



universität
wien

Diplomarbeit

Titel der Arbeit:

**Äquivalenzprüfung bei impliziten und expliziten
Motivmaßen anhand eines experimentellen Re-Test
Designs**

Verfasser

Jürgen Grafeneder

Angestrebter akademischer Grad

Magister der Naturwissenschaften (Mag. rer. nat.)

Wien, im September 2010

Studienkennzahl: A 298

Studienrichtung: Psychologie

Betreuer: Privatdoz. Mag. Dr. Michaela Wagner-Menghin

Danksagung

Rangereiht danke ich zuerst meinen Eltern, die mir das Studium und so viel mehr erst möglich gemacht haben. Sie haben mich immer unterstützt und werden das mit Sicherheit auch in Zukunft tun. Auch meinen Brüdern, Ing. Wolfgang Grafeneder und Ing. Dominik Grafeneder, möchte ich an dieser Stelle für ihre Unterstützung und den Zusammenhalt danken. Meiner gesamten Familie somit ein großes DANKESCHÖN!

Sehr großer Dank gebührt auch meiner Betreuerin, Privatdozentin Mag. Dr. Michaela Wagner-Menghin, die es mir trotz der schwierigen Rahmenbedingungen durch ihren persönlichen Einsatz ermöglichte, die Diplomarbeit so schnell abzuschließen.

Fehlen darf hier selbstverständlich auch meine Freundin, Mag. phil. Mag. rer. nat. Vanessa Maria Gruber, Bakk., nicht. Von ihr gab es laufende Unterstützung und diese nicht zu knapp. Vielen Dank fürs für mich da sein.

Speziellen Dank an meine Korrekturleser, Mag. Mag. Vanessa Maria Gruber, Bakk., Teresa Sposato, Bakk., und Ing. Wolfgang Grafeneder. Danke fürs aufmerksame Lesen der Seiten.

Abschließend noch vielen Dank an alle, die mich durchs Studium begleitet haben. „Das Schöne vergeht nie ganz, denn es geht in die Erinnerung ein.“ (Rainer Kaune)

Abstract

Wird der Vorgabemodus eines psychologisch-diagnostischen Verfahrens geändert, so muss vor der Verwendung der Normen der Ursprungsversion die Äquivalenz festgestellt werden. Ziel dieser Studie war es, die Äquivalenz einer Papier und Bleistift Testung (PBT) und einer PC-Testung bei impliziten und expliziten Motivmaßen zu untersuchen. Dafür wurde nach Wagner-Menghin (2003a) sowohl die psychometrische als auch die erfahrungspsychologische Äquivalenz ermittelt. Für die Erfassung der expliziten Motive wurde der PRF-D verwendet, die impliziten Motive wurden mit Hilfe des OMT erhoben. Zur Überprüfung der erfahrungsbezogene Äquivalenz diente eine Ratingskala. Beide Verfahren und die Ratingskala wurden insgesamt 177 Psychologie-Studenten der Universität Wien in einem experimentellen Retest Design vorgelegt. Die Ergebnisse zeigen, dass beim PRF-D nur die Reliabilität im Sinne der inneren Konsistenz einen signifikanten Unterschied ergibt (.72 in der PBT-Gruppe, .85 in der PC Gruppe). Die restlichen Werte der psychometrischen und erfahrungsbezogenen Äquivalenz weisen keinen signifikanten Unterschied auf. Beim OMT hingegen zeigt sich ein signifikanter Unterschied in der Skala „Leistung“ in den Mittelwerten, der Verteilung und der Rangreihung. Auch die erfahrungsbezogene Äquivalenz kann nicht als gegeben angenommen werden. Es zeigt sich wiederum die Notwendigkeit einer Äquivalenzprüfung bei einer Änderung des Vorgabemodus.

Abstract

When changing the mode of presentation of a psychological test, the equivalence of these two modes has to be verified before using the norms of the original. The goal of this study was to measure the equivalence between Paper and Pencil (PP) and Computer based tests (CBT) when it comes to measuring implicit and explicit motives. To do so, the approach of Wagner-Menghin (2003a) was used, hence the psychometric and the subjective equivalence were measured. For assessing the explicit motives the PRF-D was used, for assessing the implicit motives, the OMT was used. The subjective equivalence was obtained by a rating scale. 177 Psychology students were given both tests and the rating scale, using an experimental retest design. The findings show, that the PRF-D only differed in the internal consistency (.72 in the PP-Group and .85 in the CBT Group). The remaining values of the psychometric and the subjective equivalence did not differ significantly. The OMT on the other hand showed a significant difference in the scale "Achievement" in its means, the distribution and the ranked orders. The subjective equivalence cannot be assumed either. These findings again show the importance of verifying the equivalence when changing the mode of presentation.

Inhaltsverzeichnis

I. Einleitung	1
II. Theoretischer Teil	2
1. Äquivalenzmessung	2
1.1 Auswirkungen des PCs	3
1.2 Ebenen der Äquivalenz	5
1.3 Bisherige Studien zum Thema	6
2. Motivmessung	7
2.1 Explizite Motive	9
2.2 Implizite Motive	11
2.3 Zur nicht vorhandenen Korrelation	14
3. Hinführung zu den Hypothesen	16
III. Empirischer Teil	16
4. Ziele der Studie	16
4.1 Hypothese 1	17
3.2 Hypothese 2	18
3.3 Hypothese 3	18
4. Methoden	19
4.1 Untersuchungsplan	19

4.2 Verfahren	20
4.2.2 PRF-D	24
4.2.3 Fragebogen zur Erfassung der subjektiven Äquivalenz	27
4.3 Transformation der Verfahren	27
4.4 Durchführung der Untersuchung	30
4.5 Stichprobe.....	32
5. Ergebnisse	33
5.1 Deskriptive Statistiken	33
5.2 Korrelation impliziter und expliziter Motivmaße.....	35
5.3 Hypothese 1	36
5.4 Hypothese 2	39
5.5 Hypothese 3	43
6. Diskussion und Ausblick	45
7. Zusammenfassung	51
IV. Literaturverzeichnis	55
V. Anhang	63
Anhang A: Quellen	63
Anhang B: Tabellenverzeichnis	64
Anhang C: Abbildungsverzeichnis.....	66
Curriculum Vitae.....	67

I. Einleitung

Wird ein als Papier-Bleistift Test (PBT) konzipiertes psychologisch-diagnostisches Verfahren am Personal Computer (PC) vorgegeben und für dessen Auswertung die PBT - Normen herangezogen, so muss davor die Äquivalenz dieser zwei Vorgabemodi empirisch überprüft werden, da andernfalls fälschlicherweise deren Gleichwertigkeit ohne statistische Kontrolle angenommen wird (APA, 1986; de Beuckelaer & Lievens, 2009; Ford, Vitelli & Stuckless, 1996; Hofer, 1985; Kindsvater & Sturm, 2003; Kolen & Brennan, 2004; Kubinger, 1993, 2009a; Kveton, Jelinek, Voboril & Klimusova, 2004; Pommerich, 2004; Russell, Goldberg & O'Connor, 2003; Wagner-Menghin, 2003a). Diese als Äquivalenzmessung bezeichnete statistische Kontrolle muss den Kriterien der APA (1986) genügen, von der sie auch gefordert wird.

In der Motivforschung wird seit McClelland, Koestner und Weinberger (1989) von zwei verschiedenen Motiv-Systemen ausgegangen: (1) explizite, dem Bewusstsein zugängliche und damit direkt erfragbare Motive und (2) implizite, außerhalb des Bewusstseins operierende und damit nicht direkt erfragbare Motive. In der Literatur finden sich viele Belege für diese Unterteilung (Brunstein, Schultheiss & Grässmann, 1998; Spangler, 1992; Schultheiss & Brunstein 2005), wie zum Beispiel die fehlende Korrelation zwischen den beiden Motivmaßen (McClelland et al., 1989; Schultheiss et al., 2009).

Ob der Vorgabemodus (PC oder Papier und Bleistift) bei der Erhebung dieser beiden Motivmaße einen signifikanten Unterschied bewirkt, wurde bisher noch nicht empirisch geprüft. Mit Hilfe eines experimentellen Retest Designs untersucht diese Studie nun die Äquivalenz der beiden Vorgabemodi bei der Erhebung von impliziten und expliziten Motivmaßen.

II. Theoretischer Teil

1. Äquivalenzmessung

Um für die Auswertung eines als PBT konzipierten, jedoch am PC vorgegebenen psychologisch-diagnostischen Verfahrens die PBT Normen heranziehen zu dürfen, muss zuvor eine Äquivalenzmessung durchgeführt werden (APA, 1986; de Beuckelaer et al., 2009; Ford et al. 1996; Hofer, 1985; Kindsvater et al., 2003; Kolen et al., 2004; Kubinger, 1993, 2009a; Kveton et al., 2004; Pommerich, 2004; Russell et al., 2003; Wagner-Menghin, 2003a).

Die American Psychology Association fordert in ihren Richtlinien (APA, 1986, S. 19): „When interpreting scores from the computerized versions of conventional tests, the equivalence of scores from computerized versions should be established and documented before using norms or cutting scores obtained from conventional tests.“

Tatsächlich ist der Großteil der PC Tests eine eins zu eins Transformation einer PBT: die Fragen werden statt auf einem Blatt Papier auf einem Bildschirm präsentiert, als Eingabegeräte dient anstatt eines Stiftes meist die Maus und/oder die Tastatur (Ferrando & Lorenzo-Seva, 2005; Kindsvater et al., 2003; Kubinger, 1993, 2009b).

Viele dieser Test-Manuale sprechen das Thema der Äquivalenzprüfung erst gar nicht an (Merten & Ruch, 1996), manche Testautoren scheinen dieses Thema gar zu ignorieren (Kubinger & Farkas, 1991) und damit auch die von der APA (1986) geforderte Äquivalenzmessung. Es werden dadurch oft Normen verwendet, die von einer konventionellen Testvorgabe stammen (Kveton et al., 2007). Eine Metaanalyse von Mead und Drasgow (1993) zeigt aber, dass der Vorgabemodus zu einer psychometrischen Differenz in den korrespondierenden Testwerten führen kann, selbst das gemessene psychologische Konstrukt könnte demnach ein anderes sein.

1.1 Auswirkungen des PCs

Die Verwendung eines PC für die Vorgabe eines psychologisch-diagnostischen Verfahrens bewirkt unter anderem eine Verbesserung der Objektivität, da

1. sich die Testleiterunabhängigkeit erhöht: die Instruktionen werden standardisiert präsentiert, deren Verständnis kann anhand von Übungs-Items überprüft und die Instruktion gegebenenfalls wiederholt werden,
2. sich die Auswertungsobjektivität erhöht: Verrechnungsfehler fehlen und
3. sich die Interpretationsobjektivität erhöht: es stehen unmittelbare (grafische) Rückmeldungen mit leicht veränderbaren Normen zur Verfügung (Amelang & Schmidt-Atzert, 2006; Garb, 2007; Wagner-Menghin, 2003b).

Mit Hilfe eines PC kann auch die ökologische Validität einer Aussage erhöht werden, da es möglich ist, die geprüfte Eigenschaft unter komplexeren Bedingungen zu erfassen als dies bei einer klassischen PBT der Fall ist (Amelang et al., 2006; Kubinger, 2009b). Lange Durchführungszeiten sind für den Testleiter nicht automatisch unökonomisch. Die aufgrund der Ökonomie früher durchgeführten Gruppentestung können, mit Hilfe des PC, einer Einzeltestung gleich gesetzt werden, da eine Anwesenheit des Testleiters nicht zwingend erforderlich ist und jede zu untersuchende Person das psychologisch-diagnostische Verfahren in einer standardisierten Geschwindigkeit durchführen kann (Amelang, et al. 2006; Kubinger, 2009b). Die Testergebnisse sind weiters sofort verfügbar und der Psychologe kann während der Testung anderen Dingen nachgehen (Amelang, et al. 2006; Wagner-Menghin, 2003b). Des Weiteren wird das adaptive Testen, ein in der Probabilistischen Testtheorie verankertes Konzept, durch den Einsatz eines PC erleichtert und teilweise sogar erst ermöglicht (Kubinger, 2009b; Wagner-Menghin, 2003b). Die Verfälschbarkeit kann durch die technischen Möglichkeiten (z.B.: nach Ablauf der Testzeit ist ein Zurückblättern durch die Items nicht mehr möglich) ebenso erhöht werden (Wagner-Menghin, 2003b), wenngleich dies nicht automatisch bedeutet, dass ein Persönlichkeitsfragebogen am PC weniger verfälscht wird als bei einer PBT (Garb, 2007; Menghin & Kubinger, 1996).

Diese Verbesserungen des Gütekriteriums Ökonomie bringt die zwei wohl wichtigsten Gründe mit sich, ein psychologisch-diagnostisches Verfahren an einem PC anstatt auf einem Papier vorzugeben: die vereinfachte Administration und die dadurch entstehende Kostenreduktion (Amelang et al., 2006; Kubinger, 2009b; Wagner-Menghin, 2003b), auch wenn diese als alleinige Begründungen kritisiert werden (Garb, 2007; Kubinger, 2009a, 2009b).

Es sei auch noch auf die bestehende Gefahr hingewiesen, dass durch die vereinfachte, automatische Administration und gleichsam automatische Auswertung es zu einer automatischen Erstellung von Gutachten kommen könnte (Kubinger, 2009a). Dies widerspricht sowohl den Richtlinien des Berufsverbandes Deutscher Psychologen (1988) als auch den allgemeinen Regeln zur Erstellung eines psychologischen Gutachtens, welche auch die Beachtung der Begleitumstände einfordern (Amelang et al. 2006, Kubinger, 2009b).

Die Vorwürfe, die Computerdiagnostik sei nicht „human“, Befürchtungen die zu untersuchenden Personen hätten Angst vor dem Computer und eventuelle grundlegende Probleme mit den Eingabegeräten Maus und/oder Tastatur (Booth, 1995; Garb, 2007) stellen sich heute als nicht (mehr) haltbar heraus, da die meisten relevanten Populationen über ausreichend Erfahrung mit dem Computer und dessen Eingabegeräte besitzen (Amelang et al., 2006; Kubinger, 2009b). Weiters konnte gezeigt werden, dass es zu keiner unterschiedlichen Ausprägung beim Faktor Testangst durch die Verwendung eines PC kommt (Klinck, 2002). Es kann sogar vorkommen, dass es erst durch eine Testung mittels eines als „moderner“ und „zeitgemäßer“ angesehenen PC zu einer (erhöhten) Akzeptanz seitens der Testperson kommt (Amelang, et al., 2006 Kubinger, 2009a, 2009b). Zudem kann eine PC-Vorgabe Menschen mit Beeinträchtigungen (z.B. einer Sehschwäche) die Bearbeitung eines Tests mittels spezieller Eingabegeräten und den erweiterten Möglichkeiten der Präsentation des Testmaterials erst ermöglichen beziehungsweise erheblich erleichtern (Garb, 2007). Auch kann die Sprache leichter geändert werden als bei herkömmlicher PBT (Bachmann, 2003). In manchen Fachbereichen der Psychologie, wie zum Beispiel der neuropsychologischen Diagnostik, ist der Einsatz eines Computers mittlerweile nicht mehr

weg zu denken. Dies kann zugleich auch gegen die Kritik, ältere Menschen würden sich nicht für eine Computerdiagnostik eignen, eingebracht werden (Kubinger, 2009a; Willmes, 2003).

Der Kritik, bei einer Computer-administrierten Testung würden Informationen zur Körpersprache fehlen, kann mit einer speziellen Test-Administration (z.B. genaue Verhaltensbeobachtung, Gespräche vor und nach der Testung) entgegen gewirkt werden (Garb, 2007). Zudem können Testleitereffekte aufgrund der verringerten Interaktion reduziert werden (Wagner-Menghin, 2006b).

Nach Jäger (1990) gibt es keinen Prozess innerhalb des gesamten psychologisch-diagnostischen Prozesses, der ungeeignet wäre, ihn innerhalb der Computerdiagnostik zu realisieren. Dies legt den Schluss nahe, dass die meiste Kritik gegen den PC auf einer Unvertrautheit mit dem Material oder einer fehlerhaften Administration aufbaut (Booth 1995; Kubinger, 2009b).

Einzig haptische Testaufgaben, wie zum Beispiel das Arbeiten mit Bauklötzen, sind nur schwer und mit speziell entwickelter Hardware zu realisieren (Kubinger, 2009b) und zeigen, dass der Computertestung (momentan noch) Grenzen gesetzt sind. Auch benachteiligt die meist schriftliche Instruktion auditive Typen (Fuchs, 2000; Wagner-Menghin, 2003b), was aber wohl eher an der nicht mediumsgerechten Umsetzung eines psychologisch-diagnostischen Verfahrens liegt. Dies verdeutlicht, dass das Potential eines Computers noch nicht zur Gänze ausgereizt wird (Kubinger, 2009a, 2009b), wenn auch viele Verfahren dazu bereits gute Ansätze zeigen (Kubinger, 2009b; Wagner-Menghin, 2003b).

1.2 Ebenen der Äquivalenz

Es werden bei psychologisch-diagnostischen Verfahren zwei Ebenen der Äquivalenz unterschieden: (1) die erfahrungsbezogene Äquivalenz und (2) die psychometrische Äquivalenz (Rammstedt, Holzinger & Rammsayer, 2004; Wagner-Menghin, 2003a). Die erfahrungsbezogene Äquivalenz umfasst dabei die „Aspekte der Wahrnehmung, Einstellung und emotionalen Bewertung des Verfahrens durch die Testpersonen“ (Wagner-Menghin,

2003a, S. 33). Die psychometrische Äquivalenz meint „die Gleichwertigkeit der Parameter der Item- und Skalenkennwerte sowie der Reliabilitäts- und Validitätskennwerte bei verschiedenen Vorgabebedingungen“ (Wagner-Menghin, 2003a, S. 33).

Um als psychometrisch äquivalent zu gelten, muss die empirische Überprüfung eines psychologischen Verfahrens demnach folgenden Kriterien genügen: (1) die deskriptiven Statistiken (Mittelwert, Varianzen, Verteilung, Rangreihung) müssen bei beiden Vorgabemodi äquivalent sein; (2) die Konstruktvalidität muss äquivalent sein; (3) die Reliabilität muss äquivalent sein (APA, 1986; Kubinger, 2009a, 2009b; Kveton et al., 2007; Wagner-Menghin, 2003a). Für die Bestimmung der erfahrungsbezogenen Äquivalenz werden qualitative Methoden (wie z.B. Interviews oder Beobachtungen) herangezogen (Rammstedt et al., 2004; Wagner-Menghin, 2003a).

Nach Bartram (1994) empfiehlt sich ein experimentelles Retest Design mit vier Gruppen für die Überprüfung der psychometrischen Äquivalenz wobei zwei Gruppen zur Kontrolle dienen (PC-PC und PBT-PBT) und zwei als Kreuzkonditionen (PC-PBT, PBT-PC).

1.3 Bisherige Studien zum Thema

Studien zu diesem Thema zeigten, dass bei der Transformierung eines Fragebogens die beiden Vorgabemodi, Papier-Bleistift oder PC, typischerweise keine Unterschiede bewirken (Merten et al., 1996; Pineseault, 1996; Watson, Thomas & Anderson, 1992). Jedoch kann es bei der Transformation von kognitiven Leistungstests (Kubinger et al., 1991; Neubauer, Urban & Malle, 1991) sowie bei reinen Speed-Tests (Greud & Green, 1986) zu Unterschieden kommen. Dieses Bild wird auch von der von Mead et al. (1993) durchgeführten Metaanalyse gestützt, die bei sorgfältig konstruierten kognitiven Leistungstests keine, bei Speed-Tests jedoch erhebliche Unterschiede zwischen den beiden Vorgabemodi zeigte.

Zur Äquivalenz von PBT und PC Testungen bei der Erhebung von impliziten und expliziten Motivmaßen liegen noch keine Studien vor. Die Vorteile einer computerisierten Vorgabe eines psychologischen Verfahrens zur Messung von impliziten oder expliziten Motivmaßen

liegen auf der Hand. Neben den schon oben erwähnten Vorteilen, wie der vereinfachten Administration, der Erhöhung der Objektivität und dergleichen, sollte hier vor allem auf die Möglichkeit für die Auswertung, der für die Erhebung von impliziten Motivmaßen erforderlichen Geschichten mit projektiven Inhalt, hingewiesen werden. Neben dem Fehlen von Uneindeutigkeiten aufgrund der Handschrift (Garb, 2007), wäre eine automatisierte Auswertung der von der Testperson assoziierten Bildgeschichten anzudenken, welche einen großen Einfluss auf die von zum Beispiel Rollet (2006) kritisierte mangelnde Objektivität der Auswertung hätte.

2. Motivmessung

McClelland (1987) definiert ein Motiv als

“a recurrent concern for a goal state based on a natural incentive a concern that energizes, orients, and selects behavior” (p. 590).

Mit dieser Definition postuliert McClelland (1987), dass es sich bei Motiven um immer wiederkehrende Anliegen („recurrent concern“) handelt, die auf natürlichen Impulsen beruhen. Das Wort „recurrent“ meint hier, dass eine motivationale Disposition bedeutet, immer wieder über das Ziel („gold state“) nachzudenken. Der Fokus liegt also auf dem Ziel, dem „gold state“ selbst,. Die Mittel zu dessen Erreichung (also die Verhaltensweisen, die zu dem erwünschten Ergebnis führen) sind demnach nicht Teil der Definition. Mit dem Bezug auf natürliche Impulse („natural incentives“) nimmt McClelland (1987) an, dass es nur wenige grundlegende motivationale Systeme gibt, dass deren Effekt auf das Verhalten durchdringend ist und dass eine sehr enge Beziehung zum Gemütszustand („emotional state“) besteht. Auch wird durch die Verwendung des Wortes „concern“, also Belang, von McClelland (1987) angedeutet, dass nicht alle Motive notwendigerweise bewusst sein müssen. Einige wichtige Motive werden auch als unbewusst angenommen. Motive sind in dieser Definition auch breiter gefasst als eine Intention (McClelland, 1987).

Wie schon in der Motiv-Definition von McClelland (1987) angedeutet, geht man heute von zwei grundlegenden motivationalen Systemen aus, die das menschliche Verhalten formen: (1) ein explizites motivationales System, basierend auf Zielen die dem Bewusstsein zugänglich sind und (2) ein implizites motivationales System, welches auf affektiven Erfahrungen basiert und außerhalb des Bewusstseins operiert (Brunstein et al., 1998; McClelland et al. 1989; Spangler, 1992; Schultheiss et al. 2005).

Folgende Untersuchungsergebnisse stützen diese Differenzierung in zwei verschiedene motivationale Systeme: (1) implizite und explizite Motive korrelieren nicht signifikant miteinander (McClelland et al., 1989; Schultheiss et al., 2009), (2) sie sagen verschiedene Arten von Verhalten voraus (Langens, 2007; McClelland et al., 1989), (3) beide Systeme reagieren auf verschiedene Anreize (McClelland et al., 1989) und (4) sie entwickeln sich zu unterschiedlichen Zeitpunkten (Brunstein, 2006; McClelland et al., 1989).

Es wird im Allgemeinen zwischen drei Basismotiven unterschieden, auf die sich die bisherige Forschung konzentriert hat: (1) Bindung/Geselligkeit, (2) Leistung und (3) Macht/Dominanz (Atkinson, 1964; Scheffer, Kuhl & Eichstaedt, 2003; McClelland, Atkinson, Clark & Lowell, 1953; McClelland et al., 1989, Schultheiss et al., 2010).

Das Bindungsmotiv ist das Bedürfnis zur Herstellung, Aufrechterhaltung oder Wiederherstellung einer engen, freundlichen Beziehung, sowie der Wunsch zu kooperieren und andere Menschen zu lieben (Murray, 1938; Schultheiss et al., 2007). Nach Hall, Stanton und Schultheiss (2010) handelt es sich dabei um ein fundamentales, evolutionäres Bedürfnis nach vertrauten Personen (engl. „kin“), um Sicherheit und Schutz zu garantieren. Als mögliche Ziele dieses Motivs werden folgende genannt: jemandem nahe zu sein, sich auszutauschen, zu kooperieren und mit anderen befreundet zu sein (Murray, 1938).

Das Verhalten einer Person ist dann leistungsmotiviert, wenn es sich mit einem Gütestandard auseinandersetzt, dem „standard of excellence“ (McClelland et al, 1953; Schultheiss et al., 2007). Das Leistungsmotiv gilt heute als das am intensivsten untersuchte Motiv. Was jedoch genau unter Leistung zu verstehen ist, wird durch den jeweiligen

kulturellen und sozialen Kontext beziehungsweise Zugehörigkeit geprägt (Brunstein & Heckhausen, 2006). Schwierige Aufgaben zu meistern oder andere im Wettbewerb zu übertreffen sind zum Beispiel für Murray (1938) Merkmale leistungsmotivierten Handelns.

Das Machtmotiv schließlich wird als das Bedürfnis gesehen, Einfluss auf andere zu haben (Schultheiss, et al. 2007) bzw. Kontrolle über die eigene Umgebung auszuüben (Murray, 1938). Als Handlungen werden von Murray (1938) unter anderem Beeinflussen, Überzeugen, Kontrollieren und Führen genannt.

2.1 Explizite Motive

2.1.1 Definition

Explizite Motive erfassen das Selbstbild einer Person, sie stehen im Zusammenhang mit den bewussten Zielen und Erwartungen einer Person (Brunstein, 2006; McClelland et al., 1989). Es handelt sich also um explizite, selbst-attribuierte Bedürfnisse und Ziele, mit denen sich eine Person identifiziert (Brunstein, 2006; Hall, 2010; Schultheiss et al., 2007).

Diese motivationale Orientierung und Ziele schreibt sich die Person selbst zu. Sie kann diese verbal ausdrücken und hat kontrolliertes Verhalten zur Folge (Schultheiss et al., 2008). Explizite Motive sind daher bewusst zugänglich (Scheffer et al., 2007), der Fokus liegt auf dem Ergebnis einer geplanten Handlung (Bauman et al., 2005). Dieses kognitiv-basierte motivationale System wird durch persönliche Ziele repräsentiert (Brunstein et al., 1998).

Explizite Motive beeinflussen geplantes Verhalten, bewusste Einstellungen, Urteile, Wahlen und Entscheidungen (Langens, 2007). Ein hohes explizites Leistungsmotiv bedingt die Entscheidung weiter zu arbeiten, selbst wenn eine Rückmeldung andeutet, man habe eine unterdurchschnittliche Leistung im Vergleich zu Anderen erbracht (Brunstein & Maier, 2005).

2.1.2 Erhebungsform

Da die expliziten Motive bewusst (Brunstein, 2006; Langens, 2007; McClelland et al., 1989; Scheffer et al., 2007) und verbal zugänglich sind (Schultheiss et al., 2008), werden diese

auch mit Hilfe von Fragebögen erhoben (Schultheiss et al., 2001). Explizite Motive werden auch manchmal durch den Umstand, dass sie direkt (d.h. per Fragebogen) erfragt werden, definiert (Schultheiss et al., 2007). Es wird also direkt nach den Ausprägungen der Motive gefragt, wie sehr eine Person motiviert ist, Macht zu haben, zu Leisten oder Beziehungen mit anderen einzugehen (Schultheiss et al., 2007).

Die ersten Fragebögen, die explizite Motive zu erheben versuchten, basieren auf der Klassifikation der psychogenen Bedürfnisse nach Murray (1938). Der Fragebogen Personality Research Form (PRF) von Jackson (1974) ist hierbei der bekannteste und zugleich auch der sorgfältigst konstruierte (Brunstein, 2006; Schultheiss et al., 2007).

2.1.3 Korrelate:

Das explizite Leistungsmotiv konnte die Wahl, weiter zu machen, vorhersagen, speziell nach der Mitteilung, die zu untersuchende Person hätten eine Aufgabe schlechter als andere gelöst, ohne ihnen die Leistung bei der aktuellen Aufgabe zu verraten (Schultheiss, 2008). Auch Brunstein et al. (2005) zeigten, dass ein hohes explizites Leistungsmotiv die Entscheidung, weiter zu arbeiten, nach negativem Feedback vorhergesagt werden kann. Explizite Motive sagen Urteile und Wahlen eher voraus als Leistungen (Schultheiss, 2008) und eher respondentes (d.h. external kontrolliertes) denn operantes (d.h. selbst-initiiertes) Verhalten (Scheffer et al., 2007). Sie sind ein guter Prädiktor für schulische Leistungen, wenn für die Bewertung soziale Normen herangezogen werden (Brunstein, 2006). Sie beeinflussen deklaratives Verhalten, also verbal präsente Entscheidungen, Urteile und Wahlen, welche auf dem Selbstbild der Person beruhen (Schultheiss et al., 2010).

2.1.4 Kritik:

Im Vergleich zu impliziten Motiven (siehe Kapitel „2.2 Implizite Motive“), zeigt dieser Ansatz eine höhere Ökonomie (insbesondere in Bezug auf die Verrechnung), Auswertungsobjektivität und Reliabilität der Motivmessungen (Brunstein, 2006).

Jedoch sind die Messung expliziter Motive von den kognitiven Fähigkeiten, Selbstdarstellungstendenzen und situativen Einflüssen stark abhängig (Brunstein, 2006),

sodass bezweifelt wird, dass die Methode der direkten Befragung Rückschlüsse auf die Motive ermöglicht (McClelland, 1958). Beim Leistungsmotiv sei es selbst mit Leistungsmessungen und Verhaltensbeobachtungen nicht möglich die zu Grunde liegenden Motive zu erfassen (McClelland, 1958). Zudem müssen die, von Personen sich selbst zugeschriebene, Motive nicht unbedingt die sein, die ihrem Handeln zu Grunde liegen (Brunstein, 2006).

2.2 Implizite Motive

2.2.1 Definition

McClelland (1987) definiert implizite Motive als früh gelernte, emotional getönte Präferenzen, sich immer wieder mit bestimmten Formen von Anreizen auseinander zu setzen. Diese Motive werden als implizit bezeichnet, weil sie außerhalb des Bewusstseins operieren (Brunstein et al., 1998; Schultheiss, 2010, Schultheiss et al., 2008). Ein Grund dafür findet sich in der postulierten frühen Entwicklung und der dadurch nicht sprachlichen Repräsentation (McClelland & Pilon, 1983; McClelland et al., 1989). Des weiteren stellen implizite Motive ein umfangreiches Netzwerk kognitiv-emotionaler Repräsentationen von relevanten Zielen, Bedürfnissen und Verhaltensalternativen dar. Sie sind aufgrund dieses großen Umfangs und der oben erwähnten emotionalen, subkognitiven Natur, dem Bewusstsein nicht vollständig zugänglich (Scheffer et al., 2007). Sie können daher nicht, wie explizite Motive, direkt erfragt werden (McClelland, 1958, Scheffer et al., 2007, Schultheiss, 2008), da explizite Repräsentationen diese Motive reflektieren, reflektieren können oder auch nicht (Scheffer et al., 2007).

Implizite Motive bestimmen, welcher Stimulus als anziehend/aversiv interpretiert wird, orientieren das Verhalten einer Person in dessen Richtung (oder davon weg), geben Antrieb für ein Verhalten diesen Stimulus aufzusuchen (oder zu meiden) und selektieren Stimuli und Verhalten, die für die Erreichung (oder Vermeidung) instrumentell sind (McClelland, 1987). Sie orientieren, selektieren und energetisieren somit spontane Formen des Verhaltens (McClelland, 1989). Individuen unterscheiden sich nun durch ihre Präferenzen, verschiedene

Arten von Belohnungen aufzusuchen. Diese Differenzen spiegeln sich in den verschieden starken Ausprägungen der einzelnen Motive wieder (McClelland, 1989).

2.2.2 Erhebungsform

Implizite Motive reagieren eher auf nonverbale Schlüsselreize als auf verbale (Schultheiss, 2001, 2008). Wie bereits durch ihr Operieren im nicht, dem Bewusstsein, dem Selbstkonzept bzw. der verbalen Repräsentation einer Person, zugänglichen Bereich angedeutet (Brunstein et al. 1998; McClelland et al., 1989; Schultheiss, 2010, Schultheiss et al., 2008), können diese Motive nicht mit deklarativen Verfahren, also anhand von direktem Erfragen, erfasst werden (McClelland, 1958, Scheffer et al., 2007; Schultheiss, 2008; Schultheiss et al., 2007). Der Fokus liegt hier also auf der Messung und Beobachtung tatsächlicher behavioraler Phänomene, anstatt auf Messungen auf der Basis von Selbstreporten zur eigenen Persönlichkeit und Verhalten (Schultheiss, 2008).

Für die Erhebung von impliziten Motiven werden operante, nicht deklarative Verfahren verwendet (Murray, 1938; Scheffer et al., 2007; Schultheiss, 2008; Schultheiss et al., 2007). Ein Grund dafür ist die Annahme, dass implizite Motive operantes Verhalten generieren (Scheffer et al., 2007). Sie reagieren somit spontan auf Anreize, die mit einem angenehmen Affekt assoziiert sind (McClelland, 1989). Ein weiterer Grund liegt in der postulierten Auslösung eher durch nonverbale denn durch verbale Stimuli (Schultheiss, 2001, Schultheiss et al., 2007) und der Annahme, dass sie nach einer Aktivierung sehr wahrscheinlich einen Einfluss auf nicht deklarative Messungen von Motiven haben (Schultheiss, 2008).

Der Thematische Apperzeptionstest (TAT) von Murray (1943), ein Vertreter der Kategorie „Picture Story Exercise“ (PSE; McClelland et al., 1989), stellt hierbei das bekannteste operanten, nicht deklartiven Verfahren dar (Scheffer et al., 2007).

Die zu untersuchenden Personen müssen beim TAT zu einem Bild eine Geschichte verfassen, von der angenommen wird, dass sie nicht durch den Stimulus selbst, sondern durch interne Prozesse bedingt wird, wie sie beim Prozess des Schreibens von Geschichten

zu uneindeutigen Bildern entstehen (Atkinson & McClelland, 1948). Durch die Reduzierung auf verbale Assoziationen soll der Einfluss bzw. die Verzerrung des logischen Denkens und der Rationalität gering gehalten werden (Scheffer et al., 2007). Des Weiteren reduziert sich der Aufwand für die Administration einer solchen Erhebung im Vergleich zu einer Verhaltensbeobachtung signifikant (Scheffer et al., 2007). Die verfassten Geschichten werden dann anhand von empirisch fundierten und validierten inhaltlichen Kodiersystemen ausgewertet (Schultheiss et al., 2007).

2.2.3 Korrelate

Es konnte gezeigt werden, dass implizite Motive operantes Verhalten vorhersagen können (Scheffer et al., 2007). Operantes Verhalten wird von McClelland (1980) als selbst initiiertes, spontanes, ohne größere Überlegungen durchgeführtes und über einen längeren Zeitraum wiederholt gezeigtes Verhalten definiert. Implizite Motive sind des Weiteren ein wichtiger Prädiktor für das emotionale Wohlbefinden, Gesundheit und persönlichen Wachstum (Brunstein et al., 1998; Sheldon & Kasser, 1995).

Das implizite Macht Motiv konnte im Gegensatz zum expliziten Bedürfnis nach Dominanz, die Leistung bei einem Computerspiel vorhersagen, bei dem die Möglichkeit bestand, Spielstände anderer Personen zu streichen. Die zuvor von den Personen getätigten Aussagen über ihren Einsatz, konnten die tatsächliche Leistung nicht voraussagen (Schultheiss & Brunstein, 1999). Brunstein und Hoyer (2002) konnten zeigen, dass das implizite Leistungsmotiv, die Leistung von Aufmerksamkeitsaufgaben vorhersagen kann, im Speziellen nach einem negativen, auf die Personen zugeschnittenen, Feedback. Implizite Motive sagen damit eher die Leistung voraus als Wahlen und Urteile (Brunstein et al., 2002; Schultheiss, 2008).

2.2.4 Kritik

Das Groß der Kritik an impliziten Verfahren baut auf deren zur Erhebung verwendeten projektiven Verfahren auf und deren mangelnde psychometrische Qualität (Amelang et al., 2006; Kubinger, 2009b; Rollet, 2006).

Im Speziellen zeigt sich, dass die Reliabilität im Sinne einer internen Konsistenz nur sehr geringe Koeffizienten von .20 bis .50 ergibt (Schultheiss & Pang, 2007). Auch wenn diese Reliabilitäten für Forschungszwecke als ausreichend bezeichnet werden (Schultheiss, 2008), so werden diese Werte im Allgemeinen als zu gering angesehen (Kubinger, 2009b).

Der Kritik der mangelnden Reliabilität wird entgegen gehalten, dass die zur Messung verwendeten Maße der internen Konsistenzen (z.B. Cronbachs Alpha), für die Erhebung von motivationalen Prozessen nicht geeignet sind, weil deren Natur es ist, eben nicht konstant zu sein. Die Motive werden durch die Bildgeschichten nicht nur aktiviert, sondern auch gleich befriedigt, sodass ihre Ausprägung bei einer zweiten Messung geringer ausfällt und es dem zu Folge zu diesen geringen Koeffizientwerten kommt (Atkinson, 1981; Schultheiss & Pang, 2007).

Ein weiterer Erklärungsversuch liegt in der Verschiedenartigkeit der Stimuli, welche zur Folge haben sollte, dass die Validität der Messungen hoch ist (Schultheiss et al., 2007), wovon auch ausgegangen wird (Kuhl et al., 2003).

Eine Meta Analyse von Schultheiss et al. (2007) konnte zeigen, dass die Reliabilität im Sinne einer Stabilität hoch ausfällt: nach einem Tag war der durchschnittliche Stabilitätskoeffizient .71, .60 nach einer Woche, .52 nach einem Monat, .37 nach einem Jahr, .29 nach 5 Jahren und .25 nach 10 Jahren.

Kritik zur empirischen Evidenz der Genese der Motive kommt von Kuhl et al. (2003) die meint, dass diese als „insgesamt zu wenig umfangreich, wenig detailliert und theoretisch nicht zufriedenstellend fundiert“ bezeichnet werden muss.

2.3 Zur nicht vorhandenen Korrelation

Implizite und explizite Motive korrelieren nicht signifikant miteinander (Brunstein, 2006; Scheffer et al., 2007; Schultheiss, 2008; Schultheiss et al., 2001). Als ein Grund dafür gilt, dass es sich dabei um zwei verschiedene Motivsysteme handelt, die eine adaptive Funktion übernehmen (Brunstein, 2006; McClelland et al., 1989). Des Weiteren werden für die

Erhebung verschiedene Stimuli - Materialien verwendet (Brunstein, 2006; Schultheiss, 2008), daher wird ein Rückschluss auf eine Gleich- oder Verschiedenartigkeit zweier Konstrukte allein auf Basis der Korrelation zweier Tests kritisiert (Brunstein, 2006).

Brunstein (2006) weist auf die entscheidende Bedeutung der Frage, ob von den beiden Motivsystemen verschiedene Verhaltensweisen vorhergesagt werden, hin. Tatsächlich wirken die beiden Motivsysteme auf unterschiedliche Arten des Verhaltens: die expliziten Motive wirken primär auf kurzfristige, mit dem Selbstbild einer Person konforme Bewertungen und Entscheidungen, wohingegen implizite Motive primär spontanes Handeln und zeitlich überdauernde Verhaltenstrends vorhersagen (Brunstein, 2006; McClelland, 1980; Schultheiss et al., 2010). Nach der PSI-Theorie von Kuhl (2000, 2001) sind die impliziten Motive und expliziten Zielorientierungen mit zwei verschiedenen kognitiven Systemen verbunden. Schultheiss (2008) gibt eine gute Übersicht über die Moderatoren der Beziehung zwischen impliziten und expliziten Motiven. Implizite Motive reagieren eher auf nonverbale Schlüsselreize, explizite Motive reagieren eher auf verbale (Schultheiss, 2001, 2008). Auch kommt die Aktivierung eines expliziten oder impliziten Motivs auf das zuvor gegebene Feedback an: beim Fokus auf individuelle Ziele, werden implizite Motive angesprochen, beim Fokus auf Ziele im sozialen Vergleich, werden explizite Motive angesprochen. Personen mit einem hohen impliziten Leistungsmotiv messen sich tendenziell nicht an einer sozialen Norm, sondern an ihren eigenen bisherigen Leistungen (Brunstein et al., 2002; Schultheiss, 2010). Das explizite Leistungsmotiv ist demnach mit sozialen Vergleichen verbunden, wohingegen das implizite Leistungsmotiv mit aufgaben- und tätigkeitsbezogenen Anreizen verbunden ist (Brunstein, 2006).

Auch der unterschiedliche Entwicklungszeitpunkt wird als Grund für die nicht vorhandene Korrelation gesehen (Brunstein, 2006). So stehen den früh gelernten und daher nicht sprachlich repräsentierten impliziten Motiven, die im Kontext der Entwicklung und damit sprachlich repräsentierten, mit dem Selbstkonzept verbundenen expliziten Motive gegenüber (Brunstein, 2006; McClelland, 1987, 1989).

3. Hinführung zu den Hypothesen

Wie in den vorigen Kapiteln ausgeführt, zeigt sich die Notwendigkeit einer Überprüfung der Äquivalenz zwischen den Vorgabemodi „Papier und Bleistift“ und „PC“ (APA, 1986; de Beuckelaer et al., 2009; Ford et al. 1996; Hofer, 1985; Kindsvater et al., 2003; Kolen et al., 2004; Kubinger, 1993, 2009a; Kveton et al., 2004; Mead, 1993; Pommerich, 2004; Russell et al., 2003; Wagner-Menghin, 2003a). Diese muss laut Wagner-Menghin (2003a) und Ramstaedt et al. (2004) sowohl die psychometrische als auch die erfahrungsbezogene Ebene umfassen. Die psychometrische Äquivalenz wird mittels Vergleich der Rohwertmittelwerte, Varianzen, Rangreihungen der Mittelwerte und Reliabilitäten überprüft (Apa, 1986; Bartmann, 1994; Kubinger, 2009b; Kveton et al., 2007; Wagner-Menghin, 2003a), die erfahrungsbezogene Äquivalenz anhand der Befragung der subjektiven Ähnlichkeit (Wagner-Menghin, 2003a). Das von Bartram (1994) vorgeschlagene Design zur Prüfung der Äquivalenz wurde versucht in dieser Studie zu realisieren. Da diese Studie jedoch auf jener von Hofer (2009) aufbaut, konnte die Kreuzkondition „PC-PBT“ nicht realisiert werden, denn in der Studie von Hofer (2009) wurden alle zu untersuchenden Personen mittels einer PBT getestet.

Die Äquivalenz soll in dieser Studie für die scheinbar voneinander unabhängigen, weil nicht miteinander signifikant korrelierenden (Brunstein, 2006; Scheffer et al., 2007; Schultheiss, 2008; Schultheiss et al., 2001), impliziten und expliziten Motivmaßen überprüft werden.

III. Empirischer Teil

4. Ziele der Studie

Ziel dieser Studie ist es die Äquivalenz zwischen PBT und PC Testungen bei der Messung von impliziten und expliziten Motivmaßen zu überprüfen. Als Verfahren zur Erhebung

expliziter Motivmaße wurde der von Hofer (2009) aktualisierte PRF-D von Stumpf et al. (1985) verwendet (siehe Kapitel „4.2.2 PRF-D“). Für die Erhebung impliziter Motivmaße wurde der ebenfalls von Hofer (2009) adaptierte OMT von Kuhl et al. (1999) verwendet (siehe Kapitel „4.2.1 OMT“).

Nach Ramstaedt et al. (2004) und Wagner-Menghin (2003a) wurden Hypothesen sowohl zur psychometrischen als auch zur erfahrungsbezogenen Äquivalenz erstellt.

Für die Überprüfung der psychometrischen Äquivalenz wurden folgende Hypothesen gebildet (Apa, 1986; Bartmann, 1994; Kubinger, 2009b; Kveton et al., 2007; Wagner-Menghin, 2003a):

4.1 Hypothese 1

Die erste Hypothese geht der Frage nach, ob die deskriptiven Statistiken für beide Vorgabemodi äquivalent sind. Dies führt zu folgenden Subhypothesen, welche für jede Skala einzeln überprüft werden müssen:

(a) H_0 : Die Rohwertmittelwerte beider Gruppen (PBT, PC) sind äquivalent.

$$\bar{x}_{PBT} = \bar{x}_{PC}$$

H_1 : Die Rohwertmittelwerte beider Gruppen (PBT, PC) sind nicht äquivalent.

$$\bar{x}_{PBT} \neq \bar{x}_{PC}$$

(b) H_0 : Die Varianzen der Rohwerte beider Gruppen (PBT, PC) sind äquivalent.

$$s^2_{PBT} = s^2_{PC}$$

H_1 : Die Varianzen der Rohwerte beider Gruppen (PBT, PC) sind nicht äquivalent.

$$s^2_{PBT} \neq s^2_{PC}$$

(c) H_0 : Die Verteilungen der Rohwerte beider Gruppen (PBT, PC) sind äquivalent.

H_1 : Die Verteilungen der Rohwerte der beiden Gruppen (PBT, PC) sind nicht äquivalent.

(d) H_0 : Die Rangfolgen der Rohwerte der beiden Gruppen (PBT, PC) sind äquivalent.

H_1 : Die Rangfolgen der Rohwerte der beiden Gruppen (PBT, PC) sind nicht äquivalent.

3.2 Hypothese 2

Mittels der zweiten Hypothese wird überprüft, ob die Reliabilität der beiden Gruppen (PBT, PC) äquivalent ist. Dafür wurden die folgenden Subhypothesen erstellt, die wiederum für jede Skala einzeln geprüft werden müssen:

(a) H_0 : Die Split Half Reliabilität der beiden Gruppen (PBT, PC) ist äquivalent.

$$r_{t1/2 - PBT} = r_{t1/2 - PC}$$

H_1 : Die Split Half Reliabilität der beiden Gruppen (PBT, PC) ist nicht äquivalent.

$$r_{t1/2 - PBT} \neq r_{t1/2 - PC}$$

(b) H_0 : Die Retest Reliabilität der beiden Gruppen (PBT, PC) ist äquivalent.

$$r_{tt - PBT} = r_{tt - PC}$$

H_1 : Die Retest Reliabilität der beiden Gruppen (PBT, PC) ist nicht äquivalent.

$$r_{tt - PBT} \neq r_{tt - PC}$$

3.3 Hypothese 3

Die dritte Hypothese beschäftigt sich mit der erfahrungsbezogenen Äquivalenz nach Rammstaedt et al. (2004) und Wagner-Menghin (2003a). Dies führte zu folgender Formulierung:

H_0 : Die beiden Vorgabemodi (PBT, PC) werden als äquivalent bewertet.

H_1 : Die beiden Vorgabemodi (PBT, PC) werden als nicht äquivalent bewertet.

4. Methoden

4.1 Untersuchungsplan

Um die Äquivalenz der beiden Vorgabemodi überprüfen zu können, wurde ein experimentelles Retest Design mit randomisierter Zuweisung verwendet. Die erste Datenerhebung wurde von Hofer (2009) im Sommersemester 2008, also zwischen März und April 2008 durchgeführt. Sie war Teil des Projekt „PeRaS“ des Instituts für Entwicklungspsychologie und Psychologische Diagnostik der Universität Wien. Dieses Projekt versuchte die Effekte des persönlichen Umfeldes von Studierenden im zweiten Abschnitt des Diplomstudiums Psychologie auf den Studienfortschritt und Studienerfolg zu erheben (siehe Hofer, 2009). Tabelle 1 zeigt den Aufbau der Ersten, sowie der zweiten Datenerhebung.

Tabelle 1: Aufbau der ersten und zweiten Datenerhebung

	KG		VG	
	OMT	PRF-D	OMT	PRF-D
Erste Datenerhebung	PBT	PBT	PBT	PBT
Zweite Datenerhebung	PBT	PBT	PC	PC

Die zweite Datenerhebung wurde für das Wintersemester 2008/09 anberaumt, 8 Monate nach der ersten. Wie aus Tabelle 1 ersichtlich, sind bei dieser Erhebung zwei Gruppen geplant: (1) die Kontrollgruppe (KG), welche beide Verfahren noch einmal am Papier bearbeiteten und (2) die Versuchsgruppe (VG), welche beide Verfahren am PC bearbeitete. Die Zuteilung zu den Gruppen soll wiederum zufällig erfolgen.

Im Unterschied der ersten Datenerhebung wird hier außer dem OMT von Kuhl et al. (1999) und dem PRF-D von Stumpf et al. (1985) kein anderes psychologisch-diagnostisches Verfahren vorgegeben. Auch diese Datenerhebung soll im Rahmen eines Seminars, dieses Mal im Rahmen der Lehrveranstaltung „Übungen zur psychologischen Diagnostik II“

durchgeführt werden. Eine Teilnahme ist aufgrund der Anwesenheitspflicht in dieser Lehrveranstaltung verpflichtend.

4.2 Verfahren

Zur Erhebung der impliziten und expliziten Motivmaße werden zwei verschiedene psychologisch-diagnostische Verfahren verwendet: (1) der Operante Multi-Motiv Test (OMT) von Kuhl et al. (1999) für die Erhebung der impliziten Motivmaße und (2) die deutsche Version des „Personality Research Form“ (PRF-D) von Stumpf et al. (1985) für die Erhebung der expliziten Motivmaße.

4.2.1 OMT

Der OMT (Kuhl et al., 1999; Scheffer, 2005) stellt ein psychologisch-diagnostisches Verfahren zur Messung der impliziten Motive Bindung („n Bindung“), Leistung („n Leistung“) und Macht („n Macht“) dar (Kuhl et al., 1999; Kuhl, Scheffer & Eichstaedt, 2003; Scheffer et al. 2007). Der Präfix „n“ steht hierbei für das englische Wort „need“, also „Bedürfnis“ und wurde sensu Brunstein (2006) gewählt, um eine Verwechslung mit den teilweise gleich benannten expliziten Motiven zu verhindern, die nicht mit den impliziten Motiven korrelieren (Brunstein, 2006; Scheffer et al., 2007; Schultheiss, 2008; Schultheiss et al., 2001) und in der Literatur deswegen von zwei verschiedenen motivationalen Systemen ausgegangen wird (Brunstein, Schultheiss & Grässmann, 1998; McClelland et al. 1989, Schultheiss & Brunstein 2005).

Die zu untersuchende Person muss beim OMT zu einer Serie von mindestens fünfzehn uneindeutigen Bildern eine Anzahl von Fragen so spontan wie möglich beantworten (Kuhl & Scheffel, 2001). Dadurch ist der OMT ein psychologisch-diagnostisches Verfahren der Kategorie „Picture Story Exercises“ (PSE, McClelland et al., 1989) und damit ein projektives, als auch „operant inhaltsanalytisch“ bezeichnetes Verfahren (Scheffer et al., 2003).

Die drei Skalen „n Bindung“, „n Leistung“ und „n Macht“ werden im OMT nach dem Auswertungsmanual wie folgt definiert (Kuhl et al., 1999):

- *n Bindung*: es findet eine soziale Interaktion statt, die horizontal verläuft, d.h. es ist keine Hierarchie vorhanden. Die Inhalte der Antworten auf die Fragen kreisen um die „Herstellung, Aufrechterhaltung, Wiedergewinnung oder das Fehlen von Kontakt zu einer oder mehreren anderen Personen“ (S. 18)
- *n Leistung*: „ein Objekt wird in Richtung eines Gütezustandes verändert oder in Bezug zu einem bestimmten Kriterium verbessert, etwas zu lernen, oder einer Anforderung zu genügen“ (S. 32)
- *n Macht*: eine Person übt Einfluss auf andere aus, eine vertikale Hierarchie wird dabei immer implizit angesprochen

Da die erste diese Datenerhebung von Hofer (2009) durchgeführt wurde und darin eine im Sinne von Kuhl et al. (2001) verkürzte Version des OMT nach Kuhl et al. (1999) verwendet wurde, wird auch für diese Studie diese in Kuhl et al. (2001) erwähnte und von Hofer (2009) verwendete Version des OMT vorgegeben. Die zu untersuchende Person muss daher nicht wie im OMT nach Kuhl et al. (1999) vier, sondern drei Fragen beantworten (Kuhl et al., 2001).

Nach dem Mehrebenen Modell wird die Antwort auf die erste der drei Fragen („Was ist für die Hauptperson in dieser Situation besonders wichtig?“) als dominierendes Motiv des Probanden mit Hilfe des Auswertungsmanuals kodiert (Kuhl et al., 1999; Scheffer et al., 2003). Die Motivkodierung ist ähnlich der von Winter (1994) für den TAT entwickelten. Die Regulationsmechanismen der durch die erste Antwort definierten dominierenden Motive werden durch die Antworten der zweiten („Wie fühlt sich die Hauptperson dabei?“) und dritten Frage („Warum fühlt sie sich so?“) kodiert (Scheffer et al., 2003). Diese Kodierung der Regulationsmechanismen erhöht unter anderem die Reliabilität, da auch die individuellen Differenzen in der Art wie die Bedürfnisse in das Verhalten einfließen in Betracht gezogen werden (Scheffer et al., 2007).

Bevor die zu untersuchende Person diese Fragen beantwortet, soll sie einen der auf den Bildern abgebildeten Personen als Protagonisten auswählen. Durch die Identifikation der zu untersuchenden Person mit dem Protagonisten sollen persönliche Bedürfnisse und Traits

(d.h. Wesenszüge) einfließen (Scheffer et al., 2007). Im Unterschied zu anderen operant inhaltsanalytischen Verfahren wie z.B. dem TAT (Murray, 1943), sind die präsentierten Bilder nicht grundsätzlich als motivational offener Stimulus konzipiert, jedes Bild steht mit einem speziellen Motiv (wenn auch implizit) in Verbindung (Kuhl et al., 2003; Scheffer et al., 2003).

Ein weiterer Unterschied zu anderen operant inhaltsanalytischen Verfahren ist, dass beim OMT die zu untersuchende Person nicht eine ganze Geschichte niederschreiben muss. Vielmehr soll sie nur die Fragen mit spontanen Assoziationen, also kurze Sätze bzw. Stichwörter, beantworten (Kuhl et al., 1999; Scheffer, 2005). Diese Reduktion soll den Einfluss der Rationalität bzw. des logisch schlussfolgernden Denkens so gering wie möglich halten (Kuhl et al. 2001; Scheffer et al., 2007). Für eine detaillierte Auswertungsanleitung sei auf das entsprechende Manual von Kuhl et al. (1999) hingewiesen.

Gütekriterien

Obwohl aufgrund der in Hofer (2009) durchgeführten Studie eine veränderte Version des OMT verwendet wurde und daher die Gütekriterien des OMT nach Kuhl et al. (1999) für die Version des auch in dieser Studie verwendeten OMT nicht zur Gänze zutreffen, sei hier ein Überblick über die Gütekriterien gegeben, zumal auch schon Kuhl et al. (2001) auf die Möglichkeit einer Verkürzung des Original OMT hinwiesen.

Objektivität

Wie bei jedem operant inhaltsanalytischen Verfahren, werden auch hier die Antworten mittels eines empirisch fundierten und validierten inhaltlichen Kodiersystems ausgewertet (Kuhl et al., 2003; Scheffer et al., 2003; Schultheiss et al., 2007). Um die Objektivität der Messungen zu erhöhen, wird diese Kodierung von mehreren unabhängigen Ratern vorgenommen und danach eine Interrater - Reliabilität berechnet (Scheffer et al., 2003). Diese Interrater -Reliabilität wird beim OMT mit der Formel von Winter (1994) berechnet und liegt nach einigen Tagen Üben bei .85, nach gemeinsamer Auswertungszeit lassen sich annähernd 100% Übereinstimmung erzielen (Scheffer et al., 2003).

Reliabilität

Aufgrund der Mehrdeutigkeit des Stimulusmaterials (ein Item kann prinzipiell jedes Motiv messen) und der Messung an sich (z.B. Befriedigung der Bedürfnisse durch Bearbeitung eines vorhergehenden Items) ergeben sich Probleme bei der Berechnung der Reliabilität im Sinne einer internen Konsistenz (Atkinson, 1981; Schultheiss & Pang, 2007). Rollet (2006) postuliert, dass eine Berechnung aufgrund der nicht vorhandenen Homogenität des Testmaterials nicht möglich sei. Durch theoretische Überlegungen kommen Scheffer et al. (2003) jedoch zum Schluss, dass eine Berechnung der Reliabilität bei den Extremgruppen möglich ist. Die durchschnittlichen Werte des Cronbach-Alpha Koeffizienten (Scheffer et al., 2003) werden in Tabelle 2 dargestellt und gelten nach Field (2005) als ausreichend. Die Reliabilität im Sinne einer Stabilität wird bei den Grundmotiven des OMT mit .72 nach einer Woche angegeben (Scheffer et al., 2003).

Tabelle 2: Durchschnittliche Cronbach-Alpha Koeffizienten des OMT nach Scheffer et al. (2003)

	n Bindung	n Leistung	n Macht
Cronbachs Alpha	.74	.70	.78

Validität

Die diskriminative Validität scheint aufgrund der nicht vorhandenen Korrelation zu explizit erhobenen Motivmaßen gegeben. Bezüglich der konvergenten Validität wurden zwar signifikante Korrelationen zum TAT gefunden, der verwendete Bildsatz zeigte jedoch einen großen Einfluss auf diese Messung (Scheffer et al., 2003).

Zur prognostischen Validität konnten Scheffer et al. (2007) für das Motiv „n Bindung“ einen positiven Befund vorlegen: das Motiv sagte positive Bewertungen von Lehrlingen im Service Bereich durch Fachleute voraus. Scheffer et al. (2003) konnten zeigen, dass das Motiv „n Bindung“ bei Männern und Frauen als ein Faktor für eine frühe sexuelle Partnerbindung gilt. Das Motiv „n Macht“ korrelierte signifikant ($r=.32$, $p<.05$, $df=56$) mit einem durch ein Assessment-Center erstellten Summenscore der Werte Argumentationsfähigkeit,

Durchsetzungsstärke, Gruppensteuerung und Konfliktfähigkeit (Scheffer et al., 2003). Für das Motiv „n Leistung“ fanden Scheffer et al. (2003) eine signifikante Korrelation ($r=-.26$, $p<.05$, $df=98$) mit den Studienzwiischennoten.

Als Anwendungsbereiche werden vor allem die Forschung zu impliziten Motiven und die Aufdeckung der Diskrepanz zwischen expliziten und impliziten Motiven/Zielen angegeben (Kuhl et al., 1999).

4.2.2 PRF-D

Der Personality Research Form (PRF) ist ein Fragebogen von Jackson (1974), der versucht die Klassifikation der psychogenen Bedürfnisse nach Murray (1938) zu erfassen. Dieses Verfahren wurde von Stumpf, Angleiter, Wieck, Jackson & Beloch-Till (1985) ins Deutsche übersetzt und dient zur Erfassung expliziter Motivmaße (Brunstein, 2006; Schultheiss et al., 2007; Stumpf et al., 1985). Primär wurde der PRF beziehungsweise der PRF-D für die Forschung im nicht psychopathologischen Bereich entwickelt, er wird jedoch auch in der psychologischen Praxis und für die Bildungs- und Berufsberatung eingesetzt (Stumpf et al., 1985).

Der PRF-D besteht aus 234 dichotom („richtig/falsch“) zu beantwortenden Items, welche zu 15 Skalen zusammengefasst werden. Jedes Item lädt dabei nur auf eine der 15 Skalen, wovon 14 so genannte „Inhaltsskalen“ darstellen, also direkt auf die Messung von Persönlichkeitsmerkmalen abzielen und eine Skala, die „Validitätsskala“, eine unsachgemäße Bearbeitung erfassen soll (Stumpf et al., 1985).

Die einzelnen Skalen lauten: (1) Leistungsstreben, (2) Geselligkeit, (3) Aggressivität, (4) Dominanzstreben, (5) Ausdauer, (6) Bedürfnis nach Beachtung, (7) Risikomeidung, (8) Impulsivität, (9) Hilfsbereitschaft, (10) Ordnungsstreben, (11) Spielerische Grundhaltung, (12) Soziales Anerkennungsbedürfnis, (13) Anlehnungsbedürfnis und (14) allgemeine Interessiertheit.

Die Vorgabedauer dieses psychologisch-diagnostischen Verfahrens beträgt zwischen 20 und 50 Minuten, das Alter der zu untersuchenden Person sollte über 17 Jahren liegen (Stumpf et al., 1985).

Von den insgesamt 15 Skalen, die durch den PRF-D erfasst werden (Stumpf et al., 1985), wurden drei ausgewählt: (1) „Leistungsstreben“ (im englischen Original: achievement), (2) „Geselligkeit“ (affiliation) und (3) „Dominanz“ (dominance). Zum Einen weil diese Motive das explizite Gegenstück zu dem im OMT gemessenen impliziten Motiv darstellen (Brunstein, 2006; Schultheiss et al., 2001), zum Anderen weil auch bei der Studie von Hofer (2009) diese Auswahl getroffen wurde und diese Studie darauf aufbaut.

Die ausgewählten Skalen werden laut Stumpf et al. (1985) wie folgt beschrieben:

- (1) Leistungsstreben: Strebsamkeit, Beharrlichkeit, Fleiß, Erfolgsstreben, Ehrgeiz, Findigkeit, Produktivität
- (2) Geselligkeit: freundschaftliche Gefühle, Treue, Herzlichkeit, Gutmütigkeit, Freundlichkeit, Umgänglichkeit, Hilfsbereitschaft, Gastfreundlichkeit, Kontaktfreude, Wohlwollen
- (3) Dominanzstreben: bestimmende, kontrollierende, herrschende Haltung; Überzeugungskraft; Überlegenheit; Durchsetzungsfähigkeit; Macht; Gewichtigkeit

Für diese Studie wurde der von Hofer (2009) veränderte PRF-D verwendet. Folgende Änderungen wurden von Hofer (2009) vorgenommen: (1) Um die Akzeptanz seitens der Studierenden zu erhöhen, wurde das dichotome Antwortformat zu einem Multiple-Choice Antwortformat geändert. Die Antwortmöglichkeiten änderten sich also von „richtig“ bzw. „falsch“ auf: „stimmt“, „stimmt eher“, „stimmt eher nicht“ und „stimmt nicht“. (2) Der Text einiger Items musste aufgrund der nicht mehr zeitgemäßen Formulierung angepasst werden.

Gütekriterien:

Aufgrund der von Hofer (2009) durchgeführten Änderungen, können hier nicht alle Ergebnisse des PRF-D nach Stumpf et al. (1985) übernommen werden.

Objektivität:

Die Objektivität kann aufgrund der Standardisierung der Instruktion (laut Testanweisung), der Verrechnung (mit Hilfe von Schablonen) und der (durch die eindeutige Zuordnung der Rohwerte zu Prozenträngen) standardisierten Interpretation als gegeben angenommen werden (Stumpf et al., 1985). Die Änderungen von Hofer (2009) geben keinen Anlass, eine Auswirkung auf die Originalergebnisse des PRF-D anzunehmen.

Reliabilität

Der PRF-D von Hofer (2009) erzielte nach dem Guttman Split-Half Koeffizienten eine Reliabilität zwischen .78 und .81. Das Cronbach Alpha für die Originalversion nach Stumpf et al (1985) lag zwischen .66 und .82. Für die in dieser Studie verwendeten Skalen sind es .70 für die Skala „Leistung“, .75 für die Skala „Geselligkeit“, und .82 für die Skala „Dominanz“.

Die Reliabilität im Sinne einer Stabilität muss aufgrund der Änderungen von Hofer (2009) neu berechnet werden. Beim Original lagen die Werte nach zwei Wochen zwischen .85 und .95, nach einem Jahr zwischen .67 und .85 und nach zwei Jahren zwischen .61 und .81 (Stumpf et al., 1985).

Validität

Die Rohübersetzung wurde anhand von Expertenurteilen nochmals abgesichert, die Inhaltsvalidität scheint daher gegeben (Stumpf et al., 1985). Auch die Faktorenstruktur konnte repliziert werden: alle Faktoren luden positiv (deren durchschnittliche Ladungen lagen zwischen .29 und .35) und waren größer als die Ladungen auf die anderen Skalenfaktoren (Stumpf et al., 1985). Die Faktorvalidität wird von Stumpf et al. (1985) ebenfalls als hoch bezeichnet. Auch gibt es eine deutliche Konvergenz sowohl auf Skalen- als auch im

Faktorenbereich zu verwandten Faktoren und Einzelindikationen anderer Persönlichkeitsinventare (Stumpf et al., 1985).

4.2.3 Fragebogen zur Erfassung der subjektiven Äquivalenz

Es wurde ein Fragebogen zur Erfassung der subjektive Äquivalenz sensu Wagner-Menghin (2003a) zwischen der PBT und der PC-Testung erstellt. Dieser Fragebogen (siehe Abbildung 1) wurde nur der PC Gruppe vorgegeben. Er beinhaltete eine 7-stufige Rating-Skala, anhand derer die subjektive Gleichwertigkeit bewertet werden konnte. Im Falle einer Bewertung als nicht gleichwertig, wurde mittels eines leeren Feldes unter der Rating-Skala um eine Begründung gebeten. Damit sollten bei nicht vorhandener Äquivalenz die Gründe für eine qualitative Auswertung erhoben werden.

Abbildung 1: Fragebogen zur Erfassung der subjektiven Äquivalenz

Ist die PC Testung und die Papier/Bleistift Testung Deiner Meinung nach gleichwertig (macht es also keinen Unterschied, ob man den Fragebogen am PC oder per PP ausfüllt)?

OMT

nicht gleichwertig

1	2	3	4	5	6	7
---	---	---	---	---	---	---

 gleichwertig

Begründung (wenn NICHT gleichwertig): _____

PRF

nicht gleichwertig

1	2	3	4	5	6	7
---	---	---	---	---	---	---

 gleichwertig

Begründung (wenn NICHT gleichwertig): _____

4.3 Transformation der Verfahren

Die beiden Verfahren (PRF-D, OMT), wurden in dieser Studie gemäß des Untersuchungsplans (siehe Tabelle 1) sowohl in Papier und Bleistift Form wie bei Hofer (2009), als auch am PC vorgegeben. Im Folgenden wird die Transformation der beiden Verfahren von Papier und Bleistift auf einen PC beschrieben.

4.3.1 OMT

Der OMT wurde mit Hilfe der Programmiersprache PHP Version 5.0 von einem Papier Bleistift Test in einen PC Test übersetzt. Die Programmiersprache wurde gewählt, da sie stabil (das Programm wird auf einem Server ausgeführt), plattformunabhängig (benötigter Web-Browser ist auf jedem PC-System vorhanden) und der Vorgabemodus ein vertrauter ist (aufgrund der Bearbeitung des Verfahrens in einem Web-Browser).

Die elektronische Version des OMT wurde so entworfen, dass die zu untersuchende Person immer nur ein Bild sehen konnte (siehe Abbildung 2).

Abbildung 2: Beispiel Item des OMT am PC

0% 100%

1)

Was ist für die Person in dieser Situation wichtig und was tut sie?

Wie fühlt sich die Person?

Warum fühlt sich die Person so?

Zurück Weiter

Unter dem Bild konnte die zu untersuchende Person die Hauptperson auswählen, rechter Hand wurden die drei Fragen präsentiert, inklusive Antwortfeld mit unbegrenzt langer Texteingabemöglichkeit. Auch konnte sie über die Schaltflächen „Zurück“ und „Weiter“ zwischen den verschiedenen Bildern hin und her blättern. Oberhalb der Bilder war ein Balken zu sehen, der den Fortschritt der Testung anzeigte (siehe Abbildung 2).

Diese Implementierung versuchte der PBT Version so ähnlich wie möglich zu sein: In die Antwortfelder konnte eine beliebig lange Eingabe vorgenommen werden, da es auch bei PBT Version keine Beschränkung in dieser Hinsicht gibt. Die Möglichkeit des hin und her

Blätterns zwischen den Aufgaben ist bei jedem PBT gegeben und wurde dementsprechend auch hier durch die zwei Schaltflächen realisiert. Der Fortschritt kann durch die noch nicht ausgefüllten Blätter eines PBTs festgestellt werden, daher wurde hier ein Fortschritts-Balken hinzugefügt.

4.3.2 PRF-D

Für die Implementierung des PRF-D wurde das Programm LimeSurvey, Version 1.70+ (4629) verwendet, welches vom Testlabor des Arbeitsbereiches für Psychologische Diagnostik an der Universität Wien zur Verfügung gestellt wurde. Auch dieses Programm baut auf der Programmiersprache PHP auf, daher konnten auch alle im vorigen Abschnitt erwähnten Vorkehrungen, die zu einer möglichst hohen Äquivalenz der Vorgabe führen sollen, getroffen werden: Wie auch in der Papier und Bleistift Version des PRF-D, wählten die Studierenden ihre Antwort aus einer Anzahl von Antwortalternativen, welche unter dem zu beantwortenden Statement zu finden waren (siehe Abbildung 3). Dabei wählten sie also aus 5 verschiedenen Antwortalternativen, wobei zu Beginn immer die Alternative „Keine Antwort“ ausgewählt war, so wie bei einer noch nicht ausgefüllten Papier und Bleistift Version des PRF-D. Auch hier konnten die Studierenden, ähnlich der Papier und Bleistift Version, zwischen den Fragen blättern und wurden oberhalb des Items durch einen Balken über den Fortschritt informiert (siehe Abbildung 3).

Abbildung 3: Beispiel Item der PC Implementierung des PRF-D

Testung

0% 100%

PRF

1. Ich versuche so oft wie möglich, in der Gesellschaft von Freunden zu sein.

Bitte wählen Sie eine der folgenden Antworten

stimmt

stimmt eher

stimmt eher nicht

stimmt nicht

Keine Antwort

<< Zurück Weiter >>

4.4 Durchführung der Untersuchung

Die erste Datenerhebung wurde von Hofer (2009) durchgeführt. 201 Studierende bearbeiteten eine Serie von psychologisch-diagnostischen Verfahren. Die Bearbeitung dauerte knapp 2 Stunden und wurde im Rahmen des Seminars „Übungen zur Psychologischen Diagnostik I“ durchgeführt, wodurch die Studierenden zu einer Teilnahme verpflichtet waren. Zu Beginn wurden die Studierenden zufällig in zwei Gruppen von 10-15 Personen aufgeteilt. Die Erste Gruppe wurde gebeten in einen anderen Raum zu wechseln um dort die PC-Verfahren zu absolvieren, die zweite Gruppe blieb in dem Seminarraum und bearbeitete die Papier und Bleistift Tests. Nachdem die Gruppen fertig waren, wechselten sie die Räume, am Ende hatten alle Studierenden sowohl die Papier und Bleistift als auch die PC-Tests durchgeführt (siehe Tabelle 3). Die Studierenden konnten die gesamte Bearbeitung mit individueller Geschwindigkeit durchführen. Nachdem die Studierenden beide Testbatterien bearbeitet hatten, war die Bearbeitung und die Einheit des Seminars beendet.

Tabelle 3: Aufteilung der Gruppen in Hofer (2009)

	Gruppe 1	Gruppe 2
erste Testserie	PC	PBT
zweite Testserie	PBT	PC

Die zweite Datenerhebung wurde in den Monaten November und Dezember 2008 durchgeführt und fand somit 8 Monate nach der ersten Datenerhebung statt. Die Testungen fanden im Rahmen des Seminars „Übungen zur Psychologischen Diagnostik II“, eine verpflichtende Lehrveranstaltung im zweiten Abschnitt des Diplomstudiums Psychologie an der Universität Wien, statt. Sie dauerte ca. eine Stunde und wurde immer nach dem gleichen Schema durchgeführt: zu Beginn des Seminars wurden die Studierenden über den weiteren Ablauf der Einheit informiert, danach wurden sie in zwei Gruppen geteilt, in eine PBT und eine PC Gruppe. Die Randomisierung erfolgte mittels eines Körbchens in dem sich,

entsprechend der Anzahl der Studenten, mehrfach gefaltete Papierstücke befanden. Auf diesen stand entweder „PP“ (für die PBT Gruppe) oder „PC“ (für die PC Gruppe) und teilten somit die Studierenden gleich in die zwei Gruppen auf.

Wurde ein Studierender der PBT Gruppe zugewiesen, so blieb er im Seminarraum und füllte den von Hofer (2009) angepassten PRF-D und OMT aus. Wurde ein Studierender der PC Gruppe zugewiesen, so wurde er gebeten in den PC-Raum der Fakultät für Psychologie der Universität zu wechseln. Dort wurden diesen Studierenden die von Hofer (2009) angepassten Versionen des PRF-D und OMT am PC vorgegeben. Bis auf eine Gruppe wurden alle Studierenden in diesem PC-Raum getestet. Die angesprochene Gruppe musste aufgrund von Raumbelagungen im PC-Raum des Neuen Instituts Gebäudes (NIG) getestet werden. Auch mussten hier jene Studierenden, die der PBT Gruppe zugewiesen wurden den Raum wechseln, die Studierenden der PC Gruppe blieben im besagten PC-Raum.

Unabhängig von der Gruppe (PBT bzw. PC) bearbeiteten die Studierenden zuerst den OMT und dann den PRF-D. In der PC Gruppe konnten die Studierenden den PRF-D nach Beendigung des OMT durch Klick auf einen Link selbst starten, in der Papier und Bleistift Gruppe wurde vom Testleiter der OMT Bogen eingesammelt und der PRF Bogen ausgegeben.

Nach Bearbeitung der beiden Verfahren, wurde den Studierenden der Fragebogen zur Äquivalenz der beiden Testverfahren vorgelegt. Nachdem die Studierenden auch diesen Fragebogen ausgefüllt hatten, war sowohl die Testung als auch diese Einheit des Seminars abgeschlossen.

4.4.1 Instruktionen

Die Instruktionen mussten für beide Verfahren (PRF-D und OMT) adaptiert werden, um der Retest Situation Rechnung tragen zu können: Die Studierenden hatten dieselben Verfahren schon vor 8 Monaten bearbeitet und konnten sich somit möglicherweise an diese erste Datenerhebung erinnern. Diese Information musste in den Instruktionen enthalten sein um etwaige Verwirrungen zu vermeiden. Es sollte erreicht werden, dass die Studierenden so

natürlich wie möglich an die Testung herangingen und nicht versuchten, die vorangegangene zu reproduzieren oder von ihr beeinflusst zu werden. Dieser Teil der Instruktion wurde fett gedruckt um die Aufmerksamkeit der Studierenden darauf zu lenken.

Für den OMT wurde die Standard-Instruktion nach Kuhl et al. (1999) um den folgenden Absatz aus Schultheiss et al. (2007) ergänzt:

„Du kannst dich vielleicht daran erinnern, die Bilder zuvor schon einmal gesehen zu haben. Wenn dem so ist, kannst du so darauf reagieren, wie du es das letzte Mal gemacht hast, oder anders – je nach dem, wie du dich jetzt fühlst. Mit anderen Worten: erzähl die Geschichte, an die dich die Bilder denken lassen, egal ob sie die Gleiche wie beim letzten Mal ist, oder eben nicht.“

Für den PRF-D wurde, ähnlich dem OMT, die Standard-Instruktion nach Stumpf et al. (1985) um folgenden Absatz erweitert:

„Du kannst dich vielleicht daran erinnern, diese Aussagen zuvor schon einmal gelesen zu haben. Wenn dem so ist, kannst du so darauf reagieren, wie du das das letzte Mal gemacht hast, oder anders – je nach dem, wie Du dich jetzt fühlst. Mit anderen Worten: beantworte die Aussagen so, wie sie JETZT auf Dich zutreffen, egal ob es die gleichen Antworten wie beim letzten mal sind, oder eben nicht.“

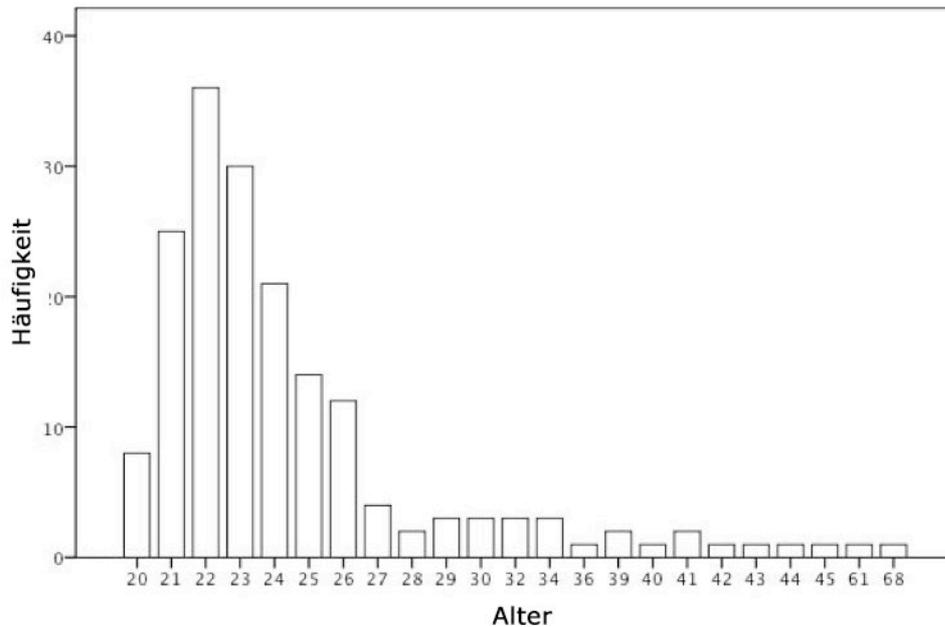
4.5 Stichprobe

Insgesamt wurden 177 Studierende getestet, 150 Frauen, 27 Männer. Nur ein Studierender brach die Testung nach der Bearbeitung des OMT am PC frühzeitig ab um einen Zug zu erwischen. Das Durchschnittsalter der Studierenden lag bei 25, der Median bei 23. Die Spannweite betrug 20 bis 68 Jahre, 82,4% der Studierenden lagen zwischen dem Alter von 20 und 26 Jahren (siehe Abbildung 4).

Wie schon erwähnt, wurden die Studierenden im Rahmen des Seminars „Übungen zur Psychologischen Diagnostik II“ getestet. Dieses Seminar stellt eine Pflichtlehrveranstaltung des zweiten Abschnitts des Diplomstudiums Psychologie an der Universität Wien dar.

Demnach waren auch alle getesteten Personen Studierende des Diplomstudiums Psychologie.

Abbildung 4: Altersverteilung der Stichprobe



5. Ergebnisse

Es folgen die Ergebnisse der statistischen Auswertung. Für alle Berechnungen wurde das Programm SPSS in der Version 16 verwendet.

5.1 Deskriptive Statistiken

Von den insgesamt 177 getesteten Studierenden waren 85% Frauen (150 Personen) und 15% Männer (27 Personen). Der Frauenanteil ergibt sich aus der hohen Frauenquote der Studierenden des Diplomstudiums Psychologie in Wien. Wie Tabelle 4 zeigt, war auch die Aufteilung der Geschlechter innerhalb der Gruppen (Papier und Bleistift bzw. PC) nicht ausgeglichen. Der Grund dafür liegt in der rein zufälligen Zuteilung, die klarerweise nicht auf das Geschlecht Rücksicht nahm. In der Versuchsgruppe, also der PC Gruppe,

bearbeiteten 90 Personen den OMT und 89 davon schlossen auch den PRF-D ab. Eine Person brach den Test frühzeitig ab.

Tabelle 4: Deskriptive Statistiken der Stichprobe

	Group	nFrauen	nMänner	N
OMT	CB	78 (90%)	9 (10%)	87
	P&P	72 (80%)	18 (20%)	90
PRF	CB	78 (90%)	9 (10%)	87
	P&P	72 (81%)	17 (19%)	89

Die Antworten des OMT wurden nach Scheffer et al. (2003) von drei Personen unabhängig voneinander kategorisiert. Anschließend wurde deren Übereinstimmung nach folgender Formel von Winter (1994) berechnet:

$$agreement = \frac{2 \times \left(\begin{array}{l} \text{Number of agreements between yourself and} \\ \text{expert on the presence of motive imagery} \end{array} \right)}{\left(\begin{array}{l} \text{Number of times you scored motive imagery} \\ \text{+} \\ \text{Number of times expert scored motive imagery} \end{array} \right)}$$

Tabelle 5 zeigt die damit berechneten Übereinstimmungen. Der Mittelwert der Übereinstimmungen liegt bei .72 und kann nach LePine (2003) und Nunally (1978) als akzeptabel interpretiert werden. Die nicht übereinstimmenden Kategorisierungen wurden mittels einer Gruppendiskussion der drei Rater angeglichen.

Tabelle 5: Übereinstimmungen der drei Bewerter des OMT

		PBT	MW	PC	MW
n Bindung	Bewerter 1 vs. Bewerter 2	.70	.74	.71	.74
	Bewerter 1 vs. Bewerter 3	.82		.70	
	Bewerter 2 vs. Bewerter 3	.71		.81	
n Leistung	Bewerter 1 vs. Bewerter 2	.57	.66	.79	.70
	Bewerter 1 vs. Bewerter 3	.61		.61	
	Bewerter 2 vs. Bewerter 3	.79		.70	
n Macht	Bewerter 1 vs. Bewerter 2	.72	.74	.75	.72
	Bewerter 1 vs. Bewerter 3	.77		.70	
	Bewerter 2 vs. Bewerter 3	.73		.72	

5.2 Korrelation impliziter und expliziter Motivmaße

Wie schon in vorangegangenen Studien (Brunstein, 2006; Scheffer et al., 2007; Schultheiss, 2008; Schultheiss et al., 2001) korrelieren auch hier die impliziten und expliziten Motivmaße nicht substantiell miteinander (siehe Tabelle 6). Einzig die Motive „n Leistung“ und „Leistung“ in der PBT Gruppe zeigen eine signifikante Korrelation ($r = -.252$, $p = .019$, $df = 86$). Dieser als gering zu bezeichnende Effekt ($r^2 = 6,4\%$) verschwindet jedoch, wenn man die gesamte Gruppe (also PBT und PC) betrachtet ($r = -.116$, $p = .125$, $df = 86$).

Tabelle 6: Korrelation der impliziten und expliziten Motivmaße

	Korrelation der Motive	r	p
PBT	n Bindung x Geselligkeit	.094	.387
	n Leistung x Leistung	-.252	.019
	n Macht x Dominanz	.135	.214
PC	n Bindung x Geselligkeit	.002	.988
	n Leistung x Leistung	-.024	.823
	n Macht x Dominanz	-.112	.295
PBT & PC	n Bindung x Geselligkeit	.048	.524
	n Leistung x Leistung	-.116	.125
	n Macht x Dominanz	.006	.933

5.3 Hypothese 1

Sub-Hypothese (a): Um zu überprüfen, ob sich die Mittelwerte der Rohwerte signifikant voneinander unterscheiden, wurde ein Welch-Test berechnet. Wie in Tabelle 7 ersichtlich ist, wurden keine Unterschiede bei den Rohwerten des PRF-D gefunden. Beim OMT zeigt nur die Skala „n Leistung“ einen signifikanten Unterschied ($t=4.21$, $p < .001$, $df=175$) mit einer Effektstärke von .63. Damit handelt es sich um einen bedeutsamen Unterschied in den Mittelwerten ($MW_{PBT}=3.1$, $MW_{PC}=2.41$). Da für diese Skala die H_0 verworfen und somit die H_1 angenommen werden muss, ist auch für den gesamten OMT die H_0 zu verwerfen und die H_1 anzunehmen. Für den PRF-D kann die H_0 beibehalten werden, da keine der Skalen einen signifikanten Unterschied bei den Rohwert-Mittelwerten zeigt.

Tabelle 7: Welch Test zu Überprüfung der Differenzen in den Rohwert-Mittelwerten der PBT und PC Gruppe für den PRF-D und den OMT

		Group	N	MW	SD	t	df	p
PRF	Geselligkeit	PBT	87	11.79	3.30	-.368	171	.714
		PC	89	11.97	2.94			
	Leistung	PBT	87	10.71	3.26	.647	174	.518
		CB	89	10.39	3.28			
	Dominanz	PBT	87	8.09	4.09	-.341	174	.734
		CB	89	8.3	4.14			
OMT	n Bindung	PBT	87	3.25	1.13	.495	175	.621
		PC	90	3.17	1.18			
	n Leistung	PBT	87	3.1	1.1	4.2*	175	.000
		PC	90	2.41	1.09			
	n Macht	PBT	87	5.54	1.73	-.523	175	.601
		PC	90	5.68	1.77			

* 5% Signifikanzniveau

Sub-Hypothese (b): Um die Gleichheit der Varianzen zu überprüfen wurde ein Levene's Test berechnet. Es zeigten sich keine signifikanten Unterschiede (siehe Tabelle 8). Die H_0 wird somit beibehalten und die Varianzen sind äquivalent.

Tabelle 8: Levene's Test zur Überprüfung der Homogenität der Varianzen

	PRF		OMT	
	F	p	F	p
Geselligkeit / n Bindung	1.461	.228	.162	.688
Leistung / n Leistung	.019	.890	.355	.552
Dominanz / n Macht	.110	.741	.210	.647

Sub-Hypothese (c): Um die Äquivalenz der Verteilungen der Rohwerte zu überprüfen, wurde ein Kolmogorov-Smirnov-Test berechnet. Die Ergebnisse, siehe Tabelle 9, zeigen ähnlich wie Tabelle 7 einen signifikanten Unterschied zwischen der Papier und Bleistift und der PC Gruppe bei der Skala „n Leistung“ des OMT ($Z=1.481$, $p=.025$). Für die Skalen des PRF-D kann somit jeweils die H_0 beibehalten werden. Beim OMT muss für die Skala „n Leistung“ die H_0 verworfen und die H_1 angenommen werden. Bei den Skalen „n Bindung“ und „n Macht“ hingegen wird die H_0 beibehalten.

Tabelle 9: Kolmogorov-Smirnov Test zur Überprüfung der Gleichheit der Rohwertverteilung

	PRF		OMT	
	Z	p	Z	p
Geselligkeit / n Bindung	.397	.998	.609	.852
Leistung / n Leistung	.403	.997	1.481*	.025
Dominanz / n Macht	.494	.967	.415	.995

* 5% Signifikanzniveau

Sub-Hypothese (d): Um die Äquivalenz der Rangreihung der Rohwerte zu überprüfen, wurde ein Mann-Whitney Test berechnet. Das Ergebnis (siehe Tabelle 10) zeigt wiederum einen signifikanten Unterschied in der OMT Skala „n Leistung“ ($Z=-3.977$, $p < .001$). Diese Ergebnisse unterstützen die Sub-Hypothesen (a) und (c). Für diese Skala wird wiederum die

H_0 verworfen und die H_1 angenommen. Für die restlichen Skalen des OMT, sowie für alle Skalen des PRF-D, wird die H_0 beibehalten.

Tabelle 10: Mann-Whitney Test zur Überprüfung der Äquivalenz der Rohwert-Rangreihung

		Group	N	Mean Rank	Z	p
PRF	Geselligkeit	PC	87	87.94	-.908	.884
		PBT	89	89.05		
	Leistung	PC	87	90.79	-3.977	.553
		PBT	89	86.26		
	Dominanz	PC	87	87.03	-.241	.704
		PBT	89	89.94		
OMT	n Bindung	PC	87	92.44	-.908	.364
		PBT	90	85.68		
	n Leistung	PC	87	103.96	-3.977*	.000
		PBT	90	74.54		
	n Macht	PC	87	88.07	-.241	.809
		PBT	90	89.90		

* 5% Signifikanzniveau

5.4 Hypothese 2

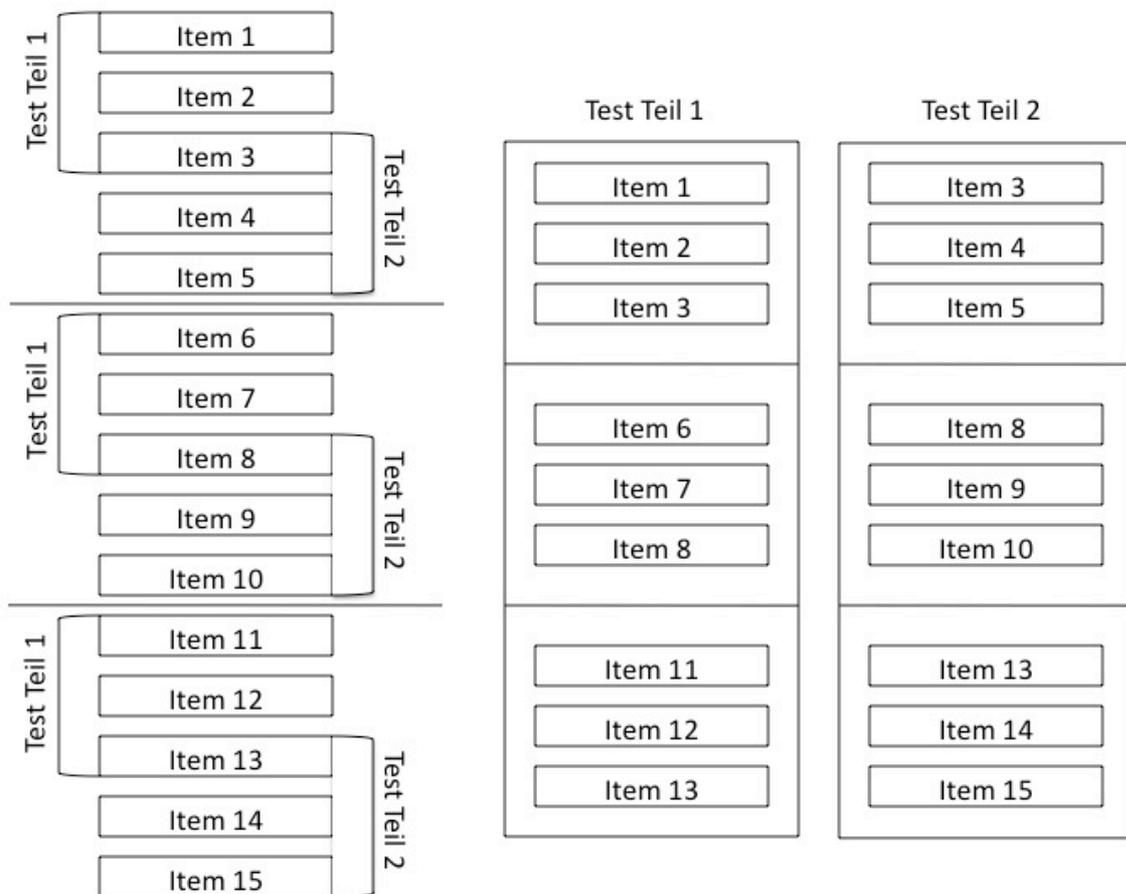
Sub-Hypothese (a): Für die Berechnung der Split-Half Reliabilität wurde jeweils der Guttman Split-Half Koeffizient berechnet (siehe Tabelle 11).

Um die Split-Half Reliabilität des OMT zu berechnen wurde folgendermaßen vorgegangen: Aufgrund des Testdesigns messen die ersten fünf Items tendenziell „n Bindung“, die nächsten fünf „n Leistung“ und die letzten fünf „n Macht“. Es sei darauf hingewiesen, dass

es nicht „falsch“ ist, wenn die zu untersuchende Person eine Geschichte schreibt deren Motivkodierung nicht zu der genannten passt. Diese wird außerdem nicht gesondert verrechnet (siehe Kuhl et al., 1999). Die Items haben wie gesagt nur die Tendenz diese Motive zu messen. Demnach kann jedes Item auch jedes Motiv messen. Eine Halbierung in der Mitte könnte somit zu einer negativen Korrelation führen. Jedoch nur die fünf Items, die ein Motiv erheben sollen zu messen, würde auch nicht korrekt sein, da alle 15 Items jedes Motiv erheben können.

Nach diesen Überlegungen wurde der Test wie folgt aufgeteilt (siehe Abbildung 5): Die ersten drei Items jedes Motives wurden zum Test-Teil 1 zusammengefügt, die letzten drei zum Test-Teil 2. Dadurch wurde der Testkonstruktion optimal Rechnung getragen und die Reliabilitäten wie in Tabelle 11 angeführt berechnet.

Abbildung 5: Aufteilung der OMT-Items zur Berechnung der Split-Half Reliabilität



Die für den von Hofer (2009) adaptierten PRF-D mit Hilfe des Gutman Split-Half Koeffizienten berechneten Reliabilitäten (siehe Tabelle 11) sind vergleichbar mit denen der Originalversion, welche bei .66 bis .87 lagen (Stumpf et al., 1985). Die von Hofer (2009) mit Hilfe des Gutman Split-Half Koeffizienten berechnete Reliabilität für den PRF-D lag bei .76 bis .81. Die Reliabilitäten in Tabelle 11 zeigen weiters einen signifikanten Unterschied zwischen den Vorgabemodi in der Skala „Dominanz“ des PRF-D. Alle restlichen Skalen zeigen keine signifikanten Unterschiede in den Reliabilitäten zwischen den beiden Vorgabemodi. Für alle Skalen kann somit die H_0 beibehalten werden. Nur für die Skala „Dominanz“ des PRF-D muss die H_0 verworfen und die H_1 angenommen werden.

Tabelle 11: Gutman Split-Half Koeffizienten und Signifikanztest zweier Korrelationen nach Fisher Z-Transformation

	PRF			OMT				
		r	N	p	r	N	p	
Geselligkeit / n Bindung	PBT	.68	92	.8	PBT	.64	92	.91
	PC	.70	90		PC	.65	90	
Leistung / n Leistung	PBT	.84	92	.244	PBT	.53	92	.856
	PC	.78	90		PC	.51	90	
Dominanz / n Macht	PBT	.72	92	.021*	PBT	.70	92	.898
	PC	.85	90		PC	.69	90	

* 5% Signifikanzniveau

Sub-Hypothese (b): Um die Reliabilität im Sinne der Stabilität zu berechnen, wurde eine bivariate Pearson Korrelation berechnet (siehe Spalte „ r_{tt} “ in Tabelle 12). Die Retest Reliabilität für den von Hofer (2009) adaptierten PRF-D ist sehr hoch (.70 und darüber) und bei jedem Motiv und in jeder Gruppe signifikant.

Die Retest Reliabilität für den OMT ist gering und mit Ausnahme des Motives „n Macht“ ($r_{tt}=.36$, $p =.002$) nicht signifikant. Diese Retest Reliabilität ist geringer als die in einer Metaanalyse von Schultheiss (2007) gefundene $r_{tt}=.37$ nach einem Jahr.

Tabelle 12: Stabilitäten des PRF und OMT nach 8 Monaten und Signifikanztest zweier Korrelationen nach Fisher Z-Transformation.

			r_{tt}	p_{rtt}	N	p_{korr}
PRF	Geselligkeit	PBT	.81	.000	92	.712
		PC	.79	.000	90	
	Leistung	PBT	.77	.000	92	.31
		PC	.70	.000	90	
	Dominanz	PBT	.77	.000	92	.266
		PC	.83	.000	90	
OMT	n Bindung	PBT	.17	.118	92	
		PC	.00	.971	90	
	n Leistung	PBT	.18	.102	92	
		PC	.16	.126	90	
	n Macht	PBT	.02	.849	92	
		PC	.36	.001	90	

* 5% Signifikanzniveau

Wiederum konnten eine signifikante Differenzen zwischen den beiden Vorgabemodi gefunden werden (siehe Spalte „ p_{korr} “ in Tabelle 12) und zwar bei der Skala „n Macht“ des OMT. Für alle Skalen kann somit die H_0 beibehalten werden. Nur für die Skala „n Macht“ des OMT muss die H_0 verworfen und die H_1 angenommen werden. Dieses Ergebnis hängt klarerweise mit der für die PBT - Gruppe nicht signifikanten, jedoch in der PC-Gruppe

signifikanten Korrelation zusammen. Da hier die PC-Testung mit einer PBT von vor acht Monaten korreliert wurde, kann dieses Ergebnis nur eingeschränkt interpretiert werden.

5.5 Hypothese 3

Die subjektive Äquivalenz wurde anhand einer sieben stufigen Rating Skala von „nicht gleichwertig“ bis „gleichwertig“ gemessen, eine für den PRF-D und eine für den OMT. Um zu berechnen, ob die Studierenden diese zwei Vorgabemodi (Papier und Bleistift vs. PC) als gleichwertig empfinden oder nicht, wurde die Skala in zwei Gruppen geteilt. Die Bewertungen von 1-4 wurden in die Gruppe 0, mit dem Titel „nicht äquivalent“ zusammengefasst und die Bewertungen von 5-7 in die Gruppe 1, mit dem Titel „äquivalent“.

Beim PRF enthielt die PRF „Gruppe 0“ 17 Studierende und die „Gruppe 1“ enthielt 70 Studierende (siehe Tabelle 13). Ein Binomial Test ergab eine signifikante Differenz ($p < .001$) und zeigt somit, dass signifikant mehr Studierende den Test für äquivalent hielten. Die H_0 wird daher beibehalten.

Tabelle 13: Binomial-Tests zur subjektiven Äquivalenz

	PRF		OMT _{Gesamt}		OMT _{Gefiltert}	
	N	p	N	p	N	p
Gruppe 0	17	<.001*	51	.203	32	.550
Gruppe 1	70		38		38	

*5% Signifikanzniveau

Für den OMT waren die Ergebnisse nicht so eindeutig (Tabelle 13). Es befanden sich mehr Studierende in der Gruppe 0 als in der Gruppe 1 (51 bzw. 38 Personen). Der durchgeführte Binomial Test zeigte keinen signifikanten Unterschied zwischen den Gruppen ($p = .203$). Die subjektive Äquivalenz kann somit nicht als gegeben angenommen werden. Daher wird die H_0 verworfen und die H_1 angenommen.

Um dieses Ergebnis genauer zu betrachten wurden die Gründe für das unterschiedliche Empfinden untersucht, welche mittels Fragebogen erhoben wurden (siehe Abbildung 1). Nachdem jene Studierenden, die als Grund für die mangelnde Äquivalenz nur „das Schreiben mit einer Tastatur“ angegeben hatten, herausgefiltert wurden, enthielt die Gruppe 0 („nicht äquivalent“) 32 Studierende und die Gruppe 1 („äquivalent“) 38 Studierende. Wie Tabelle 13 zeigt, war auch hier die Differenz nicht signifikant ($p=.550$) und daher wird die H_0 verworfen und die H_1 angenommen. Daraus folgt, dass es für den OMT einen Unterschied in der subjektiven Äquivalenz zwischen den beiden Vorgabemodi gibt.

Um zu überprüfen, ob die Ergebnisse der Sub-Hypothese 1 (a) durch die fehlende subjektive Äquivalenz hervorgerufen wurden, wurde wiederum ein Welch Test berechnet. Diesmal jedoch nur für die Gruppe 1 („äquivalent“). Die Ergebnisse (siehe Tabelle 14) sind mit jenen der Sub-Hypothese 1 (a) vergleichbar (siehe Tabelle 7): es gibt einen signifikanten Unterschied in der Skala „n Leistung“ ($t=2.68$, $p=.009$, $df=70$) mit einer wiederum hohen Effektstärke von .53. Dieses Ergebnis scheint daher unabhängig von der subjektiven Äquivalenz zu sein.

Tabelle 14: Welch-Test zur Überprüfung der Äquivalenz der Rohwerte bei jenen Studierenden, die die beiden Vorgabemodi (PBT vs. PC) des OMT als subjektiv Äquivalent bewerteten

	Group	N	Mean	SD	t	df	p
n Bindung	PBT	87	3.25	1.13	1.10	66	.277
	PC	38	3.00	1.21			
n Leistung	PBT	87	3.10	1.10	2.68*	70	.009
	PC	38	2.53	1.11			
n Macht	PBT	87	5.54	1.73	.112	73	.903
	PC	38	5.50	1.67			

* 5% Signifikanzniveau

6. Diskussion und Ausblick

In der vorliegenden Studie wurde die Äquivalenz von Papier und Bleistifttestungen (PBT) und PC-Testungen bei Verfahren zur Erhebung von impliziten und expliziten Motivmaßen untersucht. Dabei wurde sowohl die psychometrische als auch die erfahrungsbezogene Äquivalenz sensu Wagner-Menghin (2003a) überprüft.

Die Ergebnisse für die psychometrische Äquivalenz des PRF-D fallen fast durchwegs positiv aus: Die Rohwerte der PBT-Gruppe der drei Skalen „Geselligkeit“, „Leistung“ und „Dominanz“ unterscheiden sich nicht signifikant von denen der PC-Gruppe. Auch die Varianzen unterscheiden sich nicht signifikant. Dieses Ergebnis ist mit den in der Literatur zur Äquivalenzmessung eines Fragebogens gefundenen Ergebnissen konform (Merten et al., 1996; Pinsoneault, 1996; Watson et al., 1992). Auch die erfahrungsbezogene Äquivalenz, erhoben durch die subjektive Bewertung der Studierenden über die Gleichheit der PC-Testung gegenüber einer PBT, zeigte ein positives Ergebnis: Nur 17 Studierende hielten den PRF-D am PC für nicht äquivalent, 70 hingegen hielten die beiden Vorgabemodi für äquivalent. Ein Binomial Test ergab einen signifikanten Unterschied ($p < .001$). Die erfahrungsbezogene Äquivalenz scheint damit gegeben. Die Kennwerte der Reliabilität im Sinne der Stabilität fallen in der PBT Gruppe mit .81 für die Skala „Geselligkeit“ und .77 für die beiden Skalen „Leistung“ und „Dominanz“, im Vergleich zu den von .67 bis .85 liegenden Werten der Originalversion des PRF-D nach Stumpf et al. (1985), nahezu gleich aus. Die Werte der PC Gruppe sind ähnlich hoch. Dieses Ergebnis ist jedoch aufgrund des Untersuchungs-Designs mit Vorsicht zu genießen: es wurde eine PBT mit einer PC-Testung korreliert, da in der ersten Datenerhebung von Hofer (2009) nur PBT durchgeführt wurden. Dies schränkt die Interpretation der Ergebnisse erheblich ein.

Die bisherigen Ergebnisse würden demnach für eine psychometrische Äquivalenz sprechen, jedoch unterscheidet sich bei der Skala „Dominanz“ die Reliabilität im Sinne der inneren Konsistenz nach dem Gutman Split-Half Koeffizienten zwischen der PBT-Gruppe und der PC-Gruppe signifikant (siehe Tabelle 11). Sie steigt von .72 in der PBT-Gruppe auf .85 in der

PC-Gruppe. Im Vergleich zum Wert von .82 für das Cronbach Alpha der Skala „Dominanz“ aus dem Manual des PRF-D nach Stumpf et al. (1985) liegt dieser Wert höher. Die übrigen Werte für den in dieser Studie am PC vorgegebenen PRF-D sind ähnlich den Originalwerten und schwanken um maximal .05 um sie herum. Diese an sich erfreuliche Steigerung der Reliabilität, die eine Steigerung der Messgenauigkeit des Tests zur Folge hat, bewirkt jedoch, dass die Skala „Dominanz“ für die in dieser Studie am PC vorgegebene Version des PRF-D nicht als psychometrisch äquivalent bezeichnet werden kann. Dieses Ergebnis widerspricht der Literatur zur Transformierung von Fragebögen von Paper auf den PC, die üblicherweise keinen Effekt findet (Merten et al., 1996; Pineseault, 1996; Watson et al., 1992). Die anderen Skalen, „Leistung“ und „Geselligkeit“, scheinen aufgrund der vorliegenden Ergebnisse psychometrisch äquivalent zu sein. Die erfahrungsbezogene Äquivalenz scheint den Ergebnissen dieser Studie zufolge jedoch für alle Skalen zu gelten. Ein möglicher Grund für die erhöhte Reliabilität innerhalb der Skala „Dominanz“ könnte in der von Kubinger (2009b) durch die Verwendung eines PC postulierten Erhöhung der Akzeptanz seitens der zu untersuchenden Person liegen.

Bei dem für diese Studie verwendeten OMT zeigen die vorliegenden Ergebnisse, dass sowohl die psychometrische als auch die erfahrungsbezogene Äquivalenz nicht als gegeben angenommen werden kann. Die Rohwert-Mittelwerte der Skala „n Leistung“ unterscheiden sich signifikant zwischen den beiden Vorgabemodi ($t=4.2$, $p<.001$, $df=175$) und die Ausprägung des Leistungsmotivs ist in der PBT-Gruppe mit einem Mittelwert von 3.1 höher als in der PC-Gruppe, deren Mittelwert bei 2.41 liegt. Die Effektstärke ist mit .63 als hoch zu beurteilen (Field, 2005) und zeigt somit, dass dieser Unterschied von großer Bedeutung ist. Weitere signifikante Unterschiede finden sich in der Verteilung ($Z=1.481$, $p=.025$) und der Rangreihungen der Mittelwerte ($Z=-3.977$, $p<.001$).

Die Reliabilitätswerte im Sinne der inneren Konsistenz nach dem Gutman Split-Half Koeffizienten differieren zwischen den beiden Vorgabemodi für keine der Skalen signifikant (siehe Tabelle 11). Die Stabilitätswerte (siehe Tabelle 12) unterscheiden sich zwar signifikant in der Skala „n Macht“, aber dieser Effekt hat zwei erhebliche Einschränkungen:

Zum Einen wurden die Ergebnisse durch eine Korrelation zweier unterschiedlicher Vorgabemodi erzielt und zum Anderen ist nur eine der beiden Korrelationen signifikant. Dies verringert die Aussagekraft dieses Ergebnisses. Bis auf dieses fragwürdige Resultat, zeigten sich für die Skalen „n Bindung“ und „n Macht“ keine signifikanten Unterschiede zwischen den beiden Vorgabemodi. Deren psychometrische Äquivalenz kann somit als gegeben bezeichnet werden.

Die erfahrungsbezogene Äquivalenz kann jedoch nicht als gegeben angenommen werden, da selbst nach herausfiltern jener Studierenden, die als Grund für eine mangelnde subjektive Äquivalenz nur „das Schreiben mit einer Tastatur“ angaben, zwei laut Binomial - Test nicht signifikante unterschiedliche Gruppen von 32 Studierenden mit der Bewertung „nicht äquivalent“ und 38 Studierende mit der Bewertung „äquivalent“ resultieren ($p=.550$, siehe Tabelle 13). Vor dieser Filterung waren es sogar 51 Studierende mit einer Bewertung als „nicht äquivalent“. Auch dies ist laut Binomial - Test wiederum ein nicht signifikanter Unterschied in der Bewertung ($p=.203$, siehe Tabelle 13). Der oben beschriebene signifikante Unterschied in den Rohwert-Mittelwerten der Skala „n Leistung“ bleibt auch dann erhalten, wenn nur Rohwerte von jenen Studierenden, die den OMT subjektiv als äquivalent einstufen, herangezogen werden ($t=2.68$, $p=.009$, $df=70$, siehe Tabelle 14). Auch die Effektstärke bleibt mit $.52$ hoch. Die implizite Leistungsmotivation der Studierenden war somit bei der PBT deutlich höher als bei der PC-Testung, unabhängig von der subjektiven Bewertung der Äquivalenz.

Die erfahrungsbezogene Äquivalenz für den in dieser Studie vorgelegten OMT kann, wie oben erwähnt, nicht angenommen werden. Ob dies nur durch die psychometrisch nicht äquivalente Leistungsskala zustande gekommen ist, konnte nicht geklärt werden, da eine Messung der subjektiven Äquivalenz je Skala nicht vorgenommen wurde und aufgrund der operant inhaltsanalytischen Erhebungsmethode auch nur schwer zu realisieren ist. Auch der Grund für diese ungleiche Ausprägung des Motives „n Leistung“ scheint nur schwer zu erklären, da die von Kubinger (2009b) postulierte Erhöhung der Akzeptanz seitens der untersuchenden Person einen Effekt in die entgegengesetzte Richtung zeigen müsste. Die

übrigen Skalen des OMT („n Bindung“ und „n Macht“) scheinen zumindest psychometrisch äquivalent zu sein allerdings kann auch hier die erfahrungsbezogene Äquivalenz nicht als gegeben angenommen werden.

Vermutlich liegt die Erklärung für den Unterschied in der Skala „n Leistung“ im Test-Design des OMT beziehungsweise dessen Bearbeitung: Zwar konnten die Studierenden mit Hilfe der Tabulator - Taste von Textfeld zu Textfeld springen, worauf sie auch in der Instruktion speziell hingewiesen wurden, allerdings verwendeten viele die Maus um zwischen den Textfeldern zu wechseln. Dies bedeutete eine wiederholte Unterbrechung des Schreibflusses und könnte somit zu Frustration und einer verringerten Bereitschaft zu Leisten geführt haben. Diese These der Frustrationserzeugung wird durch Gelegenheitsbeobachtungen seitens der Testleiter unterstützt. Da erst die Items 6-10 das Leistungsmotiv zu erheben versuchen, könnte sich die Frustration bis dahin aufgebaut haben und den signifikanten Unterschied bewirkt haben. Dieser Effekt könnte beim PRF durch das ständige Verwenden der Maus verhindert worden sein. Hier musste nichts getippt werden und daher kam es zu keiner Unterbrechung bei der Bearbeitung des Verfahrens. Das Antwortformat des OMT erfordert aber eine Bearbeitung mit der Tastatur. Eine andere Implementierung des OMT scheint technisch (derzeit) nicht möglich zu sein. Zudem zeigten sich manche Studierende über die Bearbeitung mittels Tastatur erfreut, da diese den in Hofer (2009) angesprochenen Schreibaufwand reduziert beziehungsweise angenehmer macht. Dies schließt allerdings nicht den Aufbau einer Frustration und damit verminderten Leistungsbereitschaft durch das ständige Wechseln von der Tastatur zur Maus und zurück aus. Zusätzliche Frustration könnte durch die konkrete Durchführung der Untersuchung erzeugt worden sein: bis auf eine Gruppe mussten alle Studierenden, die per Zufall in die PC-Gruppe eingeteilt wurden, den Raum verlassen und in den Computersaal der Fakultät für Psychologie wechseln. Führt man einen Welch-Test für die PC-Gruppe durch, die den Raum nicht verlassen musste, so verschwindet das signifikante Ergebnis in der Skala „n Leistung“ ($t=1.748$, $p=.096$, $df=19$). Zwar wird die Frustrationstoleranz, also das Aushalten eines Zustandes reduzierten positiven Affektes nach der PSI Theorie von Kuhl (2000, 2001) als zentral für das

Leistungsmotiv angesehen, jedoch wird dabei von einem für die Person wichtigen Ziel ausgegangen. Da die Bearbeitung der Verfahren im Rahmen eines Seminars mit Anwesenheitspflicht durchgeführt wurden und die Studierenden somit weder gänzlich freiwillig teilnahmen noch deren Ergebnisse irgendeinen Einfluss auf Noten oder ähnlich wichtiges hatten, scheint es plausibler, dass eine Frustration in einer mangelnden Leistungsbereitschaft resultiert.

Konträr dazu könnte der Grund für die geringere Leistungsmotivation auch in der angenehmeren Bearbeitung durch das Schreiben mittels Tastatur liegen. Von den 49 Begründungen, die bei einer subjektiven Bewertung als nicht äquivalent gegeben wurden, standen 20 in Verbindung mit der Tatsache, sich am PC aufgrund des Schreibens auf einer Tastatur weniger anstrengen zu müssen. Dadurch könnte das Leistungsmotiv absinken, da im Vergleich zum aufwändigen Schreiben mit der Hand die notwendige Anstrengung sinkt. Von den 17 Personen die den PRF-D als nicht äquivalent eingestuft hatten, wurde keine einzige Begründung abgegeben. Eine leichtere Bearbeitung scheint daher und aufgrund der erfahrungsbezogenen Äquivalenz nicht gegeben. Das Leistungsmotiv ist somit auch nicht von der PBT - Gruppe verschieden.

Diese scheinbar konträren Erklärungen – Frustration auf der einen Seite und leichtere Bearbeitung auf der anderen Seite – könnten zusammen für die Größe des gefundenen Effekts verantwortlich sein.

Die Konsequenz aus den Ergebnissen zur Äquivalenz des OMT ist, dass der Vorgabemodus bei der Messung des impliziten Leistungsmotivs eine erhebliche Rolle zu spielen scheint. Dies sollte auch bei zukünftigen Messungen, die womöglich ohne Äquivalenzprüfung durchgeführt werden, beachtet werden.

Da die Literatur zur Äquivalenzmessung bei Fragebögen zum Schluss kommt, dass eine Transformierung eines Fragebogens üblicherweise keinen Effekt zeigt (Merten et al., 1996; Pineseault, 1996; Watson et al., 1992), könnten die operant – inhaltsanalytischen Verfahren hier eine Sonderstellung einnehmen. Dem Ergebnis dieser Studie nach zu

urteilen, scheinen solche Verfahren nicht ohne Äquivalenzprüfung auf einen PC transformierbar zu sein. Vielleicht ist auch hier, wie oben erwähnt und ähnlich wie bei Kubinger et al. (1991), nicht der Test selbst sondern die Umsetzung das Problem. Eine nachfolgende Studie sollte daher die oben angesprochenen Probleme der Frustration weiter untersuchen und auch der Hypothese wonach die geringere Leistungsmotivation aufgrund der leichteren Bearbeitung zustande gekommen ist nachgehen.

Weiters muss die Erhebung der impliziten Motive stark kritisiert werden, da diese mittels projektiver Verfahren erfolgt (Rollett, 2006), auch wenn diese mittlerweile als operant inhaltsanalytisch bezeichnet werden.

Offen bleibt auch die Frage, ob die Reihenfolge der Vorgabe der beiden psychologisch-diagnostischen Verfahren (PRF-D bzw. OMT) einen Einfluss auf die Ergebnisse dieser Studie hatte. Eine nachfolgende Studie sollte daher auch die Reihenfolge zufällig variieren.

Es konnte auch in dieser Studie keine substantielle Korrelation zwischen den impliziten und expliziten Motivmaßen gefunden werden. Dies ist ein häufig in der Literatur gefundenes Phänomen (Brunstein, 2006; Scheffer et al., 2007; Schultheiss, 2008; Schultheiss et al., 2001). Die einzige signifikante Korrelation findet sich in der PBT - Gruppe ($r=-.252$, $p=.019$), mit einem geringen Bestimmtheitsmaß von nur 6% erklärter Varianz. Über beide Gruppen hinweg verschwindet dieser Effekt und die Korrelationen sind allesamt nicht signifikant (siehe Tabelle 6). Damit unterstützt diese Studie die Annahme zweier unabhängiger motivationaler Systeme (vgl. Brunstein, 2006; McClelland et al., 1989).

Es sei noch darauf hingewiesen, dass die vorliegenden Ergebnisse auf die Stichprobe beschränkt sind. Es wurden nur Psychologie-Studierende des zweiten Abschnittes des Diplomstudiums Psychologie an der Universität Wien getestet, wobei 80% davon Frauen waren. Für eine Generalisierung der Ergebnisse wäre daher eine erweiterte Stichprobe mit einem ausgeglichenen Geschlechteranteil von Nöten. Des Weiteren sollte eine zukünftige Studie auch das von Bartram (1994) vorgeschlagene 4-Gruppen Design verwenden, damit auch die Äquivalenz der Reliabilität im Sinne einer Stabilität überprüft werden kann.

Schlussendlich zeigen die vorgelegten Ergebnisse, dass eine Äquivalenzmessung, wie in der Literatur beschrieben (APA, 1986; de Beuckelaer et al., 2009; Ford et al. 1996; Hofer, 1985; Kindsvater et al., 2003; Kolen et al., 2004; Kubinger, 1993, 2009a; Kveton et al., 2004; Pommerich, 2004; Russell et al., 2003; Wagner-Menghin, 2003a), zwingend notwendig ist.

7. Zusammenfassung

Eine Äquivalenzmessung eines am PC vorgegebenen jedoch mit Hilfe von Papier und Bleistift Normen ausgewerteten Verfahrens, wird von vielen Autoren gefordert und dient einer empirischen Kontrolle der Gleichwertigkeit (APA, 1986; de Beuckelaer et al., 2009; Ford et al. 1996; Hofer, 1985; Kindsvater et al., 2003; Kolen et al., 2004; Kubinger, 1993, 2009a; Kveton et al., 2004; Pommerich, 2004; Russell et al., 2003; Wagner-Menghin, 2003a).

In der Motivforschung werden explizite (d.h. dem Bewusstsein zugängliche und daher direkt erfragbar) und implizite (d.h. dem Bewusstsein nicht direkt zugängliche und daher auch nicht direkt erfragbar) Motive unterschieden (McClelland et al., 1989).

In dieser Studie wurde die Äquivalenz der Vorgabemodi „Papier und Bleistift“ und „PC“ bei der Erhebung impliziter und expliziter Motivmaße überprüft. Die aufgestellten Hypothesen bezogen sich dabei sowohl auf die psychometrische Äquivalenz (d.h. Gleichheit der Mittelwerte, Varianzen, Reliabilitäten, Rangfolgen, Verteilung) als auch auf die erfahrungsbezogene Äquivalenz (d.h. die subjektive Bewertung der Äquivalenz der beiden Vorgabemodi).

Für die Überprüfung der Hypothesen wurde ein experimentelles Retest Design erstellt. Zur Erhebung der expliziten Motive wurde der PRF-D verwendet und die impliziten Motive wurden mit Hilfe des OMT erhoben. Beide Verfahren wurden in der von Hofer (2009) adaptierten Version vorgegeben. Die Testungen fanden im Rahmen des Seminars „Übungen zur Psychologischen Diagnostik II“ statt und die Studierenden wurden zufällig in zwei

Gruppen aufgeteilt. Die erste Gruppe bearbeitete beide Verfahren auf Papier-Bleistift Basis, die zweite Gruppe bearbeitete beide Verfahren am PC. Der zweiten Gruppe wurde nach der Bearbeitung noch ein Fragebogen zur Erfassung der erfahrungsbezogenen Äquivalenz vorgelegt.

Die Stichprobe bestand aus 177 Studierenden (150 Frauen, 27 Männer) aus dem zweiten Abschnitts des Psychologiestudiums an der Universität Wien. Das Durchschnittsalter lag bei 25 Jahren, die Spannweite von 20 bis 86 Jahren. Somit lagen 82,4% der Studierenden im Alter zwischen 20 und 26 Jahren.

Die Ergebnisse des PRF-D zeigen nur bei der Reliabilität im Sinne der inneren Konsistenz einen signifikanten Unterschied: in der Papier und Bleistift Gruppe liegt der Gutman Split-Half Koeffizient bei .72, in der PC Gruppe bei .85. Eine Fisher Z-Transformation ergibt ein signifikantes Ergebnis ($p = .021$). Die restlichen Werte der psychometrischen Äquivalenz zeigen keine signifikanten Unterschiede. Die erfahrungsbezogene Äquivalenz kann mit 17 Studierenden mit einer Bewertung als „nicht äquivalent“ zu 70 Studierenden mit der Bewertung als „äquivalent“ - ein laut Binomial - Test signifikantes Ergebnis ($p < .001$) - angenommen werden.

Bei Überprüfung der psychometrischen Äquivalenz des OMT zeigten sich mehrere signifikante Unterschiede: der Rohwert-Mittelwerte ($t=4.2$, $p < .001$, $df=175$), die Verteilung ($Z=1.481$, $p=.025$) und die Rangreihung ($Z=-3.977$, $p < .001$) der Skala „n Leistung“ zeigten einen signifikanten Unterschied zwischen den beiden Vorgabemodi. Die beiden anderen Skalen, „n Bindung“ und „n Anschluss“, wiesen keine signifikanten Unterschiede auf. Auch die erfahrungsbezogene Äquivalenz kann den vorliegenden Ergebnissen nach nicht angenommen werden - 51 Studierende bewerteten die Vorgabemodi als „nicht äquivalent“, 38 Studierende gaben die Bewertung „äquivalent“ ab - ein nach Binomial - Test nicht signifikantes Ergebnis ($p=.203$).

Beide Verfahren erwiesen sich als nicht äquivalent. Der PRF-D zeigte einen signifikanten Unterschied bei der Reliabilität im Sinne der inneren Konsistenz. Die PC Version erzielte mit

einem Gutman Split-Half Koeffizienten von .85 im Vergleich zur Papier Bleistift Gruppe, mit einem Gutman Split-Half Koeffizienten von .72, eine deutlich höhere Messgenauigkeit. Grund dafür könnte die von Kubinger (2009b) postulierte höhere Akzeptanz einer PC-Testung seitens der zu untersuchenden Person sein, wenngleich die erfahrungsbezogene Äquivalenz gegeben scheint und dieses Ergebnis nur auf die Skala „Dominanz“ beschränkt bleibt.

Die signifikant niedrigeren Rohwerte des Leistungsmotivs der PC-Gruppe des OMT könnten mit dem Test-Design (Frustration durch ständigen Wechsel von Tastatur auf Maus und zurück), dem Untersuchungsplan (Frustration durch den Wechsel in einen anderen Raum) oder durch den von vielen Studierenden genannte verringerte Aufwand bei der Testbearbeitung (durch das Schreiben mittels Tastatur) begründbar sein. Der letztgenannte Punkt war wohl auch Grund für die nicht vorhandene erfahrungsbezogene Äquivalenz des Verfahrens.

Abschließend sei drauf hingewiesen, dass wiederum die Wichtigkeit einer Äquivalenzprüfung gezeigt wurde, wie sie auch in der Literatur gefordert wird (APA, 1986; de Beuckelaer et al., 2009; Ford et al. 1996; Hofer, 1985; Kindsvater et al., 2003; Kolen et al., 2004; Kubinger, 1993, 2009a; Kveton et al., 2004; Pommerich, 2004; Russell et al., 2003; Wagner-Menghin, 2003a). Beide Verfahren zeigten zwischen den Vorgabemodi trotz sorgfältig durchgeführter Transformierung einen signifikanten Unterschied. Das Ergebnis des PRF-D überrascht, da die Transformierung eines Fragebogens typischerweise keinen Effekt zeigt (Merten & Ruck, 1996; Pinsoneault, 1996; Watson et al., 1992).

Weitere Studien sollten sich mit den Gründen für die niedrigere Ausprägung des Leistungsmotivs eines am PC vorgegebenen OMT beschäftigen und versuchen, die oben genannten Begründungen zu falsifizieren. Auch sollte die Erfassung der erfahrungsbezogenen Äquivalenz genauer erfolgen und das Thema Tastatur eventuell speziell behandelt werden. Die Generalisierbarkeit der Ergebnisse ist aufgrund der Stichprobe eingeschränkt. Zukünftige Studien sollten daher versuchen, den

Geschlechteranteil auszugleichen und zumindest mehrere Studienrichtungen in Betracht ziehen.

IV. Literaturverzeichnis

Amelang, M. & Schmidt-Atzert, L. (2006). *Psychologische Diagnostik und Intervention* (4. Aufl.). Heidelberg: Springer.

Atkinson, J. W. & McClelland, D. C. (1948). The projective expression of needs: II. The effect of different intensities of the hunger drive on thematic apperception. *Journal of Experimental Psychology*, 33, 643-658.

Atkinson, J.W. (1964). *An introduction to motivation*. Princeton, NJ: Van Nostrand.

Atkinson, J. W. (1981). Studying personality in the context of an advanced motivational psychology. *American Psychologist*, 36, 117-128.

APA (1986). *Guidelines for computer-based tests and interpretations*. Washington, DC: Author.

Bachman, J. W. (2003). The patient-computer interview: A neglected tool that can aid the clinician. *Mayo Clinic Proceedings*, 78(1), 67-78.

Bartram, D. (1994). Computer-based assessment. In C. L. Cooper & I. T. Robertson (Hrsg.), *International review of industrial and organizational psychology* (9. Aufl., S. 31- 69). Chichester: Wiley.

Berufsverband Deutscher Psychologen. (1988). *Richtlinien für die Erstellung psychologischer Gutachten*. Bonn: Deutscher Psychologen Verlag.

Booth, J. (1995). Computerdiagnostik. In R.S. Jäger & F. Petermann (Hrsg.) *Psychologische Diagnostik* (3. Aufl., S. 186-197). Weinheim: PVU.

- Brunstein, J. C., Schultheiss, O. C. & Grässmann, R. (1998). Personal goals and emotional well-being: The moderating role of motive dispositions. *Journal of Personality and Social Psychology*, 75(2), 494-508.
- Brunstein, J. C. & Hoyer, S. (2002). Implizites und explizites Leistungsstreben: Befunde zur Unabhängigkeit zweier Motivationssysteme. *Zeitschrift für Pädagogische Psychologie*, 16, 51-62.
- Brunstein, J. C. & Maier, G. W. (2005). Implicit and self-attributed motives to achieve: two separate but interacting needs. *Journal of Personality and Social Psychology*, 89(2), 205-222.
- Brunstein, J. C. (2006). Implizite und explizite Motive. In J. Heckhausen & H. Heckhausen (Hrsg.), *Motivation und Handeln* (3. Aufl., S. 235-253). Berlin: Springer.
- Brunstein, J. C. & Heckhausen, H. (2006). Leistungsmotivation. In J. Heckhausen & H. Heckhausen (Hrsg.), *Motivation und Handeln* (3. Aufl., S. 235-253): Berlin. Springer.
- de Beuckelaer, A. , Lievens, F. (2009). Measurement Equivalence of Paper-and-Pencil and Internet Organisational Surveys: A Large Scale Examination in 16 Countries. *Journal of Applied Psychology*, 58, 336-361.
- Ferrando, P. J. & Lorenzo-Seva, U. (2005). IRT-based factor analytic procedures for testing the equivalence of paper-and-pencil and internet-administered questionnaires. *Psychological Methods*, 10, 193-205.
- Field, A. P. (2005). *Discovering statistics using SPSS: and sex and drugs and rock 'n' roll* (2. Aufl.). London: Sage publications.
- Ford, B. D., Vitelli, R. & Stuckless, N. (1996). The effects of computer vs. paper and pencil administration on measures of anger and revenge with an inmate population. *Computers in Human Behavior*, 12(1), 159-166.

- Fuchs, J. (2000). *Ein Vergleich der Bewertung von visuellen und audiovisuellen Instruktionen bei Computertests*. Unveröffentlichte Diplomarbeit, Universität Wien.
- Garb, H. N. (2007). Computer-Administered Interviews and Rating Scales. *Psychological Assessment*, 19(1), 4-13.
- Greaud, V. A. & Green, B. F. (1986). Equivalence of conventional and computer presentation of speed tests. *Applied Psychological Measurement*, 10(1), 23-34.
- Hall, J. L., Stanton, S. J. & Schultheiss, O. C. (2010). Biopsychological and neural processes of implicit motivation. In O. C. Schultheiss & J. C. Brunstein (Hrsg.), *Implicit motives* (S. 279-307). New York: Oxford University Press.
- Stumpf, H., Angleitner, A., Wieck, T., Jackson, D.N. & Beloch-Till, H. (1985). Deutsche Personality Research Form (PRF). Göttingen: Hogrefe .
- Hofer, P. J. (1985). Developing Standards for Computerized Psychological Testing. *Computers in Human Behavior*, 1(1), 301-315.
- Hofer, S. (2009). *Die Entwicklung und Erprobung einer multimethodischen Batterie zur Erhebung des Leistungsmotivs bei Studierenden*. Unveröffentlichte Diplomarbeit, Universität Wien.
- Jackson, D. N. (1974). *Manual for the Personality Research Form*. Goshen: Research Psychology Press.
- Jäger, R. S. (1990). Computerdiagnostik - eine Einführung. *Diagnostica*, 36, 91-95.
- Kindsvater, S. & Sturm, W. (2003). Computer- vs. Papier-Bleistiftvorgabe: Äquivalenzstudie zum nonverbalen Lerntest (NVLT). *Zeitschrift für Neuropsychologie*, 14(1), 13-21.
- Klinck, D. (2002). *Computergestützte Diagnostik*. Göttingen: Hogrefe.

- Kolen, M. J. & Brennan, R.L. (2004). *Test equating, scaling, and linking: Methods and practices*. New York: Springer.
- Kubinger, K. D. & Farkas, M. G. (1991). Die Brauchbarkeit der Normen von Papier-Bleistift-Tests für die Computer-Vorgabe: Ein Experiment am Beispiel der SPM von Raven als kritischer Beitrag. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 12, 257–266.
- Kubinger, K. D. (1993). Testtheoretische Probleme der Computerdiagnostik. *Zeitschrift Für Arbeits- Und Organisationspsychologie*, 37, 130-137.
- Kubinger, K. D. (2009a). Psychologische Computerdiagnostik. *Zeitschrift für Psychiatrie, Psychologie und Psychotherapie*, 57(1), 23-32.
- Kubinger, K. D. (2009b). *Psychologische Diagnostik – Theorie und Praxis psychologischen Diagnostizierens* (2. Aufl.). Göttingen: Hogrefe.
- Kuhl, J. & Scheffer, D. (1999). *Auswertungsmanual für den Operanten Motiv-Test (OMT)*. Universität Osnabrück.
- Kuhl, J. (2000). A functional-design approach to motivation and volition: The dynamics of personality systems interactions. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Hrsg.), *Self-regulation: Directions and challenges for future research* (S. 111-169). New York: Academic Press.
- Kuhl, J. (2001). *Motivation und Persönlichkeit: Interaktionen psychischer Systeme*. Göttingen: Hogrefe.
- Kuhl, J., Scheffer, D. & Eichstaedt, J. (2003). Der Operante Motiv. Test (OMT): Ein neuer Ansatz zur Messung impliziter Motive. In J. Stiensmeier-Pelster & F. Rheinberg (Hrsg.), *Diagnostik von Motivation und Selbstkonzept* (S. 129-150). Göttingen: Hogrefe.
- Kveton, P., Jelinek, M., Voboril, D. & Klimusova, H. (2007). Computer-based tests: The impact of test design and problem of equivalency. *Computers in Human Behavior*, 23(1) Jan 2007, 32-51.

- Langens, T. A. (2007). Congruence between implicit and explicit motives and emotional wellbeing: The moderating role of activity inhibition. *Motivation and Emotion, 31*, 49-59.
- McClelland, D. C. & Pilon, D. A. (1983). Sources of adult motives in patterns of parent behavior in early childhood. *Journal of Personality and Social Psychology, 44*, 564-574.
- McClelland, D. C., Atkinson, J. W., Clark, R. A. & Lowell, E. L. (1953). *The achievement motive*. New York: Appleton-Century-Crofts.
- McClelland, D. C. (1958). Methods of measuring human motivation. In J. W. Atkinson (Hrsg.), *Motives in fantasy, action, and society* (S. 7-42). Princeton: Van Nostrand.
- McClelland, D. C. (1987). *Human motivation*. New York: Cambridge University Press.
- McClelland, D. C., Koestner, R. & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review, 96*, 690-702.
- Mead, A. D. & Drasgow, F. (1993). Equivalence of computerized and paper-and-pencil cognitive ability tests: A meta-analysis. *Psychological Bulletin, 114*, 449-458.
- Menghin, S. & Kubinger, K.D. (1996). Zur Legende: „Test-personen beantworten dem Computer persönliche und intime Fragen offener als einem Testleiter“ – Ergebnisse eines Experimentns. *Zeitschrift für Differentielle und Diagnostische Psychologie, 17*, 163-169.
- Merten, T. & Ruch, W. (1996). A comparison of computerized and conventional administration of the German versions of the Eysenck Personality Questionnaire and the Carroll Rating Scale for Depression. *Personality & Individual Differences, 20*, 281-291.
- Murray, H. A. (1938). *Explorations in personality*. New York: Oxford University Press.
- Murray, H. A. (1943). *Thematic Apperception Test Manual*. Cambridge: Harvard University Press.

- Neubauer, A. C., Urban, E. & Malle, F. M. (1991). Raven's Advanced Matrices: Computerunterstützte Präsentation versus Standardvorgabe. *Diagnostica*, 37, 204–212.
- Pinsonneault, T. B. (1996). Equivalency of computer-assisted and paper-and-pencil administered versions of the Minnesota Multiphasic Personality Inventory-2. *Computers in Human Behavior*, 12(2), 291–300.
- Pommerich, M. (2004). Developing computerized versions of paper-and-pencil tests: Mode effects for passage-based tests. *Journal of Technology, Learning, and Assessment*, 2(6), 3-44.
- Rammstedt, B., Holzinger, B. & Rammsayer, T. H. (2004). Zur Äquivalenz der Papier-Bleistift- und einer computergestützten Version des NEO-Fünf-Faktoren-Inventars (NEO-FFI). *Diagnostica*, 50, 88 - 97.
- Rollet, B. (2006). Projektive Verfahren. In K.D. Kubinger & R.S. Jäger (Hrsg.), *Schlüsselbegriffe der Psychologischen Diagnostik* (S. 340-348). Weinheim: PVU.
- Russell, M., Goldberg, A. & O'Connor, K. (2003). Computer-based testing and validity: A look back into the future. *Assessment in Education*, 10, 278-293.
- Scheffer, D., Kuhl, J. & Eichstaedt, J. (2003). Der Operante Motiv-Test (OMT): Inhaltsklassen, Auswertung, psychometrische Kennwerte und Validierung. In F. Rheinberg, und J. Stiensmeier-Pelster (Hrsg.), *Diagnostik von Motivation und Selbstkonzept*, (S. 151-168). Göttingen: Hogrefe.
- Scheffer, D., Eichstaedt, J., Chasiotis, A. & Kuhl, J. (2007). Towards an integrated measure of need affiliation and agreeableness derived from the Operant Motive Test. *Psychology Science*, 49(4), 308-324.
- Schultheiss, O. C. & Brunstein, J. C. (1999). Goal imagery: Bridging the gap between implicit motives and explicit goals. *Journal of Personality*, 67, 1-38.

- Schultheiss, O. C. (2001). An information processing account of implicit motive arousal. In M. L. Maehr & P. Pintrich (Hrsg.), *Advances in motivation and achievement* (12. Aufl., S. 1-41). Greenwich: JAI Press.
- Schultheiss, O. C. & Brunstein, J. C. (2001). Assessment of implicit motives with a research version of the TAT: Picture profiles, gender differences, and relations to other personality measures. *Journal of Personality Assessment*, *77*(1), 71-86
- Schultheiss, O. C. & Brunstein, J. C. (2005). An implicit motive approach to competence. In A. J. Elliot & C. S. Dweck (Hrsg.), *Handbook of competence and motivation* (S. 31-51). New York: Guilford.
- Schultheiss, O. C. & Pang, J. S. (2007). Measuring implicit motives. In R. W. Robins, R. C. Fraley & R. Krueger (Hrsg.), *Handbook of Research Methods in Personality Psychology* (S. 322-344). New York: Guilford.
- Schultheiss, O. C. (2008). Implicit motives. In O. P. John, R. W. Robins & L. A. Pervin (Hrsg.), *Handbook of Personality: Theory and Research* (3. Aufl., S. 603-633). New York: Guilford.
- Schultheiss, O. C., Wirth, M. M., Waugh, C. E., Stanton, S. J., Meier, E. & Reuter-Lorenz, P. (2008). Exploring the motivational brain: Effects of implicit power motivation on brain activation in response to facial expressions of emotion. *Social Cognitive and Affective Neuroscience*, *3*(4), 333-343.
- Schultheiss, O. C., Yankova, D., Dirlikov, B. & Schad, D. J. (2009). Are Implicit and Explicit Motive Measures Statistically Independent? A Fair and Balanced Test Using the Picture Story Exercise and a Cue- and Response-Matched Questionnaire Measure. *Journal of Personality Assessment*, *91*, 72-81.
- Sheldon, K. M. & Kasser, T. (1995). Coherence and congruence: Two aspects of personality integration. *Journal of Personality and Social Psychology*, *68*, 531-543.

- Spangler, W. D. (1992). Validity of questionnaire and TAT measures of need for achievement: Two meta-analyses. *Psychological Bulletin*, 112(1), 140-154.
- Stumpf, H., Angleitner, A., Wieck, T., Jackson, D. N. & Beloch-Till, H. (1985). *Deutsche Personality Research Form (PRF)*. Göttingen: Hogrefe.
- Wagner-Menghin, M. M. (2003a). Äquivalenzprüfung. In K.D. Kubinger & R.S. Jäger (Hrsg.), *Schlüsselbegriffe der Psychologischen Diagnostik* (S. 68-82). Weinheim: PVU.
- Wagner-Menghin, M. M. (2003b). Computerdiagnostik. In K.D. Kubinger & R.S. Jäger (Hrsg.), *Schlüsselbegriffe der Psychologischen Diagnostik* (S. 68-82). Weinheim: PVU.
- Watson, Ch. G., Thomas, D. & Anderson, P. E. D. (1992). Do computer administered Minnesota Multiphasic Personality Inventories underestimate booklet-based scores? *Journal of Clinical Psychology*, 48(6), 744-748.
- Willmes, K. (2003). Neuropsychologische Diagnostik. In K.D. Kubinger & R.S. Jäger (Hrsg.), *Schlüsselbegriffe der Psychologischen Diagnostik* (S. 287-97). Weinheim: PVU.

V. Anhang

Anhang A: Quellen

Ich habe mich bemüht, sämtliche Inhaber der Bildrechte ausfindig zu machen und ihre Zustimmung zur Verwendung der Bilder in dieser Arbeit eingeholt. Sollte dennoch eine Urheberrechtsverletzung bekannt werden, ersuche ich um Meldung bei mir.

Anhang B: Tabellenverzeichnis

Tabelle 1: Aufbau der ersten und zweiten Datenerhebung	19
Tabelle 2: Durchschnittliche Cronbach-Alpha Koeffizienten des OMT nach Scheffer et al. (2003)	23
Tabelle 3: Aufteilung der Gruppen in Hofer (2009).....	30
Tabelle 4: Deskriptive Statistiken der Stichprobe.....	34
Tabelle 5: Übereinstimmungen der drei Bewerter des OMT	35
Tabelle 6: Korrelation der impliziten und expliziten Motivmaße	36
Tabelle 7: Welch Test zu Überprüfung der Differenzen in den Rohwert-Mittelwerten der PBT und PC Gruppe für den PRF-D und den OMT	37
Tabelle 8: Levene's Test zur Überprüfung der Homogenität der Varianzen	38
Tabelle 9: Kolmogorov-Smirnov Test zur Überprüfung der Gleichheit der Rohwertverteilung	38
Tabelle 10: Mann-Whitney Test zur Überprüfung der Äquivalenz der Rohwert-Rangreihung	39
Tabelle 11: Gutman Split-Half Koeffizienten und Signifikanztest zweier Korrelationen nach Fisher Z-Transformation	41
Tabelle 12: Stabilitäten des PRF und OMT nach 8 Monaten und Signifikanztest zweier Korrelationen nach Fisher Z-Transformation.....	42
Tabelle 13: Binomial-Tests zur subjektiven Äquivalenz.....	43

Tabelle 14: Welch-Test zur Überprüfung der Äquivalenz der Rohwerte bei jenen Studierenden, die die beiden Vorgabemodi (PBT vs. PC) des OMT als subjektiv Äquivalent bewerteten 44

Anhang C: Abbildungsverzeichnis

Abbildung 1: Fragebogen zur Erfassung der subjektiven Äquivalenz	27
Abbildung 2: Beispiel Item des OMT am PC.....	28
Abbildung 3: Beispiel Item der PC Implementierung des PRF-D	29
Abbildung 4: Altersverteilung der Stichprobe	33
Abbildung 5: Aufteilung der OMT-Items zur Berechnung der Split-Half Reliabilität	40

Curriculum Vitae

JÜRGEN GRAFENEDER

PERSONAL INFORMATION

Date and Place of Birth: 22.05.1985 in Amstetten, Austria

Citizenship: Austria

Address: Schaltberg 33a
3323 Neustadt/Donau
Austria

E-Mail: Juergen.Grafeneder@gmx.at

EDUCATION

1991-1995	Elementary School in Neustadt/D.
1995-1999	Comprehensive School School in Neustadt/D.
1999-2004	HTL St. Pölten, dept. Data Processing & Organisation
2005-2007	Psychology at the University of Vienna (Level One)
2007-2010	Psychology at the University of Vienna (Level Two)
07/2009 – 12/2009	Exchange semester, MacQuarie University, Sydney

WORK EXPERIENCE

08/2001	Data processing & administration at the magistrate in Amstetten
---------	-----------------------------------------------------------------

07/2002	Shift work at the company “Umdasch” in Amstetten, Austria
08/2002	Data processing & administration at the magistrate of Neustadt/D.
08/2003 - 09/2003	Shift work at the company “Neusiedler” in Neuhofen/Ybbs, Austria
07/2006 - 08/2006	Shift work at the company “Buntmetall” in Amstetten, Austria
10/2006 - 10/2008	Data processing part time job (TYPO3, JavaScript, HTML)
08/2007	Shift work at the company “Buntmetall” in Amstetten, Austria
03/2008 - 06/2008	Internship at the Department of Developmental Psychology and Psychological Assessment at the University of Vienna
07/2008 - 08/2008	Full-time job at the Red Cross Austria in Melk
10/2008 – 01/2009	Assistant of Professor Kubinger at the Department of Developmental and Psychological Assessment at the University of Vienna
10/2008 - 01/2009	Student Mentor for first-year Psychology students
03/2009 – 06/2009	Teaching Assistant at the Department of Developmental and Psychological Assessment at the University of Vienna
03/2010 – 06/2010	Tutor & Teaching Assistant at the Department of Developmental and Psychological Assessment at the University of Vienna

05/2010 -

EU Task Force

Research Services and International Relationships

University of Vienna

OTHER SKILLS

Student Mentor and Tutor

Expert knowledge in HTML, Typo3 and CSS

Good knowledge in Adobe CS3

EMT Intermediate

Java-Certification (SCJD, SCWCD)

Languages

English (business fluent, TOEFL-iBT: 110/120 points, exchange semester at the MacQuarie University in Sydney)

HOBBIES

Bouldering, Climbing, Hiking, voluntary service at the Red Cross Austria in Amstetten and Melk, Swimming, Belletristic Literature