



universität
wien

Diplomarbeit

Wiener Studieneignungs-Persönlichkeitsinventar (WSP)

Formulierung eines Regelkataloges zur Itemkonstruktion und
dessen Anwendung zur Erstellung eines Itempools

Verfasserin

Martina Maurer

Angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag. rer. nat.)

Wien, im Dezember 2009

Studienkennzahl: 298

Studienrichtung: Psychologie

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

Inhaltsverzeichnis

1	Einleitung	5
2	Theoretischer Rahmen	9
2.1	Begriffsbestimmung	9
2.1.1	Begriffsbestimmung Persönlichkeitsfragebogen	9
2.1.2	Begriffsbestimmung Item	10
2.2	Problematik von Persönlichkeitsfragebögen	11
2.3	Stufen der Itempoolkonstruktion	12
2.3.1	Anforderungsanalyse und Problemstellung	12
2.3.2	Eingrenzung des Merkmals und Arbeitsdefinition	13
2.3.3	Festlegung der Zielgruppe, Zweck und Art der Information	13
2.3.4	Wahl der Testart	13
2.3.5	Itemauswahl und -entwicklung	14
2.3.6	Revision des Itempools	14
3	Anforderungsanalyse – allgemeine Studierfähigkeit	15
3.1	Einleitung	15
3.2	Begriffsbestimmung Anforderungsanalyse	16
3.3	Ziel der Untersuchung	17
3.4	Untersuchungsdesign	17
3.5	Auswahl der Studienrichtungen	18
3.6	Methode	21
3.7	Material	24
3.7.1	Fragebogen	24
3.7.2	Prätest	26
3.7.3	Interviewleitfaden	26
3.8	Durchführung	29
3.8.1	Fragebogenerhebung Studierende	30
3.8.2	Experteninterviews Lehrpersonen	32
3.9	Auswertung	33
3.9.2	Schritt 1: Dokumentation	33
3.9.3	Schritt 2: Kategorienbildung	34
3.9.4	Schritt 3: Interpretation	35

3.10	Ergebnis Anforderungsanalyse	39
4	Regelkatalog zur Fragebogen-Itemkonstruktion	43
4.1	Einleitung	43
4.2	Problematik Itemkonstruktion	44
4.3	Methode	46
4.3	Quantifizierbare Kriterien zur Überprüfung der Itemqualität	47
4.3.1	Objektivität von Items	47
4.3.2	Reliabilität von Items.....	48
4.3.3	Validität von Items	48
4.3.4	Itemschwierigkeit	48
4.3.5	Eindimensionalität von Items	49
4.4	Formale Gestaltung von Items	49
4.4.1	Gestaltung der Antwortformate.....	49
4.4.1.1	Antworttendenzen.....	50
4.4.1.1.1	Tendenz zu Extremurteilen.....	50
4.4.1.1.2	Tendenz zur Mitte	50
4.4.1.1.3	Akquieszenz, Zustimmung- oder Ja-Sage Tendenz.....	51
4.4.1.1.4	Reihenfolgeeffekte.....	51
4.4.1.1.5	Positionseffekte	52
4.4.1.1.6	Reaktanz.....	52
4.4.1.1.7	Soziale Erwünschtheit	52
4.4.1.2	Antwortformate.....	53
4.4.1.2.1	Freies Antwortformat	53
4.4.1.2.2	Gebundenes Antwortformat.....	54
4.4.1.2.2.1	Ordnungsaufgaben.....	54
4.4.1.2.2.2	Auswahlaufgaben.....	55
4.4.1.2.2.3	Beurteilungsaufgaben	57
4.4.1.2.3	Aufgaben mit atypischem Antwortformat.....	59
4.4.2	Wahl des Antwortformates.....	59
4.5	Sprachliche Itemgestaltung	60
4.5.1	Kategorisierung von Itemarten.....	60
4.5.2	Sprachliche Verständlichkeit.....	62
4.5.3	Eindeutigkeit des Iteminhalts	63
4.6	Itempolung	64
4.7	Experteninterview.....	65
4.8	Ergebnis Regelkatalog	66

5	Itempool des Wiener Studieneignungs-Persönlichkeitsinventar (WSP)	71
5.1	Einleitung	71
5.2	Auswahl und Definition der Merkmale	72
5.2.1	Auswahl der Merkmale (Dimensionen)	72
5.2.2	Definition der Merkmale	74
5.2.2.1	Arbeitsdefinitionen	74
5.2.2.2	Definitionen von Merkmalen im Testmanual	75
5.3	Zielgruppe	75
5.4	Zweck des WSP	76
5.5	Art der Information	76
5.6	Testlänge	77
5.7	Testadministration	77
5.8	Konstruktion des Itempools	78
5.8.1	Methode – Situative Fragen	79
5.8.2	Itemstruktur WSP	81
5.9	Durchführung	82
5.9.1	Instrumente zur Qualitätssicherung	84
5.9.1.1	Anwendung des ‚Regelkataloges zur Fragebogen-Itemkonstruktion‘	84
5.9.1.2	Internetplattform	84
5.9.1.3	Workshop	87
5.9.2	Probleme bei der Itementwicklung und Ergänzung des Regelkataloges	88
5.10	Itemverrechnung	91
5.11	Ergebnis Itempool	91
6	Diskussion	95
7	Zusammenfassung	99
8	Bibliographie	103
9	Abbildungs- und Tabellenverzeichnis	111
10	Anhang	113
	Abstract	153
	Curriculum Vitae	155

1 Einleitung

Seit dem Hochschulrahmengesetz 1998 wird die Auswahl von Studierenden stärker in den Zuständigkeitsbereich der Universitäten verlagert. (Hell, Ptok & Schuler, 2007; Schmitt, 2005; Zimmerhofer & Hornke, 2005). Dadurch wurde gleichzeitig eine Entwicklung initiiert, welche die Konstruktion psychologisch-diagnostischer Verfahren zur Selektion von Studierenden forciert.

Im Gegensatz zum allgemeinen Trend, soll der hier zu entwickelnde Itempool des Wiener Studieneignungs-Persönlichkeitsinventar (WSP) nicht in Selektionssituationen, sondern als Instrument zur Verbesserung der Beratung von StudienbeginnerInnen und StudentInnen eingesetzt werden. Der Persönlichkeitsfragebogen WSP verfolgt demnach einen förderdiagnostischen Ansatz - ein Stärken- Schwächenprofil aufzuzeigen – um anhand detaillierter Rückmeldungen bei Bedarf Interventionsmöglichkeiten setzen zu können.

Um die Beratung allen StudieninteressentInnen zugänglich zu machen, entwickelt der Arbeitsbereich Diagnostik des Institutes für Entwicklungspsychologie und Psychologische Diagnostik (Fakultät für Psychologie, Universität Wien) ein webbasiertes Beratungskonzept für ein psychologisches Self-Assessment. Interessierte Personen können Psychologisch-diagnostische Verfahren zum Thema Studienwahl und -eignung via Internet durchführen und erhalten anonym Feedback zu ihrem Stärken- und Schwächenprofil. Derzeit befindet sich das Self-Assessment in Erprobung. Nähere Informationen dazu unter: <http://www.studienberatungstest.at/> (Frebort & Kubinger, 2006b). Geplant ist, den hier zu entwickelnde Persönlichkeitsfragebogen als Teil dieser psychologischen Testbatterie online, anonym und gratis allen InteressentInnen zur Verfügung gestellt werden.

Der diesem diagnostischen Vorgehen zugrunde liegende Gedanke ist, durch ein auf den Studienalltag zugeschnittenes Verfahren die Beratung von StudienbeginnerInnen zu verbessern und eine Entscheidungshilfe für die individuelle Laufbahnentwicklung zu bieten. Eine erhöhte Passung von StudienanfängerInnen und Studienrichtung wirkt sich dabei sowohl auf die Optimierung des Studienerfolgs, die erlebte Studienzufriedenheit als auch auf die Verringerung der Abbruchrate in den Studiengängen positiv aus (Hell, Ptok & Schuler, 2007).

Die Items bezüglich der Iteminhalte optimal auf die gewünschte Zielgruppe zuzuschneiden, stellt ein ausgewiesenes Qualitätskriterium des zu entwickelnden Itempools dar. Zielgruppe sind MaturantInnen und StudienbeginnerInnen, die Beratung bezüglich Studienwahl und Studieneignung suchen, sowie StudentInnen, welche Schwierigkeiten während des Studiums erleben und hinsichtlich eines möglichen Studienwechsels beraten werden wollen.

Um dies zu erreichen, wurde in einem ersten Schritt eine Anforderungsanalyse zur allgemeinen Studierfähigkeit nach der Methode der Critical Incident Technique (Flanagan, 1954) in Zusammenarbeit mit Khorramdel und Kubinger durchgeführt. Dazu wurden 208 Studierende und 71 Lehrpersonen mittels Fragebogenerhebung und Experteninterviews befragt. Die Verwendung der Critical Incident Daten ermöglicht sowohl eine Konkretisierung der Iteminhalte als auch einen Praxisbezug zur interessierenden Fragestellung – der Studieneignung und -passung – herzustellen. Die Anforderungsanalyse wurde als Poster beim Internationalen Psychologen Kongress in Berlin vorgestellt¹ und wird nun in dieser Diplomarbeit erstmalig beschrieben.

Das Ziel dieser Diplomarbeit ist die Erstellung eines Itempools für einen Persönlichkeitsfragebogen zur Erfassung studienrelevanter Eigenschaften – das Wiener Studieneignungs-Persönlichkeitsinventar (WSP).

Zentrale Fragestellung ist dabei, wie die Qualität während des Erstellungsprozesses bestmöglich sichergestellt werden kann. Neben der ausführlichen Dokumentation der Vorgehensweise der Itempoolkonstruktion wird ein Instrumentarium entwickelt, welches Regeln für die Itementwicklung zur Verfügung stellt. Darüber hinaus werden zusätzliche Instrumente vorgestellt, die im Rahmen der Itemkonstruktion die Qualität der Items erhöhen können.

Als grundlegend für die Konstruktion des angestrebten psychologisch-diagnostischen Verfahrens wurden folgende drei Prozesse definiert, welche zeitlich und inhaltlich aufeinander aufbauen (siehe Tab. 1):

Darstellung des Arbeitsprozesses		
Schritt 1	Anforderungsanalyse	Erstellung eines Anforderungsprofils für die universitäre Hochschulausbildung in Wien
Schritt 2	Regelkatalog zur Fragebogen-Itemkonstruktion	Vor der Entwicklung des Itempools wird als Maßnahme zur Qualitätssicherung ein Regelkatalog entwickelt, der Handlungsschritte und Formulierungshinweise zur Itemkonstruktion enthält.

¹ Khorramdel, L., Maurer, M. & Kubinger, K.D. (2008). *A requirement analysis of study specific demands – What requirements of ability and personality do students need to be successful?* 29th International Congress of Psychology, Berlin, 20th-25th July 2008, Unveröffentlichtes Poster.

Schritt 3	Konstruktion des Itempools	Auf Basis des Regelkataloges werden die Items für das Wiener Studieneignungs-Persönlichkeitsinventar entwickelt. Inhaltliche Grundlage dafür sind die Daten der Anforderungsanalyse.
-----------	----------------------------	--

Tab. 1 Darstellung des Arbeitsprozesses

Der strukturelle Aufbau der Diplomarbeit orientiert sich an diesen drei Arbeitsprozessen (Anforderungsanalyse, Regelkatalog zur Fragebogen-Itemkonstruktion, Entwicklung des Itempools). Jeder dieser Arbeitsprozesse wird nach der I-M-R-Methode (Introduction – Method – Result) beschrieben.

Zur theoretischen Einbettung der Vorgehensweise wird vorab eine Fundierung der Begriffe *Persönlichkeitsfragebogen* und *Item* vorgenommen. Ebenso werden die Stufen zur Itemkonstruktion theoretisch skizziert. Diese Skizzierung der Abfolge bei der Itemkonstruktion hilft der Eingliederung der einzelnen Bausteine der Diplomarbeit zu einem Gesamtkonstrukt und bildet dadurch eine theoretische Klammer vom Beginn bis zum Ende des Forschungsprozesses.

2 Theoretischer Rahmen

Im folgenden Kapitel werden die Begriffe des Tests und des Persönlichkeitsfragebogens theoretisch fundiert und auf die Problematik von Persönlichkeitsfragebögen als psychologisch-diagnostische Verfahren eingegangen. Anschließend werden die Stufen zur Konstruktion eines Itempools vorgestellt. Neben der theoretischen Fundierung liefert dieses Kapitel damit eine Übersicht über die Arbeitsschritte, welche im Anschluss bei der Itempoolkonstruktion des WSP durchgeführt werden.

2.1 Begriffsbestimmung

2.1.1 Begriffsbestimmung Persönlichkeitsfragebogen

Um den Begriff des Persönlichkeitsfragebogens zu bestimmen, ist es ratsam sich vorerst mit dem Begriff des Tests auseinander zu setzen, um den Persönlichkeitsfragebogen als Sonderform des psychologisch-diagnostischen Verfahrens anschließend davon abzugrenzen.

Nach Lienert und Raatz (1998, 1) ist „ein Test [...] ein wissenschaftliches Routineverfahren zur Untersuchung eines oder mehrerer empirisch abgrenzbarer Persönlichkeitsmerkmale mit dem Ziel einer möglichst quantitativen Aussage über den relativen Grad der individuellen Merkmalsausprägung.“

Mit der Eingrenzung auf wissenschaftliche Routineverfahren unterscheidet sich der Begriff des Tests von dessen umgangssprachlichen Gebrauch, der vielfach mit Schularbeiten im herkömmlichen Sinn des Schulalltages verbunden wird und meist beliebig ist, sowohl in der Entwicklung als auch der Interpretation der „Test“-Ergebnisse. Demgegenüber wird im Kontext der psychologischen Diagnostik von wissenschaftlich fundierten Gütekriterien ausgegangen. Zu den Hauptgütekriterien zählen Objektivität, Reliabilität und Validität. Die Nebengütekriterien umfassen nach Kubinger (2006) Normierung, Skalierung, Ökonomie, Nützlichkeit, Zumutbarkeit, Unverfälschbarkeit und Fairness. Näheres hierzu bei Kubinger (2006) oder Lienert und Raatz (1998). Darüber hinaus bezieht sich der Begriff der Wissenschaftlichkeit ganz im Gegensatz zur angesprochenen Beliebigkeit auf die theoretische Fundierung und testtheoretischen Orientierung desselben.

Der Begriff des Routineverfahrens impliziert außerdem, dass der Test an einer repräsentativen Stichprobe geeicht sein muss (Eichung) und aufgrund einer detaillierten

Beschreibung zur Durchführung und Auswertung unter Standardbedingungen auch von anderen TestleiterInnen angewandt werden kann (Objektivität).

Ein psychologisch-diagnostisches Verfahren prüft empirisch abgrenzbare Persönlichkeitsmerkmale, d.h. verhaltens- und erlebnisanalytische, phänomenologische Eigenschaften, Verhaltensdispositionen, Fähigkeiten, Fertigkeiten oder Kenntnisse. Der von Lienert und Raatz (1998) genannte relative Grad der individuellen Merkmalsausprägung definiert die Einordnung eines Individuums innerhalb einer Gruppe von Individuen oder in Bezug auf ein bestimmtes Kriterium.

Die hier definierten Kriterien eines Tests können für einen Persönlichkeitsfragebogen, der den Anspruch erhebt, psychologische Merkmale quantitativ zu erfassen, übernommen werden.

Darüber hinaus sind Persönlichkeitsfragebögen dadurch charakterisiert, dass sie im Gegensatz zu Leistungstests keine Leistung, sondern eine Selbstauskunft von den befragten Personen verlangen. Die Antworten der befragten Personen werden dabei nicht nach Korrektheit der Antwort in Bezug auf eine eindeutige Lösung kategorisiert, sondern danach, ob und in welcher Ausprägung das interessierende Merkmal vorhanden ist. Neben Persönlichkeitsfragebögen zur Selbstbeurteilung gibt es auch Fragebögen, die zur Fremdbeurteilung entwickelt werden, d.h. andere Personen schätzen die untersuchte Person hinsichtlich der Merkmalsausprägungen ein (Jonkisz & Moosbrugger, 2007).

2.1.2 Begriffsbestimmung Item

Das Grundgerüst des Persönlichkeitsfragebogens wird durch die Items gebildet. Nach Rost (2004, 55) ist ein „Item [...] die kleinste Beobachtungseinheit in einem Test, sozusagen der elementare Baustein, aus dem ein Test gebaut ist. An einem Item lassen sich zwei Komponenten unterscheiden, nämlich der so genannte Itemstamm und das Antwortformat.“

Der Itempool ist die Gesamtanzahl von Items, welche für die Konstruktion eines Persönlichkeitsfragebogens entwickelt werden. Aus diesem Itempool werden die Items für den Persönlichkeitsfragebogen ausgewählt. In der Mehrzahl der Fragebögen werden alle Items vorgegeben. Beim Adaptiven Testen hingegen, werden nur jene Items vorgegeben, die dem jeweiligen Schwierigkeitsgrad der Person entsprechen. Der Itempool eines adaptiven psychologisch-diagnostischen Verfahrens ist somit um ein Vielfaches größer als die eingeschränkte Anzahl bei der Testvorgabe (Kubinger, 2006).

2.2 Problematik von Persönlichkeitsfragebögen

Das psychologisch-diagnostische Verfahren des Persönlichkeitsfragebogens steht vermehrt in der Kritik der psychologisch-diagnostischen Forschungslandschaft. Vor der Entwicklung eines Persönlichkeitsfragebogens ist es daher ratsam, sich mit den Kritikpunkten des Verfahrens auseinander zu setzen und gegebenenfalls Rückschlüsse auf den Entwicklungsprozess und das Einsatzgebiet zu ziehen. Im Folgenden wird kurz auf die wichtigsten Kritikpunkte eingegangen.

Ein gängiger Kritikpunkt am Persönlichkeitsfragebogen ist die hohe Augenscheinvalidität und Durchschaubarkeit von Persönlichkeitsfragebögen (Kubinger, 2002). D.h. den Testpersonen ist leicht ersichtlich, welche Merkmalsdimensionen durch den Persönlichkeitsfragebogen erfasst werden und können eine bewusste Verfälschung ihrer Selbstdarstellung erwirken (faking bad). Der Persönlichkeitsfragebogen ist somit nicht resistent gegenüber einer möglichen Verfälschbarkeit. Eine Variante, die Verfälschbarkeit bei der Erhebung von Persönlichkeitseigenschaften zu reduzieren, stellen Objektive Persönlichkeitstests dar (Kubinger, 2006). Von einer Anwendung des Persönlichkeitsfragebogens im Rahmen einer Selektionssituation (z.B. Bewerbung oder Studierendenauswahl) wird daher abgeraten.

Obwohl Persönlichkeitsfragebögen sowohl zur Selbst- als auch Fremdbeurteilung konzipiert werden können, bedient sich der Großteil der Fragebögen der Selbstauskunft durch die befragte Person. Der Persönlichkeitsfragebogen als subjektives Messverfahren steht dabei mehrfach in der Kritik. Subjektive Informationen, welche von der Person selbst gegeben werden, unterliegen mehreren Fehlerquellen. So ist beispielsweise zu berücksichtigen, ob die befragten Personen die kognitiven Kompetenzen zur Einschätzung ihrer eigenen Persönlichkeitsstruktur mitbringen oder nicht.

Eine weitere Voraussetzung ist die Bereitschaft zur Offenbarung. Problematisch kann dies einerseits in Selektionssituationen sein, wenn die Person gemäß eines erwarteten Idealbildes sozial erwünscht antwortet. Andererseits wird das Gütekriterium der Zumutbarkeit verletzt, wenn Fragen zu stark auf den Intimbereich einer Person abzielen. Die Bereitschaft zur Offenbarung ist somit auch abhängig von den Iteminhalten (Rost, 2004).

Die Verwendung von Persönlichkeitsfragebogen ist somit von der Art der Fragestellung und dessen Einsatzbereich abhängig. Bei Zielgruppen, bei denen eine der oben dargestellten Voraussetzungen nicht gegeben sind (z.B. Kindern, spezifische klinische Fragestellungen) ist ein Persönlichkeitsfragebogen zur Selbstauskunft nicht zielführend. Neben der Zielgruppe

ist darüber hinaus auch auf die Zumutbarkeit der Iteminhalte bei der Erstellung von Fragebogenitems zu achten.

2.3 Stufen der Itempoolkonstruktion

Im Zentrum dieser Diplomarbeit steht die Entwicklung des Itempools für das Wiener Studieneignungs-Persönlichkeitsinventar (WSP). Dieses Unterkapitel liefert eine theoretische Übersicht über die Arbeitsschritte, welche für die Itempoolerstellung durchlaufen werden müssen. Dazu werden die Stufen zur Konstruktion eines Itempools in Anlehnung an Bühner (2006) beschrieben.

Die Reihenfolge der Stufen wurde aufgrund theoretischer Überlegungen und empirischer Erfahrung leicht abgeändert, um die praktische Umsetzbarkeit zu erhöhen. Beispielsweise ist es sinnvoll die Testart festzulegen, bevor die Entscheidung über Fragetyp und Antwortformat getroffen werden.

2.3.1 Anforderungsanalyse und Problemstellung

Zu Beginn der Entwicklung eines Persönlichkeitsfragebogens steht die Klärung der zu erhebenden psychologischen Konstrukte. Psychometrische Tests beziehen ihre Legitimation „aus der Annahme, dass das Verhalten von Personen mit Hilfe von Persönlichkeitsmerkmalen (Konstrukten) erklärt werden kann. Diese Konstrukte sollen mit Hilfe des Tests erfasst werden. Am Beginn der Entstehung steht deshalb die Wahl des zu messenden Persönlichkeitsmerkmals bzw. im Fall einer Testbatterie des zu messenden Merkmalsspektrums (Jonkisz & Moosbrugger, 2007, 28).“

Auch nach den Qualitätsanforderungen der DIN 33430 ist die Definition der zu erhebenden Merkmale in Form eines Anforderungsprofils zentrales Element im qualitätsgesicherten psychologisch-diagnostischen Prozess der Fragebogenentwicklung (Heyse & Kersting, 2004; Kersting, 2008).

Am Beginn der Entwicklung des Wiener Studieneignungs-Persönlichkeitsinventars stellten sich daher folgende Fragen: Welche Persönlichkeitsmerkmale sind entscheidend für die erfolgreiche Bewältigung einer Hochschulausbildung und in welchem Ausmaß? Worin unterscheiden sich erfolgreiche von weniger erfolgreichen Studierenden aus psychologischer Sicht?

2.3.2 Eingrenzung des Merkmals und Arbeitsdefinition

Ein Anforderungsprofil kann in den seltensten Fällen als Ganzes in ein Testinventar übernommen werden. Vielmehr bedarf es der Auswahl bestimmter Merkmale, welche sinnvollerweise durch eine Verfahrensart erfasst werden können. Daher ist es notwendig die zu erhebenden Konstrukte einzugrenzen und genau zu definieren. Erst auf Basis einer genauen Arbeitsdefinition ist es möglich, präzise Items zu formulieren. An die Formulierung eines Anforderungsprofils schließt somit die Frage an, welche der im Anforderungsprofil definierten psychologischen Konstrukte sind für die Fragestellung der Studieneignung hinreichend relevant und mithilfe eines Persönlichkeitsfragebogens messbar?

2.3.3 Festlegung der Zielgruppe, Zweck und Art der Information

Nach der Festlegung des Messgegenstandes wird überlegt, welcher Zielgruppe der Test vorgelegt und welchen Zweck er verfolgen soll. Die Formulierung der Zielgruppe ist von der Art der Fragestellung abhängig. Durch die Definition der Zielgruppe wird das Vorwissen der zu befragenden Personen spezifiziert und die Iteminhalte können auf dieses Vorwissen abgestimmt werden. Je nach Zweck des psychologisch-diagnostischen Verfahrens unterscheidet Bühner (2006) zwischen der Intention der Trennung von Personengruppen bezüglich des Vorhandenseins eines Merkmals (z.B. Hyperaktivität) und der Beschreibung von Eigenschaften, Fähigkeiten oder Fertigkeiten einer Person.

Neben der Zielgruppe und dem Zweck des Persönlichkeitsfragebogens wird die Art der zu erhebenden Information bestimmt. Dabei kann zwischen subjektiven und objektiven Informationen unterschieden werden. Während subjektive Informationen von der Person selbst in Form einer Selbstbeurteilung gegeben werden, bauen objektive Informationen auf eine Fremdbeurteilung durch andere Personen auf. Auf den missverständlichen Gebrauch des Wortes „objektiv“ sei an dieser Stelle hingewiesen, denn auch Fremdbeurteilungen bauen auf der subjektiven Einschätzung und Wahrnehmung anderer Personen auf und unterliegen damit gewissen Verzerrungen.

2.3.4 Wahl der Testart

Nach der Festlegung der Zielgruppe, des Zwecks und der Art der Information wird die Testart festgelegt. Die Testart bezieht sich auf die Form des Verfahrens (z.B. Objektiver Persönlichkeitstest, Objektiver Leistungstest, Persönlichkeitsfragebogen, Projektives Verfahren etc.). Die Auswahl der Testart ist abhängig von der Art und der Anzahl der zu

erhebenden Konstrukte (z.B. kognitiven Fähigkeiten vs. Persönlichkeitseigenschaften), der Zielgruppe (z.B. Kinder vs. Erwachsene, klinische vs. eignungsdiagnostische Fragestellungen) und der Testsituation (Selektionsstichprobe vs. Freiwilligkeit der Testung, Gruppen- oder Einzeltestung), und bedarf reiflicher Überlegung.

2.3.5 Itemauswahl und -entwicklung

Die Itemauswahl bezieht sich einerseits auf die Form der Fragestellung und des Antwortformates und andererseits auf die in diesem Format dargebotene Itemmenge.

Sowohl der Itemstamm als auch das Antwortformat unterliegen vielfältigen Überlegungen. Zentrale Fragestellung ist dabei, wie das zu erhebende Konstrukt operationalisiert werden kann. Beispielsweise besteht die Aufgabe der Testpersonen im Test d2 (Brickenkamp, 2002) zur Erhebung der Aufmerksamkeit (Konzentration) darin, innerhalb einer vorgegebenen Zeit nur die doppelt unterstrichenen Buchstaben „d“ in einer Buchstabenreihenfolge bestehend aus „d“s und „p“s (mit unterschiedlich vielen Strichen) durchzustreichen.

Die Itemmenge legt die Länge des Tests fest. Davon abhängig ist die Testzeit, die sich aus der für die Beantwortung der Testaufgaben benötigten oder vorgegebenen Zeit ergibt (Lienert & Raatz, 1998). Um genügend Informationen über die Ausprägung des zu erhebenden Konstruktes bei einer Testperson zu gewinnen ohne gleichzeitig die Zumutbarkeit des Verfahrens (aufgrund zu langer Bearbeitungsdauer) zu strapazieren, kommen den Überlegungen zur Itemanzahl große Bedeutung zu.

2.3.6 Revision des Itempools

Der fertig gestellte Itempool wird vor der empirischen und statistischen Überprüfung einer nochmaligen Revision unterzogen. Besonderes Augenmerk wird dabei auf die sprachliche Formulierung der Items gelegt. Die Verständlichkeit der Darstellung komplizierter Sachverhalte, die Vermeidung verschachtelter und umständlicher Satzkonstellationen ebenso wie die Verwendung von Fachausdrücken, aber auch die Polung der Items (Fragerichtung) werden überprüft (Mummendey, 2008).

In der vorliegenden Diplomarbeit werden die oben dargestellten Schritte am Beispiel der Itempoolerstellung des Wiener Studieneignungs-Persönlichkeitsinventars (WSP) beschrieben.

3 Anforderungsanalyse – allgemeine Studierfähigkeit

3.1 Einleitung

Die DIN 33430 dient der Qualitätssicherung und –optimierung in der berufsbezogenen Eignungsdiagnostik. Nach der DIN 33430, definiert als Prozessnorm, ist die Gewährleistung von Qualität nur dann gegeben, wenn der gesamte Prozess über alle Stufen der beruflichen Eignungsbeurteilung hinweg bestimmten Qualitätsanforderungen entspricht (Reimann, 2005). Die Verfahrensentwicklung beginnt somit nicht erst mit der Entwicklung des psychologisch-diagnostischen Verfahrens, sondern bei der Festlegung der zu erhebenden Anforderungen.

Die Fragestellung zu Beginn des Forschungsvorhabens war entsprechend jene nach einem Anforderungsprofil für die universitäre Hochschulausbildung in Wien. In Zusammenarbeit mit Khorramdel und Kubinger wurde in einer einjährigen Studie ein Anforderungsprofil für die allgemeine Studierfähigkeit nach der Critical Incident Technique (Flanagan, 1954) erstellt. Dazu wurden StelleninhaberInnen und Vorgesetzte – Studierende und Lehrpersonen – an drei Wiener Universitäten mittels Fragebogenerhebung und Experteninterviews befragt.

Das Ergebnis und die Daten der Anforderungsanalyse stellen die Grundlage für die anschließende Verfahrensentwicklung dar:

- Einerseits ist das Anforderungsprofil Ausgangspunkt für die Festlegung und Definition der Merkmale, die mit dem WSP erfasst werden sollen.
- Andererseits stellen die in der Anforderungsanalyse gewonnenen Daten die Grundlage für die Entwicklung des Itempools, wodurch die geforderte Realitätsnähe der Items sichergestellt werden kann. Unter Realitätsnähe wird dabei verstanden, dass die Items spezifisch auf den Studienalltag zugeschnitten und somit für die Testpersonen verständlich und nachvollziehbar sind. Die Qualität der Daten aus der Anforderungsanalyse spiegelt sich in der Qualität der Items wieder.
- Drittens – und dem gilt besonderes Augenmerk – ist die Verschränkung dieser Arbeitsprozesse ein wesentliches Qualitätsmerkmal dieser Arbeit und entspricht den Qualitätsanforderungen der DIN 33430 (Heyse & Kersting, 2004; Kersting, 2008).

3.2 Begriffsbestimmung Anforderungsanalyse

„Anforderungen einer Tätigkeit an den Menschen sind die Gesamtheit der für das forderungsgerechte Ausführen benötigten körperlichen und geistigen Leistungsvoraussetzungen (Hacker, 2005, 74).“ Das Anforderungsspektrum reicht somit von physischen bis zu psychologischen Leistungsvoraussetzungen, wobei diese wiederum in kognitive Fähigkeiten und Persönlichkeitseigenschaften differenziert werden können.

Ein Anforderungsprofil fasst nach Ortner (2003) die Menge aller Anforderungen zusammen und beschreibt die notwendigen Ausprägungen in den relevanten Leistungs- und Persönlichkeitsdimensionen. Während Ortner bei der Begriffsbestimmung größeres Augenmerk auf die Benennung und Gewichtung der Anforderungen in einem Anforderungsprofil legt, setzt Reimann den Bezug zu der zu besetzenden Stelle: „[Das Anforderungsprofil] enthält alle benötigten und wünschenswerten Voraussetzungen und Kompetenzen einer Person für den zu besetzenden Arbeitsplatz, das Aufgabenfeld, eine Ausbildung oder einen Beruf einschließlich der Merkmale, die für die berufliche Zufriedenheit wichtig sind (2005, 116).“

Grundlage für die Bestimmung von Anforderungen ist nach der DIN 33430 die Arbeits- und Anforderungsanalyse. Die DIN 33430 definiert die Arbeitsanalyse als „Methode der Identifizierung der an einem Arbeits-/Ausbildungsplatz oder in einem Beruf auszuführenden Aufgaben oder auszuübenden Tätigkeiten, ihrer Ausführungsbedingungen sowie ihrer psychischen, physischen und sozialen Umfeldbedingungen und Organisationsmerkmale (DIN Norm 3.2, zit.n. Heyse & Kersting, 2004, 30).“

Laut Dorsch Psychologischem Wörterbuch ist eine Arbeits(platz)analyse „in der Arbeitswissenschaft ein Oberbegriff für «alle Methoden, Verfahren und Instrumente, die dazu dienen, Informationen über die Arbeitstätigkeiten, die organisatorisch-technischen Arbeitsbedingungen, die Arbeitsmittel und Werkzeuge sowie deren Auswirkungen auf den Menschen zu sammeln, zu verarbeiten und zu interpretieren» (Frieling & Buch, 2006, zit.n. Häcker & Stapf, 2009, 63).“

An die Arbeitsanalyse schließt die Anforderungsanalyse an. Dabei werden die Arbeitstätigkeiten, Arbeitsbedingungen und verhaltensbezogenen Anforderungen in psychologische Konstrukte übersetzt. Sie dient der „Ermittlung von personrelevanten psychischen und psychophysischen Voraussetzungen [...] für den zu besetzenden Arbeitsplatz, das Aufgabenfeld, die Ausbildung bzw. das Studium oder den Beruf, für das/den/die die Eignung eines Kandidaten festgestellt werden soll (DIN Norm 3.1, zit.n

Heyse & Kersting, 2004, 30).“ Nach Kubinger (2006) geht es neben den Kompetenzen wie Fach- und Methodenkompetenz (Fachkenntnisse und Berufserfahrung), gelegentlich wichtigen körperlichen Voraussetzungen und formalen Qualifikationen, vor allem um psychologische Kompetenzen, insbesondere kognitive Fähigkeiten, Persönlichkeitseigenschaften und Sozialkompetenz.

3.3 Ziel der Untersuchung

Ziel der Anforderungsanalyse war die Erstellung eines Anforderungsprofils für die allgemeine Studierfähigkeit für die Hochschulausbildung in Wien. Anhand von Expertenbefragungen mit Studierenden und Lehrpersonen nach der Critical Incident Technique (Flanagan, 1954) wurden an drei Universitäten Wiens Anforderungsprofile für sieben Studienrichtungen erstellt. Die Anforderungsprofile werden nach Integration bestehender Anforderungsprofile zu einem standortspezifischen Anforderungsprofil über die allgemeine Studierfähigkeit zusammengefasst.

3.4 Untersuchungsdesign

Die Anforderungsanalyse folgte sechs Arbeitsschritten, die der Übersicht halber in einem Prozessdiagramm dargestellt werden. Die einzelnen Arbeitsschritte werden im Anschluss diskutiert.

Untersuchungsdesign Anforderungsanalyse		
Schritt 1	Formulierung des Ziels der Untersuchung	Entwicklung eines standortspezifischen Anforderungsprofils für die universitäre Hochschulausbildung in Wien. Charakterisierung von erfolgreichen und weniger erfolgreichen Studierenden durch Operationalisierung des Studienerfolgs anhand erfolgsförderlicher und –hinderlicher Verhaltensweisen in kritischen Situationen.
Schritt 2	Auswahl der Studienrichtungen	Die Hauptstudienrichtungen wurde nach dem RIASEC Modell von Holland 1977 (zit. n. Bergmann & Eder, 2005) und der Einteilung von Forschungsbereichen der Österreichischen Akademie der Wissenschaften (ÖAW,

		2006) nach inhaltlicher Zusammengehörigkeit geclustert. Aus den Clustern wurden die sieben Studienrichtungen für die Anforderungsanalysen ausgewählt.
Schritt 3	Planung der Expertenbefragung	Entwicklung der Erhebungsmaterialien für die Fragebogenerhebung und Experteninterviews mit Studierenden und Lehrpersonen nach der Critical Incident Technique (Flanagan, 1954).
Schritt 4	Durchführung der Untersuchung	Es wurden 208 Studierende und 71 LehrveranstaltungsleiterInnen der sieben Studienrichtungen befragt (WS und SS 2007/08).
Schritt 5	Datenanalyse	Für jede Studienrichtung wurde ein Anforderungsprofil ermittelt. Dazu wurden die von den Studierenden und Lehrpersonen genannten kritischen Ereignisse in Situationsklassen zusammengefasst. Die Verhaltensnennungen wurden von zwei unabhängigen Beurteilerinnen nach einem induktiv ermittelten Kategorieschema kategorisiert und von diesem auf die zugrunde liegenden Anforderungen geschlossen. Grundlage für die Datenanalyse ist die Inhaltsanalyse nach Mayring (2008) und das Konstruktschema von Frebort und Kubinger (2006). Das ermittelte Kategorieschema wird in ein Anforderungsprofil übersetzt.
Schritt 6	Ergebnisdarstellung und Formulierung des Anforderungsprofils	Die Anforderungsprofile der sieben Studienrichtungen werden zu einem gemeinsamen Anforderungsprofil über die allgemeine Studierfähigkeit akkumuliert, ergänzt um ein achttes bereits bestehendes Anforderungsprofil von Frebort und Kubinger (2006).

3.5 Auswahl der Studienrichtungen

Nachdem sich die Universitäten Österreichs sowohl hinsichtlich ihrer organisatorischen Strukturen als auch ihrer inhaltlichen Ausrichtungen einzelner Studienrichtungen unterscheiden (Schwerpunktsetzungen), wurde ein Anforderungsprofil standortspezifisch für die Hochschulausbildung in Wien entwickelt. Es wird daher empfohlen, vor einer

Übertragbarkeit auf andere Studienorte die Passung des Anforderungsprofiles zu prüfen.

Um ein Anforderungsprofil für die allgemeine Studierfähigkeit zu entwickeln, ist die Auswahl der Studienrichtungen, die zu einem gemeinsamen Anforderungsprofil zusammengezogen werden, elementar. Die Auswahl repräsentativer und unterschiedlicher Studienrichtungen, stellvertretend für die Breite möglicher Ausbildungswege an Universitäten, war ein wesentliches Kriterium zu Beginn des Forschungsvorhabens. Die Studienrichtungen sollten aus möglichst unterschiedlichen inhaltlichen Ausrichtungen und Orientierungen stammen. Dazu war es notwendig Studienrichtungen in möglichst unterschiedliche, weit voneinander entfernte inhaltliche Klassen einzuordnen, aus denen jeweils eine Studienrichtung ausgewählt wurde. Grundlage für die Einteilung der Studienrichtungen in inhaltliche Cluster war das RIASEC-Modell des Allgemeinen-Interessen-Struktur-Test (Holland, 1973, zit.n. Bergmann & Eder, 2005) ergänzt um die Einteilung von Forschungsbereichen der Österreichischen Akademie der Wissenschaften (ÖAW, 2006).

Holland (1973, 1985a, 1997, zit.n. Bergmann & Eder, 2005) teilt in seinem RIASEC-Modell die Interessen, welche er als grundlegende Wesensmerkmale beschreibt, in sechs Orientierungen ein:

- Praktisch-technische Orientierung (Realistic): mechanischer, technischer, elektrotechnischer und landwirtschaftlicher Bereich
- Intellektuell-forschende Orientierung (Investigative): mathematischer und naturwissenschaftlicher Bereich
- Künstlerisch-sprachliche Orientierung (Artistic): Bereich Sprache, Kunst, Musik, Schauspiel und Schriftstellerei
- Soziale Orientierung (Social): Bereich der zwischenmenschlichen Beziehungen
- Unternehmerische Orientierung (Enterprising): Bereich der Führungs- und Überzeugungsqualität
- Konventionelle Orientierung (Conventional): rechnerischer und geschäftlicher Bereich

Anhand dieser Einteilung wurden die größten und bekanntesten Studienrichtungen den Clustern R, I, A, S, E und C zugeteilt. Das Studium der Architektur wurde nach Sichtung des Lehrplanes dem artistischem Cluster (Kulturwissenschaften) zugeordnet, gegensätzlich zur Einteilung des Berufes des/der ArchitektIn zum Cluster I (Code IAR) nach Holland.

Die Clusterung und Zuteilung wurde von ExpertInnen im Bereich der Psychologischen Diagnostik begutachtet. Der Anspruch zu Beginn des Forschungsvorhabens, die gesamte Fülle von Studienrichtungen dem Modell zuzuordnen, wurde aufgrund der Komplexität und des großen notwendigen Zeitaufwandes unterlassen. Ziel wäre es gewesen, anhand des Modells eine Aussage darüber zu machen, welches Anforderungsprofil für welche

Studienrichtungen übernommen werden könne. Dieser Anspruch war letztendlich nicht umsetzbar. Darüber hinaus ist es auch wahrscheinlich, dass Studienrichtungen innerhalb eines Clusters eine gewisse Heterogenität der Anforderungsprofile vorweisen, d.h. dass das Anforderungsprofil einer Studienrichtung nicht für eine andere Studienrichtung desselben Clusters in identer Form übertragen werden kann.

In Abb. 1 sind schematisch die größten und bekanntesten Studienrichtungen nach Einteilung in das RIASEC-Modell, entnommen aus dem Testmanual des Allgemeinen Interessen-Struktur-Test/Umwelt-Struktur-Test (Holland, 1973, zit.n. Bergmann & Eder, 2005) (Zeile 1) und der Einteilung von Forschungsbereichen der Österreichischen Akademie der Wissenschaften (ÖAW, 2006) (Zeile 2) dargestellt.

Intellektuell-Forschende Orientierung (I) bzw. Naturwissenschaftliche Orientierung				Soziale Orientierung (S)	Künstlerisch- Sprachliche Orientierung (A)			Praktisch-Technische Orientierung (R)	Unternehmerische Orientierung (E)	Konventionelle Orientierung (C)
Biologie & Medizin & Umwelt	Formalwissenschaften	Physik & Weltraumforschung	Erdwissenschaften	Sozialwissenschaften	Kulturwissenschaften	Sprach- & Literaturwissenschaften	Historische Wissenschaften			
Medizin	Mathematik	Physik	Erdwissenschaften	Psychologie	Architektur	Publizistik und Kommunikationswissenschaft	Geschichte	Maschinenbau	Betriebswirtschaft (BWL/IBWL)	Rechtswissenschaften
Biologie	Statistik	Astronomie	Geographie (und Wirtschaftskunde)	Soziologie	Theater-, Film- und Medienwissenschaft	Philosophie		Elektrotechnik, Informatik, Technische Informatik		Volkswirtschaftslehre (VWL)
						Anglistik		Bauingenieurwesen		

Abb. 1 Clusterung von Studienrichtungen

Nach der Gruppierung der Studienrichtungen folgte eine umfassende Recherche zu bestehenden methodisch fundierten Anforderungsprofilen dieser nunmehr zweiundzwanzig Studienrichtungen im deutschsprachigen Raum. Folgende Fragestellungen bildeten den Rahmen für die Recherche:

- Welche Anforderungsprofile zu welchen Studienrichtungen sind bereits vorhanden?
- Auf welcher Methode bauen diese auf?
- Sind sie nach Abgleich der Studienpläne auf den Standort Wien übertragbar?
- Welche Anforderungsprofile können für die Erstellung eines Anforderungsprofils über die allgemeine Studierfähigkeit herangezogen werden?

Es gibt zahlreiche Anforderungsprofile im deutschsprachigen Raum für die oben eingetragenen Studienrichtungen. Der Großteil davon ist jedoch nicht methodisch fundiert. Für die Studienrichtung Medizin im Cluster I (Intellektuell-forschende Orientierung) gibt es ein Anforderungsprofil, welches die Grundlage für den in der Schweiz und in Österreich durchgeführten Eignungstest für das Medizinstudium (EMS Test, Zentrum für Testentwicklung und Diagnostik, 1995) darstellt. Schüpbach, Pixner und Zapf (2005) entwickelten ein Anforderungsprofil für die Studienrichtung Anglistik (Cluster A). Das

Anforderungsprofil von Zimmerhofer und Hornke (2005) bezieht sich auf die Studienrichtungen Informatik, Elektrotechnik und technische Informatik (Cluster R). Hell, Ptok und Schuler (2007) entwickelten auf Basis der Critical Incident Technique ein Anforderungsprofil für die Studienrichtungen Wirtschaftswissenschaften (Cluster E und C). Im Cluster S wurde von Frebort und Kubinger (2006) ein Anforderungsprofil für Psychologie an der Universität Wien nach der Critical Incident Technique erstellt.

In einem letzten Schritt wurde je Cluster eine Studienrichtung nach dem Kriterium der Studierendennachfrage (bestehende Studierendenzahl) ausgewählt. Das Cluster I wurde aufgrund des Vorhandenseins eines Anforderungsprofils für Medizin und der geringen Studierendennachfrage in den anderen Studienrichtungen des Clusters nicht erhoben. Nach dem Kriterium der Studierendennachfrage ergaben sich in den übrigen Clustern R A S E und C folgende sieben Studienrichtungen (siehe Tab. 2), welche an drei Wiener Universitäten erhoben wurden.

Hochschule	Studienrichtung
Universität Wien	Publizistik- und Kommunikationswissenschaft (A) Rechtswissenschaften (C) Soziologie (S) Volkswirtschaftslehre (C)
Technische Universität	Architektur (A) Maschinenbau (R)
Wirtschaftsuniversität Wien	Betriebswirtschaftslehre und Internationale Betriebswirtschaftslehre (E)

Tab. 2 Studienrichtungen nach Universitäten

3.6 Methode

Nach Schuler (2002, 112) gibt es zur Bestimmung von Anforderungen drei methodische Zugänge:

1. Erfahrungsbasierte-intuitive Methode, bei der die Einschätzung der Anforderungen auf einer erfahrungsgeliteten Beurteilung der Tätigkeit, den erforderlichen Arbeitsmitteln und –gegenständen, den Umweltbedingungen und Qualifikationserfordernissen etc. beruht.
2. Arbeitsanalytisch-empirische Methoden, bei denen teil- oder vollstandardisierte Fragebogen und Arbeitsanalyseverfahren zum Einsatz kommen. Tätigkeitselemente werden ermittelt und in Personenanforderungen übersetzt.

3. Personenbezogene-empirische Methoden, bei denen die statistischen Zusammenhänge zwischen Personenmerkmalen und der erbrachten Leistung sowie Tätigkeitsfolgen (z.B. Belastung, Leistungshöhe, Zufriedenheit etc.) zur Bestimmung der Anforderungen genutzt werden.

Weiters kann die Datengewinnung bei Anforderungsanalysen nach dem Standardisierungsgrad unterschieden werden, wobei un-, teil- oder vollstandardisierte Untersuchungsmethoden eingesetzt werden können.

Für die Anforderungsanalyse wurde die Methode der Critical Incident Technique (CIT; Flanagan, 1954) in Anlehnung an die Studien von Westhoff (2005) sowie Frebort und Kubinger (2006) verwendet.

Die Critical Incident Technique (Flanagan, 1954) ist eine teilstandardisierte arbeitsanalytische-empirische Methode zur Erhebung von erfolgskritischen Situationen, den sogenannten *critical Incidents*. Die zugrunde liegende Annahme ist, dass sich gute und schlechte StelleninhaberInnen anhand unterschiedlicher Verhaltensweisen in kritischen Situationen erkennen lassen (Reimann, 2005).

Dazu werden ExpertInnen – StelleninhaberInnen und Vorgesetzte – zu den in ihrem Berufsfeld kritischen Situationen (Ereignissen) befragt. Unter kritischen Ereignissen werden all jene Situationen verstanden, deren erfolgreiche bzw. weniger erfolgreiche Bewältigung Aufschluss über die Fähigkeiten einer Person für die zu besetzende Stelle geben können.

„By an incident is meant any observable human activity that is sufficiently complete in itself to permit inferences and predictions to be made about the person performing the act (Flanagan, 1954, 327).“

Koch, Kici, Strobel und Westhoff (2006, 87) definieren erfolgskritische Situationen als „solche, in denen sich leistungsstarke Stelleninhaber im Vergleich zu leistungsschwachen durch effizientere Verhaltensweisen unterscheiden lassen.“

Zu jeder dieser kritischen Ereignisse werden die Befragten gebeten, Verhaltensweisen zu nennen, die zur erfolgreichen Bewältigung dieser Situation beitragen und solche, die für die erfolgreiche Bewältigung hinderlich sind. Aus den erhobenen kritischen Situationen und Verhaltensweisen (manifeste Elemente) kann auf die zugrunde liegenden notwendigen physischen und psychischen Leistungsvoraussetzungen (latente Elemente) induktiv geschlossen werden.

“[...] the critical incident technique, rather than collecting opinions, hunches, and estimates, obtains a record of specific behaviours from those in the best position to make the necessary observations and evaluations. The collection and tabulation of

these observations make it possible to formulate the critical requirements of an activity (Flanagan, 1954, 354).”

Der Vorteil dieser Methode liegt darin, dass die Nachvollziehbarkeit zwischen den Elementen (Situationen – Verhaltensweisen – Anforderungen) erhalten bleibt. Umgekehrt werden die Anforderungen durch die Verhaltensweisen explizit definiert; die Arbeitsdefinition der Anforderungen besteht aus der für diese Anforderungen kennzeichnende Sammlung von Verhaltensweisen.

Vielfältiges kann beispielsweise unter der Anforderung „Lernbereitschaft“ verstanden werden. Nach der CIT bestehen die Definitionen der Anforderungen aus den diese Anforderungen beschreibenden manifesten Inhalte: den kritischen Situationen und den dazugehörenden erfolgreichen und hinderlichen Verhaltensweisen. Beispielhaft sei hier die Definition der Anforderung „Lernbereitschaft“ aus der Anforderungsanalyse dargestellt:

<p>Lernbereitschaft: Kritik positiv sehen und als schöpferisches Potenzial verstehen, sich weiterbilden, etwas lernen wollen, Hilfe z.B. von LehrveranstaltungsleiterInnen annehmen können,... +</p>

Dadurch „verbessert [sich] in hohem Maße die praktische Handhabung des Anforderungsprofils. Aber auch die Objektivität der weiteren Nutzung des Anforderungsprofils wird verbessert, da keine nachträgliche – und damit für Verfälschungen anfällige – Interpretation der Anforderungen stattfindet (Koch et al., 2006, 87).“

Die Methode der CIT wurde auch aus jener Überlegung als geeignetste Methode ausgewählt, dass dieser methodische Zugang dem „Arbeitsplatz“ Studium in seiner Komplexität und Breite gerecht wird. Die Befragung von StelleninhaberInnen (Studierende) und Vorgesetzten (Lehrpersonen) anhand teilstandardisierter Befragungstechniken lässt die Auswahl und Beschreibung der Situationen den befragten Personen offen und ermöglicht dadurch ein möglichst die gesamte Vielfalt der praktischen Ausprägungen abdeckendes weites Spektrum an kritischen Situationen zu erfassen; dies nicht nur bezogen auf die Anforderungen der einzelnen Studienfächer sondern auch auf die im Studienalltag zu bewerkstellenden Herausforderungen (z.B. organisatorische Bewältigung, Mangel an Seminarplätzen).

Als Erhebungsinstrumente dienen üblicherweise Fragebögen, Interviews oder Workshops, die für das jeweilige Berufsfeld eigens entwickelt werden. Es gibt mehrere Varianten, wie die Critical Incident Technique durchgeführt werden kann. Je nach Forschungsfeld werden die Kriterien der Erhebung und Erhebungsmaterialien adaptiert.

“From the foregoing discussion, it is clear that the critical incident technique is essentially a procedure for gathering certain important facts concerning behavior in defined situations. It

should be emphasized that the critical incident technique does not consist of a single rigid set of rules governing such data collection. Rather it should be thought of as a flexible set of principles which must be modified and adapted to meet the specific situation at hand (Flanagan, 1954, 335)."

3.7 Material

Für die Erhebung der critical incidents wurden zweierlei Erhebungsmaterialien entwickelt, um die Vorteile zweier Erhebungstechniken zu nützen:

- Fragebogen für StelleninhaberInnen (Studierende)
- Interviewleitfaden für Experteninterviews mit Vorgesetzten (Lehrpersonen)

3.7.1 Fragebogen

Die StelleninhaberInnen (Studierende) der sieben Studienrichtungen wurden mittels Fragebogen befragt. Dafür wurden eigens Fragebögen nach der Critical Incident Technique (Flanagan, 1954) konzipiert.

Für die Umfrage unter den Studierenden eignete sich dieses Erhebungsinstrument, weil damit eine größere Anzahl von Studierenden erreicht werden konnte.

Der Umfang des Fragebogens beträgt vier Seiten. Exemplarisch sind hier die Fragen aus dem Fragebogen für Studierende der Publizistik- und Kommunikationswissenschaft dargestellt (siehe Box 1). Die Studierenden aller Studienrichtungen wurden mit denselben Fragen befragt; lediglich die Bezeichnung des Studiums wurde geändert. Der komplette Fragebogen für die Studienrichtung der Publizistik- und Kommunikationswissenschaft befindet sich im Anhang.

Im Fragebogen wurden die deskriptiven Daten zur Person (Alter und Geschlecht) und der belegte Studienabschnitt getrennt nach den Studienplänen (alter, neuer Studienplan) erhoben.

Nach den Daten zur Person folgte ein standardisierter Einleitungstext. Dieser erläutert das Ziel der Untersuchung mit dem Hinweis darauf, dass die Eignung zum Studium (Studienerfolg) und nicht eine an das Studium anschließende berufliche Eignung erhoben wird. Ebenso wurde spezifiziert, dass sich die Fragen auf eine spezifische Studienrichtung und nicht auf das Studium im Allgemeinen beziehen.

Die Einstiegsfrage des Fragebogens verfolgte den Zweck, die Auseinandersetzung mit den Fragen nach der CIT zu erleichtern, indem auf einem allgemeinen Abstraktionsniveau die

Studierenden dazu angeregt wurden, über die Kennzeichen eines/r erfolgreichen Studierenden nachzudenken.

Die darauf folgenden Fragen bauen auf der Methode der CIT auf. Die erste Frage bezieht sich auf die Nennung erfolgskritischer Situationen, die zweite auf die Nennung erfolgsförderlicher Verhaltensweisen zu den zuvor genannten kritischen Situationen, und die dritte auf die Nennung erfolgshinderlicher Verhaltensweisen. Um stereotype (mit dem Studium im Allgemeinen verbundene) Situationsnennungen zu vermeiden, wurde nach Situationen gefragt, welche die Studierenden selbst erlebt oder selbst beobachtet haben.

Es wurde den Studierenden offen gelassen, wie viele erfolgskritische Situationen sie nennen. Insgesamt wurden ihnen 16 freie Zeilen (Antwortfelder) zur Beschreibung der Situationen zu Verfügung gestellt.

Box 1: Fragebogen für Studierende nach der CIT

Standardisierter Einleitungstext:

Die folgenden 4 Fragen dienen der Erstellung eines Anforderungsprofils für Studierende Publizistik- und Kommunikationswissenschaft an der Universität Wien. Sie beziehen sich nicht auf beruflichen Erfolg bzw. berufliche Eignung, sondern immer auf den Studienerfolg bzw. Eignung zum Studium. Bedenken Sie, welche Eigenheiten das Studium der PKW hat und welche Hürden dabei genommen werden müssen.

Frage 1 – Einstiegsfrage

Beschreiben Sie eine/n erfolgreiche/n Studierende/n der Publizistik- und Kommunikationswissenschaft. Was zeichnet sie/ihn aus?

Frage 2 - kritische Situationen

Erinnern Sie sich bitte an typische und wichtige Situationen im Studium der Publizistik und Kommunikationswissenschaften, die Sie selbst erlebt oder beobachtet haben und deren erfolgreiche Bewältigung für den Fortgang des Studiums wichtig ist.

Welche Situationen waren/sind das? (Situationen können auch Prüfungen oder Lehrveranstaltungen sein.)

Bitte gehen Sie konkret auf die Situation ein und beschreiben Sie diese genau.

Tragen Sie bitte jede Situation in eine neue Zeile in der Tabelle ein. Wenn Ihnen noch mehr Situationen einfallen, schreiben Sie einfach unter der Tabelle weiter.

Frage 3 – Erfolgsförderliche Verhaltensweisen

Welche Verhaltensweisen sollte man in diesen Situationen zeigen, um sie erfolgreich zu bewältigen?

Berichten Sie aus Ihrer eigenen Erfahrung und/oder denken Sie an eine KollegIn, die Sie für sehr geeignet für das Studium der Publizistik und Kommunikationswissenschaften halten und die diese Situationen erfolgreich bewältigt hat.

(Bitte geben Sie dabei keine Eigenschaften an wie z.B. „ist ehrgeizig“, sondern Verhaltensweisen wie z.B.: „arbeitet fleißig mit“.)

Tragen Sie bitte für jede oben genannte Situation die dazugehörigen erfolgreichen Verhaltensweisen in die entsprechend nummerierte Zeile in der Tabelle ein.

Frage 4 – Erfolgshinderliche Verhaltensweisen

Welche Verhaltensweisen sollte man in diesen Situationen auf keinen Fall zeigen?

Berichten Sie aus Ihrer eigenen Erfahrung und/oder denken Sie an eine KollegIn, die Sie für ungeeignet für das Studium der Publizistik und Kommunikationswissenschaften halten (die das Studium eventuell sogar abgebrochen hat).

(Bitte geben Sie dabei wieder keine Eigenschaften an wie z.B. „ist nicht ehrgeizig“, sondern Verhaltensweisen wie z.B.: „arbeitet nicht fleißig mit“).

Tragen Sie bitte für jede oben genannte Situation die dazugehörigen nicht erfolgreichen Verhaltensweisen in die entsprechend nummerierte Zeile in der Tabelle ein.

3.7.2 Prätest

Vor der Fragebogenerhebung wurde ein Prätest durchgeführt. Der Fragebogen wurde fünf Studierenden vorgelegt mit der Aufgabenstellung den Fragebogen auf Verständlichkeit der Fragen und Begrifflichkeiten zu beurteilen. Anhand der rückgemeldeten Kritikpunkte wurde der Fragebogen adaptiert.

Die Änderungen bezogen sich auf die Ergänzung der ersten Frage um ein Antwortfeld, die Erläuterung was unter „Situation“ beispielhaft verstanden werden könnte, die geschlechterneutrale Schreibweise sowie auf die Hervorhebung der zentralen Aussagen mittels Markierung „unterstrichen“ und „fett“.

Durch die Hervorhebungen wurde leichter erkennbar, dass sich die Fragen drei und vier hinsichtlich ihrer Ausprägung „positiv“ und „negativ“ unterscheiden und, dass sich die Verhaltensnennungen von Frage 3 und 4 auf die in Frage 2 genannten Situationsnennungen beziehen.

Trotz Hervorhebung der Instruktion, dass in Frage 3 und 4 keine Eigenschaften erhoben werden, zeigten sich in der Untersuchung dennoch vereinzelt Schwierigkeiten bei der Differenzierung in Verhaltensweisen und Eigenschaften. Von einem geringfügigen Teil der befragten Personen wurden Eigenschaften angegeben. Die Eigenschaftsnennungen wurden in weiterer Folge nicht ausgewertet.

3.7.3 Interviewleitfaden

Die Lehrpersonen wurden mittels halbstrukturierter Interviews mit Interviewleitfaden befragt. Der Befragung der „Vorgesetzten“ liegt die Annahme zugrunde, dass aufgrund des anderen Blickwinkels sowohl andere Situationen als auch dieselben Situationen in einer anderen

Häufigkeit genannt werden wie von den Studierenden selbst. Als Untersuchungsmethode wurde das Experteninterview in Ergänzung zur Fragebogenerhebung mit den Studierenden gewählt, um den Vorteil des detaillierteren Nachfragens nutzen zu können. Nach Reimann ist der Vorteil teilstandardisierter Verfahren, dass „spezifischere Fragestellungen untersucht werden und zusätzlich breit gestreute Informationen erlangt werden (2005, 118).“

Der Interviewleitfaden richtete sich nach dem inhaltlichen Aufbau der CIT. Im Unterschied zum Fragebogen konnte der Vorteil des detaillierteren Nachfragens zu den einzelnen Situationen genutzt werden. Es wurde ausschließlich nach Ereignissen gefragt, die die InterviewpartnerInnen selbst erlebt oder beobachtet haben. Von Personen abstrahierte Situationsbeschreibungen wurden nicht erfragt, um analog zu der Fragebogenerhebung stereotype (mit dem Studium im Allgemeinen verbundene) Situationsnennungen zu vermeiden. Hypothetische, in der Zukunft liegende Situationen wurden erst am Ende des Interviews erfragt.

“Usually the interviewer should apply certain criteria to the incidents while they are being collected. Some of the more important criteria are: (a) is the actual behavior reported; (b) was it observed by the reporter; (c) were all relevant factors in the situation given; (d) has the observer made a definite judgment regarding the criticalness of the behavior; (e) has the observer made it clear just why he believes the behavior was critical (Flanagan, 1954, 342).”

Die Entwicklung des Interviewleitfadens orientierte sich an den oben genannten Kriterien von Flanagan (1954). Die zentralen Fragen des Interviewleitfadens nach der CIT werden in Box 2 exemplarisch dargestellt und anschließend erläutert. Der vollständige Interviewleitfaden (beispielhaft für die Studienrichtung Volkswirtschaftslehre) befindet sich im Anhang.

Box 2: Interviewleitfaden für Lehrende nach der CIT

Einstiegsfragen:

Wann ist Ihrer Meinung nach ein Studierender im Studium „x“ erfolgreich?

Woran erkennen Sie erfolgreiche / wenig erfolgreiche Studierende bzw. was unterscheidet einen erfolgreichen und einen wenig erfolgreichen Studierenden? (Welche Verhaltensweisen machen sie/ihn erfolgreich / wenig erfolgreich?)

Beispielfrage CIT:

Können Sie sich an das letzte Mal erinnern, als ein (von Ihnen betreuter) Studierender etwas geleistet oder getan hat, das Sie besonders gut fanden?

Fragenkatalog zur detaillierten Situationsbeschreibung:

- Was war das für eine Situation / unter welchen Umständen kam es dazu?
- Was war das Besondere an dieser Situation?
- Welche Aufgabe genau erhielt der Studierende in dieser Situation?

- Was hat dieser Studierende getan, welches Verhalten hat er/sie gezeigt, das zu dieser außerordentlichen Leistung führte?
- Wann im Semester/im Laufe der Lehrveranstaltung war dieses Ereignis genau?
- Er/sie war Studierende/r in welchem Semester / Studienabschnitt?
- Welches Verhalten sollte ein/e Studierende/r in einer solchen Situation auf keinen Fall zeigen bzw. wie hätte sich ein/e wenig erfolgreiche/r Studierende/r verhalten?

Zusatzfrage zur Änderung der Studienpläne:

Denken Sie, es wird in Zukunft noch andere typische und wichtige Situationen / Umstände geben, in denen sich zeigt, ob ein Studierender erfolgreich oder wenig erfolgreich ist? Denken Sie dabei an mögliche Änderungen im Studienplan (z.B. Bakkalaureat und Master).

Um den Einstieg in die Fragen nach der CIT zu erleichtern, wurden zu Beginn des Interviews zwei Einstiegsfragen gestellt, die den InterviewpartnerInnen die Auseinandersetzung und thematische Eindenken in das Thema erleichtern sollten.

Die daran anschließende Interviewfrage nach der CIT fragte nach jenem erfolgskritischen Ereignis, an welches sich der/die InterviewpartnerIn in der jüngsten Vergangenheit erinnern konnte, mit dem Zweck die Erinnerungsgenauigkeit der Beschreibung zu erhöhen. Nach Frebort und Kubinger sind „eher kürzlich zurückliegende Beobachtungen üblicherweise leichter und genauer erinnerbar (2006, 441).“ Die Aktualität der kritischen Situationen war auch in Hinblick auf die sich ändernden Studienpläne von Bedeutung.

Der/die InterviewerIn erhielt die Instruktion erst mit den Fragen des Leitfadens fortzufahren, wenn der/die InterviewpartnerIn eine konkrete Situation vor Augen hatte. Daran anschließend folgte ein Fragenkatalog zur detaillierten Situationsbeschreibung. Dieser Fragenpool wiederholte sich (aufgrund der Redundanz der Fragen war eine Variation der sprachlichen Formulierung der Fragen erwünscht) in der Interviewsituation solange, bis die Lehrperson keine weiteren erfolgskritischen Ereignisse mehr nennen konnte.

An den österreichischen Universitäten werden aufgrund der Ratifizierung des Bologna-Vertrages 1999 in einem noch andauernden Prozess Änderungen der Studienpläne (Umstellung auf Bakkalaureats- und Masterstudiengänge) durchgeführt (Bundesministerium für Wissenschaft und Forschung, 2009). Um Auswirkungen dieser Änderungen auf die in Zukunft notwendigen Anforderungen abschätzen zu können, wurden die Lehrpersonen darum gebeten, sich auch zukünftige Situationen vorzustellen, die sich durch die Änderung der Studienpläne ergeben könnten. In jenen Studienrichtungen, in denen bereits Bachelor- und Masterstudiengänge eingerichtet wurden, wurde spezifisch nachgefragt, ob Veränderungen hinsichtlich der geforderten Leistungen erkennbar geworden sind.

Die Zusatzfrage zur Änderung der Studienpläne diente in erster Linie der Qualitätssicherstellung der erhobenen Daten unter Berücksichtigung des

Veränderungsprozesses an den österreichischen Hochschulen. Die in dieser Frage genannten Situationen wurden daraufhin analysiert, ob Änderungen der Anforderungsstruktur zu erwarten sind.

Am Ende des Interviewleitfadens wurde eine Checkliste für den/die Interviewerin integriert, um die Vollständigkeit der Daten vor Beendigung des Interviews zu überprüfen (siehe Box 3).

Box 3: Checkliste für den/die InterviewerIn

- Wurden die Verhaltensweisen zu den entsprechenden Situationen berichtet?
- Wurden die Verhaltensweisen vom interviewten Experten/der interviewten Expertin selbst beobachtet?
- Wurden alle relevanten Faktoren der Situationen beschrieben?
- Hat der/die interviewte Experte/Expertin klar definiert, welches die erfolgreiche und welches die wenig erfolgreiche Verhaltensweise ist?
- Hat der/die interviewte Experte/Expertin erklärt, warum welches Verhalten erfolgreich oder wenig erfolgreich ist?

3.8 Durchführung

In der Anforderungsanalyse wurden zwei Zielgruppen - Studierende (n = 208) sowie Lehrpersonal (n = 71) - in den Studienrichtungen Architektur, (Internationale) Betriebswirtschaftslehre, Maschinenbau, Publizistik- und Kommunikationswissenschaft, Rechtswissenschaften, Soziologie und Volkswirtschaftslehre an drei Wiener Universitäten befragt. In Abb. 2 sind die Studienrichtungen und die jeweilige Anzahl der befragten StudentInnen und Lehrpersonen aufgeschlüsselt.

Die Befragung erfolgte schriftlich und mündlich. Die Studierenden wurden mittels Fragebogen befragt. Mit den Lehrpersonen wurden halbstrukturierte Experteninterviews mit Interviewleitfaden durchgeführt. In diesem Unterkapitel wird auf die Vorgehensweise bei der Erhebung und die Erhebungsstichproben eingegangen.

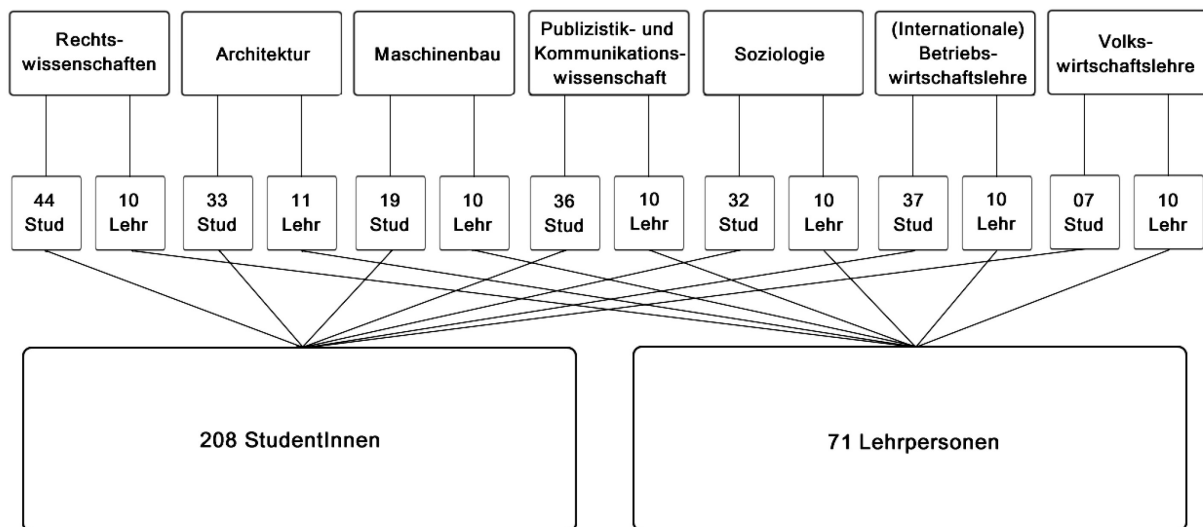


Abb. 2 Stichprobe
 (Quelle: Sample, Khorramdel, Maurer & Kubinger, 2008, Übersetzung und Reduktion um die Studienrichtung Psychologie, Anm.d.A.)

3.8.1 Fragebogenerhebung Studierende

Die Erhebung erfolgte vor Ort an den Fakultäten und Instituten der Universitäten. Damit sollte sichergestellt werden, dass Studierende erreicht werden, die aktiv in den Studienalltag eingebunden sind. Unsicherheiten in Bezug auf die weitere Verwendung der Ergebnisse – mögliche Verwendung des Anforderungsprofils zur Generierung eines Studierenden-Auswahlverfahrens – konnten dadurch minimiert und die Motivation zum sorgfältigen Ausfüllen des Fragebogens erhöht werden (Hinweis auf den Beratungs- und Förderansatz der Studie).

Das Verständnis der Instruktion ist essentiell für die Qualität der Fragebogenerhebung (Flanagan, 1954). Das persönliche Austeilen der Fragebögen ermöglichte das Klären von Verständnisschwierigkeiten anhand einer standardisierten Zusatzinformation. Verständnisschwierigkeiten gab es trotz Beispielvorgaben und Hervorhebungen wiederholt beim Begriff der „typischen und wichtigen Situationen“. Standardisierte Zusatzinstruktion: „Bitte lesen Sie die Instruktion sorgfältig durch. Die im Text genannten Beispiele bieten Hilfestellung zur jeweiligen Aufgabenstellung.“

Die Fragebögen wurden an Studierende ausgegeben, die anhand eines externen Kriteriums als erfolgreich im Studienfortschritt betrachtet werden können. Als Kriterium galt mindestens Absolvierung des ersten Studienabschnittes nach dem alten Studienplan bzw. mindestens zweites Studienjahr nach dem neuen Studienplan. Der Zugang zu den Noten der Studierenden an den verschiedenen Instituten war aufgrund des Datenschutzes nicht gegeben. Das Außenkriterium der Noten konnte daher nicht herangezogen werden. Eine andere Möglichkeit wäre gewesen, AbsolventInnen zu befragen, die das Studium bereits

positiv abgeschlossen haben. Gegen dieses Kriterium sprach jedoch, dass nach Abschluss des Studiums die kritischen Situationen im Studium im Rückblick verzerrt wahrgenommen werden könnten und die Aktualität der Studiensituation nicht mehr gegeben wäre.

Es wurden 208 Studierende in den Studienrichtungen Architektur, (Internationale) Betriebswirtschaftslehre, Maschinenbau, Publizistik- und Kommunikationswissenschaft, Rechtswissenschaften, Soziologie und Volkswirtschaftslehre befragt. Zu jeder Studienrichtung sollten mindestens zwanzig Fragebögen gesammelt werden. In der Studienrichtung Maschinenbau (n=19) wurden vier Fragebögen aufgrund unvollständiger Angaben ausgeschlossen. Die größte Anzahl gibt es in der Studienrichtung Rechtswissenschaften (n=44), gefolgt von (Internationale) Betriebswirtschaftslehre (n=37) und Publizistik- und Kommunikationswissenschaft (n=36). Die kleinste Stichprobe ist die der Studienrichtung Volkswirtschaftslehre. Hier wurden lediglich sieben Fragebögen gesammelt. Die Studienrichtung Volkswirtschaftslehre wurde erst am Ende des Forschungsprozesses mit aufgenommen, um das konventionelle Cluster (C) zu ergänzen. Aufgrund des geplanten zeitlichen Abschlusses der Studie und bestehender Ferienzeiten am Institut musste auf eine größere Stichprobe in der Volkswirtschaftslehre verzichtet werden. Das Cluster C ist jedoch durch die Studierendenbefragung der Studienrichtung Rechtswissenschaften vertreten.

187 StudentInnen erfüllten das Außenkriterium des fortgeschrittenen Studienabschnittes (mind. 2. Abschnitt/Studienjahr) (siehe Tab. 3). Die Stichprobe von 21 Studierenden des ersten Abschnittes/Studienjahres wurde mit der Intention befragt, mögliche Schwierigkeiten, die insbesondere während der Phase des Studienbeginns auftreten können, mit zu erheben. In dieser Stichprobe zeigte sich eine stärkere Konzentration auf kritische Situationen in Zusammenhang mit der Anmeldung zu Seminaren und „Knock-out“-Prüfungen. In der Interpretation der Daten zeigten sich jedoch keine Unterschiede bezüglich der notwendigen Anforderungen zum Studium und die Stichproben wurden gemeinsam ausgewertet.

Das durchschnittliche Alter der befragten Personen beträgt 23,6 Jahre. Die Studierenden der Studienrichtung Maschinenbau sind die Stichprobe mit dem höchsten Durchschnittsalter (26,1), die Studienrichtung Rechtswissenschaften die jüngste Stichprobe (22,5).

Studienrichtung	n	Studienplan		Abschnitt		Alter
		a	n	1.	2.	
Architektur	33	10	23	1	32	23,6
Maschinenbau	19	16	3	4	15	26,1
Publizistik	36	9	27	1	35	23,7
Rechtswissenschaften	44	29	15	6	38	22,5
Soziologie	32	3	29	0	32	24
VWL	7	7	0	0	7	24
BWL/IBWL	37	28	9	9	28	23,1
Gesamt	208	102	106	21	187	23,6

Tab. 3 Deskriptive Statistik Fragebogenerhebung - Studienplan und Alter

Von den 208 befragten StudentInnen waren 113 weiblich (entspricht 55,4%) und 91 männlich (4 Fragebögen ohne Angabe zum Geschlecht) (siehe Tab. 4). Das entspricht der geschlechtsspezifischen Verteilung der Grundgesamtheit der Studierenden in Österreich. (Anzahl Studierende an Österreichischen Hochschulen insgesamt: Frauen 142.000 (entspricht 52,3%) und Männer 129.503) (Statistik Austria, 2008, Daten aus 2007/08).

Studienrichtung	Geschlecht		
	m	w	k.A.
Architektur	18	13	2
Maschinenbau	18	1	0
Publizistik	9	27	0
Rechtswissenschaften	12	30	2
Soziologie	12	20	0
VWL	5	2	0
BWL/IBWL	17	20	0
Gesamt	91	113	4

Tab. 4 Deskriptive Statistik Fragebogenerhebung – Geschlecht

3.8.2 Experteninterviews Lehrpersonen

In der Anforderungsanalyse wurden 71 Lehrpersonen aus sieben Studienrichtungen mittels halbstrukturierter Interviews befragt. Pro Studienrichtung wurden mindestens zehn Lehrpersonen aus verschiedenen Fachbereichen interviewt. Die Lehrpersonen wurden via Email darum gebeten, an der Studie teilzunehmen. Ein Auswahlkriterium beim Versand der Emails war das Lesen mindestens einer Lehrveranstaltung im zweiten Abschnitt.

Die durchschnittliche Dauer der Lehr- und Forschungstätigkeit der befragten Personen beträgt 17,8 Jahre. Von den befragten ExpertInnen sind 38 Personen UniversitätsprofessorInnen, 7 AssistenzprofessorInnen, 19 UniversitätsassistentInnen, 4 externe UniversitätslektorInnen, zwei emeritierte Universitätsprofessoren in aktiver Lehre und ein Tutor.

Von den 71 Lehrpersonen waren 52 männlich und 19 weiblich. Dies liegt mit 26,7% Frauenanteil knapp unterhalb des Durchschnitts der geschlechtsspezifischen Verteilung des Lehrpersonals an den drei befragten Universitäten. (Anteil der Frauen in Lehre an der Universität Wien 2004: 30,36%, Wirtschaftsuniversität Wien 2007: 38,73%, Technische Universität Wien 2006: 16,6%) (Referat Frauenförderung und Gleichstellung der Universität Wien, 2004; Parnigoni, 2008; Ratzer, 2006, entn. aus Abb. 12, p. 28).

Die Länge der Interviews betrug durchschnittlich 44 Minuten. Die Interviews wurden digital aufgezeichnet und vollständig transkribiert, ehe die Daten für die weitere Interpretation in einem nächsten Dokumentationsschritt in Exceltabellen aufbereitet wurden.

3.9 Auswertung

Die Daten aus den Fragebögen und Interviewtranskripten wurden für jede Studienrichtung getrennt nach Stichproben (LehrveranstaltungsleiterInnen und Studierende) in Exceltabellen übertragen. Grundlage für die Dokumentation und Gruppierung des Materials in Auswertungseinheiten war die Qualitative Inhaltsanalyse nach Mayring (2008). Folgende Schritte wurden bei der Analyse der Daten durchgeführt (Tab. 5).

Datenauswertung		
Schritt 1	Dokumentation	Übertragung der Daten in Exceltabellen >> Pool von Situationen (Studierende n = 876, Lehrende n = 377) und Verhaltensweisen
Schritt 2	Kategorienbildung	Gruppierung ähnlicher Situationen zu Situationsklassen (Kategorien n = 62) Beispiele: Prüfungsvorbereitung, Prüfungssituation schriftlich, Präsentationen, Gruppenarbeiten, negative Beurteilungen, Abgabefristen etc.
Schritt 3	Interpretation	Kodierung der Verhaltensweisen nach Konstruktzugehörigkeit (Konstruktschema als Heuristik) >> Induktives Schließen von den Verhaltensweisen im situationalen Kontext auf die zugrunde liegenden Anforderungen.

Tab. 5 Vorgangsweise bei der qualitativen Datenanalyse

3.9.2 Schritt 1: Dokumentation

Theoretische Basis für die Dokumentation und Kategorisierung der erhobenen Daten sind die Anforderungsanalysen von Koch et al. (2006) und Frebort und Kubinger (2006) unter Anwendung der qualitativen Inhaltsanalyse nach Mayring (2008).

Die Daten aus den Fragebögen und Interviewtranskripten wurden in 14 Excelmasken übertragen (7 Studienrichtungen x 2 Erhebungsarten). Die Regeln für die Dokumentation sind durch die Frageeinheiten der Critical Incident Technique vorab festgelegt. Wie in Tab. 6 (Auszug aus einer Tabelle der Studienrichtung Soziologie) dargestellt, enthalten die Tabellen spaltenweise

- die deskriptiven Daten,
- die critical incidents,
- die dazugehörigen erfolgsförderlichen Verhaltensweisen und
- die erfolgshinderlichen Verhaltensweisen.

Die Verhaltensweisen wurden auf separate Zeilen aufgeteilt, um die anschließende Kodierung und Interpretation zu erleichtern. Die Excelmasken der Stichprobe Lehrende enthielten zusätzlich eine Spalte für die Einstiegsfragen, die Frage zur Hochschulreform und ein optionales freies Feld für Anmerkungen zur Interviewsituation oder für ergänzende Informationen (z.B. allgemeine Aussagen zur Studiensituation in Österreich).

FB-Nr.:	Abschnitt	Dipl. od. Diss.	Studienplan	sex	Alter	kritische Situation	günstige Verhaltensweisen - konkrete Nennungen	ungünstige Verhaltensweisen - konkrete Nennungen
414	3. Jahr	-	n	w	25	Bakkalaureatsarbeit	<p>Viel Literatur lesen, bis man von dem groben Thema zur eigentlichen Forschungsfrage kommt.</p> <p>Gute Fragestellung entwickeln. Reflektieren, was ist eine gute Fragestellung</p> <p>Roter Faden durch die Bakk-Arbeit wichtig. Auch die Fragen vom Anfang am Ende beantworten</p> <p>Möglichst genaues Exposee erstellen</p> <p>sich konsequent hinsetzen. Nicht zu lange Pausen machen, sonst hat man das Gefühl, wieder von vorne anzufangen</p>	<p>Zusammenstückeln von Literatur ohne eine eigene Argumentation und Schreibweise zu finden</p> <p>Exposé nicht absprechen mit Betreuer</p> <p>Falsch oder nachlässig zitieren</p> <p>Zu spät beginnen - gibt eine Frist.</p>

Tab. 6 Darstellung einer CIT-Situationsnennung der Studienrichtung Soziologie

3.9.3 Schritt 2: Kategorienbildung

Von den 208 Studierenden wurden insgesamt 876 Situationen und von den 71 Lehrpersonen 377 erfolgskritische Situationen genannt. Im Durchschnitt nannte somit jede/r StudentIn 4,2 und jede Lehrperson 5,3 erfolgskritische Situationen.

Nach der Übertragung in Excelmasken wurden die „critical incidents“ in Kategorien (Situationsklassen) gruppiert. Grundlage für das Kategoriensystem war jenes von Frebort und Kubinger (2006) für die Studienrichtung Psychologie. Das daraus konzipierte Kategoriensystem umfasst 62 inhaltlich verschiedene Klassen von Situationen (z.B. Prüfungsvorbereitung, Prüfungssituation, Präsentationen, Gruppenarbeiten, negative Beurteilungen etc.). Die insgesamt 1253 Situationsnennungen wurden den 62 Situationsklassen zugeordnet.

3.9.4 Schritt 3: Interpretation

Nach der Zuordnung der kritischen Situationen zu Situationsklassen, wird von den dazugehörigen erfolgsförderlichen und –hinderlichen Verhaltensweisen auf die zugrunde liegenden Anforderungen geschlossen. Die Interpretation wurde von zwei unabhängigen Beurteilerinnen vorgenommen.

Um bei diesem Interpretationsschritt theorie- und regelgeleitet vorzugehen, wurde ein Konstruktschema erstellt, welches als Heuristik bei der Zuordnung der Anforderungen zu Konstruktklassen diene. Das Konstruktschema ist eine Erweiterung des Konstruktschemas von Frebort und Kubinger (2006), das seinerseits eine Erweiterung des Konstruktschemas von Westhoff (2005) darstellt (siehe Abb.3).

Ein Konstrukt ist „ein gedankliches Konzept, das aus Überlegungen und Erfahrungen abgeleitet worden ist, um beobachtbares Verhalten zu erklären, z.B. Intelligenz, Angst (DIN 33430, B.28, 21, zit.n. Westhoff, 2005, 128).“ Ein Konstruktschema fasst die Fülle von Anforderungen zu wenigen Klassen von Anforderungen zusammen. Dadurch bietet es sowohl eine Strukturierungshilfe bei der Einordnung von Anforderungen, als auch eine Möglichkeit zur Überprüfung, ob alle Aspekte physischer und psychologischer Anforderungen im Interpretationsprozess berücksichtigt wurden (Westhoff, 2005).

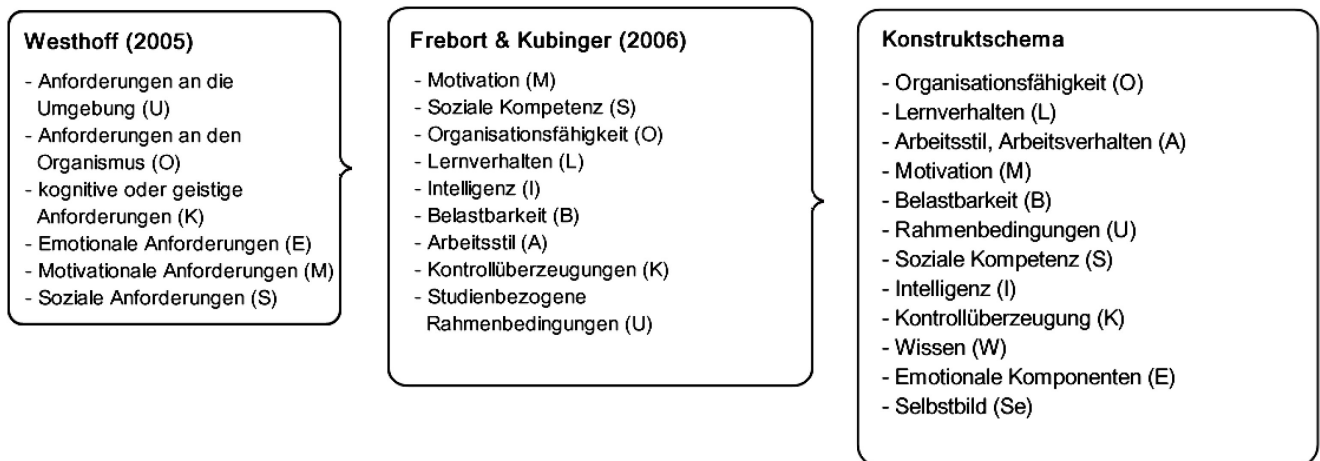


Abb. 3 Konstruktschemata und Ausdifferenzierung
(Quelle: Khorramdel, Maurer & Kubinger, 2008, Übersetzung aus dem Englischen, Anm.d.A.)

Um alle ermittelten Anforderungen zuordnen zu können, bedurfte das Konstruktschema von Frebort und Kubinger (2006) einer zusätzlichen Differenzierung. Das resultierende Kategorienschema ist in Abb. 3 dargestellt, und enthält die zwölf Konstruktklassen:

- Arbeitsstil und Arbeitshaltungen (A)
- Belastbarkeit (B),
- Emotionale Komponenten (E)

- Intelligenz (I)
- Kontrollüberzeugung (K)
- Lernverhalten (L)
- Motivation (M)
- Motorik (Mk)
- Organisationsfähigkeit (O)
- Soziale Kompetenz (S)
- Selbstbild (Se)
- Rahmenbedingungen (U)
- Wissen (W)

Das erweiterte Konstruktschema wurde von zwei weiteren MitarbeiterInnen des Fachbereichs für Psychologische Diagnostik auf Vollständigkeit und Sinnhaftigkeit geprüft.

In der Critical Incident Technique wird von dem wahrnehmbaren Verhalten induktiv auf die latenten Anforderungen geschlossen. Die induktiv ermittelten Anforderungen wurden den zwölf Konstruktklassen anhand einer Kodierung (A, B, E, I, K, L, M, Mk, O, S, Se, U, W) zugeordnet. Beispielsweise zählen die Anforderungen *Gewissenhaftigkeit* und *Selbstständiges Arbeiten* zur Konstruktklasse Arbeitshaltungen (A), *Zeitmanagement* und *Informiertheit* zur Konstruktklasse Organisationstalent (O), *Frustrationstoleranz* und *Arbeiten unter Zeitdruck* zur Konstruktklasse Belastbarkeit (B) etc. Exemplarisch sei der Interpretationsschritt in Tab. 7 für eine Situations- und Verhaltensnennung der Studienrichtung Soziologie dargestellt.

Die Interpretation wurde von zwei unabhängigen Beurteilerinnen vorgenommen. Zur Qualitätssicherung wurde eine zweifache Kontrollschleife eingebaut. Jede Beurteilerin überprüfte die Interpretationsschritte der anderen und meldete die Änderungen rück. Die Änderungen wurden ihrerseits von der ursprünglichen Beurteilerin angenommen oder bei Divergenz die ursprüngliche Zuordnung argumentiert. Dadurch wurden verschiedene Blickwinkel berücksichtigt und gleichzeitig ein gemeinsames Vorgehen bei der Interpretation gewährleistet.

Je Verhaltensweise waren Mehrfachnennungen von Anforderungen möglich. Beispielsweise erfordert die Verhaltensweise „mit anderen Studierenden austauschen und lernen“ sowohl die Anforderungen *Kontaktfähigkeit*, *Kommunikationsfähigkeit* und *Lernfähigkeit*.

Es wurden insgesamt 76 Anforderungen ermittelt und den zwölf Konstruktklassen (A, B, E, I, K, L, M, Mk, O, S, SE, U, W) zugeordnet. Die Arbeitsdefinitionen der Anforderungen ergaben sich aus den sie kennzeichnenden Verhaltensweisen. Wenn eine neue eine Anforderung

eindeutig kennzeichnende Verhaltensweise interpretiert wurde (keine Verhaltensweise mit Mehrfachnennung), wurde sie der Arbeitsdefinition hinzugefügt. (siehe Unterkapitel «Methode»).

Situation	günstige Verhaltensweisen - konkrete Nennungen	A/B/ E//K /L/M/ Mk/O /S/Se /U/W	daraus geschlossene Eigenschaft	ungünstige Verhaltensweisen - konkrete Nennungen	A/B/ E//K /L/M/ Mk/O /S/Se /U/W	daraus geschlossene Eigenschaft
Bakkalaureatsarbeit	Viel Literatur lesen, bis man von dem groben Thema zur eigentlichen Forschungsfrage kommt.	A, M	Gewissenhaftigkeit, Engagement	Zusammenstückeln von Literatur ohne eine eigene Argumentation und Schreibweise zu finden	I	Verbale Intelligenz
	Gute Fragestellung entwickeln. Reflektieren, was ist eine gute Fragestellung	I, I	schlussfolgerndes Denken, kritisches Denken	Exposé nicht absprechen mit Betreuer	S	Kommunikationsfähigkeit
	Roter Faden durch die Bakk-Arbeit wichtig. Auch die Fragen vom Anfang am Ende beantworten	I, I	Verbale Intelligenz, schlußfolgerndes Denken	Falsch oder nachlässig zitieren	AI	Gewissenhaftigkeit, Verbale Intelligenz
	Möglichst genaues Exposé erstellen	A, A	genaues Arbeiten, Gewissenhaftigkeit	Zu spät beginnen - gibt eine Frist.	O	Zeitmanagement
	sich konsequent hinsetzen. Nicht zu lange Pausen machen, sonst hat man das Gefühl, wieder von vorne anzufangen	M	Durchhaltevermögen			

Tab. 7 Darstellung des induktiven Interpretationsschluss bei der Ermittlung der Anforderungen (Studienrichtung Soziologie)

Auswertung Zusatzfrage: Änderung der Studienpläne

Im Experteninterview wurde eine Zusatzfrage gestellt, welche sich auf die aktuell stattfindenden Änderungen der Studienpläne (Umsetzung des Bologna Prozesses) bezieht. Diese diente in erster Linie der Qualitätssicherstellung der erhobenen Daten unter Berücksichtigung des Veränderungsprozesses an den österreichischen Hochschulen. Die in dieser Frage genannten Situationen wurden daraufhin analysiert, ob Änderungen der Anforderungsstruktur zu erwarten sind. Sie wurde unabhängig von den übrigen Fragen nach der Critical Incident Technique ausgewertet.

In den Studienrichtungen (Internationale) Betriebswirtschaftslehre, Architektur, Publizistik- und Kommunikationswissenschaft, Soziologie und Volkswirtschaftslehre) wurden die Bachelor-Studiengänge bereits eingeführt. Hier konnten die Lehrpersonen direkt über ihre Erfahrungen mit den neuesten Studienjahren sprechen. Den Lehrpersonen der Rechtswissenschaft fiel es schwer, sich hypothetische zukünftige Situationen vorzustellen.

Sie beantworteten die Frage dahingehend, ob sie Änderungen in den Anforderungen, die an die Studierenden gestellt werden, erwarten. Wirkliche Änderungen in den Anforderungen, seien nicht sichtbar. Es könnte hingegen Unterschiede in der Gewichtung der Anforderungen geben.

So werde beobachtet, dass die Konkurrenz und die „Ellbogentechnik“ unter den Studierenden zunehme. Insgesamt gehe die Tendenz stärker in Richtung schnelles, effizientes Studieren und Ergebnisorientierung.

Kritische Anmerkungen zu den Änderungen im Rahmen des Bologna Prozesses bezogen sich einerseits auf den vorzeitigen Bachelor-Abschluss nach drei Jahren (bisherige Mindeststudiedauer: 8-10 Semester) und die damit verbundene befürchtete mangelnde Berufsvorbildung. Zusätzlich werde der zeitliche Druck auf die Studierenden erhöht, weil versucht wird, denselben Umfang in einem kürzeren Zeitraum zu vermitteln.

Weitere Bedenken bezogen sich auf die „Verschulung“ des Universitätssystems. Bisher sei es notwendig gewesen, den gesamten Studienverlauf größtenteils eigenständig zu planen und zu organisieren. Durch die Verschulung gehe die Selbstständigkeit der Studierenden bis zu einem gewissen Grad verloren. Von einzelnen Lehrpersonen wurden Zweifel angesprochen, ob die bisherige zentrale Anforderung des kritischen Denkens bei jungen Menschen in Zukunft weiterhin derart gefördert werden könne, wie sie in einer Universitätsstruktur, die sehr starke Eigenverantwortung den Studierenden überlässt, möglich war.

Box 4: Beispielauszug Interview (Lehrperson, Studienrichtung Architektur)

„Seit 2004 gibt es den Bakkalaureatsstudienplan. Es gibt eine striktere Trennung und eine große Verschulung. Für manche hat das Vorteile, aber im Bereich der Architektur stellt sich die Sinnhaftigkeitsfrage. Die Schule hat sicher Vorteile für viele Dinge, aber auf der Universität ist es ein wenig Fehl am Platz. Es geht nicht darum hinzugehen, etwas eingetrichtert zu bekommen, zu repetieren, sondern es geht ums eigenständige Denken. Nicht nur etwas zu wiederholen, sondern etwas zu verstehen und damit weiterzuarbeiten. Das geht dadurch verloren, durch den ganzen Zeitdruck. Die ganzen Studierenden arbeiten nebenher. Es gibt viele, die würden in den Übungen mehr machen, müssen aber aus sehr pragmatischen Gründen darin zurückstecken.“

In den Studiengängen, in denen der Bachelor- und Masterstudienplan bereits umgesetzt wurde, kann davon ausgegangen werden, dass die Anforderungsprofile keiner Änderung bedürfen. Die Fragen der Critical Incident Technique bezogen sich auf jene Situationen, an die sie sich die Lehrpersonen als letztes erinnern konnten, wodurch die Aktualität sichergestellt werden konnte. Lediglich die kritischen Situationen am Ende des Studiums

(z.B. Diplomarbeit, Diplomprüfung, Wahlfächer) bezogen sich noch auf Studierende der alten Studiengänge.

Das Anforderungsprofil der Studienrichtung Rechtswissenschaften wurde nach den oben genannten möglichen Änderungen überprüft und mit den Anforderungsprofilen der anderen Studiengänge verglichen. Es konnten keine Auswirkungen zukünftiger Entwicklungen auf das Anforderungsprofil abgeschätzt werden.

3.10 Ergebnis Anforderungsanalyse

Nach Benennung und Kodierung der 76 Anforderungen wurden die Häufigkeiten der Anforderungen je Studienrichtung ermittelt und in eine Rangordnung gesetzt. Daraus ergaben sich sieben Anforderungsprofile für die Studienrichtungen (Internationale) Betriebswirtschaftslehre, Maschinenbau, Publizistik- und Kommunikationswissenschaft, Rechtswissenschaften, Soziologie und Volkswirtschaftslehre, die sich hinsichtlich der Gewichtung der Anforderungen unterscheiden.

Inhaltlich gab es große Überschneidungen zwischen den Anforderungsprofilen der einzelnen Studienrichtungen. Einzig im Studium der Architektur wurde eine Anforderung ermittelt, die in keiner anderen Studienrichtungen eine Nennung erbrachte (Feinmotorik (Mk)). Unterschiede gab es zwischen den Anforderungsprofilen insbesondere in der Gewichtung der Anforderungen, nicht jedoch hinsichtlich ihrer Ausprägung (durchschnittlich, über- oder unterdurchschnittlich).

Die Anforderungen der sieben Studienrichtungen wurden zu einem gemeinsamen Anforderungsprofil über die allgemeine Studierfähigkeit zusammengefasst. Dazu wurden die Häufigkeitszählungen der Anforderungen addiert, um eine Gewichtung über alle Studienrichtungen hinweg zu erzielen.

Das zusammengefasste Anforderungsprofil wurde um das Anforderungsprofil für Psychologie von Frebort und Kubinger (2006) ergänzt. Das Anforderungsprofil für Psychologie wurde von Frebort und Kubinger (2006) standortspezifisch für Wien entwickelt und baut ebenso auf der Methode der Critical Incident Technique nach Flanagan (1954) auf. Die Häufigkeitszählungen der Merkmalsnennungen wurden für die Zusammenfassung der Anforderungsprofile von Frebort und Kubinger (2006) zur Verfügung gestellt.

Ursprünglich war geplant, auch das Anforderungsprofil der Studienrichtung Medizin, welches die Basis für den in der Schweiz und in Österreich durchgeführten Eignungstest für das Medizinstudium (EMS-Test, Zentrum für Testentwicklung und Diagnostik, 1995) darstellt, mit in das Anforderungsprofil zur allgemeinen Studierfähigkeit aufzunehmen. Aufgrund der

fehlenden Kenntnis über die Häufigkeitszählungen pro Anforderung, konnte das Anforderungsprofil für Medizin jedoch nicht integriert werden. Es gab keine Information darüber, wie das Anforderungsprofil in die gewichtete Rangreihe des hier entwickelten Anforderungsprofiles zur allgemeinen Studierfähigkeit übernommen werden könnte.

Ebenso wurde davon Abstand genommen, die Anforderungsprofile der Studienrichtungen Anglistik (Schüpbach, Pixner & Zapf, 2005), Informatik, Elektrotechnik und technische Informatik (Zimmerhofer & Hornke, 2005) und Wirtschaftswissenschaften (Hell, Ptok & Schuler, 2007) in ein standortspezifisches Anforderungsprofil für Wien zu übernehmen. Die Anforderungsprofile wurden in Deutschland entwickelt und die Übertragbarkeit der Anforderungsprofile auf diesen Standort konnte aufgrund der Unterschiede in den Lehrplänen nicht gestützt werden.

Das Anforderungsprofil zur allgemeinen Studierfähigkeit setzt nicht den Anspruch, spezifische Anforderungsprofile einzelner Studienrichtungen zu ersetzen. Es geht hier vielmehr um eine Sammlung zentraler Anforderungen, die sich in mehreren inhaltlich divergierenden Studienrichtungen als notwendig erwiesen haben. Auf dieser Basis ist es möglich, StudienbeginnerInnen hinsichtlich ihrer Stärken und Schwächen in Bezug auf die Herausforderungen des Studiums zu beraten. Ebenso erlaubt es allgemeine Fragestellungen zur Studienreife beantworten zu können.

Der Rückschluss auf einzelne Studienrichtungen ist jedoch als problematisch zu werten. Einzelne Studienrichtungen können sowohl hinsichtlich der Benennung und Gewichtung der Anforderungen von diesem allgemeinen Anforderungsprofil abweichen. Beispielsweise ist es wahrscheinlich, dass im Rahmen von Linguistikstudiengängen, die Anforderungen *verbale Intelligenz* und das *Wissen um eine Fremdsprache* von höherer Bedeutsamkeit sind. Die Übernahme des allgemeinen Anforderungsprofils ist in Abhängigkeit der Fragestellung jeweils kritisch zu prüfen. In Zusammenhang mit der Auswahl von Studierenden für eine Studienrichtung ist eine spezifischere Analyse unerlässlich.

Das Ergebnis ist ein standortspezifisches Anforderungsprofil zur allgemeinen Studierfähigkeit in Wien. Das Anforderungsprofil setzt sich aus 12 Konstrukten zusammen, denen 75 Merkmale zugeordnet wurden. Das Anforderungsprofil inklusive Häufigkeitszählung ist in Tab. 14 im Anhang zu entnehmen. Das Anforderungsprofil zur allgemeinen Studierfähigkeit beinhaltet folgende Konstrukte und Anforderungen:

Das Konstrukt Motivation (M) setzt sich aus den 15 Anforderungen *Durchhaltevermögen, Lernbereitschaft, Engagement, Aktive Problembearbeitung, Interesse Zielorientierung, Zielsetzung, Wissensdurst/Neugier, Ehrgeiz, Gestaltungsmotivation, Anpassungsfähigkeit,*

Erfolgszuversicht, Wettbewerbsorientierung, Schwierigkeitspräferenz und Leistungsstolz zusammen.

Organisationsfähigkeit (O) beinhaltet die fünf Merkmale *Zeitmanagement, Prioritäten setzen, Planen, Organisationstalent* und *Informiertheit*.

Das Konstrukt Intelligenz (I) besteht aus 13 Anforderungen: *Schlussfolgerndes Denken, Kritisches Denken, verbale Intelligenz, numerische Intelligenz, Flexibilität im Denken, Raumvorstellung, Kreativität, figurale Fähigkeiten, praktisch-technisches Verständnis, technisches Verständnis, Umsetzung/Anwendung von Erlerntem, physikalisches Grundverständnis* und *Informationsverarbeitungsgeschwindigkeit*.

Das Konstrukt Arbeitsstil und Arbeitsverhalten (A) ist zusammengesetzt aus acht Anforderungen: *Gewissenhaftigkeit, Reflexivität, genaues Arbeiten, realistisches Einschätzen der eigenen Leistungsfähigkeit / Fähigkeiten, realistisches Anspruchsniveau, hohes Anspruchsniveau, Eigeninitiative* und *selbstständiges Arbeiten*.

Zur Sozialen Kompetenz (S) zählen zehn Merkmale: *Durchsetzungsstärke, Überzeugungsstärke, Autoritätshörigkeit, Kommunikationsfähigkeit, Teamfähigkeit, Teamorientierung, Kollegialität, Kompromissbereitschaft, Kontaktfähigkeit* und *Sensitivität*.

Das Konstrukt Lernverhalten besteht aus vier Merkmalen: *Lernstil, Lernfähigkeit, Merkfähigkeit* und *Konzentration bzw. zielgerichtete Aufmerksamkeit*.

Belastbarkeit (B) setzt sich aus sechs Anforderungen zusammen: *Frustrationstoleranz, Irritierbarkeit und Nervosität bei Leistungserbringung vor anderen, Arbeiten (bzw. Belastbarkeit) unter Zeitdruck, Bewältigung von Mehrfachanforderungen, Fähigkeit zur Entspannung* und *Belastbarkeit und Flexibilität bei Hindernissen*.

Das Konstrukt Emotionalität (E) beinhaltet die sechs Merkmale *Distanzierungsfähigkeit, Offenheit, Standhaftigkeit, Extraversion vs. Introversion, Selbstständigkeit* und *Ehrlichkeit*.

Das Konstrukt Wissen (W) gliedert sich in vier Merkmale: *fachspezifisches Wissen, Fremdsprachenkenntnisse, Vorwissen* und *Allgemeinwissen*.

Das Konstrukt Selbstbild (Se) beinhaltet die Anforderung *Selbstbewusstsein*.

Zum Konstrukt Rahmenbedingungen (U) zählen sowohl *studienbezogene* als auch *umweltbezogene Rahmenbedingungen*.

Das Konstrukt Kontrollüberzeugungen (K) beinhaltet das Merkmal *Internalität*.

Das Anforderungsprofil zur allgemeinen Studierfähigkeit bildet die Grundlage für die Konstruktion des Wiener Studieneignungs-Persönlichkeitsinventars (WSP). Vor der Beschreibung der Verfahrensentwicklung wird im anschließenden Kapitel auf die Entwicklung des ‚Regelkataloges zur Fragebogen-Itemkonstruktion‘ eingegangen, welcher der Konstruktion des Itempools des WSP zugrunde liegt.

4 Regelkatalog zur Fragebogen-Itemkonstruktion

4.1 Einleitung

Ziel der Diplomarbeit ist die Erstellung eines Itempools für einen Persönlichkeitsfragebogen zur Erfassung studienrelevanter Eigenschaften. Die zentrale Fragestellung neben der Entwicklung des Itempools ist die Qualitätssicherung des Erstellungsprozesses anhand einer ausführlichen Dokumentation der Vorgehensweise und die Entwicklung eines Instrumentariums, welches Handlungsanleitungen und Regeln für die Itementwicklung beschreibt.

Es gibt eine nahezu unbegrenzte Möglichkeit der Konstruktion und Formulierung von Items, denn sowohl die Art der Fragestellung (Itemstamm) als auch das Antwortformat unterliegen vielfältigen Überlegungen. Um diese Überlegungen zu bündeln wird ein standardisiertes Regelwerk entwickelt, welches die theoretischen Überlegungen zur Itemkonstruktion und die durch den empirischen Erfahrungsgewinn gewonnenen Erkenntnisse in einem ‚Regelkatalog zur Fragebogen-Itemkonstruktion‘ zusammenfasst. Am Ende dieses Kapitels wird ein Katalog präsentiert, der Regeln und Formulierungshinweise zur Itemgenerierung enthält, um die Qualität der Items durch Objektivierung der Vorgehensweise zu sichern.

Das Vorhaben der Erstellung eines ‚Regelkataloges zur Itemkonstruktion‘ ist nicht neu. Edwards (1957, zit.n. Mummendey, 2008) stellte bereits vor gut fünfzig Jahren vierzehn Regeln für die Formulierung von Items auf. Ein Teil davon hat auch heute noch Gültigkeit. In letzter Zeit gingen die Bemühungen eher in die Richtung, Konstruktionsrationale zu entwickeln, welche es ermöglichen Items regelgeleitet anhand der Variation einzelner Strukturmerkmale zu konstruieren (Reif, 2008; Sonnleitner, 2007).

Diese Diplomarbeit setzt den Schwerpunkt nicht auf Konstruktionsrationale, die für jede Merkmalsdimension spezifisch entwickelt werden müssen und daher nicht beliebig auf andere neu zu konzipierende Persönlichkeitsfragebögen angewandt werden können. Es stellt vielmehr den Versuch dar, die zentralen Regeln für die Fragebogen-Itemgenerierung aus Literatur und empirischer Erfahrung zu sammeln und damit ein praktisch-anwendbares Instrument zur Qualitätssteigerung von (Persönlichkeitsfragebogen)-Items im Allgemeinen zur Verfügung zu stellen.

Vorab wird auf die Problematik bei der Itementwicklung eingegangen. Was ist an der Itementwicklung derart kritisch, dass ihr sowohl in Lehrbüchern als auch in der empirischen

Forschung so viel Aufmerksamkeit entgegen gebracht wird? Forschungsergebnisse geben über die Auswirkungen mangelhafter Items auf Testergebnisse Aufschluss.

Daran anschließend wird die bestehende Literatur zur Itementwicklung und Fragebogenkonstruktion gesichtet. Dies bezieht sich sowohl auf die quantifizierbaren Kriterien zur Überprüfung der Itemqualität, als auch die Wahl des Antwortformates und nicht zuletzt auf die sprachliche Gestaltung des Itemstamms.

Zusätzlich zur Literaturrecherche wird ein Experteninterview mit zwei Mitarbeitern des Arbeitsbereichs für Diagnostik am Institut für Entwicklungspsychologie und Diagnostik (Fakultät für Psychologie, Universität Wien) geführt. Die wesentlichsten Aussagen aus der Recherche und dem Experteninterview, d.h. jene die auf eine Vielzahl von Itemarten übertragen werden können, finden sich in Form von Regeln im ‚Regelkatalog zur Fragebogen-Itemkonstruktion‘ wieder.

4.2 Problematik Itemkonstruktion

Die Güte eines Fragebogens ist zu einem Großteil durch die ihm zugrunde liegenden Items bestimmt. Becker (2003) nennt vier Arten von Fehlerquellen, die sich auf die Güte des Fragebogens auswirken können:

- Fehlerquelle im Medium der Sprache
- Fehlerquelle im Fragebogen selbst
- Fehlerquelle auf Seiten der Testperson
- Fehlerquelle in der Untersuchungssituation

Für die Itemkonstruktion wird besonders der Fehlerquelle im „Medium Sprache“ Relevanz zugewiesen. Ergebnisse über die Variation im Antwortverhalten abhängig von der sprachlichen Formulierung wurde von Schwarz (1999) zusammengetragen. „Self-reports of behavior and attitudes are strongly influenced by features of the research instrument, including question wording, format and context (Schwarz, 1999, 93).“ Zentrales Element ist die Verständlichkeit der Items (Strack & Martin, 1984, zit.n. Schwarz, 1999). Versteht die Testperson unter dem Item dasselbe wie der/die TestkonstrukteurIn? Dies ist insbesondere in Hinblick auf die Mehrdeutigkeit des Iteminhalts und die mangelnde Präzision (z.B. durch die Verwendung unklarer Häufigkeitsangaben „häufig“ „manchmal“ „selten“) zu berücksichtigen.

„Suppose that respondents are asked how frequently they felt „really irritated“ recently. To provide an informative answer, respondents have to determine what the researcher means with “really irritated”. Does this term refer to major or minor annoyances? To identify the intended

meaning of the question, they may consult the response alternatives provided by the researcher. If the response alternatives present low-frequency categories, for example, ranging from “less than once a year” to “more than once a month”, respondents may conclude that the researcher has relatively rare events in mind (Schwarz, 1999, 95).”

Auch nach Konsultation der ‚response options‘ ist jedoch nicht eindeutig klar, was die Testkonstrukteure unter dem Begriff der Irritation verstehen. Neben der semantischen Nachvollziehbarkeit von Items können sich Fehler im Medium der Sprache auch in Verständnisprobleme bei der Verwendung von Fachausdrücken und Fremdwörtern, doppelten Verneinungen oder langen verschachtelten Sätzen äußern (Becker, 2003; Bühner, 2006; Jonkisz & Moosbrugger, 2007).

Beispielitem für eine doppelte Verneinung (Bühner, 2006, 67):

„Ich würde mich nicht als jemanden bezeichnen, der nicht traurig ist.“

Trifft nicht zu trifft eher nicht zu trifft eher zu trifft zu

Den Einfluss der Angemessenheit von Items auf die Validität von Persönlichkeitsfragebögen wurde von Horlacher und Rodewald (2002) untersucht. Unter Angemessenheit wird dabei verstanden, ob die Iteminhalte oder das Antwortformat dem subjektiven Erleben der befragten Personen entsprechen. Ihre Annahme war, dass Personen bei subjektiv unpassenden Items gar nicht in der Lage sind, Informationen zu liefern, die im Sinne der Messintention liegen. Obwohl keine eindeutigen Rückschlüsse aus den Ergebnissen gezogen werden konnten, konnten sie dennoch Hinweise finden, wonach als unpassend empfundene Items eine deutlich geringere Validität aufwiesen als angemessen eingeschätzte Items.

Eine Fehlerquelle auf Seiten der Testperson, die vorab bei der Itemkonstruktion und der Wahl des Antwortformats zu berücksichtigen ist, sind die Antworttendenzen (response sets). Untersuchungen über Antworttendenzen in Abhängigkeit von der Gestaltung der Antwortformate werden im Unterkapitel «Antworttendenzen» vorgestellt. An dieser Stelle sei darauf hingewiesen, dass die Wahl des Antwortformates sowohl auf die Gütekriterien der Reliabilität und Validität, die Eindimensionalität und die Zumutbarkeit von Fragebögen haben (Borkenau & Amelang, 1985; Hambros, 2002; Karner, 1999 & 2002; McKelvie, 1978; Rost, 2002).

4.3 Methode

Um einen Regelkatalog zu generieren, welcher Handlungsanleitungen und Regeln für die Fragebogen-Itemerstellung enthalten soll, bietet sich eine Vielzahl von Quellen an, die zur Konstruktion herangezogen werden können:

- Forschungsergebnisse aus Studien zur Itemqualität
- Literatur zur Test- und Fragebogenkonstruktion
- Experteninterviews mit TestkonstrukteurInnen
- Empirische Erfahrungen während der Testkonstruktion

Einerseits befasst sich eine große Anzahl von Studien mit der Qualität von Items und deren Auswirkungen auf das Antwortverhalten der befragten Personen. Andererseits gibt es eine umfangreiche Literatur zur Test- und Fragebogenkonstruktion. Darüber hinaus wird die Itemqualität auch in Lehrbüchern zur diagnostischen Psychologie thematisiert.

Neben der Literaturrecherche bietet es sich an, ExpertInnen aus dem Feld, d.h. Personen, die selbst bereits Tests und Fragebögen entwickelt haben, zu ihren Erfahrungen während der Itementwicklung zu befragen. Ergänzend zu den a-priori zusammengetragenen Regeln, können die eigenen Erfahrungen im empirischen Prozess der Fragebogenkonstruktion dokumentiert werden und in den Regelkatalog einfließen.

Für die Entwicklung des Regelkataloges wurden alle vier Quellen herangezogen und werden auf den nächsten Seiten diskutiert. Die Ergänzungen des Regelkataloges aus der vierten Quelle – empirische Erfahrungen während der Konstruktion des Itempools des WSP – werden im Unterkapitel «Probleme bei der Itementwicklung und Ergänzung des Regelkataloges» dargestellt.

Abgesehen von der Frage nach der theoretischen Fundierung der Regeln stellt sich die Frage, nach welchen Kriterien Aussagen zur Test- und Fragebogenkonstruktion in den Regelkatalog aufgenommen werden sollen. Der Anspruch des Regelkataloges als leicht handhabbares Instrument impliziert eine gute Übersichtlichkeit und daher auch eine gewisse Kürze und Prägnanz der Aussagen und -menge. Als Kriterien für die Aufnahme von Regeln in den Regelkatalog wurde daher festgelegt, dass es sich bei den Aussagen um Regeln handelt, die sich spezifisch auf Persönlichkeitsfragebögen – im Gegensatz zu anderen Verfahrensarten – beziehen. Zweitens sollen die Regeln insofern allgemeiner Natur sein, dass sie für eine Vielzahl von Itemarten übertragen werden können, wobei sich der Regelkatalog auf verbale Items bezieht. Drittens, die Regeln wurden in der Literatur bzw. von ExpertInnen hinreichend argumentiert und stellen nach heutigem Wissensstand aktuelle Kenntnisse zur Fragebogenkonstruktion dar.

Die Regeln, welche am Ende des Kapitels in Form des Regelkataloges gesammelt dargestellt werden, sind – um die Nachvollziehbarkeit zu gewährleisten – in der folgenden Literaturrecherche jeweils im Text separat vermerkt und bereits mit jener Nummer gekennzeichnet, mit welcher sie im Regelkatalog beziffert sind.

Auch wenn der Regelkatalog Empfehlungen zur Itemkonstruktion im Allgemeinen enthält, ist er dennoch eng mit der Vorgehensweise der Gesamtarbeit verknüpft. Es ist daher möglich, dass der Regelkatalog gewisser Adaptierungen bedarf, wenn er aus dem hier vorliegenden methodischen Vorgehen herausgelöst wird. Der Großteil der Regeln sollte jedoch unabhängig von der gewählten Itemart Gültigkeit besitzen.

Ergänzend sei darauf hingewiesen, dass der ‚Regelkatalog zur Fragebogen-Itemkonstruktion‘ lediglich ein Instrument zur Qualitätssicherung neben anderen darstellt. Weitere Instrumente zur Qualitätssicherung werden im Unterkapitel «Instrumente zur Qualitätssicherung» vorgestellt.

4.3 Quantifizierbare Kriterien zur Überprüfung der Itemqualität

Damit ein Test oder ein Persönlichkeitsfragebogen die Gütekriterien der Objektivität, Reliabilität und Validität erfüllen, müssen die Testaufgaben selbst diesen Gütekriterien genügen. Neben der Objektivität, Reliabilität und Validität von Items lässt sich auch ihr Schwierigkeitsgrad quantifizieren.

4.3.1 Objektivität von Items

„Eine Aufgabe ist dann objektiv, wenn ihre Beantwortung von verschiedenen Beurteilern übereinstimmend als richtig oder falsch, als kennzeichnend für das Vorhandensein oder Fehlen bzw. einen stärkeren oder schwächeren Ausprägungsgrad des untersuchten Persönlichkeitsmerkmals gewertet wird (Lienert & Raatz, 1998, 29).“ Bei der Itemkonstruktion ist somit darauf zu achten, dass die Eindeutigkeit der Messrichtung bei der Interpretation der Aufgabe durch verschiedene BeurteilerInnen gegeben ist. (Regel 13 des Regelkataloges) Unter Messrichtung wird bei Persönlichkeitsfragebögen nicht von richtig oder falsch gesprochen sondern von ‚im Sinne der Erwartung (in Schlüsselrichtung)‘ bzw. ‚gegen die Erwartung (gegen die Schlüsselrichtung)‘. Die Eindeutigkeit der Messrichtung eines Items (Itempolung), ist insbesondere bei der Verrechnung der Items ausschlaggebend.

4.3.2 Reliabilität von Items

Die Reliabilität von Items kann ebenso wie bei den Verfahren selbst mittels Retest-Reliabilität oder Paralleltest-Reliabilität überprüft werden. Im ersteren Fall werden dieselben Items nach einem angemessenen Zeitintervall der Person nochmals zur Beantwortung vorgegeben. Voraussetzung dafür ist, dass das Item einen situationsüberdauernde Disposition (trait) und keinen momentanen Zustand (state) erhebt. Ein Item ist dann retest-reliabel, wenn die befragte Person die Frage bei der wiederholten Vorgabe in derselben Weise beantwortet wie bei der erstmaligen Vorgabe. Die Paralleltest-Reliabilität ist dann gegeben, wenn ein Item in der Testserie A auf die gleiche Weise beantwortet wird, wie ein streng vergleichbares Item in der Testserie B (Lienert & Raatz, 1998).

4.3.3 Validität von Items

„Eine Aufgabe ist dann valide, wenn sie von dem Pb mit starker Merkmalsausprägung häufiger im Sinne der Erwartung beantwortet wird als von dem Pb mit geringer Merkmalsausprägung (Lienert & Raatz, 1998, 30).“ Die Validität von Aufgaben wird meist am Gesamtpunktwert des Tests ermittelt (Trennschärfe). Die Höhe der Korrelation einer Aufgabe in einer Skala (z.B. Durchhaltevermögen) mit dem Gesamtwert, den die Person in dieser Skala erreicht, wird ermittelt. Dabei wird die Validität der Gesamtskala vorausgesetzt.

4.3.4 Itemschwierigkeit

Ein weiteres wichtiges Kriterium ist die Schwierigkeit einer Testaufgabe. Die Begriffe Leichtigkeit und Schwierigkeit von Items sind statistische Begriffe und gelten nicht nur für Leistungstestitems, sondern drücken prinzipiell die Wahrscheinlichkeit aus, mit der auf ein dichotomes Item eine mit 1 kodierte Antwort gegeben wird, d.h. im Sinne der Erwartung (in Schlüsselrichtung) (Rost, 2004).

Zur Feststellung interindividueller Unterschiede zwischen Personen sind Aufgaben erforderlich, die von unterschiedlichen Personen unterschiedliche Ausprägungen in den einzelnen Merkmalen erzeugen. Die Itemformulierung muss daher entsprechend der Varianz des Antwortverhaltens gewählt werden. Extrem schwer bzw. extrem leicht zu bejahende Fragen bringen wenig bis keinen Informationsgewinn (Jonkisz & Moosbrugger, 2007). (Regel 14) Die Entscheidung über die Varianz der Itemschwierigkeiten ist in Abhängigkeit von der definierten Zielgruppe zu treffen. Beispielsweise könnte die Zielgruppe StudentInnen, die

bereits erste Erfahrungen im Studienalltag gesammelt haben, ein Item als sehr leicht empfinden, während es in der Zielgruppe der MatruantInnen wesentlich stärker differenziert.

4.3.5 Eindimensionalität von Items

Grundsätzlich müssen die Testaufgaben dem Anspruch der psychologischen Testtheorie entsprechen. Eine Testtheorie beschreibt als statistisches Modell, „wie aufgrund des Testergebnisses auf ein diesem Ergebnis zugrunde liegendes Persönlichkeitsmerkmal (hypothetisches Konstrukt) geschlossen werden kann (Moosbrugger & Rauch, 2005, 182).“

Die probabilistische Testtheorie (Item-Response-Theorie) im Sinne des (dichotom) logistischen Testmodells von Rasch (1960) setzt Eindimensionalität der Testitems voraus. D.h. die Items erfassen eine einzige Eigenschaftsdimension. (Regel 8) Wenn ein Item in einem kognitiven Leistungstest anhand zweier verschiedener Fähigkeiten gelöst werden kann, wird mit diesem Item nicht mehr nur ein einziges Konstrukt erfasst. Das Testergebnis ist nicht mehr eindeutig interpretierbar (Kubinger & Proyer, 2005; Kubinger, 2006). Dasselbe gilt für Items in Persönlichkeitsfragebögen. Nach Rost (2002) kann die Verletzung der Eindimensionalität sowohl im Item selbst, in der Person, dem Antwortformat oder der latenten Variable liegen.

4.4 Formale Gestaltung von Items

Ein Item setzt sich aus dem Itemstamm und dem Antwortformat zusammen. Je nach Testart und Merkmal kann der Itemstamm aus einer Frage, einer Aussage, einem Bild, einer Geschichte, einer Zeichnung oder einer Rechenaufgabe bestehen. Nach Seiwald (2003) haben Items in Persönlichkeitsfragebögen typischerweise die Form von Fragen zur Selbsteinschätzung oder bestehen aus Statements, zu denen die befragte Person Zustimmung oder Ablehnung ausdrücken kann.

Krauth (1995) strukturiert die Itemgestaltung in Abhängigkeit möglicher Reaktionsklassen. Unter Reaktionsklassen wird die Gestaltung der Antwortkategorien (response options) verstanden.

Bevor auf die sprachliche Gestaltung des Itemstamms eingegangen wird, werden mögliche Antwortformate und ihre Vor- und Nachteile und mögliche Antworttendenzen thematisiert.

4.4.1 Gestaltung der Antwortformate

Prinzipiell hängt der Einsatz der Antwortformate von der Fragestellung ab, die mit dem Verfahren beantwortet werden möchte. Die Antwortformate bieten verschiedene Vor- und Nachteile, die hinsichtlich der Fragestellung abgewogen werden müssen. Die Wahl des Antwortformats hat darüber hinaus weitreichende Implikationen auf die Qualität der Items, wie dies von mehreren empirischen Untersuchungen nachgewiesen wurde. Die Auswirkung des Antwortformats auf die Verletzung der Eindimensionalität von Persönlichkeitsfragebögen wurde beispielsweise von Rost (2002) untersucht. Lechner (2008) überprüfte in seiner Sinnhaftigkeitsprüfung von Multiple Choice Items die Qualität der Distraktoren bei Leistungstests. McKelvie (1978) untersuchte die Reliabilität und Validität in Abhängigkeit des Antwortformats. Ebenso gibt es zahlreiche Untersuchungen über mögliche Antworttendenzen in Abhängigkeit von der Gestaltung der Antwortformate (Borkenau & Amelang, 1985; Dörner 2009; Hambros, 2002; Karner, 1999 & 2002).

Die Wahl des Antwortformats ist von kritischer Bedeutung und soll an dieser Stelle ausführlich diskutiert werden. Vorab werden ausgewählte Antworttendenzen (response sets) vorgestellt, die Auswirkungen auf die Wahl der Antwortformate haben. Anschließend werden die Antwortformate und ihre Vor- und Nachteile vorgestellt.

4.4.1.1 Antworttendenzen

Bei der Formulierung von Items sind mögliche Antworttendenzen (response sets) vorab zu berücksichtigen. Unter Antworttendenzen werden stereotype Reaktionsweisen auf Fragebogenitems (oder Testitems) verstanden (Bortz & Döring, 2003). Die Tendenz zur Mitte, die Tendenz zu Extremurteilen und die Akquieszenz treten besonders häufig bei Antwortformaten im Multiple-Choice-Format (z.B. Ratingskala) auf. Wenn Personen die Items nicht nur abhängig von ihrem Iteminhalt, sondern aufgrund individueller Itemdarbietungsstile beantworten, ist die Eindimensionalität der Items nicht mehr gegeben. Sie messen neben dem zu erhebenden Merkmal ein zweites Artefakt. Daher sind die Überlegungen zur Vermeidung von „response sets“ von zentraler Bedeutung.

4.4.1.1.1 Tendenz zu Extremurteilen

Unter der Tendenz zu Extremurteilen wird jenes Verhalten von Testpersonen verstanden, bei welchem bei der Beantwortung der Items die äußeren Kategorien (z.B. einer Ratingskala) bevorzugt werden (Krauth, 1995).

4.4.1.1.2 Tendenz zur Mitte

Die Tendenz zur Mitte entspricht dem Gegenteil des oben dargestellten Antwortstils. Hier wählen die Testpersonen verstärkt die mittlere, meist neutrale Kategorie und vermeiden Extremurteile (Krauth, 1995). Dies kann unterschiedliche Ursachen haben. Neben der eigentlichen Messintention der neutralen Beurteilung, kann das Wählen der mittleren Kategorie aber auch bedeuten, dass sich die Testpersonen in ihrem Urteil unsicher sind oder den Iteminhalt unpassend finden und daher die Antwort verweigern. Letzteres wurde von Borkenau und Amelang (1985) untersucht. Sie stellten fest, dass bei Items, die von Testpersonen für nicht angemessen empfunden werden, tendenziell häufiger die neutrale Kategorie angegeben wird, sodass in der gesamten Merkmalsausprägung eine weniger extreme Ausprägung vorliegt. Dadurch wird jedoch nicht die eigentliche Messintention erhoben und ist besonders in Bezug auf die Eindimensionalität von Items kritisch.

4.4.1.1.3 Akquieszenz, Zustimmung- oder Ja-Sage Tendenz

Die Akquieszenz (Ja-Sage-Tendenz) bezeichnet jene Tendenz von Personen, unabhängig vom Iteminhalt zustimmend zu antworten. Ein umgekehrt stereotyp ablehnendes Antwortverhalten wird als Nein-Sage-Tendenz bezeichnet. Um der Akquieszenz und der Nein-Sage-Tendenz entgegenzuwirken, empfiehlt sich eine wechselnde Itempolung (Krauth, 1995, Mummendey, 2008). Auf die unterschiedliche Itempolung zur Vermeidung der Akquieszenz wird im Unterkapitel «Itempolung» eingegangen.

4.4.1.1.4 Reihenfolgeeffekte

Unter Reihenfolgeeffekten wird jenes Phänomen beschrieben, bei dem die unmittelbar vorangehenden Items Auswirkungen auf die Beantwortung der Folgeitems haben (Rost, 2004). Damit in Zusammenhang steht der von Thorndike (1920, zit.n. Krauth, 1995) beschriebene Halo-Effekt. Hier wird von einem einzelnen Merkmal einer Person auf weitere Merkmale geschlossen bzw. das einzelne Merkmal (z.B. das Aussehen einer Person) überstrahlt ein anderes (z.B. die Wahrnehmung der Intelligenz ebenjener). Bei psychologisch-diagnostischen Verfahren kann sich dies unterschiedlich äußern. Beispielsweise könnte die Person das Bedürfnis haben, nicht immer dieselbe Kategorie anzukreuzen. Das inhaltliche Beurteilungssystem einer Person kann sich durch vorangegangene Items verschieben. Beispielsweise könnte ein Item mit einer extremen Aussage, ein nachfolgendes mit einer tendenziell weniger extremen Aussage als neutral erscheinen lassen und die Person könnte dadurch eher zustimmen (Krauth, 1995). Bei Leitungstests gilt es zu vermeiden, dass ein Item Lösungshinweise für Folgeitems enthält.

Bühner (2006) empfiehlt für Fragebögen eine zufällige Itemreihenfolge. Der Vorteil einer gleich verteilten Itemvorgabe gegenüber einer nach Subskalen geblockten Itemvorgabe wurde auch von Dörrer (2009) nachgewiesen.

4.4.1.1.5 Positionseffekte

Von den Reihenfolgeeffekten sind die Positionseffekte zu unterscheiden. Sichtbar wird dieser Effekt, wenn Items abhängig von der Position im Fragebogen unterschiedlich beantwortet werden. Einen Einfluss darauf können Verständnisschwierigkeiten der Instruktion zu Beginn der Testung sein oder Ermüdungs- und Motivationseinbußen gegen Ende der Testung (Rost, 2004).

4.4.1.1.6 Reaktanz

Nach der Reaktanztheorie von Brehm (1966) ist der Mensch bestrebt, seine Handlungsfreiheit beizubehalten. Bei Bedrohung dieser Handlungsfreiheit, kann es zu Reaktionen führen, um die Handlungsfreiheit wiederherzustellen.

Diese sogenannten Reaktanzphänomene sind auch in der psychologischen Diagnostik sichtbar und können sich in untypischen oder willkürlichem Antwortverhalten äußern, wenn die Reaktionsmöglichkeiten der Testperson durch die response options eingeschränkt werden (Karner, 1999). Hambros (2002) untersuchte das Reaktanzphänomen in Zusammenhang mit dem dichotomen (zweikategoriellen) Antwortformat. Die Hypothese, nach welcher das dichotome Antwortformat aufgrund der Freiheitseingrenzung zu untypischem Antwortverhalten führt, konnte nicht nachgewiesen werden. Die Testpersonen äußerten während der Bearbeitung des dichotomen Antwortformates jedoch Unmut und Kritik an den begrenzten Antwortmöglichkeiten.

4.4.1.1.7 Soziale Erwünschtheit

Im Gegensatz zu den oben dargestellten Antworttendenzen, die sich auf formale Aspekte der Itemvorgabe beziehen, stellt die Soziale Erwünschtheit eine inhaltliche Antworttendenz dar. Mummendey (2008) beschreibt diese Antworttendenz folgendermaßen:

„Die Person bildet eine Annahme darüber, was entweder für sie selbst oder für den Versuchsleiter oder für die Gemeinschaft schlechthin (beispielsweise für eine bestimmte gesellschaftliche Gruppe oder für den Versuchsleiter als Repräsentanten einer solchen

Gruppe) als positiv bzw. günstig erscheint, und sie bemüht sich somit, im sozial erwünschten Sinne zu antworten – in diesem Falle haben wir es mit einer Antworttendenz der sozialen Erwünschtheit (Social Desirability) zu tun (Mummendey, 2008, 166).“

Items aus Persönlichkeitsfragebögen sind in ihrer großen Mehrzahl entweder sozial erwünscht (hoher Social-Desirability-Wert) oder sozial nicht erwünscht. Sie werden selten als neutral hinsichtlich ihrer sozialen Erwünschtheit eingestuft. Damit in Zusammenhang stehen die Antworttendenzen der Simulation und Dissimulation. Dabei geht es um die bewusste Verfälschung (faking bad) des Fragebogens insbesondere in Bewerbungs- und Testsituationen. Dies hat insbesondere Auswirkungen auf die Einsatzmöglichkeiten des Persönlichkeitsfragebogens. Karner (2002) konnte eine Verfälschungstendenz (faking) des MBTI (Myers-Briggs Typenindikator, Briggs & Briggs Myers, 1995) in Selektionsstichproben nachweisen.

Um die inhaltlichen Antworttendenzen der sozialen Erwünschtheit und der Simulation und Dissimulation bestmöglich zu reduzieren, wird vom Einsatz von Persönlichkeitsfragebögen in Selektionssituationen Abstand genommen. Die Freiwilligkeit der Testsituation und die Aussicht auf eine detaillierte Rückmeldung über die Testergebnisse erhöhen die Bereitschaft zur Offenbarung und reduzieren die Verfälschungstendenz bei Persönlichkeitsfragebögen (Karner, 2002).

4.4.1.2 Antwortformate

Die Antwortformate unterscheiden sich hinsichtlich ihres Strukturierungsgrades. Es kann zwischen Aufgaben mit freiem, gebundenem und atypischem Antwortformat unterschieden werden (Bühner, 2006; Jonkisz & Moosbrugger, 2007; Lienert & Raatz, 1998; Rost, 2004).

4.4.1.2.1 Freies Antwortformat

Bei freien Antwortformaten sind keine Antwortkategorien vorgegeben, d.h. die Person formuliert ihre Antwort frei, wobei die Instruktion eine gewisse Strukturierung vorgibt. Eingesetzt wird dieses Antwortformat insbesondere bei Projektiven Verfahren, zur Erfassung kreativer Leistungen oder bei Assoziationstests. Der Einsatz bei Leistungstests liegt darin begründet, dass das Wissen beim freien Antwortformat reproduziert und die Ratewahrscheinlichkeit dadurch minimiert werden kann. Rost (2004) unterscheidet drei Arten von freien Antwortformaten:

- Keine Vorgabe: Es wird keine Strukturierung vorgegeben, wie dies beispielsweise vermehrt bei projektiven Verfahren Anwendung findet. In Familie-in-Tieren (Brehm-Gräser, 1970) wird die befragte Person gebeten, ihre Familie als Tiere zu zeichnen. Außer der Größe des Blattes und der Stiftauswahl, die zur Verfügung gestellt wird, gibt es keine Einschränkung bei der Aufgabenerfüllung.
- Formale Vorgabe: In der Instruktion erfolgt eine Formvorgabe, z.B. ein Wort, ein Satz etc.
- Lückenvorgabe: Die befragte Person wird gebeten, eine Lücke im oder im Anschluss an den Itemstamm zu ergänzen.

Aufgrund des zeitlichen Aufwandes ist das freie Antwortformat wenig ökonomisch und die Auswertungsobjektivität ist gering. Bei Persönlichkeitsfragebögen ist es insbesondere im qualitativen Bereich sinnvoll, wie z.B. zur Erfassung von Motiven. Aufgrund der Nachteile des freien Antwortformats wird bei Persönlichkeitsfragebögen am häufigsten das gebundene Antwortformat verwendet (Jonkisz & Moosbrugger, 2007).

4.4.1.2.2 Gebundenes Antwortformat

Aufgaben mit gebundenem Antwortformat sind dadurch charakterisiert, dass Antwortkategorien vorgegeben sind, an welche die befragte Person gebunden ist. Die einfachste Einteilung ist jene in Antwortformate mit zwei oder mehr Antwortmöglichkeiten (Krauth 1995). Jonkisz und Moosbrugger (2007) wiederum unterscheiden die Antwortformate nach Ordnungsaufgaben (Zuordnungs- und Umordnungsaufgaben), Auswahlaufgaben (dichotome und Mehrfachwahl-Aufgaben) und Beurteilungsaufgaben (Analogskala und Ratingskalen).

4.4.1.2.2.1 Ordnungsaufgaben

Ordnungsaufgaben zeichnen sich dadurch aus, dass die einzelnen Bestandteile einer Aufgabe umgeordnet oder einander zugeordnet werden. Die Aufgabe gilt als richtig gelöst, wenn die Bestandteile in die richtige Ordnung gebracht werden. Der Einsatz von Umordnungsaufgaben bietet im Leistungsbereich eine Möglichkeit, schlussfolgerndes Denken, Ursache-Wirkungs-Zusammenhänge oder die Abstraktionsfähigkeit der Probanden zu prüfen. Bei der Verwendung von Bildmaterial kann darüber hinaus auf verbales Material verzichtet werden, wenn die Lesefähigkeit nicht als gegeben angenommen werden kann. Die Einsatzmöglichkeiten des Antwortformates sind jedoch beschränkt und finden daher nur für wenige Fragestellungen Anwendung (Jonkisz & Moosbrugger, 2007).

4.4.1.2.2.2 Auswahlaufgaben

Häufiger Anwendung bei Persönlichkeitsfragebögen finden Auswahlaufgaben. Unterschieden werden können Auswahlaufgaben hinsichtlich der Anzahl der gebotenen response options (zwei- oder mehrkategorial). Die Aufgabe der Testperson besteht darin, die zutreffende oder richtige zu identifizieren.

4.4.1.2.2.2.1 Dichotomes Antwortformat

Das zweikategorielle oder dichotome Antwortformat bietet den Testpersonen die Wahlmöglichkeit zwischen zwei response options. Aufgrund der einfachen Gestaltungsweise findet es sehr häufig Anwendung bei Persönlichkeitsfragebögen, wie beispielsweise im FPI-R (Freiburger Persönlichkeitsinventar. Revidierte Fassung. Fahrenberg, Hampel & Selg, 2001) oder im PRF (Deutsche Personality Research Form. Stumpf, Angleitner, Wieck, Jackson & Belloch-Till, 1985). Die befragte Person gibt ihre Zustimmung oder Ablehnung anhand von „richtig/falsch“, „ja/nein“, „trifft zu/trifft nicht zu“ etc. zum Ausdruck

Das dichotome Antwortformat ist sehr ökonomisch in der Instruktion, Bearbeitung und Auswertung. Die Bearbeitungsdauer kann reduziert werden, weil sich die Testperson nicht zwischen einer Vielzahl von response options entscheiden muss (Jonkisz & Moosbrugger, 2007). Gleichzeitig wird aber vermehrt vom dichotomen Antwortformat Abstand genommen.

In Zusammenhang mit dem dichotomen Antwortformat besteht die Annahme, dass die Einschränkung der Reaktionsmöglichkeiten der Testpersonen auf zwei Optionen zu Reaktanzphänomenen führen kann. Nach der Reaktanztheorie von Brehm (1966) ist der Mensch bestrebt, seine Handlungsfreiheit beizubehalten. Bei Bedrohung dieser Handlungsfreiheit, kann es zu Reaktionen führen, um die Handlungsfreiheit wiederherzustellen.

Wie im Unterkapitel «Antworttendenzen» dargestellt, können diese sogenannten Reaktanzphänomene auch in der psychologischen Diagnostik sichtbar werden und sich in untypischen oder willkürlichem Antwortverhalten äußern (Karner, 1999).

Die Vernünftigkeit der Anwendung von dichotomen Antwortformaten wurde von Hambros (2002) untersucht: Eine mangelnde „reasonableness“ des dichotomen Antwortformates konnte zwar nicht nachgewiesen werden, aber bei der Anwendung der Methode des lauten Denkens, konnte bei jenen Testpersonen, welche das Testmaterial mit dichotomen Antwortformat ausfüllten, die kritische Einstellung gegenüber der Einschränkung der Wahlmöglichkeiten festgestellt werden.

Krosnick (1999) stellte darüber hinaus Hinweise auf eine erhöhte Akquieszenz (Ja-Sage-Tendenz) bei dichotomen Antwortformat fest. Die Akquieszenz beschreibt das Verhalten der

Testpersonen, eher zustimmende als ablehnende Antworten zu geben (Krauth, 1995, Mummendey, 2008).

4.4.1.2.2.2 Mehrkategorielles Antwortformat

Zu den Auswahlaufgaben werden neben dem dichotomen Antwortformat auch mehrkategorielle Antwortformate gezählt. Hier können sich die befragten Personen zwischen mehr als zwei response options entscheiden.

Im Unterschied zu den Rating Skalen geben die Personen jedoch nicht ihre Zustimmung oder Ablehnung, sondern identifizieren eine response option als die richtige/zutreffende. Dieser Aufgabentypus wird auch als „Multiple-Choice-Aufgabe“ bezeichnet. Kennzeichnend dafür ist, dass die Identifikation der richtigen Antwort auf Wiedererkennung und nicht auf dem Reproduzieren von Wissen beruht.

Um mehrkategorielle Auswahlaufgaben zu formulieren, ist die Gestaltung der Distraktoren von entscheidender Bedeutung. Die richtigen oder Bestantworten werden vielfach als „Attraktoren“, die übrigen Antwortmöglichkeiten als „Alternativ-Antworten“ oder „Distraktoren“ bezeichnet (Lienert & Raatz, 1998). Je mehr Distraktoren zur Verfügung gestellt werden, desto geringer ist die Ratewahrscheinlichkeit. D.h. jene Wahrscheinlichkeit mit der zufällig die richtige Lösung gefunden wird.

Bei der Gestaltung von Distraktoren müssen mehrere Kriterien beachtet werden. Kriterien für die Güte von Distraktoren sind ihre Attraktivität (Auswahlwahrscheinlichkeit), ihre Ähnlichkeit mit der richtigen Antwortmöglichkeit (Attraktor) und ihre Plausibilität (Jonkisz & Moosbrugger, 2007). Es stellt eine besondere Schwierigkeit dar, Distraktoren zu entwickeln, die gleich attraktiv sind wie die richtige Lösung. Bei der Gestaltung von Fragebogenitems zur Erfassung von Persönlichkeitseigenschaften oder Einstellungen ist insbesondere auf die Exhausivität der Antwortalternativen zu achten. Werden alle sinnhaften Verhaltensweisen durch die Alternativ-Antworten abgedeckt? Ist dies nicht der Fall, kann es dazu kommen, dass die befragte Person keine Antwort identifizieren kann, die für sie zutreffend ist und zu einem Reaktanzphänomen (siehe oben) führen.

Bei Leistungstests ist darüber hinaus die Disjunktheit der Alternativ-Antworten vielfach nicht gegeben. Disjunktheit bedeutet, dass sich die Alternativ-Antworten gegenseitig ausschließen und die Schnittmenge leer ist. Der Großteil der Distraktoren wird jedoch ausgehend vom Attraktor konstruiert. Lechner (2008) beschreibt in seiner Studie zur Sinnhaftigkeitsprüfung von Multiple Choice Items, wie anhand der Fähigkeit zum logischen Schlussfolgern, einzig aufgrund der Betrachtung der Attraktoren und Distraktoren (unter Ausblendung des Itemstamms) die richtige Lösung bei Leistungstests mit einer hohen Wahrscheinlichkeit erkannt werden kann. Dies geschieht aufgrund der sich überschneidenden Schnittmenge von Kriterien, die mehrere Distraktoren mit dem Attraktor gemeinsam haben. Jene Alternativ-

Antwort, deren Schnittmenge mit den anderen Alternativ-Antworten am größten ist, ist am wahrscheinlichsten die richtige Antwort.

Der Vorteil des MC-Formates liegt in der Ökonomie der Auswertung. Die Auswertungsobjektivität ist bei gebundenen Antwortformaten im Allgemeinen ebenso höher als beim freien Antwortformat. Die erhöhte Auswertungsobjektivität kann jedoch zu Lasten der Validität gehen, wenn die Antwortkategorien nicht alle Reaktionsmöglichkeiten ausschöpfen. Problematisch ist die sorgfältige Konstruktion der Distraktoren im Sinne der gleichen Attraktivität und der Exhaustivität (Jonkisz & Moosbrugger, 2007; Rost, 2004).

4.4.1.2.2.3 Beurteilungsaufgaben

Zu den Beurteilungsaufgaben zählen Aufgaben mit den Antwortformaten Ratingskala und Analogskala. Dabei wird von den befragten Personen der Grad der Zustimmung oder Ablehnung einer Aussage ermittelt.

4.4.1.2.2.3.1 Ratingskala

Ratingskalen sind itemunspezifisch, d.h. die Antwortkategorien können für mehrere oder alle Items eines Persönlichkeitsfragebogens übernommen werden. Die Antwortkategorien sind mehrkategorial und stehen in einer Rangordnung (ordinal). Gängige Ratingskalen sind zum Beispiel „trifft nicht zu, trifft eher nicht zu, trifft eher zu, trifft zu“ oder „fast nie, manchmal, oft, fast immer“.

Es kann zwischen unipolaren und bipolaren Ratingskalen unterschieden werden. Während unipolare Ratingskalen nur eine Richtung abdecken (z.B. Häufigkeit: „täglich, 2-3 mal pro Woche, 1 mal pro Woche, 1 mal pro Monat, noch seltener“), decken bipolare Ratingskalen sowohl die zustimmende als auch ablehnende Richtung ab (z.B. „gar nicht, eher nicht, kaum, etwas, sehr stark“) (Jonkisz & Moosbrugger, 2007; Rost, 2004).

Darüber hinaus wird zwischen verbalen und numerischen Ratingskalen unterschieden. Im einen Fall werden die Abstufungen verbal („gar nicht, eher nicht, kaum, etwas sehr stark“) gekennzeichnet, im zweiten Fall stehen Zahlen stellvertretend für die verbalen Verankerungen (1, 2, 3, 4, 5). Während die Richtung der Antwortskalierung (von positiv zu negativ und umgekehrt) keine Auswirkungen auf das Antwortverhalten zeigt, konnte Rammstedt und Krebs (2007) nachweisen, dass der positive Pol mit hohen Zahlen assoziiert wird und der negative Pol mit niedrigen Zahlen. Bei der Verwendung einer numerischen Ratingskala sollte dies Beachtung finden. Es ist auch eine kombinierte Vorgabe (verbal und numerisch) möglich.

Im Vergleich zum dichotomen Antwortformat sind Ratingskalen informativer, weil sie eine größere Abstufung der Zustimmung oder Ablehnung zulassen. Es gilt jedoch abzuwägen, wie differenziert die Antwort erhoben werden möchte, d.h. wie viele Stufen vorgegeben werden. Das hängt einerseits vom kognitiven Differenzierungsgrad der Zielgruppe ab, andererseits auch von der Vermeidung von Antworttendenzen. Um die Tendenz zum extremen Urteil zu reduzieren, ist es sinnvoll die Anzahl der Stufen gering zu halten. Eine drei oder vierstufige Anzahl (im Gegensatz zu sieben Abstufungen) hat sich bewährt. Bei einer ungeraden Anzahl einer Ratingskala (z.B. 3, 5, 7) ist eine mittlere/neutrale Kategorie vorhanden. Dies hat sich mehrfach als ungünstig erwiesen, weil damit nicht immer eine neutrale Position zum Ausdruck gebracht wird, sondern die mittlere Kategorie wird auch verwendet, wenn die Person das Item für unpassend hält und die Antwort verweigert. Damit misst das Item jedoch nicht mehr jene Eigenschaft, die es zu messen vorgibt, was eine Verletzung der Eindimensionalität bewirkt. Das Weglassen der mittleren Kategorie (Verwendung einer geraden Anzahl von Stufen) verhindert darüber hinaus die „Tendenz zur Mitte“ (Rost, 2004).

Das Ratingformat ermöglicht es sehr differenzierte Informationen über die Ausprägung eines Merkmals zu erhalten. Der Vorteil dieses Antwortformates liegt darin, dass es für mehrere Items übernommen werden kann und nicht spezifisch für jedes einzelne Item entwickelt werden muss. Sowohl die Durchführung als auch die Auswertung sind ökonomisch (Bühner, 2006). Zusätzlich werden die Probleme, die bei der Konstruktion von Distraktoren auftreten können, ausgeschaltet.

Die Testpersonen können dieselben Maßstäbe für alle Items übernehmen. Daraus ergibt sich jedoch der Nachteil, dass sich Antworttendenzen während der Bearbeitung des Persönlichkeitsfragebogens herausbilden können. Wie einzelne dieser Antworttendenzen minimiert werden können wurde oben dargestellt.

4.4.1.2.2.3.2 Analogskala

Die Analogskala gehört ebenfalls zu den Beurteilungsaufgaben. Sie unterscheidet sich von der Ratingskala dadurch, dass sie keine konkreten Skalenabstufungen hat, sondern eine kontinuierliche Skala darstellt, auf welcher die Testpersonen an einer beliebigen Stelle ihre Antwort vermerken können. Bei Paper-Pencil-Tests ist dies möglich, indem sie an die Stelle ein Kreuz setzen, bei computer-basierten Tests können die Personen einen Schieberegeler hin- und herschieben. Bei computer-basierten Tests ist es auch möglich, die Tendenz zur Mitte dadurch auszuschalten, dass der Schieberegeler (in der Mitte positioniert) von der Stelle weg bewegt werden muss, damit die Frage als beantwortet gezählt wird.

Die Vorteile entsprechen jener der Ratingskala. McKelvie (1978) erkannte eine Präferenz der Testpersonen für die Analogskala gegenüber stetig abgestuften Antwortskalen. Hirsch (2006) kam zu einem anderen Befund. In ihrer Untersuchung sprechen die Testpersonen eine Präferenz für die fünfstufige verbale Ratingskala aus. Die Analogskala wurde zweitgerichtet. Folgende Antwortformate wurden geboten: dichotom, drei-, fünf- sechskateoriell und analog.

Ein Nachteil der Analogskala ist, dass die Differenziertheit der Messung in der Mehrzahl der Items nicht der Differenziertheit des Urteils entspricht (Jonkisz & Moosbrugger, 2007).

4.4.1.2.3 Aufgaben mit atypischem Antwortformat

Jonkisz und Moosbrugger (2007) benennen neben den oben dargestellten Antwortformaten, atypische Antwortformate, die sich nicht in die oberen Kategorien einordnen lassen. Dazu zählen die Aussagen-Vergleichs-Aufgabe, welche hauptsächlich zur Prüfung des logischen Urteilvermögens Verwendung findet. Ein anderes atypisches Antwortset bietet der „Zahlen-Verbindungstests“ von Oswald und Roth (ZVT, 1987), bei dem die Zahlen 1 bis 8, die beliebig in einem Kästchen gestreut sind, mit einem Stift in der aufsteigenden Reihenfolge verbunden werden müssen. Für die Gestaltung von Fragebogenitems sind sie jedoch von geringerem Belang.

4.4.2 Wahl des Antwortformates

Die empirischen Befunde ermöglichen keine eindeutige Beantwortung der Frage nach dem optimalen Antwortformat. Die Wahl der Antwortformate ist abhängig von der Fragestellung, die mit dem Verfahren beantwortet werden möchten. Es kann daher nur eine Empfehlung basierend auf den oben dargestellten Vor- und Nachteilen von Antwortformaten und möglichen Fehlerquellen aufgrund der Antworttendenzen gegeben werden. Nach Jonkisz und Moosbrugger (2007) bieten sich für Persönlichkeitsfragebögen insbesondere das gebundene Antwortformat an, das eine höhere Auswertungsobjektivität und Ökonomie besitzt als das freie Antwortformat. Aufgrund der Schwierigkeiten bei der Konstruktion von Distraktoren (Jonkisz & Moosbrugger, 2007; Lechner, 2008) wird darüber hinaus das Antwortformat der Beurteilungsaufgaben den Auswahlaufgaben vorgezogen.

Obwohl es bisher keine hinreichenden empirische Forschungsergebnisse für das Auftreten von Reaktanzphänomenen bei dichotomen Antwortformaten gibt, wird empfohlen ein mehrkategorielles oder analoges Antwortformat zu wählen, um die Handlungsfreiheit der

befragten Personen zu erhöhen. Ein zu großes Antwortspektrum (z.B. sechs- oder achtkategoriell) könnte jedoch die befragten Personen überfordern und zu einer zufälligen Beantwortung der Items führen (Karner, 1999). Bühner (2006) spricht davon, dass sich für Persönlichkeitsfragebögen größtenteils ein fünfkategoriales Antwortformat anbiete. Auch McKelvie (1978) konnte in zwei Experimenten nachweisen, dass das fünf-kategoriale Antwortformat gegenüber einem sieben oder elf-kategoriellen Antwortformat eine höhere Reliabilität besitzt.

Der Tendenz zur Mitte kann dadurch Vorschub geleistet werden, dass keine mittlere Kategorie vorgegeben wird, was dafür spricht, ein vier- oder sechskategoriales Antwortformat zu verwenden. Die vierkategoriale Ratingskala wurde weder von McKelvie (1978) noch von Hirsch (2006) untersucht. Der Ausschluss der mittleren Kategorie bedeutet für die Analogskala, dass der Schieberegeler, welcher in seiner Ausgangsposition in der Mitte des Balkens positioniert ist, verschoben werden muss, damit die Frage als beantwortet gilt. Auch die anschließende Prüfung der Items nach dem dichotom logistischen Modell nach Rasch (1960) (probabilistische Testtheorie) legt eine geradzahlige Anzahl der Antwortstufen – aufgrund der notwendigen Dichotomisierung der Antworten – nahe.

Zusammenfassend wird für die Itempoolentwicklung des WSP das Itemformat der Beurteilungsaufgaben gewählt. Es werden Items entwickelt, welche sowohl eine vierstufige Ratingskala oder eine Analogskala ermöglichen. Aufgrund der geplanten Überprüfung des Itempools mit dem dichotom logistischen Modell nach Rasch (1960), welches die Dichotomisierung der Itemantworten voraussetzt, wird eine vierstufige verbale Ratingskala aus verrechnungstechnischen Gründen bevorzugt. (Regel 11)

4.5 Sprachliche Itemgestaltung

Die sprachliche Itemgestaltung bildet den Kern der Itemkonstruktion. Dabei sind vielfältige Regeln zu beachten, die sich auf die Kategorisierung der Items, ihre sprachliche Verständlichkeit und die Eindeutigkeit des Iteminhaltes beziehen. Empfehlungen aus der Literatur werden im folgenden Unterkapitel dargestellt.

4.5.1 Kategorisierung von Itemarten

Nach Jonkisz und Moosbrugger (2007) können die Items anhand von vier Gesichtspunkten kategorisiert werden. Diese richten sich an folgenden Fragen aus: Wird nach dem Merkmal

direkt oder indirekt gefragt? Beziehen sich die Fragen auf das tatsächliche Verhalten in der Vergangenheit oder auf ein situationsbezogenes hypothetisches Verhalten? Sind die Iteminhalte konkret formuliert, sodass sie für die Testpersonen nachvollziehbar sind? Beziehen sich die Fragen auf die Testperson selbst oder fragen sie nach allgemeinen personenunabhängigen Regeln und Einstellungen?²

Direkte vs. indirekte Formulierung: Die Items des WSP werden *indirekt formuliert*. Das bedeutet, dass z.B. nicht direkt nach dem zu messenden Merkmal gefragt wird z.B. mit ‚Sind sie belastbar?‘, sondern indirekt mit ‚Denken Sie ans Aufgeben, wenn Sie mehrmals an einer Prüfung scheitern?‘, da bei einer direkten Befragung nicht immer von einer interindividuellen Übereinstimmung bezüglich der Bedeutung der Frage ausgegangen werden kann. (Regel 21)

Hypothetische vs. biographiebezogene Itemformulierung: Es können Items mit *hypothetischem* Charakter von *biographischen* Fragen unterschieden werden. Im ersten Fall werden die Testpersonen darum gebeten, sich eine Situation vorzustellen („Was würden Sie tun wenn...“ ‚Stellen Sie sich vor...‘). Im zweiten Fall wird nach einem tatsächlichen Verhalten in Situationen gefragt, von denen angenommen werden kann, dass sie die Person bereits erlebt hat. Die Items des WSP lehnen sich aufgrund der Datensammlung nach der Critical Incident Technique an konkrete Situationen aus dem Studienalltag an, von denen ausgegangen werden kann, dass die Testpersonen diese Situationen bereits erlebt haben bzw. sich in diese hineinversetzen können.

Konkrete vs. abstrakte Itemformulierung: Konkrete Itemformulierungen erhöhen die Nachvollziehbarkeit des Iteminhalts. Dadurch wird die Gefahr der Fehleinschätzung von Items reduziert. Die Items werden auf Basis der Critical Incident Daten möglichst *konkret formuliert*. Beispielsweise ist die konkrete Frage („Geben Sie einem Studienkollegen Ihre Notizen, wenn er nicht zur Vorlesung kommen kann?“) für befragte Personen leichter vorstellbar, als die abstrakte Frage („Wie hoch schätzen Sie die Hilfsbereitschaft im Studienkontext ein?“). (Regel 10)

Personalisierte vs. depersonalisierte Formulierung: Depersonalisierte Items („Sollten Menschen gegenteilige Meinungen akzeptieren?“) liefern im Vergleich zu *personalisierten Items* („Fällt es Ihnen schwer, eine gegenteilige Meinung zu akzeptieren.“) wenig Informationsgehalt in Bezug auf die befragte Person. Bei personalisierten Fragen besteht zwar im Allgemeinen die Gefahr, dass sie von Testpersonen als Verletzung der Intimsphäre

² Die beispielhaften Ausführungen stammen von der Autorin der vorliegenden Diplomarbeit.

gewertet werden, jedoch ist das bei den Items des hier zu entwickelnden Fragebogens für die Studienwahlberatung nicht zu erwarten. (Regel 20)

4.5.2 Sprachliche Verständlichkeit

Bei der Formulierung von Items steht die sprachliche Verständlichkeit und Klarheit im Vordergrund. Wenn Items zu kompliziert formuliert sind und ein mehrmaliges Durchlesen der Frage notwendig ist, besteht die Gefahr eines Motivationsverlustes und falscher Interpretationsschlüsse, was mit dem Item erfragt wird. Vorrangiges Ziel ist es daher, die Items möglichst einfach und klar zu formulieren.

Einfache Satzformulierungen: Es wird empfohlen, die Items so kurz und klar wie möglich zu gestalten. Einfache Sätze ohne Verschachtelungen, d.h. mit wenigen Nebensätzen, sind leichter verständlich als Sätze mit komplexem Satzaufbau (Edwards, 1957, zit.n. Mummendey, 2008; Jonkisz & Moosbrugger, 2007). (Regel 1)

Positive Formulierung von Items: Mehrere Autoren (Bühner, 2006; Edwards, 1957, zit.n. Mummendey, 2008; Jonkisz & Moosbrugger, 2007; Krauth, 1995) weisen darauf hin, dass Doppelverneinungen missverständlich sein können und die Testpersonen in eine von ihnen nicht intendierte Richtung antworten. Jonkisz und Moosbrugger (2007) gehen sogar soweit, dass sie generell davon abraten, Verneinungen in Items zu verwenden. Dies ist jedoch häufig nicht praktisch umsetzbar. (Regel 5)

Fachsprache/Jargon: Fachtermini oder Formulierungen können die Nachvollziehbarkeit der Items beeinträchtigen und sollten daher vermieden werden. Es können allerdings Begriffe angewandt werden, die der gesamten Zielgruppe bekannt sind (z.B. Vorlesung, Sprechstunde bei StudienbeginnerInnen). Die TestkonstrukteurInnen sollten sich die Frage stellen, welche Sprache (Jargon) in der gesamten Zielgruppe gebräuchlich ist (Edwards, 1957, zit.n. Mummendey, 2008; Bühner, 2006; Krauth, 1995). Falls Definitionen nötig sind, sollten diese noch vor der eigentlichen Frage gegeben werden (Jonkisz & Moosbrugger, 2007). (Regel 2)

Telegraphische Kürzungen: Um die Verständlichkeit der Iteminhalte sicherzustellen, empfehlen Jonkisz und Moosbrugger (2007) telegraphische Kürzungen, z.B. u.a., usw. evt. zu vermeiden. (Regel 16)

Angaben zur Intensität oder Häufigkeit von Verhaltensweisen können unklare Antworten erzeugen. Z.B.: die Verneinung der Frage „Ich gehe häufig in Vorlesungen“ könnte bedeuten, dass die Person nur selten oder nie Vorlesungen besucht oder aber, dass sie dies nach fast ausnahmslos tut. Die Bedeutung von Worten wie „viel, wenig, weniger, mehr, manchmal, häufig, oft, selten oder meistens“ etc. ist aufgrund der Relativität der Zeitangabe für einzelne Personen unterschiedlich (Jonkisz und Moosbrugger, 2007). (Regel 7)

Verallgemeinerungen: Ebenso sollte von Angaben wie „alle“, „immer“, „keine“ oder „nie“ Abstand genommen werden. Die dadurch hervorgerufene Absolutheit von Aussagen lässt keinen Raum für mögliche Ausnahmen, wodurch bei Personen, die der prinzipiellen Tendenz zwar zustimmen würden, Ambiguität ausgelöst werden könnte (Edwards, 1957, zit.n. Mummendey, 2008; Jonkisz & Moosbrugger, 2007). Ein Beispiel dafür liefert Bühner (2006, 69): „Alle Kinder machen Lärm.“

Präzise Formulierungen: Bei der Konstruktion von Items ist die präzise und wortgemäße Verwendung von Begriffen unabdingbar. Wenn es z.B. um die Bezeichnung von Schwierigkeit vs. Einfachheit von Prüfungen geht, ist das Wort „schwierig“ und nicht „schwer“ zu verwenden (Kubinger, 2007). (Regel 15)

4.5.3 Eindeutigkeit des Iteminhalts

Sprachliche Eindeutigkeit des Iteminhaltes liegt dann vor, wenn alle Testpersonen diesen in gleicher Weise verstehen. Eine Mehrdeutigkeit des Iteminhaltes kann zu unterschiedlichen Interpretationen und somit zu unterschiedlichen Antwortverhalten bei gleicher Disposition führen. Hier gibt es mannigfaltige Fehlerquellen, auf die in diesem Unterkapitel eingegangen werden. Nur wenn die Eindeutigkeit des Iteminhaltes gegeben ist, können die Testpersonen unter vergleichbaren Bedingungen getestet werden und es liegt eine Vergleichbarkeit der Messung vor (Jonkisz & Moosbrugger, 2007).

Erfragung eines einzigen Aspektes: In Zusammenhang mit der Mehrdeutigkeit des Iteminhalts steht darüber hinaus, ob im Item nur ein einziger oder mehrere Aspekte angesprochen werden (Bühner, 2006; Edwards, 1957, zit.n. Mummendey, 2008; Rost, 2004). Wenn das Item mehrere Aspekte offeriert, ist nicht sicher, worauf sich die Antwort bezieht. Darüber hinaus steht dies im Gegensatz zur geforderten Eindimensionalität von Items. (Regel 6)

Konkrete Situationen: Um den Interpretationsspielraum einzuschränken, wird empfohlen, die Iteminhalte auf konkrete Situationen zu beziehen (Jonkisz & Moosbrugger, 2007). (Regel 10)

Unterschiedliches Vorwissen darf keine Vor- oder Nachteile bei der Beantwortung der Frage bedingen (Jonkisz & Moosbrugger, 2007). Dies ist auch in Hinblick auf die Zielgruppe MaturantInnen bis Personen, die erste Erfahrungen im Studium gesammelt haben, zu berücksichtigen. (Regel 22)

Suggestive Inhalte können nach Jonkisz und Moosbrugger (2007) zu einer deutlichen Verzerrung von Antwortmustern führen. (Regel 19)

Aktualität der Iteminhalte: Bei der Entwicklung von Items ist ferner auf die Aktualität der Iteminhalte zu achten, um ein schnelles „veralten“ zu vermeiden (Jonkisz & Moosbrugger, 2007). (Regel 17)

Wertungen: Persönlichkeitsfragebögen werden auch dazu eingesetzt, Einstellungen von Testpersonen zu erfragen. Bei Persönlichkeitsfragebögen, die Auskunft über Persönlichkeitseigenschaften (im Gegensatz zu „Attitudes“) geben wollen, wird dazu geraten, keine Werthaltungen in die Items zu integrieren. Wertorientierungen können zu Verzerrungen des Antwortmusters führen (Jonkisz & Moosbrugger, 2007). (Regel 18)

4.6 Itempolung

Bei der Formulierung von Items sind mögliche Antworttendenzen (siehe Unterkapitel «Antworttendenzen») vorab zu berücksichtigen. Die Akquieszenz (Ja-Sage-Tendenz) bezeichnet jene Tendenz von Personen, unabhängig vom Iteminhalt zustimmend zu antworten. Ein umgekehrt stereotyp ablehnendes Antwortverhalten wird als Nein-Sage-Tendenz bezeichnet. Um der Akquieszenz und der Nein-Sagetendenz entgegenzuwirken, empfiehlt sich eine wechselnde Itempolung (Krauth, 1995; Mummendey, 2008; Bortz & Döring, 2003). Das heißt „die Items sollten so formuliert werden, dass zu gleichen Teilen eine Itembejahung und eine Itemverneinung für das Vorhandensein des geprüften Merkmals sprechen (Bortz & Döring, 2003, 236).“ (Regel 12)

4.7 Experteninterview

Im Anschluss an die Sichtung der Literatur wurden mit zwei Mitarbeitern des Instituts für Entwicklungspsychologie und Diagnostik (Arbeitsbereich Diagnostik) ein Experteninterview geführt. Fragestellung war die Herangehensweise an die Entwicklung von Fragebogenitems (Ablauf des Prozesses) und Regeln zu deren Formulierung. Die beiden Experten haben einschlägige Erfahrung in der Konstruktion von Persönlichkeitsfragebögen. Mag. Philipp Sonnleitner arbeitete im Rahmen des Projektes „Standard-Tests zu den Bildungsstandards in Österreich“ (Kubinger, Frebort, Holoher-Ertl & Pletschko, 2006) an der Entwicklung der Standard-Tests für Deutsch in den 4. und 8. Schulstufen. In seiner Diplomarbeit (Sonnleitner, 2007) entwickelte er ein Konstruktionsrational für den Fähigkeitsbereich des Leseverständnisses. Mag. Manuel Reif entwickelte im Rahmen seiner Diplomarbeit einen adaptiven Persönlichkeitsfragebogen für die Persönlichkeitsdimension „Extraversion“ unter Anwendung eines Konstruktionsrationalen. Dazu setzte er die Idee der Itemkonstruktion aufgrund „itemgenerierender Regeln“, die bislang nur für Leistungstests angewandt wurde, erstmals für Persönlichkeitsfragebögen um (Reif, 2008).

Der Großteil der Angaben aus dem Interview deckt sich mit den Regeln aus der Literaturrecherche. Die zentralsten Aussagen sollen hier dennoch festgehalten werden:

Die **sprachliche Eindeutigkeit des Iteminhalts** hat bei der Entwicklung oberste Priorität. Die Frage im Itemstamm darf keinen breiten Interpretationsspielraum für die befragten Personen eröffnen.

Genauigkeit der Aussage: Situationen (nach der CIT) sind im Itemstamm so genau wie möglich zu beschreiben. (Regel 10) Dabei steht jedoch die **Einfachheit des Ausdrucks** im Vordergrund. Verschachtelte Sätze führen zu Missverständnissen darüber, was mit dem Item gemeint ist. Es ist ein Balanceakt zwischen genauer Situations- und Verhaltensbeschreibung und Einfachheit des Satzbaus. (Regel 1)

Keine Verwendung möglicherweise **missverständlicher oder doppeldeutiger Begriffe**, z.B. fremde Menschen. Fremd könnte im Sinne von „unbekannt“ aber auch im Sinne von „andersartig“ oder auch „AusländerInnen“ verstanden werden. (Regel 3)

Konstruktion von Distraktoren: Besonderes Augenmerk ist auf die Konstruktion der Distraktoren zu legen, d.h. jene Antwortmöglichkeiten (response options), die zusätzlich zur richtigen vorgegeben werden um die Identifikation der richtigen Antwort zu erschweren.

Antwortformat: Die Anzahl der Antwortkategorien muss in Hinblick auf die Akzeptanz der befragten Personen, Reaktanz, Informationsgehalt und anschließende Modellprüfung gewählt werden. Zur Minimierung der Reaktanz ist die Verwendung einer Analogskala oder eines gebundenen Antwortformats mit mehrkategorialen Kategorien empfehlenswert. Die anschließende Modellprüfung nach dem dichotom logistischen Modell von Rasch (dichotome Verrechnung) und die Vermeidung der Antworttendenz „Tendenz zur Mitte“ legt eine geradzahlige Anzahl von Kategorien nahe. (Regel 11)

Bekanntheit der Iteminhalte: Um die Nachvollziehbarkeit der Iteminhalte zu gewährleisten, sollen Situationen und Verhaltensweisen erfragt werden, von denen angenommen werden kann, dass die gesamte Zielgruppe (und nicht nur einzelne Personen) damit vertraut sind bzw. die zumindest leicht vorstellbar sind. z.B. Statt von der Mechanik I Prüfung zu sprechen, allgemein von einer Prüfung sprechen. (Regel 4)

Konkretisierung der Zeitangabe: Wird ein Zeitpunkt oder eine Zeitspanne im Item angegeben, ist eine eindeutige Definition notwendig, um den Iteminhalt zu konkretisieren. Z.B.: statt „In letzter Zeit war ich.....“ besser „In den letzten zwei Wochen war ich...“.

Eine zusätzliche Regel, die in den Regelkatalog aufgenommen wurde, stammt von Mag. Lale Khorramdel. Sie arbeitet am Arbeitsbereich Diagnostik (Institut für Entwicklungspsychologie und psychologische Diagnostik, Fakultät für Psychologie an der Universität Wien) und hat gemeinsam mit Philipp Sonnleitner an der Testentwicklung für die Standardtests in Englisch und Deutsch (4. und 8 Schulstufe) mitgearbeitet:

Attraktivität der Itembeantwortung gegen die Messrichtung: Items sind so zu formulieren, dass es der Testperson leicht fällt, sie als zutreffend anzukreuzen, wenn sie tatsächlich zutreffen. (also auch wenn das Items einen vermeintlich „negativen“ Inhalt hat). Und umgekehrt Items derart formulieren, dass auch Items mit vermeintlich positiven Inhalt verneint werden können, wenn dies tatsächlich zutrifft (Auszug aus einem Gespräch mit Khorramdel, L. am 17.09.2008). (Regel 9)

4.8 Ergebnis Regelkatalog

Aus der bestehenden Literatur, den Ergebnissen des Experteninterviews und den empirischen Erfahrungen bei der Entwicklung des WSP (siehe Unterkapitel «Probleme bei

der Itementwicklung und Ergänzung des Regelkataloges») wurde folgender ‚Regelkatalog zur die Fragebogen-Itemkonstruktion‘ konzipiert.

Der Regelkatalog bezieht sich auf verbale Items von Persönlichkeitsfragebögen und dient dabei als Handlungsanleitung. Er fasst die zentralen Aspekte der Itementwicklung zu den Themen der sprachlichen Verständlichkeit, inhaltlichen Eindeutigkeit, Itempolung, Itemschwierigkeit, Eindimensionalität, Antwortformat und sprachlichen Variabilität zusammen. Die vorliegenden Regeln sind an die jeweilige Verfahrensart, den Zweck des Verfahrens und das Datenmaterial anzupassen.

Regelkatalog zur Fragebogen-Itemkonstruktion

- 1) Items so einfach wie möglich formulieren, damit der Inhalt leicht verständlich ist. Verschachtelte Sätze, insbesondere viele Nebensätze, sollten vermieden werden.
- 2) Keine Verwendung von Fachwörtern. Begriffe aus dem Alltag der Zielgruppe sind jedoch möglich, z.B. Prüfung, Semester, Seminararbeit, Kursanmeldung, etc.
- 3) Keine Verwendung missverständlicher oder doppeldeutiger Begriffe, z.B. ‚fremde‘ Person kann sowohl im Sinne von ‚unbekannt‘ oder auch ‚von anderer Herkunft‘ verstanden werden.
- 4) Situationen und Verhaltensweisen erfragen, von denen angenommen werden kann, dass die gesamte Zielgruppe (und nicht nur einzelne Personen) damit vertraut ist bzw. die zumindest leicht vorstellbar sind. Unterschiedliches Vorwissen darf keine Vor- oder Nachteile bei der Beantwortung der Fragen bedingen. Z.B. Statt von der ‚Prüfung Mechanik I‘ zu sprechen, ist der allgemeine Begriff ‚Prüfung‘ vorzuziehen.
- 5) Keine doppelten Verneinungen.
- 6) Ein Item soll nur eine einzige Aussage enthalten.
- 7) Worte wie „viel, wenig, weniger, mehr, manchmal, selten, oft, häufig etc.“ vermeiden bzw. anhand eines Beispiels erklären, damit jede/r dasselbe darunter verstehen kann.

Z.B. „Wenn etwas schwieriger wird als zunächst gedacht...“ statt „Wenn etwas schwierig ist....“

- 8) Eindimensionalität, d.h. das Item soll einer einzigen Skala eindeutig zuordenbar sein
- 8b) Ad Eindimensionalität, Einschränkung dieser Bedingung während des Arbeitsprozesses der Itemgenerierung: Aufgrund der Verwendung der Critical Incident Daten kann bei einigen Items die Trennung der Dimensionen nicht aufrechterhalten werden. Für eine Verhaltensweise sind z.T. mehrere Eigenschaften notwendig. Wenn die Eindimensionalität bei der Itemkonstruktion nicht eingehalten werden kann, ist zunächst anzugeben, welche Dimensionen erhoben werden und zu vermerken, welche davon im Vordergrund steht. Anschließende Überprüfung erforderlich.
- 9) Items sind so zu formulieren, dass es der Testperson leicht fällt, sie als zutreffend anzukreuzen, wenn sie tatsächlich zutreffen. (also auch wenn das Items einen vermeintlich „negativen“ Inhalt hat). Und umgekehrt Items derart formulieren, dass auch Items mit vermeintlich positiven Inhalt verneint werden können, wenn dies tatsächlich zutrifft.
- 10) Durch situative Einengungen kann der Interpretationsspielraum der Items reduziert werden. D.h. Formulierung konkreter Situationen und Verhaltensweisen.
- 11) Antwortformat: sämtliche Items sollten so konzipiert sein, dass sowohl eine Analogskala als auch ein mehrkategorielles Antwortformat (geradzahlig) einsetzbar ist, z.B. eine 4-stufige Ratingskala (ohne Mittelkategorie). Dadurch können Antworttendenzen wie Reaktanz, Akquieszenz und Tendenz zur Mitte vermieden werden. Die Dichotomisierung zur anschließenden Prüfung nach dem dichotom logistischen Modell von Rasch (1960) kann durchgeführt werden. Die Erstellung von Distraktoren und deren ‚Problemanfälligkeit‘ werden vermieden.
- 12) Variation der Itempolung: Wechseln der Antwortrichtung zur Vermeidung von Akquieszenz.
- 13) Die Eindeutigkeit der Messrichtung eines Items muss für die BeurteilerInnen gegeben sein. D.h. es muss nachvollziehbar sein, in welche Richtung das Item gepolt ist.
- 14) Die Items sollten weder zu leicht noch zu schwer zu beantworten sein. Items, auf die jede/r mit Ja oder Nein antwortet, sind nicht informativ.

- 15) Bei der Konstruktion von Items ist die präzise und wortgemäße Verwendung von Begriffen unabdingbar. Wenn es z.B. um die Bezeichnung von Schwierigkeit vs. Einfachheit von Prüfungen geht, ist das Wort „schwierig“ und nicht „schwer“ zu verwenden.
 - 16) Vermeidung umständlicher Fragen und telegraphischer Kürzungen wie: usw., usf., u.U., u.a....
 - 17) Aktualität: Die Items sollten so formuliert sein, dass sie nicht schnell „veralten“.
 - 18) Wertungen sollten bei Persönlichkeitsfragebögen zur Erfassung von Persönlichkeitseigenschaften (vs. Einstellungen) vermieden werden
 - 19) Vermeidung suggestiver Inhalte.
 - 20) Personalisierung von Items, d.h. die Testperson sollte direkt angesprochen werden.
 - 21) Es soll nicht direkt nach dem zu messenden Merkmal gefragt werden. Anstatt zu fragen „Sind sie belastbar?“, ist eine indirekte Variante vorzuziehen ‚Denken Sie ans Aufgeben, wenn Sie mehrmals an einer Prüfung scheitern?‘
 - 22) Die Begriffe „im Allgemeinen“ und „grundsätzlich“ können zu unterschiedlichen Interpretationen auf Seiten der befragten Person führen. „Im Allgemeinen“ legt eine Depersonalisierung nahe (betrifft alle und nicht nur die befragte Person). Das Wort „grundsätzlich“ kann je nach Betonung im Sinne von „ausschließlich“ oder als „grundsätzlich ja, aber Ausnahmen bestätigen die Regeln“ interpretiert werden.
 - 23) Variation der Fragestellungen im Itemstamm, z.B. nicht nur „Es fällt mir schwer/leicht, ...“, sondern auch „Es fällt mir auf, dass ich ...“, „Ich wäre gerne besser darin, ...“, Ich neige dazu,....“, Ich versuche laufend,....“
-

5 Itempool des Wiener Studieneignungs-Persönlichkeitsinventar (WSP)

5.1 Einleitung

Das vorliegende Kapitel beschreibt die Entwicklung des Itempools für das Wiener Studieneignungs-Persönlichkeitsinventar (WSP) unter Anwendung des ‚Regelkataloges zur Fragebogen-Itemkonstruktion‘. Zunächst wird dargestellt, wie die Dimensionen des WSP ausgewählt und definiert sowie die Zielgruppe, Zweck und Art der Information, Testlänge und Testadministration beschrieben wurden. In einem weiteren Schritt wird der Itempool entwickelt und die zur Anwendung kommenden Instrumente zur Qualitätssicherung während des Itemerstellungprozesses dokumentiert.

Das Wiener Studieneignungs-Persönlichkeitsinventar (WSP) wird als Persönlichkeits-Strukturtest mit dem Anspruch entwickelt, die Persönlichkeit in zentralen Strukturdimensionen umfassend zu beschreiben. D.h. er erfasst keine unimodalen Persönlichkeitsdimension (z.B. Zielorientierung), sondern eine Vielzahl von Persönlichkeitsdimensionen, die gemeinsam ein Stärken- und Schwächenprofil von Persönlichkeitseigenschaften bilden, welches sich aufgrund der vorangegangenen Anforderungsanalyse als relevant für die Studieneignung dargestellt haben. Der WSP ist multidimensional konzipiert, wobei die einzelnen Subtests unidimensional erstellt werden.

Die zentrale Fragestellung bei der Entwicklung des Itempools ist die Dokumentation des Prozesses der Itemerstellung unter Einbezug mehrerer Instrumente zur Qualitätssicherung. Wesentliches Element der Qualitätssicherung nach der DIN 33430 ist die Verschränkung der Itementwicklung mit dem zugrunde liegenden Anforderungsprofil (Heyse & Kersting, 2004; Kersting, 2008). Ausgangspunkt für die Itementwicklung des WSP ist daher die Verknüpfung der Iteminhalte mit den Daten der Anforderungsanalyse nach der Critical Incident Technique (Flanagan, 1954), welche im vorherigen Kapitel beschrieben wurde. Die in der Anforderungsanalyse genannten critical incidents bilden die Grundlage für die Iteminhalte. Durch die direkte Verwertung der Daten wird gewährleistet, dass die Items aufgrund der inhaltlichen Nähe diejenigen Persönlichkeitseigenschaften erheben, welche für die Bewältigung des Studienalltages tatsächlich relevant sind. Die Validität des Persönlichkeitsfragebogens kann durch die Verwendung der CIT-Daten erhöht werden (Koch et al., 2006). Ein weiterer wesentlicher Effekt ist, dass die Iteminhalte auf die Zielgruppe zugeschnitten werden. Die Iteminhalte beziehen sich auf Situationen und Verhaltensweisen,

die für die Fragestellung der befragten Person (Studieneignung) einen unmittelbaren Praxisbezug beinhalten.

Vor der Itementwicklung wurde ein Regelkatalog zur Entwicklung von Fragebogenitems erstellt und im Zuge des Itemerstellungprozesses ergänzt. Darüber hinaus werden webbasierte Rückkopplungsschleifen in den Entwicklungsprozess (in Form von Diskussionsforen und Workshops) eingebaut. Ebenso werden Kriterien für die Qualitätsabnahme von Items festgelegt.

5.2 Auswahl und Definition der Merkmale

Das Wiener Studieneignungs-Persönlichkeitsinventar wird mit dem Anspruch entwickelt, Persönlichkeitseigenschaften, die für die erfolgreiche Studienbewältigung relevant sind, zu erfassen. Ausgangspunkt zur Auswahl der zu erhebenden Merkmale bildet das Anforderungsprofil zur Allgemeinen Studierfähigkeit. Das Anforderungsprofil ist Tab. 14 im Anhang entnehmen.

5.2.1 Auswahl der Merkmale (Dimensionen)

Die Auswahl der Merkmale, welche durch das Persönlichkeitsinventar erfasst werden sollen, erfolgt anhand zweier Kriterien. Diese Kriterien sind einerseits die Art des Merkmals (Persönlichkeitseigenschaft vs. kognitive Fähigkeiten) und andererseits die hinreichende Relevanz des Merkmals für die Fragestellung (Häufigkeit der Nennungen).

Das Anforderungsprofil zur allgemeinen Studierfähigkeit enthält sowohl psychologische als auch nicht-psychologische Merkmale, kognitive Fähigkeitsmerkmale sowie Persönlichkeitseigenschaften. Für den WSP wurden jene Merkmale ausgewählt, die sinnvollerweise durch einen Persönlichkeitsfragebogen ermittelbar sind. Nicht-psychologische und kognitive Fähigkeitsmerkmale werden ausgeschlossen.

Damit das Persönlichkeitsinventar Persönlichkeitseigenschaften erfasst, welche für die Studieneignung hinreichend relevant sind, werden von den verbleibenden Persönlichkeitseigenschaften nur jene ausgewählt, deren Häufigkeitswert oberhalb eines festgelegten Wertes liegen. Dieser „Cut-Off-Wert“ wurde mit 40 Nennungen festgesetzt. In den Persönlichkeitsfragebogen werden somit jene Merkmale aufgenommen, die 40 Mal oder häufiger von den StelleninhaberInnen genannt wurden.

Die Merkmale, welche die oben beschriebenen Kriterien erfüllen, bilden die Subskalenstruktur des WSP. Die Konstrukte, denen die Merkmale zugeordnet sind, bilden

die übergeordnete Skalenstruktur. Daraus ergibt sich eine Skalenstruktur, bestehend aus sieben Skalen und 22 Subskalen (siehe Tab. 8).

Die Skala Organisationsfähigkeit besteht aus den fünf Subskalen *Zeitmanagement*, *Prioritäten setzen*, *Planen*, *Organisationstalent* und *Informiertheit*. Die Skala Arbeitshaltungen setzt sich aus den vier Subskalen *Gewissenhaftigkeit*, *Genaues Arbeiten*, *Eigeninitiative* und *Selbstständiges Arbeiten* zusammen. Die Skala Motivation baut auf den sechs Subskalen *Durchhaltevermögen*, *Lernbereitschaft*, *Engagement*, *Zielorientierung*, *Zielsetzung* und *Erfolgszuversicht* auf. Vom Konstrukt Belastbarkeit konnte nur das Merkmal *Frustrationstoleranz* die oben angegebenen Kriterien erfüllen. Unter die Skala Soziale Kompetenz fallen die drei Subskalen *Kommunikationsfähigkeit*, *Kontaktfähigkeit* und *Teamfähigkeit*. Dabei wurden in die Subskala *Teamfähigkeit* aus theoretischen Überlegungen die Merkmale Teamfähigkeit, Teamorientierung, Kollegialität und Kompromissbereitschaft subsumiert. Die Skala Emotionale Komponenten setzt sich aus den zwei Subskalen *Distanzierungsfähigkeit* und *Offenheit* zusammen. Die Skala Selbstbild besteht aus der Subskala *Selbstbewusstsein*.

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

Tab. 8 Skalen und Subskalen des WSP

5.2.2 Definition der Merkmale

Ein Vorteil der Anforderungsanalyse nach der CIT (Flanagan, 1954) ist, dass die Definitionen der zu erhebenden Merkmale nicht theoriegeleitet erstellt werden, sondern auf Befragungsergebnissen der Vorgesetzten und StelleninhaberInnen (Lehrende und Studierende) basieren. Die gewonnenen Dimensionen des Anforderungsprofils sind durch die Situations- und Verhaltensbeschreibungen definiert. Bühner (2006) spricht in diesem Zusammenhang von einer induktiven Fragebogenkonstruktion. Vorteil dieser Methode ist die höhere Validität und der höhere Praxisbezug der Arbeitsdefinitionen (Koch et al., 2006). Eine detaillierte Definition des zu erfassenden Konstruktes ist dabei nicht nur für die Validität, sondern auch für die interne Konsistenz von Bedeutung, denn durch eine detaillierte Definition kann die inhaltliche Breite der Skala und damit die Itemanzahl vergrößert werden, was positiv mit der Reliabilität der Skala korreliert (Inceoglu & Bartram, 2007).

5.2.2.1 Arbeitsdefinitionen

Nach der Critical Incident Technique (Flanagan, 1954) bestehen die Arbeitsdefinitionen aus der Sammlung der für dieses Merkmal kennzeichnenden Verhaltensweisen. Dadurch bleibt die Nachvollziehbarkeit zwischen der Datengrundlage und dem Anforderungsprofil nach Vollzug des Interpretationsschrittes erhalten. Für die Entwicklung des WSP bieten die Verhaltensbeschreibungen in den Arbeitsdefinitionen zusätzlich den Vorteil, dass sie gleichzeitig die Inhalte der Items bedingen.

Exemplarisch werden die Arbeitsdefinitionen für die Dimensionen (Subskalen) *Zeitmanagement* (Skala Organisation), *Gewissenhaftigkeit* (Skala Arbeitshaltungen) und *Engagement* (Skala Motivation) angeführt. Die weiteren Arbeitsdefinitionen sind im Anhang aus Tab. 15 zu entnehmen.

Subskala	Arbeitsdefinition
Zeitmanagement (Skala Organisation)	Einteilen, wie viel Zeit Arbeitsschritte benötigen, zeitliche Koordination von Arbeitsschritten, zeitliche Koordination von Mehrfachanforderungen, Einteilung Arbeitszeit und Freizeit/Studium und Arbeit +
Gewissenhaftigkeit (Skala Arbeitshaltungen)	Termine regelmäßig wahrnehmen, Lehrveranstaltungen regelmäßig besuchen, immer gut vorbereitet sein, immer auf dem aktuellen Wissensstand sein, Pünktlichkeit, immer mitschreiben, eine „saubere“ Mitschrift haben, alle wichtigen Punkte berücksichtigen, die Hausübung selber machen und nicht abschreiben, sich an den eigenen Zeitplan halten, Kontakte nicht abreißen lassen sondern pflegen, Rücksprache halten, um sich zu vergewissern, dass alles so passt, diszipliniertes Arbeiten +
Engagement (Skala Motivation)	gut mitarbeiten, aus Interesse immer zu den Lehrveranstaltungen gehen, zusätzliche Praktika machen, wenn nicht nötig oder Stoff vertiefen bzw. mehr lernen als nötig, sich intensiv mit etwas beschäftigen, Anstrengungsbereitschaft +

5.2.2.2 Definitionen von Merkmalen im Testmanual

Für die benutzerfreundliche Handhabung im Manual des Persönlichkeitsfragebogens wird eine Umformulierung der Arbeitsdefinitionen empfohlen. Folgende Variante wäre denkbar: Eine hohe Ausprägung in der Subskala x weist darauf hin, dass... Eine niedrige Ausprägung in der Subskala x, weist darauf hin, dass... Dies soll anhand der Subskala *Distanzierungsfähigkeit* (Skala Emotionale Komponenten) illustriert werden (siehe Tab. 9).

Distanzierungsfähigkeit (Skala Emotionale Komponenten)	
Arbeitsdefinition	in Bezug auf die eigenen Arbeiten neutral bleiben, objektiv beurteilen, Kritik nicht persönlich nehmen, etwas nicht zu wichtig nehmen, loslassen können +
Definition Manual	Eine hohe Ausprägung in der Subskala Distanzierungsfähigkeit weist darauf hin, dass bei der Beurteilung eigener Arbeiten ein möglichst neutraler Standpunkt eingenommen werden kann, welcher eine möglichst objektive Beurteilung erlaubt. Kritik an der eigenen Arbeit wird auf einer sachlichen Ebene aufgenommen. Eine niedrige Ausprägung weist darauf hin, dass Kritik persönlich und arbeitsbezogene Themen zu wichtig genommen werden. Bei Missglücken einer Aufgabe fällt das Loslassen des Geschehenen schwer.

Tab. 9 Umwandlung Arbeitsdefinition für Testmanual

5.3 Zielgruppe

Zielgruppe des WSP sind StudienbeginnerInnen und StudentInnen, die das Studium wechseln möchten und Beratung suchen. Die Items werden prototypisch erstellt, d.h. sie werden direkt aus den erfolgskritischen und –hinderlichen Verhaltensweisen, die bei der CIT genannt wurden, abgeleitet. Es ist möglich, dass die Items daher studienspezifische Situationen beinhalten, die SchulabgängerInnen noch unbekannt sind. Es ist anschließend zu prüfen, inwiefern die Iteminhalte MaturantInnen bereits bekannt sind bzw. sie sich in hypothetische Studiensituationen hineinversetzen können.

Die Breite des Schwierigkeitsgrades hängt von der Homogenität der Zielgruppe ab (Jonkisz & Moosbrugger, 2007). Die Begriffe Leichtigkeit und Schwierigkeit einer Testaufgabe drücken prinzipiell die Wahrscheinlichkeit aus, mit der auf ein dichotomes Item eine mit 1 kodierte Antwort gegeben wird, d.h. im Sinne der Erwartung (in Schlüsselrichtung) geantwortet wird (Rost, 2004).

Es wird davon ausgegangen, dass MaturantInnen und StudienwechslerInnen unterschiedliche Vorkenntnisse und Erfahrungen in den einzelnen Merkmalsdimensionen mitbringen. StudienwechslerInnen werden aufgrund der Einbindung in den Hochschulbetrieb bereits mehr Erfahrungen z.B. in den Bereichen der Organisation und Arbeitshaltungen gewonnen haben. Die Items sollen daher inhaltlich und nach ihrer Schwierigkeit breit

gefächert sein, um möglichst viele Merkmalsausprägungen abdecken zu können. Eine Normierung des WSP in Hinblick auf Studienfortschritt und Studienrichtungen ist vom Arbeitsbereich Diagnostik in Planung.

5.4 Zweck des WSP

Seit dem Hochschulrahmengesetz 1998 und der verstärkten Verlagerung der Auswahl von Studierenden in den Zuständigkeitsbereich der Universitäten geht der allgemeine Trend in Richtung der Entwicklung von Verfahren zur Auswahl von Studierenden (Hell, Ptok & Schuler, 2007; Schmitt, 2005; Zimmerhofer & Hornke, 2005). Die Intention der Verfahrensentwicklung des WSP ist ein förderdiagnostischer Ansatz. Der WSP soll – zum gegensätzlichen Trend – nicht in Selektionssituationen sondern als Instrument zur Verbesserung der Beratung von StudienbeginnerInnen und StudentInnen eingesetzt werden. Vom Einsatz des Persönlichkeitsfragebogens in Selektionssituationen wird allein aufgrund der möglichen Verfälschbarkeit von Fragebögen, welche eine hohe Augenscheinvalidität, i.S.v. Durchschaubarkeit der Messintention besitzen, abgeraten (Kubinger, 2002).

Im Rahmen des vom Institut für Entwicklungspsychologie und Diagnostik, Arbeitsbereich Diagnostik (Fakultät für Psychologie, Universität Wien) geplanten Self-Assessments wird das Ergebnis des Persönlichkeitsfragebogen den Testpersonen in Form eines Stärken- und Schwächenprofils individuell und anonym rückgemeldet mit dem Angebot einer ergänzenden Beratung durch TestpsychologInnen (Frebort & Kubinger, 2006b).

5.5 Art der Information

Bei der Art der Information kann zwischen subjektiven und objektiven Informationen unterschieden werden. Während subjektive Informationen von der Person selbst in Form einer Selbstbeurteilung gegeben werden, bauen objektive Informationen auf eine Fremdbeurteilung durch andere Personen auf (Bühner, 2006). Auf den missverständlichen Gebrauch des Wortes „objektiv“ sei an dieser Stelle hingewiesen, denn auch Fremdbeurteilungen bauen auf der subjektiven Einschätzung und Wahrnehmung anderer Personen auf und unterliegen damit gewissen Verzerrungen.

Der WSP wird als Instrument zur Selbstbeurteilung konzipiert. Die Art der erfassten Information ist subjektiv, d.h. die Testpersonen geben über sich selbst Auskunft. Dies setzt von den Testpersonen die Fähigkeit zur Metakognition voraus, d.h. die Person besitzt Einsicht in die eigenen kognitiven Prozesse, welche die Grundlage für das Vorhandensein einer Selbstkenntnis darstellt. Um von den Ergebnissen Rückschlüsse über einen Ausschnitt

der Persönlichkeit der Person ziehen zu können, wird zusätzlich die Bereitschaft zur Offenbarung vorausgesetzt. Die Person ist bereit, Auskunft über ein reales Selbstbild im Gegensatz zu ihrem Idealbild zu geben (Rost, 2004). Dies sollte unter dem Aspekt der Freiwilligkeit der Testung im Rahmen des Self-Assessments und der individuellen anonymen Rückmeldung gegeben sein.

5.6 Testlänge

Die Testlänge wird als Anzahl der Items definiert und wird auf Basis der Subskalenanzahl und testtheoretischer Überlegungen festgelegt.

Der Persönlichkeitsfragebogen wird in einer Folgeuntersuchung auf Modellgültigkeit nach der probabilistischen Testtheorie im Sinne des (dichotom) logistischen Testmodells nach Rasch (1960) überprüft. Für die Überprüfung der Modellgültigkeit ist es sinnvoll, eine genügend große Menge an Items zu entwickeln um die Wahrscheinlichkeit zu erhöhen, nach Ausschluss nicht Rasch-Modell-konformer Items noch genügend Items beibehalten zu können.

Je Subskala wird daher eine Anzahl von 10 bis 20 Items angestrebt, abhängig von der Datenfülle der Anforderungsanalyse. Daraus ergibt sich eine Gesamt-Itemanzahl von 220 bis 440.

Angestrebt wird, dass nach der Durchführung der Analyse auf Modellgültigkeit der Großteil der Skalenstruktur und mindestens 150 Items erhalten werden können. Die Testzeit kann vorab aufgrund der bevorstehenden Analysen nicht festgelegt werden. Die Bearbeitungsdauer sollte jedoch 40 Minuten nicht überschreiten. Studieninteressierten Personen kann zwar durchaus eine längere Testzeit zugemutet werden. Es gilt jedoch zu vermeiden, dass testfremde Variablen wie Konzentration oder Motivation die Messung beeinträchtigen (Jonkisz & Moosbrugger, 2007).

5.7 Testadministration

Der WSP wird in einem Online-Self-Assessment eingesetzt und daher als computerunterstützter Test entwickelt. Dieses Format ist ökonomischer und die Auswertungsobjektivität ist höher als beim Papier-Bleistift-Format, da die Auswertung weniger Fehlerquellen unterliegt und die Testergebnisse sofort rückgemeldet werden können (Jonkisz & Moosbrugger, 2007). Je nach Fragestellung kann der WSP als Einzel- oder in Gruppentestungen vorgegeben werden.

Das Online-Self-Assessment sieht eine Einzeltestung vor. Eine Beobachtung der befragten Person in der Testsituation ist aufgrund der Online-Testung nicht möglich. Zu hinterfragen gilt, welche Auswirkungen webbasierte Testungen (nicht-standardisierte Teststuiation) auf die Testergebnisse haben. Studien (Hertel, 2002; Miller, 2009; Buchanan & Smith, 1999) konnten bisher die Anwendbarkeit webbasierter Testungen und die Vergleichbarkeit der Ergebnisse mit Papier-Bleistift-Testungen nachweisen. Miller (2009) untersuchte die Vergleichbarkeit einer webbasierten Testung mit einer herkömmlichen Testvorgabe. In der Studie wurden Personen randomisiert drei Versuchsbedingungen zugewiesen (Papier-Bleistift, web-basierte Testung und webbasierte Testung mit Unterbrechung). Die Ergebnisse zur Validität und Reliabilität weisen darauf hin, dass eine Online-Testung eine gute Alternative zur herkömmlichen Papier-Bleistift Testung ist. Die Auswertungsobjektivität und Ökonomie sind deutlich höher, ebenso der Zugang zu einer größeren Population. Darüber hinaus stellten Hertel et al. (2002) eine hohe Akzeptanz der Internettestung fest.

5.8 Konstruktion des Itempools

Die Fragebogenitems für das Wiener Studieneignungs-Persönlichkeitsinventar werden entlang des zuvor formulierten ‚Regelkataloges zur Fragebogen-Itemkonstruktion‘ entwickelt. Grundlage für die Itemkonstruktion bilden die Daten der Anforderungsanalyse. Die Anforderungsanalyse wurde im Kapitel «Anforderungsanalyse – allgemeine Studierfähigkeit» beschrieben, weshalb hier nur verkürzt darauf eingegangen wird. Durch die direkte Verarbeitung der Daten aus der Anforderungsanalyse werden die Items situations- und verhaltensspezifisch entwickelt und damit maßgeschneidert auf die Zielgruppe abgestimmt. Die Datenstruktur der Anforderungsanalyse gliedert sich nach der Critical Incident Technique (Flanagan, 1954) in Nennungen erfolgskritischer Situationen und dazugehöriger erfolgsförderlicher und -hinderlicher Verhaltensweisen. Jede dieser Situations-Verhaltensweisen-Kombinationen wurde im Zuge der Interpretation (nach Mayring 2008; Frebort & Kubinger, 2006) mit einer oder mehreren zugrunde liegenden Eigenschaften (Merkmalen) kodiert. Diese nach Merkmalen kodierte Datensammlung bildet die inhaltliche Grundlage für die Itementwicklung.

Die bestehende CIT-Datenbasis legt die Verwendung von Items in Form situativer Fragen nahe. Im folgenden Unterkapitel werden die wesentlichen Merkmale der situativen Fragen nach Latham, et al. (1980) und Schuler (2002) beschrieben. Diese Form der Items wird in einer abgeänderten Version für die Items des WSP übernommen.

5.8.1 Methode – Situative Fragen

Das Datenmaterial der Anforderungsanalyse, bestehend aus Situations- und Verhaltensbeschreibungen, nach der Critical Incident Technique (Flanagan, 1954) bietet die Möglichkeit, die Items des WSP in Form situativer Fragen (Latham et al. 1980) aufzubauen. Situative Fragen sind dadurch gekennzeichnet, dass sie hypothetische Aufgaben- oder Problemsituationen enthalten und die befragten Personen angeben, wie sie in den geschilderten Situationen verhalten würden. Vorteil der situativen Fragen ist, dass sie möglichst eng an die berufliche Praxis angelehnt sind und dadurch ein hoher Praxisbezug erreicht werden kann.

Box 5: Beispiele für zwei situative Fragen in Einstellungsinterviews

„Sie kommen als Auszubildender in eine neue Abteilung. Der Abteilungsleiter ist nicht da und der Nebenausbilder hat keine Zeit und keine Arbeit für Sie. Sie setzen sich an einen freien Schreibtisch und lesen eine Zeitschrift, die dort liegt. Plötzlich erscheint der Abteilungsleiter und kritisiert Sie. Was tun Sie (Schuler 2002, 174)?“

„Your spouse and two teenage children are sick in bed with a cold. There are no relatives or friends available to look in on them. Your shift starts in 3 hours. What would you do in this situation (Latham et al. 1980, 424)?“

Die Anwendung situativer Fragen in Einstellungsinterviews wurde von Latham et al. (1980) und Schuler (2002) beschrieben. Schuler (2002) nennt die situativen Fragen auch „mentale Tätigkeitssimulationen“, d.h. die befragte Person versetzt sich in eine beschriebene Situation hinein und gibt an, welches Verhalten sie in dieser Situation zeigen würde. Situative Fragen beziehen sich folglich auf hypothetische, in der Zukunft liegende Problemstellungen, während sich biographische Fragen dadurch auszeichnen, dass nach Verhalten in tatsächlich stattgefundenen Situationen gefragt wird.

Die von Latham et al. (1980) und Schuler (2002) beschriebenen situativen Fragen sind für Einstellungsinterviews im Personalwesen konzipiert. Offene Fragestellungen ohne Vorgabe von Antwortkategorien – wie in den oben genannten Beispielen – erfordern in Persönlichkeitsfragebögen ein freies Antwortformat. Ein freies Antwortformat besitzt jedoch eine geringe Auswertungsobjektivität, ist wenig ökonomisch und für ein Online-Self-Assessment nicht praktikabel umsetzbar (Jonkisz & Moosbrugger, 2007).

Um ein freies Antwortformat zu vermeiden, ist eine adaptierte Variante situativer Fragen für Persönlichkeitsfragebögen denkbar. Dazu wird die Problemsituation im Itemstamm beschreiben und mehrere Verhaltensalternativen als Antwortkategorien vorgegeben, wie dies beispielsweise von Schuler und Funke (1995) für ein Einstellungsinterview umgesetzt wurde (siehe Box 6).

Box 6: Beispiel für eine situative Frage mit Antwortkategorien
(Schuler & Funke, 1995, 248)

Die Leistung eines Ihrer Mitarbeiter hat nachgelassen. Anlässlich Ihrer jährlichen Gehaltsgespräche müssen Sie ihm erklären, daß seine Gehaltserhöhung geringer ausfällt als die Zulage, die die meisten seiner Kollegen bekommen. Wie gehen Sie vor?

Beispielantwort 0 Punkte:

Ich sage dem Mitarbeiter, daß ich ihm gerne mehr gegeben hätte, daß aber die Geschäftsleitung keinen weiteren Rahmen offenläßt.

Beispielantwort 2 Punkte:

Ich sage dem Mitarbeiter, daß er seine Ziele nicht erreicht hat, und stelle ihm bei Verbesserung eine Gehaltsüberprüfung in Aussicht.

Beispielantwort 4 Punkte:

Ich sage dem Mitarbeiter, daß ich mir Gedanken über seine nachlassende Leistung mache, derentwegen die Zulage geringer ausfällt. Ich versuche, gemeinsam mit ihm die Gründe herauszufinden. Dann besprechen wir Maßnahmen, die Leistung wieder zu verbessern, und vereinbaren neue Ziele.

Die Situationsbeschreibung im Itemstamm und die Verhaltensweisen als Antwortkategorien bergen in der Umsetzung für einen Persönlichkeitsfragebogen allerdings mehrere Nachteile. So kann beispielsweise die Einschränkung der Handlungsoptionen auf einzelne Alternativen zu Reaktanz (Brehm, 1966) vonseiten der befragten Personen führen.

Darüber hinaus ist es schwierig, sämtliche prinzipiell möglichen Verhaltensweisen in den Antwortkategorien festzuhalten. Die Verletzung der Exhaustivität kann ebenfalls zu Reaktanzphänomenen führen, wenn die Person eine andere Verhaltensweise als die angegebenen zeigen wollen würde (Jonkisz & Moosbrugger, 2007).

Das Lesen der ausführlichen Beschreibung der Situation und Antwortmöglichkeiten (response options) und die Abwägung der zutreffenden Antwort erfordert viel Zeit, weshalb nur eine begrenzte Anzahl von Situationsbeschreibungen möglich ist, ohne das Gütekriterium der Zumutbarkeit zu verletzen. Für ein multidimensionales

Persönlichkeitsinventar ist jedoch eine Vielzahl von Items notwendig, um die Aussagekraft in den einzelnen Subskalen zu erhöhen.

Zudem stellt die Entwicklung von Distraktoren eine besondere Herausforderung dar. Je mehr Distraktoren vorgegeben werden, desto geringer ist die Ratewahrscheinlichkeit (Wahrscheinlichkeit des zufälligen Auffindens der richtigen Lösung). Kriterien für die Güte von Distraktoren sind ihre Attraktivität (Auswahlwahrscheinlichkeit), ihre Ähnlichkeit mit der richtigen Antwortmöglichkeit (Attraktor) und ihre Plausibilität (Jonkisz & Moosbrugger, 2007). Lechner (2008) beschreibt in seiner Studie zu Sinnhaftigkeitsprüfung von Multiple Choice Items, wie anhand der Fähigkeit zum logischen Schlussfolgern, einzig aufgrund der Betrachtung der Attraktoren und Distraktoren (unter Ausblendung des Itemstamms) die richtige Lösung bei Leistungstests mit einer hohen Wahrscheinlichkeit erkannt werden kann. Eine qualitativ hochwertige Konstruktion von Distraktoren ist daher mit zahlreichen Schwierigkeiten verbunden.

5.8.2 Itemstruktur WSP

Um diese möglichen Fehlerquellen weitestgehend auszuschließen, setzte die Itementwicklung des WSP eine andere Variante situativer Fragen um. Diese unterscheidet sich von den oben genannten dadurch, dass sowohl die Situationsangabe als auch die Verhaltensweise im Itemstamm enthalten sind. Die befragte Person drückt anhand eines mehrkategorialen Antwortformates ihre Zustimmung oder Ablehnung dazu aus. Dadurch wird das Risiko von Reaktanzphänomen reduziert und die Schwierigkeit der Entwicklung geeigneter Distraktoren ausgeschaltet. Anhand mehrerer Beispielitems der Subskala *Selbstbewusstsein* sei dies dargestellt:

Itembeispiele der Subskala *Selbstbewusstsein*:

Ich vermeide es in Vorlesungen durch mehrmaliges Nachfragen aufzufallen.

Ja *eher ja* *eher nein* *nein*

Wenn dem Lehrveranstaltungsleiter meiner Ansicht nach ein Fehler unterläuft, mache ich ihn darauf aufmerksam.

Ja *eher ja* *eher nein* *nein*

Bei Gruppenarbeiten äußere ich meine Meinung, selbst wenn ich damit anderen widerspreche.

Ja *eher ja* *eher nein* *nein*

Ich melde mich ungerne freiwillig, wenn eine Aufgabe an der Tafel zu bearbeiten ist.

Ja *eher ja* *eher nein* *nein*

Bei mündlichen Prüfungen trete ich selbstsicher auf.

Ja *eher ja* *eher nein* *nein*

Die Situationen sind in den vorliegenden Beispielen eine *Vorlesung*, der *Fehler eines/einer LehrveranstaltungsleiterIn*, eine *Gruppenarbeit*, die *Aufforderung an die Tafel zu kommen* und eine *mündliche Prüfung*. Die Verhaltensweisen sind: *vermeiden aufzufallen*, *den/die LehrveranstaltungsleiterIn auf den Fehler aufmerksam zu machen*, *die eigene den anderen widersprüchliche Meinung zu äußern*, *sich freiwillig zu melden* und *selbstsicher aufzutreten*.

Ein weiterer Unterschied zu den klassischen situativen Fragen nach Latham et al. (1980) und Schuler (2002) besteht darin, dass die Situationsbeschreibungen in der sprachlichen Formulierung zwar einen hypothetischen Charakter verfolgen, aber im Indikativ verfasst sind. Die zugrunde liegende Überlegung für diese Darstellungsform der Fragen ist die Reduktion der Komplexität der Itemformulierung und die Annahme, dass sich auch Personen, welche sich noch nicht in einer derartigen Situation befunden haben, sich hypothetisch darin hineinversetzen können. Es wird nahe gelegt, diese Annahme in Folgeuntersuchungen zu prüfen. Der Großteil der Items bezieht sich jedoch auf Situationen, die sowohl StudienbeginnerInnen als auch StudentInnen, die Beratung in Bezug auf einen Studienwechsel suchen, bekannt sind.

5.9 Durchführung

In einem ersten Schritt wurden für jede der 22 Subskalen die spezifischen Situationsbeschreibungen und Verhaltensweisen in Exceltabellen gesammelt. Dadurch entstanden je Subskala umfangreiche Beschreibungen von kritischen Situationen und Verhaltensoptionen.

Die Situationsbeschreibungen und Verhaltensweisen liefern die Inhalte für die Items, d.h. die Situationen und Verhaltensweisen werden in Items umgewandelt. Die Vorgehensweise bei der Umwandlung der Daten aus der Anforderungsanalyse in Items ist in Box 7 dargestellt.

Box 7: Beispiel für die Umwandlung der Daten der Anforderungsanalyse in Items der Subskala *Planen* und zwei Items der Subskala *Gewissenhaftigkeit*. Die Daten stammen aus der Fragebogenerhebung der Studienrichtung Rechtswissenschaften .

Situation	günstige Verhaltensweisen - konkrete Nennungen	Kodierung	daraus geschlossene Eigenschaft
Teildiplomprüfung "Einführung in die Rechtswissenschaften und ihre Methoden"	lange im Voraus "Lernplan" erstellen, jeden Tag ein gewisses Pensum an Lernstoff erfüllen, sich dabei aber nicht die Ziele zu hoch stecken, da es sonst sehr frustrierend ist	O, A, A	Planen, Gewissenhaftigkeit, realistisches Anspruchsniveau

Für die Subskala *Planen* wurde folgendes Item aus dieser Situations- und Verhaltensbeschreibung erstellt:

Bei umfangreichen Prüfungen erstelle ich mir einen genauen Lernplan (Einteilung des Prüfungsstoffes in Lerneinheiten).

Ja eher ja eher nein nein

Für die Subskala *Gewissenhaftigkeit* wurden folgende zwei Items erstellt:

Wenn ich einen Zeitplan erstelle, halte ich mich daran.

Ja eher ja eher nein nein

Wenn ich mir pro Lerntag eine bestimmte Lernmenge (z.B. Seitenanzahl) vorgenommen habe, dann halte ich mich daran.

Ja eher ja eher nein nein

Das Merkmal *realistisches Anspruchsniveau*, welches sich auf die Verhaltensweise „nicht zu hohe Ziele stecken“ bezieht, bildet keine Subskala des WSP. Das Merkmal wurde aufgrund des Kriteriums „hinreichende Relevanz des Merkmals für die Fragestellung (Häufigkeit der Nennungen)“ ausgeschlossen.

5.9.1 Instrumente zur Qualitätssicherung

Die Fragestellung der Diplomarbeit ist, wie die Qualität bei der Entwicklung von Items sichergestellt werden kann. Dieses Unterkapitel stellt drei Instrumente zur Qualitätssicherung vor, die im Rahmen der Itementwicklung eingesetzt wurden.

Die Anwendung des ‚*Regelkatalog zur Fragebogen-Itemkonstruktion*‘ (siehe Unterkapitel «Ergebnis Regelkatalog») wird dargestellt. Zusätzlich werden die Items auf ihre Entsprechung gemäß dem Regelkatalog hin überprüft. Die Items werden in einer *Internetplattform* mehreren ExpertInnen vorgelegt. Ihre Aufgabe ist die Überprüfung der Items anhand mehrerer Fragestellungen und das Üben konstruktiver Kritik. Entsprechend der Kritik werden die Items überarbeitet und erneut der Prüfung unterzogen. Insgesamt wurden dadurch mehrfache Kontrollschleifen eingebaut.

Jene Items, welche trotz Überarbeitung nicht den Qualitätskriterien entsprechen, aber aufgrund ihrer Iteminhalte in den Fragebogen aufgenommen werden sollten, werden im Rahmen von *Workshops* in einer Arbeitsgruppe diskutiert und revidiert bis ein Konsens erzielt wird. Wenn kein Konsens erreicht werden kann, werden die Items aus dem Itempool ausgeschlossen.

5.9.1.1 Anwendung des ‚Regelkataloges zur Fragebogen-Itemkonstruktion‘

Die Items werden entlang des ‚*Regelkataloges zur Fragebogen-Itemkonstruktion*‘ erstellt. D.h. die Items müssen den Kriterien des Regelkataloges entsprechen. Der komplette Regelkatalog ist im Unterkapitel «Ergebnis Regelkatalog» dargestellt.

5.9.1.2 Internetplattform

Zur Sicherstellung der Qualität der Items wurden in den Itemerstellungprozess webbasierte Feedback- und Kontrollschleifen eingebaut. Die Items wurden mehreren ExpertInnen zur Begutachtung vorgegeben. Dazu wurde eigens eine Internet-Plattform eingerichtet, auf die die ausgewählte ExpertInnengruppe Zugriff hatte.

Die ExpertInnengruppe setzte sich aus sechs MitarbeiterInnen des Instituts für Entwicklungspsychologie und Diagnostik (Arbeitsbereich Diagnostik, Fakultät für Psychologie, Universität Wien) und elf DiplomandInnen, die ihre Diplomarbeit am Arbeitsbereich für Diagnostik verfassen, zusammen. Für jene DiplomandInnen, die noch keine Vorerfahrungen in der Itemkonstruktion mitbrachten, wurden Hintergrundinformationen

zu den den Items zugrunde liegenden Daten und eine Anleitung zur Itemkommentierung zur Verfügung gestellt (siehe Anhang). Inhaltlich gliederte sich die Plattform in zwei Bereiche.

Der erste Bereich stellte ein allgemeines Diskussionsforum zur Verfügung. Hier waren der Regelkatalog, die Anleitung zur Itemkommentierung und die Arbeitsdefinitionen der Skalen zugänglich. Außerdem bot es Raum für Diskussionen über Fragestellungen, die in der Itemkommentierung wiederholt aufgeworfen wurden und einer gesammelten Diskussion bedurften.

Der zweite Bereich gliederte sich in die Skalen und Subskalen des WSP. Zu jeder Subskala wurden die neu entwickelten Items von der Autorin gepostet und zur interaktiven Diskussion gestellt. Auf diese Weise konnten die Items parallel von verschiedenen BeurteilerInnen kritisiert und kommentiert werden. Die ExpertInnengruppe überprüfte die Items anhand des entwickelten Regelkataloges. Zusätzlich zum Regelkatalog leiteten folgende Fragen die konstruktive Kritik an:

1. Steht die angegebene Dimension im Vordergrund?
D.h. misst ein Item der Subskala Zeitmanagement wirklich Zeitmanagement oder eine andere Dimension? Wenn ja welche?
Eindimensionalität ist schwierig zu erreichen. Wenn ein Item mehr als eine Dimension zu messen scheint, z.B. Zeitmanagement und Planen -> Welche Dimension steht im Vordergrund? (hierzu siehe „Arbeitsdefinition der Merkmale“)
2. Stimmt die Formulierung des Items?
Sind die Begriffe zu kompliziert? Gibt es zu viele Nebensätze? Könnte man es einfacher ausdrücken? Ist es missverständlich? (Kriterien siehe Regelkatalog) Wenn ja, wie würdet ihr das Item umformulieren?
3. Versetzt euch in die Testperson hinein!
Fällt es euch schwer, die Frage zu beantworten? Wenn ja, wieso?
Kann ich mich eindeutig für eine Antwort entscheiden. Wenn nein, wieso nicht?
Ist es ein Item, bei dem jede/r zusagen oder es negieren würde? -> Kein Informationsgewinn?

Jedes Item wurde mehrfach von unterschiedlichen Personen kommentiert. Dazu werden beispielhaft Itemkommentierungen zu den Subskalen *Zeitmanagement* und *Eigeninitiative* in Box 8 dargestellt.

Box 8: Kommentierung von Itemvorschlägen durch die ExpertInnengruppe
(Auszug aus der Internetplattform)

Item der Subskala *Zeitmanagement*:

- Ursprungitem: An einem lang andauernden Projekt plane ich grundsätzlich zu wenig Zeit für die einzelnen Arbeitsschritte ein.
- Kommentar 1: Ich glaube ehrlich gesagt nicht, dass jemand bei diesem Item wirklich zustimmen würde. Als Vorschlag: Von Anfang an den Zeitaufwand einzelner Arbeitsschritte abschätzen zu können, fällt mir leicht.
- Kommentar 2: Ich finde auch, dass man da schwer zustimmen kann, und das führt uns aber eher schon in die Grundsatzdiskussion zum Wort „grundsätzlich“, denn zumindest im Zusammenhang mit dem Wort „planen“ drückt „grundsätzlich“ für mich eine Absicht aus (das mach ich grundsätzlich so!) und nicht eine Gewohnheit., leider hier kein Alternativvorschlag...
- Kommentar 3: Vorschläge: Es fällt mir leicht/schwer, den Zeitaufwand der einzelnen Arbeitsschritte eines langwierigen Projektes abzuschätzen. Bzw. Für die einzelnen Arbeitsschritte eines lang andauernden Projektes plane ich meistens zu wenig Zeit ein.
- Kommentar 4: Die Frage ist, ob man wahrnimmt, dass man für einzelne Arbeitsschritte zu wenig Zeit einplant oder am Ende einfach nur merkt, dass man nicht fertig geworden ist. Sollten wir nicht besser danach fragen, ob jemand im Rahmen eines Projektes Termine einhalten kann oder einzelne Arbeiten termingerecht fertig stellen kann oder das Gefühl hat, sich gegen Ende sehr „stressen“ zu müssen, da die Zeit knapp wird?
- Überarbeitung: Die Fragestellungen, ob jemand termingerecht fertig wird, ist bereits in anderen Items umgesetzt. Der Schwerpunkt der Frage liegt auf dem Abschätzen einzelner sequentieller Arbeitsschritte in einem Arbeitsprozess.
Aufgrund der mehrfachen Kritikpunkte am Wort „grundsätzlich“ und seiner Mehrdeutigkeit wird von dem Begriff Abstand genommen.
Die Frage nach dem Wort langwierig ist berechtigt. Ich halte den Begriff für emotional negativ besetzt. Ich schlage daher den Begriff langfristig vor.
- Alt:** An einem lang andauernden Projekt plane ich grundsätzlich zu wenig Zeit für die einzelnen Arbeitsschritte ein.
Überarbeitet: Es fällt mir leicht, den Zeitaufwand einzelner Arbeitsschritte bei einem langfristigen Projekt abzuschätzen.
- Kommentar 5: Passt von meiner Seite
- Kommentar 6: Ok
- Kommentar 7: Ok
- Endversion: **OK:** Es fällt mir leicht, den Zeitaufwand einzelner Arbeitsschritte bei einem langfristigen Projekt abzuschätzen.

Item der Subskala *Eigeninitiative*:

Ursprungsitem: Ich suche aktiv nach Kritik, um meine Arbeit zu verbessern.

Kommentar 1: Ich brauche lange, um mir unter „aktiv Kritik suchen“ etwas vorstellen zu können. Vielleicht wäre in diesem Zusammenhang „Feedback“ besser. Ich nehme an, dass mit aktiv Kritik suchen gemeint ist, jemanden um seine Meinung zu fragen? Kann man gleich diese Formulierung verwenden? Ansonsten schlage ich vor: Ich suche aktiv nach Feedback, um meine Arbeit zu verbessern.

Kommentar 2: Ich muss mich der Kritik von x anschließen und würde auch sagen, dass Feedback besser passt bzw. man sich mehr darunter vorstellen kann.

Kommentar 3: Ja, entweder „Feedback“ oder „konstruktive Kritik“, wobei Feedback sicher positiver ist und damit für das Item besser passt.

Überarbeitung: Ich schließe mich dem Einwand an. Feedback ist positiver besetzter und leichter nachvollziehbar.

Alt: Ich suche aktiv nach Kritik, um meine Arbeit zu verbessern.

Überarbeitet: Ich suche aktiv nach Feedback, um meine Arbeit zu verbessern.

Kommentar 4: Ok

Kommentar 5: Ok

Kommentar 6: Ok

Endversion: **OK:** Ich suche aktiv nach Feedback, um meine Arbeit zu verbessern.

Erst wenn ein Itemvorschlag mindestens drei Kritiken beinhaltet, wurde das Item von der Autorin anhand der Kommentare überarbeitet. Voraussetzung dafür war, dass die Autorin bei der Umformulierung oder Neuentwicklung Stellung dazu nahm, welche Kritikpunkte bei der Umformulierung berücksichtigt werden und welche nicht (mit Begründung). Nach der Überarbeitung wurde das Item mit **Ü** (für überarbeitet) neu kodiert und erneut zur Diskussion gestellt. Es durchlief erneut eine Feedback- und Korrekturschleife. Die alte Version des Items und die ursprünglichen Kommentare blieben erhalten und wurden mit **A** (für alt) kodiert. Damit blieb der durchlaufene Entwicklungsprozess der Items nachvollziehbar.

Das Qualitätskriterium zur Endabnahme eines Items wurde vorab mit einer Zustimmung durch mindestens drei ExpertInnen (z.B. Kommentar: **OK**) festgelegt.

5.9.1.3 Workshop

Einige Items konnten trotz zweifacher Überarbeitung nicht den Qualitätskriterien genügen. Jene kritischen Items, deren Inhalte bereits durch andere Items abgedeckt wurden, wurden

archiviert und nicht weiter verwendet. Einzelne Items enthielten jedoch Situationen und Verhaltensweisen, die noch nicht von anderen Items erfragt wurden. Diese Items wurden gesammelt und in zweimalig stattfindenden Workshops diskutiert. Dazu trafen sich jeweils zehn der ExpertInnengruppe mit der Autorin. Die Themeninhalte der Workshops waren einerseits die Grenzen des Regelkataloges und dessen Ergänzung und andererseits die Überarbeitung kritischer Items. Ein Großteil der Item-Kritikpunkte bezog sich auf

- die inhaltliche Zuordnung der Items zu den Subskalen (Eindimensionalität) und Abgrenzung der Merkmale. Besondere Schwierigkeit stellten jene Iteminhalte (Situationen und Verhaltensweisen) dar, die in der Anforderungsanalyse mit zwei oder mehr Eigenschaften kodiert wurden.
- die Nachvollziehbarkeit der Situationen und Verhaltensweisen für die Zielgruppen. Es wurde die Nachvollziehbarkeit der Situationen für MaturantInnen ohne Studienerfahrung diskutiert. Auch wurde sichergestellt, dass nur jene Situationen in Items integriert wurden, welche in allen Studienrichtungen vorkommen.
- die umständliche Formulierung aufgrund einer notwendigen genauen Situationsbeschreibung.

In der Diskussion gelang es ca. zwei Drittel der kritischen Items umzuformulieren und zu verwerten.

5.9.2 Probleme bei der Itementwicklung und Ergänzung des Regelkataloges

Im Unterkapitel «Workshops» wurden bereits drei Schwierigkeiten während des Itementwicklungsprozesses angesprochen. Diese bezogen sich auf die Nachvollziehbarkeit der Situations- und Verhaltensbeschreibungen für die Zielgruppen (insbesondere der StudienbeginnerInnen), die Komplexität der Formulierung aufgrund der genauen Situationsbeschreibung zur Gewährleistung der Eindeutigkeit sowie die inhaltliche Zuordnung der Items zu den Subskalen.

Um den Anforderungen der Nachvollziehbarkeit und Eindeutigkeit des Itemmaterials zu entsprechen (Regel 4 und 10 des Regelkataloges), wurden die Situationsbeschreibungen aus den Critical Incident Daten derart verallgemeinert, dass sie für die gesamte Zielgruppe nachvollziehbar sind, und gleichzeitig so spezifisch gestaltet, dass eine Mehrdeutigkeit der Interpretation des Iteminhaltes weitestgehend ausgeschlossen werden konnte.

Beispielsweise wurde die Situation Mechanik 1 Prüfung allgemein mit der Bezeichnung Prüfung übersetzt. Die Sprechstunde von Prof. XY wurde geändert in eine Sprechstunde oder in ein persönliches Gespräch mit einer Lehrperson. Es sind jedoch nicht alle Situationen

StudienbeginnerInnen aus ihrem Schulalltag bekannt, auch wenn der Großteil der Items sich auf Situationen bezieht, die sowohl StudienbeginnerInnen als auch StudentInnen, die Beratung in Bezug auf einen Studienwechsel suchen, im Laufe ihrer schulischen Karriere erfahren haben.

Es wird daher geraten, bei der Normierung des WSP darauf zu achten, ob Personen ohne Studienerfahrung Schwierigkeiten bei der Beantwortung von Items aufgrund mangelnden Vorwissens haben.

Schwierigkeiten bereitete ebenso die Notwendigkeit der Eindimensionalität von Items (Regel 8 des Regelkataloges), weil einzelne Verhaltensweisen in der Anforderungsanalyse mit mehreren Merkmalsdimensionen kodiert wurden. Beispielsweise wurde die Verhaltensweise „Ohne aufgefordert worden zu sein zu irgendwelchen Themen, die in der Lehrveranstaltung besprochen wurden, Fragen stellen oder selbst Informationen dazu in die Lehrveranstaltung mitbringen“ mit *Eigeninitiative* (A), *Engagement* (M), und *Wissensdurst/Neugier* (M) kodiert. Daraus ergab sich die Schwierigkeit, dass einige Items nicht eindeutig einer Subskala zugeordnet werden konnten. Die Items wurden folglich jener Subskala zugeordnet, der sie am ehesten entsprachen mit dem Hinweis ihre Passung in einer weiterführenden Analyse zu prüfen. In einer anschließenden Analyse der Items nach dem (dichotomisierten) logistischen Modell nach Rasch (1960) sollte die Passung der Items zur primären und zu sekundären Skalen überprüft werden. Die Regel 8 des Regelkataloges wurde um den Zusatz 8b ergänzt (siehe Box 9).

Eine sprachliche Variabilität der Itemformulierung erhöht das Commitment und die Motivation während der Aufgabenbearbeitung. D.h. nicht jedes Item sollte mit demselben Wortlaut beginnen. Neben dem Wortlaut „Es fällt mir leicht/schwer...“ können auch noch weitere Darstellungsformen verwendet werden. Auch wenn dies oftmals schwer umzusetzen ist, wird eine Variationen der Itemformulierung empfohlen. Dies könnte beispielsweise sein: „Es fällt mir auf, dass ich...“, „Ich wäre gerne besser darin, ...“, „Es zählt zu meinen Stärken, dass ich...“, „Ich neige dazu,...“, „Ich versuche laufend,...“ Die sprachliche Variabilität der Itemformulierung wurde daraufhin in Regel 23 des Regelkataloges festgehalten.

Unabhängig von diesen standardisierten Satzformulierungen, wurde beim Großteil der Items darauf geachtet, die Formulierung an den Verhaltensweisen selbst auszurichten. Z.B. „Ich suche den Erfahrungsaustausch mit anderen.“ „Ich führe einen Kalender, um wichtige Termine zu notieren.“ oder „Beim Lernen lasse ich mich wiederholt von unbedeutenden Dingen ablenken.“

Besonderes Diskussionspotenzial auf der Internetplattform boten die Begriffe „im Allgemeinen“ und „grundsätzlich“. Während der Begriff „im Allgemeinen“ eine Depersonalisierung bewirkt (betrifft alle und nicht nur die befragte Person), ist das Wort grundsätzlich doppeldeutig. Je nach Betonung kann das Wort im Sinne von „ausschließlich“ oder im Sinne von „grundsätzlich ja, aber es gibt Ausnahmen von der Regel“ verwendet werden. Schlussendlich wurde von beiden Varianten Abstand genommen. Nur in Einzelfällen und nach kritischer Prüfung der Interpretationsbreite wurde das Wort „grundsätzlich“ in die Items integriert. Dies wurde im Regelkatalog unter dem Punkt 22 vermerkt.

Box 9: Ergänzung des ‚Regelkataloges zur Fragebogen-Itemkonstruktion‘
um Regel 8b, 22 und 23

- 8) Eindimensionalität, d.h. das Item soll einer einzigen Skala eindeutig zuordenbar sein
- 8b) Ad Eindimensionalität, Einschränkung dieser Bedingung während des Arbeitsprozesses der Itemgenerierung: Aufgrund der Verwendung der Critical Incident Daten kann bei einigen Items die Trennung der Dimensionen nicht aufrechterhalten werden. Für eine Verhaltensweise sind z.T. mehrere Eigenschaften notwendig. Wenn die Eindimensionalität nicht eingehalten werden kann, ist zunächst anzugeben, welche Dimensionen erhoben werden und zu vermerken, welche davon im Vordergrund steht.
- 22) Die Begriffe „im Allgemeinen“ und „grundsätzlich“ können zu unterschiedlichen Interpretationen auf Seiten der befragten Person führen. „Im Allgemeinen“ legt eine Depersonalisierung nahe (betrifft alle und nicht nur die befragte Person). Das Wort „grundsätzlich“ kann je nach Betonung im Sinne von „ausschließlich“ oder als „grundsätzlich ja, aber Ausnahmen bestätigen die Regeln“ interpretiert werden.
- 23) Die Fragestellungen im Itemstamm sollen variiert werden. Z.B. nicht nur „Es fällt mir schwer/leicht, ...“, sondern auch „Es fällt mir auf, dass ich ...“, „Ich wäre gerne besser darin, ...“, „Ich neige dazu,...“, „Ich versuche laufend,...“

5.10 Itemverrechnung

Die Items, welche den Qualitätskriterien entsprachen, wurden in den Itempool aufgenommen. Um die Items als psychologisch-diagnostisches Verfahren vorzugeben, ist die Verrechnung der Items anzugeben, d.h. welche Antwortkategorien sind als Lösung (mit 1 kodiert) zu werten.

Aufgrund der anschließenden Überprüfung mit dem (dichotom) logistischen Modell nach Rasch (1960) ist die Vorgabe der Items mit einem vier-kategoriellen Antwortformat mit einer dichotomisierten Verrechnung (0 und 1) geplant. Die dichotomisierte Itemverrechnung sei an drei Items exemplarisch dargestellt:

Items der Subskala <i>Genauigkeit</i> (Arbeitsstil und Arbeitshaltungen)	ja	eher ja	nein	eher nein
Ich investiere viel Zeit in die Formatierung meiner Arbeiten.	1	1	0	0
Ich neige zu Flüchtigkeitsfehlern beim Schreiben und/oder Rechnen.	0	0	1	1
Nach Möglichkeit vermeide ich detailreiche Aufgaben, die Präzision erfordern.	0	0	1	1

Tab. 10 Itemverrechnung

Im ersten Beispiel wären die zustimmenden response options (☐ Ja und ☐ eher ja) mit 1 (Lösung) kodiert und die ablehnenden response options (☐ nein und ☐ eher nein) mit 0 kodiert. Das zweite und dritte Beispielitem hat eine gegensätzliche Verrechnung. Die zustimmenden response options sind mit null kodiert und die ablehnenden mit eins (Lösung).

Der Begriff Lösung bedeutet in diesem Falle jedoch nicht, dass es bei den Items eine richtige oder falsche Antwort gibt, sondern ob eine response options für die Ausprägung des zu erhebende Merkmal spricht oder nicht.

5.11 Ergebnis Itempool

Zu den sieben Skalen und 22 Subskalen wurden insgesamt 351 Items entwickelt. Tab. 11 zeigt eine Auflistung der Itemanzahl je Subskala.

Aufgrund der unterschiedlichen Fülle und inhaltlichen Qualität der Daten aus den Fragebögen und Interviews der Anforderungsanalyse je Subskala, ist die Varianz der

Itemanzahl in den einzelnen Subskalen beträchtlich. Die Subskala mit der geringsten Anzahl (Informiertheit) enthält lediglich neun Items. Die größte Subskala (Gewissenhaftigkeit) enthält 25 Items. Aufgrund des Überwiegens positiver Verhaltensweisen in den Critical Incident-Nennungen sind mehr Items positiv gepolt (n=206) als negativ (n=145).

Skala	Subskala	Positive Itempolung	Negative Itempolung	Anzahl Items (gesamt)
Organisation (78 Items)	Zeitmanagement	16	8	24
	Prioritäten setzen	5	8	13
	Planen	11	6	17
	Organisationstalent	10	5	15
	Informiertheit	6	3	9
Arbeitsstil und Arbeitsverhalten (72 Items)	Gewissenhaftigkeit	20	5	25
	Genaueres Arbeiten	13	6	19
	Eigeninitiative	11	5	16
	Selbstständiges Arbeiten	7	5	12
Motivation (94 Items)	Durchhaltevermögen	6	6	12
	Lernbereitschaft	7	9	16
	Engagement	15	3	18
	Zielorientierung	10	8	18
	Zielsetzung	8	4	12
	Erfolgsoptimismus	6	12	18
Belastbarkeit (21 Items)	Frustrationstoleranz	8	13	21
Soziale Kompetenz (36 Items)	Kommunikationsfähigkeit	4	5	9
	Kontaktfähigkeit	6	3	9
	Teamfähigkeit (Teamfähigkeit, Teamorientierung, Kollegialität, Kompromissbereitschaft)	11	7	18
Emotionale Kompetenzen (32 Items)	Distanzierungsfähigkeit	6	9	15
	Offenheit	13	4	17
Selbstbild (18 Items)	Selbstbewusstsein	7	11	18
Gesamt		206	145	351

Tab. 11 Itemanzahl nach Skalen und Subskalen des WSP

Es ist davon auszugehen, dass aufgrund der Modellgültigkeitsprüfung nach der probabilistischen Testtheorie Items, welche nicht dem Modell nach Rasch (1960) entsprechen, aus dem Itempool iteriert (ausgeschlossen) werden müssen. Die Zusammensetzung des endgültigen Itempools und die Skalenstruktur des WSP werden sich daher erst nach Prüfung der Itemqualität mit dem dichotom logistischen Modell nach Rasch zeigen. Ziel war es vorerst, einen möglichst großen Itempool zu erstellen, der sämtliche Daten aus der Anforderungsanalyse verwertet.

In Tab. 12 sind exemplarisch für jede der 22 Subskalen ein Beispielitem angeführt. Der vollständige Itempool (n=351) inklusive Verrechnung befindet sich im Anhang.

Skalen und Subskalen	Items	Verrechnung			
Organisationsfähigkeit					
Zeitmanagement	Ich teile mir den Lernstoff vor einer Prüfung so ein, dass am Ende genug Zeit für eine Wiederholung des Lernstoffs bleibt.	1	1	0	0
Prioritäten setzen	Beim Lernen lasse ich mich wiederholt von unbedeutenden Dingen ablenken.	0	0	1	1
Planen	Ich entscheide zu Beginn des Semesters eher spontan, welche Lehrveranstaltung ich besuche.	0	0	1	1
Organisations-talent	Bevor ich mit dem Schreiben einer Arbeit beginne, überlege ich mir den strukturellen Aufbau der Arbeit.	1	1	0	0
Informiertheit	Vor Studienbeginn habe ich mich intensiv mit dem Studienplan auseinandergesetzt.	1	1	0	0
Arbeitsstil und Arbeitshaltungen					
Gewissenhaftigkeit	Ich achte darauf, dass meine Mitschrift vollständig ist.	1	1	0	0
Genaueres Arbeiten	Ich lese Texte genauer durch als die meisten anderen, damit mir wichtige Details nicht entgehen.	1	1	0	0
Eigeninitiative	Bei der Themenauswahl in Arbeitsgruppen halte ich mich mit Vorschlägen eher zurück.	0	0	1	1
Selbstständiges Arbeiten	Auch bei kleineren Problemen ist für die mich die Meinung anderer unverzichtbar.	0	0	1	1
Motivation					
Durchhalte-vermögen	Wenn erforderlich, lerne ich für Prüfungen von morgens bis abends.	1	1	0	0
Lernbereitschaft	Solange meine eigene Vorgehensweise funktioniert, lerne ich ungern eine neue (eventuell effizientere).	0	0	1	1
Engagement	Auf Prüfungen lerne ich gerade genug, um sie bestehen zu können.	0	0	1	1
Zielorientierung	Wenn ich mir etwas vorgenommen habe, bringe ich es auch zu Ende.	1	1	0	0
Zielsetzung	Ich habe eine konkrete Vorstellung davon, was ich während meines Studiums erreichen möchte.	1	1	0	0
Erfolgszuversicht	Vor neuen Aufgaben überkommen mich Zweifel, ob ich diese bewältigen kann.	0	0	1	1
Belastbarkeit					
Frustrations-toleranz	Ich arbeite unbeirrt weiter, auch wenn mich zwischendurch der Wunsch überkommt, alles hinzuschmeißen.	1	1	0	0
Soziale Kompetenz					

Kommunikationsfähigkeit	Es fällt mir eher schwer, andere Personen in einer Diskussion ausreden zu lassen.	0	0	1	1
Kontaktfähigkeit	Ich knüpfe sehr schnell neue Bekanntschaften mit Studienkollegen.	1	1	0	0
Teamfähigkeit	Es fällt mir grundsätzlich leicht, gemeinsam mit anderen an einem Projekt zu arbeiten. = <i>Teamfähigkeit</i>	1	1	0	0
	Wenn ich einen Prüfungsplatz nicht in Anspruch nehme, melde ich mich rechtzeitig ab. = <i>Kollegialität</i>	1	1	0	0
	In Gruppenarbeiten fällt es mir schwer, von meiner Meinung abzukommen und Kompromisse zu schließen. = <i>Kompromissbereitschaft</i>	0	0	1	1
	Die Arbeit in einer Gruppe macht mir mehr Spaß als die Arbeit alleine. = <i>Teamorientierung</i>	1	1	0	0
Emotionale Komponenten					
Distanzierungsfähigkeit	Nach Erhalt eines negativen Feedbacks hindern mich die Gedanken daran einzuschlafen.	0	0	1	1
Offenheit	Ich suche Gespräche mit Personen, die eine andere Auffassung haben als ich.	1	1	0	0
Selbstbild					
Selbstbewusstsein	Richtet ein Lehrveranstaltungsleiter vor versammelter Gruppe eine Frage an mich, werde ich nervös.	0	0	1	1

Tab. 12 Itembeispiele nach Skalen und Subskalen des WSP

6 Diskussion

Aufgabe der Diplomarbeit war es, einen Itempool für das Wiener Studieneignungs-Persönlichkeitsinventar (WSP) zu entwickeln. Die Entwicklung schloss dabei mehrere Instrumente zur Qualitätssicherung mit ein.

Zentrales Element der Itemkonstruktion war die Verknüpfung der Iteminhalte mit den zugrunde liegenden Daten aus der Anforderungsanalyse. Die Daten der Anforderungsanalyse wurden nach der Methode der Critical Incident Technique (Flanagan, 1954) erhoben und beinhalten die Benennung erfolgskritischer Situationen und Verhaltensweisen im Studienalltag. Durch die Verwendung dieser Daten konnte der Praxisbezug der Items erhöht werden. Die Items wurden spezifisch auf die Zielgruppe und die zu beantwortende Fragestellung zur Studieneignung zugeschnitten. Die Nachvollziehbarkeit zwischen den Elementen der Anforderungsanalyse und der Itemkonstruktion (Situationen – Verhaltensweisen – Anforderungen – Iteminhalte) blieb erhalten, was darüber hinaus zur Erhöhung der Validität beitragen kann.

Vor der Entwicklung der Items wurde des Weiteren ein ‚Regelkatalog zur Fragebogen-Itemkonstruktion‘ erstellt. Dieser fasst die wesentlichen Regeln und Handlungsanleitungen zur Fragebogen-Itemkonstruktion nach aktuellem Wissensstand zusammen. Damit wurde ein Instrument entwickelt, das kurz und prägnant die wichtigsten Hinweise aus Forschung und Literatur in dreiundzwanzig Regeln formuliert.

Die Idee des Regelkataloges zur Itemkonstruktion wurde von Edwards 1957 (zit.n. Mummendey, 2008) zum ersten Mal umgesetzt. Der 1957 konzipierte Regelkatalog wird jedoch den neuesten Entwicklungen in der Testtheorie – insbesondere der Probabilistischen Testtheorie (Item-Response-Theorie) – und dem heutigen methodischen Anspruch nicht mehr gerecht.

Der in dieser Arbeit entwickelte Regelkatalog enthält Anleitungen zu den Themen der sprachlichen Verständlichkeit, inhaltlichen Eindeutigkeit, Itempolung, Itemschwierigkeit, Eindimensionalität, Antwortformat und zur sprachlichen Variabilität.

Das Ziel, ein kompaktes Instrument zu entwickeln, steht im Gegensatz zur Informationsdichte im Bereich der Fragebogenkonstruktion. Der Regelkatalog stellt daher nicht den Anspruch auf Vollständigkeit. Es ging vielmehr darum, die zentralen Regeln und Anleitungen zusammenzutragen und ein Basisinstrument zu Verfügung zu stellen, das als Hilfestellung bei der Itemerstellung als auch zur Überprüfung von Items dienen soll. Bei der

Anwendung im Rahmen eines anderen methodischen Vorgehens als dem hier vorgestellten kann es notwendig sein, die Regeln der jeweiligen Problemstellung anzupassen.

Im Anschluss an den Regelkatalog wurde der Itempool des WSP konstruiert. Dabei wurden mehrfache Rückkopplungsschleifen zur Überprüfung der Items eingebaut. Diese Feedback-Schleifen durch ExpertInnen fanden sowohl webbasiert als auch im direkten Austausch (Workshops) statt. Durch diese Vorgehensweise konnten die neu entwickelten Items gemäß dem Regelkatalog begutachtet und bei Widerspruch mit einer der Regeln kritisiert werden. Nach Überarbeitung der Items wurde eine erneute Rückkopplungs- und Korrekturschleife gezogen. Die Endabnahme von Items wurde vorab definiert und erreicht, wenn mind. drei Personen die Letztversion der Items guthießen.

Ergebnis ist der Itempool des Wiener Studieneignungs-Persönlichkeitsinventars (WSP). Er besteht aus 351 Items zu sieben Skalen und 22 Subskalen.

Im Anschluss an die Itemkonstruktion sind weiterführende Untersuchungen zur Überprüfung der Güte des Itempools geplant. Eine dieser Nachfolgeuntersuchungen wurde parallel zur Dokumentation dieser Arbeit von Dörrer (2009) durchgeführt. Dörrer (2009) überprüfte den Itempool des WSP auf Gültigkeit nach dem „Rasch-Modell“. Das (dichotom) logistische Modell nach Rasch (1960) basierend auf der probabilistischen Testtheorie ist im Gegensatz zur klassischen Testtheorie ein prüfbares Modell. Gültigkeit nach dem (dichotom) logistischen Modell nach Rasch (1960) bedeutet, dass die Itembeantwortung einzig auf die Parameter Personenfähigkeit ξ_v und Itemschwierigkeit σ_i (und den Zufall) zurückzuführen sind. Der Itempool wurde von Dörrer (2009) in zwei Versionen (gleichverteilte Itemreihenfolge versus inhaltshomogene Itemblockbildung) 520 Personen vorgegeben und in jeder Teilstichprobe wurden Modelltests (Likelihood-Ratio-Test von Anderson, 1974) und graphische Modellkontrollen durchgeführt.

Eindimensionalität (im Sinne von Konformität mit dem „Rasch-Modell“) konnte in beiden Stichproben auf Ebene der Skalen festgestellt werden, weshalb weitere Modellkontrollen auf Ebene der Subskalen nicht durchgeführt wurden. Im Zuge der Modellüberprüfung wurden unter der Bedingung „gleichverteilte Itemreihenfolge“ von den 351 Items 7 Items ausgeschlossen und unter der Bedingung „inhaltshomogene Itemblockbildung“ 22 Items. Das ist verglichen mit der Gesamt-Itemmenge eine relativ geringe Anzahl, was für eine hohe Qualität auf Itemebene spricht.

Aus dieser sehr umfangreichen Itemmenge sind nach deskriptivstatistischer Itemanalyse diejenigen auszuwählen, die am geeignetsten für die Testendversion des WSP erscheinen.

Aufgrund des genügend großen Itempools sind auch unterschiedliche Versionen des WSP (z.B. differenziert nach Studienrichtungen) denkbar. Neben der Überprüfung der Reliabilität und Validität sollen die Testendform/en des WSP sowohl nach Geschlecht, Alter, Studienerfahrung (MaturantInnen und StudentInnen) und bestimmten Studienrichtungen normiert werden.

Eine Normierung des WSP nach dem Studienfortschritt ist insbesondere in Hinblick auf jene Gruppe der Testpersonen ohne Studienvorkenntnis von Interesse. Durch die Verwendung der Daten aus der Anforderungsanalyse beziehen sich die Iteminhalte auf Situationen und Verhaltensweisen im Studienalltag. Die Items wurden zwar derart entwickelt, dass davon ausgegangen werden kann, dass MaturantInnen die Situationen, welche in den Items beschrieben sind, bereits kennen oder sich zumindest in diese hineinversetzen können. Dennoch wird nahe gelegt, die Iteminhalte auf ihre Bekanntheit für die Zielgruppe der SchulabgängerInnen zu prüfen.

Das Wiener Studieneignungs-Persönlichkeitsinventar (WSP) liefert als Persönlichkeitsfragebogen zur Erfassung studienrelevanter Eigenschaften einen wichtigen Beitrag zur Studienberatung. Ein Großteil der am Markt befindlichen Verfahren konzentriert sich auf die kognitiven Fähigkeiten von StudienbeginnerInnen. Der Bereich der Persönlichkeitseigenschaften, denen im Studium eine entscheidende Rolle zukommt, wurde vergleichsweise vernachlässigt.

Immer mehr SchulabsolventInnen (national und international) drängen an österreichische Hochschulen. Sowohl die Frage nach der persönlichen Studieneignung als auch der Beratungsbedarf bezüglich der Wahl der Fachrichtung werden aufgrund der steigenden Studierendenselektion an Fakultäten und der damit einhergehender Unsicherheit, zukünftig einen Studienplatz zu erhalten, an Bedeutung gewinnen.

Der WSP ist ein multidimensionaler Persönlichkeitsfragebogen, der jene Eigenschaften erhebt, die sich für die Studienbewältigung als erfolgskritisch erwiesen haben. Seine Iteminhalte sind auf den Studienalltag zugeschnitten, wodurch sowohl die Akzeptanz des psychologisch-diagnostischen Verfahrens als auch das Commitment bei der Aufgabenbearbeitung als hoch eingeschätzt werden können.

Die Testvorgabe im Rahmen des webbasierten Self-Assessments des Arbeitsbereichs Diagnostik (Institut für Entwicklungspsychologie und psychologische Diagnostik, Fakultät für Psychologie, Universität Wien) (Frebort & Kubinger, 2006b) soll darüber hinaus vielen InteressentInnen den Zugang zum Verfahren ermöglichen und insgesamt zu einer Verbesserung der Orientierung und Beratung in der österreichischen Studienlandschaft beisteuern.

7 Zusammenfassung

Ziel der Diplomarbeit war die Entwicklung eines Itempools für einen Persönlichkeitsfragebogen zur Erfassung studienrelevanter Persönlichkeitseigenschaften – das Wiener Studieneignungs-Persönlichkeitsinventar (WSP) – unter Einbezug mehrerer Instrumente zur Qualitätssicherung während der Itemkonstruktion.

Das Ergebnis dieser Diplomarbeit ist der Itempool des Wiener Studieneignungs-Persönlichkeitsinventars (WSP), ein ‚Regelkatalog zur Fragebogen-Itemkonstruktion‘ und die Dokumentation möglicher Instrumente zur Qualitätssicherung während des Itemerstellungprozesses.

Ausgangspunkt und erster zentraler Baustein zur Qualitätssicherung der Fragebogenitems war die enge Verschränkung der Itementwicklung mit dem zugrunde liegenden Anforderungsprofil für die universitäre Hochschulausbildung in Wien. Die Anforderungsanalyse zur allgemeinen Studierfähigkeit wurde nach der Methode der Critical Incident Daten (Flanagan, 1954) durchgeführt. Dazu wurden 278 StelleninhaberInnen und Vorgesetzte (StudentInnen und Lehrpersonen) mittels Fragebogenerhebung und Experteninterviews befragt. Die Ergebnisse und Daten der Anforderungsanalyse stellen die Grundlage für die Festlegung der zu erhebenden Merkmale durch den WSP und die Iteminhalte.

Im Anschluss an die Anforderungsanalyse wurde ein Regelkatalog konzipiert, welcher Regeln und Formulierungshinweise zur Itemgenerierung enthält. Dazu wurde eine umfassende Literaturrecherche und Experteninterviews mit zwei Mitarbeitern des Arbeitsbereichs für Diagnostik (Instituts für Entwicklungspsychologie und Diagnostik, Fakultät für Psychologie, Universität Wien) mit einschlägiger Erfahrung in der Konstruktion von Persönlichkeitsfragebögen geführt. Ergänzt wurde der Regelkatalog anhand der Erfahrungen in der empirischen Umsetzung des WSP.

Ein Großteil der Regeln des Regelkataloges war bereits in Lehrbüchern zur Test- und Fragebogenkonstruktion festgehalten. Neu ist jedoch, dass diese umfangreichen Anleitungen in einer aktuellen Version in wenigen Regeln kurz zusammengefasst wurden. Dadurch ist eine Übersicht über empfohlene Regeln zur Itemkonstruktion gegeben und die Qualität der Items und deren Entwicklung werden anhand der formulierten Kriterien leichter nachvollziehbar und überprüfbar.

Der konzipierte Regelkatalog enthält 23 Regeln zu den Themen der sprachlichen Verständlichkeit, inhaltlichen Eindeutigkeit, Itempolung, Itemschwierigkeit, Eindimensionalität der Items, empfohlene Antwortformate und sprachlichen Variabilität.

Im Anschluss daran wurden die Items des WSP in Anlehnung an die situativen Fragen nach Latham et al. (1980) und Schuler (2002) unter Anwendung des Regelkataloges entwickelt. Die Datenlage der Anforderungsanalyse nach der Critical Incident Technique (Flanagan, 1954) legte die Darstellung in Form situativer Fragen nahe. Dadurch war es möglich, die Situations- und Verhaltensbeschreibungen aus den Befragungen der StelleninhaberInnen und Vorgesetzten direkt in die Items zu übernehmen, was den Praxisbezug der Iteminhalte und die Validität der Items (aufgrund der gegebenen Nachvollziehbarkeit zwischen kodierter Verhaltensweise und Merkmalsdimension) erhöht.

In den Itemerstellungprozess wurden webbasierte Feedback- und Korrekturschleifen eingebaut. Die Items wurden insgesamt siebzehn ExpertInnen (MitarbeiterInnen und DiplomandInnen des Arbeitsbereichs Diagnostik) vorgelegt, welche die Items auf ihre Entsprechung gemäß dem Regelkatalog überprüften. Dazu wurde eigens eine Internet-Plattform eingerichtet, auf die die ausgewählte ExpertInnengruppe Zugriff hatte.

Die Items wurden von den TestpsychologInnen entlang des Regelkataloges und zusätzlich formulierter Fragen kommentiert und anschließend durch die Autorin überarbeitet. Die überarbeiteten Items wurden erneut zur Diskussion gestellt. Ebenso wurde ein Kriterium zur Endabnahme von Items festgelegt.

Einige Items konnten trotz zweifacher Überarbeitung nicht den festgelegten Qualitätskriterien genügen. Jene Items, deren Inhalte nicht bereits durch andere Items abgedeckt wurden, wurden gesammelt und in zweimalig stattfindenden Workshops diskutiert. Dazu trafen sich jeweils zehn der ausgewählten ExpertInnengruppe mit der Autorin. Die Themeninhalte der Workshops waren einerseits die Grenzen des Regelkataloges und dessen Ergänzung sowie andererseits die Überarbeitung kritischer Items. In der Diskussion gelang es ca. zwei Drittel der kritischen Items umzuformulieren und zu verwerten.

Das Ergebnis ist ein Itempool bestehend aus 351 Items, die sieben Skalen und 22 Subskalen zugeordnet sind. Die ursprüngliche Zielvorgabe bestand aus 10-20 Items je Subskala (insgesamt 220-440 Items), abhängig von der Datenfülle der Anforderungsanalyse in den einzelnen Merkmalsdimensionen. Um einen möglichst großen Itempool zu erstellen, wurden sämtliche Daten (die eine Kodierung gemäß den Subskalen des WSP aufwiesen) aus der Anforderungsanalyse verwertet.

Aufgrund des unterschiedlichen Datenumfanges (Anzahl von Situations- und Verhaltensbeschreibungen nach der CIT) in den Merkmalsdimensionen setzen sich die kleinste Subskala (Informiertheit) aus neun Items und die größte Subskala aus 25 Items (Gewissenhaftigkeit) zusammen. In Tab. 13 ist die Itemanzahl je Subskala aufgelistet. Aufgrund des Überwiegens verwertbarer positiver Verhaltensweisen in den Critical Incident-Nennungen sind mehr Items positiv gepolt (n=206) als negativ (n=145).

Skala	Subskala	Positive Itempolung	Negative Itempolung	Anzahl Items (gesamt)
Organisation (78 Items)	Zeitmanagement	16	8	24
	Prioritäten setzen	5	8	13
	Planen	11	6	17
	Organisationstalent	10	5	15
	Informiertheit	6	3	9
Arbeitsstil und Arbeitsverhalten (72 Items)	Gewissenhaftigkeit	20	5	25
	Genaueres Arbeiten	13	6	19
	Eigeninitiative	11	5	16
	Selbstständiges Arbeiten	7	5	12
Motivation (94 Items)	Durchhaltevermögen	6	6	12
	Lernbereitschaft	7	9	16
	Engagement	15	3	18
	Zielorientierung	10	8	18
	Zielsetzung	8	4	12
	Erfolgszuversicht	6	12	18
Belastbarkeit (21 Items)	Frustrationstoleranz	8	13	21
Soziale Kompetenz (36 Items)	Kommunikationsfähigkeit	4	5	9
	Kontaktfähigkeit	6	3	9
	Teamfähigkeit (Teamfähigkeit, Teamorientierung, Kollegialität, Kompromissbereitschaft)	11	7	18
Emotionale Kompetenzen (32 Items)	Distanzierungsfähigkeit	6	9	15
	Offenheit	13	4	17
Selbstbild (18 Items)	Selbstbewusstsein	7	11	18
Gesamt		206	145	351

Tab. 13 Itemanzahl nach Skalen und Subskalen des WSP

Zusammenfassend wurden im Zuge dieser Arbeit folgende Instrumente zur Qualitätssicherung angewandt:

- Orientierung an den Qualitätsstandards der DIN 33430,
- Verwendung der Critical Incident Daten der Anforderungsanalyse zur allgemeinen Studierfähigkeit
- Literaturrecherche zur Test- und Fragebogenkonstruktion
- Durchführung von Experteninterviews zur Fragebogenkonstruktion
- Entwicklung und Anwendung des ‚Regelkataloges zur Fragebogen-Itemkonstruktion‘
- Webbasierte Rückkopplungs- und Kontrollschleifen (e-Plattform)

- Workshops zur Diskussion und Überarbeitung kritischer Items
- Formulierung einer Kontrolle zur Qualitätsabnahme von Items

Die endgültige Zusammensetzung des Itempools und die Skalenstruktur des WSP sind durch die Ergebnisse der anschließenden Modellgültigkeitsprüfung nach der probabilistischen Testtheorie bedingt. An dieser Stelle sei auf die Nachfolgeuntersuchung von Dörrer (2009) verwiesen. Im Zuge der von Dörrer (2009) durchgeführten Modellgültigkeitsprüfung nach dem (dichotom) logistischen Testmodells von Rasch (1960) erwiesen sich die Items des WSP als konform nach dem Rasch-Modell. Bei der gleichverteilten Vorgabe (d.h. die Items wurden nicht nach Subskalen geblockt, sondern nach einer möglichst großen inhaltlichen Distanz gemischt vorgegeben) mussten lediglich sieben Items aus dem Itempool ausgeschlossen. Der Itempool erwies sich somit als robust gegenüber der Modellprüfung. Die Ergebnisse der Modellprüfung sind in Dörrer (2009) dargestellt.

8 Bibliographie

Anderson, N. (1974). Cognitive algebra. In L. Berkowitz (Hrsg.), *Advances in experimental social psychology*. New York: Academic Press, 1-101.

Becker, P. (2003). Persönlichkeitsfragebogen. In K. D. Kubinger & R. S. Jäger (Hrsg.), *Schlüsselbegriffe der Psychologischen Diagnostik*. Weinheim: Beltz, 332-336.

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

Borkenau, P. & Amelang, M. (1985). Individuelle Angemessenheit von Eigenschaftskonstrukten als Moderatorvariable für die Übereinstimmung zwischen Selbst- und Bekannten-Ratings. *Diagnostica*, 31, 105–118.

Bortz, J. & Döring, N. (2003). *Forschungsmethoden und Evaluation* (3. überarbeitete Auflage). Berlin, Heidelberg u.a.: Springer.

Brehm, J. W. (1966). *A theory of psychological reactance*. New York, London: Academic Press.

Brehm-Gräser, L. (1970). *Familie in Tieren. Die Familiensituation im Spiegel der Kinderzeichnung. Entwicklung eines Testverfahrens*. München, Basel: Reinhardt.

Brickenkamp, R. (2002). *Aufmerksamkeits- Belastungs-Test d2*. Göttingen: Hogrefe.

Briggs, K. C. & Briggs Myers, I. (1995). *Myers-Briggs Typenindikator (MBTI): Deutsche Bearbeitung von R. Bents und R. Blank* (2. Aufl.). Weinheim: Beltz.

Buchanan, T. & Smith, J. L. (1999). Using the internet for psychological research: personality testing on the world wide web. *British Journal of Psychology*, 90, 125-144.

Bühner, Markus (2006). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson Studium.

Bundesministerium für Wissenschaft und Forschung (BM.WF) (2009). *Der Bologna Prozess. Schaffung eines Europäischen Hochschulraumes*. Zugriff am 14. Juni 2009 unter <http://www.bmwf.gv.at/euinternationales/bolognaprozess/>

Dörner, N. (2009). *Über den Einfluss der inhaltshomogenen Itemblockbildung auf die Dimensionalität eines Persönlichkeitsfragebogens*. Unveröffentlichte Diplomarbeit, Universität Wien.

Fahrenberg, J., Hampel, R. & Selg, H. (2001). *Das Freiburger Persönlichkeitsinventar – Revidierte Fassung (FPI-R) (7., überarbeitete und neu normierte Auflage)*. Göttingen: Hogrefe.

Flanagan, J. C. (1954) The critical incident technique. *Psychological-Bulletin*, 51, 327-358.

Frebort, M. & Kubinger, K. D. (2006). Ermittlung eines Anforderungsprofils für Studierende der Psychologie. In B. Gula, R. Alexandrovicz, S. Strauß, E. Brunner, B. Jenufl-Schiefer & O. Vitouch (Hrsg.), *Perspektiven Psychologischer Forschung in Österreich. Proceedings zur 7. Wissenschaftlichen Tagung der Österreichischen Gesellschaft für Psychologie*. Lengerich: Pabst, 440-446.

Frebort, M. & Kubinger, K. D. (2006b). *Webbasierter Studienberatungstest: „Self-Assessment“ für StudienanfängerInnen*. Zugriff am 12. Juni 2009 unter <http://www.studienberatungstest.at/>

Häcker, H. & Stapf K. (2009). *Dorsch Psychologisches Wörterbuch*. Bern: Huber.

Hacker, W. (2005). *Allgemeine Arbeitspsychologie: Psychische Regulation von Wissens-, Denk- und körperlicher Arbeit*. Bern: Huber.

Hambros, K. (2002). On reasonableness of personality inventories with dichotomous item response format. *Psychologische Beiträge*, 44, 126-135.

Hell B., Ptok C. & Schuler, H. (2007). Methodik zur Ermittlung und Validierung von Anforderungen an Studierende (MEVAS). Anforderungsanalyse für das Fach Wirtschaftswissenschaften. *Zeitschrift für Arbeits- und Organisationspsychologie*, 51, 88-95.

Hertel, G., Naumann, S., Konradt, U., & Batinic, B. (2002). Person assessment via Internet: Comparing online and paper-and-pencil questionnaires. In B. Batinic, U. Reips & M. Bosnjak (Hrsg.), *Online Social Sciences*. Berlin: Hogrefe, 115-133.

Heyse, H. & Kersting, M. (2004). Anforderungen an den Prozess der Eignungsbeurteilung. In L. Hornke & U. Winterfeld (Hrsg.), *Eignungsbeurteilungen auf dem Prüfstand: DIN 33430 zur Qualitätssicherung*. München: Elsevier, 29-43.

Hirsch, C. (2006). *Inwiefern hängt die Präferenz eines multiple-choice Antwortformats, mit wenigen (bis dichotom) oder mehr (bis Analogskala bei Persönlichkeitsfragebögen) Antwortmöglichkeiten mit der Persönlichkeit der Testperson zusammen?* Unveröffentlichte Diplomarbeit, Universität Wien.

Horlacher, K. D. & Rodewald, R. (2002). Subjektive Angemessenheit von Items und die Validität von Persönlichkeitsfragebögen. *Zeitschrift für differentielle Psychologie*, 23, 13-27.

Inceoglu, I. & Bartram, D. (2007). Die Validität von Persönlichkeitsfragebögen. Zur Bedeutung des verwendeten Kriteriums. *Zeitschrift für Personalpsychologie*, 6, 160-173.

Jonkisz, E. & Moosbrugger, H. (2007). Planung und Entwicklung von psychologischen Tests und Fragebogen. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion*. Heidelberg: Springer Medizin Verlag, 27-72.

Karner, T. (1999). *Eine systematische Untersuchung der Auswirkungen verschiedener Antwortmodi auf die Qualität psychologischer Fragebögen*. Unveröffentlichte Dissertation, Universität Wien.

Karner, T. (2002). The volunteer effect of answering personality questionnaires. *Psychologische Beiträge*, 44, 42-49.

Khorramdel, L., Maurer, M. & Kubinger, K.D. (2008). *A requirement analysis of study specific demands – What requirements of ability and personality do students need to be successful?* 29th International Congress of Psychology, Berlin, 20th-25th July 2008, Unveröffentlichtes Poster.

Kersting, M. (2008). *Qualität in der Diagnostik und Personalauswahl – der DIN-Ansatz*. Göttingen: Hogrefe.

Koch, A., Kici, G., Strobel, A. & Westhoff, K. (2006). Anforderungsanalysen nach DIN 33430: exemplarisch für die Position eines Dozenten im Arbeitsschutz. In K. Westhoff (Hrsg.), *Nutzen der DIN 33430. Praxisbeispiele und Checklisten*. Wien u.a.: Pabst Science Publ, 85-93.

Krosnick, J. A. (1999). Survey research. *Annual review of Psychology*, 50, 537-567.

Kubinger, K. D. (2002). On faking personality inventories. *Psychologische Beiträge*, 44, 10-16.

Kubinger, K. D. & Proyer, R. (2005). Item-Response-Theorie. 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*. Lengerich u.a.: Pabst Science Publ. 186-191.

Kubinger, K. D. (2006). *Psychologische Diagnostik. Theorie und Praxis psychologischen Diagnostizierens*. Göttingen: Hogrefe.

Kubinger, K.D., Frebort, M., Holocher-Ertl, S. & Pletschko, T. (2006). *Standard-Tests zu den Bildungsstandards in Österreich. Wissenschaftlicher Hintergrund und Hinweise zur Interpretation der Ergebnisse der Standardtests*. Unveröffentlichter Bericht. Herstellung: bmbwk.

Kubinger, K. D. (2007). *Zur Terminologie in psychologischer Diagnostik*. Zugriff am 04. Juli 2009 unter <http://www.univie.ac.at/psychologie/diagnostik/files/Terminologie.pdf>

Latham, G. P., Saari, L. M., Pursell, E. D. & Campion, M. A. (1980). The situational interview. *Journal of applied psychology* 65, 422-427.

Lechner, M. (2008). *Sinnhaftigkeitsprüfung bei MC-Leistungstests*. Unveröffentlichte Diplomarbeit, Universität Wien.

Lienert, G. & Raatz, U. (1998). *Testaufbau und Testanalyse*. Weinheim: Psychologie-Verlag-Union.

Mayring, P. (2008). *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. Weinheim: Beltz.

McKelvie, S.J. (1978). Graphik rating scales – how many categories? *British Journal of Psychology*, 69, 185-202.

Miller, E., T. Neal, D. J., Roberts, L. J., Boer, J.S., Cresskr, S. O., Metrik, J. Marlett, G. (2009). Test-retest reliability of alcohol measures: Is there a difference between internet-based assessment and traditional measures? In Witkiewitz, K. (Hrsg.), *Addictive behaviors: New readings on etiology, prevention, and treatment*. Washington DC: American Psychological Association, 323-341.

Moosbrugger, H. & Rauch W. (2005). Klassische Testtheorie. 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*. Lengerich u.a.: Pabst Science Publ, 182-186.

Mummendey, H. D. (2008). *Die Fragebogen-Methode* (5 überarbeitete und erweiterte Auflage). Göttingen: Hogrefe.

Ortner, T. M. (2003). Anforderungsprofil. In K. D. Kubinger & R.S. Jäger (Hrsg.), *Schlüsselbegriffe der Psychologischen Diagnostik*. Weinheim: Beltz, 20-23.

Österreichische Akademie der Wissenschaften (ÖAW) (2006). *Fachbereiche*. Zugriff am 05 Juli 2009 unter <http://www.oeaw.ac.at/deutsch/forschung/einrichtungen/fachbereiche.html>

Oswald, W. D. & Roth, E. (1987). *Zahlen-Verbindungs-Test (ZVT)* (2. überarbeitete und erweiterte Auflage). Göttingen: Hogrefe.

Paragoni, B. (2008). *Frauenbericht 2007 gemäß Frauenförderplan der WU*. Zugriff am 15. Juni 2009 unter <http://www.wu.ac.at/portal/publikationen/frauenbericht/frauenbericht07.pdf>

Pixner, J. & Zapf, S. (2005). *Welche Anforderungen stellt das Anglistik-Studium?* Unveröffentlichter Bericht. Freiburg: Psychologisches Institut der Universität Freiburg.

Rammstedt, B. & Krebs, D. (2007). Does response scale format affect the answering of personality scales? Assessing the Big 5 dimensions of personality with different response scales in a dependent sample. *European Journal of Psychological Assessment*, 23, 32-28.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.

Ratzer, B., Hnilica, S. & Szalai, E. (2006). *Erster Zwischenbericht. Gender in der Lehre (GIL)*. Zugriff am 26. Mai 2009 unter http://www.tuwien.ac.at/typo3conf/ext/user_tuwien_links/download.php?cuid=13370&file=fileadmin%2Ft%2Fgender%2Fprojekte%2FGiL%2FGiL_Zwischenbericht1.pdf

Referat Frauenförderung und Gleichstellung der Universität Wien (Hrsg.) (2004). *Hinter den Kulissen. Frauen und Männer an der Universität Wien*. (im Auftrag von M. Sebök). Kindberg: Druck-Express W. Tösch.

Reif, M. (2008). *Zur Konstruktion eines adaptiven Persönlichkeitsfragebogens auf Basis eines Konstruktionsrationalen*. Unveröffentlichte Diplomarbeit, Universität Wien.

Reimann, G. (2005). Arbeits- und Anforderungsanalyse. 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*. Lengerich: Pabst, 111-127.

Rost, J. (2002). When personality questionnaires fail to be unidimensional. *Psychologische Beiträge*, 44, 108-125.

Rost, J. (2004). *Lehrbuch Testtheorie und Testkonstruktion* (2. vollständig überarbeitete und erweiterte Auflage). Bern: Verlag Hans Huber

Schmitt, M. (2005). Auswahl von Studierenden: Historie, Zweck und Aufbau dieses Diskussionsforums. *Psychologische Rundschau*, 56, 123-154.

Schuler, H. (2002). *Das Einstellungsinterview*. Göttingen u.a.: Hogrefe.

Schuler, H. & Funke, U. (1995). Diagnose beruflicher Eignung und Leistung. In H. Schuler (Hrsg.), *Lehrbuch Organisationspsychologie* (2. korrigierte Auflage). Bern u a.: Verlag Hans Huber, 235 – 284.

Schwarz, N. (1999). Self reports. How questions shape the answers. *American Psychologist*, 54, 93-105.

Seiwald, B.B. (2003). Item. In K. D. Kubinger & R.S. Jäger (Hrsg.), *Schlüsselbegriffe der Psychologischen Diagnostik*. Weinheim: Beltz, 20-23.

Sonnleitner, P. (2007). Versuch eines Konstruktionsrationalis zum Leseverständnis: Entwicklung & Evaluation auf Basis kognitiver Modelle. Unveröffentlichte Diplomarbeit, Universität Wien.

Statistik Austria (2008, 7. August). *Hochschulstatistik. Ergebnisse im Überblick: Studieren* Zugriff am 17. Juni 2009 unter http://www.statistik.at/web_de/statistiken/bildung_und_kultur/formales_bildungswesen/universitaeten_studium/index.html.

Stumpf, H., Angleitner, A., Wieck, T., Jackson, N. D. & Belloch-Till, H. (1985). *Deutsche Personality Research Form (PRF)*. Göttingen: Hogrefe.

Westhoff, K. (2005). Konstrukte und Operationalisierungen. 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*. Lengerich u.a.: Pabst, 128-161.

Zimmerhofer, A. & Hornke, L. (2005). Konzeption einer webbasierten Studienberatung für Interessierte der Studienfächer Informatik, Elektrotechnik und Technische Informatik. In K. Renner, A. Schütz & F. Machilek (Hrsg.), *Internet und Persönlichkeit: differentiell-psychologische und diagnostische Aspekte der Internetnutzung*. Göttingen: Hogrefe, 269-284.

Zentrum für Testentwicklung und Diagnostik (1995). *Eignungstest für das Medizinstudium in der Schweiz (EMS)*. Zugriff am 30. Juni 2009 unter <http://www.unifr.ch/ztd/>

9 Abbildungs- und Tabellenverzeichnis

Abbildungsverzeichnis

Abb. 1 Clusterung von Studienrichtungen.....	20
Abb. 2 Stichprobe.....	30
Abb. 3 Konstruktschemata und Ausdifferenzierung	35

Tabellenverzeichnis

Tab. 1 Darstellung des Arbeitsprozesses.....	7
Tab. 2 Studienrichtungen nach Universitäten	21
Tab. 3 Deskriptive Statistik Fragebogenerhebung - Studienplan und Alter.....	31
Tab. 4 Deskriptive Statistik Fragebogenerhebung – Geschlecht	32
Tab. 5 Vorgangsweise bei der qualitativen Datenanalyse	33
Tab. 6 Darstellung einer CIT-Situationsnennung der Studienrichtung Soziologie	34
Tab. 7 Darstellung des induktiven Interpretationsschluss bei der Ermittlung der Anforderungen.....	37
Tab. 8 Skalen und Subskalen des WSP	73
Tab. 9 Umwandlung Arbeitsdefinition für Testmanual.....	75
Tab. 10 Itemverrechnung	91
Tab. 11 Itemanzahl nach Skalen und Subskalen des WSP	92
Tab. 12 Itembeispiele nach Skalen und Subskalen des WSP	94
Tab. 13 Itemanzahl nach Skalen und Subskalen des WSP	101
Tab. 14 Anforderungsprofilsprofil und Häufigkeitsnennungen	120
Tab. 15 Arbeitsdefinitionen der Subskalen des WSP.....	122

Boxenverzeichnis

Box 1: Fragebogen für Studierende nach der CIT	25
Box 2: Interviewleitfaden für Lehrende nach der CIT	27
Box 3: Checkliste für den/die InterviewerIn	29
Box 4: Beispielauszug Interview (Lehrperson, Studienrichtung Architektur).....	38
Box 5: Beispiele für zwei situative Fragen in Einstellungsinterviews	79
Box 6: Beispiel für eine situative Frage mit Antwortkategorien	80
Box 7: Beispiel für die Umwandlung der Daten der Anforderungsanalyse in Items	83
Box 8: Kommentierung von Itemvorschlägen durch die ExpertInnengruppe	86
Box 9: Ergänzung des ‚Regelkataloges zur Fragebogen-Itemkonstruktion‘.....	90

10 Anhang

CIT-Fragebogen für Studierende der Publizistik- und Kommunikationswissenschaft.....	114
CIT-Interviewleitfaden für Lehrpersonen der Volkswirtschaftslehre.....	118
Anforderungsprofil zur allgemeinen Studierfähigkeit.....	120
Arbeitsdefinitionen der Subskalen des WSP.....	122
Anleitung zur Kommentierung der Fragebogen-Items.....	124
Itempool des WSP und Itemverrechnung.....	126

CIT- Fragebogen für Studierende der Publizistik- und Kommunikationswissenschaft

Code: (vom Testleiter auszufüllen)

Datum: _____	Alter _____	Geschlecht: <input type="checkbox"/> männlich <input type="checkbox"/> weiblich
Alter Studienplan:		
Ich habe den 1. Studienabschnitt abgeschlossen	<input type="checkbox"/> Ja	<input type="checkbox"/> Nein
Ich habe den 2. Studienabschnitt abgeschlossen	<input type="checkbox"/> Ja	<input type="checkbox"/> Nein
Ich arbeite an meiner Diplomarbeit	<input type="checkbox"/> Ja	<input type="checkbox"/> Nein
Ich arbeite an meiner Dissertation	<input type="checkbox"/> Ja	<input type="checkbox"/> Nein
Neuer Studienplan:		
Ich befinde mich im Bakkalaureatsstudium in folgendem Abschnitt	<input type="checkbox"/> 1. Studienjahr	<input type="checkbox"/> 2. Studienjahr
	<input type="checkbox"/> 3. Studienjahr	
Ich befinde mich im Magisterstudium	<input type="checkbox"/> Ja	<input type="checkbox"/> Nein
Ich arbeite an meiner Diplomarbeit	<input type="checkbox"/> Ja	<input type="checkbox"/> Nein
Ich arbeite an meiner Dissertation	<input type="checkbox"/> Ja	<input type="checkbox"/> Nein

Die folgenden **4 Fragen** dienen der Erstellung eines Anforderungsprofils für **Studierende der Publizistik- und Kommunikationswissenschaft**. Sämtliche Fragen beziehen sich auf Studierende an der Fakultät für Sozialwissenschaften der Universität Wien. Sie beziehen sich nicht auf beruflichen Erfolg bzw. berufliche Eignung, sondern immer auf **Studienerfolg bzw. Eignung zum Studium**. Bedenken Sie, welche Eigenheiten das Studium der Publizistik und Kommunikationswissenschaften hat und welche Hürden dabei genommen werden müssen.

1) Beschreiben Sie eine/n erfolgreiche/n Publizistik-Studierende/n. Was zeichnet sie/ihn aus?

- 2) Erinnern Sie sich bitte an **typische und wichtige Situationen** im Studium der Publizistik und Kommunikationswissenschaften, die Sie selbst erlebt oder beobachtet haben und deren erfolgreiche Bewältigung für den Fortgang des Studiums wichtig ist.

Welche Situationen waren/sind das? (Situationen können auch Prüfungen oder Lehrveranstaltungen sein.)

Bitte gehen Sie konkret auf die Situation ein und beschreiben Sie diese genau.

Tragen Sie bitte jede Situation in eine neue Zeile in der Tabelle ein. Wenn Ihnen noch mehr Situationen einfallen, schreiben Sie einfach unter der Tabelle weiter.

	Typische und wichtige Situationen:
01	
02	
03	
04	
05	
06	
07	
08	
09	
10	
11	
12	
13	
14	
15	
...	

- 3) Welche **Verhaltensweisen** sollte man in diesen Situationen zeigen, um sie erfolgreich zu bewältigen?

Berichten Sie aus Ihrer eigenen Erfahrung und/oder denken Sie an eine KollegIn, die Sie für **sehr geeignet** für das Studium der Publizistik und Kommunikationswissenschaften halten und die diese Situationen erfolgreich bewältigt hat.

(Bitte geben Sie dabei keine Eigenschaften an wie z.B. „ist ehrgeizig“, sondern Verhaltensweisen wie z.B. „arbeitet fleißig mit“.)

Tragen Sie bitte für jede oben genannte Situation die **dazugehörigen erfolgreichen Verhaltensweisen** in die entsprechend nummerierte Zeile in der Tabelle ein.

	Erfolgreiche Verhaltensweisen:
01	
02	
03	
04	
05	
06	
07	
08	
09	
10	
11	
12	
13	
14	
15	
...	

4) Welche **Verhaltensweisen** sollte man in diesen Situationen auf keinen Fall zeigen?

Berichten Sie aus Ihrer eigenen Erfahrung und/oder denken Sie an eine KollegIn, die Sie für **ungeeignet** für das Studium der Publizistik und Kommunikationswissenschaften halten (die das Studium eventuell sogar abgebrochen hat).

(Bitte geben Sie dabei wieder keine Eigenschaften an wie z.B. „ist nicht ehrgeizig“, sondern Verhaltensweisen wie z.B: „arbeitet nicht fleißig mit“).

Tragen Sie bitte für jede oben genannte Situation die **dazugehörigen nicht erfolgreichen Verhaltensweisen** in die entsprechend nummerierte Zeile in der Tabelle ein.

	Nicht erfolgreiche Verhaltensweisen:
01	
02	
03	
04	
05	
06	
07	
08	
09	
10	
11	
12	
13	
14	
15	
...	

Vielen Dank für die Bearbeitung!

CIT-Interviewleitfaden für Lehrpersonen der Volkswirtschaftslehre

Code:

Datum: _____	Anzahl der Lehrjahre: _____	Geschlecht: männlich <input type="checkbox"/> weiblich <input type="checkbox"/>	Anstellung/Beruf: Univ-Prof <input type="checkbox"/> Univ-Ass <input type="checkbox"/> Univ-Lekt <input type="checkbox"/>
---------------------	------------------------------------	---	---

Die Folgenden Fragen dienen der Erstellung eines Anforderungsprofils für Studierende der Volkswirtschaftslehre. Sämtliche Fragen beziehen sich auf Studierende am Institut für Volkswirtschaftslehre an der Universität Wien. Sie beziehen sich nicht auf beruflichen Erfolg bzw. berufliche Eignung, sondern immer auf **Studienerfolg bzw. Eignung zum Studium**. Bedenken Sie, welche Eigenheiten das Studium der Volkswirtschaftslehre hat und welche Hürden dabei genommen werden müssen.

- 1) Wann ist Ihrer Meinung nach ein Studierender im Studium der Volkswirtschaftslehre erfolgreich?
- 2) Woran erkennen Sie erfolgreiche / wenig erfolgreiche Studierende bzw. was unterscheidet einen erfolgreichen und einen wenig erfolgreichen Studierenden? (Welche Verhaltensweisen machen sie/ihn erfolgreich / wenig erfolgreich?)

- 3) Können Sie sich an das letzte Mal erinnern, als ein (von Ihnen betreuter) Studierender etwas geleistet oder getan hat, das Sie besonders gut fanden? [*Instruktion von Flanagan: nach dieser Frage kurz warten, bis der Interviewte ein solches Ereignis im Kopf hat*]
- 4) Was war das für eine Situation / unter welchen Umständen kam es dazu?
- 5) Was war das Besondere an dieser Situation?
- 6) Welche Aufgabe genau erhielt der Studierende in dieser Situation?
- 7) Was hat dieser Studierende getan, welches Verhalten hat er/sie gezeigt, das zu dieser außerordentlichen Leistung führte?
- 8) Wann im Semester/im Laufe der Lehrveranstaltung war dieses Ereignis genau?
- 9) Er/sie war Studierender in welchem Semester / Studienabschnitt?
- 10) Welches Verhalten sollte ein Studierender in einer solchen Situation auf keinen Fall zeigen bzw. wie hätte sich ein wenig erfolgreicher Studierender verhalten?

- 11) Fallen Ihnen weitere typische und wichtige Situationen ein (*bei Nachfragen des Interviewten: d.s. Situationen, deren erfolgreiche Bewältigung für den Fortgang des Studiums wichtig sind*), in denen ein Studierender etwas geleistet oder getan hat, das Sie besonders gut fanden? Oder: Fallen Ihnen weitere typische und wichtige Situationen ein, in denen sich zeigt bzw. zeigen könnte, ob ein Studierender erfolgreich oder wenig erfolgreich ist?
- 12) Was waren / sind das für Situationen?
- 13) Was war / ist das besondere an diesen Situationen?
- 14) Was hat der Studierende getan, welches Verhalten hat er/sie gezeigt, das zu dieser außerordentlichen Leistung führte? Oder: Wie würde sich ein erfolgreicher Studierender in solchen Situationen verhalten?

- 15) Welche Aufgabe /Aufgaben genau erhielt / erhält der Studierende in diesen Situationen?
- 16) Er/sie war Studierender in welchem Semester bzw. wann im Laufe des Semesters zeigte er/sie diese außergewöhnliche Leistung? Oder: Wann genau im Laufe des Studiums (Semester / Studienabschnitt / Lehrveranstaltung) zeigen sich solche Situationen?
- 17) Welches Verhalten sollte ein Studierender in solchen Situationen auf keinen Fall zeigen bzw. wie hätte / würde sich ein wenig erfolgreicher Studierender verhalten?

18) Denken Sie, es wird in Zukunft noch andere typische und wichtige Situationen / Umstände geben, in denen sich zeigt, ob ein Studierender erfolgreich oder wenig erfolgreich ist? Denken Sie dabei an mögliche Änderungen im Studienplan (z.B. Bakkalaureat).

19) Was für Situationen / Umstände werden das sein?

20) Welches Verhalten würde ein erfolgreicher Studierender in diesen Situationen zeigen?

21) Welches Verhalten würde ein wenig erfolgreicher Studierender in dieser Situation zeigen?

22) Welche Eigenschaften hat ein erfolgreicher Studierender (Ihrer Meinung nach)?

23) Welche Eigenschaften fehlen einem wenig erfolgreichen Studierenden (Ihrer Meinung nach) um im Studium erfolgreich zu sein?

24) Denken Sie an einen Ihrer Kollegen, den Sie sehr schätzen.

- Was meinen Sie, woran erkennt sie/er erfolgreiche / wenig erfolgreiche Studierende?

- Was meinen Sie: unterscheiden sich Ihre Kriterien von denen Ihres Kollegen? Falls ja, inwiefern?

25) Denken Sie an einen Ihrer Kollegen, den Sie nicht so sehr schätzen.

- Was meinen Sie, woran erkennt sie/er erfolgreiche / wenig erfolgreiche Studierende?

- Was meinen Sie: unterscheiden sich Ihre Kriterien von denen Ihres Kollegen? Falls ja, inwiefern?

Checkliste für den Interviewer:

- Wurden die Verhaltensweisen zu den entsprechenden Situationen berichtet?
- Wurden die Verhaltensweisen vom interviewten Experten selbst beobachtet?
- Wurden alle relevanten Faktoren der Situationen beschrieben?
- Hat der interviewte Experte klar definiert, welches die erfolgreiche und welches die wenig erfolgreiche Verhaltensweise ist?
- Hat der interviewte Experte erklärt, warum welches Verhalten erfolgreich oder wenig erfolgreich ist?

Vielen Dank für das Interview!

Anforderungsprofil zur allgemeinen Studierfähigkeit

Tab. 14 Anforderungsprofilsprofil und Häufigkeitsnennungen

Konstrukt	Hfkt	Fähigkeiten und Eigenschaften	Hfkt
Motivation (M)	1343	Durchhaltevermögen	146
		Lernbereitschaft	130
		Engagement	440
		Aktive Problembearbeitung	15
		Interesse	155
		Zielorientierung	81
		Zielsetzung	49
		Wissensdurst/Neugier	34
		Ehrgeiz	34
		Gestaltungsmotivation	14
		Anpassungsfähigkeit	06
		Erfolgszuversicht	73
		Wettbewerbsorientierung	04
		Schwierigkeitspräferenz	10
		Leistungsstolz	06
Organisationsfähigkeit (O)	776	Zeitmanagement	137
		Prioritäten setzen	78
		Planen	76
		Organisationstalent	202
		Informiertheit	171
Intelligenz (I)	738	schlussfolgerndes Denken	199
		Kritisches Denken	128
		Verbale Intelligenz	90
		Numerische Intelligenz	45
		Flexibilität im Denken	62
		Raumvorstellung	14
		Kreativität	59
		figurale Fähigkeiten	13
		praktisch-technisches Verständnis	24
		technisches Verständnis	17
		Umsetzung/Anwendung von Gelerntem	34
		physikalisches Grundverständnis	03
		Informationsverarbeitungsgeschwindigkeit	04
		Arbeitsstil und Arbeitsverhalten (A)	591
Reflexivität	16		
genaues Arbeiten	55		
realistisches Einschätzen der eigenen Leistungsfähigkeit / Fähigkeiten	35		
realistisches Anspruchsniveau	06		
hohes Anspruchsniveau	03		
Eigeninitiative	49		
Selbstständiges Arbeiten	86		
Soziale Kompetenz (S)	567	Durchsetzungsstärke	23
		Überzeugungsstärke	36
		Autoritätsgläubigkeit	09
		Kommunikationsfähigkeit	195
		Teamfähigkeit	43
		Teamorientierung	05
		Kollegialität	15
		Kompromissbereitschaft	14
		Kontaktfähigkeit	92
		Sensitivität	18

Fortsetzung: Anforderungsprofilsprofil und Häufigkeitsnennungen

Lernverhalten (L)	224	Lernstil	51
		Lernfähigkeit	54
		Merkfähigkeit	13
		Konzentration bzw. zielgerichtete Aufmerksamkeit	34
Belastbarkeit (B)	229	Frustrationstoleranz	113
		Irritierbarkeit und Nervosität bei Leistungserbringung vor anderen	15
		Arbeiten (bzw. Belastbarkeit) unter Zeitdruck	25
		Bewältigung von Mehrfachanforderungen	07
		Fähigkeit zur Entspannung	27
		Belastbarkeit und Flexibilität bei Hindernissen	06
		Emotionale Komponenten (E)	174
Offenheit	55		
Standhaftigkeit	36		
Extraversion vs. Introversion	07		
Selbständigkeit	32		
Ehrlichkeit	02		
Wissen (W)	78	fachspezifisches Wissen	53
		Fremdsprachenkenntnisse	14
		Vorwissen	08
		Allgemeinwissen	03
Selbstbild (Se)	76	Selbstbewusstsein	76
Rahmenbedingungen (U)	38	studienbezogene Rahmenbedingungen	20
		umweltbezogene Rahmenbedingungen (muss neben dem Studium arbeiten)	15
Kontrollüberzeugung (K)	38	Internalität	32

Arbeitsdefinitionen der Subskalen des WSP

Tab. 15 Arbeitsdefinitionen der Subskalen des WSP

Skala	Subskala	Arbeitsdefinition
Organisation (O)	Zeitmanagement	Einteilen, wie viel Zeit Arbeitsschritte benötigen, zeitliche Koordination von Arbeitsschritten, zeitliche Koordination von Mehrfachanforderungen, Einteilung Arbeitszeit und Freizeit/Studium und Arbeit +
	Prioritäten setzen	Prioritäten so setzen, dass alles zum richtigen Zeitpunkt fertig wird, abwägen, was wichtiger ist, effizientes Arbeiten, sich auf das Wesentliche konzentrieren, Wichtiges von Unwichtigem trennen: was ist wirklich notwendig? +
	Planen	Planen von Arbeitsschritten, Semesterplan erstellen, im Voraus planen, was man alles benötigen wird, damit man es rechtzeitig organisieren kann, planen, wann welche Prüfungen man am besten macht: z.B. die Prüfung gleich nach der entsprechenden LV, weil da das Wissen noch frisch ist ...+
	Organisationstalent	z.B. Organisieren von Informationen, Organisieren von (Zusatz-)Terminen, sich für bestimmte LVs anmelden i.S. von Umsetzung des Geplanten, den Überblick behalten, Organisieren von Kontakten, Ordnung ins „Chaos“ bringen, strukturiertes Arbeiten, mehrere Dinge gleichzeitig bedenken +
	- Informiertheit	vorab informieren, über bestimmte Themen genauer informieren, sich bestimmtes „nötiges“ Wissen aneignen, realistische Vorstellung vom Studium und dessen Inhalten, realistische Erwartungen an das Studium +
Arbeitsstil, Arbeitsverhalten (A)	Gewissenhaftigkeit	Termine regelmäßig wahrnehmen, Lehrveranstaltungen regelmäßig besuchen, immer gut vorbereitet sein, immer auf dem aktuellen Wissensstand sein, Pünktlichkeit, immer Mitschreiben, eine „saubere“ Mitschrift haben, alle wichtigen Punkte berücksichtigen, die Hausübung selber machen und nicht abschreiben, sich an den eigenen Zeitplan halten, Kontakte nicht abreißen lassen sondern pflegen, Rücksprache halten, um sich zu vergewissern, dass alles so passt, diszipliniertes Arbeiten, +
	genaues Arbeiten	nur genauer Arbeitsstil ohne sonstige Einschränkungen) +
	Eigeninitiative	von sich aus Feedback einholen, von sich aus Literatur zu einem bestimmten Thema recherchieren, wenn eine Führung ausfällt, von sich aus ein Alternativprogramm zusammenstellen +
	Selbstständiges Arbeiten	alleine vorankommen, ohne Gruppe und Hilfe, eigenständiges Erlernen von z.B. Computerprogrammen; oder „Kleingruppe“ mit 2-3 Studierenden, die sehr selbstständig und ohne viel Hilfe der LV-Leiter arbeiten +

Fortsetzung: Arbeitsdefinitionen der Subskalen des WSP

Motivation (M)	Durchhaltevermögen	nicht aufgeben, z.B. bei langweiligen oder langwierigen Aufgaben/Tätigkeiten,... +
	Lernbereitschaft	Kritik positiv sehen und als schöpferisches Potenzial verstehen, sich weiterbilden, etwas lernen wollen, Hilfe z.B. von LV-Leitern annehmen können,... +
	Engagement	gut mitarbeiten, aus Interesse immer zu den LVs gehen, zusätzliche Praktika machen, wenn nicht nötig oder Stoff vertiefen bzw. mehr Lernen als nötig, sich intensiv mit etwas beschäftigen, Anstrengungsbereitschaft +
	Zielorientierung	das Ziel nicht aus den Augen verlieren +
	Zielsetzung	sich überhaupt Ziele setzen, sich bestimmte Ziele setzen, eine konkrete Vorstellung haben, wissen, was man will und wo man hin will, z.B. auch mit einem Konzept +
	Erfolgszuversicht	Zuversicht, dass die eigene Handlung sehr wahrscheinliches erfolgreich sein wird, auch bei neuen Herausforderungen, Hoffnung auf Erfolg + vs. <u>Misserfolgsängstlichkeit</u> : Furcht vor Misserfolg –
Belastbarkeit (B)	Frustrationstoleranz	z.B. kann jemand gut mit negativer Rückmeldung umgehen, gibt jemand nach frustrierenden Erlebnissen oder negativer Rückmeldung sein Ziel/seine Idee nicht gleich auf; Fähigkeit, so mit Kritik oder Misserfolg umzugehen, dass diese keinen Einfluss auf die Leistungsbereitschaft bzw. Motivation hat. Hier ist also eine motivationale Komponente dabei. +
Soziale Kompetenz (S)	Kommunikationsfähigkeit	= Wissen um Kommunikationsregeln und Bereitschaft zur Kommunikation, angemessene Kommunikation, Höflichkeit, Bereitschaft miteinander zu reden, ... +
	Teamfähigkeit	in Gruppen zusammenarbeiten können: z.B. Kommunikationsfähigkeit, Kontaktfähigkeit, Kollegialität, Kompromissbereitschaft, ... +
	Kontaktfähigkeit	auf andere Personen aktiv zugehen, ansprechen, Gespräch suchen, Kontakt knüpfen +
Emotionale Komponenten (E)	Distanzierungsfähigkeit	in Bezug auf die eigenen Arbeiten neutral bleiben, objektiv beurteilen, Kritik nicht persönlich nehmen, etwas nicht zu wichtig nehmen, loslassen können +
	Offenheit	für neue Eindrücke und Erfahrungen, für neue/andere Ideen und Meinungen + Abgrenzung Lernbereitschaft: Bei Lernbereitschaft muss nicht nur die Bereitschaft vorhanden sein, sondern auch der Wille es umzusetzen. Bei der Offenheit geht es nur mal darum, sich etwas anzuhören, neue Eindrücke zu suchen.
Selbstbild (SE)	Selbstbewusstsein	zu eigenen Ideen/Meinung stehen, sich nicht verunsichern lassen, nicht an sich selbst zweifeln, an sich selbst glauben, den Mut haben, sich in Lehrveranstaltungen zu melden +

Anleitung zur Kommentierung der Fragebogen-Items

Was gibt es Wissenswertes zum Diplomarbeitshintergrund?

Kurzer Überblick über die Vorstudie

In einer Vorstudie wurden Anforderungsprofile zu 8 Studienrichtungen an Wiener Universitäten mittels der Methode der Critical Incident Technique, CIT (Flanagan, 1954) erstellt.

Es wurde eine Fragebogenerhebung an Studierenden und Experteninterviews mit Lehrpersonen durchgeführt. Dabei wurden kritische Situationen während des Studiums beschrieben und zu den Situationen erfolgreiche und nicht erfolgreiche Verhaltensweisen benannt.

Diese Daten wurden in Anlehnung an die Qualitative Inhaltsanalyse nach Mayring interpretiert.

Daraus ergaben sich die Nennungen und Definitionen der studienrelevanten Fähigkeiten und Eigenschaften.

Fragestellung der Diplomarbeit

Die Diplomarbeit umfasst die Erstellung eines Regelkataloges zur Itemkonstruktion und die Konzeption eines **Persönlichkeitsfragebogens zur Erfassung studienrelevanter Eigenschaften**. Welche Eigenschaften relevant sind, wurde durch die Häufigkeitszählung (Cutt-off Wert bei > 40) festgelegt.

Zielgruppe: Studierende, die sich ihres Studiums unsicher sind/Probleme haben/das Studium wechseln möchten etc. und Beratung suchen. In weiterer Folge wäre es interessant, die Items derart zu „vereinfachen“, dass auch MaturantInnen das notwendige Hintergrundwissen mitbringen. Im Moment werden die Items jedoch sehr nahe am Studienalltag konzipiert.

Auswahl der Skalen und Subskalen:

Organisation:

- Zeitmanagement
- Prioritäten setzen
- Planen
- Organisationstalent
- Informiertheit

Arbeitsstil, Arbeitsverhalten:

- Gewissenhaftigkeit
- genaues Arbeiten
- Eigeninitiative
- Selbstständiges Arbeiten

Motivation:

- Durchhaltevermögen
- Lernbereitschaft
- Engagement
- Zielorientierung
- Zielsetzung
- Erfolgszuversicht vs. Misserfolgsängstlichkeit

Belastbarkeit:

- Frustrationstoleranz

Soziale Kompetenz:

- Kommunikationsfähigkeit
- Kontaktfähigkeit
- Teamfähigkeit

Emotionale Komponenten:

- Distanzierungsfähigkeit
- Offenheit

Selbstbild:

- Selbstbewusstsein

Ziel: Je Subskala 15 – 20 qualitativ hochwertige Items.

Wobei könnt ihr mir helfen?

Es wurde eine Internetplattform eingerichtet, auf der die Testitems zugänglich sind und jede/r jederzeit seine Kommentare anbringen kann. Jede/r von euch bekommt einen eigenen Besucheraccount.

Was ist auf der Internetplattform zu finden:

- Definitionen der Subskalen
- Die aktualisierten Items (wird regelmäßig von mir ergänzt – bin selbst gerade im Prozess der Erstellung).
- Vorläufiger Itemregelkatalog – dient als Anleitung, worauf bei der Itemkommentation geachtet werden muss (Anmerkungen zum Regelkatalog sind ebenfalls erwünscht).

Anleitende Fragen, die euch beim Kommentieren helfen:

4. Steht die angegebene Dimension im Vordergrund?
D.h. misst ein Item der Subskala Zeitmanagement wirklich Zeitmanagement oder eine andere Dimension? Wenn ja welche?
Manchmal ist die Eindimensionalität schwer zu gewährleisten. Wenn ein Item mehr als eine Dimension zu messen scheint, z.B. Zeitmanagement und Planen -> Welche Dimension steht im Vordergrund? (siehe Definition der Subskalen)
5. Passt die Formulierung des Items?
Sind die Begriffe zu kompliziert? Zu viele Nebensätze? Könnte man es einfacher ausdrücken? Ist es missverständlich? (Kriterien siehe Regelkatalog) Wie würdet ihr es umformulieren?
6. Versetzt euch in die Testperson hinein!
Fällt es euch schwer die Frage zu beantworten? Wenn ja, wieso?
Kann ich mich eindeutig für eine Antwort entscheiden. Wenn nein, wieso nicht?
Ist es ein Item, bei dem jede/r zusagen oder es negieren würde? -> Kein Informationsgewinn.

Was es noch anzumerken gibt:

Das Kommentieren der Items sollte möglichst interaktiv vor sich gehen. Kommentiert nicht nur meine Items sondern auch die Kommentare untereinander. Wenn je Item 3-4 Kommentare sind, überarbeite ich es. Dann bitte nochmals anschauen und das überarbeitete Item kommentieren.

Jedes Mal, wenn ich die Items zu einer neuen Subskala auf die Plattform gestellt habe, schicke ich euch ein Rundmail. Ziel wäre es, die Itementwicklung bis Ende November abzuschließen.

Schreibt euch die Anzahl der Stunden, die ihr benötigt, auf und teilt es mir am Schluss mit. Ich leite sie weiter.

Wichtig: Die Kritik sollte konstruktiv sein, aber seid nicht zimperlich mit mir ☺. Je besser die Kritik, umso besser die Qualität der Items.

Danke für deine Hilfe!

Itempool des WSP und Itemverrechnung

Skalen und dazugehörige Items

Verrechnung pro Item

	Ja	Eher ja	Eher nein	nein
Organisationsfähigkeit				
O1 Zeitmanagement				
Arbeitsdefinition: Einteilen, wie viel Zeit Arbeitsschritte benötigen, zeitliche Koordination von Arbeitsschritten, zeitliche Koordination von Mehrfachanforderungen, Einteilung Arbeitszeit und Freizeit/Studium und Arbeit +				
Ich teile mir meine Arbeitszeit so ein, dass ich Abgabetermine erst gar nicht verschieben muss.	1	1	0	0
Wenn eine Arbeit mehrere Wochen bis zur Fertigstellung benötigt, erstelle ich einen Zeitplan.	1	1	0	0
Meine Arbeitszeit teile ich mir so ein, dass ich zwischendurch Pausen machen kann und trotzdem zeitgerecht fertig werde.	1	1	0	0
Wenn bei einer schriftlichen Prüfung die Zeit knapp ist, teile ich mir die Zeit für die Beantwortung pro Frage genau ein.	1	1	0	0
Es zählt zu meinen Stärken, dass ich Aufgaben innerhalb der vorgegebenen Zeit erledige.	1	1	0	0
Ich wäre gerne besser darin, neue Arbeitsaufträge in meine bestehende zeitliche Planung zu integrieren.	0	0	1	1
Auch bei einer Zunahme des Stoffumfanges im Vergleich zu bisherigen Prüfungen, schätze ich die benötigte Dauer für die Prüfungsvorbereitung richtig ein.	1	1	0	0
Wenn mehrere Prüfungen in derselben Woche stattfinden, fühle ich mich mit der Einteilung der Lernzeit überfordert.	0	0	1	1
Ich teile mir den Lernstoff vor einer Prüfung so ein, dass am Ende genug Zeit für eine Wiederholung des Lernstoffs bleibt.	1	1	0	0
Die zeitliche Koordination mehrerer Aufgaben fällt mir leicht.	1	1	0	0
Bei langfristigen Projekten unterschätze ich den zeitlichen Arbeitsaufwand.	0	0	1	1
Ich plane zu wenig Zeit für Prüfungsvorbereitungen ein.	0	0	1	1
Bei der Prüfungsvorbereitung plane ich genügend Pufferzeit für unerwartete Ereignisse ein (z.B. Krankheitsfall).	1	1	0	0
Ist die Prüfungszeit sehr knapp bemessen, halte ich mich nicht zu lange bei einzelnen Fragen auf.	1	1	0	0
Andere fangen mit ihren Aufgaben üblicherweise später an als ich.	1	1	0	0

Fortsetzung: Itempool des WSP und Itemverrechnung

Skalen und dazugehörige Items	Verrechnung pro Item			
	Ja	Eher ja	Eher nein	nein
Für wichtige Termine melde ich mich so früh wie möglich an. -> <i>auch Gewissenhaftigkeit prüfen</i>	1	1	0	0
Bei der Organisation von Sozialleistungen (z.B. Familienbeihilfe, Studienbeihilfe) informiere ich mich so früh wie möglich. -> <i>auch Informiertheit prüfen</i>	1	1	0	0
Wie viel Zeit ich für einzelne Aufgaben benötige, kann ich sehr gut einschätzen.	1	1	0	0
Ich plane die Zeit für anstehende Aufgaben zu knapp ein.	0	0	1	1
Wenn ich an einem Tag mehrere Aufgaben zu bearbeiten habe, fällt es mir schwer, den Tag so zu gestalten, dass ich zumindest die wichtigsten Anliegen erledige. -> <i>auch auf Prioritäten setzen prüfen</i>	0	0	1	1
Während der Prüfungsvorbereitung überprüfe ich regelmäßig, ob ich noch im vorgesehenen Zeitplan bin.	1	1	0	0
Es fällt mir leicht, meine Zeitpläne an die Änderung äußerer Rahmenbedingungen anzupassen.	1	1	0	0
Bei einem umfangreichen Prüfungsstoff werde ich grundsätzlich nicht mit dem Lernen fertig.	0	0	1	1

O2: Prioritäten setzen

Arbeitsdefinition: Prioritäten so setzen, dass alles zum richtigen Zeitpunkt fertig wird, abwägen, was wichtiger ist, effizientes Arbeiten, sich auf das Wesentliche konzentrieren, Wichtiges von Unwichtigem trennen: was ist wirklich notwendig? +

Um im Studium gut voranzukommen, verzichte ich notfalls auch auf Freizeitaktivitäten, auf die ich mich schon lange gefreut habe.	1	1	0	0
Wenn ich kurz vor einer Prüfung den Stoff noch nicht ausreichend beherrsche, verzichte ich auf meine Freizeit, um mehr lernen zu können	1	1	0	0
Ich neige grundsätzlich dazu, alles gleichzeitig machen zu wollen. = <i>kritisches Item (für RM vormerken)</i>	0	0	1	1
Wenn ich mehrere Aufgaben zu erledigen habe, fällt es mir schwer zu entscheiden, welche Aufgabe am dringendsten ist.	0	0	1	1
Ich bearbeite zuerst die interessanten Aufgaben, auch wenn diese nicht vorrangig sind. = <i>kritisches Item (für RM vormerken)</i>	0	0	1	1

Fortsetzung: Itempool des WSP und Itemverrechnung

Skalen und dazugehörige Items	Verrechnung pro Item			
	Ja	Eher ja	Eher nein	nein
Nachdem ich mir einen Überblick über anstehende Aufgaben verschafft habe, fällt es mir leicht, die Aufgaben ihrer Wichtigkeit nach zu reihen.	1	1	0	0
Zu Beginn eines Projektes fällt es mir schwer zu bestimmen, was als erstes zu tun ist.	0	0	1	1
Kurz vor einer Prüfung sind Dinge des Alltags zweitrangig. = <i>kritisches Item (für RM vormerken)</i>	1	1	0	0
Treffen mit Freunden sage ich auch ab, wenn ich für eine Prüfung noch nicht genügend gelernt habe.	1	1	0	0
Wenn ich mehrere Aufgaben zu erledigen habe, dann weiß ich nicht, mit welcher ich beginnen soll.	0	0	1	1
Beim Lernen lasse ich mich wiederholt von unbedeutenden Dingen ablenken.	0	0	1	1
Weil ich mich vom Lernen ständig ablenken lasse, schaffe ich nicht den gesamten Lernstoff.	0	0	1	1
Wichtige Aufgaben kommen wiederholt zu kurz, weil ich nicht mit den wichtigsten, sondern mit den angenehmsten Aufgaben beginne.	0	0	1	1

O3: Planen

Arbeitsdefinition Planen: Planen von Arbeitsschritten, Semesterplan erstellen, im Voraus planen, was man alles benötigen wird, damit man es rechtzeitig organisieren kann, planen, wann welche Prüfungen man am besten macht: z.B. die Prüfung gleich nach der entsprechenden LV, weil da das Wissen noch frisch ist ... +

Um sicher in genügend Lehrveranstaltungen einen Platz zu bekommen, recherchiere ich vor der Seminaranmeldung mehrere alternative Lehrveranstaltungen.	1	1	0	0
Ich recherchiere mögliche Überschneidungen von Seminaren bereits mehrere Tage vor der Anmeldung. -> <i>auch auf Gewissenhaftigkeit prüfen</i>	1	1	0	0
Zu Semesterbeginn lege ich fest, welche Prüfungen ich zu Semesterende machen werde.	1	1	0	0
Zu Semesterbeginn versäume ich es darauf zu achten, dass ich genügend ECTS-Punkte für den Bezug der Familienbeihilfe belege. = <i>kritisches Item, RM-Analyse vormerken</i>	0	0	1	1

Fortsetzung: Itempool des WSP und Itemverrechnung

Skalen und dazugehörige Items	Verrechnung pro Item			
	Ja	Eher ja	Eher nein	nein
Bei umfangreichen Prüfungen erstelle ich mir einen genauen Lernplan (Einteilung des Prüfungsstoffes in Lerneinheiten).	1	1	0	0
Bei der Prüfungsvorbereitung setze ich mir pro Tag eine bestimmte Lernmenge (z.B. Seitenanzahl) fest. -> auch auf Zeitmanagement prüfen	1	1	0	0
Ich entscheide zu Beginn des Semesters eher spontan, welche Lehrveranstaltung ich besuche.	0	0	1	1
In sehe in meiner Studienplanung ein großes Verbesserungspotential.	0	0	1	1
Ich wäre gerne besser darin, mein Semester entlang des Studienplans zu planen.	0	0	1	1
Ich plane sehr ungern lange im Voraus.	0	0	1	1
Ich plane so kurzfristig, dass ich wichtige Termine (z.B. Prüfungsanmeldung) auch mal versäume.	0	0	1	1
Um in meiner Arbeit zügig voran zu kommen, nehme ich mir für jeden Tag bestimmte Tagesziele vor.	1	1	0	0
Ich plane meine nächsten Arbeitsschritte lange im Voraus.	1	1	0	0
Noch bevor ich für eine Prüfung zu lernen beginne, setze ich die zu lernende Stoffmenge für jeden Tag fest.	1	1	0	0
Bei neuen Aufgaben überlege ich immer genau, was wann zu tun ist.	1	1	0	0
Ich überlege mir in der Früh genau, was ich alles während des Tages zu erledigen habe.	1	1	0	0
Ich plane mein Semester mit der Kenntnis davon, welche Lehrveranstaltungen Voraussetzung für darauf folgende Seminare sind. -> auch auf Informiertheit prüfen	1	1	0	0

O4: Organisationstalent

Arbeitsdefinition Organisationstalent: z.B. Organisieren von Informationen, Organisieren von (Zusatz-)Terminen, sich für bestimmte LVs anmelden i.S. von Umsetzung des Geplanten, den Überblick behalten, Organisieren von Kontakten, Ordnung ins „Chaos“ bringen, strukturiertes Arbeiten, mehrere Dinge gleichzeitig bedenken +

Es fällt mir schwer, administrative Informationen von Ämtern (z.B. Öffnungszeiten, Zuständigkeiten, Formulare etc.) herauszufinden.	0	0	1	1
---	---	---	---	---

Fortsetzung: Itempool des WSP und Itemverrechnung

Skalen und dazugehörige Items	Verrechnung pro Item			
	Ja	Eher ja	Eher nein	nein
Bei unbekanntem Internetseiten finde ich mich auf Anhieb zurecht.	1	1	0	0
Bei unbekanntem Internetseiten finde ich schnell die Informationen, die ich brauche.	1	1	0	0
Die Koordination mehrerer dringender Aufgaben überfordert mich.	0	0	1	1
Bei neuen Aufgaben gewinne ich rasch einen Überblick darüber, welche Unterlagen ich benötige.	1	1	0	0
Meine Freunde würden mich als sehr strukturiert in der Bearbeitung von Aufgaben bezeichnen.	1	1	0	0
Bevor ich eine Seminararbeit schreibe, strukturiere ich den Textinhalt nach Kapiteln.	1	1	0	0
Bevor ich mit dem Schreiben einer Arbeit beginne, überlege ich mir den strukturellen Aufbau der Arbeit.	1	1	0	0
Wenn mehrere Aufgaben gleichzeitig zu bearbeiten sind, behalte ich trotzdem mühelos den Überblick über diese.	1	1	0	0
Beim gleichzeitigen Bearbeiten mehrerer Aufgaben verliere ich leicht den Überblick über den Fortschritt der einzelnen.	0	0	1	1
Beim Stellen von Anträgen muss ich Unterlagen nachreichen, da ich zu Beginn nicht immer alle wichtigen dabei habe.	0	0	1	1
Ich führe einen Kalender, um wichtige Termine zu notieren.	1	1	0	0
Bei der Aufbewahrung von Dokumenten halte ich eine exakte Ordnung ein (z.B. Dokumentenmappe/Ordner). -> auch auf Gewissenhaftigkeit prüfen	1	1	0	0
Wenn ich eine Vorlesung nicht besuchen kann, organisiere ich mich vorab mit einem Studienkollegen um die Mitschriften an Semesterende auszutauschen.	1	1	0	0
Beim gleichzeitigen Bearbeiten mehrerer Aufgaben verliere ich leicht den Überblick, was wann fertig sein sollte. -> auch auf Zeitmanagement/Organisationstalent prüfen	0	0	1	1

Fortsetzung: Itempool des WSP und Itemverrechnung

Skalen und dazugehörige Items

Verrechnung pro Item

Ja Eher ja Eher
nein nein

O5: Informiertheit

Informiertheit: vorab informieren, über bestimmte Themen genauer informieren, sich bestimmtes „nötiges“ Wissen aneignen, realistische Vorstellung vom Studium und dessen Inhalten, realistische Erwartungen an das Studium +

Ich bin über das Serviceangebot der Universität (z.B. Student Point, ÖH, Studienrichtungsververtretung) informiert.	1	1	0	0
Nach der Inskription musste ich feststellen, dass sich meine Erwartungen vor Studienbeginn nicht mit den tatsächlichen Inhalten des Studiums decken.	0	0	1	1
Andere sind früher über Abgabetermine/Fristen informiert als ich.	0	0	1	1
Andere sind über Terminänderungen meist früher informiert als ich.	0	0	1	1
Vor Studienbeginn habe ich mich intensiv mit dem Studienplan auseinandergesetzt.	1	1	0	0
Ich informiere eher die anderen über organisatorische Fragen zum Studium als umgekehrt.	1	1	0	0
Vor Studienbeginn habe ich mich ausführlich über die einzelnen Inhalte der Lehrveranstaltungen informiert.	1	1	0	0
Sind verpflichtende Wahlfächer im Studium zu belegen, informiere ich mich frühzeitig über die Möglichkeiten.	1	1	0	0
Ich bin über das Studienangebot der Universität gut informiert.	1	1	0	0

Arbeitshaltungen

A1: Gewissenhaftigkeit

Arbeitsdefinition Gewissenhaftigkeit: Termine regelmäßig wahrnehmen, Lehrveranstaltungen regelmäßig besuchen, immer gut vorbereitet sein, immer auf dem aktuellen Wissensstand sein, Pünktlichkeit, immer Mitschreiben, eine „saubere“ Mitschrift haben, alle wichtigen Punkte berücksichtigen, die Hausübung selber machen und nicht abschreiben, sich an den eigenen Zeitplan halten, Kontakte nicht abreißen lassen sondern pflegen, Rücksprache halten, um sich zu vergewissern, dass alles so passt, diszipliniertes Arbeiten, +

Ich versäume grundsätzlich keine Lehrveranstaltung.	1	1	0	0
---	---	---	---	---

Fortsetzung: Itempool des WSP und Itemverrechnung

Skalen und dazugehörige Items	Verrechnung pro Item			
	Ja	Eher ja	Eher nein	nein
Ich treffe eine Entscheidung erst nach reiflicher Überlegung. = <i>Kritisches Item, für RM-Analyse vormerken</i>	1	1	0	0
Beim Einreichen von Anträgen kümmere ich mich erst sehr spät um die notwendigen Dokumente. -> <i>auch auf Organisationstalent/Informiertheit prüfen</i>	0	0	1	1
Wenn ich mir einmal einen bestimmten Zeitpunkt zur Fertigstellung einer Aufgabe gesetzt habe, halte ich diesen ein.	1	1	0	0
Selbst gesetzte Fristen verschiebe ich regelmäßig.	0	0	1	1
Es fällt mir schwer, mich an meine eigenen Zeitpläne zu halten.	0	0	1	1
Ich achte darauf, dass meine Mitschrift vollständig ist.	1	1	0	0
Auch wenn eine Arbeit eher uninteressant ist, bringe ich sie zu Ende, wenn ich sie zugesagt habe.	1	1	0	0
Für Abgabetermine bin ich regelmäßig zu spät dran. -> <i>auch auf Zeitmanagement prüfen</i>	0	0	1	1
Bei der Informationssuche im Internet achte ich auf die Aktualität der Informationen. -> <i>auch auf Informiertheit prüfen</i>	1	1	0	0
Wichtige Unterlagen sortiere ich zuhause sorgfältig ein (z.B. Ringmappen/Ordner). -> <i>auch auf Organisationstalent prüfen</i>	1	1	0	0
Auch wenn mein Erscheinen nicht verpflichtend ist, besuche ich Vorlesungen regelmäßig. -> <i>auch auf Engagement prüfen</i>	1	1	0	0
Wenn ich einen Zeitplan erstelle, halte ich mich daran.	1	1	0	0
Wenn ich mir pro Lerntag eine bestimmte Lernmenge (z.B. Seitenanzahl) vorgenommen habe, dann halte ich mich daran.	1	1	0	0
Ich bereite mich ausreichend gut auf eine Prüfung vor und benötige daher keine unerlaubten Hilfsmittel. = <i>kritisches Item, für RM-Analyse vormerken</i>	1	1	0	0
Vereinbarte Termine nehme ich wahr, auch wenn ich die Zusage im Nachhinein gerne rückgängig machen würde.	1	1	0	0
Wenn es nicht kontrolliert wird, mache ich Hausaufgaben unvollständig.	0	0	1	1

Fortsetzung: Itempool des WSP und Itemverrechnung

Skalen und dazugehörige Items

Verrechnung pro Item

	Ja	Eher ja	Eher nein	nein
Vereinbarte Termine nehme ich wahr, auch wenn sich zwischenzeitlich interessantere Optionen bieten würden.	1	1	0	0
In schriftlichen Arbeiten verwende ich nur Quellen, die wissenschaftlichen Standards genügen.	1	1	0	0
Bei Lehrveranstaltungen mit E-Learning Unterstützung besuche ich die Internetplattform regelmäßig.	1	1	0	0
Wird der Hörsaal aller Voraussicht nach überfüllt sein, gehe ich früher hin und sichere mir einen Platz.	1	1	0	0
Wenn ich einmal meine Hilfe zusagt habe, dann ziehe ich dieses Angebot nicht zurück. -> auch auf Kollegialität prüfen	1	1	0	0
Vor Seminarbeginn besorge ich mir die empfohlene Literatur/Unterlagen. -> auch auf Informiertheit prüfen	1	1	0	0
Nach dem Fertigstellen einer schriftlichen Arbeit (z.B. Seminararbeit) lese ich den Text ein weiteres Mal Korrektur. -> auch auf Genauigkeit prüfen	1	1	0	0
Literatur recherchiere ich so ausführlich wie möglich.	1	1	0	0

A2: Genaues Arbeiten

Arbeitsdefinition: nur genauer Arbeitsstil ohne sonstige Einschränkungen +

Beim Verfassen einer schriftlichen Arbeit achte ich auf das äußere Erscheinungsbild der Arbeit (z.B. Deckblatt).	1	1	0	0
Arbeitsaufträge führe ich mit hoher Genauigkeit aus.	1	1	0	0
Bei der Aufgabenbearbeitung achte ich sehr auf die Details.	1	1	0	0
Arbeiten, die eine hohe Genauigkeit in der Ausführung erfordern, fallen mir leicht.	1	1	0	0
Beim Verfassen schriftlicher Arbeiten beachte ich vorgegebene Formatvorschriften genau.	1	1	0	0
Ich versuche so genau wie möglich zu arbeiten, da auch mehrere kleine Fehler zu Problemen führen können.	1	1	0	0
Offizielle Dokumente fülle ich detaillierter aus als andere.	1	1	0	0

Fortsetzung: Itempool des WSP und Itemverrechnung

Skalen und dazugehörige Items	Verrechnung pro Item			
	Ja	Eher ja	Eher nein	nein
Für Aufgaben, die eine hohe Genauigkeit erfordern, bin ich zu ungeduldig.	0	0	1	1
Beim Zitieren achte ich penibel auf die idente Formulierung des übernommenen Zitats.	1	1	0	0
Meine schriftlichen Arbeiten müssen perfekt gelayoutet sein. = <i>kritisches Item wg. Wort gelayoutet</i>	1	1	0	0
Ich investiere viel Zeit in die Formatierung meiner Arbeiten.	1	1	0	0
Ich neige zu Flüchtigkeitsfehlern beim Schreiben und/oder Rechnen.	0	0	1	1
Ich lerne eher überblicksmäßig und gehe wenig ins Detail. -> <i>auch auf Gewissenhaftigkeit prüfen</i>	0	0	1	1
Ich lese Texte genauer durch als die meisten anderen, damit mir wichtige Details nicht entgehen.	1	1	0	0
Es ist mir nicht so wichtig, ob in meinen Seminararbeiten die Kapitelüberschriften einheitlich formatiert sind.	0	0	1	1
Bei Arbeiten gehe ich sehr bedacht vor, sodass ich Flüchtigkeitsfehler bestmöglich vermeide. -> <i>auch auf Gewissenhaftigkeit prüfen</i>	1	1	0	0
Ich versuche grundsätzlich alles perfekt zu erledigen und Fehler zu vermeiden. -> <i>auch auf Gewissenhaftigkeit prüfen</i>	1	1	0	0
Meine Antworten auf Prüfungsfragen sind mitunter nicht vollständig, da ich die Angabe nicht genau genug gelesen habe.	0	0	1	1
Nach Möglichkeit vermeide ich detailreiche Aufgaben, die Präzision erfordern.	0	0	1	1

A3: Eigeninitiative

Arbeitsdefinition Eigeninitiative: von sich aus Feedback einholen, von sich aus Literatur zu einem bestimmten Thema recherchieren, wenn eine Führung ausfällt, von sich aus ein Alternativprogramm zusammenstellen +

Ich versuche laufend, meine Ideen und Meinungen in die Lehrveranstaltung einzubringen. -> <i>Engagement im Hintergrund</i>	1	1	0	0
---	---	---	---	---

Fortsetzung: Itempool des WSP und Itemverrechnung

Skalen und dazugehörige Items	Verrechnung pro Item			
	Ja	Eher ja	Eher nein	nein
Ich versuche laufend, neue Ideen und Möglichkeiten zur Optimierung meiner Arbeit zu finden.	1	1	0	0
Wenn ich in meiner Arbeit nicht weiterkomme, hole ich mir neue Ideen von außen (z.B. vom LV-Leiter oder von Kollegen).	1	1	0	0
Ich suche aktiv nach Feedback, um meine Arbeit zu verbessern.	1	1	0	0
Wenn ein Gruppentreffen ausfällt, kümmere ich mich noch vor den anderen um einen Ersatztermin. -> <i>auch auf Engagement prüfen</i>	1	1	0	0
Ich äußere mich zu einem bestimmten Thema erst, wenn ich durch den Lehrveranstaltungsleiter dazu aufgefordert werde.	0	0	1	1
Bei einem negativen Prüfungsausgang gehe ich zur Prüfungseinsicht.	1	1	0	0
Wenn ich Fragen habe, besuche ich die Sprechstunde. -> <i>auch negativ auf Selbstständig Arbeiten prüfen</i>	1	1	0	0
Wenn mir von den empfohlenen Praktikumsstellen keine zusagt, suche ich mir auf eigene Faust einen andere.	1	1	0	0
Wenn das Angebot ausdrücklich besteht, hole ich mir zu Semesterende Feedback zu meiner geleisteten Arbeit ein.	1	1	0	0
Ich wäre gerne besser darin, auch ohne Aufforderung meine eigenen Ideen für Projekte zu entwickeln.	0	0	1	1
Es fällt mir schwer, eigene Ideen zu entwickeln sondern halte mich lieber an die Vorgaben von außen. = <i>Kritisches Item – für RM-Analyse vormerken. Vielleicht eher Kreativität.</i>	0	0	1	1
Bei der Themenauswahl in Arbeitsgruppen halte ich mich mit Vorschlägen eher zurück.	0	0	1	1
Bei der Gruppenbildung bringe ich meine eigenen Ideen ein, wie die Arbeit aufgeteilt werden könnte.	1	1	0	0
Bei der Problembearbeitung in Arbeitsgruppen warte ich eher darauf, dass andere Lösungsvorschläge bringen.	0	0	1	1
Wenn Literaturquellen für eine Seminararbeit vorgegeben werden, suche ich unaufgefordert mehr. -> <i>auch auf Engagement prüfen</i>	1	1	0	0

Fortsetzung: Itempool des WSP und Itemverrechnung

Skalen und dazugehörige Items

Verrechnung pro Item

Ja	Eher ja	Eher nein	nein
----	---------	-----------	------

A4: Selbstständiges Arbeiten

Selbstständiges Arbeiten: alleine vorankommen, ohne Gruppe und Hilfe, eigenständiges Erlernen von z.B. Computerprogrammen; oder „Kleingruppe“ mit 2-3 Studierenden, die sehr selbständig und ohne viel Hilfe der LV-Leiter arbeiten +

Lieber arbeite ich selbständig an einer Aufgabe, als diese regelmäßig mit anderen abzusprechen. -> <i>auch negativ auf Teamorientierung prüfen</i>	1	1	0	0
Ich hole mir erst dann Unterstützung für die Lösung eines Problems, wenn ich es selbst mehrfach nicht geschafft habe.	1	1	0	0
Wenn ich mir für eine Arbeit zusätzliche Fertigkeiten aneignen muss, bringe ich es mir ohne Unterstützung anderer bei.	1	1	0	0
Bei Problemen in Gruppenarbeiten ziehe ich es vor, diese ohne der Hilfe des LV-Leiters zu lösen.	1	1	0	0
Lieber arbeite ich in einer Gruppe mit der Unterstützung anderer, als alleine an einer Aufgabe zu arbeiten. -> <i>auch auf Teamorientierung prüfen</i>	0	0	1	1
Zu Semesterbeginn fällt es mir schwer, ohne Unterstützung (z.B. durch Studienkollegen) alle nötigen Informationen zu sammeln. -> <i>auch auf Organisationstalent prüfen</i>	0	0	1	1
Auch bei kleineren Problemen ist für die mich die Meinung anderer unverzichtbar.	0	0	1	1
Ich ziehe es im Studium vor, eigenständig und autonom zu arbeiten. <i>Könnte aber auch negativ auf Teamfähigkeit, i.S. von Teamorientierung laden.</i>	1	1	0	0
Nach Erhalt einer Aufgabe fällt es mir leicht, diese alleine zu auszuführen.	1	1	0	0
Während der Bearbeitung einer Aufgabe suche ich regelmäßig die Unterstützung durch die Lehrperson. -> <i>auch positiv auf Eigeninitiative prüfen</i>	0	0	1	1
Selbst wenn ich nach fixen Vorgaben arbeiten kann, sichere ich mich durch mehrmaliges Nachfragen regelmäßig ab.	0	0	1	1
Bei Problemen frage ich erst dann nach, wenn ich selbst keine Lösung mehr sehe.	1	1	0	0

Fortsetzung: Itempool des WSP und Itemverrechnung

Skalen und dazugehörige Items

Verrechnung pro Item

Ja	Eher ja	Eher nein	nein
----	---------	-----------	------

M: Motivation

M1: Durchhaltevermögen

Arbeitsdefinition: nicht aufgeben, z.B. bei langweiligen oder langwierigen Aufgaben/Tätigkeiten,... +

Auch wenn ich viele Anstrengungen auf mich nehmen muss, lasse ich mich von meinen Plänen nicht abbringen.	1	1	0	0
Wenn erforderlich, lerne ich für Prüfungen von morgens bis abends.	1	1	0	0
Beim Einreichen von Anträgen (z.B. Stipendium) bleibe ich beharrlich und reiche diese notfalls mehrfach ein.	1	1	0	0
Wenn ich mich einmal für ein Seminar entschieden habe, dann bringe ich es trotz großen Aufwandes zu Ende.	1	1	0	0
Anfängliche Schwierigkeiten halten mich nicht von der Umsetzung meiner Pläne ab.	1	1	0	0
Wenn der Arbeitsaufwand für ein Referat größer ist als erwartet, suche ich mir ein weniger umfangreiches Thema.	0	0	1	1
Wenn ich das Interesse an einer Aufgabe verliere, lasse ich mich leicht ablenken.	0	0	1	1
Wenn eine Aufgabe uninteressant wird, lasse ich sie sehr schnell links liegen.	0	0	1	1
Wenn etwas schwieriger wird als zunächst gedacht, suche ich mir nach Möglichkeit eine leichtere Aufgabe.	0	0	1	1
Wenn eine Gruppenarbeit nicht schnell genug weitergeht, reduziere ich meinen eigenen Arbeitseinsatz.	0	0	1	1
Es kommt vor, dass ich Seminararbeiten aus mangelndem Interesse nicht fertig stelle.	0	0	1	1
Eine begonnene Aufgabe bringe ich grundsätzlich zu Ende, auch wenn ich mehr Zeit und Mühe investieren muss als geplant.	1	1	0	0

M2: Lernbereitschaft

Arbeitsdefinition: Kritik positiv sehen und als schöpferisches Potenzial verstehen, sich weiterbilden, etwas lernen wollen, Hilfe z.B. von LV-Leitern annehmen können,... +

Fortsetzung: Itempool des WSP und Itemverrechnung

Skalen und dazugehörige Items	Verrechnung pro Item			
	Ja	Eher ja	Eher nein	nein
Auf die Ratschläge anderer höre ich sehr ungern. -> auch auf Offenheit prüfen.	0	0	1	1
Bevor ich die Ratschläge anderer annehme, versuche ich es erst auf meine Weise.	0	0	1	1
Verbesserungsvorschläge von einer Lehrperson berücksichtige ich auch dann, wenn es für die Benotung keinen Unterschied macht.	1	1	0	0
Solange meine eigene Vorgehensweise funktioniert, lerne ich ungern eine neue (eventuell effizientere).	0	0	1	1
Auch wenn Kritik eine Möglichkeit zur Weiterentwicklung darstellt, fällt es mir schwer sie anzunehmen.	0	0	1	1
Rückmeldungen anderer versuche ich zu vermeiden, weil sie einen erhöhten Arbeitsaufwand bedeuten.	0	0	1	1
Es fällt mir schwer, Verbesserungsvorschläge anderer anzunehmen.	0	0	1	1
Ich schätze Kritik an meiner Arbeit und suche sie ausdrücklich. -> auch auf Eigeninitiative prüfen	1	1	0	0
Wenn eine Kritik an meiner Arbeit geäußert wird, frage ich interessiert nach.	1	1	0	0
Nach schlechten Noten gehe ich zur Prüfungseinsicht um herauszufinden, wie ich mich beim nächsten Mal verbessern kann. -> auch auf Eigeninitiative prüfen	1	1	0	0
Habe ich mir eine bessere Note erwartet, möchte ich meine Fehler wissen.	1	1	0	0
Auch bei einer guten Prüfungsnote, möchte ich meine Fehler wissen.	1	1	0	0
Routinetätigkeiten machen mir mehr Spaß als neue Aufgaben, in die ich mich erst einarbeiten muss.	0	0	1	1
Ich ziehe die Bearbeitung vertrauter Aufgaben vor, als mich auf neue Themengebiete einzulassen.	0	0	1	1
Kritik an einer von mir bereits abgeschlossenen Arbeit interessiert mich nicht.	0	0	1	1

Fortsetzung: Itempool des WSP und Itemverrechnung

Skalen und dazugehörige Items	Verrechnung pro Item			
	Ja	Eher ja	Eher nein	nein
Mir ist das Verständnis schwieriger Themen wichtig, auch wenn sie nicht prüfungsrelevant sind. -> auch auf Engagement prüfen	1	1	0	0
M3: Engagement:				
Arbeitsdefinition: gut Mitarbeiten, aus Interesse immer zu den LVs gehen, zusätzliche Praktika machen, wenn nicht nötig oder Stoff vertiefen bzw. mehr Lernen als nötig, sich intensiv mit etwas beschäftigen, Anstrengungsbereitschaft +				
Neben dem Studium sammle ich mehr Praxiserfahrung als vorgeschrieben ist.	1	1	0	0
Bei Lehrveranstaltungen mit E-Learning Unterstützung poste ich mehr Beiträge als notwendig wäre.	1	1	0	0
Ich versuche auch in meiner Freizeit mehr zu den Seminarinhalten zu erfahren.	1	1	0	0
Auch wenn es für die Absolvierung einer Lehrveranstaltung nicht notwendig ist, lese ich die empfohlene Literatur.	1	1	0	0
Ich versuche den Vorlesungsstoff durch zusätzliche Literatur zu vertiefen.	1	1	0	0
In investiere viel Zeit und Mühe in die Aufgabenstellungen der einzelnen Lehrveranstaltungen.	1	1	0	0
Wenn in den Ferien Kurse angeboten werden, die eine thematische Ergänzung zu meinem Studium darstellen, belege ich sie nach Möglichkeit.	1	1	0	0
Auf Prüfungen lerne ich gerade genug, um sie bestehen zu können.	0	0	1	1
Ich versuche im Studium den Aufwand zu minimieren, indem ich für Lehrveranstaltungen nur die Mindestanforderungen erbringe.	0	0	1	1
Ich nehme eine längere Studienzeit in Kauf, wenn ich dafür mein Wissen in dem Fach noch mehr vertiefen kann.	1	1	0	0
Ich lerne nicht nur für Prüfungen, sondern lese auch in der Freizeit Zusatzliteratur.	1	1	0	0
Zusatzaufgaben, die für den Studienfortschritt nicht notwendig sind, vermeide ich.	0	0	1	1

Fortsetzung: Itempool des WSP und Itemverrechnung

Skalen und dazugehörige Items	Verrechnung pro Item			
	Ja	Eher ja	Eher nein	nein
Für eine Präsentation ziehe ich mehr Literatur heran als notwendig wäre.	1	1	0	0
Bei einer Seminararbeit zitiere ich mehr Literaturquellen als notwendig wäre.	1	1	0	0
Aus Interesse besuche ich auch Vorlesungen anderer Studienrichtungen.	1	1	0	0
Aus Interesse belege ich mehr Fächer als gefordert.	1	1	0	0
Hausaufgaben bearbeite ich ausführlicher als andere. -> <i>auch auf Gewissenhaftigkeit prüfen</i>	1	1	0	0
In schriftliche Arbeiten zitiere ich mehr Literaturquellen als erwartet wird.	1	1	0	0

M4: Zielorientierung

Zielorientierung: das Ziel nicht aus den Augen verlieren +

Rückschläge und Verzögerungen können mich nicht von der konsequenten Verfolgung meiner langfristigen Ziele abbringen. -> <i>auch auf Frustrationstoleranz prüfen</i>	1	1	0	0
Ich setze mir zwar Tagesziele, verliere diese aber sehr schnell aus den Augen.	0	0	1	1
Nach einer Pause der Arbeitsgruppe, rege ich die Gruppe wieder zur gemeinsamen Arbeit an. -> <i>auch auf Gewissenhaftigkeit prüfen</i>	1	1	0	0
Wenn ich ein Ziel schwer erreichen kann, verfolge ich es nicht weiter.	0	0	1	1
Ich ändere meine Ziele laufend (noch vor der eigentlichen Zielerreichung).	0	0	1	1
Ich wäre gerne besser darin, meine Ziele hartnäckig zu verfolgen.	0	0	1	1
Ein Ziel aufzugeben, kommt für mich nicht in Frage.	1	1	0	0
Wenn ich mir ein Ziel gesetzt habe, möchte ich es um jeden Preis erreichen.	1	1	0	0

Fortsetzung: Itempool des WSP und Itemverrechnung

Skalen und dazugehörige Items	Verrechnung pro Item			
	Ja	Eher ja	Eher nein	nein
Wenn ich mir etwas vorgenommen habe, bringe ich es auch zu Ende.	1	1	0	0
Ein einmal gesetztes Ziel verliere ich nicht aus den Augen, selbst wenn die Zielerreichung länger dauert als geplant. -> <i>auch auf Durchhaltevermögen prüfen</i>	1	1	0	0
Langfristige Ziele verliere ich sehr schnell aus den Augen.	0	0	1	1
Wenn ich mich für etwas entschieden habe, verfolge ich das Ziel konsequent.	1	1	0	0
Meine Ziele zu erreichen fällt mir schwer, da immer etwas dazwischen kommt.	0	0	1	1
Von meinen Zielen lasse ich mich leider immer wieder abbringen.	0	0	1	1
Wenn ich mir für den nächsten Tag eine Aufgabe vornehme, bleibe ich dabei. -> <i>auch auf Gewissenhaftigkeit prüfen</i>	1	1	0	0
Es fällt mir schwer, langfristige Ziele tatsächlich in Angriff zu nehmen.	0	0	1	1
Wenn ich mir ein Lernziel setze (z.B. eine bestimmte Stoffmenge pro Tag zu lernen), dann erreiche ich es. -> <i>auch auf Gewissenhaftigkeit prüfen</i>	1	1	0	0
Ich setze mir sehr hohe Ziele und halte an deren Erreichung fest. -> <i>auch auf Zielsetzung prüfen</i>	1	1	0	0

M5: Zielsetzung

Arbeitsdefinition Zielsetzung: sich überhaupt Ziele setzen, sich bestimmte Ziele setzten, eine konkrete Vorstellung haben, wissen, was man will und wo man hin will, z.B. auch mit einem Konzept +

Bevor ich eine schriftliche Arbeit verfasse, lege ich das Ziel der Arbeit genau fest.	1	1	0	0
Ich habe eine konkrete Vorstellung davon, was ich erreichen möchte.	1	1	0	0
Meinen späteren Berufswunsch habe ich für mich bereits sehr genau definiert.	1	1	0	0

Fortsetzung: Itempool des WSP und Itemverrechnung

Skalen und dazugehörige Items	Verrechnung pro Item			
	Ja	Eher ja	Eher nein	nein
Zu Beginn eines Projektes habe ich eine klare Zielvorstellung davon, was ich erreichen möchte.	1	1	0	0
Ich setze mir keine Tagesziele, sondern lebe möglichst spontan in den Tag hinein. <i>Evt. auch negativ für Organisationsfähigkeit</i>	0	0	1	1
Ich setze mir keine langfristigen Ziele, sondern gehe sehr intuitiv in meiner Planung vor.	0	0	1	1
Ich habe mich noch nicht für eine bestimmte Ausbildung entschieden.	0	0	1	1
Ich verfolge ein für mich klar definiertes Ausbildungsziel.	1	1	0	0
Ich habe eine konkrete Vorstellung davon, was ich während meines Studiums erreichen möchte.	1	1	0	0
Es fehlt mir ein Ziel, auf das ich hinarbeiten möchte.	0	0	1	1
Ich weiß bereits sehr genau, was ich in den nächsten fünf Jahren erreichen möchte.	1	1	0	0
Ich setze mir sehr hohe Ziele und halte an deren Erreichung fest. <i>-> auch auf Zielorientierung prüfen</i>	1	1	0	0

M6: Erfolgsoversicht vs. Misserfolgsoversichtlichkeit

Erfolgsoversicht: Zuversicht, dass die eigene Handlung sehr wahrscheinliches erfolgreich sein wird, auch bei neuen Herausforderungen, Hoffnung auf Erfolg + vs. Misserfolgsoversichtlichkeit: Furcht vor Misserfolg –

Aus Furcht eine Prüfung nicht zu bestehen, lerne ich mehr als die meisten anderen.	0	0	1	1
Auch wenn ich den Prüfungsstoff bestmöglich gelernt habe, überkommen mich kurz vor der Prüfung Zweifel.	0	0	1	1
Anfängliche Schwierigkeiten lassen mich daran zweifeln, ein Projekt erfolgreich meistern zu können.	0	0	1	1
Ich bin zuversichtlich, mein Studium erfolgreich abschließen zu können.	1	1	0	0

Fortsetzung: Itempool des WSP und Itemverrechnung

Skalen und dazugehörige Items	Verrechnung pro Item			
	Ja	Eher ja	Eher nein	nein
Ich bin grundsätzlich zuversichtlich, die an mich gestellten Aufgaben bewältigen zu können.	1	1	0	0
Ich bin zuversichtlich, neue herausfordernde Aufgaben bewältigen zu können.	1	1	0	0
Vor neuen Aufgaben überkommen mich Zweifel, ob ich diese bewältigen kann.	0	0	1	1
Wenn ich ein Problem nicht gleich lösen kann, werde ich dies auch später nicht mehr schaffen.	0	0	1	1
Trotz Lernens bin ich wiederholt unsicher, Prüfungen bestehen zu können.	0	0	1	1
Wenn eine Prüfung meine bisherigen Lernmengen übersteigt, schüchtert mich das ein.	0	0	1	1
Neue Aufgaben sehe ich als eine Herausforderung, die ich gerne annehme.	1	1	0	0
Selbst wenn ich den Stoff in- und auswendig kenne, bin ich vor Prüfungen unsicher.	0	0	1	1
Wenn eine Arbeit schwieriger ist, als zunächst gedacht, werde ich unsicher.	0	0	1	1
Bei einer negativen Prüfungsbeurteilung bleibe ich zuversichtlich, sie beim nächsten Antritt zu bestehen. -> auch auf Frustrationstoleranz prüfen	1	1	0	0
Schwierige Prüfungen schiebe ich aus Respekt vor dem Prüfungsumfang möglichst lange hinaus.	0	0	1	1
Ich bin zuversichtlich, anfängliche Misserfolge am Ende doch noch in Erfolge verwandeln zu können.	1	1	0	0
Wenn meine Prüfungsnote schlechter ausfällt als zunächst gedacht, zweifle ich an meinen Fähigkeiten. -> auch auf Frustrationstoleranz prüfen	0	0	1	1
Selbst wenn Kritik konstruktiv und sachlich ist, fühle ich mich durch sie verunsichert. -> auch auf Frustrationstoleranz prüfen	0	0	1	1

Fortsetzung: Itempool des WSP und Itemverrechnung

Skalen und dazugehörige Items

Verrechnung pro Item

Ja	Eher ja	Eher nein	nein
----	---------	-----------	------

B Belastbarkeit

B 1: Frustrationstoleranz

Arbeitsdefinition: z.B. kann jemand gut mit negativer Rückmeldung umgehen, gibt jemand nach frustrierenden Erlebnissen oder negativer Rückmeldung sein Ziel/seine Idee nicht gleich auf) +

Scheitere ich an einer Prüfung, verliere ich schnell das Interesse an dem Fach.	0	0	1	1
Wenn ich an der Umsetzung einer Seminararbeit scheitere, brauche ich einige Anlaufzeit, um ein neues Thema in Angriff zu nehmen.	0	0	1	1
Selbst wenn ich viel Zeit und Mühe in eine Arbeit investiert habe, bin ich bereit, nochmals von vorne zu beginnen wenn nötig.	1	1	0	0
Wenn ich eine Prüfung beim wiederholten Antritt nicht bestanden habe, macht es wenig Sinn, es noch einmal zu versuchen.	0	0	1	1
Scheitere ich mehrmals an einer Prüfung, denke ich ans Aufgeben.	0	0	1	1
Bevor ich eine mühsam erstellte schriftliche Arbeit umschreibe, nehme ich die schlechte Note in Kauf.	0	0	1	1
Wenn ich die Kritik an meiner Arbeit nicht gerechtfertigt finde, melde ich mich auch mal von einem Seminar ab.	0	0	1	1
Kritik durch Lehrpersonen spornen mich grundsätzlich dazu an, mehr Leistung zu erbringen.	1	1	0	0
Nach einer negativen Prüfungsnote trete ich zum nächstmöglichen Prüfungstermin erneut an.	1	1	0	0
Wenn ich eine Prüfung nicht schaffe, ärgere ich mich noch lange darüber. -> auch auf Distanzierungsfähigkeit prüfen	0	0	1	1
Negatives Feedback bremst meinen Arbeitseifer.	0	0	1	1
Von negativem Feedback (z.B. von Lehrpersonen) lasse ich mich schnell entmutigen.	0	0	1	1
Wenn ich von einer Lehrperson negativ beurteilt werde, wähle ich im darauf folgenden Semester ein Seminar bei jemand anderem.	0	0	1	1

Fortsetzung: Itempool des WSP und Itemverrechnung

Skalen und dazugehörige Items	Verrechnung pro Item			
	Ja	Eher ja	Eher nein	nein
Durch Kritik an meiner Arbeit lasse ich mich nicht verunsichern.	1	1	0	0
Rückschläge in meinem Studium bringen mich nicht dazu aufzugeben.	1	1	0	0
Auch wenn mich zwischendurch Zweifel an der Sinnhaftigkeit meines Studiums überkommen, halte ich an meinem Ziel (dem Studienabschluss) fest.	1	1	0	0
Ich arbeite unbeirrt weiter, auch wenn mich zwischendurch der Wunsch überkommt, alles hinzuschmeißen.	1	1	0	0
Selbst wenn das Studium länger dauert als geplant, gebe ich nicht vor dem Abschluss auf.	1	1	0	0
Wenn organisationale Voraussetzungen an der Uni ein schnelles Voranschreiten im Studium verhindern, bin ich versucht, das Studium abzubrechen.	0	0	1	1
Wenn die Studiensituation ein schnelles Weiterkommen im Studium verhindert, breche ich das Studium möglicherweise frühzeitig ab.	0	0	1	1
Die Verzögerungen aufgrund der überfüllten Studiensituation lassen mich ans Aufgeben denken.	0	0	1	1

SK Soziale Kompetenz

SK1: Kommunikationsfähigkeit

Arbeitsdefinition: Kommunikationsfähigkeit = Wissen um Kommunikationsregeln und Bereitschaft zur Kommunikation: angemessene Kommunikation, Höflichkeit, Bereitschaft miteinander zu reden, ...+

Bei Meinungsverschiedenheiten bleibe ich freundlich und respektvoll.	1	1	0	0
Um wichtige Fragen für meine Seminararbeit zu klären, bespreche ich diese auch mit anderen Studienkollegen.	1	1	0	0
Ich wäre gerne besser darin, Kritik mit einer angemessenen Sachlichkeit zu äußern.	0	0	1	1
Ich wäre gerne besser darin, in Diskussionen anderen zuzuhören bevor ich mich zu Wort melde.	0	0	1	1
Es fällt mir auf, dass ich andere Personen oft unterbreche.	0	0	1	1

Fortsetzung: Itempool des WSP und Itemverrechnung

Skalen und dazugehörige Items	Verrechnung pro Item			
	Ja	Eher ja	Eher nein	nein
Ich unterbreche meine Gesprächspartner, wenn mir im Laufe eines Gesprächs etwas anderes interessant erscheint.	0	0	1	1
Probleme anzusprechen ohne jemanden zu kränken, sollte ich noch üben.	0	0	1	1
Es fällt mir eher schwer, andere Personen in einer Diskussion ausreden zu lassen.	0	0	1	1
In Gruppenarbeiten informiere ich die anderen regelmäßig über meinen Arbeitsfortschritt.	1	1	0	0

SK2: Kontaktfähigkeit

Arbeitsdefinition: auf andere Personen aktiv zugehen, ansprechen, Gespräch suchen, Kontakt knüpfen +

Es ist mir unangenehm, fremde Personen anzusprechen, wenn ich Informationen benötige.	0	0	1	1
Wenn nach dem Seminar noch Fragen offen sind, suche ich das Gespräch mit dem Lehrveranstaltungsleiter.	1	1	0	0
Ich knüpfe sehr schnell neue Bekanntschaften mit Studienkollegen.	1	1	0	0
Ich muss mich dazu überwinden, auf neue (unbekannte) Kollegen aktiv zuzugehen.	0	0	1	1
Anstatt auf andere zuzugehen, warte ich darauf, von anderen angesprochen zu werden.	0	0	1	1
Es fällt mir leicht, mit anderen Studenten in ein Gespräch zu kommen.	1	1	0	0
Es fällt mir leicht, mit fremden (unbekannten) Personen ein Gespräch zu beginnen.	1	1	0	0
Wenn ich in einer Lehrveranstaltung niemanden kenne, dann suche ich von mir aus Kontakt zu anderen Personen. -> auch auf Eigeninitiative prüfen	1	1	0	0
In Gruppenarbeiten fällt es mir leicht, mich auf neue Gruppenmitglieder einzustellen.	1	1	0	0

Fortsetzung: Itempool des WSP und Itemverrechnung

Skalen und dazugehörige Items	Verrechnung pro Item			
	Ja	Eher ja	Eher nein	nein
SK3: Teamfähigkeit				
Arbeitsdefinition: in Gruppen zusammenarbeiten können: z.B. Kommunikationsfähigkeit, Kontaktfähigkeit, Kollegialität, Kompromissbereitschaft, ... +				
Es fällt mir grundsätzlich leicht, gemeinsam mit anderen an einem Projekt zu arbeiten. = <i>Teamfähigkeit als solches</i>	1	1	0	0
Wenn sich eine Gruppenarbeit als zeitintensiv herausstellt, lasse ich eher die anderen arbeiten. = <i>Kollegialität</i>	0	0	1	1
In Gruppenarbeiten bin ich eher nachlässig und bearbeite meinen Teil unvollständig. = <i>Kollegialität</i>	0	0	1	1
Es ärgert mich, anderen Kollegen bei ihren Problemen helfen zu müssen. = <i>Kollegialität</i>	0	0	1	1
In Lerngruppen helfe ich gerne anderen. = <i>Kollegialität</i>	1	1	0	0
Ich könnte meine Studienkollegen bei der Prüfungsvorbereitung mehr unterstützen, tue es aber nicht. = <i>Kollegialität</i>	0	0	1	1
Wenn ich einen Prüfungsplatz nicht in Anspruch nehme, melde ich mich rechtzeitig ab. = <i>Kollegialität</i>	1	1	0	0
Ich borge meine Prüfungsunterlagen ungern anderen. = <i>Kollegialität</i>	0	0	1	1
Wenn ein Studienkollege nicht zur Vorlesung kommen kann, gebe ich ihm meine Notizen. = <i>Kollegialität</i>	1	1	0	0
Ich sehe es nicht als meine Aufgabe, anderen bei schwierigen Aufgaben zu helfen. = <i>Kollegialität</i>	0	0	1	1
Ich biete anderen meine jederzeit Hilfe an, wenn sie sich bei einer Aufgabe schwerer tun. = <i>Kollegialität</i>	1	1	0	0
Selbst verfasste Prüfungsskripten stelle ich nach bestandener Prüfung auch anderen Studienkollegen zur Verfügung. = <i>Kollegialität</i>	1	1	0	0

Fortsetzung: Itempool des WSP und Itemverrechnung

Skalen und dazugehörige Items	Verrechnung pro Item			
	Ja	Eher ja	Eher nein	nein
In Gruppenarbeiten nehme ich auf die Wünsche anderer Rücksicht. = <i>Kompromissbereitschaft</i>	1	1	0	0
In Gruppenarbeiten fällt es mir schwer, von meiner Meinung abzukommen und Kompromisse zu schließen. = <i>Kompromissbereitschaft</i>	0	0	1	1
Meinungsverschiedenheiten in der Gruppe diskutiere ich aus, auch wenn die Mehrheit auf meiner Seite ist. = <i>Kompromissbereitschaft und Kommunikationsbereitschaft</i>	1	1	0	0
Ich arbeite lieber alleine als in Gruppen. = <i>Teamorientierung</i>	0	0	1	1
Die Arbeit in einer Gruppe macht mir mehr Spaß als die Arbeit alleine. = <i>Teamorientierung</i>	1	1	0	0
Ich arbeite gerne in Teams. = <i>Teamorientierung</i>	1	1	0	0

E Emotionale Komponenten

E1 Distanzierungsfähigkeit

Arbeitsdefinition: in Bezug auf die eigenen Arbeiten neutral bleiben, objektiv beurteilen, Kritik nicht persönlich nehmen, etwas nicht zu wichtig nehmen, loslassen können +

Durch Kritik an meiner Arbeit fühle ich mich sehr schnell persönlich angegriffen.	0	0	1	1
Nach Erhalt eines negativen Feedbacks gelingt es mir, meine eigene Arbeit daraufhin kritisch zu prüfen.	1	1	0	0
Kritik an meiner Arbeit werte ich sehr schnell als ein Infragestellen meiner Fähigkeiten.	0	0	1	1
Es fällt mir schwer, Kritik an meiner Arbeit nicht persönlich zu nehmen.	0	0	1	1
Wenn ich mit einer Prüfungsnote nicht zufrieden bin, ärgere ich mich noch sehr lange darüber.	0	0	1	1
Eine missglückte Aufgabe beschäftigt mich im Nachhinein noch sehr lange.	0	0	1	1
Nach Erhalt eines negativen Feedbacks hindern mich die Gedanken daran einzuschlafen.	0	0	1	1
Es fällt mir schwer, mich auf eine neue Aufgabe einzulassen, wenn die letzte enttäuschend ausgefallen ist.	0	0	1	1

Fortsetzung: Itempool des WSP und Itemverrechnung

Skalen und dazugehörige Items	Verrechnung pro Item			
	Ja	Eher ja	Eher nein	nein
Bei einer missglückten Arbeit fällt es mir schwer, loszulassen und mich auf eine neue Aufgabe zu konzentrieren.	0	0	1	1
Negative Emotionen kann ich bei Seite schieben, um mich auf meine Arbeit zu konzentrieren.	1	1	0	0
Bei fachlichen Diskussionen gelingt es mir, meine Emotionen außen vor zu lassen, auch wenn mir das Thema sehr am Herzen liegt.	1	1	0	0
In Seminaren gelingt es mir, auch für mich sehr emotionale Themen möglichst sachlich zu bearbeiten.	1	1	0	0
Nach Erhalt einer Rückmeldung kann ich meine eigene Arbeit daraufhin möglichst objektiv prüfen.	1	1	0	0
Ich wäre gerne besser darin, meine eigene Arbeit objektiv zu beurteilen.	0	0	1	1
Ich kann meine eigene Leistung von einem möglichst objektiven Standpunkt aus beurteilen.	1	1	0	0

E2 Offenheit

Arbeitsdefinition: für neue Eindrücke und Erfahrungen, für neue/andere Ideen und Meinungen + Abgrenzung Lernbereitschaft: Bei Lernbereitschaft muss nicht nur die Bereitschaft vorhanden sein, sondern auch der Wille es umzusetzen. Bei der Offenheit geht es nur mal darum, sich etwas anzuhören, neue Eindrücke zu suchen.

Hinweis LK: Nähe mit Lernbereitschaft beachten.

Offenheit: eher die prinzipielle Bereitschaft für andere Meinungen und Ideen

Lernbereitschaft: eher diese zusätzlich zu übernehmen und auszuprobieren

Ich lege wenig Wert darauf, Erfahrungen in fachfremden Fächern zu sammeln.	0	0	1	1
Ich finde es interessant, andere (ungewohnte) Sichtweisen kennen zu lernen.	1	1	0	0
Ich suche Gespräche mit Personen, die eine andere Auffassung haben als ich.	1	1	0	0
Das Verständnis für andere Kulturen ist mir wichtig.	1	1	0	0
Ich nehme möglichst jede Gelegenheiten wahr, neue Arbeitsweisen und –techniken kennen zu lernen. -> auch auf Lernbereitschaft prüfen	1	1	0	0
Ich suche laufend Anregungen von außen und nehme sie in meine Arbeit auf. -> auch auf Lernbereitschaft prüfen	1	1	0	0

Fortsetzung: Itempool des WSP und Itemverrechnung

Skalen und dazugehörige Items	Verrechnung pro Item			
	Ja	Eher ja	Eher nein	nein
Ich suche den Erfahrungsaustausch mit anderen.	1	1	0	0
Andere würden mich als offen für die Ideen und Meinungen anderer beschreiben.	1	1	0	0
Kenne ich die Lösung für ein Problem, ziehe ich keine alternativen Vorschläge mehr in Betracht.	0	0	1	1
Ideen von Kollegen wie ich meine Arbeit verbessern kann, höre ich mir gerne an. -> <i>auch auf Lernbereitschaft prüfen</i>	1	1	0	0
Ich finde es positiv, wenn der Studienplan Kurse von anderen Studienrichtungen vorsieht.	1	1	0	0
Kurse anderer Studienrichtungen wecken mein Interesse.	1	1	0	0
Ich bin an Erkenntnissen anderer Studienrichtungen interessiert.	1	1	0	0
Ich suche nach Gelegenheiten, um neue Kulturen kennen zu lernen.	1	1	0	0
Es fällt mir schwer, eine gegenteilige Meinung zu akzeptieren. -> <i>auch auf Lernbereitschaft prüfen</i>	0	0	1	1
Wenn ich erst einmal einen Standpunkt vertrete, lasse ich mich schwer zu einer neuen Ansicht bewegen.	1	1	0	0
Andere Sichtweisen zu einem mir bekannten Thema sind mir wenig wichtig.	0	0	1	1

S Selbstbild

S1: Selbstbewusstsein

Arbeitsdefinition: zu eigenen Ideen/Meinung stehen, sich nicht verunsichern lassen, nicht an sich selbst zweifeln, an sich selbst glauben, den Mut haben sich in Lehrveranstaltungen zu melden +

Bei Diskussionen im Seminar äußere ich selbstbewusst meine Ideen.	1	1	0	0
Ich lasse mich durch die Meinung anderer leicht verunsichern.	0	0	1	1
Vertrete ich in einer Gruppe eine andere Meinung als die meisten anderen, fällt es mir schwer diese zu äußern.	0	0	1	1

Fortsetzung: Itempool des WSP und Itemverrechnung

Skalen und dazugehörige Items	Verrechnung pro Item			
	Ja	Eher ja	Eher nein	nein
Bei Gruppenarbeiten äußere ich meine Meinung, selbst wenn ich damit anderen widerspreche.	1	1	0	0
Es verunsichert mich, wenn andere bessere Leistungen erbringen als ich.	0	0	1	1
Wenn ich eine andere Meinung als die des Lehrveranstaltungsleiters vertrete, dann äußere ich diese.	1	1	0	0
In Gesprächssituationen mit Lehrpersonen fühle ich mich verunsichert.	0	0	1	1
Es ist mir unangenehm nachzufragen, wenn ich etwas nicht sofort verstehe.	0	0	1	1
Kritische Anmerkungen anderer verunsichern mich, sodass ich ein geplantes Vorhaben schnell wieder aufgebe.	0	0	1	1
Wenn ich den Prüfungsstoff gut beherrsche, fühle ich mich bei mündlichen Prüfungen selbstsicher. -> <i>auch auf Erfolgsoversicht prüfen</i>	1	1	0	0
Bei mündlichen Prüfungen trete ich selbstsicher auf.	1	1	0	0
Angesichts hervorragender Leistungen von Kollegen, beginne ich an meinen eigenen Stärken zu zweifeln. -> <i>auch negativ auf Erfolgsoversicht prüfen</i>	0	0	1	1
Der Gedanke etwas präsentieren zu müssen, verunsichert mich nicht.	1	1	0	0
Ich vermeide es in Vorlesungen durch mehrmaliges Nachfragen aufzufallen.	0	0	1	1
Ich melde mich ungern freiwillig, wenn eine Aufgabe an der Tafel zu bearbeiten ist.	0	0	1	1
Richtet ein Lehrveranstaltungsleiter vor versammelter Gruppe eine Frage an mich, werde ich nervös.	0	0	1	1
Wenn mich in einer Vorlesung etwas genauer interessiert, lese ich lieber in einem Buch nach statt Zwischenfragen zu stellen.	0	0	1	1
Wenn dem Lehrveranstaltungsleiter meiner Ansicht nach ein Fehler unterläuft, mache ich ihn darauf aufmerksam.	1	1	0	0

Abstract

Zusammenfassung. Die Güte eines Persönlichkeitsfragebogens ist zu einem Großteil durch die ihm zugrunde liegenden Items bestimmt. Zur Test- und Fragebogenkonstruktion gibt es eine Vielzahl von Literatur und Studien, jedoch kaum Instrumente, welche die zentralen Regeln kurz und prägnant zusammenfassen. Ein solches Instrument wird in dieser Arbeit präsentiert. Es wurde ein ‚Regelkatalog zur Fragebogen-Itemkonstruktion‘ entwickelt und Instrumente zur Qualitätssicherung während des Itemstellungsprozesses vorgestellt.

Anwendung und praktische Überprüfung findet der Regelkatalog in der Konstruktion eines Itempools für das Wiener Studieneignungs-Persönlichkeitsinventar (WSP). Der WSP dient der Erfassung studienrelevanter Eigenschaften und verfolgt einen förderdiagnostischen Ansatz zur Verbesserung der Studienberatung.

Grundlage für die Itemerstellung ist die vorangegangene Anforderungsanalyse zur allgemeinen Studierfähigkeit nach der Methode der Critical Incident Technique (Flanagan, 1954). Aufgrund der engen Verknüpfung der Daten aus der Anforderungsanalyse mit den Inhalten der Items konnte ein Itempool entwickelt werden, der optimal auf die gewünschte Zielgruppe zugeschnitten ist. Der Itempool besteht aus 351 Items zu sieben Skalen und zweiundzwanzig Subskalen. In den Prozess der Itemkonstruktion wurden zusätzliche Instrumente zur Qualitätssicherung integriert. Damit liefert diese Arbeit sowohl praktische Regeln als auch Handlungsanleitungen zur Konstruktion von Itempools.

Schlüsselwörter: Persönlichkeitsfragebogen, Item, Regelkatalog, Konstruktion, Critical Incident Technique, Anforderungsanalyse

Abstract. The quality of a personality questionnaire is to a great extent determined by the items it consists of. There exists a multitude of literature and studies about the construction of psychological tests and questionnaires, however, barely any instrument which summarizes the central rules in a short and concise manner. Such an instrument is presented in this paper. A ‘catalogue of rules for item construction’ has been developed and various instruments to secure the quality during the item construction process are discussed.

The catalogue was applied and reviewed by developing an itempool for the ‘Vienna study-aptitude personality inventory (WSP). The personality inventory measures study relevant personality traits and pursues an approach of furtherance to improve the counselling of study beginners.

The basis for the itempool construction are the critical incident data (Flanagan, 1954) from the preceding requirement analyses on general study-relevant requirements. Due to the close connection of this data with the contents of the items it was possible to develop an itempool specifically designed for the target group. The completed itempool consists of 351 items in seven scales and 22 subscales.

Additional instruments of quality assurance have been integrated into the process of item construction. Hence, this paper offers practical rules for item construction as well as suggestions for its application.

Key words: personality questionnaire, item, catalogue, construction, critical incident technique, requirement analysis

Curriculum Vitae

Akademischer Titel: Bakk. phil.

Geburtsort: Bludenz
Staatsbürgerschaft: Österreich

Universitäre Ausbildung

11/2007 Universität Wien
Abschluss Bakkalaureatsstudium Soziologie

02 – 07/2006 Universidad de Deusto, Bilbao, Spanien
Auslandssemester Erasmus

10/2001 Universität Wien
- Diplomstudium Psychologie
- Bakkalaureatsstudium Soziologie

06 – 08/2001 Sommeruniversität der Southwestern University, Oaxaca, Mexiko
Spanisch Intensivkurs

2000 – 2001 Southwestern University, Georgetown, USA,
Auslandsstudium
- Major Communication
- Major Spanish

1991 – 2000 Realgymnasium, Bludenz

Berufserfahrung

Seit 12/2008 Arbeitsmarktservice Österreich, Bundesgeschäftsstelle
Abteilung Arbeitsmarktpolitik für Frauen
Qualifizierte Sachbearbeitung, Fachgebiet:
Genderfragen, Arbeitsmarktdaten, Forschung und Controlling

06/2007 – 07/2008 Institut für Entwicklungspsychologie und Psychologische Diagnostik,
Fakultät für Psychologie, Universität Wien
Projektmitarbeit, Projekt:
Anforderungsanalyse nach der „Critical Incident Technique“ (Khorramdel,
Maurer & Kubinger, 2008)

07 – 09/2005 Personalberatung Lehmann GmbH, Frankfurt am Main
Psychologiepraktikum

01 – 12/2004 Österreichischer Versöhnungsfond (ÖVF), Wien
Mitarbeit
Registrierung, Aufbereitung, und Dokumentation von Akten ehemaliger
Zwangsarbeiter des nationalsozialistischen Regimes auf dem Gebiet der
heutigen Republik Österreich

- 05 – 06/2004 Institut für Jugendforschung, Wien
Interviewtätigkeit, Projekt:
Einflussfaktoren beim Umwelthandeln von Kindern und Jugendlichen unter Berücksichtigung der Rolle von Bildungsinstitutionen
- 07 – 09/2003 Österreichisches Bundesinstitut für Gesundheitswesen (ÖBIG), Wien
Soziologiepraktikum, Projektassistenz
- Evaluierung der Einrichtungen für Akutgeriatrie/Remobilisation und Palliativmedizin in Wien
 - Akutgeriatrie/Remobilisation in Österreich – Begleitung und Steuerung im Jahr 2003
- 08/2000 – 05/2001 Sprachlabor, Southwestern University, Georgetown, USA
Tutorin

Außeruniversitäre Weiterbildung

- 10/2009 Konferenz der schwedischen EU-Präsidentschaft, Stockholm
Kongressbeitrag:
Nationale Programmvorstellung: Frauen in Handwerk und Technik (FiT) (Session)
- 10/2009 Tagung Wege zur Entgeltgleichheit, Berlin
Workshopteilnahme
- 07/2008 International Congress of Psychology (ICP 2008), Berlin
Kongressbeitrag:
Frebort, M., Khorramdel, L., Kubinger, K.D. & Maurer, M. (2008). *An instruction guide for non psychologists to administrate a test: Experiences of a large scale assessment using teachers*. 29th International Congress of Psychology, Berlin, 20th-25th July 2008 (Symposium).
Khorramdel, L., Maurer, M. & Kubinger, K.D. (2008). *A requirement analysis of study specific demands – What requirements of ability and personality do students need to be successful?* 29th International Congress of Psychology, Berlin, 20th-25th July 2008 (Poster).
- 11/2003 Kongress „Integrating Europe“ der Österreichischen Gesellschaft für Soziologie (ÖGS), Wien
Teilnahme

Wien, Dezember 2009

Martina Maurer