



universität
wien

Diplomarbeit

Titel der Arbeit

Ein Beitrag zur Konstruktvalidierung einer Kurzversion des
WSP aus dem Wiener Self-Assessment Psychologie: Sind
Objektive Persönlichkeitstests zur Konstruktvalidierung
eines Fragebogens geeignet?

Verfasserin

Katarina Ifcic

Angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag. rer. nat.)

Wien, im April 2012

Studienkennzahl: 298

Studienrichtung: Psychologie

Betreuer: Univ.-Prof. Dr. Lale Khorramdel

Ich möchte diese Arbeit meinen Eltern und Großeltern widmen.

Abstract

Das Ziel der vorliegenden Arbeit war die Konstruktvalidierung der Kurzversion des *Wiener Studieneignungs-Persönlichkeitsinventars* (WSP, Khorramdel & Maurer, unpubl.), und zwar der vier Skalen *Organisation*, *Motivation*, *Frustrationstoleranz* und *Selbstbewusstsein* wie sie im **Wiener Self-Assessment Psychologie**® zu finden sind, und auch die Beantwortung der Frage, ob sich Objektive Persönlichkeitstests zur Konstruktvalidierung von Persönlichkeitsfragebogen eignen. Die Stichprobe setzte sich aus 180 Personen zusammen, die im Zeitraum zwischen 29.08.2011 und 04.12.2011 das **Wiener Self-Assessment Psychologie**® abgeschlossen haben.

Die korrelative Überprüfung ergab, dass es, bis auf einen, keine signifikanten Zusammenhänge zwischen den betreffenden Skalen und den Testkennwerten der Objektiven Persönlichkeitstests gab. Die korrelative Auswertung der MTMM-Matrix ergab keine konvergente oder diskriminante Validität. Eine Faktorenanalyse zeigte, dass der Persönlichkeitsfragebogen und die Objektiven Persönlichkeitstests nicht auf dieselben Faktoren laden.

Die vorliegende Arbeit konnte zeigen, dass zwischen dem Wiener Studieneignungs-Persönlichkeitsinventar und den Objektiven Persönlichkeitstests aus dem **Wiener Self-Assessment Psychologie**® kein Zusammenhang besteht. Objektive Persönlichkeitstests eignen sich nicht zur Konstruktvalidierung von Persönlichkeitsfragebogen.

Abstract

The aim of this work was a construct validation of the short version of the *Wiener Studieneignungs-Persönlichkeitsinventar* (WSP, Khorramdel & Maurer, unpubl.), precisely of the four scales *organizational capability*, *motivation*, *frustration tolerance* and *self-confidence*, as found in the **Wiener Self-Assessment Psychologie**®. Also, the question, whether objective personality tests can be used for a construct validation of a personality questionnaire, should be answered. The sample consisted of 180 participants who completed the **Wiener Self-Assessment Psychologie**® between August 29th 2011 and December 4th 2011.

Just one significant correlation, among many non significant ones, could be found between the scales in question and the test values of the objective personality tests. A correlation based analysis of the MTMM-matrix showed neither convergent nor discriminant validity coefficients. A following factor analysis showed that the personality questionnaire and the test values of the objective personality tests don't load on common factors.

This work showed, that there is no connection between the *Wiener Studieneignungs-Persönlichkeitsinventar* and the objective personality tests used in the **Wiener Self-Assessment Psychologie**®. Objective personality tests are not suited to be used for a construct validation of personality questionnaires.

Inhaltsverzeichnis

I Einleitung und Fragestellung der Arbeit.....	3
II Theoretischer Teil.....	7
1. DIN 33430 und das Testgütekriterium der Validität.....	7
1.1 DIN 33430 und die Validität	7
1.2 Geschichtliche Entwicklung des Gütekriteriums Validität.....	7
1.3 Inhaltliche Gültigkeit	9
1.3.1 Definition und Überprüfung der Inhaltlichen Gültigkeit	9
1.3.2 Exkurs: Augenscheinvalidität.....	10
1.4 Kriteriumsvalidität	11
1.4.1 Definition; Übereinstimmungs- und prognostische Validität	11
1.4.2 Überprüfung	12
1.4.3 Exkurs: Ökologische Validität	12
1.5 Konstruktvalidität	13
1.5.1 Der Begriff „Konstrukt“, Definition und das nomologische Netz	13
1.5.2 Überprüfung und Gefahrenquellen bei der Überprüfung	15
2. Das Wiener Studieneignungs-Persönlichkeitsinventar (WSP; Khorramdel & Maurer, unpubl., siehe auch Khorramdel und Maurer, in Druck)	19
2.1 Allgemeine Beschreibung.....	19
2.2 Die Entwicklung des WSP (Maurer, 2010)	19
3. Das Wiener Self-Assessment Psychologie ®	24
3.1 Das diagnostische Instrument des Self-Assessments.....	24
3.2 Entwicklung des Wiener Self-Assessments Psychologie ®	25
3.3 Exkurs: Objektive Persönlichkeitstests.....	29
3.4 Psychologisch diagnostische Verfahren im Wiener Self-Assessment Psychologie ®	29
3.4.1 LAMBDA-neu (Kubinger & Maryschka, 2001).....	29
3.4.2 „Symbole kodieren“ (Kodiertest) (aus <i>Arbeitshaltungen</i> , Kubinger & Ebenhöf, 1996)	32
3.4.3 WITE (Friebert, unpubl.)	34
3.4.4 FIP – Fragebogen zum Interesse im Psychologiestudium (Undeutsch, 2010, unpubl., adaptiert 2010 von Useini) und SKIP – Selbstkonzept im Psychologiestudium (Useini, in Vorb.)	35
3.4.5 FMT – Farbiger Matrizentest (Undeutsch, 2010)	36
3.4.6 Das <i>Wiener Studieneignungs-Persönlichkeitsinventar</i> (WSP; Khorramdel & Maurer, unpubl.).....	36
3.4.7 Persönliche Einschätzung.....	37
3.5 Bisherige Forschungsergebnisse bezüglich Validität und Akzeptanz zum Wiener Self-Assessment Psychologie ®	37
4. Theorie zur statistischen Auswertung der Daten – Methodik.....	39
4.1 Einführung in die Faktorenanalyse und Konfirmatorische FA.....	39
4.2 Das Multi-Trait-Multi-Method-Modell	41
4.2.1 Grundgedanken des Multitrait-Multimethod-Ansatzes/Konvergente und diskriminante Validität.....	41

4.2.2 MTMM-Matrix – Aufbau und Struktur.....	43
4.2.3 Die korrelationsbasierte Auswertung der MTMM-Matrix/Kritikpunkte und Probleme der korrelationsbasierten Auswertung.....	44
III Empirischer Teil.....	49
1. Vorangehende Überlegungen und Grundlegendes zur empirischen Untersuchung.....	49
1.1 Konkrete Umschreibung der relevanten Dimensionen des WSP	50
1.2 Hypothesen und Darstellung der MTMM-Matrix	51
2. Die Stichprobe.....	57
2.1 Akquirierung von Testpersonen.....	57
2.2 Beschreibung der Datenlage	58
2.3 Deskriptive Beschreibung der Stichprobe	59
3. Darstellung der Ergebnisse der Hypothesen	65
3.1 Ergebnisse der Hypothesen zu „ <i>Organisation</i> “	66
3.2 Ergebnisse der Hypothesen zu „ <i>Motivation</i> “	70
3.3 Ergebnisse der Hypothesen zu „ <i>Frustrationstoleranz</i> “	78
3.4 Ergebnisse der Hypothesen zu „ <i>Selbstbewusstsein</i> “	82
4. Darstellung der Ergebnisse der MTMM-Matrix	83
5. Exploratorische Faktorenanalyse	86
IV Zusammenfassung der Ergebnisse und Interpretation.....	90
V Diskussion der Ergebnisse und Kritik.....	93
VI Zusammenfassung	99
VII Literaturverzeichnis.....	101
VIII Anhang	109
1. Der WSP im Rahmen des Wiener Self-Assessment Psychologie ®	109
2. Beschreibung der Datenlage.....	111
3. Darstellung der Ergebnisse der Hypothesen	113
3.1 Prüfung der Voraussetzungen zu <i>Organisation</i>	113
3.2 Prüfung der Voraussetzungen zu <i>Motivation</i>	115
3.3 Prüfung der Voraussetzungen zu <i>Frustrationstoleranz</i>	117
IX Tabellenverzeichnis	119
X Abbildungsverzeichnis	121
XI Lebenslauf	123

I Einleitung und Fragestellung der Arbeit

Das Ziel der vorliegenden Arbeit ist die Konstruktvalidierung der Kurzversion des *Wiener Studieneignungs-Persönlichkeitsinventars* (WSP, Khorramdel & Maurer, unpubl.) aus dem **Wiener Self-Assessment Psychologie®**. Gleichzeitig soll die Frage beantwortet werden, ob sich Objektive Persönlichkeitstests zur Konstruktvalidierung von Persönlichkeitsfragebogen, in diesem Fall des WSP, eignen.

Psychologisch diagnostische Verfahren werden anhand von bestimmten Qualitätsanforderungen konstruiert, angewandt und bezüglich ihrer Qualität beurteilt; den Testgütekriterien (Moosbrugger & Kevala, 2007). Nach Lienert und Raatz (1994) werden die Hauptgütekriterien Objektivität, Reliabilität und Validität, und die Nebengütekriterien Skalierung, Normierung, Ökonomie, Nützlichkeit, Zumutbarkeit, Unverfälschbarkeit und Fairness unterschieden. Dabei stellt die Validität das wichtigste Gütekriterium dar, und ist aber gleichzeitig auch am schwersten zu überprüfen (Kubinger, 2009). Diese Gütekriterien wurden auch, mit etlichen anderen Forderungen, vom Deutschen Institut für Normierung e. V. im Jahr 2002 zur DIN 33430 zusammengefasst (vgl. Reimann, 2005). Die DIN 33430 stellt eine Liste an Anforderungen an die Auswahl, Zusammenstellung, Durchführung, Auswertung und Interpretation der Verfahren zur berufsbezogenen Eignungsbeurteilung dar (Reimann, 2005).

Dies soll nun die Bedeutung der Überprüfung der Konstruktvalidität des vorliegenden Verfahrens im Rahmen des **Wiener Self-Assessment Psychologie®** verständlich machen.

Diese Herangehensweise an die Konstruktvalidierung eines Persönlichkeitsfragebogens, und zwar mittels Objektiver Persönlichkeitstests, ist kritisch zu hinterfragen, da bisherige Studien (Litzenberger & Kubinger, 2003) gezeigt haben, dass diese beiden Erfassungsmethoden scheinbar nicht dasselbe messen. Da dies aber nur für einzelne Objektive Persönlichkeitstests und nur für bestimmte Fragebögen gezeigt wurde, soll in dieser Diplomarbeit nochmals geprüft werden, ob Objektive Persönlichkeitstests zur Konstruktvalidierung von Fragebögen geeignet sind und zwar mit den Verfahren aus dem **Wiener Self-Assessment Psychologie®**.

Das *Wiener Studieneignungs-Persönlichkeitsinventar*, kurz WSP (Khorramdel & Maurer, unpubl.), siehe auch Khorramdel und Maurer (2012), beabsichtigt in seiner Gesamtheit sieben studienrelevante Skalen zu messen. Für die vorliegende Validierung sind jedoch nur vier Skalen von Bedeutung, da sich die Validierung nur auf jene Teile bezieht, die im Rahmen des

Wiener Self-Assessment Psychologie® zur Anwendung kommen, diese sind *Organisation, Motivation, Belastbarkeit* und *Selbstbild*, und dies nochmals beschränkt auf die jeweils für das **Wiener Self-Assessment Psychologie®** ausgewählten Items, insgesamt 53.

Das **Wiener Self-Assessment Psychologie®** wurde zwischen Herbst 2003 und Herbst 2005 von der *Test- und Beratungsstelle des Arbeitsbereichs Psychologische Diagnostik* (Institut II: Fakultät für Psychologie der Universität Wien) vom „*Wiener Autorenkollektiv Studienberatungstests*“ (Kubinger, Frebort, Khorramdel & Weitensfelder, in Druck) entwickelt und erstmals im Herbst 2005 online gestellt (Frebort & Kubinger, 2007). Seitdem wurde es weiterentwickelt, evaluiert und mit neuen psychologisch-diagnostischen Verfahren ausgestattet, sodass es nun Ende August 2011, unter anderem auch mit den betreffenden Skalen des *Wiener Studieneignungs-Persönlichkeitsinventars*, online gestellt wurde. Das **Wiener Self-Assessment Psychologie®** richtet sich erstrangig an InteressentInnen und StudienaspirantInnen des Studiengangs Psychologie an der Universität Wien.

Konkreter ausgedrückt will diese Arbeit die Konstruktvalidität der Kurzversion mit den Skalen *Organisation, Motivation, Frustrationstoleranz* und *Selbstbewusstsein* des WSP überprüfen, und zwar indem diese in Zusammenhang mit den Verfahren der experimentalpsychologischen Verhaltensdiagnostik (Objektive Persönlichkeitstests) aus dem **Wiener Self-Assessment Psychologie®** gesetzt werden. Die Personen der Validierungsstichprobe sollen die gesamte Testbatterie des **Wiener Self-Assessment Psychologie®** bearbeiten und schließlich soll die Validität mit Hilfe von Korrelationen zwischen den einzelnen Testkennwerten und Unterschieden zwischen verschiedenen Gruppen nach zuvor aufgestellten Hypothesen und einer Multitrait-Multimethod-Matrix beurteilt werden. Anschließend folgt eine exploratorische Faktorenanalyse, um zu prüfen, auf welche Faktoren die jeweiligen Kennwerte laden.

Die hier bearbeitete Forschungsfrage soll dementsprechend lauten: „Können Verfahren der experimentalpsychologischen Verhaltensbeobachtung (Objektive Persönlichkeitstest) zur Konstruktvalidierung des *Wiener Studieneignungs-Persönlichkeitsinventars* herangezogen werden?“ Wie kann der Nutzen von Objektiven Persönlichkeitstests in Bezug auf eine Konstruktvalidierung von Persönlichkeitsfragebögen bewertet werden, beziehungsweise umgekehrt; sind Persönlichkeitsfragebögen zur Validierung von Objektiven Persönlichkeitstests geeignet?

Die vorliegende Arbeit wird im theoretischen Teil (Teil II) in den Kapiteln 1-4 zu Beginn den Begriff der Validität erläutern, mit besonderem Augenmerk auf der Konstruktvalidität, dann soll das *Wiener Studieneignungs-Persönlichkeitsinventar*, inklusive Konstruktion und Skalen, beschrieben werden. Darauf folgt eine Darstellung des **Wiener Self-Assessment Psychologie®**. Abschließend zum theoretischen Teil soll auf die Methodik der statistischen Auswertung der Daten eingegangen werden, wobei hier besonderer Wert auf die Faktorenanalyse und die Konzeption und Auswertung einer Multitrait-Multimethod-Matrix gelegt wird.

Im Kapitel III Empirischer Teil folgen die Darstellung der Grundannahmen, sowie schließlich die Prüfung der Hypothesen. Es werden die Ergebnisse der Hypothesen dargestellt, schließlich die Auswertung der MTMM-Matrix und auch die Ergebnisse der folgenden exploratorischen Faktorenanalyse dargestellt.

Abschließend folgt eine Zusammenfassung der Ergebnisse mit Interpretation (Kapitel IV), schließlich folgt in Kapitel V die Diskussion und kritische Auseinandersetzung mit der Vorgehensweise und den Ergebnissen und schließlich die Zusammenfassung der Arbeit in Kapitel VI.

II Theoretischer Teil

1. DIN 33430 und das Testgütekriterium der Validität

1.1 DIN 33430 und die Validität

Das Deutsche Institut für Normierung e. V. (DIN) hat eine Norm (DIN 33430) mit den „Anforderungen an Verfahren und deren Einsatz bei berufsbezogenen Eignungsbeurteilungen“ beschlossen und festgelegt (Reimann, 2005, S. 17). Inhalt dieser Norm sind Qualitätsstandards zum Inhalt und Ablauf von Personalauswahlverfahren, dies bezieht sich auch auf die dabei verwendeten psychologisch-diagnostischen Verfahren (Reimann, 2005). Unter anderem, nachzulesen bei Reimann (2005), fordert die DIN 33430, dass die Validität der eingesetzten Verfahren empirisch nachgewiesen ist.

Die englische Fassung der DIN 33430 findet sich in Westhoff, K., Hellfritsch, L. J., Hornke, L. F., Kubinger, K. D., Lang, F., Moosbrugger, H., Püschel, A. und Reimann G. (2005).

Unter Validität versteht man den Grad der Genauigkeit, mit dem das psychologisch diagnostische Verfahren das, was es messen soll oder zu messen behauptet, misst (Lienert, 1989). Wenn ein Verfahren das Merkmal, das es erfassen soll, auch wirklich erfasst, und nicht ein beliebig anderes, dann spricht man von einem validen Test (Moosbrugger & Kevala, 2007).

1.2 Geschichtliche Entwicklung des Gütekriteriums Validität

Anastasi (1986) skizziert in ihrem Artikel die Geschichte der Validierung, und beschreibt die Anfänge des standardisierten Testens, bei denen die Validität durch verschiedene Verfahren überprüft wurde und verschiedene Namen trug. Die Beweise für eine allfällige Validität wurden auf unterschiedliche Weise erbracht, abhängig vom Zweck des Tests, der theoretischen Orientierung, den Grundannahmen des Testautors und von den bereits verfügbaren Daten (Anastasi, 1986).

Anastasi (1986) führt an, dass Binet und Simon (Binet & Simon, 1908, zit. n. Anastasi, 1986) durch die Annahme, dass kognitive Fähigkeiten mit dem Alter zunehmen, eine erste empirische Annäherung an die Evaluierung von Testitems durchgeführt haben. Sie wählten Aufgaben aus, deren Lösungswahrscheinlichkeit mit zunehmendem Alter stieg, und ordneten

so jede Aufgabe einem Alter zu, bei dem ein bestimmter Prozentsatz der Kinder das Item lösen konnte (Binet & Simon, 1908, zit. n. Anastasi, 1986). So fanden sie, durch die Differenzierung aufgrund des Alters, die validesten Items (Binet & Simon, 1908, zit. n. Anastasi, 1986). Diese Vorgehensweise diente dann schließlich auch in Zukunft zur Konstruktion von Intelligenztests, die Intelligenz auf Basis des mentalen Alters erfassten (Binet & Simon, 1908, zit. n. Anastasi, 1986).

Bald wurden die Gesamtergebnisse von Tests nicht nur in Hinblick auf das Alter, sondern auch mit Hilfe von Fremdurteilen, wie Beurteilungen der Leistungen der Schüler von Lehrern und anderer Beurteilungskriterien des täglichen Lebens, evaluiert, ebenso wie Einzelfallbeschreibungen und psychiatrische Diagnosen bei Persönlichkeitstests und bei Tests zur Erfassung einer kognitiven Behinderung (Anastasi, 1986). Mit dem Fortschritt der statistischen Methoden war es möglich die einzelnen Items eines Tests in Bezug auf den Gesamtscore und in Bezug auf externe Kriterien zu überprüfen (Anastasi, 1986). Mit der Faktorenanalyse war es schließlich möglich, eine Untersuchung einzelner Items, Subtestergebnisse und Gesamtergebnisse mit Ergebnissen anderer Tests durchzuführen (Anastasi, 1986). Dabei wurden verschiedene Bezeichnungen für die gefundenen Validitäten angeführt, wie „*face validity, validity by definition, intrinsic validity, and logical validity, empirical validity and factorial validity*“ (Anastasi, 1986, S. 2).

In der Veröffentlichung der „*Technical Recommendations for Psychological Tests and Diagnostic Techniques*“ (APA, 1954 zit. n. Anastasi, 1986, S. 2) durch die American Psychological Association wurden die Inhaltsvalidität, die „predictive“, „concurrent“ und Konstruktvalidität definiert und formal unterschieden. Durch eine in den „*Standards for Educational and Psychological Testing*“ (APA, 1985, zit. n. Anastasi, 1986) im Jahr 1985 folgende Zusammenfassung der „predictive“ und „concurrent“ Validitäten zur „Kriteriumsvalidität“, wurde die bis heute gängige Dreiteilung begründet, nämlich in Inhaltsvalidität, Kriteriumsvalidität und Konstruktvalidität. Zu Beginn wurden diese als Arten der Validität angesehen, später dann als essentielle Komponenten der Validität (Anastasi, 1986).

Dementsprechend spricht auch Kubinger (2009) davon diese drei Konzepte, beziehungsweise Aspekte der Validität, zu unterscheiden: Inhaltliche Gültigkeit, Konstruktvalidität und Kriteriumsvalidität.

Im Gegensatz dazu nennen Moosbrugger und Kevala (2007) zusätzlich die Augenscheinvalidität und unterscheiden sie klar von der inhaltlichen Validität, und sehen sie als gleichwertig neben der Inhalts-, Konstrukt- und Kriteriumsvalidität.

In dieser Diplomarbeit soll allerdings die Validität nach der Unterscheidung und Einteilung von Kubinger (2009) dargestellt werden.

1.3 Inhaltliche Gültigkeit

1.3.1 Definition und Überprüfung der Inhaltlichen Gültigkeit

Definition

Inhaltliche Gültigkeit, nach einer unmittelbaren Übersetzung aus dem Englischen „*Kontentvalidität*“ (Kubinger, 2009, S. 55), oder „*Inhaltsvalidität*“ nach Moosbrugger und Kevala (2007, S. 15), bezeichnet, inwieweit ein Verfahren oder ein Item das zu messende Merkmal repräsentativ erfasst (Moosbrugger & Kevala, 2007). Ein Test gilt als inhaltlich valide, wenn er selbst das beste und am besten passende Kriterium für das Merkmal darstellt (Kubinger, 2009), also wenn der Test oder die Items das Merkmal oder das Verhalten ideal repräsentieren (Lienert, 1989). Lienert (1989) nennt als Beispiel dafür einen Schulkenntnistest für Geographie, der eine repräsentative Auswahl an Fragen aus dem Unterrichtsstoff der Geographie enthalten muss um inhaltlich valide zu sein, und nicht etwa aus dem Fach Biologie.

Überprüfung

Inhaltliche Validität wird durch ein Expertenrating überprüft (Lienert, 1989). Hierbei wird keine statistische Maßzahl ermittelt, sondern auf Basis der Sachkompetenz von Experten und laut einer vorgegebenen operationalen Definition geprüft, ob das Testmaterial das misst was es messen soll (Kubinger, 2003). Nach Fisseni (2004) ist ein Experte in diesem Zusammenhang eine Person, die mit dem Zielmerkmal vertraut ist.

Die Überprüfung der inhaltlichen Gültigkeit ist nach Fisseni (2004) der Überprüfung der beiden anderen Validitäten vorgeordnet. Fisseni (2004) begründet dies mit den beiden folgenden Argumenten: Wenn ein Test A bei der Überprüfung der Kriteriumsvalidität am Kriterium A validiert wird, so muss auch das Kriterium A valide sein, um damit an einem

Kriterium B validiert werden. Dies würde sich bis in die Unendlichkeit fortsetzen und es würde nie zu einer Erklärung führen, deswegen muss zumindest ein Glied in dieser Kette auf eine andere Weise validiert werden. Das zweite Argument (Fisseni, 2004) sagt aus, dass die Konstruktvalidität die Inhaltsvalidität und die Kriteriumsvalidität mit einschließt, und damit auch das Problem der Kriteriumsvalidität vom ersten Argument. Damit muss auch hier auf eine andere Art validiert werden. Also ist die inhaltliche Validierung im Laufe der Kriteriums- und Konstruktvalidierung erforderlich.

Guion (1977) publizierte fünf Kriterien die bei der Beurteilung der Inhaltsvalidität zu bedenken sind. Erstens muss sich der Inhalt des Tests auf eine Reihe von Verhaltensweisen stützen, deren Bedeutung allgemein anerkannt ist und nennt dafür als Beispiel die Prüfung zur Erlangung des Führerscheins, bei der auf Basis einzelner Teilkompetenzen (rechts abbiegen, einparken, ...) auf die Fahrfähigkeit des Kandidaten geschlossen wird, ein Fahrzeug sicher lenken zu können, also es wird vom Besitz oder Nicht-Besitz des Führerscheines auf die zugrundeliegenden Kompetenzen geschlossen, die allgemein bekannt sind (Guion, 1977). Der Autor (Guion, 1977) nennt als zweite Bedingung, dass der Umfang und damit auch die Grenzen des Inhalts eindeutig definiert sein sollen, damit man zweifellos erkennen kann, ob ein Item oder eine Beobachtung innerhalb oder außerhalb dieser Grenzen liegt. Drittens muss der Inhalt relevant für die Ziele der vorliegenden Messung sein, was bedeutet, dass ein Wissenstest zum Beispiel dann auch die Inhalte des Lehrplanes des betreffenden Faches abfragen muss, und nicht andere Inhalte (Guion, 1977). Als viertes Kriterium nennt Guion (1977), dass qualifizierte Beurteiler bestätigen und einer Meinung sein müssen, dass der Inhalt durch die Items gut repräsentiert wird, und fünftens müssen die Ergebnisse verlässlich erfasst und berechnet werden, die Vorgabe des Tests muss also standardisiert sein, genauso wie die Berechnung der Ergebnisse.

1.3.2 Exkurs: Augenscheinvalidität

Inhaltliche Gültigkeit soll nicht mit Augenscheinvalidität verwechselt werden, die ihrerseits meint, dass es der zu testenden Person, also einem Laien, klar ist, was gemessen werden soll (Kubinger, 2009). Inhaltlich gültigen Verfahren wird häufig auch Augenscheinvalidität zugesprochen, umgekehrt müssen augenscheinvalide Tests aber nicht inhaltlich valide sein (Kersting, 2003).

1.4 Kriteriumsvalidität

1.4.1 Definition; Übereinstimmungs- und prognostische Validität

Definition

Ein psychologisch diagnostisches Verfahren ist kriteriumsvalide, wenn von den Testscores erfolgreich auf ein Außenkriterium, also eine als relevant angesehene Variable außerhalb der Testsituation, geschlossen werden kann (vgl. Kubinger, 2009, oder Moosbrugger & Kevala, 2007).

Übereinstimmungs- und prognostische Validität

Man unterscheidet innerhalb der Kriteriumsvalidität zwischen Übereinstimmungsvalidität (konkurrente Validität), wenn das Kriterium bereits in der Gegenwart vorliegt und durch einen anderen Test repräsentiert ist, und Vorhersagevalidität (prognostische Validität), wenn das Kriterium ein Ereignis in der Zukunft ist (vgl. Moosbrugger & Kevala, 2007, oder Kubinger, 2003).

Die Logik der Übereinstimmungsvalidität liegt darin, dass „*bekannt [ist], welches Verhalten der Kriterium-Score repräsentiert. Darum ist der Test-Score in dem Umfang, den die Höhe des Korrelationskoeffizienten angibt, Indikator für das Merkmal, das im Kriterium zum Ausdruck kommt*“ (Fisseni, 2004, S. 66).

Bei der Vorhersagevalidität wird die Übereinstimmung zwischen den zu untersuchenden Testscores und Kriterienscores bestimmt, die ihrerseits erst später erhoben werden (Fisseni, 2004).

Es ist zu beachten, dass bei der Übereinstimmungsvalidität vom Inhalt des Kriteriums auf den Inhalt des Tests geschlossen wird (Fisseni, 2004). Allerdings kann man bei der Vorhersagevalidität nicht aufgrund des Kriteriums auf den Inhalt des Tests schließen, denn bei der prognostischen Validität wird nur festgestellt, ob dieses gemessene Merkmal einen guten Prädiktor für das Kriterium darstellt (Fisseni, 2004).

1.4.2 Überprüfung

In der Regel erfolgt die Prüfung dieses Gütekriteriums durch Bestimmung der Korrelation zwischen den Testergebnissen und der als Außenkriterium ernannten, relevanten Variable (Moosbrugger & Kevala, 2007).

Mögliche Fehlerquellen dabei sind einerseits die Tatsache, dass die Außenkriterien selbst nicht absolut messgenau sind, und andererseits, dass man auf Basis eines Tests nicht ein konkretes, punktuell Geschehen in der Zukunft vorhersagen kann (Kubinger, 2003). Darüber hinaus ist zu beachten, dass Personen über Eigenschaften, Dispositionen verfügen, die lediglich ein bestimmtes Verhalten, eine Handlung, in einem psychologisch diagnostischen Verfahren wahrscheinlicher oder eben unwahrscheinlicher machen, und unvorhergesehene Faktoren die Leistung und damit das Ergebnis beeinflussen können, was lediglich eine Diagnostik der zugrundeliegenden Disposition sinnvoll erscheinen lässt (Kubinger, 2003).

Bei der praktischen Überprüfung der Übereinstimmungsvalidität stellt sich die Frage, wie weit sich die Generalisierung der Test- und Kriteriumsmerkmale erstrecken kann, also wie weit sich die von ihnen getroffenen Aussagen erweitern lassen, sodass man das eine mit Hilfe des anderen erklären kann (Fisseni, 2004). Ebenso ist zu beachten, dass das Ergebnis dieser Überprüfung vom Ausmaß der Erfüllung der Gütekriterien des Tests, der als Außenkriterium dient, abhängt, sowie wenn dieser die Gütekriterien zufriedenstellend erfüllt, wird die Frage nach dem Bedarf für das neue, gerade überprüfte Verfahren aufgeworfen (Kubinger, 2003). Eine mögliche Anwendung kann es sein, wenn zum Beispiel eine aufwändige Arbeitsprobe durch einen Test, eventuell im Multiple-Choice Format, ersetzt werden soll, so würde die konkurrente Validität berechnet werden, um die Gültigkeit des neuen Tests zu überprüfen (Cronbach & Meehl, 1955).

1.4.3 Exkurs: Ökologische Validität

Einen weiteren Aspekt der Validität stellt die ökologische Validität dar (Kubinger, 2009). Ökologisch valide Ergebnisse sind solche, die auch im realen Leben, in der praktischen Anwendung, also außerhalb des Forschungskontextes, gültig sind (Wolf, 1998). Es geht um den Anspruch das natürliche Verhalten unter natürlichen Bedingungen zu erforschen (Wolf, 1998), sodass die Ergebnisse nicht nur im Labor, sondern auch für die Realität außerhalb

repräsentativ sind (Jäger & Petermann, 1999). Das Ziel ist es, ein optimales Gleichgewicht zwischen einer möglichst guten Erfüllung der Gütekriterien, für die eine Objektivität, Reliabilität und Standardisierung der Testsituation erforderlich sind, was unter Laborbedingungen einfacher zu realisieren ist, und einer Repräsentativität und Generalisierbarkeit für die Praxis zu finden, sodass die Daten auf die psychosoziale Umwelt übertragbar sind (Jäger & Petermann, 1999). Auf der einen Seite könnte die artifizielle Laborsituation die Beziehungen zwischen den Merkmalen verzerren, auf der anderen Seite werden die gewonnenen Daten eingeschränkter und weniger generalisierbar (Jäger & Petermann, 1999).

1.5 Konstruktvalidität

1.5.1 Der Begriff „Konstrukt“, Definition und das nomologische Netz

Der Begriff Konstrukt

Nach Anastasi (1986) sind Konstrukte theoretische Konzepte verschiedener Abstraktionsgrade und Generalisierbarkeit, die ein Verständnis der empirischen Daten ermöglichen und erleichtern. Konstrukte werden von empirisch beobachtbarem Verhalten und dessen Stetigkeit abgeleitet, und sie sind gekennzeichnet und definiert durch ein Netzwerk von beobachteten Zusammenhängen (Anastasi, 1986). Einfacher ausgedrückt sind Konstrukte allgemein anerkannte, aber nicht direkt beobachtbare Erscheinungen, wie Intelligenz, Angst oder Stress (Kubinger, 2009). Eine handlungsorientierte Definition von Cronbach und Meehl (1955) erklärt Konstrukte als postulierte Attribute von Personen, die sich in Testergebnissen widerspiegeln, es sind also die Eigenschaften über die wir Hypothesen zur Interpretation des Tests aufstellen.

Ein anschauliches Beispiel für ein Konstrukt zeigen Murphy und Davidshofer (1991), indem sie die Schwerkraft anführen. Wenn ein Apfel vom Baum fällt, liefert das Konstrukt „Schwerkraft“ eine Erklärung und eine Vorhersage dieses Geschehens beziehungsweise Verhaltens. Man kann die Schwerkraft nicht sehen, man sieht nur den fallenden Apfel, analog zu einer Disposition einer Person und ihren Ergebnissen in einem psychologisch-diagnostischen Verfahren (Murphy & Davidshofer, 1991). Trotzdem ist es sinnvoll die Schwerkraft zu erfassen und Theorien auf Basis der Schwerkraft aufzustellen (Murphy & Davidshofer, 1991). Auch erscheint es sinnvoll sich mit dem Begriff der Schwerkraft zu

beschäftigen, anstatt Theorien nur über fallende Äpfel aufzustellen (Murphy & Davidshofer, 1991). Analog dazu ist zum Beispiel das Konstrukt Ehrlichkeit zu verstehen (Murphy & Davidshofer, 1991). Ehrlichkeit kann nicht direkt beobachtet werden, nur anhand von Aussagen oder Taten von Personen, und es ist ein hilfreiches Konzept um menschliches Verhalten zu verstehen, beschreiben und vorherzusagen (Murphy & Davidshofer, 1991).

Definition

Bei der Konstruktvalidität geht es darum zu prüfen, ob ein psychologisch-diagnostisches Verfahren tatsächlich zum Beispiel das Konstrukt, die latente Variable „Intelligenz“, misst oder nicht doch ein anderes Konstrukt stattdessen erfasst wird, wie zum Beispiel Sprachverständnis (Moosbrugger & Kevala, 2007). Wenn ein Test also bestimmte theoriegeleitete Ansprüche in Bezug auf ein Konstrukt erfüllt, dann ist dieser als konstruktvalide zu bezeichnen (Kubinger, 2009).

Mit einem konstruktvaliden Test ist der Schluss von den Testergebnissen auf das zugrundeliegende Persönlichkeitsmerkmal, eine Fähigkeit, eine Disposition, einen Charakterzug oder eine Einstellung möglich (Moosbrugger & Kevala, 2007). Dies wäre eine „erklärende Interpretation“, bei der Aussagen über die individuelle Ausprägung eines Merkmals bei einer Person mit einem bestimmten Testwert möglich sind, und das theoretische Konstrukt wird für die beobachteten Unterschiede zwischen verschiedenen Personen und ihren Testergebnissen verantwortlich gemacht (Hartig, Frey & Jude, 2007). Ebenfalls ist mit einem konstruktvaliden Test eine „extrapolierende Interpretation“ möglich, bei der von den Testergebnissen auf andere Variablen geschlossen wird, mit denen das Konstrukt in einem bestimmten Zusammenhang steht (Hartig, Frey & Jude, 2007).

Das nomologische Netz

Das Konzept der Konstruktvalidität ist konkret die Übereinstimmung zwischen den Ergebnissen im zu validierenden Test und einem Bündel an anderen Scores oder Aussagen (Fisseni, 2004). Diese (latenten) Konstrukte und die beobachtbaren Variablen werden durch bestimmte empirische Gesetze beschrieben und durch Korrespondenzregeln miteinander verbunden, was als „nomologisches Netz“ bezeichnet wird (Moosbrugger & Kevala, 2007).

Dieses besteht aus dem Testverhalten und theoretisch verwandten oder entfernten Konstrukten (Fisseni, 2004). Damit ein Konstrukt akzeptabel und annehmbar ist, muss es in einem nomologischen Netz eingebettet sein (Cronbach & Meehl, 1955). Ein nomologisches Netz kann zum einen beobachtbare Eigenschaften oder Maße miteinander, zum anderen theoretische Konstrukte und beobachtbare Maße, oder verschiedene theoretische Konstrukte verbinden (Cronbach & Meehl, 1955). Für eine Konstruktvalidierung muss jedoch das nomologische Netz mit Beobachtungen verbunden werden und klar definierte Schlussfolgerungen enthalten und aus diesen müssen sich bestimmte Relationen für die Ergebnisse ableiten lassen (Cronbach & Meehl, 1955).

1.5.2 Überprüfung und Gefahrenquellen bei der Überprüfung

Überprüfung

Einleitend soll noch hervorgehoben werden, dass theoretisch alle mit dem Verfahren erhobenen Daten zum Verständnis, was der Test misst, beitragen und mit der praktischen Anwendung diese Erklärung ständig verbessert, verfeinert und bereichert wird (Anastasi, 1986). Daraus ergibt sich einerseits, dass Konstruktvalidierung ein endloser Prozess ist (Anastasi, 1986) und andererseits dass man auf Basis der Konzeption der Konstruktvalidität keine konkrete Handlungsanweisung einer Validierung ableiten kann (Fisseni, 2004). Stattdessen stellt der Begriff der Konstruktvalidität ein Prinzip dar, ein Programm, und bietet dem Forscher verschiedene Herangehensweisen (Fisseni, 2004).

Cronbach und Meehl (1955) unterstreichen, dass Konstruktvalidierung nicht durch einen einfachen, numerischen Koeffizienten bestimmt und interpretiert werden kann, sondern dass durch die erhobenen Daten nur die obere und untere Grenze der Konstruktvalidität bestimmt werden kann, und die Konstruktvalidierung um eine Integration der verschiedenen Daten zu deren Interpretation nicht herum kommt.

Murphy und Davidshofer (1991) stellen eine Erklärung des Konstrukts an den Beginn der Konstruktvalidierung. Diese erfolgt in drei Schritten: Zuerst muss das Verhalten das mit dem Konstrukt zusammenhängt identifiziert und bestimmt werden, dann werden andere Konstrukte gefunden, die entweder mit dem betreffenden Konstrukt zusammenhängen oder nicht, und schließlich werden Verhaltensweisen ermittelt, die mit diesen anderen Konstrukten zusammenhängen und auf Basis der Relationen zwischen diesen Konstrukten wird bestimmt ob die einzelnen Verhaltensweisen mit dem Konstrukt zusammenhängen. Daraus resultiert

schließlich eine genaue Beschreibung der Beziehungen zwischen den Konstrukten und Verhaltensweisen, was als nomologisches Netz bezeichnet wird.

Cronbach und Meehl (1955, S. 287ff) führen verschiedene Methoden zur Konstruktvalidierung an, die im Folgenden kurz plakativ dargestellt werden sollen:

- Berechnung von Differenzen zwischen verschiedenen Gruppen: Diese Methode kommt zur Anwendung, wenn man auf Basis des Konstruktes erwartet, dass sich zwei oder mehr Gruppen unterscheiden.
- Korrelations-Matrizen: Korrelations-Matrizen werden berechnet, wenn erwartet wird, dass der zu validierende Test ein gleiches Konstrukt wie ein anderer misst.
- Faktorenanalyse.
- Überprüfung der inneren Struktur: Dies kommt zur Anwendung, wenn für die Beurteilung der Validität die Homogenität innerhalb des Tests bedeutend ist, es wird mit Interkorrelationen zwischen den Items überprüft, ob sie alle das eine Merkmal erfassen.
- Erfassung von Veränderungen in verschiedenen Situationen.
- Erfassung von Veränderungsprozessen: Eine mögliche Validierungsmethoden für einen Test kann es sein, die Veränderung eines Merkmals über die Zeit, auch zum Beispiel nach einer Maßnahme, oder nach dem Lernen, zu überprüfen.

Lienert beschreibt in Anlehnung an Cronbach und Meehl (1955, zit. nach Lienert, 1989, S. 262) die wichtigsten methodischen Wege:

- Korrelation des Tests mit Außenkriterien.
- Korrelation des Tests mit Tests ähnlichen Validitätsanspruchs.
- Korrelation mit Tests, die andere Persönlichkeitsmerkmale erfassen.
- Faktorenanalyse des zu validierenden Tests gemeinsam mit Außenkriterien, validitätsverwandten und validitätsdivergenten Tests.
- Analyse interindividueller Unterschiede in den Testresultaten.
- Analyse interindividueller Veränderungen bei wiederholter Durchführung mit und ohne systematische Variation der Durchführungsbedingungen.
- Inhaltlich-logische Analyse der Testelemente.

Dieterich (1973) präsentiert eine Aufteilung in theoretische Maßnahmen, empirische Maßnahmen und Maßnahmen, die Theorie und Empirie integrieren und subsummiert damit Wege zur Konstruktvalidierung, die der Überprüfung der inhaltlichen Gültigkeit und der Kriteriumsvalidität entsprechen. Auch Fisseni (2004) und Michel und Conrad (1982) sehen die Konstruktvalidität als Oberbegriff, der die anderen Validitätsarten mit einschließt. Hier sehen die Autoren (Michel & Conrad, 1982) auch den Wert der Konstruktvalidität, und zwar in der Verknüpfung einer pragmatischen Kriteriumsvalidierung, mit wissenschaftstheoretischen Überlegungen auf Basis von Grundlagenforschung, Messtheorien und Testkonstruktion mit einer praktisch-psychologischen Anwendung.

Nach Moosbrugger und Kevala (2007) sind bei der Überprüfung der Konstruktvalidität zwei Ansätze zu unterscheiden; struktursuchende Ansätze, die durch eine exploratorische Faktorenanalyse realisiert werden, und strukturprüfende Ansätze, die durch eine konfirmatorische Faktorenanalyse und eine Multitrait-Multimethod-Analyse bestimmt werden. Diese strukturprüfenden Verfahren, die in der vorliegenden Konstruktvalidierung zur Anwendung kommen, werden in Kapitel 4 erläutert.

Da es jedoch häufig in der Praxis unmöglich ist, im Zuge der Entwicklung eines psychologisch diagnostischen Verfahrens alle zur Verfügung stehenden Validierungsmethoden einzusetzen, empfehlen Hartig, Frey und Jude (2007) die vorherige Überlegung, welche Verwendung und Interpretation für einen Test besonders wichtig erscheint, und dann die sinnvollste Validierungsmethode zu wählen. Die Autoren (Hartig, Frey & Jude, 2007) nennen dabei drei mögliche Interpretationsschwerpunkte und geben jeweils die empfohlene Validierungsmethode an. Wenn der Schwerpunkt eines Tests auf dem Repräsentationsschluss liegt, also man überprüfen möchte, ob die Aufgaben des Tests etwas optimal repräsentieren und damit dann eine Verallgemeinerung zulässig erscheint, so empfehlen die Autoren ausgewiesene Experten ein Urteil abgeben zu lassen (Hartig, Frey & Jude, 2007). Wurde ein Test entwickelt um dann auf Basis dieser Ergebnisse auf in Zukunft liegende Ereignisse zu schließen, wie zum Beispiel einen Auswahltest für Studienbewerber, empfehlen die Forscher (Hartig, Frey & Jude, 2007) den Zusammenhang zwischen den Testergebnissen und den Außenkriterien empirisch zu bestimmen. Und schließlich, wenn bei der Entwicklung eines Tests die Erfassung eines Konstruktes das Ziel darstellte, so befürworten Hartig, Frey und Jude (2007) die Validität durch eine Vorhersage

experimenteller Effekte und korrelativer Zusammenhänge mit anderen theoretisch relevanten Variablen, zu überprüfen.

Die Zielsetzung des in dieser Arbeit zu validierenden Fragebogens (Kurzversion des *Wiener Studieneignungs-Persönlichkeitsinventars*) im Rahmen des **Wiener Self-Assessment Psychologie®** ist die Erfassung der Konstrukte *Organisation, Motivation, Belastbarkeit und Selbstbewusstsein*. Demzufolge ist in diesem Fall laut Hartig, Frey und Jude (2007) die Validierungsmethode der Wahl die Vorhersage experimenteller Effekte und Überprüfung der korrelativen Zusammenhänge mit anderen theoretisch relevanten Variablen.

Gefahrenquellen bei der Überprüfung

Messick (1995, S. 742) weist auf zwei zentrale Gefahrenquellen für eine zufriedenstellende Validität hin: „*construct underrepresentation*“ und „*construct-irrelevant variance*“. „*Construct underrepresentation*“ meint, dass ein Test zu eng konzipiert ist und wichtige Dimensionen oder Facetten des Konstrukts nicht mit einschließt und damit auch nicht erfasst (Messick, 1995). „*Construct-irrelevant variance*“ wird erfasst, wenn das Verfahren zu weit angelegt ist, sodass einerseits auch Facetten anderer Konstrukte erhoben werden und auch durch die spezifische Testkonstruktion andere Variablen, wie bestimmte Antworttendenzen oder Rateeffekte auftreten, die Interpretation des Konstrukts stören (Messick, 1995).

Der Autor (Messick, 1995, S. 742) unterteilt die *construct-irrelevant variance* weiter in „*Construct-irrelevant difficulty*“ und „*Construct-relevant easiness*“. Wenn bestimmte Eigenschaften des psychologisch-diagnostischen Verfahrens es für bestimmte Personen oder Personengruppen unnötig schwer machen die richtigen oder individuell zutreffenden Antworten zu geben, spricht man von *construct-irrelevant difficulty* (Messick, 1995). Ein Beispiel dafür wäre, wenn bei einem Test zum Allgemeinwissen die Fragen sehr kompliziert gestellt werden, sodass auch andere Faktoren wie Lesefähigkeit und Sprachverständnis zwangsläufig mit einspielen (Messick, 1995). *Construct-irrelevant difficulty* führt zu schlechteren Ergebnissen für die Betroffenen, zu einer systematischen Verzerrung der Gesamtergebnisse und der Interpretation, und zusätzlich kann der Test nicht als fair bezeichnet werden (Messick, 1995). *Construct-irrelevant easiness* passiert, wenn irrelevante Hinweise in den Items oder dem Aufgabenformat es manchen Personen erlaubt, richtig oder

angemessen zu antworten, was aber irrelevant in Bezug auf das Konstrukt ist (Messick, 1995). Ein anderer Fall tritt auf, wenn das spezifische Testmaterial absichtlich oder unabsichtlich manchen Testpersonen bekannter ist als anderen (Messick, 1995). Construct-irrelevant easiness führt zu guten Ergebnissen für die einzelnen Personen, aber in Bezug auf eine Gesamtinterpretation des Konstruktes macht es diese schwierig (Messick, 1995).

Allgemein lässt sich sagen, dass „*construct underrepresentation*“ „und *construct-irrelevant variance*“ nicht nur eine Gefahr für die Validität darstellen, sondern auch für die darauf basierenden Entscheidungen und Konsequenzen (Messick, 1995).

Ein weiteres Problem ist die Interpretation der Ergebnisse der Konstruktvalidierung (Kane, 2001). Wenn die Ergebnisse der Untersuchung mit der Theorie übereinstimmen, so gelten die Theorie und die Erfassungsmethode als valide, wenn sie jedoch nicht übereinstimmen, so ist nicht eindeutig ersichtlich, ob das Problem im definierten Konstrukt, im nomologischen Netz oder in Details der Messung liegt (Kane, 2001).

2. Das Wiener Studieneignungs-Persönlichkeitsinventar (WSP; Khorramdel & Maurer, unpubl., siehe auch Khorramdel und Maurer, 2012)

2.1 Allgemeine Beschreibung

Das *Wiener Studieneignungs-Persönlichkeitsinventar* ist ein Persönlichkeitsfragebogen zur Studienwahlberatung und umfasst im **Wiener Self-Assessment Psychologie®** die Skalen Organisation, Motivation, Frustrationstoleranz und Selbstbewusstsein. Die Teilnehmer beantworten die jeweiligen Statements und haben vier Antwortmöglichkeiten zur Wahl („Ja“, „Eher ja“, „Eher nein“ und „Nein“).

2.2 Die Entwicklung des WSP (Maurer, 2010)

Erste Schritte bei der Entwicklung

Das *Wiener Studieneignungs-Persönlichkeitsinventar* wurde im Rahmen der Diplomarbeit von Maurer (2010), siehe auch Khorramdel und Maurer (2012), entwickelt, und zwar wurde ein erster Itempool für diesen erstellt. Die Erstellung erfolgte nach den Anforderungen der

DIN 33430, welche 2002 vom Deutschen Institut für Normierung e. V. festgelegt wurde (zit. n. Reimann, 2005).

Die Basis für die Entwicklung des *Wiener Studieneignungs-Persönlichkeitsinventars* legten Khorramdel und Maurer (2012) mit einer Anforderungsanalyse zur allgemeinen Studierfähigkeit nach der Methode der Critical Incident Technique (Flanagan, 1954). Dieses Anforderungsprofil (Khorramdel & Maurer, 2012) diente als Ausgangspunkt für die Bestimmung und Definition der Eigenschaften, die mit dem WSP erfasst werden sollten.

Zur Bestimmung, welche Skalen des WSP für das Psychologiestudium relevant sind, wurde auch das Anforderungsprofil von Frebort und Kubinger (2006, S. 411) herangezogen, welches sich rein auf die Anforderungen des Psychologiestudiums konzentrierte. Mit Hilfe dessen sollten aus dem Itempool schließlich diejenigen Items ausgewählt werden, die eine passende Kurzversion für das **Wiener Self-Assessment Psychologie®** darstellten.

In Tabelle 1 sind in der linken Spalte die induktiv ermittelten Kategorien für die Verhaltensnennungen zu sehen, in der rechten die entsprechenden Kompetenzen des Anforderungsprofils:

Kategorien (induktiv ermittelt)	Anforderungsprofil
Motivation	Hohe Leistungsmotivation und Zielorientierung
Soziale Kompetenz	Realistisches Konzept eigener Fähigkeiten
Organisationsfähigkeit	Planerisches Vorgehen
Lernverhalten	Effizienter Lernstil
Intelligenz	Fähigkeit zum folgerichtigen Denken
Belastbarkeit	Hohe Frustrationstoleranz, Durchhaltevermögen
Arbeitsstil	Angemessenes Anspruchsniveau, selbständige Informationssuche
Kontrollüberzeugungen	Leistungsförderlicher Attributionsstil
Studienbezogene Rahmenbedingungen	---

Tabelle 1: Anforderungsprofil von Frebort & Kubinger (2006, S. 411)

An drei Universitäten Wiens (Universität Wien, Wirtschaftsuniversität Wien und Technische Universität Wien) wurden so für insgesamt sieben Studienrichtungen Anforderungsprofile erstellt, die schließlich zusammengefasst wurden zu einem standortspezifischen Anforderungsprofil für die allgemeine Studierfähigkeit in Wien.

Zur Anforderungsanalyse wurde dann die Methode der Critical Incident Technique (Flanagan, 1954) verwendet, es wurde ein Fragebogen für StelleninhaberInnen erstellt, in diesem Fall für die Studierenden, und ein Interviewleitfaden für Experteninterviews mit Vorgesetzten, in diesem Fall für die Lehrpersonen. Die genauen Inhalte kann man bei Maurer (2010) und Khorramdel und Maurer (2012) nachlesen.

Nach Durchführung, Auswertung und Interpretation wurde das resultierende Anforderungsprofil mit dem von Frebort und Kubinger (2006) zusammengefasst und enthielt in der Endfassung 12 Konstrukte mit insgesamt 75 Merkmalen; diese können bei Maurer (2010) nachgelesen werden.

Nach Maurer (2010) ist es an dieser Stelle wichtig zu bemerken, dass mit diesem Anforderungsprofil nicht spezifische Anforderungsprofile für einzelne Studienrichtungen ersetzt werden wollen, und damit auch kein Rückschluss auf einzelne Studienrichtungen möglich ist. Das Anforderungsprofil hat zum Ziel die zentralen Anforderungen in inhaltlich unterschiedlichen Studienrichtungen aufzulisten, was eine Spiegelung der Stärken und Schwächen von möglichen Studierenden erlaubt, sowie auf die Studienreife schließen lässt. Der Rückschluss auf einzelne Studienrichtungen wird nicht empfohlen, da in manchen Studienrichtungen einzelne Bereiche wichtiger sind als in anderen, deswegen ist für die definitive Auswahl für ein Studium oder einen Studierenden eine spezifische Auswahl der Fähigkeiten nötig.

Erstellung eines Regelkatalogs zur Fragebogen-Itemkonstruktion

Als nächster Schritt folgte die Erarbeitung eines Regelkatalogs zur Formulierung der Items. Zuvor sollte aber entschieden werden, welches Fragen- und Antwortformat verwendet werden sollte. Als Itemformat wurden Beurteilungsaufgaben gewählt, und die Items sollten sowohl eine vierstufige Ratingskala als auch eine Analogskala bei der Beantwortung ermöglichen, wobei schließlich, aufgrund der geplanten Überprüfung des Itempools mit dem dichotom logistischen Modell nach Rasch (1960), eine vierstufige verbale Ratingskala gewählt wurde, da in diesem Modell eine Dichotomisierung der Itemantworten vorausgesetzt wird.

Der resultierende Regelkatalog wurde auf der Basis der bestehenden Literatur, den Ergebnissen der Experteninterviews, die mit Personen geführt wurden, die schon aktiv an der Erstellung von Items für Fragebögen beteiligt waren, und den empirischen Erfahrungen bei

der Entwicklung des WSP, erstellt, besteht aus insgesamt 23 Punkten und ist bei Maurer (2010) genauer nachzulesen.

Auswahl der Skalen und Subskalen, sowie genaue Arbeitsdefinition

Als weiteren Schritt war die Auswahl der im *Wiener Studieneignungs-Persönlichkeitsinventar* zu erfassenden Merkmale. Da sich das Anforderungsprofil aus Persönlichkeitseigenschaften, kognitiven Fähigkeiten, psychologischen und nicht-psychologischen Merkmalen zusammensetzt, sollten nur Merkmale durch den WSP erfasst werden, die sinnvollerweise durch einen Persönlichkeitsfragebogen erfasst werden können; also Persönlichkeitseigenschaften und psychologische Merkmale. Um dann schließlich nur mehr die relevantesten Eigenschaften zu wählen, wurde ein Cut-Off Wert festgesetzt, über dem die Anzahl der Nennungen der StelleninhaberInnen, also Studenten, liegen musste, sodass sieben Skalen mit 22 Subskalen resultierten. Diese werden im Folgenden angeführt:

Skalen	Subskalen
Organisationsfähigkeit	Zeitmanagement, Prioritäten setzen, Planen, Organisationstalent, Informiertheit
Arbeitsstil und Arbeitsverhalten	Gewissenhaftigkeit, Genaues Arbeiten, Eigeninitiative, Selbständiges Arbeiten
Motivation	Durchhaltevermögen, Lernbereitschaft, Engagement, Zielorientierung, Zielsetzung, Erfolgszuversicht
Belastbarkeit	Frustrationstoleranz
Soziale Kompetenz	Kommunikationsfähigkeit, Kontaktfähigkeit und Teamfähigkeit (Teamfähigkeit, Teamorientierung, Kollegialität und Kompromissbereitschaft)
Emotionale Kompetenz	Distanzierungsfähigkeit und Offenheit
Selbstbild	Selbstbewusstsein

Tabelle 2: Skalen und Subskalen des WSP (Maurer, 2010)

Eine genaue inhaltliche Beschreibung der Subskalen kann bei Maurer (2010) und Khorramdel und Maurer (2012) nachgelesen werden.

Die Zielgruppe sind StudienbeginnerInnen und StudentInnen die das Studium wechseln möchten und für sich persönlich eine Beratung in Bezug auf ihre Stärken und Schwächen

suchen. Es geht nicht darum, den WSP in Selektionssituationen anzuwenden, sondern um einen förderdiagnostischen Ansatz: es sollen die individuellen Stärken und Schwächen erfasst werden und somit soll die Beratung der StudienbeginnerInnen verbessert werden. Die Teilnehmer sollen sich selbst beurteilen, im Rahmen eines Computertests. Bezüglich der Testlänge wurden zehn bis 20 Items pro Subskala angestrebt. Das Antwortformat sollte vierkategorial mit den Antwortmöglichkeiten „ja“, „eher ja“, „eher nein“ und „nein“ enthalten, wobei bei der Verrechnung für die Antworten „ja“ und „eher ja“ ein Punkt vergeben werden sollte, für die Antworten „eher nein“ und „nein“ werden keine Punkte vergeben.

Erstellung des Itempools

Zur Qualitätssicherung der zu erstellenden Items wurde einerseits der Regelkatalog zur Fragebogen-Itemkonstruktion verwendet, und weiter wurde ein Expertenteam aus Mitarbeitern des Instituts für Entwicklungspsychologie und Diagnostik und DiplomandInnen vom Arbeitsbereich für Diagnostik, zusammengestellt, das im Rahmen von webbasierten Feedback- und Kontrollschleifen die einzelnen Items begutachtete (siehe Maurer, 2010).

Der resultierende Itempool setzt sich aus 351 Items zusammen, die sich auf die oben beschriebenen 22 Subskalen und sieben Skalen aufteilen. Dieser Itempool kann in der Diplomarbeit von Maurer (2010) nachgelesen werden.

Resultierende Kennwerte

Es werden prinzipiell für alle Skalen und Subskalen, sowie genaue Arbeitsdefinition, erwähnten Skalen und Subskalen Testkennwerte berechnet. Diese resultierenden Werte werden so berechnet, indem jeweils für zwei Antwortmöglichkeiten (je nach Polung positiv oder negativ), ein Punkt vergeben wird. Daraus ergeben sich insgesamt vier Kennwerte für die Skalen Organisation, Motivation, Belastbarkeit und Selbstbewusstsein.

Skalierung

Der Itempool, bestehend aus 351 Items wurde von Dörner (2009) einer Prüfung zur Rasch-Modell-Konformität unterzogen, die eine Gültigkeit des Rasch-Modells für die

übergeordneten Skalen ergab. Im Zuge dessen wurden auch nicht Rasch-Modell konforme Items entfernt, sodass schließlich 287 Rasch-Modell konforme Items übrig blieben (Dörner, 2009), aus denen dann schließlich die betreffenden 53 Items für das **Wiener Self-Assessment Psychologie**® ausgewählt wurden (Khorramdel & Maurer, 2012).

3. Das **Wiener Self-Assessment Psychologie**®

3.1 Das diagnostische Instrument des Self-Assessments

Unter einem Self-Assessment versteht man facheinschlägige psychologisch-diagnostische Verfahren, die als Selbsttests durchgeführt werden, ohne Beaufsichtigung und/oder Kontrolle durch einen Testleiter (Kubinger, Moosbrugger, Frebort, Jonkisz & Reiß, 2007). Charakteristisch ist, dass die Personen die Testung aus eigenem Antrieb und im eigenen Interesse durchführen (Kubinger, 2010), mit dem Ziel die Frage, inwieweit man in Bezug auf bestimmte Anforderungen für einen Beruf oder eine Ausbildung geeignet ist, zu beantworten (Kubinger, Moosbrugger, Frebort, Jonkisz & Reiß, 2007).

Self-Assessments sind von Fachpsychologen entwickelt, sie wurden hinsichtlich der üblichen Testgütekriterien überprüft, ihre Durchführung erfolgt nach einer psychologisch ausgereiften Instruktion, und auf Basis einer Berechnung und Interpretation der Testwerte erfolgt eine berufs- oder ausbildungsbezogene Eignungsbeurteilung (Kubinger, Moosbrugger, Frebort, Jonkisz & Reiß, 2007).

Konkreter werden Self-Assessments als internetbasierte Selbsttests, also Online Self-Assessments, StudieninteressentInnen angeboten, sie enthalten bestimmte fach- und/oder standortspezifische Fragen und Aufgaben, und es folgt eine automatisch erstellte, individuelle Rückmeldung (Rudinger & Hörsch, 2009).

Wichtig zu erwähnen ist, dass es sich nicht um Instrumente zur Studierendenauswahl im Sinne einer Zulassung oder Nicht-Zulassung zu dem Studium handelt, sondern um eine Entscheidungshilfe für die Personen selbst, in der der teilnehmenden Person die individuelle Passung für das Studienfach aufgezeigt werden soll (Rudinger & Hörsch, 2009). Self-Assessments dienen also nur der persönlichen Auskunft über die eigenen Stärken und Schwächen in Bezug auf das interessierende Studienfach, weswegen auch die leicht verfälschbaren Persönlichkeitsfragebögen eingesetzt werden können (Kubinger, 2010).

Self-Assessments zeichnen sich ebenfalls dadurch aus, dass sie ein breites Spektrum an Variablen berücksichtigen, es werden kognitive Aspekte, also zum Beispiel logisches Denken, mathematisches Verständnis, Textverstehen, und nicht-kognitive Aspekte, wie Interesse, Motivation oder Selbstwirksamkeit, durch verschiedene Verfahren erfasst (Zimmerhofer, Heukamp & Hornke, 2006).

3.2 Entwicklung des **Wiener Self-Assessment Psychologie®**

Das **Wiener Self-Assessment Psychologie®** wurde von der Test- und Beratungsstelle des Arbeitsbereichs Psychologische Diagnostik an der Fakultät für Psychologie an der Universität Wien entwickelt, angeboten und wurde und wird laufend evaluiert (Sonnleitner, Kubinger & Frebort, 2009).

Gemäß DIN 33430 (Kersting, 2005, zit. n. DIN Deutsches Institut für Normierung e. V., 2002) müssen die eingesetzten psychologisch-diagnostischen Verfahren nach einem Anforderungsprofil ausgewählt werden, das wiederum nach einschlägigen Methoden erstellt wurde, und die Verfahren müssen auch den üblichen Gütekriterien (vgl. Kubinger, 2010) entsprechen.

Das von Frebort und Kubinger (2006, S. 411) mit der Methode der Critical Incident Technique nach Flanagan (1954) erstellte Anforderungsprofil, kann in Tabelle 1 nachgesehen werden.

Für die Punkte „Realistisches Selbstkonzept eigener Fähigkeiten“, „Angemessenes Anspruchsniveau“, „Effizienter Lernstil“, „Reasoning“ (Fähigkeit zum folgerichtigen Denken) und „Belastung“ (Belastbarkeit: hohe Frustrationstoleranz, Durchhaltevermögen) kommen psychologisch-diagnostische Verfahren zum Einsatz, die der experimentalpsychologischen Diagnostik entsprechen, also objektive Persönlichkeitstests (siehe Kubinger, 2009), die persönlichkeitsabhängige Fähigkeiten erfassen sollen, die mit klassischen Persönlichkeitsfragebögen nicht problemlos, beziehungsweise unverfälscht, erhoben werden können (Sonnleitner, Kubinger & Frebort, 2009).

Im Folgenden wird auf die Eigenschaften, die mit Verfahren der experimentalpsychologischen Diagnostik erfasst werden, näher eingegangen.

Realistisches Selbstkonzept eigener Fähigkeiten:

Um festzustellen wie sehr die potentiellen PsychologiestudentInnen zur systematischen Über- oder Unterschätzung der eigenen Fähigkeiten neigen wurde in Anlehnung an den Lexikon-Wissens-Test (LEWITE, Wagner-Menghin, 2004) von Frebort (unpubl.) der Wissenstest Psychologie (WITE) entwickelt.

Der genaue Testablauf und die Testkennwerte werden im Kapitel 3.4.3 im Theoretischen Teil erläutert. Es soll hier noch erwähnt werden, dass auf diese Art das relevante Vorwissen in Bezug auf das Psychologiestudium erfasst werden soll und auch eine realistische Vermittlung der schwerpunktmäßigen Inhalte im Studium erfolgen soll, ebenso wie das Selbstkonzept der eigenen Fähigkeiten, was bedeutet über eine realistische Einschätzung über das vorhandene Wissen zu haben (Sonnleitner, Kubinger & Frebort, 2009).

Angemessenes Anspruchsniveau:

Das Anspruchsniveau ist definiert als „*diejenige Leistungsgüte, die eine Person bei einer Aufgabe explizit zu erreichen versucht und die sie mindestens erreichen muss, um zufrieden mit dieser Leistung zu sein*“ (Rudolph, 2003, S. 108). Das bedeutet, dass das individuelle Erleben von Erfolg oder Misserfolg nicht vom objektiven Ausmaß und der Güte der erbrachten Leistung abhängt, sondern davon ob die Person die eigenen Ambitionen als erreicht oder nicht erreicht ansieht (Rudolph, 2003).

Bei zu leichten Aufgaben erfährt man kein Gefühl der Befriedigung, bei zu schweren Aufgaben wird Enttäuschung ausgelöst, deswegen wählen die meisten Menschen Aufgaben mit mittlerer subjektiv eingeschätzter Schwierigkeit, was einem realistischen Anspruchsniveau entspricht und zu Stolz führt (Edelmann, 2000, zit. n. Atkinson, 1957).

Demnach erfolgt, auf das Studium bezogen, bei einem niedrigen Anspruchsniveau, die Vorbereitung auf Prüfungen nicht gut genug, weil man sich die Ziele selbst nicht so hoch steckt, und auch Prüfungen werden häufig verschoben, da man an sich nicht so hohe Ansprüche stellt, diese in einer bestimmten Zeit (gut) zu absolvieren (Sonnleitner, Kubinger & Frebort, 2009). Hingegen bei einem zu hohen Anspruchsniveau strebt man zu hohe Ziele an, die schwer zu erreichen sind und dies resultiert zum Beispiel in dem Streben zu viele Prüfungen in einer zu kurzen Zeit absolvieren zu wollen (Sonnleitner, Kubinger & Frebort, 2009).

Das individuelle Anspruchsniveau wird bei Erfolg gehoben und bei Misserfolg gesenkt (Rudolph, 2003). Ein angemessenes Anspruchsniveau wird dann erreicht, wenn auch das Feedback, das man, wie in diesem Fall im Studium bekommt, adäquat verarbeitet (Sonnleitner, Kubinger & Frebort, 2009).

Der Untertest „Symbole kodieren“ der Testbatterie *Arbeitshaltungen* (Kubinger & Ebenhö, 1996), soll das individuelle Anspruchsniveau im Rahmen eines objektiven Persönlichkeitstests erfassen (Sonnleitner, Kubinger & Frebort, 2009). Der genaue Testablauf und die Testkennwerte werden im Kapitel 3.4 im Theoretischen Teil dargestellt.

Effizienter Lernstil:

Ein effizienter Lernstil zeigt sich in der Fähigkeit, mehr oder weniger unstrukturierte Informationen, die auch oft nicht logisch schlüssig sind, zu lernen und sich langfristig zu merken (Sonnleitner, Kubinger & Frebort, 2009). Diese Fähigkeit soll im Self-Assessment Psychologie mittels des LAMBDA-neu (Lernen auswendig, Merken, Belastbarkeit, Denken analytisch, Karolyi, 2009) erfasst werden.

In diesem psychologisch-diagnostischen Verfahren geht es um die Diagnostik von intellektueller Lernfähigkeit, es geht um das herkömmliche (Auswendig-) Lernen von einem unbekannten, simplen und logisch nicht schlüssigen Lernstoffs, dabei wird auch initiiert, dass bestimmte dafür notwendige Bearbeitungsstrategien ausgebildet und angewandt werden (Kubinger, 2009). Die genaue Beschreibung des Testablaufs und der Testkennwerte erfolgt in Kapitel 3.4.1 im Theoretischen Teil.

Reasoning und Belastung:

Reasoning, also schlussfolgerndes Denken, ist „die Fähigkeit, Gesetzmäßigkeiten oder logisch zwingende Zusammenhänge [zu] erkennen und zweckentsprechend verwerten zu können“ (Kubinger, 2009, S. 206).

Der Begriff Belastbarkeit bezieht sich auf die Fähigkeit unter Zeitdruck und mit Aufgabenkollision, also mehreren Anforderungen gleichzeitig, umzugehen und trotzdem noch eine entsprechende Leistung zu erbringen (Sonnleitner, Kubinger & Frebort, 2009). Es ist das Ausmaß, in dem es einer Person gelingt, den Belastungen mit denen sie gerade konfrontiert

wird, standzuhalten, was sich darin zeigt, dass Leistung und Befindlichkeit unbeeinflusst bleiben und die subjektive Beanspruchung niedrig bleibt (Kubinger, 2009).

Diese beiden Punkte sollen ebenfalls mit Hilfe des LAMBDA-neu erfasst werden, wobei hier im Untertest „Rechnen in Symbolen“ (Schmotzer, Kubinger & Maryschka, 1994) das logisch-schlussfolgernde Denken erfasst werden soll, und schließlich im letzten Teil (siehe Kapitel 3.6.1 im Theoretischen Teil) wird eine Beanspruchungssituation provoziert, die dann die Belastbarkeit festhalten soll (Sonnleitner, Kubinger & Frebort, 2009).

Diese drei Verfahren der experimentalpsychologischen Leistungsdiagnostik, WITE, LAMBDA-neu und „Symbole kodieren“ (im Folgenden „Kodiertest“ genannt) aus der Testbatterie *Arbeitshaltungen* erfassen die Punkte Realistisches Selbstkonzept eigener Fähigkeiten, angemessenes Anspruchsniveau, effizienter Lernstil und Reasoning und Belastung.

Die übrigen Punkte im Anforderungsprofil, hohe Leistungsmotivation und Zielorientierung, Planerisches Vorgehen, hohe Frustrationstoleranz/Durchhaltevermögen, selbständige Informationssuche und leistungsförderlicher Attributionsstil, werden dann schließlich vom Persönlichkeitsfragebogen WSP und der persönlichen Einschätzung erfasst.

Schließlich wurden die folgenden Skalen und Subskalen für das Wiener Self-Assessment ausgewählt:

Skalen	Subskalen
Organisation (18 Items)	Zeitmanagement (4 Items)
	Prioritätensetzung (3 Items)
	Planen (4 Items)
	Organisieren (3 Items)
	Informiertheit (4 Items)
Motivation (22 Items)	Durchhaltevermögen (4 Items)
	Lernbereitschaft (4 Items)
	Engagement (4 Items)
	Zielorientierung (3 Items)
	Zielsetzung (3 Items)
	Erfolgsoversicht vs. Misserfolgsängstlichkeit (4 Items)
Belastbarkeit (7 Items)	Frustrationstoleranz (7 Items)
Selbstbild (6 Items)	Selbstbewusstsein (6 Items)

Tabelle 3: Skalen und Subskalen des WSP im **Wiener Self-Assessment Psychologie**® verwendet

3.3 Exkurs: Objektive Persönlichkeitstests

Objektive Persönlichkeitstests, wie sie unter anderem im **Wiener Self-Assessment Psychologie®** zu finden sind, basieren auf der Forderung, aus einem direkt beobachtbaren Verhalten einer Person auf besondere, persönliche (Persönlichkeits-) Merkmale einer Person zu schließen (Ortner, Horn, Kersting, Krumm, Kubinger, Proyer, Schmidt-Atzert, Schuhfried, Schütz, Wagner-Menghin & Westhoff, 2007). Die Praxis sieht dabei meist so aus, dass in einer standardisierten, computerisierten Testsituation auf Basis des (Leistungs-)Verhaltens der Testperson auf die zugrundeliegenden Persönlichkeitsmerkmale geschlossen wird, dies ganz ohne Selbstbeurteilung, und somit auch ohne Introspektionsfähigkeit und sozialer Erwünschtheit (Ortner et. al., 2007).

Aus diesen Punkten folgt, dass objektive Persönlichkeitstests nicht zielgerichtet verfälscht werden können, da der Testperson das Messprinzip nicht bekannt ist und somit nicht augenscheinlich klar ist, was dieser Test zu erfassen beabsichtigt (Kubinger, 2006). Dies ist der klare Vorteil den objektive Persönlichkeitstests gegenüber Persönlichkeitsfragebogen haben, und zwar die Tatsache, dass sie als unverfälschbar gelten (Kubinger, 2006). Bei Persönlichkeitsfragebogen, bei denen der Testperson die Messintention augenscheinlich klar ist, wird, besonders in Situationen in denen die Person mit Konsequenzen rechnet, verfälscht (Kubinger, 2002, zit. nach Kubinger, 2006).

3.4 Psychologisch diagnostische Verfahren im **Wiener Self-Assessment Psychologie®**

3.4.1 LAMBDA-neu (Kubinger & Maryschka, 2001)

Verfahrensbeschreibung

Lernen auswendig, Merken, Belastbarkeit, Denken-analytisch, kurz LAMBDA-neu wurde von Kubinger und Maryschka (2001) entwickelt und von Haiden (2009) und Karolyi (2009) weiterentwickelt und ist in einer adaptierten Version im **Wiener Self-Assessment Psychologie®** zu finden. Die Informationen dazu wurden aus der Arbeit von Karolyi (2009) entnommen. Es ist ein computergestützter Lern- und Merkfähigkeitstest und erfasst in drei Phasen (Lernphase, Prüfphase, Stressphase mit und ohne Belastung) die Lernfähigkeit, die mittelfristige Merkfähigkeit, die Belastbarkeit und das analytische Denken.

In der ersten Lernphase werden der Testperson fünf Unternehmen in Form von Organigrammen mit insgesamt 20 Informationen vorgegeben. Die Organigramme enthalten die Namen, die Branchen, die Standorte der Unternehmen und die Jahresumsätze, beinhalten also verbale und numerische Informationen. Nun soll sich die Testperson diese Informationen einprägen, ohne zu wissen, wie diese später abgeprüft werden. Dafür hat die Person so viel Zeit, wie sie selbst glaubt zum Einprägen zu benötigen, und leitet dann schließlich selbst die erste Prüfphase ein.

Hier werden der Person die Organigramme wieder vorgegeben, allerdings enthalten sie drei bis sechs falsche Informationen, die die Testperson finden und korrigieren soll, in einem Multiple-Choice Auswahlssystem. Danach beendet die Person die erste Prüfphase selbst, wenn sie glaubt, alle Informationen richtig ausgewählt zu haben, und kommt in die zweite Lernphase, in der sie sich die Organigramme weiter einprägen kann.

Das Ziel ist es, fünf Mal hintereinander die fehlerhaften Organigramme auszubessern, wobei im oberen Eck zu sehen ist, wie viele der fünf Phasen man richtig absolviert hat, indem die jeweilige Anzahl an Häkchen zu sehen ist.

In der Stressphase ohne Belastung folgt der Subtest „Rechnen in Symbolen“ (Schmotzer, Kubinger, Maryschka, 1994). Hier hat die Testperson die Aufgabe, einfache Rechenaufgaben zu lösen, bei denen die Ziffern durch Symbole ersetzt sind, und die Person zuerst erkennen muss, für welche Ziffer das jeweilige Symbol steht, um dann das Rechenbeispiel lösen zu können. Hier wird das rechnerisch-schlussfolgernde Denken erfasst, und in der Phase ohne Stress sorgt dieser Subtest für eine Zeitverzögerung, bevor die Bearbeitung des Organigramms weiter gehen kann.

In der Stressphase mit Belastung steuert das Programm den Wechsel zwischen dem Subtest Rechnen in Symbolen und die Bearbeitung des Organigramms, sodass die Unterbrechungen bei der Testperson Stress auslösen.

Zur besseren Anschaulichkeit sei hier ein Screenshot vom LAMBDA-neu dargestellt (siehe Abb. 1). Zur Vermeidung ungebührlicher Verbreitung des Testmaterials wurden hier einige Informationseinheiten unkenntlich gemacht.

Sie haben bereits 2 min gearbeitet

Mutter	1.Tochter	3.Sohn	4.Tochter
RENAST	MOLGANE	Tampere	CARPELL
RENAST	MOLGANE	Tampere	CARPELL
Tiefziehbleche	2.7 Mrd. EUR	51.7 Mio. Yen	65 Mio. EUR
Tiefziehbleche	2.7 Mrd. EUR	51.7 Mio. Yen	65 Mio. EUR
12 Mio.			
420 Mio. Yen			
Plymouth			
Plymouth			
Papierwaren			
Papierwaren			

Weiter

Abbildung 1: Screenshot LAMBDA-neu (Kubinger, Haiden, Karolyi & Maryschka, 2012, S. 73)

Operationalisierung der relevanten Testkennwerte

Die relevanten resultierenden Kennwerte des LAMBDA-neu sind wie folgt bezeichnet und berechnet (Karolyi, 2009):

Kriterium erreicht vs. nicht erreicht	Eine Testperson erreicht das Kriterium, wenn sie fünf fehlerhafte Organigramme in Folge richtig ausbessert. Dieser Wert spiegelt die Lern- bzw. Testbearbeitungsstrategie wider, die bei Erreichen des Kriteriums erfolgreich war.
Anzahl der Prüfphasen	Dieser Wert zeigt an, wie viele Prüfmasken, Organigramme, die Person insgesamt bearbeitet hat, bevor sie das Kriterium erreicht hat oder bevor diese Phase automatisch wegen Nichterreichen, beendet wurde.
Ungültige Null-Fehler-Durchgänge	Ungültige Null-Fehler-Durchgänge ergeben sich, wenn die Testperson einen Fehler macht, und somit die vorherigen, richtig gelösten Organigramme ungültig werden. Aufgrund einer mangelnden Konsolidierung oder Konzentrationsproblemen kommt es zu ungültigen Null-Fehler-Durchgängen.
Gesamtlerndauer	Diese ergibt sich aus der Zeit, die in der ersten Lernphase verbracht wurde plus der Zeit, die nach jedem Fehler wieder fürs Lernen verwendet wurde. Eine Testperson mit einer guten Merkfähigkeit weist einen geringen Wert bei der Gesamtlerndauer auf.
Relative Lernzeit 1	Die Relative Lernzeit 1 berechnet sich aus der Dauer der ersten Lernphase dividiert durch die Gesamtlernzeit, und stellt einen guten Prädiktor für die Gesamtlernleistung und das Erreichen des Kriteriums dar.

Lernstrategie	Durch Vergleich der Dauer der ersten und zweiten Lernphase mit der Dauer der restlichen Lernphasen ergibt sich ein Indikator für die Lernstrategie, wobei ein hoher Wert (über 0,5) in Kombination mit einem erreichten Kriterium für eine gute Lernstrategie steht.
Belastbarkeitsquotient	Durch dividieren der Gesamtleistung im Subtest Rechnen in Symbolen durch die Leistung in der Stressphase ergibt sich der Belastbarkeitsquotient, der bei geringeren Leistungen unter Stress unter einem Wert von 0,5 liegt.

In der Arbeit von Karolyi (2009) konnten die Lerntypen in vier Cluster aufgeteilt werden, in den impulsiven Lerntyp, den unsicheren Lerntyp, den sich überschätzenden Lerntyp und in den erfolgreichen Lerntyp, die genauen Beschreibungen sind bei Karolyi (2009) nachzulesen.

3.4.2 „Symbole kodieren“ (Kodiertest) (aus *Arbeitshaltungen*, Kubinger & Ebenhö, 1996)

Verfahrensbeschreibung

Der Kodiertest der im **Wiener Self-Assessment Psychologie**® zur Anwendung kommt ist Teil der objektiven Testbatterie *Arbeitshaltungen* (Kubinger & Ebenhö, 1996). Der gesamte Test erfasst in seinen drei Untertests Impulsivität vs. Reflexivität, Entschlussfreudigkeit, Exaktheit, die Leistungsmotivation, und speziell im Untertest „Symbole kodieren“ („Kodiertest“ im **Wiener Self-Assessment Psychologie**®) das Anspruchsniveau, die Zieldiskrepanz und das Leistungsniveau.

Der Test beinhaltet die Aufgabe abstrakte, bunte Symbole anderen abstrakten aber schwarz-weißen Symbolen zuzuordnen. Die Person sieht eine Reihe von vier, sich in der Farbe und Form unterscheidende Symbole, darunter wird jedem Symbol ein abstraktes, schwarz-weißes Symbol zugeordnet. Diese Zuordnung bleibt über alle Durchgänge gleich.

Schließlich erscheint eines der schwarz-weißen Symbole unterhalb dieser Zuordnung, und die Aufgabe der Testperson ist es nun, dieses dem bunten Symbol zuzuordnen, durch Mausklick auf eben dieses, welches sich in einer Reihe unterhalb befindet. In jedem Durchgang hat die Testperson 50 Sekunden Zeit, so viele Symbole wie möglich zu kodieren. Dabei bleiben die zwei Reihen mit der Zuordnung immer am Bildschirm.

Nach jedem der drei Durchgänge kommt ein Feedback über die Anzahl der richtig kodierten Symbole, mit der Anweisung eine Schätzung über die Anzahl der im nächsten Durchgang erwarteten kodierten Symbole abzugeben. In der Vollversion, also nicht im **Wiener Self-Assessment Psychologie**[®], wird der Testperson ab dem dritten Durchgang, vor der Abgabe der Prognose, eine Falschmeldung angezeigt, bei der ihr berichtet wird, dass andere Teilnehmer hier durchschnittlich mehr Symbole kodieren konnten, was eine Frustrationsbedingung darstellt.

In Abbildung 2 ist das Layout des Kodiertests mit den Symbolen zu sehen. Zur Vermeidung ungebührlicher Verbreitung des Materials wurden hier einige Informationseinheiten unkenntlich gemacht.

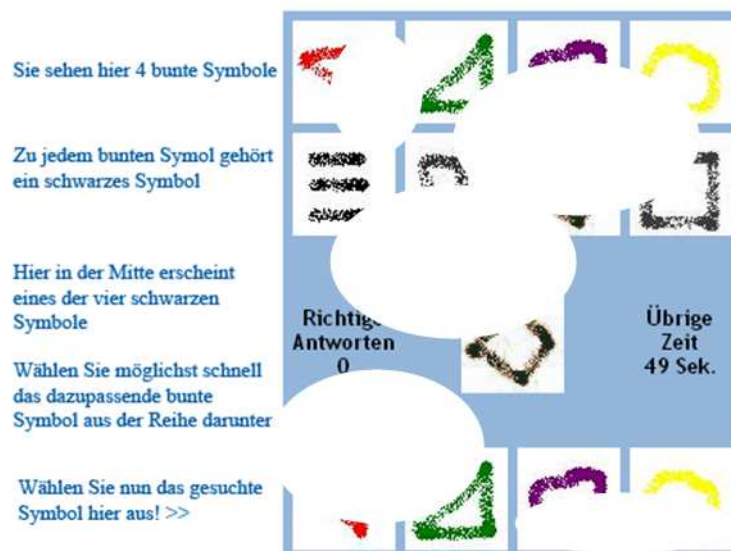


Abbildung 2: Screenshot Kodiertest (Kubinger & Ebenhö, 2012, S. 95)

Operationalisierung der relevanten Testkennwerte

Die berechneten Testwerte aus dem Kodiertest sind folgende (Kubinger & Ebenhö, 1996):

Anspruchsniveau	Berechnet wird dieser Wert aus der ersten Prognose minus der Anzahl der richtigen Kodierungen im zweiten Durchgang, dies dann dividiert durch die Anzahl der richtigen Kodierungen im zweiten Durchgang.
Zieldiskrepanz	Alle Abweichungen zwischen Prognose und folgender Leistung werden in absoluten Werten addiert und ergeben die Zieldiskrepanz, also die Fähigkeit zur Selbsteinschätzung.

Zeitpunkt der Leistungsspitze	Hier resultiert jener Durchgang, bei dem die Testperson die größte Anzahl an Symbolen richtig kodiert hat und sollte möglichst gegen Ende sein. Dieser Wert gibt Auskunft über die Frustrationstoleranz, weil Personen mit einer hoch ausgeprägten Frustrationstoleranz im Zuge des Untertests, unter Einbezug des Übungseffektes, ihre beste Leistung gegen Ende erreichen.
--------------------------------------	--

Zusätzlich wird für den Kodiertest die Summe aller richtig kodierten Symbole berechnet, diese Variable heißt „**AHAInsgesamtrichtigbearbeitet**“.

3.4.3 WITE (Frebort, unpubl.)

Verfahrensbeschreibung

Der Wissenstest Psychologie (WITE), wie er im **Wiener Self-Assessment Psychologie**® zu finden ist, basiert auf dem von Wagner entwickelten Lexikon-Wissen-Test (LEWITE), (Wagner, 1999), und ist ein computergestütztes Verfahren zur Erfassung der Intelligenzdimension „Wortschatz“ und des Persönlichkeitsaspekts „Selbsteinschätzung“ und wurde von Frebort (unpubl.) zu einem Wissenstest rein in Bezug auf das spezifische Wissen im Bereich Psychologie für die Anwendung im **Wiener Self-Assessment Psychologie**® adaptiert.

Das Verfahren ist aus zwei Teilen aufgebaut, zuerst kommt die „Wortliste“, bei der die Testperson antworten soll, ob ihr ein bestimmtes Wort bekannt ist und ob sie es verstanden hat, beziehungsweise erklären kann. Unabhängig von der Antwort auf dieses Element kommt der zweite Testteil, „Lückentext“, in dem die Testperson aufgefordert wird, das betreffende Wort mit Hilfe eines zweiteiligen Multiple-Choice-Lückentextes zu erklären. Das bedeutet, auch wenn eine Person im ersten Teil geantwortet hat, sie kenne das Wort nicht, so wird sie trotzdem aufgefordert, dieses zu definieren zu versuchen.

Als Beispiele lassen sich die Begriffe Abwehrmechanismen, Neuron, Klassische Konditionierung und Hormon zu nennen. Insgesamt enthält das Verfahren 24 Begriffe, also Items. Ein Item ist in Abbildung 3 dargestellt, es ist ein Screenshot aus dem **Wiener Self-Assessment Psychologie**®. Zur Vermeidung ungebührlicher Verbreitung wurden des Materials hier einige Informationseinheiten unkenntlich gemacht.

Unter dem Halo-Effekt versteht man die Tendenz , sich einer Person leiten zu lassen.

mögliche Antworten

- eines Beamtens
- eines Kindes
- eines Kindes
- eines Kindes
- von Live-Events oder ungewöhnlichen Schicksalsschlägen

Abbildung 3: Screenshot WITE (Frebort, Gleeson & Weitensfelder, 2012, S. 84)

Operationalisierung der relevanten Testkennwerte

Die relevanten Kennwerte des WITE sind wie folgt berechnet und beschrieben (Frebort, unpubl):

Gelöst und gewusst/ Aktivwissen:	Dieser Wert errechnet sich aus allen richtig ausgefüllten Lückentexten, bei denen aber davor angegeben wurde, dass man diese kennt. In Wagner (1999) wird dieses Wissen als „aktives Wissen“ bezeichnet, also Wissen über das eine Testperson verfügt, und sich dessen auch bewusst ist.
Gelöst und nicht gewusst/ Passivwissen:	Wenn eine Testperson im ersten Teil angibt, einen Begriff nicht zu kennen, dann den Lückentext aber doch fehlerfrei vervollständigen kann, so zählt das zu dem „passiven Wissen“ der Person. Die Person unterschätzt sich also, und gibt lieber an, einen Ausdruck nicht zu kennen.
Nicht gelöst und gewusst/ Scheinwissen:	Dieser Kennwert berechnet sich aus den Items bei denen die Testperson angibt, den Begriff zu kennen, den Lückentext dann aber fehlerhaft bearbeitet. Personen mit hohen Werten hier überschätzen ihr Wissen und ihre Fähigkeiten.

3.4.4 FIP – Fragebogen zum Interesse im Psychologiestudium (Undeutsch, 2010, unpubl., adaptiert 2010 von Useini) und SKIP – Selbstkonzept im Psychologiestudium (Useini, in Vorb.)

Die beiden psychologisch diagnostischen Verfahren FIP und SKIP sind im **Wiener Self-Assessment Psychologie®** zu einer Einheit zusammengefasst und sollen nun im Folgenden kurz beschrieben werden, der interessierte Leser möge bei Useini (in Vorb.) nachlesen.

Wie aus dem Namen hervorgeht, zielt der Fragebogen zum Interesse im Psychologiestudium (FIP) auf die Erhebung des Interesses für die Inhalte im Psychologiestudium ab und richtet sich an die Teilnehmer des **Wiener Self-Assessment Psychologie®**. Das Konzept des Fragebogens ist dergestalt, dass es sich um einen subjektiven Interessensfragebogen handelt. Das Verfahren wurde von Undeutsch (2010) entwickelt und von Useini im Jahr 2010 (Useini, unpubl.) adaptiert.

3.4.5 FMT – Farbiger Matrizentest (Undeutsch, 2010)

Der Farbige Matrizentest wurde für die Anwendung im **Wiener Self-Assessment Psychologie®** konstruiert. Demzufolge ist er als Computertest entwickelt worden.

Das Ziel des Tests ist die Messung des logisch schlussfolgernden Denkens, indem Items vorgegeben werden, die sich aus neun Feldern zusammensetzen, die eine Matrix bilden. Ein Feld dabei bleibt frei, und die Aufgabe der Testperson ist es, in dieses Feld eine logisch sinnvolle Figur einzusetzen. Im Gegensatz zur üblichen Vorgabe, bei der das einzusetzende Feld rechts unten ist, ist der FMT so konzipiert, dass die gesuchte Figur sich in der Mitte der Matrix befindet. Die Figuren die einzusetzen sind, sind farbig.

Das verwendete Antwortformat ist das Multiple-Choice Format, dabei werden der Testperson sieben Distraktoren und eine richtige Lösung präsentiert, also „eins aus acht“. Es gibt hier weder die Antwortmöglichkeit „ich weiß es nicht“ noch „unlösbar“.

Der FMT wird den TeilnehmerInnen im **Wiener Self-Assessment Psychologie®** mit acht Items vorgegeben, zusätzlich gibt es zwei Warm-Up-Items, die den Teilnehmern das Prinzip des Verfahrens nochmal nahe bringen.

Es resultiert ein Gesamtwert, für den alle richtig gelösten Items mit 1 kodiert werden und dann addiert werden.

3.4.6 Das Wiener Studieneignungs-Persönlichkeitsinventar (WSP; (Khorramdel & Maurer, unpubl.)

Das *Wiener Studieneignungs-Persönlichkeitsinventar* ist in diesem Kapitel nur der Vollständigkeit wegen angeführt, er wurde bereits im Kapitel 3 Das Wiener Studieneignungs-

Persönlichkeitsinventar (WSP) im Theoretischen Teil beschrieben. Die Items sind im Anhang im Kapitel 1 zu finden.

3.4.7 Persönliche Einschätzung

Diese abschließenden Fragen dienen der Erfassung des leistungsförderlichen Attributionsstils und der persönlichen Einschätzung der eigenen Passung zum Psychologiestudium an der Universität Wien. Die Aufgabe der Teilnehmer ist es hier, jeweils einen Prozentwert zwischen 0 und 100 Prozent einzufügen (in Zehnerschritten).

Ich denke... %, der anderen StudienanfängerInnen ist ein erfolgreicher Studienabschluss wichtiger als mir.

Die Wahrscheinlichkeit, dass ich mein Studium erfolgreich zu Ende bringe, beträgt%.

Wenn ich ein Ziel erreiche, so liegt dies zu ... % an meinen eigenen Fähigkeiten.

Erfolg hängt zu ... % davon ab, wie sehr ich mich anstrengte.

Mein Studienerfolg wird zu ... % von anderen Personen beeinflusst.

Glück und Zufall beeinflussen zu ... % meinen Studienerfolg.

Ich denke, meine Interessen sind zu ... % in diesem Studium abgedeckt.

Ich bringe mehr günstige Eigenschaften mit als ... % der anderen StudienanfängerInnen.

Weitere zwei Fragen sollen die Pläne nach einem Studienabschluss erfassen, sowie auch einen „Plan B“, wenn man das Psychologiestudium dann doch nicht absolvieren kann.

3.5 Bisherige Forschungsergebnisse bezüglich Validität und Akzeptanz zum Wiener Self-Assessment Psychologie®

In der Arbeit von Fritsch (2008) wurde versucht, anhand von später im Studium erreichten Prüfungsnoten eine Validierung des **Wiener Self-Assessment Psychologie®** durchzuführen. Es wird lediglich die Leistung im Untertest des LAMBDA „Rechnen in Symbolen“ als Prädiktorvariable für die späteren Leistungen im Fach Psychologische Methodenlehre und Statistik gefunden (Fritsch, 2008). Auch eine darauf folgende hierarchische Clusteranalyse zur Zuordnung der Teilnehmer zu verschiedenen Typen (Typ 1-Hohes Interesse, Typ 2-Mischtyp, Typ 3-Mischtyp und Typ 4-Überwiegend Schwächen) und dann eine Überprüfung, ob bestimmte Typen dann bei bestimmten Prüfungen erfolgreicher sind, zeigt sich als wenig

erfolgreich (Fritsch, 2008). Das bedeutet, dass aufgrund der Typenzugehörigkeit einer Person keine zuverlässige Aussage über den späteren Studienerfolg, gemessen an den Prüfungsnoten, gemacht werden kann (Fritsch, 2008). Lediglich ein knapp signifikanter Zusammenhang zwischen der Zugehörigkeit zu Typ 2 beziehungsweise Typ 4 und den Prüfungsleistungen in den Fächern Ringvorlesung 1, Ringvorlesung 2 und Ringvorlesung Gesamt konnte gefunden werden, der aber nach der Autorin nicht als gewichtig zu interpretieren ist (Fritsch, 2008). Eine Person, die zu Typ 2-Mischtyp gezählt wird, mit teils Schwächen und Stärken im Profil, besteht die drei oben genannten Prüfungen signifikant häufiger als eine Person die dem Typ 4-Überwiegend Schwächen, zugeordnet wird (Fritsch, 2008).

Eine weitere Validierungsstudie des **Wiener Self-Assessment Psychologie®** wurde von Graß (2011) durchgeführt. Mit Hilfe einer Prototypenanalyse sollten Variablen und Testkennwerte des **Wiener Self-Assessment Psychologie®** gefunden werden, die als prototypisch für exzellente Studenten des Faches Psychologie an der Universität Wien gesehen werden können (Graß, 2011). Die genauen prototypischen Werte können in (Graß, 2011) nachgesehen werden. Es sei hier nur ein kurzer Abriss dargestellt:

Diese Ergebnisse zeigen, dass im Studium exzellente Studenten in den Testkennwerten „LAMBDA-Rechnen in Symbolen-ohne Stress“, „LAMBDA-Rechnen in Symbolen-mit Stress“, „LEWITE-Vorwissen“ und „FIP-Gesamtinteresse“ durchschnittlich hohe Werte zeigen, und über eine gute Fähigkeit zum logisch-schlussfolgernden Denken verfügen, belastbar sind unter Stress, ein hohes Vorwissen haben im Fach Psychologie, über ein hohes Interesse für einzelne Fächer verfügen und über eine hohe Organisationsfähigkeit, Motivation, Frustrationstoleranz und hohes Selbstbewusstsein verfügen.

Bezüglich der Akzeptanz des **Wiener Self-Assessment Psychologie®** wurde in der Untersuchung von Fritsch (2008) gefunden, dass sich die Testbatterie einer hohen Akzeptanz seitens der Teilnehmer erfreuen kann. Die Teilnehmer haben zuerst das **Wiener Self-Assessment Psychologie®** durchgeführt, bevor sie anschließend gebeten wurden einen Fragebogen zur Erfassung der Akzeptanz auszufüllen (Fritsch, 2008).

Für die Mehrheit der Teilnehmer des **Wiener Self-Assessment Psychologie®** gelten die folgenden Ergebnisse (Fritsch, 2008): sie würden das Self-Assessment weiterempfehlen, sie fühlten sich vor der Testung gut mit Informationen versorgt, ebenso wie nach der Testung. Es war ihnen der Zusammenhang zwischen den einzelnen Testteilen und dem Psychologiestudium klar und

sie waren mit dem Ablauf der Testung und den Inhalten zufrieden. Die Mehrheit zeigte sich mit der Ergebnismeldung zufrieden, das Feedback wurde als hilfreich bewertet.

4. Theorie zur statistischen Auswertung der Daten – Methodik

4.1 Einführung in die Faktorenanalyse und Konfirmatorische FA

Die Faktorenanalyse allgemein

Das Ziel der Faktorenanalyse ist die Identifikation von latenten Dimensionen (Fischer, 1974), es geht um die Auffindung einer ordnenden Struktur für einen größeren Variablensatz und darum, die vorhandenen Daten zu reduzieren (Bortz, 2005). Auf Basis der korrelativen Beziehungen zwischen den Variablen werden diese in wenige, voneinander unabhängige Variablengruppen geordnet, sodass man dann daraus schließen kann, dass diese zusammengefassten Variablen gemeinsame Informationen erfassen und die anderen unterschiedliche Informationen (Bortz, 2005).

Das Ergebnis einer Faktorenanalyse: Faktoren

Das rechnerische Ergebnis einer Faktorenanalyse sind Indexzahlen, Ladungen genannt, die Auskunft darüber geben, wie gut die Variable zu einer Gruppe von Variablen passt. Es ergeben sich voneinander unabhängige Faktoren, die die Zusammenhänge zwischen den Variablen einer Gruppe erklären sollen (Bortz, 2005).

In Bortz (2005) wird die Zielsetzung der Faktorenanalyse auf folgende Art verdeutlicht: Wenn man von nur zwei gemessenen Variablen x und y ausgeht, und diese eine hohe Korrelation haben, so kann man sich berechnen, ob diese hohe Korrelation durch eine dritte Variable z erklärt werden kann, und zwar ob diese Variable z die Variable x und/oder y beeinflusst. Diesen Rechenvorgang nennt man partielle Korrelation, und ist zum Beispiel in Bortz (2005) näher nachzulesen.

Wenn es nun der Fall ist, dass Variable z die Variablen x und y beeinflusst, so werden alle drei Variablen hoch miteinander korrelieren, und man hat die Möglichkeit die Variable z aus den beiden anderen herauszupartialisieren. Das Ergebnis sind unbedeutende partielle

Korrelationen weil die Information schon durch Variable z hinreichend gut dargestellt wird (Bortz, 2005). Das verdeutlicht, dass wenn Variablen hoch miteinander korrelieren, die von ihnen erfasste Information sich umso ähnlicher ist, was bedeutet, dass wenn man eine Variable misst, eigentlich die Messung der anderen, bei vorliegender hoher Korrelation, redundant wird (Bortz, 2005).

Bei der Faktorenanalyse wird also auf Basis der Korrelationen zwischen den erfassten Variablen eine „synthetische“ (Bortz, 2005, S. 386) Variable erschaffen, die mit den Variablen möglichst hoch korrelieren soll, diese Variable wird dann als Faktor bezeichnet. Also ist ein Faktor eine theoretische, im Geiste existierende Variable beziehungsweise ein Konstrukt, das miteinander korrelierende Variablen erklären kann. Diese werden zusammengefasst zu einer Gruppe, und die restlichen, nicht zusammenhängenden Variablen werden vom Einfluss dieses ersten Faktors bereinigt und mit ihnen wird eine partielle Korrelation durchgeführt, mit der nicht durch diesen ersten Faktor erklärte Zusammenhänge erfasst werden sollen. Dazu wird ein zweiter Faktor bestimmt, der unabhängig von diesem ersten ist und diese Restkorrelationen relativ gut erklären kann. Dies führt zur Bildung einer zweiten Gruppe und einer weiteren Reduktion der Variablen.

Die konfirmatorische Faktorenanalyse

Ebenso wie bei der exploratorischen Faktorenanalyse wird bei der konfirmatorischen Faktorenanalyse eine Reduktion der Daten vorgenommen, wobei hier dieser Reduktion einer vorher genau definierte Theorie zugrunde liegt (Moosbrugger & Schermelleh-Engel, 2007). Diese Theorie spezifiziert die resultierende Anzahl der Faktoren, die Beziehungen zwischen den Variablen und den Faktoren und ebenso wie Beziehungen zwischen den einzelnen Faktoren untereinander (Moosbrugger & Schermelleh-Engel, 2007). Mit der konfirmatorischen Faktorenanalyse wird somit eine Prüfung der zuvor aufgestellten, theoriebasierten Hypothesen vorgenommen (Moosbrugger & Schermelleh-Engel, 2007). Die zentrale Frage ist, ob mit den empirisch erhobenen Daten das theoretische Modell bestätigt werden kann (Moosbrugger & Schermelleh-Engel, 2007).

Eine konfirmatorische Faktorenanalyse wäre, so wie bei Kubinger und Litzenberger (2003) sicherlich die Methode der Wahl der Auswertung einer MTMM-Matrix, da dies jedoch den Rahmen dieser Diplomarbeit übersteigen würde, soll hier auf die Berechnung dieses Verfahrens verzichtet werden. Es sollte deswegen hier nur auf die korrelationsbasierte

Auswertung eingegangen werden; die Berechnung einer konfirmatorischen Faktorenanalyse soll späteren Untersuchungen vorbehalten bleiben. In dieser Studie soll die Validität anhand von zuvor aufgestellten Hypothesen über etwaige Zusammenhänge zwischen den Kennwerten und Unterschiede zwischen Gruppen, einer MTMM-Matrix und einer exploratorischen Faktorenanalyse beurteilt werden.

4.2 Das Multi-Trait-Multi-Method-Modell

4.2.1 Grundgedanken des Multitrait-Multimethod-Ansatzes/Konvergente und diskriminante Validität

Grundgedanken des Multitrait-Multimethod-Ansatzes

Der Multitrait-Multimethod-Ansatz wurde von Campbell und Fiske (1959) entwickelt. Die Autoren präsentierten dieses Verfahren als Königsweg zur Konstruktvalidierung eines Verfahrens (Campbell & Fiske, 1959). Bis dahin hat man beim Nachweis der Konstruktvalidität eines psychologisch-diagnostischen Verfahrens die Korrelation zwischen dem zu untersuchenden Test und anderen psychologisch-diagnostischen Verfahren, die dasselbe Merkmal erfassen sollen, bestimmt (Schermelleh-Engel & Schweizer, 2007). Es wurden also verschiedene Scores von psychologisch-diagnostischen Verfahren, die postulierten, dasselbe Merkmal zu erfassen, analysiert, dies war vom heutigen Standpunkt die konvergente Validität (Schermelleh-Engel & Schweizer, 2007).

Schließlich haben Campbell und Fiske (1959) begründet warum auch Korrelationen von Messungen von verschiedenen Merkmalen geringer miteinander korrelieren sollten als Messungen desselben Merkmals, sie haben also die Bedeutung der diskriminanten Validität hervorgehoben. *„Diskriminante Validität liegt vor, wenn Messungen verschiedener Konstrukte mit derselben Methode nicht oder nur gering miteinander korrelieren“* (Schermelleh-Engel & Schweizer, 2007, S. 326).

Ergänzend haben Campbell und Fiske (1959) erkannt, dass die bei der Validierung verwendeten psychologisch-diagnostischen Verfahren eine Rolle spielen und der Einfluss dieser Methodeneffekte kontrolliert werden soll. Die Daten enthalten also Informationen über die Ausprägung des Merkmals, das ist die systematische Trait-Methoden-Einheit, und einen Fehleranteil, welcher unsystematisch ist (Campbell & Fiske, 1959). Diese verschiedenen systematischen Varianzquellen, wie zum Beispiel spezielle Charakteristika der Verfahren

oder die Situation in der es durchgeführt wurde, werden als Methodeneffekte bezeichnet (Schermerle-Engel & Schweizer, 2007). Das bedeutet, dass sich das individuelle Ergebnis eines Verfahrens aus der Ausprägung des Merkmals bei der Person und aus Methodeneffekten durch das verwendete Messinstrument zusammensetzt (Schermerle-Engel & Schweizer, 2007). Bei der Konstruktvalidierung soll deswegen auch die verwendete Erfassungsmethode, wie zum Beispiel ein Persönlichkeitsfragebogen, ein objektiver Persönlichkeitstest, Selbst- oder Fremdeinschätzung, berücksichtigt werden (Campbell & Fiske, 1959).

Der Ansatz geht davon aus, dass mit einem Test und damit mit einer bestimmten Methode ein Merkmal erfasst wird, aber es wird nicht „*das Merkmal an sich*“ wirklich erfasst (Fisseni, 2004, S. 69). Die resultierenden Daten enthalten also Informationen über das Merkmal, und auch Angaben über die Methode mit der es erfasst wurde (Fisseni, 2004).

Die Multitrait-Multimethod-Analyse hat es zum Ziel, diese beiden Komponenten zu trennen, und zwar das Merkmal und die Methode, deswegen braucht man dafür zumindest zwei verschiedene Merkmale und zwei verschiedene Erfassungsmethoden, dabei sollen sich sowohl die Merkmale deutlich voneinander unterscheiden als auch die Erhebungsmethoden (Fisseni, 2004). Bezüglich der Erfassungsmethoden sollten strukturell unterschiedliche Methoden verwendet werden, wie zum Beispiel verschiedene Messmethoden, also beispielsweise Fragebogen und Objektiver Persönlichkeitstest, sprachgebundene und sprachfreie Tests, verschiedene Beurteiler, also Selbst- und Fremdbeurteilung, oder die Messung soll zu verschiedenen Zeitpunkten stattfinden, zum Beispiel vor und nach einer bestimmten, (wirkungsvollen) Trainingsmaßnahme (Schermerle-Engel & Schweizer, 2007).

Es wird also getestet, wie gut sich die Erhebungsmethoden zur Erfassung der Konstrukte eignen und ob sich die Merkmale beziehungsweise Konstrukte tatsächlich voneinander unterscheiden (Bortz & Döring, 2006). Als Beispiel nennt Fisseni (2004) drei verschiedene Konstrukte, wie Dominanz, Depression und Offenheit und dazu drei verschiedene Methoden, und zwar Tests, Exploration und Verhaltensbeobachtung.

Konvergente und diskriminante Validität

Die oben als Beispiel genannten Konstrukte Dominanz, Depression und Offenheit sollen über alle drei Methoden hinweg gemessen, hoch miteinander korrelieren, weil das Merkmal ja gleich hoch beziehungsweise niedrig ausgeprägt ist, egal mit welcher Methode es erfasst wird,

dies bezeichnet man als konvergente Validität (Fisseni, 2004). Hingegen sollten aber die Scores für die einzelnen Konstrukte die mit einer Methode erfasst wurden, niedrig miteinander korrelieren, was als diskriminante Validität bezeichnet wird (Fisseni, 2004), also sollten zum Beispiel bei der Erfassung der oben als Beispiel genannten Konstrukte (Dominanz, Depression und Offenheit) mit einem Test niedrig oder gar nicht korrelieren. Das bedeutet, dass eine valide Messung dann zustande kommt, wenn konvergente Validität (Messung eines Konstruktes mit verschiedenen Methoden führt zu hohen Korrelationen) und diskriminante Validität (Messungen unterschiedlicher Konstrukte mit einer Methode und mit verschiedenen Methoden führen zu niedrigen Korrelationen) besteht (Schermerle-Engel & Schweizer, 2007).

4.2.2 MTMM-Matrix – Aufbau und Struktur

Bei der Erstellung der Matrix werden die verschiedenen Methoden auf die Spalten aufgeteilt, die Zeilen ergeben sich aus der Anzahl der erfassten Merkmale (Fisseni, 2004). Es werden alle erfassten Merkmale die jeweils mit allen Methoden erfasst wurden, miteinander korreliert (Schermerle-Engel & Schweizer, 2007).

Die Matrix enthält zwei Arten von Blöcken, die Monomethod-Blöcke, bei denen unterschiedliche Merkmale korreliert werden, die mit der gleichen Methode erfasst wurden, und die Heteromethod-Blöcke, die die Korrelationen der mit verschiedenen Methoden erfassten Merkmale enthalten (Schermerle-Engel & Schweizer, 2007).

Der Monomethod-Block teilt sich wiederum auf in den Monotrait-Monomethod-Block und in den Heterotrait-Monomethod-Block, und der Heteromethod-Block besteht aus dem Monotrait-Heteromethod-Block und dem Heterotrait-Heteromethod-Block (Schermerle-Engel & Schweizer, 2007).

Monotrait-Monomethod-Block:

Die Hauptdiagonale enthält die Reliabilitätskoeffizienten, also die Korrelationen der Merkmale innerhalb derselben Methode (Fisseni, 2004). Die Reliabilitätskoeffizienten werden in diesem Fall den Wert von 1 annehmen, weil hier die Werte mit sich selbst korreliert werden (Bortz & Döring, 2006). Die Hauptdiagonale wird auch als Monotrait-Monomethod-Block bezeichnet, weil hier jeweils ein Merkmal mit einer Methode erfasst und korreliert

wird, was natürlicherweise den Wert von 1 ergibt, wenn man einen Wert mit sich selbst korreliert (Bortz & Döring, 2006).

Monotrait-Heteromethod-Block:

Die Nebendiagonalen enthalten die konvergenten Validitätskoeffizienten, also die Korrelationen der Merkmale A, B, C mit denselben Merkmalen aber mit einer anderen Methode erfasst (Fisseni, 2004). Hier wird also jeweils ein Merkmal mit mehreren Methoden erfasst, deswegen wird dies auch als Monotrait-Heteromethod-Block bezeichnet (Bortz & Döring, 2006).

Heterotrait-Monomethod-Block:

In den Feldern unterhalb der Hauptdiagonale und oberhalb der oberen Nebendiagonalen stehen die Korrelationen zwischen den Merkmalen innerhalb einer Methode, dies sind die diskriminanten Validitätskoeffizienten (Fisseni, 2004). Es werden also mehrere Merkmale mit derselben Methode gemessen und die Messwerte werden dann korreliert, deswegen wird dieser Block auch als Heterotrait-Monomethod-Block bezeichnet (Bortz & Döring, 2006)

Heterotrait-Heteromethod-Block:

Ebenso finden sich in den Feldern unterhalb der Nebendiagonalen die Korrelationen eines Merkmals mit einem anderen Merkmal aber mit einer anderen Methode erfasst, auch die diskriminanten Validitätskoeffizienten (Fisseni, 2004). Hier werden mehrere Merkmale mit unterschiedlichen Methoden gemessen und korreliert, weswegen dieser Block als Heterotrait-Heteromethod-Block bezeichnet wird (Bortz & Döring, 2006).

4.2.3 Die korrelationsbasierte Auswertung der MTMM-Matrix/Kritikpunkte und Probleme der korrelationsbasierten Auswertung

Es besteht die Möglichkeit die MTMM-Matrix auf die ursprüngliche Weise, also korrelationsbasiert auszuwerten, oder mit Hilfe einer Varianzanalyse (Stanley, 1961, zit. n. Schermelleh-Engel & Schweizer, 2007), einer exploratorischen Faktorenanalyse (Schweizer, 2003, zit. n. Schermelleh-Engel & Schweizer, 2007) oder mit einer konfirmatorischen Faktorenanalyse (Moosbrugger & Schermelleh-Engel, 2007).

Im Folgenden soll auf die korrelationsbasierte Methode eingegangen werden, da dies in der vorliegenden Arbeit die Methode der Wahl ist.

Für eine erfolgreiche Multitrait-Multimethod-Validierung auf Basis der Korrelationen werden die folgenden Forderungen gestellt (vgl. Campbell & Fiske, 1959, Fisseni, 2004, Bortz & Döring, 2006, Schermelleh-Engel & Schweizer, 2007):

Nachweis der konvergenten Validität:

Die konvergenten Validitätskoeffizienten sollen sich deutlich von null unterscheiden und eine substantielle Höhe erreichen (Fisseni, 2004). Nach Bortz und Döring (2006) soll der Durchschnitt dieser Monotrait-Heteromethod-Korrelationskoeffizienten statistisch signifikant und bedeutsam größer als null sein. Nach dem Vorschlag von Campbell und Fiske (1959) liegt konvergente Validität vor, wenn die konvergenten Validitätskoeffizienten beziehungsweise deren Mittelwert signifikant größer als null sind/ist.

Nachweis der diskriminanten Validität:

Die konvergenten Validitätskoeffizienten sollen höher sein als die diskriminanten Validitätskoeffizienten innerhalb derselben Methode (Fisseni, 2004). Auch bei Bortz und Döring (2006) wird erklärt, dass dadurch, dass unterschiedliche Merkmale erfasst werden die Korrelationen nicht allzu groß sein sollen, weil diese dann auf Redundanzen in den Konstrukten (die Konstrukte überschneiden sich) oder auf unsensible Messungen (die Messmethode erfasst das Konstrukt nicht genau, sondern vielleicht mehrere) hindeuten (Bortz & Döring, 2006). Auch Campbell und Fiske (1959) haben postuliert, dass die Heterotrait-Monomethod-Korrelationen signifikant kleiner als die Monotrait-Heteromethod-Korrelationen sein sollen. Das bedeutet, dass obwohl mit einer Methode gemessen wird, dürfen die Unterschiede zwischen den Konstrukten verschwinden, hier wird also gezeigt, dass trotz Verwendung einer Operationalisierungsform die gemessenen Konstrukte unterscheidbar sind (Bortz & Döring, 2006).

Ebenso sollen die konvergenten Validitätskoeffizienten höher sein als die diskriminanten Validitätskoeffizienten bei verschiedenen Methoden (Fisseni, 2004). Weiter sollen nach Campbell und Fiske (1959) die Heterotrait-Heteromethod-Korrelationskoeffizienten signifikant kleiner sein als die Monotrait-Heteromethod-Korrelationskoeffizienten. Nach

Bortz und Döring (2006) werden hier die geringsten Korrelationen erwartet, da theoretisch weder dasselbe Konstrukt noch dieselbe Methode zur Messung zum Einsatz gekommen ist.

Die Muster der divergenten Korrelationskoeffizienten sollen innerhalb der gleichen Methode, und auch zwischen den Methoden, ungefähr gleich sein (Schermelleh-Engel & Schweizer, 2007). Das bedeutet, dass die Koeffizienten die Zuordnung der Konstrukte zueinander widerspiegeln sollen (Fisseni, 2004). Schermelleh-Engel und Schweizer (2007) beschreiben, dass es auch hier unterschiedliche Empfehlungen für die Auswertung gibt, fassen aber zusammen, dass am häufigsten überprüft wird ob die Vorzeichen in allen Heterotrait-Block-Korrelationen gleich sind, und überprüft wird, ob die Rangreihe der Korrelationen über alle Teilmatrizen hinweg konstant ist.

Nach Campbell und Fiske (1959) sind konvergente und diskriminante Validität Voraussetzungen für eine zufrieden stellende Konstruktvalidität, oder zumindest muss die gefundene Rangreihe im Nachhinein theoretisch und empirisch plausibel erklärt werden (Bortz & Döring, 2006).

Kritikpunkte und Probleme der korrelationsbasierten Auswertung der MTMM-Matrix:

Als Kritikpunkte führen Schermelleh-Engel und Schweizer (2007) die folgenden Punkte an:

Bei der korrelationsbasierten Auswertung der MTMM-Matrix werden zahlreiche Einzelvergleiche und Häufigkeitsauszählungen von Korrelationskoeffizienten vorgenommen, was nur einen groben Überblick über die Struktur der Variablen verschafft. Da man dabei nicht die Konfidenzintervalle der Korrelationen beachtet, ist es kein zufallskritisches Vorgehen. Als zweiten Punkt nennen die Autoren die Tatsache, dass die Auswertung der Matrix darin besteht, die manifesten Variablen zu analysieren, aber gleichzeitig werden die Kriterien interpretiert, indem auf die zugrunde liegenden latenten Merkmale und Methoden geschlossen wird. Auch das Fehlen von eindeutigen Entscheidungsregeln, ab wann eine konvergente oder divergente Validität als bewiesen gilt, fehlt, was eine Auswertung sehr subjektiv macht.

Ein weiteres Problem, an das man denken sollte ist die Tatsache, dass wenn sich die Methoden systematisch in der Reliabilität unterscheiden mit denen sie die Merkmale messen,

so ergeben sich auch bei Methoden mit geringerer Reliabilität höhere Korrelationen als bei Methoden mit geringerer Reliabilität, was zu bedeutenden Fehleinschätzungen führen kann (Schermerle-Engel & Schweizer, 2007).

Zu den Problemen der Multitrait-Multimethod-Validierung mit Hilfe der Korrelationsmethode gehört die Frage, ab wann eine Korrelation als „substantiell“ zu bezeichnen ist, und ab wann sie sich „deutlich“ von null unterscheidet (Fisseni, 2004). Und ebenso, wie groß muss der Unterschied zwischen den Werten der konvergenten und diskriminanten Validitätskoeffizienten sein, damit der konvergente als höher angesehen wird (Fisseni, 2004). Und nicht zuletzt stellt sich die Frage bei welchen Werten die Muster der divergenten Validitätskoeffizienten als äquivalent gelten, sodass die Zuordnung der Konstrukte zueinander widergespiegelt wird (Fisseni, 2004)?

Ein weiteres Grundproblem stellt die Tatsache dar, dass trotz zufriedenstellender konvergenter und diskriminanter Validität nicht sichergestellt werden kann, dass zweifelfrei das angestrebte Konstrukt mit den Erhebungsmethoden erfasst wird (Bortz & Döring, 2006).

Im anderen Fall, sollten die Kriterien nicht zur Gänze erfüllt werden, so geben Schermelleh-Engel und Schweizer (2007) an, dass dies trotzdem nicht in jedem Fall gegen die Konstruktvalidität spricht, und es liegt am Forscher zu entscheiden ob die Kriterien hinreichend erfüllt werden, für die Beurteilung der Abweichungen gibt es keine verbindlichen Regeln.

III Empirischer Teil

1. Vorangehende Überlegungen und Grundlegendes zur empirischen Untersuchung

Wie bereits oben erwähnt, soll in dieser Arbeit eine Konstruktvalidierung der im **Wiener Self-Assessment Psychologie®** verwendeten Skalen des *Wiener Studieneignungs-Persönlichkeitsinventars* vorgenommen werden. Dies wird mittels Korrelationen der Skalen des WSP mit relevanten Testkennwerten der anderen im **Wiener Self-Assessment Psychologie®** verwendeten Verfahren vorgenommen werden. Ebenso soll eine Multitrait-Multimethod-Matrix berechnet und korrelativ ausgewertet werden, genauso wie eine exploratorische Faktorenanalyse.

Die folgenden Annahmen basieren in Bezug auf den LAMBDA-neu auf den im Manual (Kubinger & Maryschka, 2001) angegebenen Kennwertbeschreibungen beziehungsweise Beschreibungen der Berechnung verschiedener Kennwerte. Ebenso basieren die Annahmen in Bezug auf den Kodiertest auf den im Manual für die *Arbeitshaltungen* beschriebenen Testkennwerten und deren Bedeutungen (Kubinger & Ebenhöf, 1996). Die Hypothesen in Bezug auf den WITE wurden aus der Arbeit von Wawrinec, in der eine Validierung des Überschätzungsparameters des LEWITE vorgenommen wurde (Wawrinec, 2006), abgeleitet.

Die Grundannahme ist, dass erfolgreichere Personen im LAMBDA-neu (in Bezug auf ihre Lernstrategie) höhere Werte in den WSP-Skalen *Organisation* und *Motivation* haben. In Bezug auf den Testkennwert „Belastbarkeitsquotient“ aus dem LAMBDA-neu wird angenommen, dass Personen mit einem besseren Wert des Belastbarkeitsquotienten eine höhere Frustrationstoleranz aufweisen.

Weiter wird angenommen, dass Personen, die im Kodiertest bessere Leistungen erbringen in den WSP-Skalen *Motivation* und *Frustrationstoleranz* höhere Werte aufweisen. Der Grund dafür liegt darin, dass angenommen wird, dass die Aufgabe des Symbole-Kodierens von den TeilnehmerInnen viel Motivation und eine gewisse Toleranz gegenüber eintönigen Aufgaben abverlangt, weswegen angenommen wird, dass diese Personen auch höhere Scores in den WSP-Skalen *Motivation* und *Frustrationstoleranz* aufweisen.

Bezüglich des WITE wird angenommen, dass Personen, die mehr „Scheinwissen“ und „Aktivwissen“ und weniger „Passivwissen“ aufweisen, höhere Werte in der WSP-Skala *Selbstbewusstsein* aufweisen.

Nach diesen Annahmen wurden die Hypothesen formuliert.

1.1 Konkrete Umschreibung der relevanten Dimensionen des WSP

Bevor Hypothesen aufgestellt werden können, muss zu Beginn eine genaue Umschreibung der relevanten Skalen des WSP vorgenommen werden. Dies ist wichtig um eine genaue Beschreibung der Skalen zu haben und im Falle einer erfolgreichen Validierung zu wissen, was die Skalen genau messen.

Die Skalen können wie folgt beschrieben werden:

Organisation:

Organisationsfähigkeit ist, unter Voraussetzung der vorherigen Informationseinholung über die gegebenen Strukturen, die zielführende Planung des Studiums, Semesters und Tages, dies beinhaltet das Organisieren von Informationen und Kontakten, die Gewinnung und Beibehaltung eines Überblicks über die Aufgaben und deren strukturierte und vorausplanende Bearbeitung und Regelung, unter Bedacht der Erkennung und Reihung der Prioritäten und der richtigen Einschätzung der Aufgaben und der gegebenen Zeit und damit der zielführenden und effektiven Planung. Dies beinhaltet die Fähigkeit, sich seine Arbeitszeit und Freizeit einzuteilen, das Vorausplanen und rechtzeitige Organisieren, die zeitliche Koordination von Mehrfachanforderungen und realistische Vorstellungen und Erwartungen vom Studium.

Motivation:

Motivation ist die Setzung und zielstrebige Verfolgung von Ausbildungszielen und Zukunftsplänen, die Meisterung von Aufgaben trotz Schwierigkeiten und die Akzeptanz von Kritik, Verbesserungsvorschlägen und den eigenen Fehlern und die Offenheit gegenüber neuen Herausforderungen, sowie dem Willen sich weiterzubilden und anzustrengen. Dies alles mit der Fähigkeit, bei uninteressanten, langwierigen und schwierigen Aufgaben nicht aufzugeben, sein Ziel nicht aus den Augen zu verlieren und bei vermehrtem Zeitaufwand und Überschreitung des mindestens geforderten Anspruchs, seine Aufgaben zu erfüllen mit der Zuversicht, dass die eigene Handlung erfolgreich sein wird.

Belastbarkeit bzw. Frustrationstoleranz:

Belastbarkeit (Frustrationstoleranz) ist die Eigenschaft, bei auftretenden Problemen, frustrierenden Erlebnissen, Kritik oder negativem Feedback, nicht gleich aufzugeben, trotzdem sein Ziel zu verfolgen und sich seine Leistungsbereitschaft und Motivation beizubehalten.

Selbstbild bzw. Selbstbewusstsein:

Selbstbewusstsein ist das Stehen zur eigenen Meinung und zu eigenen Ideen, es ist das an sich selbst Glauben und sich nicht verunsichern lassen und nicht an sich selbst zweifeln. Im Universitätskontext ist es weiter der Mut, sich in Lehrveranstaltungen zu melden, bei Unklarheiten nachzufragen und andere auf ihre Fehler aufmerksam machen.

1.2 Hypothesen und Darstellung der MTMM-Matrix*Annahmen in Bezug auf die Skala Organisation:*

Personen mit einer erfolgreichen Lernstrategie im LAMBDA-neu laut Manual erreichen höhere Werte in der Skala WSP-Organisation. Personen, die das Kriterium erreichen im LAMBDA-neu unterscheiden sich von Personen, die das Kriterium nicht erreichen in Bezug auf ihren Wert in der Skala WSP-Organisation. Die Dauer der in der Relativen Lernzeit 1 verbrachten Zeit korreliert mit dem Score in der Skala WSP-Organisation.

Annahmen in Bezug auf die Skala Motivation:

Personen mit einer erfolgreichen Lernstrategie laut Manual erreichen höhere Werte in der Skala WSP-Motivation. Personen, die im Kodiertest laut Manual ein angemessenes Anspruchsniveau haben, erreichen höhere Werte in der Skala WSP-Motivation. Personen, die im Kodiertest mehr Symbole richtig kodieren erreichen höhere Werte in der Skala WSP-Motivation.

Annahmen in Bezug auf die Skala Frustrationstoleranz:

Personen, die im Kodiertest mehr Symbole richtig kodieren erreichen höhere Werte in der Skala WSP-Frustrationstoleranz. Personen, die im LAMBDA-neu bessere oder gleiche Leistungen unter Stress aufweisen, erreichen höhere Werte in der Skala WSP-

Frustrationstoleranz. Personen, deren Leistungsspitze im Kodiertest im letzten Durchgang liegt, erreichen höhere Werte in der Skala WSP-Frustrationstoleranz.

Annahmen in Bezug auf die Skala Selbstbewusstsein:

Personen, die sich ihres Wissens im WITE sicherer sind, das bedeutet höhere Werte in der Skala „Aktivwissen“ und niedrigere Werte in den Skalen „Passivwissen“ und „Scheinwissen“, weisen höhere Werte in der Skala *Selbstbewusstsein* auf.

Hypothesen zu „Organisation“

H ₀ OF ₁	Personen mit einer erfolgreichen/nicht erfolgreichen Lernstrategie im LAMBDA-neu (laut Manual) unterscheiden sich nicht in Bezug auf ihre Ausprägung in der Skala Organisation des WSP.
H ₁ OF ₁	Personen mit einer erfolgreichen/nicht erfolgreichen Lernstrategie im LAMBDA-neu (laut Manual) unterscheiden sich in Bezug auf ihre Ausprägung in der Skala Organisation des WSP.
H ₀ OF ₂	Es findet sich kein Zusammenhang zwischen der Ausprägung in der Skala Organisation des WSP und dem Wert der „Relativen Lernzeit1“ im LAMBDA-neu.
H ₁ OF ₂	Es findet sich ein Zusammenhang zwischen der Ausprägung in der Skala Organisation des WSP und dem Wert der „Relativen Lernzeit1“ im LAMBDA-neu.
H ₀ OF ₃	Es findet sich kein Zusammenhang zwischen der Ausprägung in der Skala Organisation des WSP und der erfolgreichen Bewältigung des LAMBDA-neu.
H ₁ OF ₃	Es findet sich ein Zusammenhang zwischen der Ausprägung in der Skala Organisation des WSP und der erfolgreichen Bewältigung des LAMBDA-neu.

Hypothesen zu „Motivation“

H ₀ M ₁	Personen mit einer erfolgreichen/nicht erfolgreichen Lernstrategie im LAMBDA-neu (laut Manual) unterscheiden sich nicht in Bezug auf ihre Ausprägung in der Skala Motivation des WSP.
H ₁ M ₁	Personen mit einer erfolgreichen/nicht erfolgreichen Lernstrategie im LAMBDA-neu (laut Manual) unterscheiden sich in Bezug auf ihre Ausprägung in der Skala Motivation des WSP.
H ₀ M ₂	Es findet sich kein Zusammenhang zwischen der Ausprägung in der Skala

	Motivation des WSP und dem Wert der „Relativen Lernzeit1“ im LAMBDA-neu.
H ₁ M ₂	Es findet sich ein Zusammenhang zwischen der Ausprägung in der Skala Motivation des WSP und dem Wert der „Relativen Lernzeit1“ im LAMBDA-neu.
H ₀ M ₃	Es findet sich kein Zusammenhang zwischen der Ausprägung in der Skala Motivation des WSP und der erfolgreichen Bewältigung des LAMBDA-neu.
H ₁ M ₃	Es findet sich ein Zusammenhang zwischen der Ausprägung in der Skala Motivation des WSP und der erfolgreichen Bewältigung des LAMBDA-neu.
H ₀ M ₄	Es findet sich kein Zusammenhang zwischen der Ausprägung in der Skala Motivation des WSP und dem „Anspruchsniveau2“ im Kodiertest.
H ₁ M ₄	Es findet sich ein Zusammenhang zwischen der Ausprägung in der Skala Motivation des WSP und dem „Anspruchsniveau2“ im Kodiertest.
H ₀ M ₅	Es findet sich kein Zusammenhang zwischen der Ausprägung in der Skala Motivation des WSP und dem Leistungsniveau im Kodiertest.
H ₁ M ₅	Es findet sich ein Zusammenhang zwischen der Ausprägung in der Skala Motivation des WSP und dem Leistungsniveau im Kodiertest.
H ₀ M ₆	Es findet sich kein Zusammenhang zwischen der Ausprägung in der Skala Motivation des WSP und dem Zeitpunkt der Leistungsspitze im Kodiertest.
H ₁ M ₆	Es findet sich ein Zusammenhang zwischen der Ausprägung in der Skala Motivation des WSP und dem Zeitpunkt der Leistungsspitze im Kodiertest.

Hypothesen zu „Frustrationstoleranz“

H ₀ FT ₁	Es findet sich kein Zusammenhang zwischen der Ausprägung in der Skala Frustrationstoleranz des WSP und dem Wert des „Belastbarkeitsquotienten“ im LAMBDA-neu.
H ₁ FT ₁	Es findet sich ein Zusammenhang zwischen der Ausprägung in der Skala Frustrationstoleranz des WSP und dem Wert des „Belastbarkeitsquotienten“ im LAMBDA-neu.
H ₀ FT ₂	Es findet sich kein Zusammenhang zwischen der Ausprägung in der Skala Frustrationstoleranz des WSP und dem Leistungsniveau im Kodiertest.
H ₁ FT ₂	Es findet sich ein Zusammenhang zwischen der Ausprägung in der Skala Frustrationstoleranz des WSP und dem Leistungsniveau im Kodiertest.
H ₀ FT ₃	Es findet sich kein Zusammenhang zwischen der Ausprägung in der Skala Frustrationstoleranz des WSP und der Position der Leistungsspitze im Kodiertest.
H ₁ FT ₃	Es findet sich ein Zusammenhang zwischen der Ausprägung in der Skala

	Frustrationstoleranz des WSP und der Position der Leistungsspitze im Kodiertest.
--	--

Hypothesen zu „Selbstbewusstsein“

H ₀ S ₁	Es findet sich kein Zusammenhang zwischen der Ausprägung in der Skala Selbstbewusstsein des WSP und der Ausprägung des Kennwerts „Scheinwissen“ im WITE.
H ₁ S ₁	Es findet sich ein Zusammenhang zwischen der Ausprägung in der Skala Selbstbewusstsein des WSP und der Ausprägung des Kennwerts „Scheinwissen“ im WITE.
H ₀ S ₂	Es findet sich kein Zusammenhang zwischen der Ausprägung in der Skala Selbstbewusstsein des WSP und der Ausprägung des Kennwerts „Passivwissen“ im WITE.
H ₁ S ₂	Es findet sich ein Zusammenhang zwischen der Ausprägung in der Skala Selbstbewusstsein des WSP und der Ausprägung des Kennwerts „Passivwissen“ im WITE.
H ₀ S ₃	Es findet sich kein Zusammenhang zwischen der Ausprägung in der Skala Selbstbewusstsein des WSP und der Ausprägung des Kennwerts „Aktivwissen“ im WITE.
H ₁ S ₃	Es findet sich ein Zusammenhang zwischen der Ausprägung in der Skala Selbstbewusstsein des WSP und der Ausprägung des Kennwerts „Aktivwissen“ im WITE.

Die Multi-Trait-Multi-Method Matrix

Ähnlich den Hypothesen wird die folgende MTMM-Matrix postuliert (siehe Tab. 4); zum Verständnis der Abkürzungen sei auch auf die Legende in Tabelle 5 verwiesen:

Die dunkelgrauen Felder bezeichnen den Monotrait-Monomethod-Block. Hier betragen die Korrelationen 1, da jeweils ein Wert mit sich selber korreliert wird.

Die roten Felder bezeichnen den Monotrait-Heteromethod-Block. Hier wird ein Merkmal mit mehreren Methoden erfasst. Diese Felder stehen für die konvergente Validität, und sollte diese vorliegen, sollten hier hohe Korrelationen vorgefunden werden. Der Durchschnitt dieser Werte soll sich deutlich, also signifikant und bedeutsam, von null unterscheiden. Diese Korrelationen sollen, neben den Monotrait-Monomethod-Korrelationen, die höchsten sein.

Die grünen Felder bezeichnen den Heterotrait-Monomethod-Block. Hier werden jeweils mehrere Merkmale mit einer Methode erfasst. Diese Felder stehen für die diskriminante Validität und sollten kleiner sein als die Korrelationen der konvergenten Validität (rote Felder), dies alles innerhalb derselben Methode.

Die blauen Felder bezeichnen den Heterotrait-Heteromethod-Block. Hier werden mehrere Merkmale mit mehreren Methoden erfasst, und auch diese Felder stehen für die diskriminante Validität. Bei vorliegender diskriminanter Validität werden in diesem Block die geringsten Korrelationen erwartet, zumindest aber kleinere als die der konvergenten Validität (rote Felder).

	WSP OF	WSP M	WSP FT	WSP S	LAMB DA PPh	LAMB DA 0F	LAMB DA Gld	REL Lz1	LAMB DA Lstr	AHA Irb	AHA An1	AHA An2	LAMB DA Krit	AHA Zd	LAMB DA Bqu	AHA Lsp	WITE AW	WITE PW	WITE SW
WSPOF																			
WSPM																			
WSPFT																			
WSPS																			
LAMBDA PPh																			
LAMBDA 0F																			
LAMBDA Gld																			
RelLz1																			
LAMBDA Lstr																			
AHA Irb																			
AHA An1																			
AHA An2																			
LAMBDA Krit																			
AHA Zd																			
LAMBDA Bqu																			
AHA Lsp																			
WITE AW																			
WITE PW																			
WITE SW																			

Tabelle 4: Entwurf der Multitrait-Multimethod-Matrix

Legende:

WSPOF	Score in WSP-Organisation	RelLz1	LAMBDA Relative Lernzeit 1	LAMBDA Bqu	LAMBDA Belastbarkeitsquotient
WSPM	Score in WSP-Motivation	LAMBDA Lstr	LAMBDA Lernstrategie	AHA Lsp	Kodiertest Zeitpunkt d. Leistungsspitze
WSPFT	Score in WSP-Frustrationstoleranz	AHA Irb	Kodiertest Insg. richtig bearb. Symbole	WITE AW	WITE Aktivwissen
WSPS	Score in WSP-Selbstbewusstsein	AHA An1	Kodiertest Anspruchsniveau 1	WITE PW	WITE Passivwissen
LAMBDA PPh	LAMBDA Anz. d. Prüfphasen	AHA An2	Kodiertest Anspruchsniveau 2	WITE SW	WITE Scheinwissen
LAMBDA 0F	LAMBDA Anz. d. Nullfehlerdgr.	LAMBDA Krit	LAMBDA Kriterium erreicht		
LAMBDA Gld	LAMBDA Gesamtlerndauer	AHA Zd	Kodiertest Zieldiskrepanz		

Tabelle 5: Legende der Multitrait-Multimethod-Matrix

2. Die Stichprobe

2.1 Akquirierung von Testpersonen

Das **Wiener Self-Assessment Psychologie**[®] wurde am 29.08.2011 fertiggestellt und online gestellt auf der Plattform www.studienberatungstest.at. Ab diesem Zeitpunkt war es den InteressentInnen möglich, auch über einen Link von der Homepage der Test- und Beratungsstelle, die Seite aufzurufen.

Die Sammlung der Daten begann ab Ende August 2011, wobei diese ab Oktober 2011 intensiviert wurde. Zu Beginn wurde der Link und die Beschreibung des **Wiener Self-Assessment Psychologie**[®] in insgesamt zehn Internetforen gepostet. Weiter wurden Forenuser, aus deren Postings hervorging, dass sie sich für ein Psychologiestudium interessierten, angeschrieben und motiviert teilzunehmen. Ebenso sind kleine Poster vor dem AudiMax der Universität Wien, in der Fakultät für Psychologie der Universität Wien und im Neuen Institutsgebäude der Universität Wien aufgehängt worden. Kollegen des Studienfachs Psychologie wurden bezüglich einer Weiterleitung an Personen, die in die Zielgruppe fallen, angeschrieben.

Darüber hinaus wurden Flyer vor einer Vorlesung für erstsemestrige PsychologiestudentInnen verteilt und das **Wiener Self-Assessment Psychologie**[®] in drei Supervised-Orientation-Tutorien (SOT), einer Pflichtveranstaltung mit Anwesenheitspflicht, vorgestellt und versucht die StudentInnen zu motivieren daran teilzunehmen. Insgesamt erfolgten Vorstellungen des **Wiener Self-Assessment Psychologie**[®] in zwei Vorlesungen an der Sigmund Freud Universität in 1030 Wien und in einer STEOP-Vorlesung, einer großen Vorlesung für erstsemestrige PsychologiestudentInnen an der Universität Wien.

Außerdem wurde das **Wiener Self-Assessment Psychologie**[®] kurz (max. 10 Minuten) in elf Schulklassen (7. und 8. Schulstufen) in Wien (14., 15., 16. und 8. Wiener Gemeindebezirk), ebenso wie in 13 Klassen der Maturaschule Dr. Roland in 1080 Wien vorgestellt. InteressentInnen erhielten Flyer mit dem Internetlink.

Zusätzlich zu den Akquirierungen in Wien wurde das **Wiener Self-Assessment Psychologie**[®] in je zwei Vorlesungen für erstsemestrige PsychologiestudentInnen an der Universität Graz und an der Universität Salzburg vorgestellt.

2.2 Beschreibung der Datenlage

Als die Daten am 04.12.2011 vom Server der Technischen Universität Wien, über den das **Wiener Self-Assessment Psychologie**® seit August 2011 lief, exportiert wurden, hatten insgesamt 391 Personen das Symbol „Testung starten“ nach der Registrierung mit einer Emailadresse und Einloggen mittels PIN-Codes, betätigt.

In einem ersten Schritt erfolgte eine Bereinigung der Daten. Wenn Teilnehmer mehrfach teilgenommen hatten, so wurden alle Durchgänge bis auf den ersten entfernt (dies war in drei Fällen notwendig). Und schließlich mussten Personen eliminiert werden, die zu den Entwicklern und Testern des **Wiener Self-Assessment Psychologie**® gehören, beziehungsweise auch Teilnehmer eines Proseminars aus dem Psychologiestudium, die die Aufgabe hatten das **Wiener Self-Assessment Psychologie**® selbst einmal zu absolvieren.

Nach diesen Entfernungen blieben 361 Personen übrig. Im Zeitraum zwischen Ende August 2011 und 04.12.2011 hatten 361 Personen das **Wiener Self-Assessment Psychologie**® begonnen.

Von diesen Personen haben 217 Personen das erste Verfahren, den LAMBDA-neu, abgeschlossen, also rund 60,1 Prozent der Personen, die begonnen haben (siehe Tab. 44 im Anhang; 2. Beschreibung der Datenlage). Weitere sechs Personen haben während dem Kodiertest abgebrochen, den nur 211 Personen abgeschlossen haben (siehe Tab. 45 im Anhang; 2. Beschreibung der Datenlage). Weitere fünf Personen haben dann während dem WITE abgebrochen, insgesamt haben 206 Personen den WITE abgeschlossen (siehe Tab. 46 im Anhang; 2. Beschreibung der Datenlage). Die Fragebögen FIP&SKIP haben noch 197 Personen komplett bearbeitet (siehe Tab. 47 im Anhang; 2. Beschreibung der Datenlage). Während des Matrizentests haben neun weitere Personen abgebrochen, sodass diesen nur 181 Personen abgeschlossen haben (siehe Tab. 48 im Anhang; 2. Beschreibung der Datenlage). Eine weitere Person hat während dem WSP die Selbsttestung abgebrochen, sodass den WSP nur mehr 180 Personen abgeschlossen haben (siehe Tab. 49 im Anhang; 2. Beschreibung der Datenlage). Bei der Persönlichen Einschätzung hat dann kein Teilnehmer mehr abgebrochen.

Von den 361 Personen, die das **Wiener Self-Assessment Psychologie**® begonnen haben, haben es schließlich 180 Personen beendet (siehe Tab. 6), das sind 49,9 Prozent.

Toolvollstaendig

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
0	181	50,1	50,1	50,1
Gültig 1	180	49,9	49,9	100,0
Gesamt	361	100,0	100,0	

Tabelle 6: Vollständig abgeschlossene Self-Assessment Tools

2.3 Deskriptive Beschreibung der Stichprobe

Die Stichprobe bestand schließlich aus 180 Personen, die im Zeitraum von Ende August 2011 bis 04.12.2011 das **Wiener Self-Assessment Psychologie®** abgeschlossen haben.

Die Stichprobe hat sich wie folgt zusammengestellt:

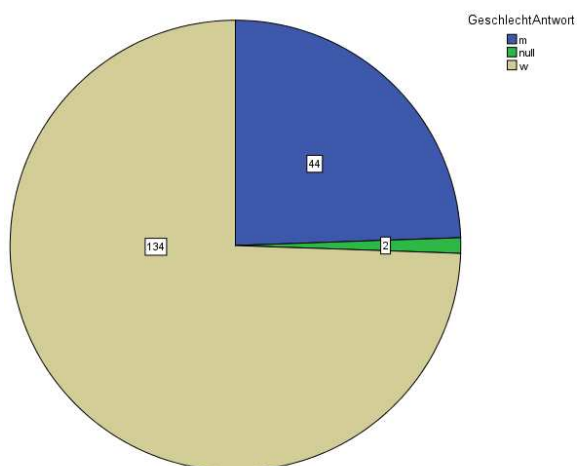


Abbildung 4: Kreisdiagramm: Geschlecht der abgeschlossenen TeilnehmerInnen

Aus dem Kreisdiagramm in Abbildung 4 geht hervor, dass 134 der TeilnehmerInnen (74,4%) weiblich waren und 44 (24,4%) männlich, zwei TeilnehmerInnen haben ihr Geschlecht nicht angegeben.

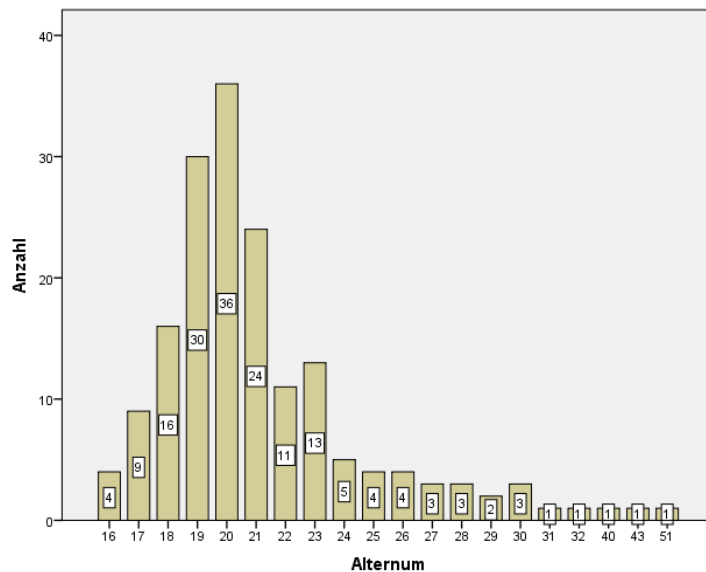


Abbildung 5: Balkendiagramm der Altersverteilung der abgeschlossenen TeilnehmerInnen

Im Balkendiagramm in Abbildung 5 ist die Altersverteilung der TeilnehmerInnen zu sehen. Das Alter der TeilnehmerInnen lag zwischen 16 und 51 Jahren, wobei die meisten TeilnehmerInnen 20 Jahre alt waren, der Mittelwert lag bei 21,36 Jahren (siehe Tab. 7).

Statistiken

Alternum

N	Gültig	172
	Fehlend	8
Mittelwert		21,36
Median		20,00
Modus		20
Minimum		16
Maximum		51

Tabelle 7: Statistiken zur Altersverteilung der TeilnehmerInnen

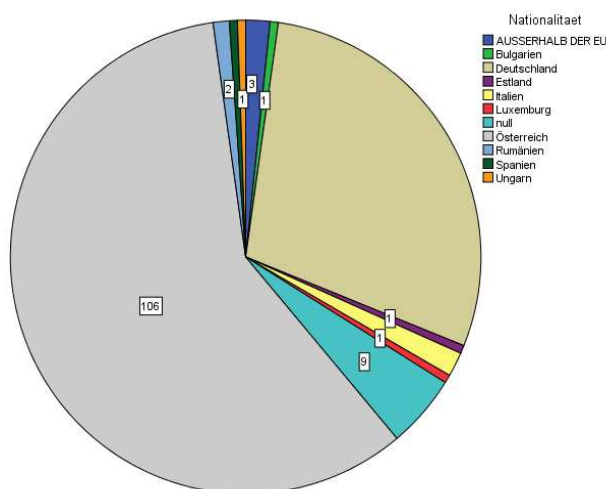


Abbildung 6: Kreisdiagramm der Herkunftsländer der abgeschlossenen TeilnehmerInnen

Wie im Kreisdiagramm in Abbildung 6 zu sehen ist, haben 106 TeilnehmerInnen (58,89%), die aus Österreich stammen, 52 TeilnehmerInnen (28,89%) aus Deutschland und insgesamt 22 weitere TeilnehmerInnen (12,22%), die aus Bulgarien, Estland, Italien, Luxemburg, Rumänien, Spanien, Ungarn und von außerhalb der Europäischen Union stammen, teilgenommen.

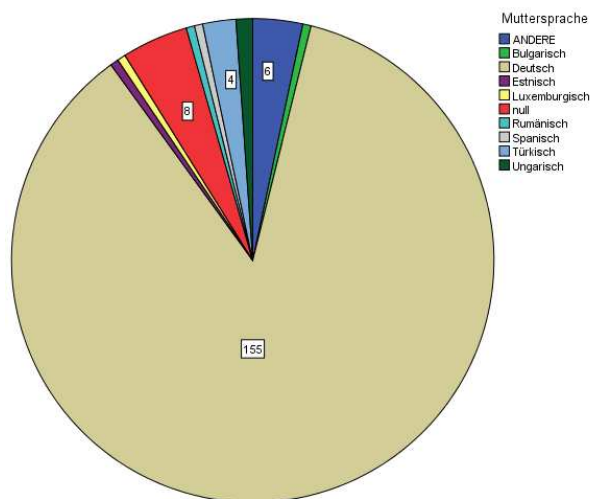


Abbildung 7: Kreisdiagramm der Muttersprachen der abgeschlossenen TeilnehmerInnen

Dementsprechend hat sich auch die Muttersprache verteilt (siehe Abb. 7), von 155 TeilnehmerInnen (86,11%) war Deutsch die Muttersprache, ansonsten gab es, passend zu den Herkunftsländern die Muttersprachen Bulgarisch, Estnisch, Luxemburgisch, Rumänisch, Spanisch, Türkisch und Ungarisch (insgesamt 13,89%).

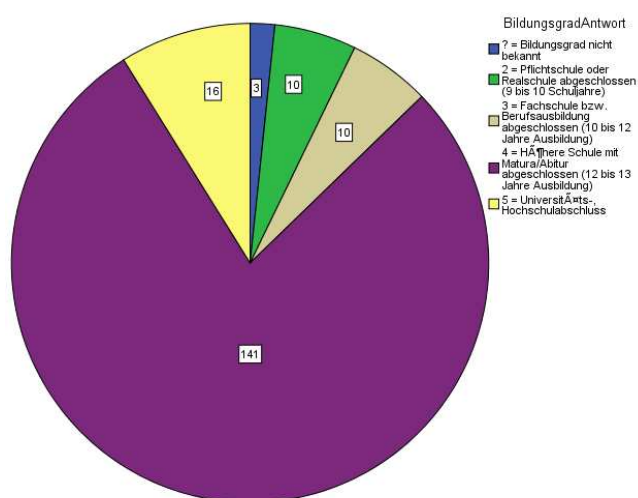


Abbildung 8: Kreisdiagramm zu der höchsten abgeschlossenen Schulbildung der abgeschlossenen TeilnehmerInnen

Bezüglich der abgeschlossenen Schulbildung haben 141 TeilnehmerInnen (78,33%) die Matura oder das Abitur abgeschlossen, 16 TeilnehmerInnen (8,89%) haben einen Universitäts- oder Hochschulabschluss, 10 TeilnehmerInnen (5,56%) haben eine Berufsausbildung abgeschlossen, 10 weitere TeilnehmerInnen (5,56%) haben die Pflichtschule abgeschlossen und drei weitere TeilnehmerInnen (1,67%) gaben an, den Bildungsgrad nicht zu kennen (siehe Abb. 8).

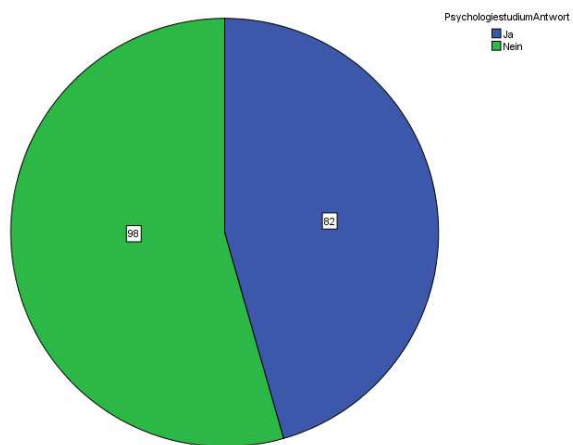


Abbildung 9: Kreisdiagramm zur Frage ob bereits ein Psychologiestudium begonnen wurde

Wie im Kreisdiagramm zu sehen ist (siehe Abb. 9) haben knapp über die Hälfte der TeilnehmerInnen (98 Personen; 54,44%) noch kein Psychologiestudium begonnen, diese haben sich also im Vorhinein über ihre persönlichen Stärken und Schwächen in Bezug auf das Psychologiestudium informiert. 82 Personen (45,56%) haben bereits ein Psychologiestudium begonnen.

Von den 82 Personen, die bereits ein Psychologiestudium begonnen haben, gaben 54 Personen (65,85%) an, an der Universität Wien zu studieren (siehe Tab. 8). Die restlichen Personen, die bereits Psychologie studieren, das waren 28 TeilnehmerInnen (34,15%) gaben an woanders zu studieren (siehe Tab. 8).

Universität Wien Antwort				
	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	98	54,4	54,4	54,4
andere Universität	28	15,6	15,6	70,0
Universität Wien	54	30,0	30,0	100,0
Gesamt	180	100,0	100,0	

Tabelle 8: Studienort der TeilnehmerInnen

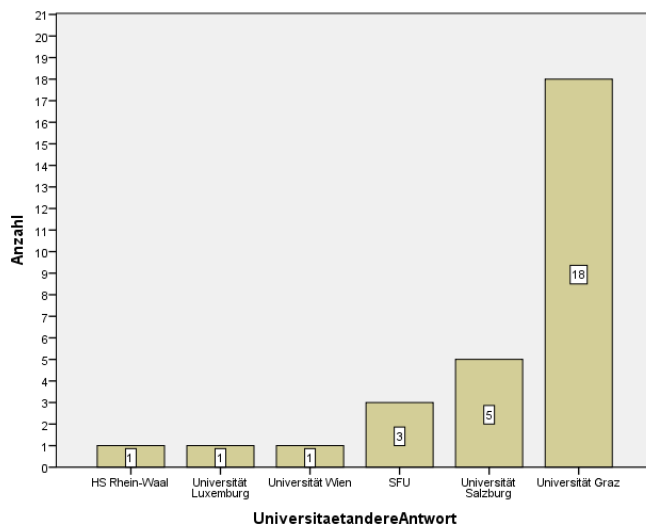


Abbildung 10: Balkendiagramm zum Studienort der TeilnehmerInnen wenn nicht Universität Wien

In Abbildung 10 ist zu sehen, dass von den Personen, die nicht an der Universität Wien Psychologie studieren, 18 an der Universität Graz, 5 an der Universität Salzburg, 3 an der Sigmund Freud Universität in Wien studieren, und jeweils ein Teilnehmer der Universität Luxemburg und ein Teilnehmer der Hochschule Rhein-Waal in Nordrhein-Westfalen in Deutschland. Zusätzlich hat eine Person nachträglich angegeben an der Universität Wien zu studieren, also sind es insgesamt 55 StudentInnen der Psychologie an der Universität Wien (67,07%), die restlichen Personen (27 Personen, 32,92%) studieren an anderen Universitäten.

An dieser Stelle sei nun auch ein kurzer Überblick über das Allgemeine Abschneiden in den in der Validierung relevanten psychologische diagnostischen Verfahren dargestellt.

Bezüglich des WSP sei auf Tabelle 9 verwiesen. Es ist zu sehen, dass die Teilnehmer in der Skala WSP-Organisation mindestens 24 Punkte erreicht haben, maximal 70 (mögliche zu erreichende Punktzahlen: 18-72), der Mittelwert lag bei 51,88 Punkten. Den Teilnehmern wurde in Bezug auf die Skalen des WSP rückgemeldet, ob sie eine Stärke oder Schwäche haben. Hierbei wurde der dichotomisierte Wert herangezogen, wobei hier Werte von 0 bis 12 als Schwäche rückgemeldet wurden, und Werte von 13 bis 18 als Stärke. In Tabelle 10 ist erkennbar, dass 46,7% der TeilnehmerInnen (84 Personen) eine Schwäche in dieser Skala rückgemeldet wurde, 53,3% (96 Personen) wurde eine Stärke rückgemeldet.

Bezüglich der Skala WSP-Motivation ist in Tabelle 9 zu sehen, dass zumindest 31 Punkte erreicht wurden, und maximal 81 (mögliche zu erreichende Punktzahlen: 22-88). Der Mittelwert lag hier bei 64,16 Punkten (siehe Tab. 9). Bei Berechnung der dichotomisierten

Scores in Bezug auf die Rückmeldung galt für die Skala Motivation ein Punktwert zwischen 0 und 15 als Schwäche und von 16 bis 22 als Stärke. Tabelle 11 zeigt, dass 89 Personen (49,4%) eine Schwäche hier rückgemeldet wurde, und 91 Personen (50,6%) eine Stärke.

In der Skala WSP-Frustrationstoleranz erreichten die TeilnehmerInnen zumindest ein Mal den Mindestwert von 10, und den Maximalwert von 28 (mögliche zu erreichende Punktzahlen: 7-28), wobei der Mittelwert bei 21,30 lag (siehe Tab. 9). Den TeilnehmerInnen wurde bei einem dichotomisierten Punktwert zwischen 0 und 4 eine Schwäche rückgemeldet, und zwischen 5 und 7 eine Stärke. In Tabelle 12 ist zu sehen, dass 28,4% der TeilnehmerInnen (51) eine Schwäche rückgemeldet wurde und 71,7% (129) eine Stärke in Bezug auf Frustrationstoleranz.

Die TeilnehmerInnen erreichten in der Skala WSP-Selbstbewusstsein den Mindestwert von 8 und zumindest ein Mal den Wert von 23 (mögliche zu erreichende Punkte: 6-24), wobei der Mittelwert bei 15,61 lag (siehe Tab. 9). 104 TeilnehmerInnen (57,8%) wurde hier eine Schwäche rückgemeldet und 76 TeilnehmerInnen (42,2%) wurde eine Stärke rückgemeldet (siehe Tab. 13), wobei hier Werte zwischen 0 und 3 als Schwäche galten und zwischen 4 und 6 als Stärke.

Deskriptive Statistik

	N	Minimum	Maximum	Mittelwert	Standardabweichung	Varianz
WSPOrganisationRWGesamt	180	24	70	51,88	7,930	62,890
WSPOrganisationRWGesamtdichotomisiert	180	2	18	12,33	3,635	13,216
WSPMotivationRWGesamt	180	31	81	64,16	8,627	74,423
WSPMotivationRWGesamtdichotomisiert	180	3	22	15,39	4,001	16,004
WSPFrustrationstoleranzRWGesamt	180	10	28	21,30	2,699	7,284
WSPFrustrationstoleranzRWGesamtdichotomisiert	180	1	7	5,39	1,309	1,714
WSPSelbstbewusstseinRWGesamt	180	8	23	15,61	3,489	12,172
WSPSelbstbewusstseinRWGesamtdichotomisiert	180	0	6	3,17	1,805	3,257
Gültige Werte (Listenweise)	180					

Tabelle 9: Abschneiden der TeilnehmerInnen im WSP

WSPOrganisationdichotomSS

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig ,00	84	46,7	46,7	46,7
1,00	96	53,3	53,3	100,0
Gesamt	180	100,0	100,0	

Tabelle 10: Anzahl der Stärke/Schwäche-Rückmeldungen für WSP-Organisation

WSPMotivationdichotomSS

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig ,00	89	49,4	49,4	49,4
1,00	91	50,6	50,6	100,0
Gesamt	180	100,0	100,0	

Tabelle 11: Anzahl der Stärke/Schwäche-Rückmeldungen für WSP-Motivation

WSPFrustrationstoleranzdichotomSS

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig ,00	51	28,3	28,3	28,3
1,00	129	71,7	71,7	100,0
Gesamt	180	100,0	100,0	

Tabelle 12: Anzahl der Stärke/Schwäche-Rückmeldungen für WSP-Frustrationstoleranz

WSPSelbstbewusstseinindichotomSS

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig ,00	104	57,8	57,8	57,8
1,00	76	42,2	42,2	100,0
Gesamt	180	100,0	100,0	

Tabelle 13: Anzahl der Stärke/Schwäche-Rückmeldungen für WSP-Selbstbewusstsein

3. Darstellung der Ergebnisse der Hypothesen

An dieser Stelle sei erwähnt, dass die Beschriftung der exportierten Daten aus dem Self-Assessment Tool noch die veralteten Testbezeichnungen enthält und diese übernommen wurden. Wenn in den Tabellen die Abkürzung „LAMBDA“ zu sehen ist, so ist jedes Mal

„LAMBDA-neu“ gemeint, ebenso wie bei „LEWITE“ „WITE“ gemeint ist. Der Kodiertest ist ebenso mit der Abkürzung „AHA“ bezeichnet, die für „Arbeitshaltungen“ steht.

3.1 Ergebnisse der Hypothesen zu „Organisation“

Überprüfung von H_0OF_1

Bei dieser Hypothese geht es um die Überprüfung der Frage, ob Personen mit einer guten Lernstrategie im LAMBDA-neu (laut Manual) durchschnittlich andere Scores in der Skala WSP-Organisation erreichen als Personen mit einer schlechten Lernstrategie.

Dafür wurden die Personen in zwei Gruppen unterteilt; in Gruppe 0 befinden sich die Personen mit einer schlechten Lernstrategie (Wert in der Lernstrategie bis 0,4999 und/oder Kriterium nicht erreicht) und Personen mit einer guten Lernstrategie (Wert in der Lernstrategie 0,5 oder höher und Kriterium erreicht) befinden sich in Gruppe 1.

Die Überprüfung der Voraussetzung der Normalverteilung in beiden Stichproben findet sich im Anhang in den Tabellen 49 und 50 (siehe Anhang Kapitel 3.1).

Die Gruppenstatistiken in Tabelle 14 zeigen, dass Personen mit schlechter Lernstrategie durchschnittlich 50,99 Punkte in der Skala WSP-Organisation erreicht haben; Personen mit guter Lernstrategie haben durchschnittlich 52,61 Punkte erreicht.

Gruppenstatistiken					
	LAMBDAErfolgreiche Lernstrategie	N	Mittelwert	Standardab- weichung	Standardfehler des Mittelwertes
WSPOrganisation	,00	81	50,99	7,969	,885
RWGesamt	1,00	99	52,61	7,863	,790

Tabelle 14: Gruppenstatistiken zu H_0OF_1

Diese Unterschiede in den Gruppen (siehe Tab. 14) sind bei einem Signifikanzwert von $p=0,174$ nicht signifikant unterschiedlich (siehe Tab. 15). Hierbei gilt die Annahme von gleichen Fehlervarianzen in beiden Gruppen, als Voraussetzung für die Berechnung dieses Verfahrens, auf Basis des nicht signifikanten Wertes beim Levene-Test; $p=0,734$ (siehe Tab. 15).

Test bei unabhängigen Stichproben

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz	
									Untere	Obere
WSPOrganisationRWG Gesamt	Varianzen sind gleich	,115	,734	-1,365	178	,174	-1,618	1,185	-3,957	,721
	Varianzen sind nicht gleich			-1,364	170,102	,175	-1,618	1,187	-3,961	,725

Tabelle 15: T-Test bei unabhängigen Stichproben zu H_0OF_1

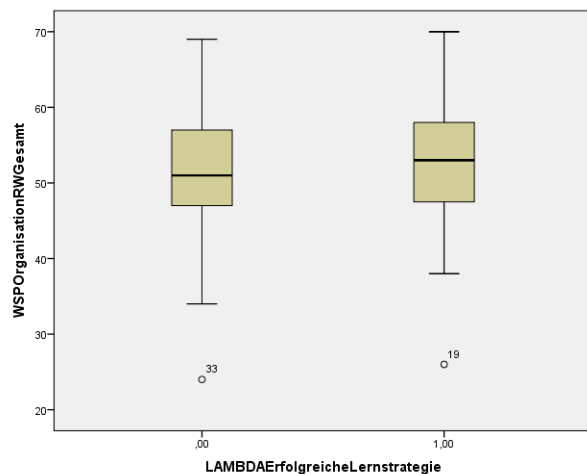


Abbildung 11: Box-Plot zu H_0OF_1

In Abbildung 11 ist der Box-Plot mit den beiden Gruppen (O: nicht erfolgreiche Lernstrategie; 1: erfolgreiche Lernstrategie) zu sehen. Man sieht, dass sich die beiden Mittelwerte der Gruppen in der Skala WSP-Organisation kaum unterscheiden.

H_0OF_2

Diese Hypothese betrifft den Zusammenhang zwischen dem Score in der Skala WSP-Organisation und der in der Relativen Lernzeit 1 verbrachten Zeit.

Da der Kennwert der Relativen Lernzeit 1 nicht intervallskaliert ist, muss hier auf ein parameterfreies Verfahren, die Korrelation nach Spearman, ausgewichen werden.

In Tabelle 16 ist zu sehen, dass der Korrelationskoeffizient zwischen den beiden betreffenden Variablen bei einem Spearman-Rho=0,055 liegt, was mit $p=0,462$ nicht signifikant ist.

Korrelationen

			WSPOrgani sationRW Gesamt	RelLernzeit1
Spearman- Rho	WSPOrganisationRWGesamt	Korrelationskoeffizient	1,000	,055
		Sig. (2-seitig)	.	,462
		N	180	180
	RelLernzeit1	Korrelationskoeffizient	,055	1,000
		Sig. (2-seitig)	,462	.
		N	180	180

Tabelle 16: Spearman-Rho Korrelation zu H_0OF_2

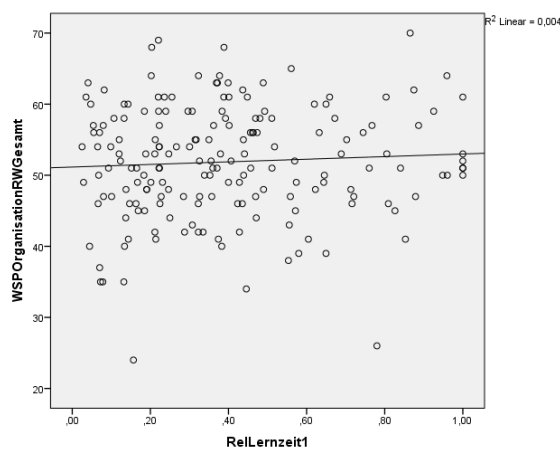


Abbildung 12: Streudiagramm zu H_0OF_2

In Abbildung 12 ist das Streudiagramm zu H_0OF_2 zu sehen; auf der X-Achse ist die Relative Lernzeit 1 und auf der Y-Achse der Score in der Skala WSP-Organisation. Weiter ist die angepasste Regressionsgerade zu sehen.

H_0OF_3

Diese Hypothese betrifft die Frage, ob sich Personen, die im LAMBDA-neu das Kriterium erreicht haben von Personen unterscheiden, die im LAMBDA-neu das Kriterium nicht erreicht haben, und zwar in Bezug auf ihren durchschnittlichen Score in der Skala WSP-Organisation. Zuerst sei die Anzahl der Personen dargestellt, die im LAMBDA-neu das Kriterium erreicht haben. Dies waren 117, also 65%, und 35%, also 63 TeilnehmerInnen haben das Kriterium nicht erreicht (siehe Tab. 17).

LAMBDAKriterr

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig				
,00	63	35,0	35,0	35,0
1,00	117	65,0	65,0	100,0
Gesamt	180	100,0	100,0	

Tabelle 17: Anzahl der Personen, die im LAMBDA-neu das Kriterium erreicht haben

Die Voraussetzung der Normalverteilung in den Gruppen wurde überprüft und ist im Anhang Kapitel 3.1 (siehe Tab. 51 und 52) nachzusehen. Die Gruppenstatistiken in Tabelle 18 zeigen, dass Personen, die das Kriterium nicht erreicht haben (Gruppe 0) im Mittel 51,08 Punkte in der Skala WSP-Organisation erzielten. Personen, die das Kriterium erfüllt haben (Gruppe 1), erreichten durchschnittlich 52,31 Punkte in der betreffenden Skala.

Gruppenstatistiken

	LAMBDAKriterr	N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
WSPOrganisationRW	,00	63	51,08	7,946	1,001
Gesamt	1,00	117	52,31	7,923	,732

Tabelle 18: Gruppenstatistiken für T-Test für unabhängige Stichproben für H_0OF_3

Der Levene-Test zur Prüfung der Varianzgleichheit ist nicht signifikant ($p=0,621$) ausgefallen (siehe Tab. 19), deswegen ist die Berechnung eines T-Tests bei unabhängigen Stichproben möglich. Der T-Test bei unabhängigen Stichproben zeigt (siehe Tab. 19), dass sich die Mittelwerte in den zwei Gruppen (Kriterium erreicht – Gruppe 1; Kriterium nicht erreicht – Gruppe 0) in Bezug auf den Score in WSP-Organisation nicht signifikant unterscheiden ($p=0,323$).

Test bei unabhängigen Stichproben

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz	
									Untere	Obere
WSPOrganisationRWGesamt	Varianzen sind gleich	,245	,621	-,991	178	,323	-1,228	1,239	-3,674	1,217
	Varianzen sind nicht gleich			-,990	126,732	,324	-1,228	1,240	-3,683	1,226

Tabelle 19: T-Test bei unabhängigen Stichproben für H_0OF_3

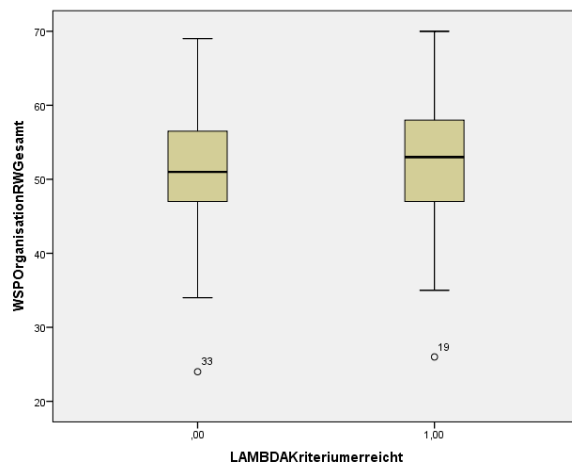


Abbildung 13: Boxplot zu H_0OF_3

Im Box-Plot in Abbildung 13 ist zu sehen, dass sich die Mittelwerte in den beiden Gruppen kaum unterscheiden. Die Gruppe 0 sammelt Personen, die das Kriterium im LAMBDA-neu nicht erreicht haben, in Gruppe 1 sind die Personen, die das Kriterium erreicht haben. In beiden Gruppen werden fast identische Mittelwerte in WSP-Organisation erreicht.

3.2 Ergebnisse der Hypothesen zu „Motivation“

H_0M_1

Diese Hypothese betrifft die Frage, ob sich Personen mit guter/schlechter Lernstrategie in Bezug auf ihren Score in WSP-Motivation unterscheiden. Die Voraussetzungen für einen T-Test für unabhängige Stichproben sind im Anhang in Kapitel 3.2 (siehe Tab. 53 und 54) nachzusehen; die Durchführung eines T-Tests für unabhängige Stichproben ist gerechtfertigt.

Die Gruppenstatistiken in Tabelle 20 zeigen, dass Personen mit einer schlechten Lernstrategie durchschnittlich 64,14 Punkte in der Skala WSP-Motivation erreicht haben; und Personen mit einer guten Lernstrategie haben durchschnittlich 64,17 Punkte hier erreicht.

Gruppenstatistiken

		N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
LAMBDAErfolgreicheLernstrategie					
WSPMotivation	,00	81	64,14	8,583	,954
RWGesamt	1,00	99	64,17	8,706	,875

Tabelle 20: Gruppenstatistiken zu H_0M_1

Die Berechnung eines T-Tests bei unabhängigen Stichproben ist auch aufgrund des nicht signifikanten Ergebnisses des Levene-Tests auf Homogenität der Fehlervarianzen ($p=0,619$) gerechtfertigt (siehe Tab. 21). Der T-Test für unabhängige Stichproben war mit einem

Signifikanzwert von $p=0,978$ nicht signifikant (siehe Tab. 21); es besteht also kein signifikanter Unterschied zwischen den beiden Gruppen.

Test bei unabhängigen Stichproben

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz	
									Untere	Obere
WSPMotivationRWGesamt	Varianzen sind gleich	,248	,619	-,028	178	,978	-,036	1,296	-2,594	2,522
	Varianzen sind nicht gleich			-,028	171,925	,978	-,036	1,294	-2,591	2,519

Tabelle 21: T-Test für unabhängige Stichproben zu H_0M_1

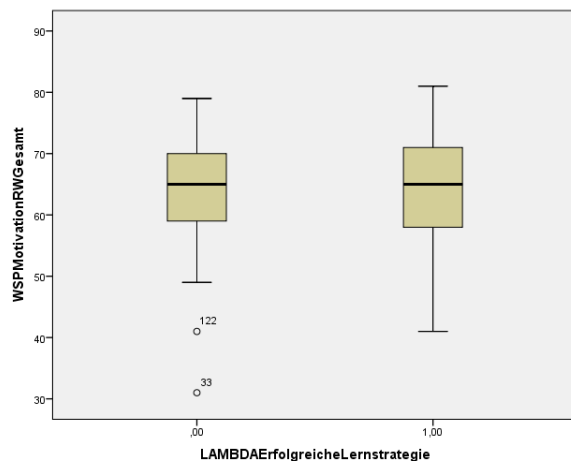


Abbildung 14: Box-Plot zu H_0M_1

In Abbildung 14, im Box-Plot, ist zu sehen, dass sich die Gruppenmittelwerte in den Gruppen 0 (schlechte Lernstrategie) und 1 (gute Lernstrategie) in Bezug auf die Skala WSP-Motivation nicht unterscheiden.

H_0M_2

Hängt die relative Dauer der ersten Lernphase mit dem Score in WSP-Motivation zusammen?

Hierfür sollte eine Korrelation nach Spearman berechnet werden. In Tabelle 22 ist zu sehen, dass der Spearman-Rho Koeffizient=0,040 und nicht signifikant ist ($p=0,597$). Es besteht also kein Zusammenhang zwischen der relativen Dauer der ersten Lernphase und dem Score in WSP-Motivation.

Korrelationen

		RelLernzeit1	WSPMotivationRWGesamt
Spearman-Rho	RelLernzeit1	Korrelationskoeffizient	1,000
		Sig. (2-seitig)	,040
		N	,597
WSPMotivationRWGesamt		180	180
	WSPMotivationRWGesamt	Korrelationskoeffizient	,040
		Sig. (2-seitig)	,597
		N	180

Tabelle 22: Spearman-Rho Korrelation zu H_0M_2

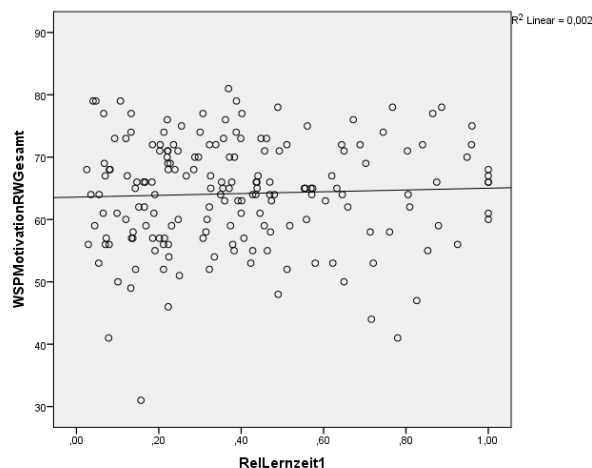


Abbildung 15: Streudiagramm zu H_0M_2

In Abbildung 15 ist das Streudiagramm der beiden Variablen Relative Lernzeit 1 und WSP-Motivation zu sehen, mit einer angepassten Regressionsgerade.

H_0M_3

Diese Hypothese betrifft die Frage, ob sich Personen, die im LAMBDA-neu das Kriterium erreicht haben in Bezug auf ihren durchschnittlichen Score in der Skala WSP-Motivation von Personen unterscheiden, die das Kriterium nicht erreicht haben.

Die Voraussetzungen für das folgende Verfahren sind im Anhang in Kapitel 3.2 in den Tabellen 55 und 56 nachzusehen; ein T-Test für unabhängige Stichproben erscheint gerechtfertigt.

In Tabelle 23 ist zu sehen, dass Personen, die im LAMBDA-neu das Kriterium nicht erfüllt haben (Gruppe 0), durchschnittlich 64,90 Punkte in der Skala WSP-Motivation erreichten, und Personen, die das Kriterium erreicht haben (Gruppe 1), erzielten im Mittel 63,75 Punkte in der betreffenden Skala.

Gruppenstatistiken

	LAMBDAKriterium	N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
WSPMotivationRWGesamt	,00	63	64,90	8,224	1,036
	1,00	117	63,75	8,844	,818

Tabelle 23: Deskriptive Statistiken zu H_0M_3

Tabelle 24 zeigt zuerst den Levene-Test auf Gleichheit der Varianzen und ist mit einem Wert von $p=0,204$ nicht signifikant, was die Berechnung eines T-Tests bei unabhängigen Stichproben ermöglicht. Der T-Test bei unabhängigen Stichproben (siehe Tab. 24) zeigt, dass die Mittelwertsunterschiede mit einem Wert von $p=0,394$ nicht signifikant sind.

Test bei unabhängigen Stichproben

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz	
									Untere	Oberere
WSPMotivationRWGesamt	Varianzen sind gleich	1,629	,204	,854	178	,394	1,153	1,349	-1,510	3,815
	Varianzen sind nicht gleich			,873	135,233	,384	1,153	1,320	-1,458	3,763

Tabelle 24: T-Test für unabhängige Stichproben zu H_0M_3

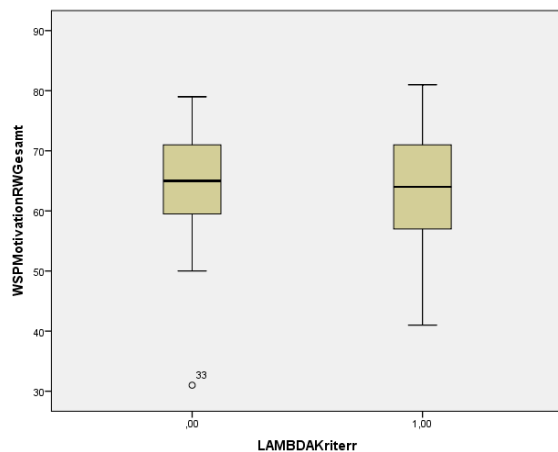


Abbildung 16: Box-Plot zu H_0M_3

Der Box-Plot in Abbildung 22 verdeutlicht, warum keine Mittelwertsunterschiede zu finden waren.

H_0M_4

Diese Hypothese betrifft die Frage, ob sich Personen, die sich unterschätzen, korrekt einschätzen oder überschätzen, in ihrem durchschnittlichen Score in der Skala WSP-Motivation unterscheiden.

Hier soll mit der zweiten Berechnung des Anspruchsniveaus („Anspruchsniveau2“) gerechnet werden. Die zweite Berechnung erfolgt so, dass die tatsächlich richtig kodierten Symbole im 3. Durchgang von der Prognose für den 3. Durchgang subtrahiert werden. (Im Anspruchsniveau1 erfolgt dies mit dem 2. Durchgang und der 2. Prognose).

Tabelle 25 zeigt die deskriptiven Statistiken der Varianzanalyse, die zeigen, dass in den Gruppen 2 und 3 die Personen durchschnittlich 64,01 bzw. 64,81 Punkte in der Skala WSP-Motivation erreicht haben und Personen der Gruppe 1 durchschnittlich 61,38 Punkte erreicht haben.

Deskriptive Statistiken

Abhängige Variable: WSPMotivationRWGesamt

AHAAnspruchsniveau2a ngemessen2	Mittelwert	Standardabweichung	N
1,00	61,38	7,367	8
2,00	64,01	8,845	139
3,00	64,81	7,748	31
Gesamt	64,03	8,587	178

Tabelle 25: Deskriptive Statistiken zur ANOVA zu H_0M_4

Eine Varianzanalyse ist hier die Methode der Wahl, da auch die Voraussetzung der Homogenität der Fehlervarianzen erfüllt wird (siehe Tab. 26; $p=0,364$). Die Mittelwertsunterschiede sind, wie in Tabelle 27, der Varianzanalyse, zu sehen ist, mit einem Wert von $p=0,603$ nicht signifikant.

**Levene-Test auf Gleichheit der
Fehlervarianzen^a**

Abhängige Variable:

WSPMotivationRWGesamt

F	df1	df2	Sig.
1,016	2	175	,364

Prüft die Nullhypothese, daß die Fehlervarianz der abhängigen Variablen über Gruppen hinweg gleich ist.

a. Design: Konstanter Term +
AHAAnspruchsniveau2angemessen2

Tabelle 26: Levene-Test zur ANOVA zu H_0M_4

Tests der Zwischensubjekteffekte

Abhängige Variable: WSPMotivationRWGesamt

Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.
Korrigiertes Modell	75,113 ^a	2	37,556	,506	,603
Konstanter Term	219969,289	1	219969,289	2966,445	,000
AHAAnspruchsniveau2angemessen2	75,113	2	37,556	,506	,603
Fehler	12976,685	175	74,152		
Gesamt	742908,000	178			
Korrigierte Gesamtvariation	13051,798	177			

a. R-Quadrat = ,006 (korrigiertes R-Quadrat = -,006)

Tabelle 27: Tests der Zwischensubjekteffekte zu ANOVA zu H_0M_4

H_0M_5

Diese Hypothese betrifft den Zusammenhang zwischen der Anzahl der richtig bearbeiteten Symbole im Kodiertest und dem Score in WSP-Motivation. Hier wurde ebenfalls eine Korrelation nach Spearman berechnet. Die Ergebnisse sind in Tabelle 28 zu sehen; der Korrelationskoeffizient beträgt 0,104 und ist nicht signifikant ($p=0,168$).

Korrelationen

			WSPMotivati onRWGesamt	AHAInsgricht igbearbeitet
Spearman- Rho	WSPMotivationRWGesamt	Korrelationskoeffizient	1,000	,104
		Sig. (2-seitig)	.	,168
		N	180	178
	AHAInsgrichtigbearbeitet	Korrelationskoeffizient	,104	1,000
		Sig. (2-seitig)	,168	.
		N	178	178

Tabelle 28: Spearman-Rho Korrelation zu H_0M_5

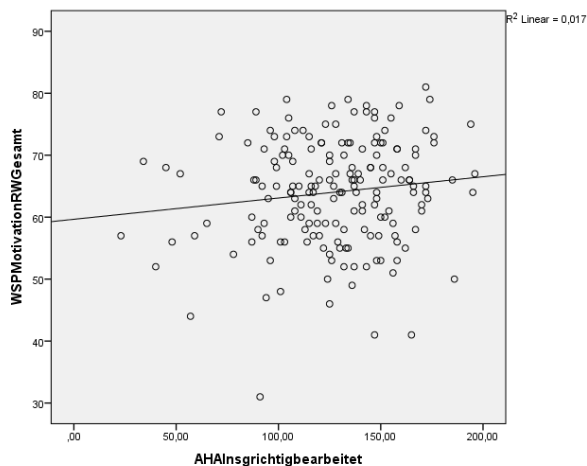


Abbildung 17: Streudiagramm zu H_0M_5

In Abbildung 17 ist zu sehen, dass die Werte kreisförmig angeordnet sind und deswegen keine bedeutende oder signifikante Korrelation ergeben haben.

H_0M_6

Hier soll die Frage beantwortet werden, ob sich Personen, die ihre Leistungsspitze einerseits im 2. Durchgang oder im 3. Durchgang erreicht haben oder in beiden Durchgängen gleich viele Symbole bearbeitet haben, in Bezug auf ihren durchschnittlichen Score in der Skala WSP-Motivation unterscheiden.

In Tabelle 29 ist zu sehen, dass 25 Personen ihre Leistungsspitze im Kodiertest im 2. Durchgang erreicht haben, 136 Personen haben sich im 3. Durchgang noch steigern können und haben im 3. Durchgang ihre Leistungsspitze erreicht, das bedeutet, die meisten Symbole kodiert. 17 Personen erbrachten dieselbe Leistung im 2. und 3. Durchgang. Weiter ist zu sehen (siehe Tab. 29), dass Personen, die ihre Leistungsspitze im 3. Durchgang erreicht haben, durchschnittlich die niedrigsten Werte in der Skala WSP-Motivation erreicht haben

(87,34) im Gegensatz zu Personen, die im 2. Durchgang ihre Leistungsspitze erreicht haben und zu Personen, die ihre Leistung konstant halten konnten.

Ränge			
	2=in 2.Dr, 3=in 3.Dg,4=beide gleich	N	Mittlerer Rang
WSPMotivationRWGesamt	2,00	25	96,78
	3,00	136	87,34
	4,00	17	96,09
	Gesamt	178	

Tabelle 29: Ränge zum Kruskal-Wallis-Test zu H_0M_6

In Tabelle 30 ist zu sehen, dass sich diese Mittelwertsunterschiede mit einem Signifikanzwert von $p=0,601$ nicht signifikant unterscheiden.

Statistik für Test ^{a,b}	
	WSPMotivationRWGesamt
Chi-Quadrat	1,018
df	2
Asymptotische Signifikanz	,601

a. Kruskal-Wallis-Test

b. Gruppenvariable: 2=in 2.Dr, 3=in 3.Dg,4=beide gleich

Tabelle 30: Statistik für Kruskal-Wallis-Test zu H_0M_6

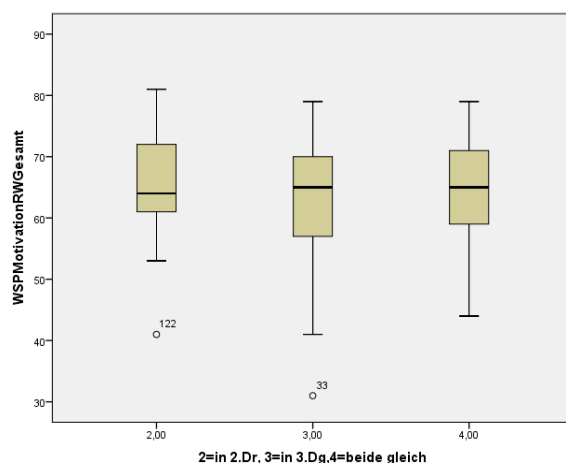


Abbildung 18: Box-Plot zu H_0M_8

In Abbildung 23 ist zu sehen, dass sich die Personen, aufgeteilt nach dem Zeitpunkt ihrer Leistungsspitze, nicht bezüglich ihrer Mittelwerte in der Skala WSP-Motivation unterscheiden.

3.3 Ergebnisse der Hypothesen zu „Frustrationstoleranz“

H_0FT_1

Diese Hypothese betrifft den Zusammenhang zwischen dem Score in der Skala WSP-Frustrationstoleranz und dem Belastbarkeitsquotienten im LAMBDA-neu. Der Belastbarkeitsquotient wird berechnet, indem die Leistung unter Stress mit der Leistung ohne Stress verglichen wird.

Die deskriptive Analyse in Tabelle 31 zeigt, dass 112 Personen eine schlechtere Leistung unter Stress gezeigt haben (Gruppe 1), 32 haben eine genau gleiche Leistung gezeigt (Gruppe 2) und 34 Personen haben sich unter Stress sogar verbessert (Gruppe 3).

LAMBDABelastbarkeit123

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	1,00	112	62,2	62,9
	2,00	32	17,8	80,9
	3,00	34	18,9	100,0
	Gesamt	178	98,9	100,0
Fehlend	System	2	1,1	
Gesamt		180	100,0	

Tabelle 31: Deskriptive Statistik zu LAMBDABelastbarkeit zu H_0FT_1

Da der Levene-Test Gleichheit der Fehlervarianzen für eine potenzielle Varianzanalyse signifikant ausgefallen ist (siehe Tab. 57 im Anhang Kapitel 3.3; $p=0,005$), darf keine Varianzanalyse gerechnet werden, sondern es muss auf ein parameterfreies Verfahren ausgewichen werden.

Tabelle 32 zeigt, dass die mittleren Ränge von 89,71 in der Gruppe 1, 88,67 in der Gruppe 2 und 89,51 in der Gruppe 3 betragen.

Ränge

	LAMBDABelastbarkeit1 23	N	Mittlerer Rang
WSPFrustrationstoleranz	1,00	112	89,73
RWGesamt	2,00	32	88,67
	3,00	34	89,51
	Gesamt	178	

Tabelle 32: Ränge zu Kruskal-Wallis-Test zu H_0FT_1

Tabelle 33 zeigt die Statistik für den Kruskal-Wallis-Test, der nicht signifikant ($p=0,995$) ausgefallen ist.

Statistik für Test ^{a,b}	
	WSPFrustrationstoleranzRWGesamt
Chi-Quadrat	,011
df	2
Asymptotische Signifikanz	,995

a. Kruskal-Wallis-Test

b. Gruppenvariable:
LAMBDABelastbarkeit123

Tabelle 33: Statistik zu Kruskal-Wallis-Test zu H_0FT_1

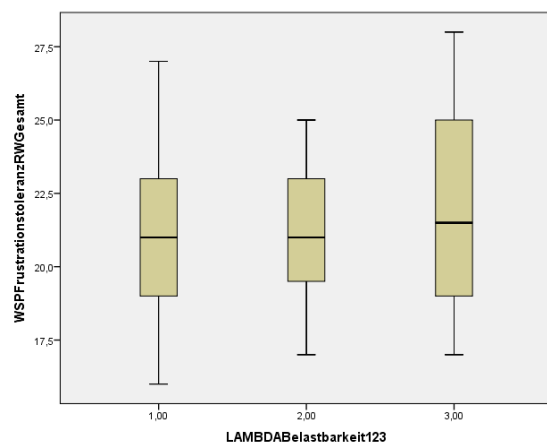


Abbildung 19: Box-Plot zu H_0FT_1

Bei Betrachtung der Box-Plots in Abbildung 19 ist ersichtlich, dass zwischen den drei Gruppen keine Mittelwertsunterschiede in Bezug auf ihren Rohwert in der Skala WSP-Frustrationstoleranz bestehen.

H_0FT_2

Diese Hypothese sollte untersuchen, ob sich Personen, die sich in Bezug auf ihren Score in der Skala WSP-Frustrationstoleranz unterscheiden, unterschiedliche viele Symbole richtig bearbeitet haben im Kodiertest.

Korrelationen

			WSPFrustrationstoleranzRWGesamt	AHAInsgrichtigbearbeitet
Spearman-Rho	WSPFrustrationstoleranzRWGesamt	Korrelationskoeffizient	1,000	,163*
		Sig. (2-seitig)	.	,030
		N	180	178
	AHAInsgrichtigbearbeitet	Korrelationskoeffizient	,163*	1,000
		Sig. (2-seitig)	,030	.
		N	178	178

*. Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig).

Tabelle 34: Spearman-Rho Korrelation zu H_0FT_2

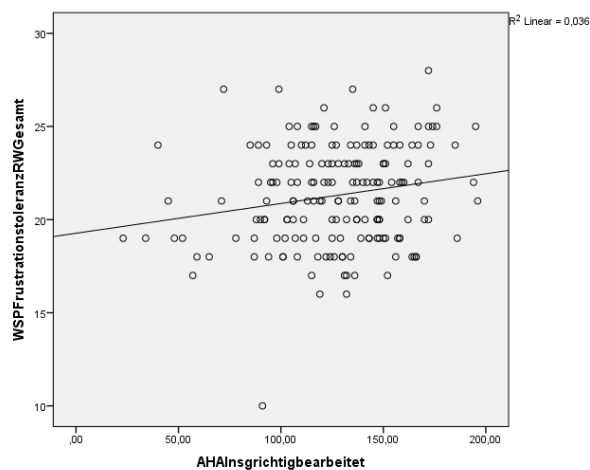


Abbildung 20: Streudiagramm zu H_0FT_2

In Abbildung 20 ist das Streudiagramm der beiden Variablen zu sehen; auch diese Werte sind kreisförmig angeordnet und ergeben somit keine signifikante Korrelation.

H_0FT_3

Hier soll die folgende Frage beantwortet werden: Unterscheiden sich die Personen, die ihrer Leistungsspitze im Kodiertest im 2. Durchgang, im 3. Durchgang oder in beiden erreicht haben, in Bezug auf ihren durchschnittlichen Score in der Skala WSP-Frustrationstoleranz?

Tabelle 35 zeigt in der Spalte (2=in 2.Dr., 3=in 3.Dg, 4=beide gleich) den Zeitpunkt der Leistungsspitze im Kodiertest. Es ist zu sehen, dass die meisten Personen (136) im 3. Durchgang ihre Leistungsspitze erreichen und die meisten Symbole kodieren, 17 Personen können in beiden Durchgängen ihre Leistung beibehalten, und 25 Personen konnten im 2. Durchgang die meisten Symbole kodieren.

Ränge

	2=in 2.Dr, 3=in 3.Dg,4=beide gleich	N	Mittlerer Rang
	2,00	25	97,58
WSPFrustrationstoleranz	3,00	136	89,49
RWGesamt	4,00	17	77,74
	Gesamt	178	

Tabelle 35: Ränge zu Kruskal-Wallis-Test zu H_0FT_3

Tabelle 36 zeigt, dass sich die drei Gruppen nicht bezüglich ihres durchschnittlich erreichten Scores in der Skala WSP-Frustrationstoleranz unterscheiden, was im Box-Plot in Abbildung 27 auch nochmals zu sehen ist.

Statistik für Test^{a,b}

	WSPFrustrationstoleranzRWGesamt
Chi-Quadrat	1,520
df	2
Asymptotische Signifikanz	,468

a. Kruskal-Wallis-Test

b. Gruppenvariable: 2=in 2.Dr, 3=in

3.Dg,4=beide gleich

Tabelle 36: Statistik zu Kruskal-Wallis-Test zu H_0FT_3

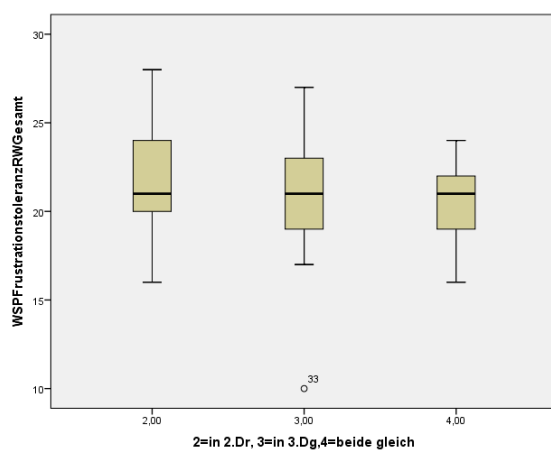


Abbildung 21: Box-Plot zu H_0FT_3

Abbildung 21 zeigt, dass sich die Mittelwerte in den drei Gruppen, bezogen auf den Zeitpunkt ihrer Leistungsspitze, in ihrem Mittelwerten in der Skala WSP-Frustrationstoleranz nicht unterscheiden.

3.4 Ergebnisse der Hypothesen zu „Selbstbewusstsein“

H_0S_1 , H_0S_2 und H_0S_3

Alle drei Hypothesen wurde mittels einer Spearman-Rho Korrelation überprüft (siehe Tab. 37), da die Kennwerte des WITE nicht intervallskaliert sind.

Die erste Hypothese bezog sich auf den Zusammenhang zwischen dem Score in der Skala WSP-Selbstbewusstsein und dem Kennwert „Scheinwissen“ im WITE. Der Kennwert „Scheinwissen“ berechnet sich aus der Summe aller Worte, die ein Teilnehmer/eine Teilnehmerin als bekannt angegeben hat, dann aber den Lückentext nicht korrekt bearbeiten konnte.

In Tabelle 37 ist zu sehen, dass zwischen den beiden Kennwerten kein Zusammenhang besteht ($Rho = 0,054$, $p = 0,472$).

Besteht ein Zusammenhang zwischen dem Score in der Skala WSP-Selbstbewusstsein und dem Wert „Passivwissen“ im WITE? Dieser Zusammenhang wurde ebenso mittels Spearman-Rho Korrelation überprüft; wie in Tabelle 37 zu sehen ist, ist diese Korrelation nicht signifikant ($p = 0,662$; $Rho = 0,035$). Zwischen dem Score in WSP-Selbstbewusstsein und dem „Passivwissen“ im WITE besteht kein korrelativer Zusammenhang.

Besteht ein Zusammenhang zwischen dem Score in der Skala WSP-Selbstbewusstsein und dem Wert „Aktivwissen“ im WITE? Auch dieser Zusammenhang sollte mit einer Spearman-Rho Korrelation überprüft werden. In Tabelle 39 ist zu sehen, dass die Korrelation sehr niedrig und nicht signifikant ausgefallen ist ($Rho = -0,079$, $p = 0,295$). Zwischen WSP-Selbstbewusstsein und dem „Aktivwissen“ im WITE besteht kein korrelativer Zusammenhang.

Korrelationen

			WSPSelbstbewusstseinRWGesamt	LEWITEScheinwissenRWGesamt	LEWITEPassivwissenRWGesamt	LEWITEAktivwissenRWGesamt
Spearman-Rho	WSPSelbstbewusstseinRWGesamt	Korrelationskoeffizient	1,000	,054	,035	-,079
		Sig. (2-seitig)	.	,472	,662	,295
		N	180	180	156	177
H ₀ S ₁	LEWITEScheinwissenRWGesamt	Korrelationskoeffizient	,054	1,000	-,105	-,432**
		Sig. (2-seitig)	,472	.	,191	,000
		N	180	180	156	177
H ₀ S ₂	LEWITEPassivwissenRWGesamt	Korrelationskoeffizient	,035	-,105	1,000	-,409**
		Sig. (2-seitig)	,662	,191	.	,000
		N	156	156	156	154
H ₀ S ₃	LEWITEAktivwissenRWGesamt	Korrelationskoeffizient	-,079	-,432**	-,409**	1,000
		Sig. (2-seitig)	,295	,000	,000	.
		N	177	177	154	177

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

Tabelle 37: Spearman-Rho Korrelationen zu H₀S₁, H₀S₂ und H₀S₃

4. Darstellung der Ergebnisse der MTMM-Matrix

In Tabelle 38 sind die Ergebnisse der Spearman-Rho-Korrelationen zu sehen. Die Legende ist in Tabelle 39 zu finden.

Die Felder der konvergenten Validität (rot) haben den Mittelwert von 0,053. Dieser Wert ist äußerst gering und unterscheidet sich nicht bedeutsam von null. Damit liegt keine konvergente Validität vor.

Die Felder der diskriminanten Validität, zuerst die grünen, haben beim WSP den Mittelwert von 0,4235. Bei vorliegender diskriminanter Validität sollte für den WSP gelten, dass der Mittelwert der konvergenten Validitätskoeffizienten höher sein sollte als der Mittelwert der diskriminanten Validitätskoeffizienten für den WSP. Dies ist hier nicht der Fall, da der Wert des diskriminanten Validitätskoeffizienten deutlich höher ist als der des konvergenten Validitätskoeffizienten und nicht umgekehrt. Deswegen kann für den Fall des WSP nicht von einer diskriminanten Validität gesprochen werden.

Die diskriminanten Validitätskoeffizienten für die anderen Methoden, LAMBDA-neu, AHA Kodiertest und WITE, sollen außen vor bleiben, da hier nur die Prüfung der Validität des WSP von Bedeutung ist.

Die Felder der diskriminanten Validitätskoeffizienten für den WSP, in Tabelle 38 blau dargestellt, haben den Mittelwert von 0,103. Im Falle einer vorliegenden diskriminanten Validität sollte dieser Wert der kleinste von allen sein, kleiner als der erste Wert der diskriminanten Validität und beide sollten kleiner sein als der Wert der konvergenten Validität. Dieser Wert ist zwar klein, jedoch doch höher als der Wert der konvergenten Validität, was auf keinen Fall sein darf.

Korrelationen

		WSPOF	WSPM	WSPFT	WSPS	LAMBDA PPh	LAMBDA 0F	LAMBDA Gld	REL Lz1	LAMBDA Lstr	AHA Irb	AHA An1	AHA An2	LAMBDA Krit	AHA Zd	LAMBDA Bqu	AHA Lsp	WITEAW	WITEPW	WITESW
Spearman-Rho	WSPOF	1																		
	WSPM	,624	1																	
	WSPFT	,329	,540	1																
	WSPS	,182	,501	,365	1															
	LAMBDA PPh	,029	-,055	,024	,064	1														
	LAMBDA 0F	,056	,053	,043	,098	,664	1													
	LAMBDA Gld	-,025	,053	,011	,093	,144	,182	1												
	REL Lz1	,055	,040	-,008	-,114	-,394	,050	-,295	1											
	LAMBDA Lstr	,036	,024	-,042	-,084	-,540	-,070	-,194	,826	1										
	AHA Irb	,074	,104	,163	,032	,087	,064	-,140	,066	,025	1									
	AHA An1	-,045	,043	,076	,082	,001	-,086	,014	-,047	-,055	,198	1								
	AHA An2	-,027	,067	,136	,079	-,007	,011	-,055	,101	,133	,328	,193	1							
	LAMBDA Krit	,071	-,061	-,035	-,154	-,201	-,249	-,387	,139	,218	,138	-,023	,053	1						
	AHA Zd	,105	-,023	-,034	-,052	,045	,133	-,094	,017	,010	,040	-,747	-,244	,018	1					
	LAMBDA Bqu	-,172	-,098	,001	-,022	-,018	-,010	-,042	,069	-,010	,060	,055	-,011	,130	,026	1				
	AHA Lsp	,019	-,017	-,091	-,040	-,024	,004	-,005	-,032	,040	-,207	,095	-,137	-,043	-,258	-,087	1			
	WITEAW	,123	,065	-,028	-,079	-,034	,070	-,025	,142	,208	,076	,025	,066	,140	-,010	,037	,050	1		
	WITEPW	-,043	-,057	-,122	,035	-,146	-,183	,036	-,121	-,054	-,041	,043	-,052	,063	-,125	,049	,067	-,409	1	
	WITESW	,033	,005	,060	,054	,222	,013	-,072	-,083	-,157	,012	-,033	,008	-,076	,096	-,184	-,040	-,432	-,105	1

Tabelle 38: Darstellung der berechneten MTMM – Matrix; Ergebnisse

Legende:

WSPOF	Score in WSP-Organisation	REL Lz1	LAMBDA Relative Lernzeit 1	LAMBDA Bqu	LAMBDA Belastbarkeitsquotient
WSPM	Score in WSP-Motivation	LAMBDA Lstr	LAMBDA Lernstrategie	AHA Lsp	Kodiertest Zeitpunkt d. Leistungsspitze
WSPFT	Score in WSP-Frustrationstoleranz	AHA Irb	Kodiertest Insg. richtig bearb. Symbole	WITEAW	WITE Aktivwissen
WSPS	Score in WSP-Selbstbewusstsein	AHA An1	Kodiertest Anspruchsniveau 1	WITEPW	WITE Passivwissen
LAMBDA PPh	LAMBDA Anz. d. Prüfphasen	AHA An2	Kodiertest Anspruchsniveau 2	WITESW	WITE Scheinwissen
LAMBDA 0F	LAMBDA Anz. d. Nullfehlerdrg.	LAMBDA Krit	LAMBDA Kriterium erreicht		
LAMBDA Gld	LAMBDA Gesamtlerndauer	AHA Zd	Kodiertest Zieldiskrepanz		

Tabelle 39: Legende zu MTMM-Matrix

Die postulierte Reihenfolge der durchschnittlichen Korrelationen sollte wie folgt aussehen (siehe Tab. 40):

Monotrait-Monomethod	>	Monotrait-Heteromethod	>	Heterotrait-Monomethod	>	Heterotrait-Heteromethod
dunkelgrau	>	rot	>	grün	>	blau

Tabelle 40: Angenommene Reihenfolge der durchschn. Korrelationen der MTMM-Matrix

Die gefundene Reihenfolge der durchschnittlichen Korrelationen sieht allerdings wie folgt aus (siehe Tab. 41):

Monotrait-Monomethod	>	Heterotrait-Monomethod	>	Heterotrait-Heteromethod	>	Monotrait-Heteromethod
dunkelgrau	>	grün	>	blau	>	rot

Tabelle 41: Gefundene Reihenfolge der durchschn. Korrelationen der MTMM-Matrix

Die verschiedenen Methoden (Persönlichkeitsfragebogen WSP, objektive Persönlichkeitstests LAMBDA-neu, Kodiertest und WITE) scheinen nicht einen Trait zu messen, wie an den niedrigen konvergenten Validitätskoeffizienten zu sehen ist. Der Persönlichkeitsfragebogen WSP scheint, aufgrund der hohen Korrelationen der Skalen untereinander, ähnliche Konstrukte zu erfassen. Außerdem scheinen die verschiedenen Methoden, obwohl sie unterschiedliche Konstrukte zu messen beabsichtigen (blaue Felder), doch ähnliche Konstrukte zu erfassen.

Somit kann weder eine konvergente noch diskriminante Validität bescheinigt werden.

5. Exploratorische Faktorenanalyse

Nachdem bei einem überwiegenden Großteil der Hypothesen die Nullhypothese beibehalten werden musste (siehe Kapitel 3; Empirischer Teil), sollte mittels einer exploratorischen Faktorenanalyse überprüft werden, ob die Objektive Persönlichkeitstests und der WSP auf unterschiedlichen Faktoren laden.

Nachdem sich in Kapitel 3; Empirischer Teil, gezeigt hat, dass die Kennwerte der Objektiven Persönlichkeitstests nicht mit den Kennwerten des Persönlichkeitsfragebogens WSP korrelieren, muss davon ausgegangen werden, dass die beiden Modalitäten unterschiedliche Konstrukte oder unterschiedliche Aspekte eines Konstruktes erfassen.

In die Faktorenanalyse wurden folgende Kennwerte der Objektiven Persönlichkeitstests und des Persönlichkeitsfragebogens einbezogen:

LAMBDAUnnoetige0FehlerDurchgänge	Anzahl der unnötigen Nullfehler-Durchgänge im LAMBDA-neu
LAMBDAAnzahlPruefphasen	Anzahl der durchlaufenen Prüfphasen im LAMBDA-neu
LAMBDAGesamtlerndauer	Gesamtlerndauer im LAMBDA-neu
RelLernzeit1	Anteil der in der ersten Lernphase verbrachten Zeit im Vergleich zur Gesamtlerndauer im LAMBDA-neu
LAMBDAKriterr	Kriterium erreicht/nicht erreicht im LAMBDA-neu
AHAAnspruchsniveaunumerischauf3Dezimalstellen gerundet	Anspruchsniveau 1 im Kodiertest
AHAAnspruchsniveau2	Anspruchsniveau 2 im Kodiertest
AHAInsgrichtigbearbeitet	Insgesamt richtig bearbeitete Symbole im Kodiertest
2=in 2.Dr, 3=in 3.Dg,4=beide gleich	Zeitpunkt der Leistungsspitze im Kodiertest
LAMBDABelastbarkeit123	Leistung unter Belastung im LAMBDA-neu
LEWITEAktivwissenRWGesamt	„Aktivwissen“ im WITE
LEWITEPassivwissenRWGesamt	„Passivwissen“ im WITE
LEWITEScheinwissenRWGesamt	„Scheinwissen“ im WITE
WSPOrganisationRWGesamt	Rohwert in WSP-Organisation
WSPMotivationRWGesamt	Rohwert in WSP-Motivation
WSPFrustrationstoleranzRWGesamt	Rohwert in WSP-Frustrationstoleranz
WSPSelbstbewusstseinRWGesamt	Rohwert in WSP-Selbstbewusstsein

Die Faktorenanalyse ergab den in Abbildung 22 dargestellten Screeplot. Es ist zu sehen, dass beim 5. Faktor ein „Knick“ ist (siehe Pfeil in Abb. 22). Dies wäre ein Abbruchkriterium für die weitere Faktorenanzahlbestimmung.

Das Kaiser-Kriterium, welches die Faktorenextraktion bei einem Eigenwert von über 1 weiterführt und sobald es unter 1 kommt, diese abbricht, kommt zu einer Faktorenlösung mit 7 Faktoren (siehe Tab. 42), die insgesamt 68,53% der Gesamtvarianz erklären.

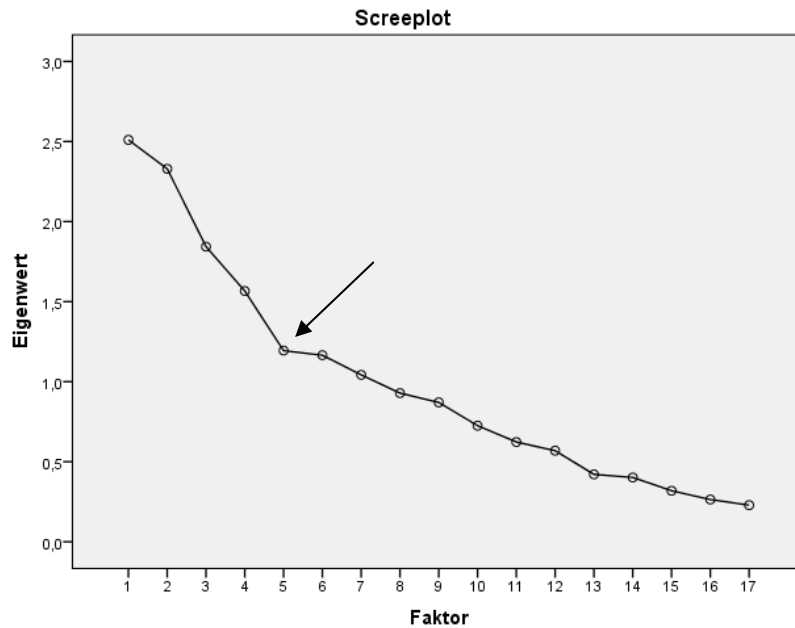


Abbildung 22: Screeplot zur Faktorenanalyse

Erklärte Gesamtvarianz						
Komponente	Anfängliche Eigenwerte			Rotierte Summe der quadrierten Ladungen		
	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %
1	2,510	14,766	14,766	2,292	13,482	13,482
2	2,330	13,703	28,469	1,791	10,533	24,015
3	1,843	10,843	39,312	1,739	10,227	34,242
4	1,566	9,212	48,524	1,709	10,053	44,295
5	1,194	7,023	55,547	1,663	9,781	54,076
6	1,165	6,854	62,400	1,231	7,241	61,317
7	1,042	6,132	68,532	1,227	7,215	68,532
8	,928	5,461	73,994			
9	,870	5,120	79,114			
10	,725	4,266	83,380			
11	,623	3,665	87,045			
12	,569	3,347	90,392			
13	,421	2,476	92,868			
14	,402	2,362	95,230			
15	,319	1,874	97,104			
16	,264	1,552	98,656			
17	,229	1,344	100,000			

Extraktionsmethode: Hauptkomponentenanalyse.

Tabelle 42: Erklärte Gesamtvarianz für die Faktorenanalyse

In Tabelle 43 ist die Rotierte Komponentenmatrix mit 7 Faktoren dargestellt. Es ist zu sehen (siehe Tab. 43), dass die vier Skalen des *Wiener Studieneignungs-Persönlichkeitsinventars* alle gemeinsam auf den ersten Faktor laden und dass kein Kennwert der Objektiven Persönlichkeitstests auf diesen ersten Faktor lädt.

Damit kann gesagt werden, dass die Kennwerte der Objektiven Persönlichkeitstests nicht auf den gleichen Faktoren laden, beziehungsweise auf gänzlich anderen, als die Kennwerte des Persönlichkeitsfragebogens.

Rotierte Komponentenmatrix ^a							
	Komponente						
	1	2	3	4	5	6	7
LAMBDAUnnoetige0FehlerDurchgaenge		,787					
LAMBDAAnzahlPruefphasen		,670		-,348			
LAMBDAGesamtlerndauer				-,731			
RelLernzeit1				,664			
LAMBDAKriterr		-,521		,521			
AHAAnspruchsniveaunumerischauf3Dezimalstellengerundet							,809
AHAAnspruchsniveau2			,761				
AHAInsgrichtigbearbeitet			,719				,472
2=in 2.Dr, 3=in 3.Dg,4=beide gleich			-,711				
LAMBDABelastbarkeit123						,852	
LEWITEAktivwissenRWGesamt					,837		
LEWITEPassivwissenRWGesamt		-,540		-,420		,376	
LEWITEScheinwissenRWGesamt					-,823		
WSPOrganisationRWGesamt	,686						
WSPMotivationRWGesamt	,885						
WSPFrustrationstoleranzRWGesamt	,734						
WSPSelbstbewusstseinRWGesamt	,641						

Extraktionsmethode: Hauptkomponentenanalyse.

Rotationsmethode: Varimax mit Kaiser-Normalisierung.

a. Die Rotation ist in 12 Iterationen konvergiert.

Tabelle 43: Rotierte Komponentenmatrix zur Faktorenanalyse

IV Zusammenfassung der Ergebnisse und Interpretation

In diesem Kapitel sollen die Ergebnisse zusammenfassend jeweils in Bezug auf die einzelnen Dimensionen und einzelnen Überprüfungsmethoden betrachtet, und in Bezug auf die Fragestellung interpretiert werden.

Bezüglich der Skala Organisation ist zu sagen, dass die ursprünglich aufgestellten Hypothesen (H_{0OF_1} bis H_{0OF_3}) alle nicht signifikant ausgefallen sind, das bedeutet, dass kein Zusammenhang zwischen dem Ergebnis in der Skala WSP-Organisation und den Kennwerten des LAMBDA-neu (Lernstrategie, Kriterium erreicht und der Relativen Lernzeit 1) besteht. Diese Ergebnisse verdeutlichen, dass es keinen Zusammenhang zwischen den Kennwerten des LAMBDA-neu und der Skala Organisation des WSP gibt. Diese sprechen dafür, dass die Kennwerte des LAMBDA-neu nicht für eine Konstruktvalidierung des vorliegenden Persönlichkeitsfragebogens geeignet sind.

Bezüglich der Skala Motivation wurden die folgenden Ergebnisse gefunden: Alle im Vorhinein aufgestellten Hypothesen (H_{0M_1} bis H_{0M_6}) sind nicht signifikant ausgefallen, was bedeutet, dass kein Zusammenhang zwischen dem Ergebnis in der Skala WSP-Motivation und den Werten im LAMBDA-neu (Lernstrategie, Relative Lernzeit 1, Kriterium erreicht) und im Kodiertest (Anspruchsniveau, Leistungsniveau und Zeitpunkt der Leistungsspitze) besteht. Offensichtlich besteht kein Zusammenhang zwischen den betreffenden Kennwerten des LAMBDA-neu und des Kodiertests und der Skala WSP-Motivation. Diese Ergebnisse sprechen ebenfalls dafür, dass die Kennwerte des LAMBDA-neu, und hier auch des Kodiertests, nicht für eine Konstruktvalidierung zweckmäßig sind.

Die aufgestellten Hypothesen zu Frustrationstoleranz H_{0FT_1} bis H_{0FT_3} waren bis auf eine nicht signifikant, was bedeutet, dass zwischen dem Score in der Skala WSP-Frustrationstoleranz und dem Belastbarkeitsquotienten im LAMBDA-neu und der Position der Leistungsspritze im Kodiertest kein signifikanter Zusammenhang besteht. Wohl besteht aber ein signifikanter Zusammenhang zwischen der Anzahl der richtig bearbeiteten Symbole im Kodiertest und dem Score in WSP-Frustrationstoleranz, wenngleich dieser Zusammenhang gering ausgefallen ist.

Die Hypothesen H_{0S_1} bis H_{0S_3} sind allesamt nicht signifikant ausgefallen, was dafür spricht, dass es keinen Zusammenhang zwischen den Kennwerten des WITE („Scheinwissen“, „Passivwissen“ und Aktivwissen“) und der Skala WSP-Selbstbewusstsein gibt. Auch hier

lautet die Schlussfolgerung, dass die Kennwerte aus dem WITE nicht geeignet sind für eine Konstruktvalidierung des vorliegenden Fragebogens.

Weiter konnte im Multitrait-Multimethod-Ansatz weder die konvergente Validität, die sehr niedrige Korrelationen hervorbrachte, anstatt hoher, noch die diskriminante Validität, die teils hohe Korrelationen beziehungsweise nicht niedrigere als die der konvergenten Validität hervorgebracht hat, bestätigt werden. Die hohen diskriminanten Korrelationen im Heterotrait-Monomethod-Block zeigen die hohen Korrelationen der Skalen des WSP untereinander, was bedeutet, dass alle Skalen hoch miteinander korrelieren, was wiederum die Skalenstruktur beziehungsweise die Aufteilung auf die vier Skalen fragwürdig erscheinen lässt. Die niedrigen Korrelationen des Monotrait-Heteromethod-Blocks wiederholen nur die niedrigen Korrelationen der einzelnen Skalen mit den Kennwerten der objektiven Persönlichkeitstests. Einzig die Heterotrait-Heteromethod-Korrelationen sind niedrig und nähern sich stark null an, was an sich für das Vorliegen einer diskriminanten Validität sprechen würde, jedoch sollten diese Korrelationen die niedrigsten von allen sein, was aber nicht der Fall ist, da die konvergenten Validitätskoeffizienten niedriger sind, was sie auf keinen Fall sein dürften, bei Vorliegen einer diskriminanten Validität.

Eine abschließende Faktorenanalyse, die prüfen sollte, ob die Kennwerte des Persönlichkeitsfragebogens auf gleiche Faktoren wie die der Objektiven Persönlichkeitstests laden, hat ergeben, dass die einzelnen Skalen des WSP auf einen Faktor laden und die Kennwerte der Objektiven Persönlichkeitstests auf jeweils andere Faktoren laden. Dies zeigt, dass die beiden Erhebungsmethoden offensichtlich etwas anderes erfassen.

Eine vorschnelle Schlussfolgerung, dass die gefundenen Ergebnisse gegen eine Konstruktvalidität der vorliegenden Skalen des *Wiener Studieneignungs-Persönlichkeitsinventars* sprechen, wäre nicht angebracht, wie später im Kapitel V Diskussion nachgelesen werden kann.

In Bezug auf die zweite Fragestellung ist eindeutig zu sagen, dass Objektive Persönlichkeitstests zur Validierung von Persönlichkeitsfragebogen nicht zurate gezogen werden können.

In Kapitel V wird diskutiert, warum das *Wiener Studieneignungs-Persönlichkeitsinventar* inhaltlich valide ist und welche Schlussfolgerungen daraus in Bezug auf die Konstruktvalidität des betreffenden Persönlichkeitsfragebogens gezogen werden können und wie der Nutzen von Objektiven Persönlichkeitsfragebogens zur Validierung eines Persönlichkeitsfragebogens bewertet werden kann.

V Diskussion der Ergebnisse und Kritik

In diesem Kapitel sollen die Ergebnisse diskutiert werden; die Beantwortungen der Fragestellungen der Arbeit sollen kritisch betrachtet werden; es soll eine kritische Auseinandersetzung mit der gewählten Validierungsmethode und deren konkreter Umsetzung vorgenommen werden, sowie auch Ideen für weitere Validierungsversuche dargestellt werden.

Es hat sich gezeigt, dass die Kennwerte der Objektiven Persönlichkeitstests und des Persönlichkeitsfragebogens bis auf eine Ausnahme nicht signifikant miteinander korrelieren. Die korrelative Auswertung der Multitrait-Multimethod-Matrix kann keine konvergente oder diskriminante Validität bestätigen. Die Faktorenanalyse hat gezeigt, dass der Persönlichkeitsfragebogen und die Objektiven Persönlichkeitstests nicht auf gleiche Faktoren laden.

Die Korrelation von Kennwerten von Objektiven Persönlichkeitstests mit Persönlichkeitsfragebögen hat sich in der vorliegenden Untersuchung als nicht erfolgreich erwiesen. Ein Großteil der gefundenen Korrelationen war nicht signifikant, geschweige denn bedeutsam. Dieses Ergebnis stimmt mit den bisher in der Validierungsforschung gefundenen Resultaten überein, dass objektive Persönlichkeitstests nicht mit konstruktnahen Persönlichkeitsfragebogen korrelieren (Ortner et. al., 2007). Dies wurde in der Validierungsstudie der „*Arbeitshaltungen*“ (Kubinger & Litzenberger, 2003) deutlich, in der ebenso keine relevanten Zusammenhänge zwischen den Kennwerten der objektiven Persönlichkeitstestbatterie „*Arbeitshaltungen*“ und dem Persönlichkeitsfragebogen NEO-FFI gefunden werden konnten. In der vorliegenden Diplomarbeit wurden kongruente Ergebnisse gefunden.

Ein grundsätzliches Problem bei der Validierung eines Tests ist die Frage, inwieweit bei einem Nichtvorfinden von Beweisen für eine Konstruktvalidität zwingend davon ausgegangen werden muss, dass diese nicht vorliegt. Wenn die Ergebnisse der Untersuchung mit der Theorie nicht übereinstimmen, ist wie Kane (2001) bereits argumentiert hat fraglich, ob das Problem beim Konstrukt liegt (in diesem Fall also in den vier Skalen des WSP, beziehungsweise der Itemauswahl, -anzahl oder -zusammensetzung) im nomologischen Netz (also hier in der angenommenen Korrelation zwischen Kennwerten objektiver Persönlichkeitstests und einem Persönlichkeitsfragebogen) oder in den Details der Messung

(hier wäre das die Vorgabe über das Internet, die damit verbundene hohe Selbstselektion der Teilnehmer und somit homogene Stichprobe, die zu geringeren Korrelationen führt).

Wenn hier der erste Fall zutreffen würde, also ein Problem im Konstrukt vorliegen würde, so ließe sich ohne weiteres von einer nicht vorhandenen Konstruktvalidität sprechen. Dagegen würden die folgenden Punkte sprechen: Die Skalen und Items wurden aus einem Itempool, eines nach strengen Kriterien (DIN 33430) erstellten Itempools von Maurer (2010), speziell für die Verwendung im **Wiener Self-Assessment Psychologie®** ausgewählt und zusammengestellt, was grundsätzlich für die Qualität desselben spricht. Man kann also aufgrund der Itemerstellung auf Basis einer empirisch fundierten Anforderungsanalyse und der Überprüfung der Sinnhaftigkeit durch Experten von einer vorliegenden inhaltlichen Validität der Items ausgehen. Der zweite Grund liegt darin, dass nach Durchsicht der Items, diese zumindest augenscheinlich größtenteils die postulierten Konstrukte messen. Weiter hat Dörner (2009) in seiner Arbeit die Gültigkeit des Rasch-Modells für einen Großteil der Items des WSP bestätigen können, aus denen dann die verwendeten Items für die Kurzversion des WSP für das **Wiener Self-Assessment Psychologie®** ausgewählt wurden. Somit ist von einer inhaltlichen Gültigkeit und einer Raschhomogenität der Items ausgehen.

Wenn man nun sagen würde, dass die gefundenen Ergebnisse nicht auf einem Problem im Konstrukt basieren, bleiben noch die beiden übrigen Möglichkeiten.

Das Problem in der vorliegenden Untersuchung könnte im nomologischen Netz liegen. Wie sich in der Vergangenheit gezeigt hat (vgl. Kubinger & Litzenberger, 2003, Skinner & Horwath, 1973) bringt die Korrelation von Kennwerten von objektiven Persönlichkeitstests mit Persönlichkeitsfragebögen nicht die gewünschten (hohen) Zusammenhänge, was wiederum an den folgenden Punkten liegen könnte: Einerseits ist es besonders schwierig objektive Persönlichkeitstests per se zu validieren. Die Schwierigkeiten liegen darin, dass die üblichen Validierungsmethoden bei Objektiven Persönlichkeitstests nicht greifen, da es bisher nur wenige konstruktnahe andere Objektive Persönlichkeitstests gibt (Ortner et. al., 2007). Ferner stellt sich die Frage, was Objektive Persönlichkeitstests, hier am Beispiel von den „*Arbeitshaltungen*“, nun genau messen, wenn sie ihrerseits nicht auf die übliche Weise validiert werden können, und ob sie einen besonderen Aspekt der Persönlichkeit erfassen, der über die inhaltliche Beschreibung in einem Fragebogen hinausgeht (Kubinger & Litzenberger, 2003). Überdies brachte eine kürzlich versuchte Extremgruppenvalidierung des LAMBDA-neu (Hauser, 2011) nicht die gewünschten Ergebnisse. Bei dieser Studie

wurde untersucht, ob das Verfahren LAMBDA-neu zwischen Personen, die typischerweise „lernendes“ und „verstehendes“ Lernverhalten zeigen, unterscheiden kann (Hauser, 2011). Es wurde gefunden, dass der LAMBDA-neu dies nicht vermag (Hauser, 2011). Dies würde bedeuten, dass der LAMBDA-neu nicht die postulierten Konstrukte erfasst, was wiederum die geringe Korrelation mit dem Persönlichkeitsfragebogen erklärt. Lediglich die Persönlichkeitskomponente „Selbstvertrauen“ des LEWITE konnte in einer Extremgruppenvalidierung (Wawrinec, 2006) validiert werden.

Daraus ergeben sich die folgenden Schlussfolgerungen: Beim Versuch einen Persönlichkeitsfragebogen mit Kennwerten Objektiver Persönlichkeitstests zu validieren stößt man auf Probleme, da diese unterschiedlichen Methoden unterschiedliche Konstrukte oder unterschiedliche Aspekte ein und desselben Konstrukts erfassen. Das Ergebnis sind sehr geringe oder keine Korrelationen zwischen Testkennwerten Objektiver Persönlichkeitstests und Fragebogenskalen. Das bedeutet, dass in dieser Untersuchung wohl falsche Annahmen über das nomologische Netz zugrunde lagen, da Objektive Persönlichkeitstests nicht dieselben Konstrukte messen wie Persönlichkeitsfragebogen. Die Annahmen über die Zusammenhänge wurden in Kapitel 3 des Empirischen Teils überprüft und dargestellt und widerlegen diese.

Ein anderer Grund für diese Ergebnisse könnte in den Details der Messung liegen, also in der computerisierten Vorgabe über das Internet. Einerseits ist bei Datenerhebungen im Internet die Repräsentativität der Stichprobe zu hinterfragen (Hussy, Schreier & Echterhoff, 2010). Die Teilnehmer selektieren sich selbst für die Erhebungen, was bedeutet, dass sich dadurch keine Zufallsstichprobe ergibt, was die Generalisierbarkeit der Ergebnisse in Frage stellt (Hussy et al., 2010). Zusätzlich stellt sich auf Basis dessen die Frage, welche Personen das **Wiener Self-Assessment Psychologie®** durchgeführt haben, oder welche grundsätzlich ein Self-Assessment durchführen, wie ihre Persönlichkeit ist. Eine Annahme wäre, dass es besonders unsichere Personen sind, die nicht wissen ob das Studium das Richtige für sie ist und die sich ihrer eigenen Fähigkeiten und ihres Könnens nicht sicher beziehungsweise bewusst sind, und gerne über das Self-Assessment mehr über sich erfahren möchten. Eine andere Annahme wäre, dass es Personen sind, die hoch motiviert, informiert und hohe und/oder geeignete Fähigkeiten besitzen und sich trotzdem noch gerne verbessern möchten, aber trotzdem wissen, dass sie bereits „gut“ sind, was auch die Bereitschaft, sich selbst zu testen erklärt, da sie vor den Ergebnissen keine Angst zu haben brauchen.

Da sich die Personen erstens selbst für die Teilnahme selektiert haben, zweitens man davon ausgehen muss, dass bestimmte Persönlichkeitstypen ein Self-Assessment durchführen und drittens aufgrund des hohen Zeitaufwands, den das **Wiener Self-Assessment Psychologie**® erfordert (Durchführungsdauer: 90-120 Minuten), folgt, dass die Stichprobe wohl homogener war als bei einer tatsächlich zufälligen Auswahl der Teilnehmer. Diese Homogenität der Stichprobe führt zu einer Minderung der Varianz in der Stichprobe, was ihrerseits wiederum zu geringeren Korrelationen führt (Kubinger, 2009). Sollte dies eingetreten sein, verwundern die geringen Korrelationen zwischen den Werten nicht.

Zusätzlich haben wohl einige technische Probleme (Nichterreichbarkeit des Self-Assessments, Abstürzen der Seite während der Durchführung) einige potenzielle TeilnehmerInnen von der (vollständigen) Teilnahme abgeschreckt. In Kombination mit der langen Durchführungsdauer, die eine hohe Motivation, Konzentration und ein hohes Engagement der TeilnehmerInnen erfordern, dann die Bemühungen bei der Testpersonenakquirierung, die nur bestimmte Personen erreicht haben und motivieren konnten am **Wiener Self-Assessment Psychologie**® teilzunehmen, haben sicherlich auch dazu beigetragen, dass die Stichprobe homogen war.

Durch die Akquirierung in Schulen sowie in einer Maturaschule in Wien, und damit der Tatsache, dass ein Teil der Stichprobe sich aus TeilnehmerInnen zusammensetzt, die noch keine Matura (Abitur), also keine Hochschulreife, haben, kann die Sinnhaftigkeit für diesen Teil der Stichprobe in Frage gestellt werden. Einige Items des *Wiener Studieneignungs-Persönlichkeitsinventars* beziehen sich auf Situationen, die schon eine gewisse Erfahrung mit dem Studienalltag voraussetzen, beziehungsweise voraussetzen, dass die Person, die die Fragen beantwortet, bereits die Matura bestanden hat und über gewisse Fähigkeiten, die damit einhergehen, bereits verfügt. Hier wird also vorausgesetzt, dass sich die TeilnehmerInnen in die Situation eines Studierenden hineinversetzen können und einschätzen können, wie sie sich in der Situation verhalten würden. Es ist fraglich, ob die Personen dies können und damit fraglich, ob die betreffenden Items angemessen sind für diese TeilnehmerInnen.

Zusammenfassend lässt sich sagen, dass die vorliegende Untersuchung vermutlich aufgrund der nun widerlegten Annahmen über die Zusammenhänge zwischen Kennwerten Objektiver

Persönlichkeitstests und Persönlichkeitsfragebogen (nomologisches Netz) und den Details der Messung (Durchführung im Internet; damit verbundene Selbstselektion der TeilnehmerInnen und damit geringere Varianz und geringere Korrelationen) zu den beschriebenen Ergebnissen geführt hat.

Nichtsdestoweniger wurden einige bisher gefundene Ergebnisse repliziert: Es besteht keine oder eine nur sehr geringe Korrelation zwischen Kennwerten Objektiver Persönlichkeitstests und Persönlichkeitsfragebogen. Es wurde gezeigt, dass die Skalen des WSP (*Organisation, Motivation, Frustrationstoleranz und Selbstbewusstsein*) hoch miteinander korrelieren, was wiederum die Frage aufwirft, ob die Skalen überlappende Konstrukte messen. Eine gewisse inhaltliche Überschneidung der Konstrukte ist jedoch inhaltlich nachvollziehbar und für eine Konstruktvalidierung nicht problematisch. Inhaltlich und auch für eine förderungsorientierte Beratung ist es sinnvoll, diese Konstrukte getrennt voneinander zu erfassen und rückzumelden.

Objektive Persönlichkeitstests messen also andere Konstrukte oder andere Aspekte derselben Konstrukte als Persönlichkeitsfragebogen. Durch die nicht gefundenen Zusammenhänge kann gesagt werden, dass Objektive Persönlichkeitstests einen besonderen Aspekt der Konstrukte erfassen, die nicht mittels einer Selbstbeschreibung in einem Persönlichkeitsfragebogen erhoben werden können.

Einerseits mag dies eine Erklärung für die nicht geglückte Konstruktvalidierung sein, andererseits zeigt die Studie auch die Bedeutung von Objektiven Persönlichkeitstests auf. Sie erfüllen einen wichtigen Zweck sowohl im **Wiener Self-Assessment Psychologie®**, als auch in der Praxis. Darüber hinaus ermöglichen sie eine Persönlichkeitsbeschreibung ohne Introspektionsfähigkeit und sozialer Erwünschtheit (Ortner et. al., 2007). Zwar möge man argumentieren, dass die Personen, die ein Self-Assessment durchführen, ihre Werte nicht verfälschen möchten, da sie eine ehrliche Selbstauskunft über sich haben wollen, andererseits sind vielleicht viele Personen nicht in der Lage sich rational einzuschätzen und können sich nicht realistisch beschreiben. Die Objektiven Persönlichkeitstests ermöglichen ihnen trotzdem eine rationale Selbstauskunft, da in einer standardisierten Testsituation auf Basis des (Leistungs-)Verhaltens der Testperson auf die zugrundeliegenden Persönlichkeitsmerkmale geschlossen wird (Ortner et. al., 2007). Somit sind Objektive Persönlichkeitstests im **Wiener Self-Assessment Psychologie®** und ganz allgemein von großer Bedeutung. Dass sie etwas anderes als Persönlichkeitsfragebogen erfassen kann eine Stärke

von ihnen sein, da sie vielleicht zusätzlich zu Leistungstests und Persönlichkeitsfragebogen einen informativen Mehrwert liefern. Interessant wäre eine Studie, die sich mit diesem Mehrwert beschäftigt und die aufzeigt, wie viel jede einzelne Methode zur Vorhersage von zum Beispiel Studienerfolg beiträgt.

In Bezug auf Anknüpfungspunkte für zukünftige Untersuchungen wäre in erster Linie zu nennen, dass es von besonderem Interesse wäre, mehr über die Persönlichkeitsstruktur von TeilnehmerInnen von Self-Assessments zu erfahren. Sind diese besonders motiviert, fähig, unsicher und/oder durchhaltefähig? Warum führen sie ein Self-Assessment durch? Sind sie sich ihres Könnens und ihrer Fähigkeiten bewusst und scheuen deswegen die Durchführung nicht, oder sind sie besonders unsicher und suchen Bestätigung? Ist die Gruppe der Responder, also Teilnehmer am Self-Assessment, repräsentativ für eine Validierungsstudie? Wie unterscheiden sich diese Personen, die das **Wiener Self-Assessment Psychologie®** fertig machen von jenen, die es abbrechen oder von Personen, die kein Interesse daran haben?

Weiter, da hier die Validierungsmethode fehlgeschlagen ist, bietet es sich an, die Skalen des WSP mit anderen Persönlichkeitsskalen zu validieren. Das *Leistungsmotivationsinventar (LMI)* (Schuler, Prochaska & Frintrup, 2001), im Besonderen die Skalen „Selbständigkeit“, „Selbstkontrolle“ und „Zielsetzung“, das *Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung (BIP)* (Hossiep, Paschen & Mühlhaus, 2003), mit den Skalen „Leistungsmotivation“, „Belastbarkeit“ und „Selbstbewusstsein“, oder das Verfahren *Arbeitsbezogenes Verhaltens- und Erlebensmuster (AVEM)* (Schaarschmidt & Fischer, 2008) mit den Skalen „Beruflicher Ehrgeiz“ und „Resignationstendenz bei Misserfolg“ erscheinen sinnvoll für eine Konstruktvalidierung.

VI Zusammenfassung

Diese Arbeit hatte es zum Ziel, die Konstruktvalidität der Kurzversion des *Wiener Studieneignungs-Persönlichkeitsinventars* zu überprüfen und damit auch die Frage zu beantworten, ob sich Objektive Persönlichkeitstests zur Konstruktvalidierung (im Speziellen die aus dem **Wiener Self-Assessment Psychologie**®) von Persönlichkeitsfragebogen eignen.

Diese Fragestellungen wurden mittels korrelativen Auswertungen und Gruppenvergleichen nach zuvor aufgestellten Hypothesen und einer korrelativen Auswertung einer Multitrait-Multimethod-Matrix und schließlich auch einer explorativen Faktorenanalyse bearbeitet. Bis auf eine signifikante Korrelation zwischen der Skala WSP-Frustrationstoleranz und dem Leistungsniveau im Kodiertest waren alle Ergebnisse nicht signifikant. Es kann also von keinem Zusammenhang zwischen den Kennwerten der verwendeten Objektiven Persönlichkeitstests (LAMBDA-neu, Kodiertest aus den *Arbeitshaltungen* und WITE) und der vorliegenden Kurzversion des Persönlichkeitsfragebogens WSP ausgegangen werden. Ebenso wurden keine konvergenten oder diskriminanten Validitäten in der MTMM-Matrix gefunden. Die explorative Faktorenanalyse ergab, dass die Skalen des Persönlichkeitsfragebogens auf anderen Faktoren laden als die Kennwerte der Objektiven Persönlichkeitstests.

Auf den ersten Blick könnte man vermuten, dass die Ergebnisse gegen eine Konstruktvalidität der Kurzversion des *Wiener Studieneignungs-Persönlichkeitsinventars* sprechen. Andererseits sind, wie in Kapitel V Diskussion der Ergebnisse und Kritik angeführt wird, die Annahmen über den Zusammenhang zwischen Persönlichkeitsfragebogen und Objektiven Persönlichkeitstests zu hinterfragen. Da man von einer inhaltlichen Validität des WSP ausgehen kann, kann eine Konstruktvalidität nicht eindeutig ausgeschlossen werden, obwohl die Konstruktvalidierung in der vorliegenden Studie nicht geglückt ist. Vielmehr muss gesagt werden, dass sich Objektive Persönlichkeitstests aus den in Kapitel V diskutierten Gründen nicht für eine Konstruktvalidierung eines Persönlichkeitsfragebogens eignen.

VII Literaturverzeichnis

- Anastasi, A. (1986). Evolving concepts of test validation. *Annual review of Psychology*, 37, 1-15.
- Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler*. Heidelberg, Springer.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4. überarb. Aufl.). Heidelberg: Springer Medizin Verlag.
- Campbell, D. T. & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 81, 81-105.
- Cronbach, L. J. & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281-302.
- Dieterich, R. (1973). *Psychodiagnostik. Grundlagen und Probleme*. (1. Aufl.). München/Basel: Ernst Reinhardt Verlag (UTB).
- Dörrer, N. (2009). *Über den Einfluss der inhaltshomogenen Itemblockbildung auf die Dimensionalität eines Persönlichkeitsfragebogens*. Unveröffentlichte Diplomarbeit: Universität Wien.
- Edelmann, W. (2000). *Lernpsychologie* (6. Aufl.). Weinheim: BeltzPVU.
- Fischer, G. H. (1974). *Einführung in die Theorie psychologischen Testens: Grundlagen und Anwendungen*. Bern: Huber.
- Fisseni, H.-J. (2004). *Lehrbuch der psychologischen Diagnostik*. (3. Aufl.). Göttingen: Hogrefe.
- Flanagan, J. C. (1954). The critical incident technique. *Psychological-Bulletin*, 51, 327-358.
- Frebort, M. (unpubliziert). *Wissenstest Psychologie*. Alle Rechte bei: Test- und Beratungsstelle, Institut für Entwicklungspsychologie und psychologische Diagnostik, Universität Wien.
- Frebort, M., Gleeson, R. & Weitensfelder, L. (2012). Wissenstests zur Erfassung bereichsspezifischen Vorwissens. In K. D. Kubinger, M. Frebort, L. Khorramdel & L. Weitensfelder („Wiener Autorenkollektiv Studienberatungstests“) (Hrsg.), *Self-*

Assessment: Theorie und Konzepte (S. 83-92). Lengerich: Pabst.

Frebort, M. & Kubinger, K. D. (2006). Ermittlung eines Anforderungsprofils für Studierende der Psychologie. In B. Gula, R. Alexandrovicz, 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. (2007). Aktueller Stand zum Self-Assessment Psychologie der Test- und Beratungsstelle des Arbeitsbereichs Psychologie. Forschungsbericht der Fakultät für Psychologie der Universität Wien, Institut für Entwicklungspsychologie und Psychologische Diagnostik.

Fritsch, E. (2008). *Das Wiener Self-Assessment Psychologie: Erste Validierungsergebnisse, Akzeptanz und Typisierung der TestteilnehmerInnen*. Unveröff. Dipl. Arbeit, Universität Wien.

Groffmann, K.-J. & Michel, L. (Hrsg.), (1982). *Enzyklopädie der Psychologie, Themenbereich B: Methodologie und Methoden, Serie II: Psychologische Diagnostik, Band 1: Grundlagen psychologischer Diagnostik*. Göttingen: Hogrefe.

Guion, R. M. (1977). Content validity – The source of my discontent. *Applied Psychological Measurement*, 1, 1-10.

Gula, B., Alexandrovicz, R., Strauß, S., Brunner, E., Jenull-Schiefer, B. & Vitouch, O. (Hrsg.), (2006). *Perspektiven Psychologischer Forschung in Österreich. Proceedings zur 7. Wissenschaftlichen Tagung der Österreichischen Gesellschaft für Psychologie*. Lengerich: Pabst Science Publishers.

Haiden, A. (2009). *Testtheoretische Analyse des psychologisch-diagnostischen Verfahrens LAMBDA-neu*. Unveröff. Dipl. Arbeit, Universität Wien.

Hartig, J., Frey, A. & Jude, N. (2007). Validität. In H. Moosbrugger & A. Kevala (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 135-164). Heidelberg: Springer Medizin Verlag.

- Hauser, M. (2011). *Extremgruppenvalidierung von LAMBDA-neu*. Unveröff. Dipl. Arbeit, Universität Wien.
- Hossiep, R., Paschen, M. & Mühlhaus, O. (2003). *Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung (BIP)* (2. Aufl.). Göttingen: Hogrefe.
- Hussy, W., Schreier, M. & Echterhoff, G. (2010). *Forschungsmethoden in Psychologie und Sozialwissenschaften*. Berlin: Springer.
- Jäger, R.S. & Petermann, F. (Hrsg.). (1999). *Psychologische Diagnostik. Ein Lehrbuch* (4. Aufl.). Weinheim: Beltz PVU.
- Kane, M. T. (2001). Current Concerns in Validity Theory. *Journal of Educational Measurement*, 38, 319-342.
- Karolyi, M. (2009). *Entwicklung und Erprobung des multifunktionalen psychologisch-diagnostischen Verfahrens LAMBDA-neu*. Unveröff. Dipl. Arbeit: Universität Wien.
- Kersting, M. (2003). Augenscheinvalidität (face validity). In K. D. Kubinger & R. S. Jäger (Hrsg.), *Schlüsselbegriffe der psychologischen Diagnostik* (S. 54-55). Weinheim: Beltz PVU.
- Kersting, M. (2005). Qualitätssichernde und qualitätsoptimierende Maßnahmen. In K. Westhoff, L. J. Hellfritsch, L. F. Hornke, K. D. Kubinger, F. Lang, H. Moosbrugger, A. Püschel & G. Reimann (Hrsg.), *Grundwissen für die berufsbezogene Eignungsbeurteilung nach DIN 33430* (S. 243-254). Lengerich: Pabst Science Publishers.
- Khorramdel, L. & Maurer, M. (unpubliziert). *WSP – Wiener Studieneignungs-Persönlichkeitsinventar*. Alle Rechte bei: Test- und Beratungsstelle, Institut für Entwicklungspsychologie und Psychologische Diagnostik, Universität Wien.
- Khorramdel, L. & Maurer, M. (2012). In K. D. Kubinger, M. Frebort, L. Khorramdel & L. Weitensfelder „Wiener Autorenkollektiv Studienberatungstests“ (Hrsg.), *Self-Assessment: Theorie und Konzepte*. (S. 103-112). Lengerich: Pabst.

- Kubinger, K. D. (2003). Gütekriterien. In K. D. Kubinger & R. S. Jäger (Hrsg.), *Schlüsselbegriffe der psychologischen Diagnostik* (S. 195-204). Weinheim: Beltz PVU.
- Kubinger, K. D. (2006). Ein Update der Definition von Objektiven Persönlichkeitstests: Experimentalpsychologische Verhaltensdiagnostik. In T. M. Ortner, R. T. Proyer & K. D. Kubinger (Hrsg.), *Theorie und Praxis Objektiver Persönlichkeitstests* (S. 38-52). Bern: Huber.
- Kubinger, K. D. (2009). *Psychologische Diagnostik: Theorie und Praxis psychologischen Diagnostizierens*. (2. Aufl.). Göttingen: Hogrefe.
- Kubinger, K. D. (2010). Die Zukunft des Verfahrensinventars psychologischen Diagnostizierens. In K. D. Kubinger & T. M. Ortner (Hrsg.), *Psychologische Diagnostik in Fallbeispielen* (S. 30-42). Göttingen: Hogrefe.
- Kubinger, K. D. & Ebenhö, J. (1996). *Arbeitshaltungen – Kurze Testbatterie Anspruchsniveau, Frustrationstoleranz, Leistungsmotivation, Impulsivität/Reflexivität*. Frankfurt/Main: Swets.
- Kubinger, K. D. & Ebenhö, J. (2012). Experimentalpsychologische Verhaltensdiagnostik des Anspruchsniveaus: Der Untertest Symbole Kodieren aus den „Arbeitshaltungen“. In K. D. Kubinger, M. Frebort, L. Khorramdel & L. Weitensfelder („Wiener Autorenkollektiv Studienberatungstests“) (Hrsg.), *Self-Assessment: Theorie und Konzepte* (S. 93-102). Lengerich: Pabst.
- Kubinger, K. D., Frebort, M., Khorramdel, L. & Weitensfelder, L. („Wiener Autorenkollektiv Studienberatungstests“) (Hrsg.) (2012). *Self-Assessment: Theorie und Konzepte*. Lengerich: Pabst.
- Kubinger, K. D., Haiden, A., Karolyi, M. (2012). Diagnostik des Lernstils: Der Lerntest LAMBDA. In K. D. Kubinger, M. Frebort, L. Khorramdel & L. Weitensfelder („Wiener Autorenkollektiv Studienberatungstests“) (Hrsg.), *Self-Assessment: Theorie und Konzepte* (S. 71-82). Lengerich: Pabst.
- Kubinger, K. D. & Jäger R. S. (Hrsg.), (2003). *Schlüsselbegriffe der psychologischen Diagnostik*. Weinheim: Beltz PVU.

- Kubinger, K. D. & Litzenberger, M. (2003). Zur Validität der Objektiven Persönlichkeits-Test-Batterie „Arbeitshaltungen“. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 24, 119-133.
- 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.
- Kubinger, K. D. & Ortner, T. M. (Hrsg.), (2010). *Psychologische Diagnostik in Fallbeispielen*. Göttingen: Hogrefe.
- Lienert, G. (1989). *Testaufbau und Testanalyse* (4. Aufl.). Weinheim: Beltz PVU.
- Lienert, G. & Ratz, U. (1994). *Testaufbau und Testanalyse*. Weinheim: Beltz PVU.
- Maurer, M. (2010). *Wiener Studieneignungs-Persönlichkeitsinventar (WSP) – Formulierung eines Regelkatalogs zur Itemkonstruktion und dessen Anwendung zur Erstellung eines Itempools*. Unveröff. Dipl. Arbeit, Universität Wien.
- Messick, S., (1995). Validity of Psychological Assessment. *American Psychologist*, 50, 741-749.
- Michel, L. & Conrad, W. (1982). Theoretische Grundlagen psychometrischer Tests. In K.-J. Groffmann & L. Michel (Hrsg.), *Enzyklopädie der Psychologie, Themenbereich B: Methodologie und Methoden, Serie II: Psychologische Diagnostik, Band 1: Grundlagen psychologischer Diagnostik* (S. 11-29). Göttingen: Hogrefe.
- Moosbrugger, H. & Kevala, A. (Hrsg.), (2008). *Testtheorie und Fragebogenkonstruktion*. Heidelberg: Springer.
- Moosbrugger, H. & Schermelleh-Engel, K. (2007). Exploratorische (EFA) und Konfirmatorische Faktorenanalyse (CFA). In H. Moosbrugger & A. Kevala (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 307-324). Heidelberg: Springer Medizin Verlag.

- Murphy, K. R. & Davidshofer, C. O. (1991). *Psychological Testing. Principles & Applications*. (2. Aufl.). New Jersey: Prentice Hall.
- Ortner, T. M., Horn, R., Kersting, M., Krumm, S., Kubinger, K. D., Proyer, R. T., Schmidt-Atzert, L., Schuhfried, G., Schütz, A., Wagner-Menghin, M. M. & Westhoff, K. (2007). Standortbestimmung und Zukunft Objektiver Persönlichkeitstests. *Report Psychologie*, 32, 60-70.
- Ortner, T. M., Proyer, R. T. & Kubinger, K. D. (Hrsg.), (2006). *Theorie und Praxis Objektiver Persönlichkeitstests*. Bern: Huber.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Kopenhagen: Danish Institute for Educational Research.
- Reimann, G. (2005). Rechtliche Rahmenbedingungen. In K. Westhoff, L. J. Hellfritsch, L. F. Hornke, K. D. Kubinger, F. Lang, H. Moosbrugger, A. Püschel & G. Reimann (Hrsg.), *Grundwissen für die berufsbezogene Eignungsbeurteilung nach DIN 33430* (S. 17-38). Lengerich: Pabst Science Publishers.
- Rost, R. D. (Hrsg.), (1998). *Handwörterbuch Pädagogische Psychologie*. Weinheim: Beltz PVU.
- Rudinger, G. & Hörsch, K. (Hrsg.), (2009). *Self-Assessment an Hochschulen: Von der Studienfachwahl zur Profilbildung (Applied Research in Psychology and Evaluation)* (1. Aufl.). Göttingen: V&R Unipress.
- Rudinger, G. & Hörsch, K. (2009). Vorwort. In G. Rudinger & K. Hörsch (Hrsg.), *Self-Assessment an Hochschulen: Von der Studienfachwahl zur Profilbildung (Applied Research in Psychology and Evaluation)* (1. Aufl.) (S. 7-8). Göttingen: V&R Unipress.
- Rudolph, U. (2003). *Motivationspsychologie*. Weinheim: BeltzPVU.
- Schaarschmidt, U. & Fischer, A. W. (2008). *Arbeitsbezogenes Verhaltens- und Erlebensmuster (AVEM)* (3. Aufl.). Frankfurt/M.: Pearson Assessment.

- Schermelleh-Engel, K. & Schweizer, K. (2007). Multitrait-Multimethod-Analysen. In H. Moosbrugger & A. Kevala (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 325-341). Heidelberg: Springer Medizin Verlag.
- Schmotzer, C., Kubinger, K. D. & Maryschka, C. (1994). *Rechnen in Symbolen* [Software und Manual]. Frankfurt: Swets.
- Schuler, H., Prochaska, M. & Frintrup, A. (2001). *Leistungsmotivationsinventar (LMI)*. Göttingen: Hogrefe.
- Skinner, N. S. F., & Horwath, E. (1973). Cross-media independence of questionnaire and objective test personality factors. *Multivariate Behavioral Research*, 8, 23-40.
- Sonnleitner, P., Kubinger, K. D., & Frebort, M. (2009). Das Wiener Self-Assessment Psychologie mit seinen Verfahren der experimentalpsychologischen Verhaltensdiagnostik zur Messung studienfachübergreifender soft skills. In G. Rudinger & K. Hörsch (Hrsg.), *Self-assessment an Hochschulen: Von der Studienfachwahl zur Profilbildung* (S. 63-72). Göttingen: V&R unipress.
- Undeutsch, N. C. (2010). *Zwei neue Subtests für das „Self-Assesment“ Psychologie durch regelgeleitete Konstruktion*. Unveröff. Dipl. Arbeit, Universität Wien.
- Useini, C. (in Vorb.). *Modifikation des Fragebogens zum Interesse am Psychologiestudium (FIP) und Erweiterung um die Dimension Selbstkonzept durch die Entwicklung des Fragebogens „Selbstkonzept im Psychologiestudium“ (SKIP) für den Einsatz im Wiener Self-Assessment Psychologie*. Unveröffentlichte Diplomarbeit: Universität Wien.
- Wagner, M. M. (1999). *Lexikon-Wissen-Test (LEWITE)*. Software und Manual. Mödling: Dr. G. Schuhfried GmbH.
- Wagner-Menghin, M. M. (2006). Multifunktionalität am Beispiel des Lexikon-Wissens-Test (LEWITE). In T. M. Ortner, R. T. Proyer & K. D. Kubinger (Hrsg.), *Theorie und Praxis Objektiver Persönlichkeitstests* (S. 204-209). Bern: Hans Huber.
- Wawrinec, A. (2006). *Selbstvertrauen in die eigene Leistung: Validierung des Überschätzungsparameters des LEWITE anhand von Extremgruppen Depressive vs. Seelisch Gesunde*. Unveröffentlichte Diplomarbeit: Universität Wien.

- Westhoff, K., Hellfritsch, L. J., Hornke, L. F., Kubinger, K. D., Lang, F., Moosbrugger, H., Püschel, A. & Reimann, G. (Hrsg.), (2005). *Grundwissen für die berufsbezogene Eignungsbeurteilung nach DIN 33430*. Lengerich: Pabst Science Publishers.
- Wolf, B. (1998). Lernumwelt/Ökologische Psychologie. In R. D. Rost (Hrsg.). *Handwörterbuch Pädagogische Psychologie* (S. 318-321). Weinheim: Beltz PVU.
- Zimmerhofer, A., Heukamp, V. M. & Hornke, L. F. (2006). Ein Schritt zur fundierten Studienfachwahl – webbasierte Self-Assessments in der Praxis. *Report Psychologie*, 31, 62-72.

VIII Anhang

1. Der WSP im Rahmen des **Wiener Self-Assessment Psychologie®**

An dieser Stelle sollen die Skalen mit den zugehörigen Items illustriert werden, wie sie im **Wiener Self-Assessment Psychologie®** zur Anwendung kommen (Khorramdel & Maurer, 2010):

Organisation (18 Items)

Zeitmanagement	Ich teile mir meine Arbeitszeit so ein, dass ich Abgabetermine erst gar nicht verschieben muss.
	Wenn eine Arbeit mehrere Wochen bis zur Fertigstellung benötigt, erstelle ich einen Zeitplan.
	Wenn ich mich auf mehrere Prüfungen gleichzeitig vorbereiten muss, fällt mir die Lernzeiteinteilung pro Prüfung schwer.
	Bei langfristigen Projekten unterschätze ich den zeitlichen Arbeitsaufwand.
Prioritäten- setzung	Um im Studium gut voranzukommen, verzichte ich notfalls auch auf Freizeitaktivitäten, auf die ich mich schon lange gefreut habe.
	Nachdem ich mir einen Überblick über anstehende Aufgaben verschafft habe, fällt es mir leicht, die Aufgaben ihrer Wichtigkeit nach zu reihen.
	Wichtige Aufgaben kommen wiederholt zu kurz, weil ich nicht mit den wichtigsten, sondern mit den angenehmsten Aufgaben beginne.
Planen	Ich recherchiere mögliche Überschneidungen von Seminaren bereits mehrere Tage vor der Anmeldung.
	Ich wäre gerne besser darin, mein Semester entlang des Studienplans zu planen.
	Um in meiner Arbeit zügig voran zu kommen, nehme ich mir für jeden Tag bestimmte Tagesziele vor.
	Ich plane mein Semester mit der Kenntnis davon, welche Lehrveranstaltungen Voraussetzung für darauf folgende Seminare sind.
Organisieren	Bei neuen Aufgaben gewinne ich rasch einen Überblick darüber, welche Unterlagen ich benötige.
	Meine Freunde würden mich als sehr strukturiert in der Bearbeitung von Aufgaben bezeichnen.
	Wenn ich eine Vorlesung nicht besuchen kann, organisiere ich mich vorab mit einem Studienkollegen, um die Mitschriften an Semesterende auszutauschen.
Informiertheit	Ich bin über das Serviceangebot der Universität (z.B. StudentPoint, ÖH, Studienrichtungsvertretung) informiert.

	Andere sind über Terminänderungen meist früher informiert als ich.
	Ich informiere eher die anderen über organisatorische Fragen zum Studium als umgekehrt.
	Ich bin über das Studienangebot der Universität gut informiert
Motivation (22 Items)	
Durchhalte- vermögen	Wenn erforderlich, lerne ich für Prüfungen von morgens bis abends.
	Wenn eine Aufgabe uninteressant wird, lasse ich sie sehr schnell links liegen.
	Wenn etwas schwieriger wird als zunächst gedacht, suche ich mir nach Möglichkeit eine leichtere Aufgabe.
	Eine begonnene Aufgabe bringe ich grundsätzlich zu Ende, auch wenn ich mehr Zeit und Mühe investieren muss als geplant.
Lernbereitschaft	Es fällt mir schwer, Verbesserungsvorschläge anderer anzunehmen.
	Wenn eine Kritik an meiner Arbeit geäußert wird, frage ich interessiert nach.
	Auch bei einer guten Prüfungsnote, möchte ich meine Fehler wissen.
	Ich ziehe die Bearbeitung vertrauter Aufgaben vor, als mich auf neue Themengebiete einzulassen.
Engagement	Für Prüfungen lerne ich gerade genug, um sie bestehen zu können.
	Ich versuche im Studium den Aufwand zu minimieren, indem ich für Lehrveranstaltungen nur die Mindestanforderungen erbringe.
	Ich nehme eine längere Studienzeit in Kauf, wenn ich dafür mein Wissen in dem Fach noch mehr vertiefen kann.
	Ich lerne nicht nur für Prüfungen, sondern lese auch in der Freizeit Zusatzliteratur.
Zielorientierung	Ich setze mir zwar Tagesziele, verliere diese aber sehr schnell aus den Augen.
	Ich wäre gerne besser darin, meine Ziele hartnäckig zu verfolgen.
	Wenn ich mir ein Lernziel setze (z.B. eine bestimmte Stoffmenge pro Tag zu lernen), dann erreiche ich es.
Zielsetzung	Ich habe eine konkrete Vorstellung davon, was ich erreichen möchte.
	Ich verfolge ein für mich klar definiertes Ausbildungsziel.
	Es fehlt mir ein Ziel, auf das ich hinarbeiten möchte.
Erfolgszuversicht vs. Misserfolgs- ängstlichkeit	Anfängliche Schwierigkeiten lassen mich daran zweifeln, ein Projekt erfolgreich meistern zu können.
	Ich bin zuversichtlich, mein Studium erfolgreich abschließen zu können.
	Wenn ich ein Problem nicht gleich lösen kann, werde ich dies auch später nicht mehr schaffen.
	Neue Aufgaben sehe ich als eine Herausforderung, die ich gerne annehme.
Frustrationstoleranz (7 Items)	

Frustrations- toleranz	Bevor ich eine mühsam erstellte schriftliche Arbeit umschreibe, nehme ich die schlechte Note in Kauf.
	Wenn ich die Kritik an meiner Arbeit nicht gerechtfertigt finde, melde ich mich auch mal von einem Seminar ab.
	Kritik durch Lehrpersonen spornen mich grundsätzlich dazu an, mehr Leistung zu erbringen.
	Nach einer negativen Prüfungsnote trete ich zum nächstmöglichen Prüfungstermin erneut an.
	Negatives Feedback bremst meinen Arbeitseifer.
	Wenn ich von einer Lehrperson negativ beurteilt werde, wähle ich im darauf folgenden Semester ein Seminar bei jemand anderem.
	Wenn die Studiensituation ein schnelles Weiterkommen im Studium verhindert, breche ich das Studium möglicherweise frühzeitig ab.

Selbstbewusstsein (6 Items)

Selbstbewusst- sein	Vertrete ich in einer Gruppe eine andere Meinung als die meisten anderen, fällt es mir schwer diese zu äußern.
	Es ist mir unangenehm nachzufragen, wenn ich etwas nicht sofort verstehe.
	Bei mündlichen Prüfungen trete ich selbstsicher auf.
	Der Gedanke etwas präsentieren zu müssen, verunsichert mich nicht.
	Richtet ein Lehrveranstaltungsleiter vor versammelter Gruppe eine Frage an mich, werde ich nervös.
	Wenn dem Lehrveranstaltungsleiter meiner Ansicht nach ein Fehler unterläuft, mache ich ihn darauf aufmerksam.

2. Beschreibung der Datenlage

In Tabelle 44 sieht man, dass 217 Personen, den LAMBDA-neu abgeschlossen haben.

RISItem1602RW

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
0	105	29,1	48,4	48,4
Gültig 1	112	31,0	51,6	100,0
Gesamt	217	60,1	100,0	
Fehlend System	144	39,9		
Gesamt	361	100,0		

Tabelle 44: Anzahl der Personen, die den LAMBDA-neu abgeschlossen haben

In Tabelle 45 sieht man, dass 211 Personen den Kodiertest abgeschlossen haben.

AHAvollstaendig

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
0	4	1,1	1,9	1,9
Gültig 1	211	58,4	98,1	100,0
Gesamt	215	59,6	100,0	
Fehlend System	146	40,4		
Gesamt	361	100,0		

Tabelle 45: Anzahl der Personen, die den Kodiertest abgeschlossen haben

In Tabelle 46 sieht man, dass 206 Personen den WITE abgeschlossen haben.

LEWITEvollstaendig

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
0	5	1,4	2,4	2,4
Gültig 1	206	57,1	97,6	100,0
Gesamt	211	58,4	100,0	
Fehlend System	150	41,6		
Gesamt	361	100,0		

Tabelle 46: Anzahl der Personen, die den WITE abgeschlossen haben

In Tabelle 47 sieht man, dass 197 Personen den FIP&SKIP abgeschlossen haben.

FIPundSKIPvollstaendig

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
0	8	2,2	3,9	3,9
Gültig 1	197	54,6	96,1	100,0
Gesamt	205	56,8	100,0	
Fehlend System	156	43,2		
Gesamt	361	100,0		

Tabelle 47: Anzahl der Personen, die den FIP&SKIP abgeschlossen haben

In Tabelle 48 sieht man, dass 181 Personen den Matrizentest abgeschlossen haben.

Matrizentestvollstaendig

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
0	9	2,5	4,7	4,7
Gültig 1	181	50,1	95,3	100,0
Gesamt	190	52,6	100,0	
Fehlend System	171	47,4		
Gesamt	361	100,0		

Tabelle 48: Anzahl der Personen, die den Matrizentest abgeschlossen haben

In Tabelle 49 sieht man, dass 180 Personen den WSP vollständig abgeschlossen haben.

WSPvollstaendig

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
0	1	,3	,6	,6
Gültig 1	180	49,9	99,4	100,0
Gesamt	181	50,1	100,0	
Fehlend System	180	49,9		
Gesamt	361	100,0		

Tabelle 49: Anzahl der Personen, die den WSP vollständig abgeschlossen haben

3. Darstellung der Ergebnisse der Hypothesen

3.1 Prüfung der Voraussetzungen zu *Organisation*

$H_0 OF_1$

Die Prüfung der Voraussetzung der Normalverteilung in den beiden Gruppen wurde mit dem K-S-Anpassungstest berechnet. In beiden Gruppen (Gruppe 0: schlechte Lernstrategie; Gruppe 1: gute Lernstrategie) kann aufgrund der nicht signifikanten Ergebnisse (siehe Tab. 49 und 50; $p=0,670$ und $p=0,971$) von einer Normalverteilung der Variablen ausgegangen werden.

Kolmogorov-Smirnov-Anpassungstest

		WSPOrganisa tionRWGesamt
N		81
Parameter der Normalverteilung ^{a,b}	Mittelwert	50,99
	Standardabweichung	7,969
Extremste Differenzen	Absolut	,081
	Positiv	,055
	Negativ	-,081
Kolmogorov-Smirnov-Z		,725
Asymptotische Signifikanz (2-seitig)		,670

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

Tabelle 49: K-S-Anpassungstest zu H_0OF_1 - Lernstrategie Gruppe 0

Kolmogorov-Smirnov-Anpassungstest

		WSPOrganisa tionRWGesamt
N		99
Parameter der Normalverteilung ^{a,b}	Mittelwert	52,61
	Standardabweichung	7,863
Extremste Differenzen	Absolut	,049
	Positiv	,035
	Negativ	-,049
Kolmogorov-Smirnov-Z		,488
Asymptotische Signifikanz (2-seitig)		,971

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

Tabelle 50: K-S-Anpassungstest zu H_0OF_1 - Lernstrategie Gruppe 1

H_0OF_3

Die Prüfung der Voraussetzung der Normalverteilung in den beiden Gruppen wurde mit dem K-S-Anpassungstest berechnet. In beiden Gruppen (Gruppe 0: Kriterium nicht erreicht; Gruppe 1: Kriterium erreicht) kann aufgrund der nicht signifikanten Ergebnisse (siehe Tab. 51 und 52; $p=0,730$ und $p=0,951$) von einer Normalverteilung der Variablen ausgegangen werden.

Kolmogorov-Smirnov-Anpassungstest

		WSPOrganisa tionRWGesam mt
N		63
Parameter der Normalverteilung ^{a,b}	Mittelwert	51,08
	Standardabweichung	7,946
Extremste Differenzen	Absolut	,087
	Positiv	,060
	Negativ	-,087
Kolmogorov-Smirnov-Z		,688
Asymptotische Signifikanz (2-seitig)		,730

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

Tabelle 51: K-S-Anpassungstest zu H_0OF_3 - Kriterium erreicht Gruppe 0

Kolmogorov-Smirnov-Anpassungstest

		WSPOrganisa tionRWGesam mt
N		117
Parameter der Normalverteilung ^{a,b}	Mittelwert	52,31
	Standardabweichung	7,923
Extremste Differenzen	Absolut	,048
	Positiv	,036
	Negativ	-,048
Kolmogorov-Smirnov-Z		,519
Asymptotische Signifikanz (2-seitig)		,951

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

Tabelle 52: K-S-Anpassungstest zu H_0OF_3 - Kriterium erreicht Gruppe 1

3.2 Prüfung der Voraussetzungen zu *Motivation*

H_0M_1

Die Prüfung der Voraussetzung der Normalverteilung in den beiden Gruppen wurde mit dem K-S-Anpassungstest berechnet. In beiden Gruppen (Gruppe 0: schlechte Lernstrategie; Gruppe 1: gute Lernstrategie) kann aufgrund der nicht signifikanten Ergebnisse (siehe Tab. 53 und 54; $p=0,784$ und $p=0,426$) von einer Normalverteilung der Variablen ausgegangen werden.

Kolmogorov-Smirnov-Anpassungstest

		WSPMotivati onRWGesamt
N		81
Parameter der Normalverteilung ^{a,b}	Mittelwert	64,14
	Standardabweichung	8,583
Extremste Differenzen	Absolut	,073
	Positiv	,042
	Negativ	-,073
Kolmogorov-Smirnov-Z		,655
Asymptotische Signifikanz (2-seitig)		,784

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

Tabelle 53: K-S-Anpassungstest zu H_0M_1 - Lernstrategie Gruppe 0

Kolmogorov-Smirnov-Anpassungstest

		WSPMotivati onRWGesamt
N		99
Parameter der Normalverteilung ^{a,b}	Mittelwert	64,17
	Standardabweichung	8,706
Extremste Differenzen	Absolut	,088
	Positiv	,046
	Negativ	-,088
Kolmogorov-Smirnov-Z		,876
Asymptotische Signifikanz (2-seitig)		,426

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

Tabelle 54: K-S-Anpassungstest zu H_0M_1 - Lernstrategie Gruppe 1

H_0M_3

Die Prüfung der Voraussetzung der Normalverteilung in den beiden Gruppen wurde mit dem K-S-Anpassungstest berechnet. In beiden Gruppen (Gruppe 0: Kriterium nicht erreicht; Gruppe 1: Kriterium erreicht) kann aufgrund der nicht signifikanten Ergebnisse (siehe Tab. 55 und 56; $p=0,859$ und $p=0,383$) von einer Normalverteilung der Variablen ausgegangen werden.

Kolmogorov-Smirnov-Anpassungstest

		WSPMotivati onRWGesamt
N		63
Parameter der	Mittelwert	64,90
Normalverteilung ^{a,b}	Standardabweichung	8,224
Extremste Differenzen	Absolut	,076
	Positiv	,066
	Negativ	-,076
Kolmogorov-Smirnov-Z		,604
Asymptotische Signifikanz (2-seitig)		,859

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

Tabelle 55: K-S-Anpassungstest zu H_0M_3 - Kriterium erreicht Gruppe 0

Kolmogorov-Smirnov-Anpassungstest

		WSPMotivati onRWGesamt
N		117
Parameter der	Mittelwert	63,75
Normalverteilung ^{a,b}	Standardabweichung	8,844
Extremste Differenzen	Absolut	,084
	Positiv	,049
	Negativ	-,084
Kolmogorov-Smirnov-Z		,907
Asymptotische Signifikanz (2-seitig)		,383

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

Tabelle 56: K-S-Anpassungstest zu H_0M_3 - Kriterium erreicht Gruppe 1

3.3 Prüfung der Voraussetzungen zu *Frustrationstoleranz*

H_0FT_1

Levene-Test auf Gleichheit der Fehlervarianzen^a

Abhängige

Variable: WSPFrustrationstoleranzRWGesamt

F	df1	df2	Sig.
5,499	2	175	,005

Prüft die Nullhypothese, daß die Fehlervarianz der abhängigen Variablen über Gruppen hinweg gleich ist.

a. Design: Konstanter Term +

LAMBDA Belastbarkeit123

Tabelle 57: Levene-Test zu H_0FT_1

IX Tabellenverzeichnis

Tabelle 1: Anforderungsprofil von Frebort & Kubinger (2006, S. 411).....	20
Tabelle 2: Skalen und Subskalen des WSP (Maurer, 2010)	22
Tabelle 3: Skalen und Subskalen des WSP im Wiener Self-Assessment Psychologie ® verwendet	28
Tabelle 4: Entwurf der Multitrait-Multimethod-Matrix.....	56
Tabelle 5: Legende der Multitrait-Multimethod-Matrix	56
Tabelle 6: Vollständig abgeschlossene Self-Assessment Tools.....	59
Tabelle 7: Statistiken zur Altersverteilung der TeilnehmerInnen	60
Tabelle 8: Studienort der TeilnehmerInnen	62
Tabelle 9: Abschneiden der TeilnehmerInnen im WSP.....	64
Tabelle 10: Anzahl der Stärke/Schwäche-Rückmeldungen für WSP-Organisation	65
Tabelle 11: Anzahl der Stärke/Schwäche-Rückmeldungen für WSP-Motivation	65
Tabelle 12: Anzahl der Stärke/Schwäche-Rückmeldungen für WSP-Frustrationstoleranz.....	65
Tabelle 13: Anzahl der Stärke/Schwäche-Rückmeldungen für WSP-Selbstbewusstsein.....	65
Tabelle 14: Gruppenstatistiken zu H_0OF_1	66
Tabelle 15: T-Test bei unabhängigen Stichproben zu H_0OF_1	67
Tabelle 16: Spearman-Rho Korrelation zu H_0OF_2	68
Tabelle 17: Anzahl der Personen, die im LAMBDA-neu das Kriterium erreicht haben	68
Tabelle 18: Gruppenstatistiken für T-Test für unabhängige Stichproben für H_0OF_3	69
Tabelle 19: T-Test bei unabhängigen Stichproben für H_0OF_3	69
Tabelle 20: Gruppenstatistiken zu H_0M_1	70
Tabelle 21: T-Test für unabhängige Stichproben zu H_0M_1	71
Tabelle 22: Spearman-Rho Korrelation zu H_0M_2	72
Tabelle 23: Deskriptive Statistiken zu H_0M_3	73
Tabelle 24: T-Test für unabhängige Stichproben zu H_0M_3	73
Tabelle 25: Deskriptive Statistiken zur ANOVA zu H_0M_4	74
Tabelle 26: Levene-Test zur ANOVA zu H_0M_4	75
Tabelle 27: Tests der Zwischensubjekteffekte zu ANOVA zu H_0M_4	75
Tabelle 28: Spearman-Rho Korrelation zu H_0M_5	76
Tabelle 29: Ränge zum Kruskal-Wallis-Test zu H_0M_6	77
Tabelle 30: Statistik für Kruskal-Wallis-Test zu H_0M_6	77

Tabelle 31: Deskriptive Statistik zu LAMBDABelastbarkeit zu H_0FT_1	78
Tabelle 32: Ränge zu Kruskall-Wallis-Test zu H_0FT_1	78
Tabelle 33: Statistik zu Kruskall-Wallis-Test zu H_0FT_1	79
Tabelle 34: Spearman-Rho Korrelation zu H_0FT_2	80
Tabelle 35: Ränge zu Kruskall-Wallis-Test zu H_0FT_3	81
Tabelle 36: Statistik zu Kruskall-Wallis-Test zu H_0FT_3	81
Tabelle 37: Spearman-Rho Korrelationen zu H_0S_1 , H_0S_2 und H_0S_3	83
Tabelle 38: Darstellung der berechneten MTMM – Matrix; Ergebnisse	85
Tabelle 39: Legende zu MTMM-Matrix	85
Tabelle 40: Angenommene Reihenfolge der durchschn. Korrelationen der MTMM-Matrix	86
Tabelle 41: Gefundene Reihenfolge der durchschn. Korrelationen der MTMM-Matrix	86
Tabelle 42: Erklärte Gesamtvarianz für die Faktorenanalyse	88
Tabelle 43: Rotierte Komponentenmatrix zur Faktorenanalyse	89
Tabelle 44: Anzahl der Personen, die den LAMBDA-neu abgeschlossen haben	111
Tabelle 45: Anzahl der Personen, die den Kodiertest abgeschlossen haben	112
Tabelle 46: Anzahl der Personen, die den WITE abgeschlossen haben	112
Tabelle 47: Anzahl der Personen, die den FIP&SKIP abgeschlossen haben	112
Tabelle 48: Anzahl der Personen, die den Matrizentest abgeschlossen haben	113
Tabelle 49: K-S-Anpassungstest zu H_0OF_1 - Lernstrategie Gruppe 0	114
Tabelle 50: K-S-Anpassungstest zu H_0OF_1 - Lernstrategie Gruppe 1	114
Tabelle 51: K-S-Anpassungstest zu H_0OF_3 - Kriterium erreicht Gruppe 0	115
Tabelle 52: K-S-Anpassungstest zu H_0OF_3 - Kriterium erreicht Gruppe 1	115
Tabelle 53: K-S-Anpassungstest zu H_0M_1 - Lernstrategie Gruppe 0	116
Tabelle 54: K-S-Anpassungstest zu H_0M_1 - Lernstrategie Gruppe 1	116
Tabelle 55: K-S-Anpassungstest zu H_0M_3 - Kriterium erreicht Gruppe 0	117
Tabelle 56: K-S-Anpassungstest zu H_0M_3 - Kriterium erreicht Gruppe 1	117
Tabelle 57: Levene-Test zu H_0FT_1	117

X Abbildungsverzeichnis

Abbildung 1: Screenshot LAMBDA-neu (Kubinger, Haiden, Karolyi & Maryschka, 2012, S. 73)	31
Abbildung 2: Screenshot Kodiertest (Kubinger & Ebenhöf, 2012, S. 95)	33
Abbildung 3: Screenshot WITE (Frebort, Gleeson & Weitensfelder, 2012, S. 84)	35
Abbildung 4: Kreisdiagramm: Geschlecht der abgeschlossenen TeilnehmerInnen	59
Abbildung 5: Balkendiagramm der Altersverteilung der abgeschlossenen TeilnehmerInnen	60
Abbildung 6: Kreisdiagramm der Herkunftsländer der abgeschlossenen TeilnehmerInnen	60
Abbildung 7: Kreisdiagramm der Muttersprachen der abgeschlossenen TeilnehmerInnen	61
Abbildung 8: Kreisdiagramm zu der höchsten abgeschlossenen Schulbildung der abgeschlossenen TeilnehmerInnen	61
Abbildung 9: Kreisdiagramm zur Frage ob bereits ein Psychologiestudium begonnen wurde	62
Abbildung 10: Balkendiagramm zum Studienort der TeilnehmerInnen wenn nicht Universität Wien	63
Abbildung 11: Box-Plot zu H_0OF_1	67
Abbildung 12: Streudiagramm zu H_0OF_2	68
Abbildung 13: Boxplot zu H_0OF_3	70
Abbildung 14: Box-Plot zu H_0M_1	71
Abbildung 15: Streudiagramm zu H_0M_2	72
Abbildung 16: Box-Plot zu H_0M_3	74
Abbildung 17: Streudiagramm zu H_0M_5	76
Abbildung 18: Box-Plot zu H_0M_8	77
Abbildung 19: Box-Plot zu H_0FT_1	79
Abbildung 20: Streudiagramm zu H_0FT_2	80
Abbildung 21: Box-Plot zu H_0FT_3	81
Abbildung 22: Screeplot zur Faktorenanalyse	88

XI Lebenslauf

Lebenslauf

Allgemeine Daten:

Name: **Katarina IFCIC**
Geburtsdatum: 18.08.1987
Geburtsort: Bratislava, Slowakei
Staatsbürgerschaft: Österreich
Wohnort: Linzerstrasse 160/7/16, 1140 Wien
Telefonnummer: 0043 650 4401 402
Email: Kati_ifcic@hotmail.com

Schulische und berufliche Ausbildung:

Oktober 2005 - laufend Studium der Psychologie an der Universität Wien
Juni 2005 Matura (ausgezeichneter Erfolg) am Goethe-Gymnasium (humanistisch-sprachlicher Zweig) in 1140 Wien

Erfahrungen im psychologischen Bereich:

04/2012 – laufend **Praktikum beim Verein LOK – Leben ohne Krankenhaus** in 1050 Wien
(<http://www.lok.at/>)
Aufgaben: Betreuung in vollbetreuter Wohngemeinschaft für Menschen mit psychischen Erkrankungen: Tagesstruktur, Freizeitplanung und –durchführung, Selbständigkeit, Selbstmanagement, Alltagsfertigkeiten

07/2011 – 09/2011 **Praktikum an der „AHG Klinik Berus/Europäisches Zentrum für Psychosomatik und Verhaltensmedizin“** in 66802 Überherrn-Berus/Deutschland
(<http://www.ahg.de/AHG/Standorte/Berus/>)
Aufgaben: Hospitieren bei Einzelpsychotherapien; Co-Leitung und Hospitieren bei Gruppentherapien; Leitung von Entspannungsgruppen (Progressive Muskelentspannung nach Jacobson, Tiefenentspannung und Entspannungstraining über fokussierte Aufmerksamkeit)

10/2010 – 11/2010 **Praktikum: „Sozialmedizinisches Zentrum Baumgartner Höhe Otto Wagner Spital und Pflegezentrum“** in 1140 Wien

(<http://www.wienkav.at/kav/ows/>)

Aufgaben: Co-Therapeutische Teilnahme an Gruppentherapien; Hospitieren und Durchführen kognitiver Trainings; Hospitieren bei Einzelpsychotherapie; Freizeitaktivitäten mit Patienten im Einzel- und Gruppensetting planen und durchführen

07/2009 – 09/2009 **Praktikum beim „Verein P.A.S.S. – Prävention, Angehörigenarbeit, Suchtbehandlung und Suchtbetreuung“** in 1080 Wien

(www.pass.at/start.htm)

Aufgaben: Persönlichkeits- und Leistungsdiagnostik; Gutachtenerstellung; Hospitieren bei Anamnesegesprächen, psychotherapeutischen Erstgesprächen und Gruppentherapien

09/2008 – 01/2009 **Praktikum an der „Austro Control Akademie“** in 1030 Wien

(<http://www.austrocontrol.at/content/atm/Akademie/aa.shtml>)

Aufgaben: Einblick in Arbeits- und Organisationspsychologie; Auswahl der Kandidaten mittels psychologisch-diagnostischer Verfahren; statistische Auswertung einer Studie („Zusammenhang: Erfolg im Auswahlverfahren und Erfolg in der Ausbildung“)

05/2008 – 06/2008 **Praktikum am „Institut für Kinderrechte und Elternbildung“** in 1010

Wien (www.kinderfreunde.at/V/Institut-fuer-Kinderrechte)

Aufgaben: Planung und Durchführung einer Studie („Desintegration“); Datenerhebung

Berufliche Tätigkeiten:

10/2009 - laufend	Mitarbeit bei Allgemein beeideten und gerichtlich zertifizierten Sachverständigen und Klinischen PsychologInnen
02/2009 – 06/2011	Kellnerin im „Replugged“, Cafe und Bar, in 1080 Wien
10/2008 – 01/2009	Promotiontätigkeiten (<i>DiemOberbauer; Easystaff</i>)
06/2005 – 09/2008	Telefoninterviewerin bei „IFES- Institut für empirische Sozialforschung“

Zusätzliche Qualifikationen:

Führerschein der Klasse B

Sehr gute Kenntnisse in der englischen Sprache in Wort und Schrift

Sehr gute Kenntnisse in der slowakischen Sprache in Wort und Schrift

Grundkenntnisse in der französischen Sprache in Wort und Schrift

Grundkenntnisse in der russischen Sprache in Wort und Schrift

