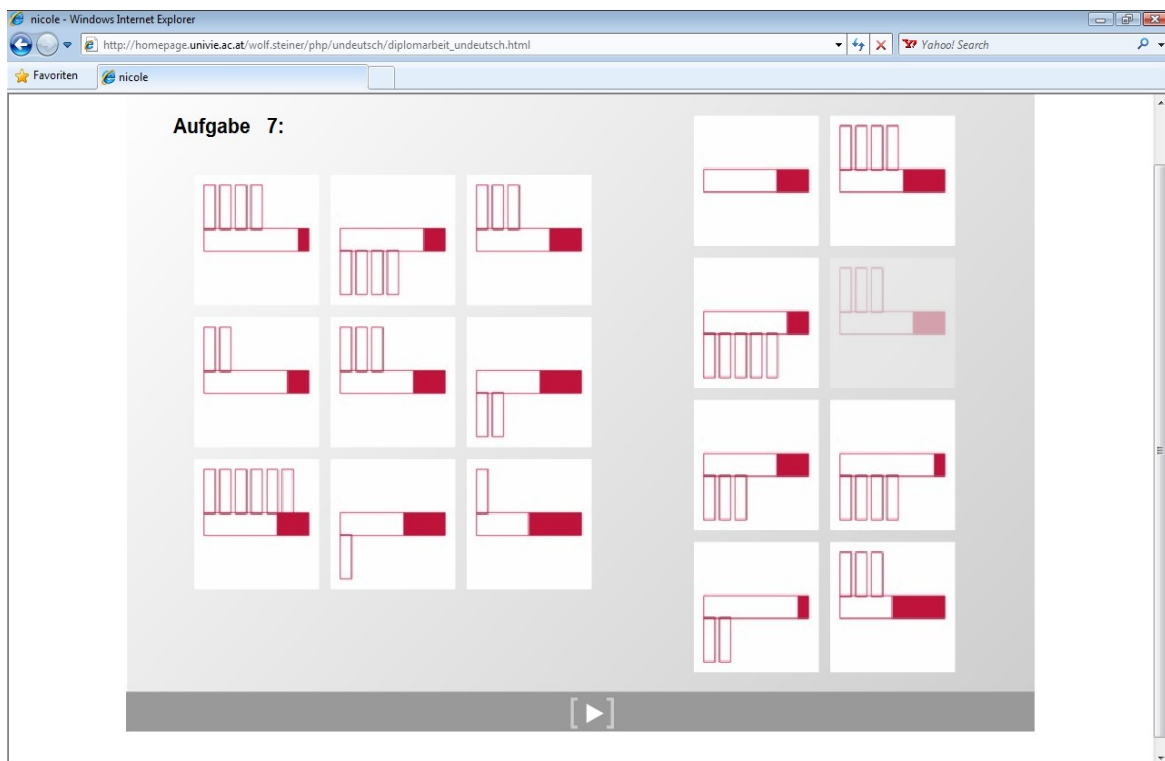


Abb. 21: Aufgabe 7 (Item a7) des FMT

Zur Erklärung: Betrachtet man das Item a7 in Abb. 21 so verändert sich im waagrechten Balken der rote Teil jeweils um eine Einheit, sowohl senkrecht als auch waagrecht (*Seriation*). Dies gilt für beide Lösungswege. Betrachtet man nun die senkrechten Balken, so sollte die Testperson schlussfolgern, dass Balken, die nach oben zeigen positiven Zahlen entsprechen und Balken, die nach unten zeigen, negativen Zahlen. Sodann ist mit diesen Zahlen senkrecht gesehen eine

Addition durchzuführen. Spalte eins: $4 + 2 = 6$; Spalte drei: $3 + (-2) = 1$ und somit würde sich für Spalte zwei ergeben: $(-4) + x = (-1)$. Die Lösung wäre somit 3, was sich durch drei nach oben zeigende Balken abbilden lässt. Diese Lösung ist in Abb. 21 bereits in der Mitte der Matrix integriert.

Es wäre jedoch noch eine andere Lösung denkbar, wenn man davon ausgeht, dass es unterschiedlich sein kann, ob man die zweite Zelle zur ersten Zelle addiert oder von ihr subtrahiert, um auf die dritte Zelle zu kommen. So könnten für die räumliche Anordnung der senkrechten Balken unterschiedliche Hypothesen aufgestellt werden. Einerseits könnte behauptet werden, dass in der ersten Spalte alle senkrechten Balken nach oben zeigen, in der zweiten alle nach unten und in der dritten abwechselnd, dies würde allerdings keiner der Regeln entsprechen, wäre jedoch für die Testperson trotzdem eine mögliche Lösung. Diese Vermutung wurde dahingehend überprüft indem die Häufigkeiten der genannten Lösungen betrachtet wurden (vgl. Tab. 15). Tatsächlich zeigte sich, dass 41 Personen (32,8%) die Antwortmöglichkeit mit den drei senkrechten Balken nach unten wählten und 51 Personen (40,8%) die richtige Antwortmöglichkeit mit den drei senkrechten Balken nach oben. Somit wurde das Item aufgrund Uneindeutigkeit der Lösung aus den weiteren Analysen ausgeschlossen.

Tabelle 15: Häufigkeiten der Antwortmöglichkeiten des Item a7

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig 1	14	11,2	11,2	11,2
2	1	,8	,8	12
3	9	7,2	7,2	19,2
4	51	40,8	40,8	60,0
5	41	32,8	32,8	92,8
6	1	,8	,8	93,6
7	7	5,6	5,6	99,2
8	1	,8	,8	100,0
Gesamt	125	100,0	100,0	

Bezüglich der Items a24 und a26 konnten keine weiteren Unstimmigkeiten aufgedeckt werden. Daher wurden für die weiteren Analysen bis auf weiteres lediglich die Items a7 und a14 ausgeschlossen.

Folglich wurden nun 25 der 27 zuvor analysierten Items erneut einer Rasch-Modell Analyse unterzogen. Tabelle 16 illustriert die χ^2 Werte der vier Likelihood Ratio Tests nach Andersen. Bei allen vier Teilungskriterien ergab sich ein nicht signifikantes Ergebnis. Es zeigten sich für die Teilungskriterien Score, Geschlecht, Schultyp und Alter, dass es zu keinen systematischen

Benachteiligungen zwischen den Gruppen kam. Somit können 24 Items als Rasch-Modell konform betrachtet werden. Item a6 kann wie bereits erwähnt lediglich bezüglich der Teilungskriterien Schultyp und Geschlecht als Modell konform betrachtet werden, da es aufgrund von unerwartetem Antwortverhalten nicht in die Analysen mit den Teilungskriterien Score und Alter miteinbezogen wurde.

Tabelle 16: Ergebnisse LRT für FMT – 25 Items (ohne a7 & a14)

Kriterium	χ^2	$\chi^2_{\text{krit } \alpha=0,01}$	df	p	
Score	37,566	41,638	23	0,028	nicht sign.
Geschlecht	29,965	42,980	24	0,186	nicht sign.
Schultyp	25,021	42,980	24	0,405	nicht sign.
Alter	25,675	41,638	23	0,316	nicht sign.

7.2 Analysen zur Rasch-Modell Konformität des FIP

Der FIP wurde im Sinne einer vierstufigen Rating-Skala entwickelt (vgl. Kapitel 5.4) und sollte mit dem Partial Credit Model – PCM nach Masters (1982) analysiert werden. Mit dem PCM sollte neben der Prüfung der Eindimensionalität ebenfalls das Interessenkonstrukt sensu Brickenkamp (vgl. Kapitel 5.3) belegt werden. Jedoch lässt die geringe Stichprobengröße keine präzisen und stabilen Itemschwellenparameterschätzungen zu. Somit wurde auf die Anwendung des PCM verzichtet und der FIP wurde im Weiteren mittels Rasch-Modell analysiert.

Es sollte festgestellt werden, ob mit dem FIP tatsächlich eine Dimension erhoben wird. Hierfür wurde das vierkategoriale Antwortformat dichotomisiert. Von der Annahme ausgehend, dass mittleres Interesse (reproduktiv) und hohes Interesse (kreativ) als positiv für das Psychologiestudium bewertet werden konnte, wurden diese Antworten mit 1 kodiert. Kein Interesse und geringes Interesse (rezeptiv) wurden demzufolge mit 0 kodiert.

Da der FIP, wie in Kapitel 5.4 erwähnt, aus drei Skalen (Psychologie Wien, Psychologie Klagenfurt und typische Irrtümer bzw. falsche Erwartungen) besteht und jede der drei Skalen eine andere Dimension eindimensional erfassen sollte, wurden alle drei Skalen einzeln auf Modell-Konformität überprüft.

7.2.1 LRT nach Andersen – FIP: Items Psychologie Wien

Die 35 Items des FIP, welche das Interesse für das Psychologiestudium Wien erheben sollen,

wurden zur Prüfung der Modellgeltung vier Likelihood Ratio Tests nach Andersen unterzogen (vgl. Tab. 17). Bei jedem der vier Teilungskriterien befand sich der χ^2 - Wert unter dem kritischen χ^2 - Wert, daher war kein Ergebnis signifikant. Folglich konnte angenommen werden, dass unter der Berücksichtigung der Teilungskriterien Median, Geschlecht, Schultyp und Alter das Konstrukt eindimensional erfasst wird. Somit konnte für 34 Items Modell Konformität angenommen werden. Einzig beim Teilungskriterium Geschlecht konnte das Item b2 nicht in die Analysen miteinbezogen werden, da in einer Gruppe nur ein Antwortverhalten zum Tragen kam. Explizit wählten alle Männer bei Item b2 die reproduktive oder kreative Antwortmöglichkeit. Für Item b2 konnten demzufolge noch keine endgültigen Aussagen getroffen werden. Betrachtet man allerdings die grafischen Modelltests (vgl. Abb. 22), so scheint sich Item b2 etwas von den anderen Items abzuheben.

Tabelle 17: Ergebnisse LRT für FIP – Items Psychologie Wien

Kriterium	χ^2	$\chi^2_{\text{krit } \alpha=0,01}$	df	p	
Score	37,057	56,061	34	0,330	nicht sign.
Geschlecht	27,996	54,776	33	0,715	nicht sign.
Schultyp	22,452	54,776	34	0,935	nicht sign.
Alter	35,596	56,061	34	0,393	nicht sign.

7.2.2 LRT nach Andersen – FIP: Items Inhalte aus Klagenfurt

Auch die fünf Items der Skala Klagenfurt, welche Inhalte aus dem Studium in Klagenfurt erheben, sollten eindimensional messen. Folglich wurden auch für diese Items Likelihood Ratio Tests nach Andersen durchgeführt, um aufzudecken, ob in einzelnen Teilstichproben die Itemparameter besser geschätzt werden können, als in anderen.

Alle vier LRT (vgl. Tab. 18) erwiesen sich als nicht signifikant. Demnach konnte angenommen werden, dass das Konstrukt in Bezug auf die Teilungskriterien Median, Geschlecht, Schultyp und Alter eindimensional misst. Somit können alle 5 Items als Rasch-homogen bezeichnet werden.

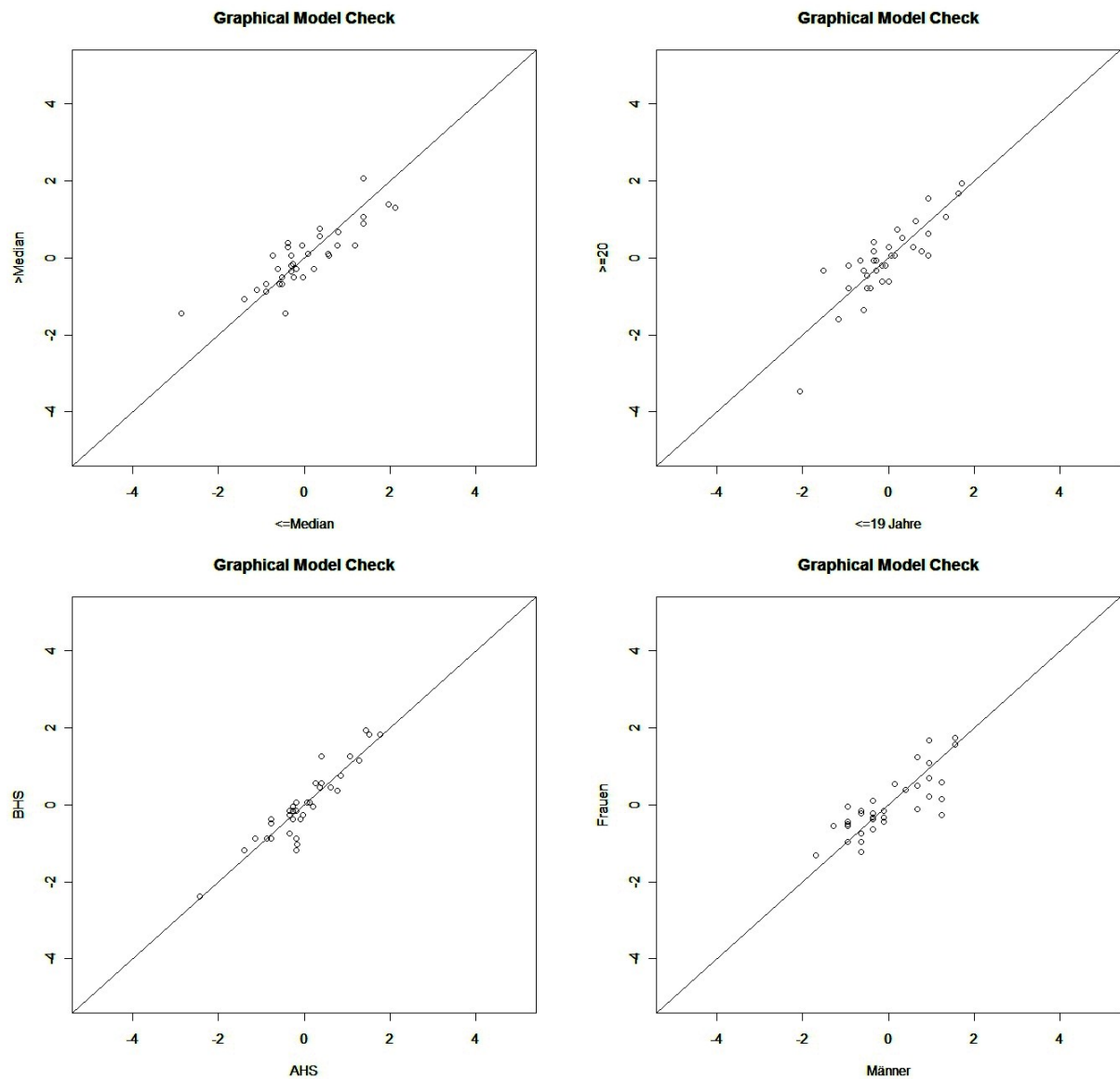


Abb. 22: grafische Modelltests der Items für Psychologie Wien

Tabelle 18: Ergebnisse LRT für FIP – Items Klagenfurt

Kriterium	χ^2	$\chi^2_{\text{krit } \alpha=0,01}$	df	p	
Score	1,870	13,277	4	0,760	nicht sign.
Geschlecht	4,6	13,277	4	0,331	nicht sign.
Schultyp	0,298	13,277	4	0,990	nicht sign.
Alter	0.98	13,277	4	0,913	nicht sign.

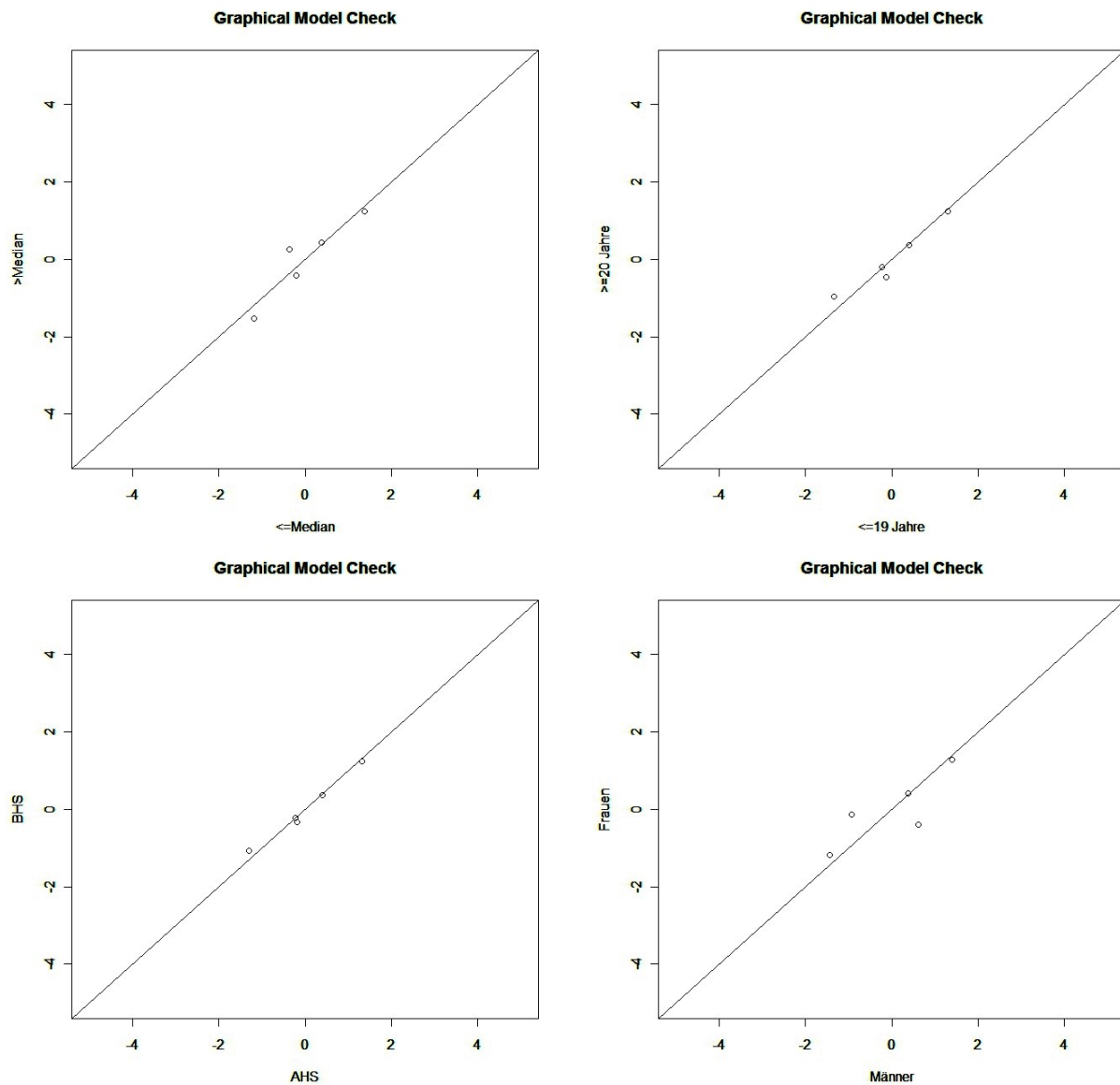


Abb. 23: grafische Modelltests der 5 Items für Psychologie Klagenfurt

7.2.3 LRT nach Andersen – FIP: Items typische Irrtümer/falsche Erwartungen

Im Weiteren sollten die fünf Items der Skala typische Irrtümer/falsche Erwartungen ebenfalls eine eigene eindimensionale Skala abbilden. Die vier χ^2 - Werte der LRTs nach Andersen (vgl. Tab. 19) lagen alle unter den kritischen χ^2 - Werten. Somit konnte die H_0 , dass die Items dem Rasch-Modell entsprechen, beibehalten werden. Dies spiegelt sich ebenfalls in den grafischen Modelltests wider (vgl. Abb. 24).

Tabelle 19: Ergebnisse LRT für FIP – Items typische Irrtümer/falsche Erwartungen

Kriterium	χ^2	$\chi^2_{\text{krit } \alpha=0,01}$	df	p	
Score	3,671	13,277	4	0,452	nicht sign.
Geschlecht	1,704	13,277	4	0,790	nicht sign.
Schultyp	3,870	13,277	4	0,424	nicht sign.
Alter	0,961	13,277	4	0,916	nicht sign.

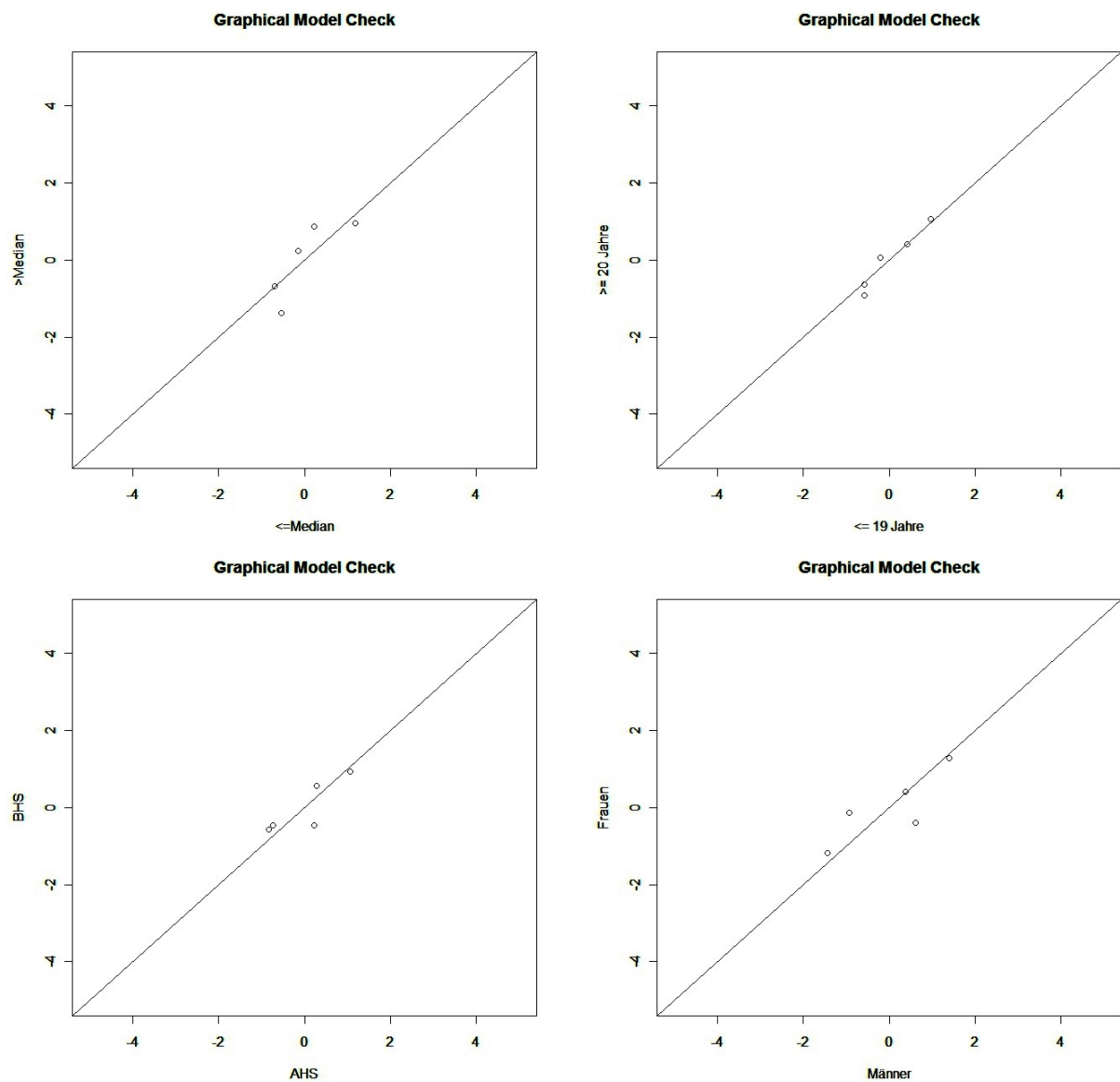


Abb. 24: grafische Modelltests der 5 Items für typische Irrtümer/falsche Erwartungen

7.3 Erstellung einer vorläufigen Eich­tabelle für Psychologieinteressenten

Da für die nun sowohl für den FMT als auch für den FIP gezeigt wurde, dass die Items Rasch-Modell-konform sind wurde in einem nächsten Schritt für die Rückmeldung an die Psychologieinteressenten eine vorläufige Eich­tabelle für Psychologieinteressenten erstellt.

Die Erstellung von Eich­tabellen erfolgt in der Item Response Theory anhand der Fähigkeitsparameter der Personen. Daher wurden zuerst die Fähigkeitsparameter aller Personen aufgelistet und anschließend mittels Häufigkeiten die kumulativen Prozente bestimmt.

Diese Eich­tabellen für den FMT und den FIP können in Anhang 2 und 3 betrachtet werden.

7.4 Analysen des FMT nach dem LLTM

Weiters sollte der FMT, genauer das Konstruktionsrational mit welchem die Items konstruiert wurden mittels LLTM (vgl. Kapitel 3.2) überprüft werden. Die Voraussetzung, dass das Rasch-Modell gilt, war den Analysen aus Kapitel 7.1. zufolge erfüllt. Daher sollte geprüft werden, ob die angenommen Basisparameter mittels LLTM die Itemschwierigkeiten genau so gut erklären wie das Rasch-Modell.

7.4.1 Überprüfung des Konstruktionsrational

Das Konstruktionsrational, auf welchem die Items des FMT konstruiert wurden, wurde bereits in Kapitel 4. ausführlich dargestellt. Ebenso entspricht die Tabelle 3 auf S. 44 bereits der Gewichtszahlen Matrix (=Q-Matrix) der einzelnen Basisparameter.

H_{0A}: Die Itemschwierigkeiten der 24 Items können durch die 12 Basisparameter ebenso gut erklärt werden, wie durch das Rasch-Modell. $\beta_i^{(LLTM)} = \beta_i^{(Rasch)}$

H_{1A}: Die Itemschwierigkeiten der 24 Items werden durch 24 unabhängige Itemparameter erklärt.

Tabelle 20: Likelihood-Quotienten Test A

$\log L_{LLTM}$	$\log L_{RM}$	χ^2	$\chi^2_{krit \alpha=0,01}$	df	
-1446,278	-1359,980	172,596	24,725	11	signifikant

Das Ergebnis des Likelihood-Quotienten Test (LQT) bezogen auf die Hypothese A war signifikant und ergab somit keine Modellpassung des LLTM (vgl. Tab. 20). Daraus lässt sich schließen, dass die 12 Basisparameter die Itemschwierigkeiten nicht genauso gut erklären wie die Itemparameter im Rasch-Modell.

Anschließend sollte in Bezug auf die Literatur die Gewichtsmatrix überarbeitet werden. Sowohl Formann (1973) als auch Hornke und Rettig (1989) zeigten, dass die Regel Identität eine geringere Schwierigkeit aufwies als die anderen Regeln. Ebenso zeigten Hornke und Rettig (1989), dass die Richtung der Regelgeltung senkrecht schwieriger sei als waagrecht. Zusätzlich besteht der Verdacht, welcher sich durch Vortestungen zeigte, dass Regeln, welche sich auf die Materialeigenschaft Farbe beziehen leichter zu erkennen seien als die übrigen Regeln. Somit wurde erneut eine Gewichtsmatrix erstellt (siehe Tabelle 22, S. 100), bei welcher diese drei Annahmen berücksichtigt wurden und folglich ein weiteres LLTM gerechnet.

H_{0B} : Die Itemschwierigkeiten der 24 Items können durch die 12 Basisparameter ebenso gut erklärt werden, wie durch das Rasch-Modell. $\beta_i^{(LLTM)} = \beta_i^{(Rasch)}$

H_{1B} : Die Itemschwierigkeiten der 24 Items werden durch 24 unabhängige Itemparameter erklärt.

Tabelle 21: Likelihood-Quotienten Test B

logL_{LLTM}	logL_{RM}	χ^2	$\chi^2_{krit \alpha=0,01}$	df	
-1438,016	-1359,980	156,072	24,725	11	signifikant

Beim LQT B lag der empirische χ^2 Wert über dem $\chi^2_{krit \alpha=0,01}$, somit war das Ergebnis auch hier signifikant (vgl. Tab. 21). Folglich können auch diese 12 Basisparameter die Itemschwierigkeiten nicht genau so gut erklären wie die Itemparameter im Rasch-Modell.

Tabelle 22: Q-Matrix der Hypothese B

Itema	Identität	Addition	Subtraktion	Seriation	Variation o	Variation g	Ident. Addition	w	s	w/s	SC	IC
Itema1	1	2	0	0	0	2	0	1	4	0	3	0
Itema2	0	0	2	0	1	0	0	1	2	0	0	2
Itema3	0	2	0	0	0	4	0	1	4	0	3	0
Itema4	1	0	2	0	0	3	0	2	2	1	1	3
Itema5	0	2	2	0	2	1	0	3	2	0	2	2
Itema8	0	0	2	0	0	4	0	2	2	0	3	0
Itema9	0	0	0	4	0	0	0	0	0	2	0	2
Itema10	0	2	0	0	0	4	0	2	2	0	1	2
Itema11	0	0	0	2	0	2	0	1	2	0	0	2
Itema12	0	2	0	0	0	1	0	1	2	0	2	0
Itema13	0,5	0	2	0	0	4	0	2	4	0	2	2
Itema15	0	0	0	0	1	2	0	1	2	0	2	0
Itema16	0	2	0	2	0	2	0	2	2	0	1	2
Itema17	0	2	0	0	0	3	0	2	2	0	0	3
Itema18	0	0	2	0	0	1	0	1	2	0	2	0
Itema19	0	0	0	0	0	1	2	1	2	0	2	0
Itema20	0	2	2	0	0	0	0	1	2	0	2	0
Itema21	0,5	0	2	0	2	2	0	2	4	0	4	0
Itema23	0	0	0	2	3	0	0	1	4	0	2	1
Itema24	0	0	0	0	3	1	0	1	4	0	2	1
Itema25	0	2	0	0	1	2	0	1	4	0	3	0
Itema26	1	0	2	0	0	1	0	2	2	0	1	2
Itema27	0	2	0	0	1	1	0	1	4	0	0	3
Itema28	1	0	0	0	1	0	2	1	4	0	1	2

7.4.1 Überprüfung des Konstruktionsrationalen mit Bezug zu Formann (1973)

Offensichtlich tragen neben den angenommenen Basisparameter weitere im Modell nicht berücksichtigte Faktoren zur Itemschwierigkeit bei. aufgrund dessen wurde eine weitere Gewichtsahlen Matrix erstellt, welche in Anlehnung an Formann (1979) den Faktor Materialeigenschaft mit einbeziehen sollte. Dabei wurde die ursprüngliche Q-Matrix um die Komponenten Form, Muster, Farbe, Anzahl und räumliche Anordnung erweitert (siehe Tabelle 24 S. 102).

H_{0C} : Die Itemschwierigkeiten der 24 Items können durch die 17 Basisparameter ebenso gut erklärt werden, wie durch das Rasch-Modell. $\beta_i^{(LLTM)} = \beta_i^{(Rasch)}$

H_{1C} : Die Itemschwierigkeiten der 24 Items werden durch 24 unabhängige Itemparameter erklärt.

Tabelle 23: Likelihood-Quotienten Test C

logL_{LLTM}	logL_{RM}	χ^2	$\chi^2_{krit \alpha=0,01}$	df	
-1411,307	-1359,980	102,654	16,812	11	signifikant

Die kritische Grenze des χ^2 - Wertes wurde überschritten womit sich der LQT C ebenfalls als signifikant erwies (vgl. Tab. 22). Folglich können auch diese 17 Basisparameter die Itemschwierigkeiten nicht genauso gut erklären, wie dies die Itemparameter im Rasch-Modell tun. Die Annahme über die 17 Basisparameter muss somit verworfen werden.

Tabelle 24: Q-Matrix der Hypothese C

Item	Identität	Addition	Subtraktion	Seriation	Variation o	Variation g	Ident. Addition	w	s	w/s	SC	IC	Form	Muster	Farbe	Anzahl	Räumliche Anordnung
Itema1	1	1	0	0	0	1	0	1	2	0	3	0	1	1	0	1	0
Itema2	0	0	1	0	1	0	0	1	1	0	0	2	1	0	1	0	0
Itema3	0	1	0	0	0	2	0	1	2	0	3	0	1	1	0	1	0
Itema4	1	0	1	0	0	2	0	2	1	1	1	3	1	1	1	1	0
Itema5	0	1	1	0	1	1	0	3	1	0	2	2	0	2	1	0	1
Itema8	0	0	1	0	0	2	0	2	1	0	3	0	2	0	0	1	0
Itema9	0	0	0	2	0	0	0	0	0	2	0	2	0	0	0	0	2
Itema10	0	1	0	0	0	2	0	2	1	0	1	2	2	1	0	0	0
Itema11	0	0	0	1	0	1	0	1	1	0	0	2	1	1	0	0	0
Itema12	0	1	0	0	0	1	0	1	1	0	2	0	1	0	1	0	0
Itema13	1	0	1	0	0	2	0	2	2	0	2	2	1	1	1	1	0
Itema15	0	0	0	0	1	1	0	1	1	0	2	0	1	0	1	0	0
Itema16	0	1	0	1	0	1	0	2	1	0	1	2	1	1	0	0	1
Itema17	0	1	0	0	0	2	0	2	1	0	0	3	1	1	1	0	0
Itema18	0	0	1	0	0	1	0	1	1	0	2	0	0	1	1	0	0
Itema19	0	0	0	0	0	1	1	1	1	0	2	0	0	0	1	1	1
Itema20	0	1	1	0	0	0	0	1	1	0	2	0	0	0	0	2	0
Itema21	1	0	1	0	1	1	0	2	2	0	4	0	1	0	1	2	0
Itema23	0	0	0	1	2	0	0	1	2	0	2	1	0	0	2	0	1
Itema24	0	0	0	0	2	1	0	1	2	0	2	1	0	0	2	1	0
Itema25	0	1	0	0	1	1	0	1	2	0	3	0	1	0	1	1	0
Itema26	1	0	1	0	0	1	0	2	1	0	1	2	1	1	1	0	0
Itema27	0	1	0	0	1	1	0	1	2	0	0	3	0	1	2	0	0
Itema28	1	0	0	0	1	0	1	1	2	0	1	2	1	0	1	1	1

8 Interpretation und Diskussion

Die Ergebnisse dieser Untersuchung zeigten, dass das grundlegende Ziel dieser Arbeit, nämlich die Entwicklung zweier Verfahren, welche nachweislich eindimensional messen, erfüllt wurde.

Beim Färbigen Matrizentest konnte nach dem Ausschluss von drei Items post-hoc Modell-Konformität für 24 Items angenommen werden. Von den ursprünglich 28 Items musste Item a22 aufgrund eines Konstruktionsfehlers bereits vor den Analysen ausgeschieden werden. Nach dem signifikanten Likelihood-Ratio Test von Andersen mit dem Teilungskriterium Score (Median) wurden im z-Test nach Fischer und Scheiblechner sechs Items signifikant. Anhand einer inhaltlichen Analyse konnte gezeigt werden, dass Item a6 zu leicht für diese Stichprobe war, Item a7 keine eindeutige Lösung hatte und Item a14 keine klare Trennung zwischen zwei gegensätzlichen Regeln vornahm. Schließlich wurden, nach dem Ausscheiden der drei Items, erneut Likelihood-Ratio Test von Andersen gerechnet, die nicht signifikant wurden. Kritisch anzumerken ist hier, dass vier der zuvor signifikanten Items (a5, a21, a24, a26) weiterhin im FMT verbleiben. Daher wäre hier für die Zukunft eine Studie zur Kreuzvalidierung angebracht um nachzuweisen, dass die Ergebnisse replizierbar sind.

Die Konstruktionsprinzipien des FMT stellten im Weiteren schwierigkeitsgenerierende Faktoren da, welche die kognitiven Anforderungen an die Testpersonen bei der Bearbeitung des FMT erklären sollten. Diese Konstruktionsprinzipien wurden teilweise aus der Literatur entnommen (vgl. Hornke & Habon 1984a) dies legte somit von Anfang an die Vermutung nahe, dass die zugrunde liegenden kognitiven Operationen mittels LLTM nachgewiesen werden können. Jedoch konnten diese Annahmen nicht bestätigt werden. Nach der Anpassung an die Gewichtung von einzelnen Faktoren (Formann, 1973; Hornke & Habon, 1989) bzw. nach der Integration von weiteren Faktoren in Anlehnung an Formann (1973) wurde der Likelihood Quotienten-Test signifikant. Somit kann man schlussfolgern, dass das Rasch-Modell die Itemschwierigkeiten besser erklärt als die angenommenen Basisparameter. Somit stellt sich die Frage, warum diese Operationen, welche teilweise bereits bei Hornke und Habon (1984a) vorkamen und dort nachweislich schwierigkeitsgenerierend wirken, hier nicht nachgewiesen werden konnten. Es liegt die Vermutung nahe, dass die Materialeigenschaft Farbe einen großen Einfluss auf die Bearbeitung der Items hat. Der Versuch diesen zu gewichten, indem Regeln welche sich auf die Materialeigenschaft Farbe beziehen nur halb gewichtet wurden, scheiterte allerdings. Ob die gestaltnahe Konstruktion der Items eventuell einen Einfluss auf die Lösung der Items hat, könnte durch eine weitere Studie, bei welcher das Lösungsverhalten anhand der Methode des lauten

Denkens untersucht wird, geklärt werden.

Beim Fragebogen zum Interesse am Psychologiestudium wurde jede einzelne Skala auf Eindimensionalität überprüft. Es konnte sowohl für die Items der Skala Psychologie Wien, die Items der Skala Psychologie Klagenfurt als auch für die Items der Skala typische Irrtümer/falsche Erwartungen Rasch-Modell-Konformität angenommen werden. Einzig Item b2 (siehe weiter unten) der Skala Psychologie Wien konnte beim Teilungskriterium Geschlecht nicht in die Analysen miteinbezogen werden, da alle Männer eine reproduktive oder kreative Antwortmöglichkeiten wählten, sowie 92 der 99 Frauen. Diese Tatsache könnte eventuell an der Formulierung dieses Items liegen. Möglicherweise ist die rezeptive Tätigkeit einfach nicht attraktiv genug, oder es ist ein Item, welches für alle Psychologieinteressenten so interessant ist, dass auch zukünftige Anwender nur die reproduktive oder die kreative Antwortmöglichkeiten wählen.

Item b2: Ich würde...

...gerne das Buch „Experimente planen“ lesen.

...gerne an einer Diskussion über die Planung von Experimenten teilnehmen.

...gerne mit einer kleinen Gruppe ein Experiment planen und durchführen.

Somit wurde beim FIP zwar Rasch-Modell-Konformität der Items erwiesen, jedoch bezüglich der zugrunde liegenden Theorie von Brickenkamp (1990) kann lediglich behauptet werden, dass sie für die Konstruktion von Items geeignet ist. Die Abstufung der Entwicklungsstufen von Interesse sollte jedoch im Weiteren an einer geeigneten und großen Stichprobe mittels Partial Credit Model überprüft werden. Erst diese Analysen können die Annahmen, welche hier getätigt wurden, bestätigen.

Durch die deskriptiven Fragen und die Fragen zur Studienwahlreife konnte die Stichprobe der Psychologieinteressenten näher beschrieben werden. Für diese Arbeit selbst hat dies neben den Analysen zudem eine Bedeutung für die Rückmeldung an die Psychologieinteressenten. Für zukünftige Arbeiten allerdings könnte dies eine Richtung vorzeigen, welche Variablen interessant bzw. relevant erscheinen, wenn es um die Typisierung von Psychologieinteressenten geht. Vorzustellen wäre dies mit Bezug zum FIP, anhand dessen mit Einbeziehung der Studienwahlreife-Fragen eine Typologie von Psychologieinteressenten erstellt werden könnte. Mit dieser Typologie wiederum könnte eine neue Arbeit begonnen werden, die versucht die Studienerfolge dieser unterschiedlichen Typen zu erfassen.

Interessant erschienen hier, dass die Mehrheit der Psychologieinteressenten (93 Personen= 74,4%) angaben, dass sie sich bereits mit den Inhalten des Studiums auseinander gesetzt haben. Dies spricht für die in dieser Arbeit getätigte Annahme, dass für das Self-Assessment Psychologie kein allgemeiner Interessenfragebogen relevant erscheint, sondern eher ein spezifischer. Aus den Fragen zur Studienwahlreife ging außerdem hervor, dass 4 der 125 Personen (3,2%) keine Angaben bezüglich ihrer Stärken und Schwächen machten und dass 12 der 125 Personen (9,6%) keine Angaben über alternative Berufswünsche machten. Dies könnte auf eine geringe Studienwahlreife hindeuten, da sich diese Personen einerseits noch keine Gedanken über ihre Stärken und Schwächen gemacht haben, wobei dies für eine Ausbildungsentscheidung als relevant angesehen werden kann, und andererseits keine beruflichen Alternativen vorstellen können. Auch Angaben bezüglich der geplanten Studiendauer in Semestern können auf Uninformiertheit und somit auf eventuelle Studienwahlreife hindeuten. Angaben in der Höhe von 10 Semestern (20 Personen= 16%) könnten auf großen Ehrgeiz oder Informiertheit hindeuten. Wohingegen Angaben über 10 Semester, folglich mit „Bufferzeiten“ auf eine realistische Einschätzung hinweisen könnten. Angaben unter 10 Semester weisen allerdings eindeutig auf Uninformiertheit hin, da die Mindeststudienzeit 10 Semester beträgt. Jedoch ist die Beurteilung der Studienwahlreife nicht alleine von einer zu beantwortenden Frage abhängig. Bei der Rückmeldung an die genannten Personen sollten demnach alle Antworten genau analysiert werden, um solche Schlüsse zu zulassen. Gegebenenfalls kann durch eine weitere Antwort der Verdacht auf fehlende Studienwahlreife gefestigt werden.

Ein wesentlicher Kritikpunkt dieser Arbeit ist die geringe Stichprobengröße. Obwohl einiges in die Wege geleitet wurde um eine angemessene Stichprobengröße zu erreichen, wie in Kapitel 6.2 ausführlich dargestellt, bestand die Stichprobe zuletzt aus 125 Personen, die zumindest eines der beiden Verfahren vollständig bearbeitet hatten. Der Zeitraum zwischen der Akquisition von Psychologieinteressenten bei der BeSt³ im März 2009 bis zum Beginn der Datenerhebung im Juli 2009 schien eindeutig viel zu lange zu sein. Man hätte hier von vornherein erwarten können, dass nach diesem langen Zeitraum nicht mehr so viele Personen Interesse an der Teilnahme dieser Untersuchung zeigen. Eine angemessen große Stichprobe wäre wünschenswert gewesen, um das Partial Credit Model anzuwenden, ohne zu befürchten, dass die Parameter in diesem nicht genau geschätzt werden können.

Die Arbeit stellt somit zwei eindimensionale psychologisch-diagnostische Verfahren für die Praxis zur Verfügung, welche im Weiteren noch hinsichtlich ihrer Gütekriterien untersucht werden sollten.

9 Zusammenfassung

Diese Arbeit hatte das Ziel einen Matrizentest und einen Interessenfragebogen für das Self-Assessment Psychologie der Universität Wien, welche dem Rasch-Modell entsprechen, zu konstruieren. Der Kritik von manchen Testautoren folgend, dass die Itemkonstruktion nicht intuitiv stattfinden sollte (vgl. Klauer, 1978, 1984; Schott & Wieberg, 1984; Hornke & Habon, 1984b, 1986), wurde für die Konstruktion der Verfahren ein regelgeleitetes Vorgehen gewählt. Dies sollte bereits in der Konstruktionsphase die Validität der Verfahren stützen.

Dieser Erhebung lag eine Stichprobe von 125 Personen im Alter zwischen 16 und 37 Jahren zugrunde, welche sich für das Psychologiestudium interessierten und den Beginn des Studiums in Erwägung zogen. Da diese Stichprobe zukünftige Anwender des Self-Assessments adäquat abbilden sollte, wurden Psychologieinteressenten als Stichprobe gewählt.

Für den Färbigen Matrizentest – FMT wurden bereits bewährte Konstruktionsprinzipien (vgl. Formann, 1973; Hornke & Habon, 1984a) näher analysiert und teilweise übernommen, sofern sie umsetzbar waren, da beim FMT das Fragezeichen in der Mitte der Matrix steht. Nach der Festlegung der Regeln bzw. Konstruktionsprinzipien für den FMT wurden 28 Items und zwei Instruktionsitems konstruiert.

Der Fragebogen zum Interesse am Psychologiestudium – FIP wurde anhand des Interessenkonstrukt sensu Brickenkamp (1990) und anhand der Schwerpunkte im Psychologiestudium Wien regelgeleitet als spezifischer Interessenfragebogen konstruiert. Es wurde versucht die Regeln weitgehend so zu entwickeln, dass, wie von Feger (1984) gefordert, andere Testautoren zu übereinstimmenden Items gelangen würden. Neben der Skala zum Interesse für Inhalte im Psychologiestudium Wien (35 Items) sollten durch zwei weitere Skalen einerseits auch Schwerpunkte in anderen Bundesländern, in diesem Fall Klagenfurt (5 Items) und andererseits typische Irrtümer bzw. falsche Erwartungen (5 Items) mit dem Fragebogen erhoben werden.

Im Weiteren wurden die beiden Verfahren Analysen zum Rasch-Modell unterzogen. Beim FMT wurden 27 der 30 Items auf Rasch-Modell-Konformität überprüft. Erste Ergebnisse deckten auf, dass sechs Items mangelnde Modellanpassung zeigten. Zuletzt konnte jedoch, nach einer

inhaltlichen Itemanalyse der auffälligen Items, für 24 Items post-hoc Modell-Konformität erreicht werden. Da angenommen wurde, dass die Konstruktionsprinzipien des FMT als schwierigkeitsgenerierende Faktoren die kognitiven Operationen bei der Lösung der Items abbilden, wurden diese mittels LLTM überprüft. Jedoch zeigte sich bei allen drei Durchgängen signifikante Likelihood Quotienten-Test. Somit konnte angenommen werden, dass das Rasch-Modell die Itemschwierigkeiten besser beschreibt als dies die angenommen Basisparameter im LLTM tun.

Beim FIP wurde für jede einzelne Skala eine Überprüfung auf Rasch-Modell-Konformität vorgenommen. Hier zeigte sich bei allen Likelihood-Ratio Tests nach Andersen, dass alle Items der drei Skalen nachweislich eindimensional messen. Nur für Item b2 der Skala Psychologie Wien konnte dies beim Teilungskriterium Geschlecht nicht nachgewiesen werden.

Der Anfangs formulierten Forderung, der Entwicklung zweier psychologisch-diagnostischer Verfahren, deren Items dem Rasch-Modell konform sind, konnte somit entsprochen werden. Es sei jedoch hier darauf verwiesen, dass noch einige Studien zur Belegung der Gütekriterien ausständig sind und erfolgen sollten.

Literaturverzeichnis

Arbeitsbereich Psychologische Diagnostik (2009, Februar). *Forschungsprofil des Arbeitsbereichs Psychologische Diagnostik, Institut für Entwicklungspsychologie und Psychologische Diagnostik, Fakultät für Psychologie* [Online im Internet]. http://psychologie.univie.ac.at/uploads/media/Forschungsprofil_Arbeitsbereich_Psychologische_Diagnostik_Feb_09.pdf [15.03.2009].

Beckmann, J. F. & Guthke, J. (1999). *Psychodiagnostik des Schlussfolgernden Denkens. Handbuch zur Adaptiven Computergestützten Intelligenz-Lerntestbatterie für Schlussfolgerndes Denken (ACIL)*. Göttingen: Hogrefe-Verlag.

Bergmann, C. (1998). Bedingungen und Auswirkungen einer interessenentsprechenden Studienwahl. In J. Abel & C. Tarnai (Hrsg.), *Pädagogisch-psychologische Interessenforschung in Studium und Beruf* (S. 29-43). Münster: Waxmann Verlag GmbH.

Bergmann, C. (2003). Interessenfragebogen. In K. D. Kubinger & R. S. Jäger (Hrsg.), *Schlüsselbegriffe der Psychologischen Diagnostik* (S. 225-229). Weinheim: Beltz Verlag.

Bergmann, C. & Eder, F. (2005). *Allgemeiner Interessen-Struktur-Test-Revision mit Umwelt-Struktur-Test-Revision (AIST-R/UST-R)*. Göttingen: Beltz Test GmbH.

Brickenkamp, R. (1990). *Die Generelle Interessen-Skala (GIS)*. Göttingen: Hogrefe.

Campbell, D. P. & Borgen, F. H. (1999). Holland's theory and the development of interest inventories. *Journal of Vocational Behavior*, 55, 86-101.

Deutsche Psychologen Akademie GmbH des BDP (2002). *DIN 33430 Portal* [Online im Internet]. URL: <http://www.din33430portal.de/> [09.09.2009].

Dilger, B., Gerholz, K.-H., Klieber, S. & Sloane, P. F. E. (2008). Studentisches Self-Assessment: Instrumente zur Unterstützung der Studienwahl. Analyse und Beurteilung von Self-Assessment-

Systemen. In P. F. E. Sloane (Hrsg.), *Higher Education 1: Texte zur Hochschuldidaktik und –entwicklung* (S. 7-76). Paderborn: Eusl-Verlagsgesellschaft mbH.

Dudenredaktion (2001). *Duden. Das Fremdwörterbuch* (7., neu bearbeitete und erweiterte Aufl.). Mannheim: Bibliographisches Institut & F. F. Brockhaus AG.

Embretson, S. E. (1999). Generating items during testing: Psychometric issues and models. *Psychometrika*, 64 (4), 407-433.

Feger, B. (1984). Die Generierung von Testitems zu Lehrtexten. *Diagnostica*, 30 (1), 24-46.

Fischer, G. H. (1974). *Einführung in die Theorie psychologischer Tests*. Bern: Verlag Hans Huber.

Fischer, G. H. (1995). The Linear Logistic Test Model. In G. H. Fischer & I. W. Molenaar (Eds.), *Rasch Models: Foundations, Recent Developments, and Applications* (S. 131-155). New York: Springer-Verlag.

Formann, A (1973). *Die Konstruktion eines neuen Matrizentests und die Untersuchung des Lösungsverhaltens mit Hilfe des linearen logistischen Testmodells*. Unveröff. Diss., Universität, Wien.

Formann, A. K. & Piswanger, K. (1979). *Wiener Matrizen – Test (WMT). Ein Rasch-skaliertes sprachfreier Intelligenztest*. Weinheim: Beltz.

Frebort, M. & Kubinger, K. D. (2006). Ermittlung eines Anforderungsprofils für Studierende der Psychologie. In B. Gula, R. Alexandrowicz, S. Strauß, E. Brunner, B. Jenull-Schiefer & O. Vitouch (Hrsg.), *Perspektiven psychologischer Forschung in Österreich. Proceedings zur 7. Wissenschaftlichen Tagung der Österreichischen Gesellschaft für Psychologie* (S. 408-414). Lengerich: Pabst Science Publishers.

Frebort, M. & Kubinger, K. D. (2008). Qualitätsansprüche an ein Self-Assessment zur Studienwahlberatung: Der Wiener Ansatz. In H. Schuler & B. Hell (Hrsg.), *Studierendenauswahl und Studienentscheidung* (S. 95-101). Göttingen: Hogrefe.

Fritsch, E. (2008). *Das Wiener Self-Assessment Psychologie*. Unveröff. Dipl.Arbeit, Universität Wien.

Fröhlich, W. D. (2002). *Wörterbuch Psychologie* (24., durchgesehene Aufl.). München: Deutscher Taschenbuchverlag.

Goswami, U. (2001). *So denken Kinder: Einführung in die Psychologie der kognitiven Entwicklung* (1. Aufl.) (M. Wengenroth Übers.). Bern: Verlag Hans Huber. (Original erschienen 1998: Cognition in Children)

Heller, K. A., Kratzmeier, H. & Lengfelder, A. (1998a). *Matrizen-Test-Manual: Band 1. Ein Handbuch mit deutschen Normen zu den Standard Progressive Matrices von J. C. Raven*. Göttingen: Beltz-Test GmbH.

Heller, K. A., Kratzmeier, H. & Lengfelder, A. (1998b). *Matrizen-Test-Manual: Band 2. Ein Handbuch mit deutschen Normen zu den Advanced Progressive Matrices von J. C. Raven*. Göttingen: Beltz-Test GmbH.

Holland, J. L. (1997). *Making Vocational Choices: A theory of vocational personalities and work environments* (3rd ed.). Odessa: Psychological Assessment Resources.

Holling, H., Blank, H., Kuchenbäcker, K. & Kuhn, J.-T. (2008). Rule-based item design of statistical word problems: A review and first implementation. *Psychology Science Quarterly*, 50, 363-378.

Hornke, L. F. & Habon, M. W. (1984a). Regelgeleitete Konstruktion und Evaluation von nicht-verbalen Denkitems. In Bundesministerium der Verteidigung – PII 4 – (Hrsg.), *Wehrpsychologische Untersuchungen* (4/84), Bonn.

Hornke, L. F. & Habon, M. W. (1984b). Erfahrungen zur rationalen Konstruktion von Testaufgaben. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 5 (3), 203-212.

Hornke, L. F. & Habon, M. W. (1986). Rule-based item bank construction and evaluation within the linear logistic framework. *Applied Psychological Measurement*, 10 (4), 369-380.

Hornke, L. F. & Rettig, K. (1989). Regelgeleitete Itemkonstruktion unter Zuhilfenahme kognitionspsychologischer Überlegungen. In K. D. Kubinger (Hrsg.), *Moderne Testtheorie* (S. 140-162) (2. verbesserte Auflage). Weinheim: Beltz.

Hornke, L. F., Etzel, S. & Rettig, K. (1999). *Adaptiver Matrizentest - AMT*. Mödling: Schuhfried GmbH.

Hornke, L. F., Küppers, A. & Etzel, S. (2000). Konstruktion und Evaluation eines adaptiven Matrizentests. *Diagnostica*, 46 (4), 182-188.

Irle, M. & Allehoff, W. (1988). *Berufs-Interessen-Test II (B-I-T. II)* (2. Auflage). Göttingen: Hogrefe Verlag.

Jörin, S., Stoll, F., Bergmann, C. & Eder, F. (2003). *EXPLORIX® - das Werkzeug zur Berufswahl und Laufbahnplanung*. [Test-Set: Ausgabe Österreich]. Bern: Verlag Hans Huber.

Jörin, S., Stoll, F., Bergmann, C. & Eder, F. (2004). *EXPLORIX® - das Werkzeug zur Berufswahl und Laufbahnplanung*. [Manual]. Bern: Verlag Hans Huber.

Klauer, K. J. (1978). Kontentvalidität. In K. J. Klauer (Hrsg.), *Handbuch der Pädagogischen Diagnostik. Band 1* (S. 225-255). Düsseldorf: Pädagogischer Verlag Schwann.

Klauer, K. J. (1984). Kontentvalidität. *Diagnostica*, 30 (1), 1-23.

Klauer, K. J. (1993). Training des induktiven Denkens. In K. J. Klauer (Hrsg.), *Kognitives Training* (S. 141-163). Göttingen: Hogrefe Verlag.

Krapp, A. (2002). Structural and dynamic aspects of interest development: Theoretical considerations from an ontogenetic perspective. *Learning and Instruction*, 12, 383-409.

Kubinger, K. D. (1989). Aktueller Stand und kritische Würdigung der Probabilistischen Testtheorie. In K. D. Kubinger (Hrsg.), *Moderne Testtheorie* (S. 19-83) (2., verbesserte Aufl.). Weinheim: Beltz.

Kubinger, K. D. (2000). Kommentar Replik auf Jürgen Rost „Was ist aus dem Rasch-Modell geworden?“. Und für die Psychologische Diagnostik hat es doch revolutionäre Bedeutung. *Psychologische Rundschau*, 51, 33-34.

Kubinger, K. D. (2003). Testtheorie, Probabilistische. In K. D. Kubinger & R. S. Jäger (Hrsg.), *Schlüsselbegriffe der Psychologischen Diagnostik* (S. 415-422). Weinheim: Beltz Verlag.

Kubinger, K. D. (2008). On the revival of the Rasch model-based LLTM: From constructing tests using item generating rules to measuring item administration effects. *Psychology Science Quarterly*, 50 (3), 311-327.

Kubinger, K. D. (2009a). Applications of the Linear Logistic Test Model in Psychometric Research. *Educational and Psychological Measurement*, 69 (2), 232-244.

Kubinger, K. D. (2009b). *Psychologische Diagnostik. Theorie und Praxis psychologischen Diagnostizierens* (2., überarbeitete und erweiterte Aufl.). Göttingen: Hogrefe.

Kubinger, K. D. & Ebenhö, J. (1996). *Arbeitshaltungen. Kurze Testbatterie: Anspruchsniveau, Frustrationstoleranz, Leistungsmotivation, Impulsivität/Reflexivität (AHA)*. Frankfurt: Swets Test Services.

Kubinger, K. D. & Maryschka, C. (2001). *LAMBDA- Lernen auswendig, Merken, Belastbarkeit, Denken-analytisch*. [Manual] 1. Konzeptentwurf. Frankfurt: Swets Test Services.

Kubinger, K. D., Moosbrugger, H., Frebort, M., Jonkisz, E. & Reiß, S. (2007). Die Bedeutung von Self-Assessments für die Studienplatzbewerbung. *Report Psychologie*, 32, 322-332.

Kuhl, J. (2005/2006). *Kurzanweisung zum Fragebogen HAKEMP 90* [Online im Internet]. URL: <http://www.psychologie.uzh.ch/fachrichtungen/motivation/Studium/lehre/alt/ws0506/methodenIWS0506/hakemp90.pdf> [1.11.2009].

Lägeler, H. (2005). Interesse und Bildung. In J. Rekus (Hrsg.), *Grundfragen der Pädagogik: Studien – Texte – Entwürfe, Band 6* (S. 3-205). Frankfurt am Main: Peter Lang GmbH.

Lehrstuhl für Betriebs- und Organisationspsychologie der RWTH Aachen (2008). *Self-Assessments der RWTH Aachen* [Online im Internet]. URL: <http://www.global-assess.rwth-aachen.de/rwth/tm/> [09.09.2009].

Mair, P. & Hatzinger, R. (2007). Extended Rasch Modeling: The eRm Package for the Application of IRT Models in R. *Journal of Statistical Software*, 20 (9), 1–20. Erhältlich von <http://www.jstatsoft.org/v20/i09>

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47, 149-174.

Müller, F. H. (2001). *Studium und Interesse. Eine empirische Untersuchung bei Studierenden*. Münster: Waxmann Verlag GmbH.

Oerter, R. & Dreher, M. (2002). Entwicklung des Problemlösens. In R. Oerter & L. Montada (Hrsg.), *Entwicklungspsychologie* (5. vollständig überarbeitete Aufl.) (S. 469-494). Weinheim: Beltz PVU.

Poinstingl, H., Mair, P. & Hatzinger, R. (2007). *Manual zum Softwarepackage eRm (extended Rasch modelling). Anwendung des Rasch-Modells (1-PL Modell)*. Lengerich: Pabst Science Publishers.

Proyer, R. T. & Häusler, J. (2008). *MOI – Multimethodische Objektive Interessenstestbatterie*. Version 22.00. [Manual]. Mödling: Dr. G. Schuhfried GmbH.

Psychologisches Team (2003). *Studienführer Psychologie 2003/04. Die Starthilfe für dein Studium (11. überarbeitete Auflage)*. Wien: Psychologisches Team (PT).

Putz-Osterloh, W. (1981). *Problemlöseprozesse und Intelligenztestleistungen*. Bern: Verlag Hans Huber.

Rasch, G. (1980). *Probabilistic Models for some intelligence and attainment tests*. Chicago: The University of Chicago Press.

Raven, J. C., Court, J. H. & Raven, J. (2001). *Raven's Coloured Progressive Matrices – Version 24.00*. [Manual]. Mödling: Dr. G. Schuhfried Ges.m.b.H.

Rayman, J. & Atanasoff, L. (1999). Holland's theory and career intervention: the power of the hexagon. *Journal of Vocational Behavior*, 55, 114-126.

Reimann, G. (2004). Arbeits- und Anforderungsanalyse. In K. Westhoff, L. J. Hellfritsch, L. F. Hornke, K. D. Kubinger, F. Lang, H. Moosbrugger, A. Püschel & G. Reimann, (Testkuratorium) (Hrsg.), *Grundwissen für die berufsbezogene Eignungsdiagnostik nach DIN 33430*, (S. 105-120). Lengerich: Pabst Science Publishers.

Schiefele, H., Prenzel, M., Krapp, A., Heiland, A. & Kasten, H. (1983). Zur Konzeption einer pädagogischen Theorie des Interesses. In H. Schiefele & A. Krapp (Hrsg.), *Arbeiten zur Empirischen Pädagogik und Pädagogischen Psychologie*, Nr. 6., (S. 1-35). München: Gelbe Reihe.

Schiefele, U., Krapp, A., Wild, K.-P. & Winteler, A. (1992). Eine neue Version des „Fragebogens zum Studieninteresse“ (FSI). Untersuchungen zu Reliabilität und Validität. In A. Krapp (Hrsg.), *Arbeiten zur Empirischen Pädagogik und Pädagogischen Psychologie*, Nr. 21., (S. 1-27). München: Gelbe Reihe.

Schiefele, U., Streblow, L. & Brinkmann, J. (2007). Aussteigen oder Durchhalten: Was unterscheidet Studienabbrecher von anderen Studierenden? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 39 (3), 127-140.

Schmotzer, C., Kubinger, K. D. & Maryschka, C. (1994). *Rechnen in Symbolen*. [Manual]. Frankfurt am Main: Swets.

Schott, F. & Wieberg, H.-J. W. (1984). Regelgeleitete Itemkonstruktion. Ein Verfahren zur Definition von Itemuniversa und deren kontextvalider Abbildung in Itemmengen für Tests und Treatments. *Diagnostica*, 30 (1), 47-66.

Steckel, M. & Peters, S. (2008). *Perspektiven auf das Moratorium Studium. Teilstudie 3: Studiengang-/ Studienfachwechsel und Studienabbruch* (Arbeitsbereich Nr. 62). Magdeburg: Otto-von-Guericke-Universität, Fakultät für Geistes, Sozial- und Erziehungswissenschaften, Institut für Berufs- und Betriebspädagogik.

Super, D. E. (1957). *The Psychology of Careers: An introduction to vocational development*. New York: Harper & Row.

Test- und Beratungsstelle des Arbeitsbereichs Psychologische Diagnostik (2005). *Studierfähigkeit zum Psychologiestudium* [Online im Internet]. URL: <http://www.studienberatungstest.at/> [29.08.2009].

The R foundation for statistical computing (2009). *R for Windows (Version 2.10.0)* [Computer software]. Wien: Institute for Statistics and Mathematics. Erhältlich von: <http://www.R-project.org>

Todt, E. (1972). *Differentieller Interessen-Test (DIT)* (1. Nachdruck). Bern: Hans Huber Verlag.

Todt, E. (1978). *Das Interesse: Empirische Untersuchungen zu einem Motivationskonzept*. Bern: Verlag Hans Huber.

Universität Hamburg – Fachbereich Psychologie (2009). *Studienkompass Psychologie* [Online im Internet]. URL: <http://webapp6.rrz.uni-hamburg.de/psykompass/> [09.09.2009].

Verbund Norddeutscher Universitäten (2009). *SelfAssessment für zukünftige Studierende* [Online im Internet]. URL: <http://www.selfassessment.uni-nordverbund.de/> [09.09.2009].

Wagner-Menghin, M. (2004). *Lexikon-Wissen-Test – Version 24.00*. [Manual]. Mödling: Schuhfried GmbH.

Westhoff, K., Hellfritsch, L. J., Hornke, L. F., Kubinger, K. D., Lang, F., Moosbrugger, H., Püschel, A. & Reimann, G. (Hrsg.) (2004). *Grundwissen für die berufsbezogene Eignungsbeurteilung nach DIN 33430*. Lengerich: Pabst Science Publishers.

Abbildungsverzeichnis

Abb. 1: Definitionssatz zum induktiven Denken von Klauer (1993, S. 143).....	27
Abb. 2: Einführungsbeispiel aus dem Aufgabenheft des WMT (Formann & Piswanger, 1979)...	30
Abb. 3: Beispiel einer Figur mit zwei Komponenten.....	31
Abb. 4: Aufgabe 1 aus dem AMT; Quelle: http://www.schuhfried.at/index.php?id=377&L=4 ...	32
Abb. 5: Aufgabe 5 des FMT.....	39
Abb. 6: Aufgabe 9 des FMT.....	40
Abb. 7: Aufgabe 28 des FMT.....	41
Abb. 8: Beispielaufgabe (Hornke & Habon, 1984a, S. 38).....	45
Abb. 9: Probleme von Todt bei der Entwicklung eines Interessenfragebogens (1978, S. 28).....	50
Abb. 10: Hexagonales Modell von Holland (1997, S. 6, S. 35).....	56
Abb. 12: Instruktionsbeispiel des FIP.....	72
Abb. 13: Seite 2 der Testbatterie dieser Diplomarbeit.....	74
Abb. 14: Übungsbeispiel 1 des FMT.....	76
Abb. 15: Übungsbeispiel 2 des FMT.....	77
Abb. 16: Übungsbeispiel des FIP.....	78
Abb. 17: Histogramm der Altersverteilung.....	81
Abb. 18: Histogramm der geplanten Semesteranzahl.....	84
Abb. 19: Graphischer Modelltest FMT – Teilungskriterium Median, Items 1-15.....	89
Abb. 20: Graphischer Modelltest FMT – Teilungskriterium Median, Items 16-27.....	89
Abb. 21: Aufgabe 7 (Item a7) des FMT.....	91
Abb. 22: grafische Modelltests der Items für Psychologie Wien.....	95
Abb. 23: grafische Modelltests der 5 Items für Psychologie Klagenfurt.....	96
Abb. 24: grafische Modelltests der 5 Items für typische Irrtümer/falsche Erwartungen.....	97

Tabellenverzeichnis

Tabelle 1: Kernaufgaben des induktiven Denkens nach Klauer (1993, S. 144).....	27
Tabelle 2: Verteilung der Regelanzahl je Item.....	42
Tabelle 3: Verteilung der Regeln je Item.....	44
Tabelle 4: Schlüssel zur Konstruktion der Distraktoren (Hornke & Habon, 1984a, S. 40).....	45
Tabelle 5: Schlüssel zur Konstruktion der Distraktoren für einzelne Regeln.....	46
Tabelle 6: Auszug aus der Charakterisierung der Dimensionen R, I, A, S, E und C.....	55
Tabelle 7: Auflistung der Inhaltsbereiche im Psychologiestudium nach Sst.....	66
Tabelle 8: Aufgliederung der Tätigkeiten zu den Verhaltensmodalitäten.....	71
Tabelle 9: Geschlechterverteilung.....	80
Tabelle 10: Häufigkeitsverteilung der kategorisierten Variable Idee.....	82
Tabelle 11: Häufigkeitsverteilung der kategorisierten Variable Wen.....	83
Tabelle 12: Gruppenstatistik der Variable Semesteranzahl.....	85
Tabelle 13: Ergebnisse LRT für FMT – kein Item ausgeschlossen.....	88
Tabelle 14: Konstruktionsregeln der im z-Test auffälligen Items.....	90
Tabelle 15: Häufigkeiten der Antwortmöglichkeiten des Item a7.....	92
Tabelle 16: Ergebnisse LRT für FMT – 25 Items (ohne a7 & a14).....	93
Tabelle 17: Ergebnisse LRT für FIP – Items Psychologie Wien.....	94
Tabelle 18: Ergebnisse LRT für FIP – Items Klagenfurt.....	95
Tabelle 19: Ergebnisse LRT für FIP – Items typische Irrtümer/falsche Erwartungen.....	97
Tabelle 20: Likelihood-Quotienten Test A.....	98
Tabelle 21: Likelihood-Quotienten Test B.....	99
Tabelle 22: Q-Matrix der Hypothese B.....	100
Tabelle 23: Likelihood-Quotienten Test C.....	101
Tabelle 24: Q-Matrix der Hypothese C.....	102
Tabelle 25: Eich Tabelle FMT - Psychologieinteressenten.....	121
Tabelle 26: Eich Tabelle FIP: Psychologie Wien - Psychologieinteressenten.....	122
Tabelle 27: Eich Tabelle FIP: Psychologie Klagenfurt - Psychologieinteressenten.....	123
Tabelle 28: Eich Tabelle FIP: typische Irrtümer/falsche Erwartungen – Psychologieinteressenten	123

Anhang

Anhang 1: Definitionen der ausgewählten Schwerpunkte

Psychologische Diagnostik

Forschungsprofil

Das Forschungsprofil des Arbeitsbereichs Psychologische Diagnostik umfasst drei Schwerpunkte:

I. Grundlagenforschung zum diagnostischen Prozess – Effizienz, Zumutbarkeit und Fairness

- (1) Adaptives Testen: Grenzen und Möglichkeiten, Anwendung auf Persönlichkeitsfragebogen
- (2) Optimierung des diagnostischen Prozesses: Anamnese/Exploration – Test – Begutachtung
- (3) Optimierung des diagnostischen Prozesses hinsichtlich Zumutbarkeit
- (4) Grenzen der Zumutbarkeit in Bezug auf Testdauer, Intimität und Frustration
- (5) Verständlichkeit von psychologischen Gutachten, Akzeptanz und Konsequenz
- (6) Umsetzung der DIN 33430
- (7) Feststellung der Testbarkeit
- (8) Testung ausländischen Klientel
- (9) Gendereffekte innerhalb des diagnostischen Prozesses: Interaktion Testleiter(in) und Testperson
- (10) Alternativen zum multiple-choice-Antwortformat
- (11) Maßnahmen zur Reduzierung der Verfälschbarkeit von Persönlichkeitsfragebogen
- (12) Fairness gegenüber Testpersonen ohne Testroutine
- (13) Äquivalenz von Computertests
- (14) Gewährleistung des Instruktionsverständnisses

II. Produktentwicklung (psychologisch-diagnostischer Verfahren)

- (1) Europäisierte/globalisierte Verfahren zur Messung kognitiver Fähigkeiten
- (2) Objektive Persönlichkeitstests (Experimentalpsychologische Verhaltensdiagnostik)
- (3) Familiendiagnostische Verfahren
- (4) Alternativen zum Markt

III. Statistische und testtheoretische Grundlagen der Psychologischen Diagnostik

- (1) Anwendungsprobleme von IRT-Modellen
- (2) Statistische Artefakte
- (3) Verfahrensentwicklung für rangskalierte Merkmale

http://psychologie.univie.ac.at/uploads/media/Forschungsprofil_Arbeitsbereich_Psychologische_Diagnostik_Feb_09.pdf, 15.03.09)

Schwerpunkte:

Psychologische Diagnostik ist die wissenschaftliche Disziplin, die psychologisches Diagnostizieren für die Praxis vorbereitet. Psychologisches Diagnostizieren ist ein Prozess, der unter Zuhilfenahme verschiedener Verfahren zielgerichtete Informationen über psychische Eigenschaften des in Betracht stehenden Menschen gewinnen will; dieser Prozess endet mit einem Maßnahmenvorschlag. Das Prüfungsfach vermittelt die entsprechende Kompetenz (u.a.: Klärung der Fragestellung, Administration und Auswertung psychologisch-diagnostischer Verfahren, Abfassung psychologischer Gutachten) inklusive der zur Beurteilung der Qualität psychologisch-diagnostischer Verfahren nötigen Kenntnisse. (Psychologisches Team, 2003, S. 30)

Anhang 2: Eichtabelle FMT

Tabelle 25: Eichtabelle FMT - Psychologieinteressenten

Score	Fähigkeitsparameter	Prozentrang
1	-3,4018342	1,6
2	-2,6369740	2,4
3	-2,1581841	8,1
4	-1,7953952	12,1
5	-1,4951424	16,9
6	-1,2334415	23,4
7	-,9973049	32,3
8	-,7787991	39,5
9	-,5725823	46,8
10	-,3747628	54,8
11	-,1822619	58,1
12	,0075001	63,7
13	,1969292	71,8
14	,3884436	78,2
16	,7885493	83,1
17	1,0041402	87,1
18	1,2367576	90,3
19	1,4943669	93,5
20	1,7899971	96,0
21	2,1476933	96,8
22	2,6209846	97,6
23	3,3800618	99,2
>23	>3,3800618	>99,2

Anhang 3: Eichtabellen FIP

Tabelle 26: Eichtabelle FIP: Psychologie Wien - Psychologieinteressenten

Score	Fähigkeitsparameter	Prozentrang
<5	<-2,03	<2,6
5	-2,03	2,6
8	-1,38834061	3,4
12	-,74952205	5,2
13	-,60771204	6,9
14	-,46983939	8,6
15	-,33484662	12,1
16	-,20181437	17,2
17	-,06981513	21,6
18	,06198239	26,7
19	,19433481	31,9
20	,32808906	41,4
21	,46413299	44,8
22	,60334832	53,4
23	,74679513	57,8
24	,89581975	62,1
25	1,05198432	69,0
26	1,21712901	75,0
27	1,39373841	78,4
28	1,58536587	82,8
29	1,79704133	85,3
30	2,03639861	88,8
31	2,31620754	93,1
32	2,66017584	95,7
34	3,87080696	99,1
>34	>3,87080696	>99,1

Tabelle 27: Eich Tabelle FIP: Psychologie Klagenfurt - Psychologieinteressenten

Score	Fähigkeitsparameter	Prozentrang
<1	<-1,5756933	<9,5
1	-1,5756933	9,5
2	-,4728604	24,1
3	,4631149	55,2
4	1,5771984	83,6
>4	>1,5771984	>83,6

Tabelle 28: Eich Tabelle FIP: typische Irrtümer/falsche Erwartungen – Psychologieinteressenten

Score	Fähigkeitsparameter	Prozentrang
<1	<-1,5043971	<14,7
1	-1,5043971	14,7
2	-,4522785	37,9
3	,4386555	60,3
4	1,5060938	84,5
>4	>1,5060938	>84,5

Lebenslauf

Nicole Undeutsch

Geburtsdatum	12. August 1982
Geburtsort	Steyr
Staatsbürgerschaft	Österreich

Ausbildung

seit Oktober 2003	Diplomstudium der Psychologie an der Universität Wien
1997 - 2002	Höhere Bundeslehranstalt für Tourismus und Wirtschaftliche Berufe, Bergheidengasse 5-19, 1130 Wien <i>Schwerpunkt::</i> Tourismusmarketing und -management <i>Abschluss:</i> Diplom- und Reifeprüfung
1992 - 1997	Gymnasium Maroltingergasse 69-71, 1160 Wien
1988 - 1992	Volksschule Grubergasse 4, 1160 Wien

Fachbezogene Berufserfahrung

seit Oktober 2007	<i>Studienassistentin</i> an der Fakultät für Psychologie, am Institut für Entwicklungspsychologie und Psychologische Diagnostik, 1010 Wien
07/2008 – 08/2008	<i>Praktikantin</i> Neurologisches Labor, Baumgartner Höhe, 1140 Wien <i>Tätigkeitsbereich:</i> psychologische Diagnostik, Fallbesprechungen

Wien am 1. März 2010