



universität
wien

Diplomarbeit

Zur Praktikabilität des IAT in der Fallbehandlung –
Machbarkeit, Akzeptanz und wiederholte Retest-
Reliabilität im Vergleich zu ILICA

Elena Simon

Angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag.^a rer. nat.)

Wien, im Jänner 2011

Studienkennzahl: 298

Studienrichtung: Psychologie

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

Danksagung

Mein Dank gilt meinen Eltern, die mir durch ihre großzügige Unterstützung das Studium ermöglichen haben, sowie meiner Schwester Lydia, die mir immer eine große Stütze war.

Besonderer Dank gebührt auch meinem Lebenspartner Mag. Bernd Plank, der mich nicht nur emotional unterstützte, sondern auch Wegbereiter eines Großteils der Testungen war.

Bei Univ.-Prof. Mag. Dr. Klaus Kubinger bedanke ich mich für die Betreuung meiner Diplomarbeit, seine Anregung, Kritik und Hilfe.

Herzlicher Dank gilt auch den SchülerInnen, DirektorInnen und LehrerInnen des Gymnasiums Sachsenbrunn, des Oberstufenrealgymnasiums Ternitz und des Bundesrealgymnasiums Gröhrmühlgasse, die mir die Durchführung der Testungen ermöglichten. Besonderer Dank gebührt meiner ehemaligen Psychologielehrerin Mag.^a Elfriede Wiedner, die mir durch ihr Engagement und ihre Kooperationsbereitschaft die ersten Testungen erleichterte.

INHALTSVERZEICHNIS

EINLEITUNG	9
THEORETISCHER TEIL	13
1. Indirekte psychologisch-diagnostische Verfahren und Objektive Persönlichkeitstests	15
1.1. Implizite Verfahren	17
1.1.1. Selbstwert-IAT	18
1.2. Simulationsverfahren	24
1.2.1. ILICA	25
2. Testgütekriterium Reliabilität	28
2.1. Bisherige Ergebnisse ILICA	28
2.2. Bisherige Ergebnisse Selbstwert-IAT	28
3. Testgütekriterium Akzeptanz	30
4. Manipulation von Verhaltensvielfalt	32
EMPIRISCHER TEIL	35
5. Fragestellung der Untersuchung	37
5.1. Hypothesen	37
6. Untersuchungsplan	40
7. Messinstrumente	42
7.1. Selbstwert-IAT	42
7.1.1. Ablauf	42
7.1.2. Verrechnungsmodus	46
7.2. ILICA	48
7.2.1. Skalen	48
7.3. AKZEPT!-L	50
7.3.1. Skalen	51
8. Durchführung der Untersuchung	53
8.1. Beschreibung der Datenerhebung	53
8.2. Instruktionen	54
9. Beschreibung der Stichprobe	56
9.1. Bestimmung des Stichprobenumfangs	56

9.2. Verteilung der TeilnehmerInnen auf die Untersuchungsbedingungen	57
10. Ergebnisse	58
10.1. Mittelwerte und Standardabweichungen der Untersuchungsbedingungen	58
10.2. Retest-Reliabilität IAT	60
10.3. Retest-Reliabilität ILICA	61
10.4. Manipulation von Verhaltensvielfalt ILICA	63
10.4.1. Vergleich der Korrelationskoeffizienten	64
10.5. Akzeptanz	65
11. Interpretation und Diskussion	70
12. Zusammenfassung	74
13. Literaturverzeichnis	78
14. Abbildungsverzeichnis	86
15. Tabellenverzeichnis	87
16. Anhang	88
16.1. Schreiben Landesschulrat	88
16.2. Elternbrief	91
<i>Abstract – Deutsch</i>	93
<i>Abstract – Englisch</i>	94
<i>Lebenslauf</i>	96

EINLEITUNG

In den letzten Jahren erfreuen sich allgemein indirekte Messmethoden und besonders Objektive Persönlichkeitstests wachsender Beliebtheit (Ortner, Horn, Kersting, Kubinger, Schmidt-Atzert, Proyer, Schuhfried, Schütz, Wagner Menghin & Westhoff, 2007). Im Gegensatz zu etablierten Verfahren der psychologischen Persönlichkeitsdiagnostik, in erster Linie Fragebogen, versuchen indirekte Verfahren, wie schon der Name impliziert, nicht durch direktes Nachfragen sondern im Sinne einer „experimentalpsychologischen Verhaltensdiagnostik“ (Kubinger, 2006, S. 50), aus sichtbaren Verhaltensweisen bei unterschiedlichen Arbeitsanforderungen theoriegeleitet Rückschlüsse auf individuelle Persönlichkeitseigenschaften zu ziehen. Ein typisches Charakteristikum dieser Verfahren ist dabei die Unverfälschbarkeit, bedingt durch Verschleierung des Messprinzips.

Die Kritik an herkömmlichen Persönlichkeitsfragebogen richtet sich gegen die Tatsache, dass Messungen vom Vermögen der getesteten Person zur Introspektion abhängig sind (Ortner et al., 2007) und das Problem der prinzipiellen Verfälschbarkeit expliziter Fragetechniken (Kubinger, 2002; Seiwald, 2002). Für Objektive Persönlichkeitstests und auch indirekte Verfahren sollen diese Einschränkungen nicht zutreffen.

Die Bezeichnung indirekte Verfahren umfasst all jene Verfahren, die interessierende Merkmale nicht durch direktes Nachfragen sondern theoriegeleitet auf experimenteller Basis erfassen. Zu den indirekten Verfahren zählen neben Objektiven Persönlichkeitstests auch implizite Messmethoden, die Zugang zu assoziativen Vorgängen schaffen sollen, wobei letztere eindeutig von der Gruppe Objektiver Persönlichkeitstests abzugrenzen sind.

Sowohl ILICA (ILICA; Möseneder & Ebenhöf 1996), ein Objektiver Persönlichkeitstest im engeren Sinne, als auch die Methode des Impliziten Assoziationstask (IAT; Greenwald, McGhee & Schwartz, 1998), ein indirektes Verfahren aus der Experimentalpsychologie, zeichnen sich durch die

Undurchschaubarkeit des Messprinzips aus (Kubinger, 2003). Beide Verfahren beanspruchen für sich die Voraussetzung der Undurchschaubarkeit zu erfüllen.

ILICA (ein Phantasiename eines Urlaubszieles), ist ein interaktives Computersimulationsverfahren. Ziel dieses Tests ist es, die so genannte Selbstverwaltungskompetenz zu erfassen. Darunter wird „die Fähigkeit mittel- und längerfristige Ziele erfolgreich zu erreichen und damit konkurrierende Ablenkungen von außen – soweit nötig – widerstehen zu können.“ (Kubinger, persönliche Mitteilung, März 1990; zit. n. Möseneder, 1992, S. 11) verstanden. Die Simulation zeichnet sich durch eine hohe Realitätsnähe aus, die Aufgaben sind komplex und dynamisch gestaltet.

Ziel des IAT, als implizites Verfahren, ist es, einen Zugang zu automatischen, assoziativen Prozessen zu schaffen. Es handelt sich bei diesem „Test“ nicht um einen spezifischen Test, sondern eher um eine generelle Vorgehensweise, die in verschiedensten Prüfungs- und Testdomänen eingesetzt werden kann, wobei der IAT weniger „im Sinne eines standardisierten Verfahrens zu betrachten ist, sondern eher als eine Messmethode“ (Gawronski & Conrey, 2004, S. 123). Es wird in diesem Zusammenhang daher auch häufig von „Implicit Association Task“ (Gawronski & Conrey, 2004) gesprochen. Ein Grund für seine große Beliebtheit ist mit Sicherheit die Möglichkeit der einfachen Adaptation für verschiedenste Konstrukte. In dieser Arbeit kommt der Selbstwert-IAT (Rudolph, Schröder-Abé & Schütz, 2006) zum Einsatz, er dient der impliziten Selbstwertschätzung.

Vergangene Untersuchungen widmeten sich in erster Linie dem Vergleich derartiger impliziter Messtechniken mit expliziten Verfahren, wie zum Beispiel herkömmlichen Persönlichkeitsfragebogen. Bisher gab es noch kaum Untersuchungen, die unterschiedliche Techniken indirekter Messmethoden miteinander verglichen.

Aufgabe dieser Arbeit soll es daher sein, die Reliabilitäten der beiden Verfahren zu vergleichen. Da es sich in diesem Fall um eine Retest-Reliabilität bei mehrmaliger Messung unmittelbar hintereinander handelt, ist ein zentrales Thema die Stabilität der durch die beiden Tests gemessenen Konstrukte (Kubinger, 2009). In diesem Zusammenhang soll auch eine Manipulation von Verhaltensvielfalt bei Bearbeitung

des Tests ILICA untersucht werden. Des Weiteren soll mit Hilfe eines Akzeptanzfragebogens, dem AKZEPT!-L (Kersting, 2009), festgestellt werden, ob der Modus der Tests die Akzeptanz bei den Testpersonen beeinflusst.

Die Ergebnisse der Untersuchung sollen Aufschluss darüber geben, ob sich IAT und ILICA auch zur Vorgabe als Individualverfahren und nicht nur als Screeningverfahren eignen (Ortner et al., 2007; Herle, 1999).

Der Theorieteil enthält, nach einer Übersicht zu indirekten Verfahren der Persönlichkeitsforschung, Objektiven Persönlichkeitstests, impliziten Verfahren und Simulationsverfahren, den theoretischen Hintergrund der in dieser Untersuchung zum Einsatz gekommenen Verfahren. Im Anschluss wird ein Überblick bisheriger Forschungsergebnisse zu ILICA und Selbstwert - IAT, mit besonderem Schwerpunkt auf Retest-Reliabilität, Akzeptanz und Manipulation von Verhaltensvielfalt gegeben.

Der empirische Teil umfasst die Hypothesen, Untersuchungsplan, Durchführung der Studie, eine genaue Schilderung der Verfahren ILICA, Selbstwert-IAT und AKZEPT!-L und eine umfassende Stichprobenbeschreibung. Anschließend werden die Ergebnisse zunächst dargestellt, darauffolgend wird die zentrale Fragestellung der Untersuchung beantwortet und die Ergebnisse interpretiert und diskutiert.

THEORETISCHER TEIL

1. Indirekte psychologisch-diagnostische Verfahren und Objektive Persönlichkeitstests

Die Idee Objektiver Persönlichkeitstests geht zurück auf Raymond B. Cattell (1957) und Hans-Jürgen Eysenck (1960), sie lieferten als erstes einen systematischen Ansatz zur Messung von Persönlichkeitsunterschieden, welcher neben Selbstbeurteilungsfragebogen auch leistungsbetonte Verfahren enthielt (Pawlik, 2006; Schmidt, 2006). Cattell erhob, im Sinne eines multimodalen Ansatzes, natürliches Verhalten (L-Daten), Selbstbeurteilungen (Q-Daten) und Daten basierend auf objektiven Tests (T-Daten), letztere definiert als standardisierte, übertragbare, Stimulussituationen, in welchen das Verhalten der betreffenden Person gemessen wird, ohne dass dieselbe weiß, was tatsächlich gemessen wird (Cattell, 1957).

Trotz durchaus positiver Ergebnisse gab es fast drei Jahrzehnte kaum neue Publikationen oder Forschungsarbeiten in diese Richtung, erst in den 90er Jahren wurde das Konzept von verschiedenen Forschergruppen wieder aufgenommen und weiterentwickelt (Pawlik, 2006). Diese lange Brachperiode lässt sich möglicherweise dadurch erklären, dass die ursprünglich von Cattell entwickelten objektiven Erhebungsmethoden sich als relativ aufwändig im Messprozedere darstellten (Pawlik, 2006). Durch die Verfügbarkeit des Mediums Computer war es mit Beginn der 90er Jahre möglich, völlig neue Testszenarien zu entwickeln (Ortner et al., 2007).

Zu diesem Zeitpunkt wurde auch die Idee des Impliziten Assoziationstasks geboren (Greenwald, McGhee & Schwartz, 1998). Der IAT zählt gemeinsam mit den Objektiven Persönlichkeitstests zu den sogenannten indirekten, genauer zu den impliziten Verfahren. Im Gegensatz zu expliziten Erhebungstechniken, wie Persönlichkeitsfragebogen, werden bei indirekten Verfahren Persönlichkeitseigenschaften nicht durch direktes Fragen erfasst (Kubinger, 2009). Wobei es zu beachten gilt, dass dieses Verfahren nicht den Objektiven Persönlichkeitstests zuzuordnen ist, handelt es sich hierbei ja um ein Verfahren, das der Experimentalpsychologie entstammt und nicht um ein „standardisiertes Verfahren“ (Gawronski & Conrey, 2004). Insofern soll in dieser Arbeit, entsprechend dem

Vorschlag von Gawronski und Conrey (2004) vom „Impliziten Assoziationstask“ gesprochen werden und nicht vom „Impliziten Assoziationstest“.

Die Bezeichnung Objektive Persönlichkeitstests bezieht sich, wie bereits eingangs erwähnt, auf das Gütekriterium Objektivität. Pawlik (2006, S.18) unterscheidet in diesem Zusammenhang „Erhebungs-, Auswertungs- und Erfassungsobjektivität“.

Unter Erhebungsobjektivität wird die Unabhängigkeit von situationsspezifischen Faktoren der Testdurchführung verstanden. Auswertungsobjektivität subsumiert laut Pawlik die Unabhängigkeit von der Person die die Auswertung vornimmt, verschiedene Untersucher sollten anhand derselben Testdaten auch zu demselben Ergebnis gelangen. Erfassungsobjektivität, oder auch Unverfälschbarkeit, beschreibt, inwiefern die getestete Person auf das Testergebnis Einfluss nehmen kann (sowohl bewusst als auch unbewusst). Objektive Tests beanspruchen für sich, besonders diese dritte Form der Objektivität zu erfüllen, was expliziten Verfahren, wie Persönlichkeitsfragebogen, nicht gelingt (Pawlik, 2006). Wobei insbesondere das so genannte „faking good“ eingeschränkt wird, wohingegen ein „faking bad“, also ein Vortäuschen niedriger Leistungen oder ein Nichteinhalten der Instruktion, auch bei Objektiven Persönlichkeitstests möglich ist (Kubinger, 2006).

Kubinger (2006) argumentiert, dass aufgrund der Missverständlichkeit der Bezeichnung „objektiv“, welche in der Psychologie verschiedenste Bedeutungen zusammenfasst, objektiv in Zusammenhang mit diesen Tests unzulänglich ist. Auch die Bezeichnung „*Persönlichkeitstest*“ ist überholt, da vielmehr persönliche Stilmerkmale erhoben werden. Aufgrund einer Analyse der Gemeinsamkeiten aktueller „Objektiver Persönlichkeitstests“ kommt er zu dem Schluss, dass die Bezeichnung „*experimentalpsychologische Verhaltensdiagnostik*“ treffender ist. Kubinger (2006) definiert: „*experimentalpsychologische Verhaltensdiagnostik* [...] bezieht sich auf Verfahren, die aus dem beobachtbaren Verhalten bei experimentell variierten Leistungsanforderungen persönliche Stilmerkmale erschließen, wobei die Registrierung der Art und Weise der Problembearbeitung der Computer übernimmt.“ (S. 50; Hervorhebungen aus dem Original übernommen).

Sowohl der Implizite Assoziationstest als auch ILICA zählen zu den indirekten Verfahren der Persönlichkeitsmessung. Das Computersimulationsverfahren ILICA ist dabei der Untergruppe Objektiver Persönlichkeitstests zuzuordnen, genauer den Computersimulationstests, wohingegen der IAT, als nicht standardisierter „Task“, nur der Obergruppe indirekter Verfahren angehört, er wird den impliziten Messmethoden zugeteilt (Kubinger, 2009; Gawronski & Conrey, 2004).

1.1. Implizite Verfahren

Implizite Verfahren basieren auf Reaktionszeitmessungen, die Rückschlüsse auf Einstellungen erlauben sollen. Diese impliziten Verfahren haben gemeinsam, dass sie behaupten, die Stärke von Assoziationen zwischen Einstellungsobjekten und Bewertungen im Gedächtnis zu prüfen und dadurch eine Einschätzung der Einstellung zu erlauben, ohne den Teilnehmer direkt nach einer verbalen Einschätzung fragen zu müssen (Degner, Wentura & Rothermund, 2006). Inzwischen gibt es eine breite Palette an angebotenen Verfahren, die hauptsächlich zur Einstellungsforschung verwendet werden. Es handelt sich hierbei nicht, wie bereits erwähnt, um standardisierte Tests, sondern vielmehr um allgemeine Screeningverfahren (Kubinger, 2009). Der wohl bekannteste Vertreter ist der Implizite Assoziationstask (IAT; Greenwald, McGhee & Schwartz, 1998), weitere Verfahren sind (ohne Anspruch auf Vollständigkeit) Affective Priming Task (AP; Fazio, Sanbonmatsu, Powell & Kardes, 1986), Go/No-go Association Task (GNAT; Nosek & Banaji, 2001), Extrinsic Affective Simon Task (EAST; De Houwer, 2003) und Affective Simon Task (AST; De Houwer & Eelen, 1998).

Die große Beliebtheit dieser Verfahren liegt vermutlich an ihrer leichten Modifizierbarkeit für verschiedenste Konstrukte, dementsprechend hoch ist auch die Anzahl der inzwischen zu diesen Verfahren veröffentlichten wissenschaftlichen Artikel (Gawronski, 2006).

1.1.1. Selbstwert-IAT

Der Implizite Assoziationstask wurde im Jahr 1998 von Greenwald, McGhee und Schwartz im *Journal of Personality and Social Psychology* eingeführt. Ziel dieses Verfahrens war es, individuelle Unterschiede in Bezug auf Einstellungen zu messen und dabei bekannte Problematiken expliziter Fragetechniken zu überwinden (Gawronski, 2006). Bestrebungen in diese Richtung gab es bereits in den 60er Jahren, die zu diesem Zeitpunkt entwickelten Verfahren waren aber äußerst unökonomisch und konnten sich nicht durchsetzen (Stieger, 2006).

Der IAT stammt ursprünglich aus der Sozialpsychologie, er gründet auf den Untersuchungen von Greenwald und Banaji (1995) zu „implicit social cognition“. Der Nachweis impliziter Stereotype gelang (Greenwald & Banaji, 1995), in weiterer Folge kam es zur Entwicklung des IAT.

Der IAT wird mittels Computer vorgegeben, im Prinzip handelt es sich um eine einfache Zuordnungsaufgabe (Greenwald, McGhee & Schwartz 1998). Es gibt jeweils ein bipolares Zielkonzept und ein bipolares Attributkonzept. Im Fall des in dieser Arbeit eingesetzten Selbstwert-IAT wären die Zielkonzepte „ich“ vs. „nicht-ich“ und die Attribute „positiv“ vs. „negativ“ (Rudolph, Schröder-Abé & Schütz, 2006). Diesen dichotomen Dimensionen müssen jeweils Stimuli (z.B.: Krieg, Frieden, mich, andere) so schnell wie möglich zugeordnet werden. Der IAT läuft üblicherweise in fünf Phasen ab, wobei jeweils nur die dritte und fünfte Phase, in welcher Ziel- und Attributkonzepte kombiniert werden, zur Berechnung herangezogen werden. Der IAT-Effekt wird aus den Reaktionszeiten berechnet, wobei immer wieder unterschiedliche Berechnungsmodi zum Einsatz kommen (eine Übersicht unterschiedlicher Berechnungsmodi liefern Greenwald, Nosek & Banaji, 2003).

Die in dieser Untersuchung verwendeten Items, Abfolge der Blöcke, Anzahl der Durchgänge in den einzelnen Blöcken sowie der genaue Verrechnungsmodus können dem Kapitel sieben entnommen werden.

Es herrscht bisher keine Einigung darüber, welche kognitiven Prozesse den IAT-Effekt verursachen (Mierke, 2004), daher soll an dieser Stelle ein kurzer Überblick über die

wichtigsten Theorien gegeben werden. Mierke (2004) unterscheidet zwei Hauptströmungen, jene welche sich mit der Frage auseinandersetzen was durch den IAT erfasst wird („Inhaltsmodelle“) und jene, die sich mit der Entstehung des IAT-Effekts beschäftigen ohne auf das erschlossene Merkmal einzugehen („Prozessmodelle“).

Inhaltsmodelle beschäftigen sich in erster Linie mit den „zugrunde liegenden mentalen Repräsentationen“ (Mierke, 2004, S. 35). Zu diesen Modellen zählt unter anderem die „Unified Theory of Social Cognition“ von Greenwald, Banaji, Rudman, Farnham, Nosek und Mellott (2002). Hier wird davon ausgegangen, dass der IAT die (relative) Stärke und die Richtung automatischer Assoziationen misst. Ausgangspunkt ist ein Modell der Darstellung sozialen Wissens, welches verschiedene sozialpsychologische Modelle integriert. Ein weiterer Ansatz der den Inhaltsmodellen zuzurechnen ist, ist die Theorie von Rothermund und Wentura (2001). Die Autoren gehen von einer „Figure-ground“ Asymmetrie aus, Unterschiede in der Salienz zwischen den Kategorien sollen den IAT-Effekt verursachen.

Die Prozessmodelle hingegen beschäftigen sich hauptsächlich mit den Prozessen „die am Zustandekommen des IAT-Effekts beteiligt sind“ (Mierke, 2004, S. 35). Hierzu zählt unter anderem das „Random Walk Modell“ von Brendl, Markman und Messner (2001). Der IAT-Effekt wird als Resultat eines Random Walk Prozesses erklärt, Reaktionen werden immer erst dann gesetzt, wenn ausreichend Information vorhanden ist, eine Antwortschwelle überschritten wird. Die Autoren nehmen an, dass im inkompatiblen Block, aufgrund der widersprüchlichen Information, die Reaktionszeiten höher als im kompatiblen Block sind. Der IAT-Effekt ist insofern Ergebnis unterschiedlicher Reaktionsschwellen.

De Houwer (2001) erklärt den IAT-Effekt als einen Stimulus-Response Kompatibilitäts-Mechanismus („stimulus-response compatibility mechanism“). In diesem Fall wird der IAT-Effekt als Resultat der Übereinstimmung zwischen Stimuli und Kategorienzuordnung gesehen, entscheidend ist also nicht das Attribut (z.B.: positiv oder negativ) der Stimuluswörter, sondern die Möglichkeit der klaren Zuteilung zu den jeweils vorgegebenen Kategoriebegriffen (z.B.: ich und nicht-ich).

Zum jetzigen Zeitpunkt ist eine schlüssige Prüfung dieses Modells allerdings ausständig.

Mierke und Klauer (2001) führten in ihrem Aufgabenwechselmodell den IAT-Effekt auf die alternierende Aufgabenstellung zurück:

„The IAT effect, according to this model, reflects general costs for switching between two different task sets that specifically affect the incompatible IAT condition.” (Mierke & Klauer, 2001, S. 112).

Wie bereits erwähnt wird der IAT zur Messung unterschiedlichster Konstrukte verwendet. Die Anwendungsgebiete erstrecken sich über verschiedene Disziplinen der Psychologie und umfassen unterschiedlichste Einsatzbereiche. So wird beispielsweise in der Persönlichkeitsforschung von Asendorpf, Banse und Mücke (2002) ein Schüchternheits-IAT eingesetzt, in der Vorurteilsforschung der Sozialpsychologie wurde unter anderem ein Homosexualitäts-IAT (Banse, Seise & Zerbes 2001) sowie ein IAT zum Thema AIDS (Neumann, Hulsenbeck & Seibt, 2004) entwickelt, in der Klinischen Psychologie gab es von Egloff und Schmukle (2002) Forschungen zu einem Angst-IAT. Eine umfassendere Übersicht liefert Stieger (2006, S. 16).

Diese Diplomarbeit widmet sich dem von Rudolph, Schröder-Abé und Schütz (2006) entwickelten Selbstwert-IAT. In seiner Konzeption orientiert er sich an der ursprünglichen Version von Greenwald et al. (1998).

Selbstwertschätzung, im Sinne der affektiven Bewertung der eigenen Person, wird häufig auf direkte Art und Weise, mit Hilfe von Fragebogen erfasst (Rudolph, et al, 2006). Die so gewonnenen Daten sind allerdings nur beschränkt valide. Einflussfaktoren, wie zum Beispiel „impression management“ (Schlenker & Weigold, 1992) und „self-deception“ (Paulhus, 1984) verfälschen die Ergebnisse. Diesen Mangel versucht man mit Hilfe des Selbstwert-IATs zu umgehen, dabei ist allerdings zu beachten, dass der implizite Selbstwert nicht unbedingt eine der Selbstbeobachtung zugängliche Haltung reflektiert.

Greenwald und Banaji (1995) fassen implizite Selbstwertschätzung folgendermaßen zusammen:

„*Implicit self-esteem* is the introspectively unidentified (or inaccurately identified) effect of the self-attitude on evaluation of self-associated and self-dissociated objects.” (S. 11).

Der Selbstwert-IAT beruht auf zwei Annahmen:

„(a) Da Menschen Objekte oder Stimuli, die mit ihnen assoziiert sind, im Durchschnitt positiver bewerten als nicht-selbstrelevante Stimuli, kann aus der Bewertung selbstbezogener Stimuli indirekt auf die Selbstbewertung geschlossen werden. [...] Und (b) bei der Präsentation von Einstellungsobjekten werden die damit verbundenen Evaluationen automatisch aktiviert. Somit können aus den Reaktionen auf evaluative und selbstbezogene Stimuli Rückschlüsse auf Selbstbewertungen gezogen werden.“
(Rudolph, Schröder & Schütz, 2006, S. 154).

Hier ist anzuführen, dass die Untersuchung impliziter Selbstwertschätzung und damit impliziter Assoziationen, nicht mit den von Freud (1933, zit. n. Rudolph, 2009) postulierten „unbewussten Assoziationen“ zu verwechseln ist. Impliziter Selbstwert wird in diesem Zusammenhang als „die automatische Bewertung der eigenen Person verstanden“ (Rudolph, 2009, S. 5). Unbewusste Vorgänge können insofern Teil des impliziten Selbstwertes sein, sind allerdings nicht prinzipielle Prämisse.

Da es sich bei den Impliziten Assoziations-„Tasks“ nicht um einheitliche, standardisierte Verfahren handelt, sondern um eine allgemeine Messprozedur und Ergebnisse bezüglich Reliabilität und Validität stark von den jeweils eingesetzten Objekt- und Attributdimensionen abhängen, sowie von den zuzuordnenden Stimuli (Gawronski, 2006), soll hier in erster Linie auf Ergebnisse zum Selbstwert-IAT eingegangen werden.

Bosson, Swenn und Pennebaker (2000) verglichen in einer umfassenden Untersuchung unterschiedlichste indirekte Verfahren, die den impliziten Selbstwert erfassen, die gefundenen Zusammenhänge waren eher gering. Auch Zusammenhänge zwischen expliziter Selbstwertschätzung und dem IAT konnten in verschiedenen Untersuchungen nur in geringem Ausmaß nachgewiesen werden (siehe z.B.: Bosson et al., 2000 und Greenwald & Farnham, 2000). Kritisch zu betrachten ist in diesem

Zusammenhang mit Sicherheit, dass, abhängig von der Höhe des gefundenen Zusammenhanges, die Korrelationen zwischen expliziten und impliziten Maßen entweder als Belege zur diskriminanten (bei geringem Zusammenhang; Gawronski & Conrey, 2004) oder als Belege für die konvergente Validität (bei starkem Zusammenhang; Greenwald et al., 2003) gewertet werden. Bisher herrscht keine Einigung bezüglich einer theoretischen Fundierung, die festlegt, wie hoch ein derartiger Zusammenhang sein sollte (Rudolph, Schröder-Abé & Schütz, 2008). In einer Studie von Rudolph Schröder-Abé, Schütz, Gregg, und Sedikides (2008), die ebenfalls die Zusammenhänge zwischen verschiedenen impliziten Verfahren zur Erfassung des Selbstwertes untersuchte, konnte man ebenfalls nur geringe Korrelationen zwischen unterschiedlichen Verfahren feststellen. Es konnte allerdings herausgearbeitet werden, dass einander in Konzept und Struktur ähnelnde Verfahren stärker zusammenhängen als Verfahren, die sich in diesen Punkten unterschieden. Die Autoren schließen daraus, dass die schlechten Befunde zur konvergenten Validität von der Strukturierung der impliziten Messtechniken abhängig sind.

Bezüglich der prädiktiven Validität des Selbstwert-IATs gab es eine umfassende Untersuchung von Rudolph, Schröder, Riketta und Schütz (2010). Ausgangspunkt der Studie war ein duales Prozessmodell, es werden zwei unterschiedliche Systeme der Informationsverarbeitung angenommen. Es wird zwischen dem impulsiven System und dem reflexiven System unterschieden, dementsprechend erfassen direkte und indirekte Messungen auch jeweils nur eines dieser informationsverarbeitenden Systeme und sagen unterschiedliche Verhaltensweisen voraus. Direkte Messungen, die Informationen aus dem reflexiven System erfassen, sollen insofern bewusst kontrollierte Verhaltensweisen vorhersagen, indirekte Messungen, die Informationen aus dem impulsiven System erfassen, sollen automatische und spontane Verhaltensweisen vorhersagen. Die Autoren konnten nachweisen, dass der IAT spontane, selbstbewusste Verhaltensweisen vorhersagen konnte ($r = .29$), die nicht durch die expliziten Messungen vorhergesagt werden konnten, wohingegen durch direkte Fragebogentechnik kontrolliertes selbstbewusstes Verhalten vorhergesagt werden konnte ($r = .35$), was durch die Ergebnisse des IAT nicht möglich war.

Da es sich bei dem Selbstwert-IAT um einen Computertest handelt, kann in Bezug auf die Objektivität davon ausgegangen werden, dass Testleiterunabhängigkeit und

Verrechnungsfairness besteht. Auch Ökonomie für die Dauer der Auswertung ist gegeben. Die Testdauer von ungefähr 15 Minuten ist durchaus ökonomisch.

Fiedler und Blümke (2005) konnten nachweisen, dass IATs zu einem geringen Grad verfälschbar sind. Allerdings muss festgehalten werden, dass die Manipulation durch die Testpersonen nicht so weit greift wie bei direkten Fragebogenverfahren. So konnte beispielsweise Steffens (2004) in einem Gewissenhaftigkeits- („conscientiousness“) IAT keine signifikanten Verfälschungseffekte nachweisen und auch Asendorpf et al. (2002) konnten in einem Schüchternheits- („shyness“) IAT keinen signifikanten Nachweis für eine Manipulation erbringen. Asendorpf et al. (2002) kommen zu dem Schluss, dass eine Verfälschung des IATs nur möglich ist, wenn das Messprinzip durchschaut wird, durch bewusste Verlangsamung der Reaktionszeiten in einem der Blöcke.

Bisher gibt es diesbezüglich keine Ergebnisse zum Selbstwert-IAT im Speziellen, dennoch kann aufgrund der Vielzahl an Belegen aus der Literatur für unterschiedliche IATs, angenommen werden, dass auch bei diesem impliziten Verfahren leichte Verfälschungen möglich sind.

Ergebnisse zu den Gütekriterien Reliabilität und Akzeptanz sind in den Kapiteln zwei und drei genauer abgehandelt.

Kritik, die immer wieder am IAT geübt wird, richtet sich gegen sein prinzipielles Konzept. So ist der erhaltene IAT-Wert (*D*-Wert) immer Ausdruck einer Beziehung, er bleibt immer ein relatives Maß, da der IAT mit der Zuordnung zu zwei Kategorien arbeitet (Rudolph et al., 2006). Der hier vorliegende IAT versucht dieses Problem zu überwinden, die Kategorien sind „unipolar“ (Rudolph et al., 2006). So wird im Fall des Selbstwert-IAT der Kategorie „ich“ die neutrale Kategorie „nicht-ich“ gegenüber gestellt. Dennoch kann ein hoher Wert auf drei unterschiedliche Arten gedeutet werden. Es kann sich (bezogen auf den Fall des Selbstwert-IATs) um eine Bevorzugung der „ich“ Kategorie handeln, oder aber um eine schlechte Bewertung der „nicht-ich“ Kategorie, beziehungsweise kann es auch zu einer Wechselwirkung beider Einflüsse kommen (Rudolph, 2009).

Kritisch zu betrachten ist die Tatsache, dass trotz der Vielzahl an vorhandenen Modellen sich bisher keine der Theorien zur Ursache des IAT-Effekts durchsetzen konnte. Grundsätzliche Voraussetzung, um festzustellen wie valide dieses Instrument misst, wäre eine theoretische Fundierung was erfasst werden soll und wie es zu dem IAT-Effekt kommt (Mierke, 2004).

Weitere Punkte, die immer wieder kritisiert werden, sind die Verfälschbarkeit des IAT, der Einfluss von Umgebungsfaktoren, die Reduktion des IAT-Effekts bei wiederholter Vorgabe und die Beeinflussung des D-Wertes durch die Reihung der Blockkombinationen (Stieger, 2006).

Dennoch gehört der IAT zu den impliziten Verfahren mit den vorteilhaftesten Qualitäten bezüglich diverser Gütekriterien (Rudolph, 2009).

1.2. Simulationsverfahren

Simulationsverfahren zeichnen sich durch eine hohe Realitätsnähe aus, aufgrund dessen ist zu erwarten, dass die „prognostische Validität“ entsprechend hoch ist (Möseneder & Ebenhöf, 1996). Die Durchführung derartiger Verfahren erfolgt meist am Computer. Es besteht eine enge Verwandtschaft zu den „Arbeitsproben“, die Durchführung mag an Computerspiele erinnern (Kubinger, 2009). „Eine gewisse Ähnlichkeit besteht auch zu den bereits als klassisch zu bezeichnenden komplexen Problemlösungssituationen *à la* Lohhausen [...], nur dass bei Computer-Simulationen im Rahmen der Psychologischen Diagnostik weniger intelligenzbezogene als vielmehr organisatorische Fähigkeiten abgeprüft werden.“ (Kubinger, 2009, S. 150).

Ein Austausch zwischen der Person, die den Test bearbeitet, und dem Computer findet statt, auf die jeweiligen Aktionen wird vom Computer unterschiedlich reagiert. Problematisch ist dabei häufig die Auswertung, bisweilen fehlt der Bezug zwischen von Personen gesetzten Aktionen und dem resultierenden Testwert. Die Zahl der derzeit erhältlichen Simulationsverfahren ist entsprechend gering. Häufig war die Verrechnung der Testwerte nicht ausgereift, bei vielen Verfahren waren die Reliabilitäten niedrig und die Normierung nicht exakt (Kubinger, 2009).

1.2.1. ILICA

ILICA ist ein interaktives Computersimulationsverfahren. Aufgabe ist es einen freien Tag vor Antritt eines Urlaubes zu planen. Ziel dieses Tests ist es, wie bereits in der Einleitung erwähnt, die so genannte Selbstverwaltungskompetenz zu erfassen. Wobei diese Fähigkeit hier als ein Persönlichkeitsmerkmal verstanden wird, nämlich die Kompetenz „mit Ablenkungen während einer angestrebten Zielerreichung, mit aktualisierten Problemen und Frustrationen umzugehen.“ (Mösener & Ebenhö, 1996, S. 5).

Testwerte werden aus den bei Bewältigung der Aufgaben eingesetzten und vom Computer registrierten Verhaltensweisen errechnet. Einerseits werden beispielsweise Anzahl der erfüllten Aufgaben oder Inanspruchnahme der vom Programm zur Verfügung gestellten Hilfen errechnet. Mit Hilfe dieser Daten sind Rückschlüsse auf individuelle Stile bei Problembearbeitung möglich. Zusätzlich erfolgt, abgeleitet vom Konzept des kognitiven Stils nach Amelang und Bartussek (1990), eine Typenzuordnung: „Reflexiver Typ“, „Impulsiver Typ“ und „Flexibler Typ“.

Während der Bearbeitung des Tests sind immer wieder unmittelbare Entscheidungen zu treffen, es stehen jeweils drei Antwortmöglichkeiten zur Verfügung, die jeweils den unterschiedlichen Personentypen zugeordnet sind. Aufgrund der gewählten Antwort erfolgt eine Typisierung (Mösener & Ebenhö, 1996).

Zur Skalierung von ILICA ist anzumerken, dass einige Werte nicht testtheoretisch begründet sind, ILICA wird daher mangelnde Verrechnungsfairness vorgeworfen (Kubinger, 1993; zit. n. Herle, 1999; S. 48). Wie bereits erwähnt, ist aufgrund der Dynamik, die derartige Simulationsverfahren kennzeichnet, eine Berechnung von eindeutigen Testkennwerten unter Umständen problematisch (Kubinger, 2009).

Diskriminante Validität konnte für den „Zahlen-Verbindungstest“ (ZVT; Oswald & Roth 1987; zit. n. Funke, 1999) nachgewiesen werden. Es besteht kein signifikanter Zusammenhang mit dem Konstrukt „Informationsverarbeitungsgeschwindigkeit“. Maslo (1994) konnte mit dem Test „Rechnen in Symbolen“ (Schmotzer, Kubinger &

Maryschka, 1994; zit. n. Maslo, 1994), der „Reasoning“ erfasst, keine signifikanten Korrelationen feststellen.

Die konvergente Validität war in den durchgeführten Untersuchungen eher gering bis nicht vorhanden. ILICA korrelierte niedrig, aber signifikant mit Skalen des SVF (Stressverarbeitungsfragebogen; Janke, Erdmann & Kallus, 1985; zit. n. Funke 1999), des IPC (IPC-Fragebogen zu Kontrollüberzeugungen; Krampen, 1981; zit. n. Funke 1999), des FSKN (Frankfurter Selbstkonzeptskalen; Deusinger, 1986; zit. n. Funke, 1999) und des MBTI (Myers-Briggs Typenindikator; Bents und Blank, 1991; zit. n. Maslo, 1994).

Bezüglich der prognostischen Validität, beziehungsweise einer Kriteriumsvalidität, wurden zwei Diskriminanzanalysen an erfolgreichen und weniger erfolgreichen PsychologiestudentInnen durchgeführt. Möseneder und Ebenhöf (1996) kommen zu dem Schluss:

„Insgesamt bewertet scheinen zwar eher gruppenspezifisch verschiedene Testkennwerte von ILICA diskriminativ in Bezug auf das Kriterium „erfolgreich“; dieses wird vermutlich auch nicht einheitlich definiert sein. Immerhin zeigen diese Ergebnisse, dass mit dem Test ILICA relevantes Verhalten erfasst werden kann. U.U. empfiehlt sich allerdings lediglich die Beschreibung des (typischen) Verhaltens einer Tp [Testperson] in Relation zu anderen.“ (S. 17).

Da es sich bei ILICA um einen Computertest handelt, kann in Bezug auf die Objektivität davon ausgegangen werden, dass Testleiterunabhängigkeit und Verrechnungsfairness besteht. Interpretationseindeutigkeit kann jedoch nicht sicher garantiert werden, da die Anweisungen im Manual eher spärlich sind (Funke, 1999).

Aufgrund der Tatsache, dass ILICA ein Computertest ist, kann Ökonomie für die Dauer der Auswertung angenommen werden. Auch die Testdauer kann mit 30 Minuten (ohne Instruktion) als ökonomisch bewertet werden (Möseneder & Ebenhöf, 1996).

Die etwas veraltete Normstichprobe ist mit Sicherheit problematisch, hinzu kommt, dass diese Normwerte nur als „vorläufig“ zu betrachten sind, da es sich um eine nicht repräsentative Stichprobe handelt (Möseneder & Ebenhöf, 1996).

Fairness bezüglich der nicht bestehenden Benachteiligung bestimmter Testpersonen wird von den Autoren als gegeben betrachtet. Allerdings wird die Fairness einiger Testkennwerte diskutiert (Funke, 1999). Funke (1999) kritisiert in diesem Zusammenhang auch die nicht eindeutige Instruktion.

Von den Testautoren wird angenommen, dass aufgrund der geringen „Augenscheinvalidität“ (Möseneder & Ebenhöf, 1996; S. 17) eine Manipulation der Testergebnisse nicht möglich ist, und insofern Unverfälschbarkeit gegeben ist. Bisher gibt es noch keine Untersuchung, die diese Voraussetzung auch empirisch prüft.

Ergebnisse zu den Gütekriterien Reliabilität und Akzeptanz sind in den Kapiteln zwei und drei genauer abgehandelt.

Der genaue Ablauf dieses Verfahrens kann dem Kapitel sieben entnommen werden.

2. Testgütekriterium Reliabilität

Die Reliabilität eines Verfahrens kann auf unterschiedliche Art und Weise erhoben werden, man differenziert in der klassischen Testtheorie Retest-Reliabilität, Paralleltest-Reliabilität und interne Konsistenz. Sie bezeichnet das Ausmaß an Präzision, mit der das jeweilige Verfahren eine Eigenschaft erfasst (Kubinger, 2009). Für die hier vorliegende Arbeit sind insbesondere bisherige Ergebnisse zur Retest-Reliabilität von Bedeutung.

2.1.Bisherige Ergebnisse ILICA

Problematisch bei der Bestimmung der Retest-Reliabilität bei Computersimulationen sind insbesondere Lerneffekte. Bei der Untersuchung der Paralleltest-Reliabilität scheitert die Durchführung meist an passenden Paralleltests. Studien zur internen Konsistenz scheitern daran, dass eine Teilung der Simulation häufig nicht möglich oder zielführend ist (Herle, 1999).

Herle (1999) untersuchte die Retest-Reliabilitäten. Diese lagen zwischen .28 und .66 für die einzelnen Skalen. Herle (1999) konnte Lerneffekte, im Sinne einer Verbesserung der Aufgabenerfüllung bei einem wiederholten Testdurchgang, feststellen. Die Retest-Reliabilität scheint sowohl durch Lerneffekte, als auch durch situative Faktoren beeinflusst zu werden (Herle, 1999).

2.2.Bisherige Ergebnisse Selbstwert-IAT

Die Ergebnisse für den Selbstwert-IAT sind insgesamt betrachtet für interne Konsistenzen durchwegs recht gut, wohingegen die Befunde zur Retest-Reliabilität weit mehr streuen.

Hofman, Gawronski, Gschwendner, Le und Schmitt (2005) berichten für verschiedene IATs eine durchschnittliche Retest-Reliabilität von .51 und eine interne Konsistenz von durchschnittlich .79.

Bosson, Swenn und Pennebaker (2000) widmeten sich verschiedenen Impliziten Assoziationstasks zur Messung des Selbstwertes, sie konnten eine Retest-Reliabilität, bei einem Zeitraum von vier Wochen zwischen den beiden Testzeitpunkten von .69 nachweisen.

In einer Untersuchung von Rudolph, Schröder-Abé und Schütz (2008) wurde für den Selbstwert-IAT ein Korrelationskoeffizient von .54 bei Testungen mit einer Woche Abstand gefunden. Die internen Konsistenzen lagen bei derselben Untersuchung zwischen .80 und .85. In einer weiteren Untersuchung von Rudolph et al. (2010) wurde eine interne Konsistenz von .78 nachgewiesen.

Die verhältnismäßig schlechten Ergebnisse bezüglich der Retest-Reliabilität werden auf unterschiedliche Ursachen zurückgeführt. Bosson et al. (2000) gehen davon aus, dass die Anzahl an Trials die Retest-Reliabilität maßgeblich beeinflusst, so sollen längere IATs bessere Ergebnisse bezüglich dieses Gütekriteriums aufweisen. Da immer wieder spekuliert wird, dass situative Faktoren ebenfalls die Retest-Reliabilität beeinflussen (Bosson et al., 2000, bzw. Gawronski, 2006), sollte anhand dieser Untersuchung, in welcher der Zeitabstand so gering wie möglich gehalten wird und damit auch die Situationsfaktoren weitgehend stabil sind, es möglich sein, einen Beitrag zu dieser Frage zu leisten.

3. Testgütekriterium Akzeptanz

Das Qualitätskriterium Akzeptanz ist prinzipiell vom Kriterium der Zumutbarkeit abzugrenzen, während Akzeptanz die subjektive Annahme eines Verfahrens durch die Testperson beschreibt (Testkuratorium, 1986, zit. n. Kubinger, 2009), zielt die Zumutbarkeit auf „eine objektive Betrachtungsweise im Sinne ethischer Richtlinien psychologischen Diagnostizierens“ ab (Kubinger, 2009; S. 120).

Akzeptanz kann durch Aufklärung der Testperson über Grund und Ziel der Testung, sowie über Rahmenbedingungen, wie voraussichtlich in Anspruch genommene Zeit, Folgen der Testung und zu bearbeitende Aufgaben (Kubinger, 2009) erzeugt werden.

Bei objektiven Verfahren wird das Messprinzip nicht offen gelegt, insofern ist es fraglich, ob diesbezüglich ethische Grundprinzipien erfüllt werden. Ortner et al. (2007) führen an, dass auch bei objektiven Tests eine Aufklärung bezüglich oben erwähnter Kriterien möglich ist, der genaue Verrechnungsmodus bleibt in diesem Fall dennoch verborgen, das Gütekriterium Unverfälschbarkeit würde also nicht verletzt werden.

Es gibt nur sehr wenige Untersuchungen, die sich mit der Akzeptanz psychologischer Tests beschäftigen. Ortner und Janous (2006) beschäftigten sich mit der Akzeptanz der Belastbarkeits-Assessment: computerisierten Objektiven Persönlichkeits-Testbatterie Deutsch (BAcO-D; Ortner, Kubinger, Schrott, Radinger, & Litzenberger, 2006), dieses Verfahren zählt ebenfalls zu den Objektiven Persönlichkeitstests. Leistungsaufgaben sind bei diesem Test unter hinderlichen Bedingungen zu bearbeiten, mit dem Ziel „emotionale Belastbarkeit“ zu erfassen. Die Ergebnisse bezüglich der Akzeptanz dieses Verfahrens waren positiv. Wahrscheinlich werden besonders Verfahren mit hohem Realitätscharakter und Aufgabenstellungen im Leistungsbereich gut akzeptiert (Ortner et al., 2007).

Bisher gibt es weder für den Impliziten Assoziationstest, noch für die Computersimulation ILICA Ergebnisse bezüglich dieses Gütekriteriums. Für das Verfahren ILICA trifft die Voraussetzung der Realitätsnähe zu, für den Selbstwert-IAT nicht, allerdings basiert auch dieser Test auf einer eindeutigen Leistungskomponente (möglichst schnelle, richtige Reaktionen bzw. Drücken der

korrekten Antworttaste). Die Erhebung von Akzeptanz mit Hilfe des AKZEPT-L (Kersting, 2009) in dieser Arbeit soll Informationen über die Akzeptanz dieser beider Verfahren liefern.

4. Manipulation von Verhaltensvielfalt

Da es sich bei ILICA um einen Objektiven Persönlichkeitstest mit stark interaktivem Charakter handelt, ist es wahrscheinlich, dass einige Testpersonen bei einem zweiten Testdurchgang ihr Verhalten variieren. Da das geschilderte Szenario des ILICA sowie das eigene Verhalten noch bekannt sind, könnte Veranlassung bestehen, das eigene Verhalten beim zweiten Testdurchgang anders zu gestalten. Da die Ziele ja weiterhin vorgegeben sind und anzunehmen ist, dass die Testpersonen nach wie vor möglichst gut in dem vorgegebenen Test abschneiden wollen, im Sinne alle Aufgaben erfüllen zu wollen, würde dies wahrscheinlich insbesondere die Skalen reflexiver, impulsiver und flexibler Typ, sowie die alternativ durchführbaren Aktivitäten betreffen. Auch die Testautoren gehen davon aus, dass „im inhaltlichen Zusammenhang die Wiederholung vermutlich ein alternatives Handeln geradezu provoziert.“ (Möseneder & Ebenhöf, 1996, S. 15).

Winter und Stewart (1977) führten eine Untersuchung für ein indirektes Verfahren, den Thematischen Apperzeptionstest (TAT; Original: Thematic Apperception Test; Murray, 1971), durch, die auf einer ähnlichen Annahme beruht.

Der TAT zählt zu den projektiven Verfahren. Beim TAT werden Testpersonen Bilder in schwarz-weiß vorgelegt, auf welchen größtenteils Personen in verschiedensten Situationen abgebildet sind. Aufgabe der TeilnehmerInnen ist es, zu jedem Bild eine Geschichte zu erzählen. Inzwischen haben sich unterschiedlichste Auswertungstechniken etabliert. Aufgrund der erzählten Geschichten, deren Handlungen und Stimmung, wird auf zugrunde liegende Motive geschlossen (Kubinger, 2009).

Winter und Stewart teilten in ihrer Untersuchung TeilnehmerInnen in drei Gruppen ein, die Autoren konnten bei einer Testwiederholung nach einer Woche, bei Aufforderung das Verhalten der letzten Testsituation zu erinnern und möglichst identische Geschichten zu erzählen, eine Retest-Reliabilität von .61 feststellen, bei Instruktion sich um die in der vorherigen Testsituation geschilderten Geschichten nicht zu kümmern, einen Korrelationskoeffizienten von .58 und bei Aufforderung möglichst andersartige Geschichten zu beschreiben, eine Retest-Reliabilität von .27 feststellen.

Diese Arbeit soll, durch eine ähnliche Manipulation der Verhaltensvielfalt einen anwendungsbezogenen Beitrag zur Klärung dieser Frage leisten.¹

¹ Umsetzung und hier verwendete Instruktion können den Kapiteln 5.1 und 8.2 entnommen werden.

EMPIRISCHER TEIL

5. Fragestellung der Untersuchung

Die hier vorliegende Arbeit hat es sich zur Aufgabe gestellt, einen Beitrag zu der Antwort auf die Frage zu leisten, ob sich der IAT nicht nur als allgemeines Screeninginstrument, sondern auch als Individualverfahren eignet. Zu diesem Zweck werden Gütekriterien des Selbstwert-IAT erhoben und mit jenen von ILICA verglichen. Die Ergebnisse zu Retest-Reliabilität und Akzeptanz dieser beiden Objektiven Persönlichkeitstest, die jedoch konzeptionell sehr unterschiedlich sind, sollen zur Klärung dieser Frage beitragen. Zusätzlich wurde die Verfälschbarkeit von ILICA untersucht.

5.1.Hypothesen

Die erste Hypothese soll die Stabilität von Selbstwert-IAT und ILICA abklären. Die Nullhypothese und Alternativhypothese werden entsprechend der Vorgangsweise, welche von Kubinger, Rasch und Simeckova (2007) vorgeschlagen wurde, formuliert, ein Ergebnis wird erst dann signifikant, wenn es den Wert von .70 überschreitet.

Nullhypothese Ia: Der Korrelationskoeffizient zwischen erstem und zweitem Testergebnis des Selbstwert-IAT ist kleiner als .70.

$$H(Ia)0: 0 < \rho \leq 0,7$$

Alternativhypothese Ia: Der Korrelationskoeffizient zwischen erstem und zweitem Testergebnis des Selbstwert-IAT ist signifikant und höher als .70.

$$H(Ia)1: \rho > 0,7$$

Nullhypothese Ib: Der Korrelationskoeffizient zwischen erstem und zweitem Testergebnis von ILICA ist kleiner als .70.

$$H(Ib)0: 0 < \rho \leq 0,7$$

Alternativhypothese Ib: Der Korrelationskoeffizient zwischen erstem und zweitem Testergebnis von ILICA ist signifikant und höher als .70.

H(Ib)1: $\rho > 0,7$

Herle (1999) konnte in der von ihr durchgeführten Untersuchung für einige Skalen von ILICA nur eine wenig zufriedenstellende Resttest-Reliabilität feststellen. Die von der Autorin festgestellten Lerneffekte boten keine hinreichende Erklärung für die aufgetretenen Abweichungen. Es besteht die Vermutung, dass beim zweiten Testdurchgang einige Testpersonen ihr Verhalten variieren. Durch eine gezielte Instruktion soll in dieser Untersuchung eine Hälfte der Testpersonen dazu aufgefordert werden, derartige Verhaltensänderungen durchzuführen, wohingegen der verbleibende Teil keine diesbezügliche Aufforderung erhält (die Instruktion zur Manipulation von Verhaltensvielfalt kann dem Kapitel 8.2. entnommen werden).

Nullhypothese II: Die Gruppe, die eine Instruktion zur Manipulation von Verhaltensvielfalt erhalten hat, unterscheidet sich in ihren Testergebnissen zu Messzeitpunkt zwei nicht signifikant von der Gruppe, die keine Verfälschungsinstruktion erhalten hat.

H(II)0 $\mu_1 = \mu_2$

Alternativhypothese II: Die Gruppe, die eine Instruktion zur Manipulation von Verhaltensvielfalt erhalten hat, unterscheidet sich in ihren Testergebnissen zu Messzeitpunkt zwei signifikant von der Gruppe, die keine Verfälschungsinstruktion erhalten hat.

H(II)1: $\mu_1 \neq \mu_2$

Bezüglich der Akzeptanz von ILICA und Selbstwert-IAT gibt es noch keine Ergebnisse. Es soll ein wichtiger Beitrag zur Akzeptanz Objektiver Persönlichkeits-tests geleistet werden, durch die Abklärung der Akzeptanz der beiden Verfahren. Die

beiden Verfahren sollen bezüglich der einzelnen Skalen des AKZEPT!-L verglichen werden. Da es bisher noch keine Ergebnisse bezüglich des Vergleichs von Computersimulationen und indirekten Verfahren bezüglich dieses Kriteriums gibt, wird zweiseitig formuliert.

Nullhypothese III: Es gibt keinen signifikanten Unterschied zwischen ILICA und IAT in den Ergebnissen der Skalen „Kontrollierbarkeit“, „Messqualität“, „Augenscheinvalidität“, „Belastungsfreiheit“, „Gesamtbeurteilung“ und „Verfahrensbezogene Selbsteinschätzung“.

$$H(II)0: \mu_1 = \mu_2$$

Alternativhypothese III: : Es gibt einen signifikanten Unterschied zwischen ILICA und IAT in den Ergebnissen der Skalen „Kontrollierbarkeit“, „Messqualität“, „Augenscheinvalidität“, „Belastungsfreiheit“, „Gesamtbeurteilung“ und „Verfahrensbezogene Selbsteinschätzung“.

$$H(II)1: \mu_1 \neq \mu_2$$

6. Untersuchungsplan

Die Vorgabe der Tests erfolgte unmittelbar hintereinander, wobei jedes Verfahren zweimal vorgegeben wurde. Dabei wurde die Reihenfolge der Vorgabe (zuerst ILICA oder IAT) zufällig erteilt, um sicherzustellen, dass keine Verzerrung aufgrund von Ermüdungs- oder Reihungseffekten eintreten würde. Nach der ersten Vorgabe jedes Tests wurde jeweils einmal der Fragebogen AKZEPT!-L vorgegeben, da durch eine Vorgabe erst nach zweimaliger Testung Reaktanzeffekte verstärkt eintreten könnten.

Zusätzlich wurde eine Einteilung in zwei weitere Untersuchungsgruppen durchgeführt, eine Gruppe erhielt bei der zweiten Vorgabe des ILICA die Instruktion ihr ursprüngliches Konzept zu verwerfen und zu versuchen, neue Strategien bei der Bearbeitung anzuwenden (die Instruktion wird in Abschnitt acht genauer beschrieben), während die andere Gruppe nur die Information erhielt, dass nun ein zweiter Testdurchgang startet. Ziel dieser Manipulation war es, Verhaltensvielfalt zu erzeugen. Auch hier erfolgte die Zuordnung zu den Gruppen randomisiert, um statistische Fehler aufgrund einer Gruppenzugehörigkeit zu vermeiden.

Die Zuweisung zu den vier Gruppen erfolgte per Los, es wurden vor Testungsbeginn Zettel gezogen, auf welchen jeweils eine dreistellige Zahl zu finden war. Die Zuteilung zu den einzelnen Untersuchungsgruppen erfolgte aufgrund der ersten Ziffer der dreistelligen Zahl, Personen mit einer Zahl zwischen 100 und 199 wurden der Gruppe eins zugewiesen, Personen mit einer Zahl zwischen 200 und 299 wurden der Gruppe zwei zugeteilt, mit einer Zahl zwischen 300 und 399 gehörte man der Gruppe drei an und mit einer Zahl zwischen 400 und 499 wurde der Teilnehmer der Gruppe vier zugeteilt. Die Einteilung und Verfahrensabfolge der einzelnen Gruppen kann der Tabelle eins entnommen werden.

Durch diese Zuteilung konnte zugleich Anonymität gewahrt werden und trotzdem waren die einzelnen zusammengehörigen Testergebnisse durch die Nummer eindeutig zuordenbar.

Tabelle 1: Einteilung der Gruppen

Gruppe 1 (Nr. 100-199)	IAT	AKZEPT!-L	ILICA	AKZEPT!-L	IAT	ILICA
Gruppe 2 (Nr. 200- 299)	IAT	AKZEPT!-L	ILICA	AKZEPT!-L	IAT	ILICA + M
Gruppe 3 (Nr. 300-399)	ILICA	AKZEPT!-L	IAT	AKZEPT!-L	ILICA	IAT
Gruppe 4 (Nr. 400-499)	ILICA	AKZEPT!-L	IAT	AKZEPT!-L	ILICA + M	IAT

Note: ILICA + M bedeutet ILICA mit Manipulation von Verhaltensvielfalt, die Untersuchungsgruppen mit Manipulation sind jeweils grau unterlegt.

7. Messinstrumente

In dieser Arbeit kommen der Selbstwert-IAT, ILICA und der Akzeptanzfragebogen AKZEPT!-L zum Einsatz, im Folgenden soll nochmals der genaue Ablauf der einzelnen Verfahren, die Items und die Skalen detailliert dargestellt werden.

7.1.Selbstwert-IAT

Der Selbstwert-IAT ist ein IAT im klassischen Sinn, es gibt zwei Attributdimensionen und zwei Zielkategorien. Die Items sollen in fünf unterschiedlichen Phasen möglichst schnell und korrekt zugeordnet werden. Der Aufbau orientiert sich an dem von Greenwald und Farnham (2000) entwickelten IAT in englischer Sprache zur Messung der Selbstwertschätzung beziehungsweise des Selbstkonzepts.

7.1.1. Ablauf

Der in dieser Arbeit verwendete Selbstwert-IAT, orientiert sich an dem von Rudolph, Schröder-Abé und Schütz (2006) nach dem Vorbild von Greenwald und Farnham (2000) entwickelten IAT. Die Programmierung des Selbstwert-IATs wurde mittels der Software Inquisit (Version 3.0.4.0; Millisecond Software LLC, 2010) vorgenommen. Der Selbstwert-IAT nach Rudolph, Schröder-Abé und Schütz (2006) prüft die implizite Selbstwertschätzung. Die Erfassung dieser Dimension erfolgt mittels Prüfung von Assoziationsstärken zwischen bipolaren Ziel- bzw. Objektdimensionen, selbstrelevant "ich" und nicht selbstrelevant "nicht ich", und bipolaren Attributkategorien, "positiv" und "negativ", die Stimuli, die zugeordnet werden sollen, werden in zufälliger Reihenfolge vorgegeben (s. Kapitel 1.1.1. und Tabelle 2).

Den Testteilnehmern wird dieses Verfahren als einfacher Arbeitsauftrag im Sinne einer Zuordnungsaufgabe vorgestellt, bei einer falschen Antwort erscheint ein rotes X in der Mitte des Bildschirms, so lange bis die korrekte Antwort gegeben wird:

Instruktion Selbstwert-IAT:

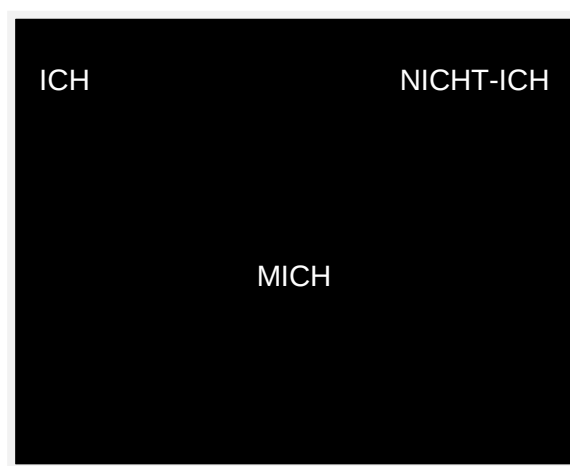
„Legen Sie Ihre Mittel- oder Zeigefinger auf die Tasten E und I Ihrer Tastatur. In der nächsten Aufgabe wird Ihnen eine Anzahl von Wörtern gezeigt, die Sie zuordnen sollen. Zwei Hinweiskategorien am oberen Bildschirmrand werden Ihnen zeigen, welche Wörter zu welcher Taste gehören. Wenn das Wort zu der Kategorie auf der linken Seite des Bildschirms gehört, drücken Sie die E Taste; wenn das Wort zu der Kategorie auf der rechten Seite gehört, drücken Sie die I Taste. Jedes Wort gehört immer nur zu einer Kategorie.

Wenn Sie einen Fehler machen erscheint ein X – um den Fehler auszubessern, drücken Sie einfach die andere Taste.

Es handelt sich um eine Zuordnungsaufgabe mit Zeitmessung. Arbeiten Sie so schnell wie möglich und versuchen Sie möglichst wenige Fehler zu machen. Wenn Sie zu langsam arbeiten oder zu viele Fehler machen, kann das Ergebnis nicht interpretiert werden. Der Test wird ungefähr 5 Minuten dauern.“

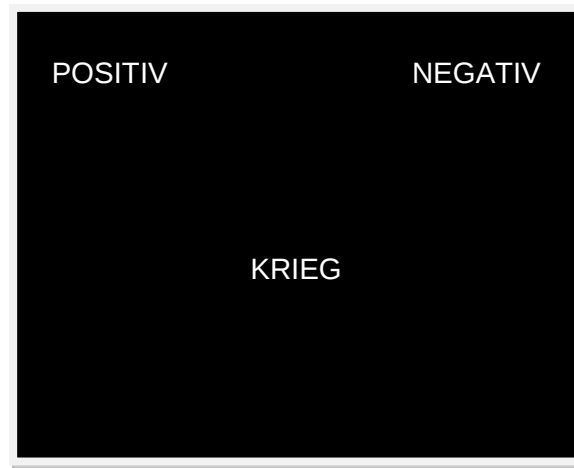
In der ersten Phase ist es Aufgabe der TeilnehmerInnen, Items den Zielkategorien so schnell und richtig wie möglich zuzuordnen.

Abbildung 1: IAT Bildschirm in Block 1



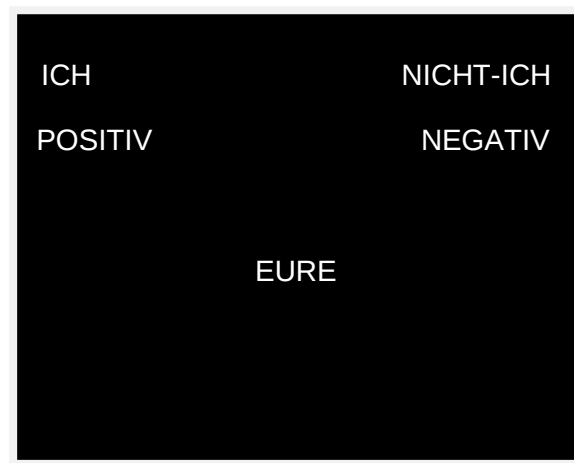
Im zweiten Block besteht die Aufgabe darin, den beiden Attributdimensionen die einzelnen Stimuli so schnell und richtig wie möglich zuzuordnen.

Abbildung 2: IAT Bildschirm in Block 2



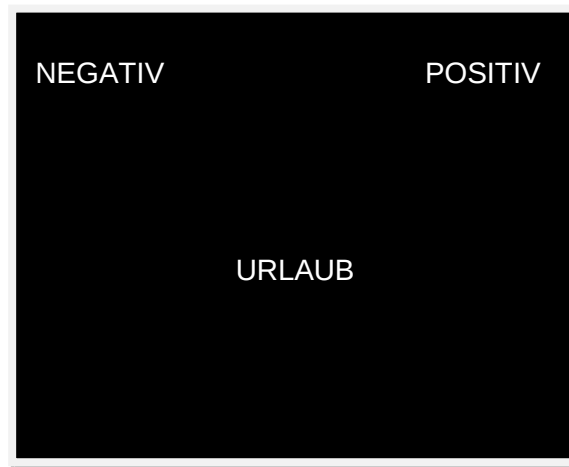
Der dritte Block ist der erste, der sowohl Attribut- als auch Zielkategorien vereint, Ziel ist es, Stimuli beider Kategorien sowohl Attribut- als auch Zieldimension richtig zuzuordnen.

Abbildung 3: IAT Bildschirm in Block 3



In der vierten Phase wird die Seite der Darstellung der Attributdimension umgedreht, Ziel ist es wiederum die Items den Kategorien „nicht ich“ und „ich“ so schnell wie möglich zuzuordnen.

Abbildung 4: IAT Bildschirm in Block 4



In der letzten, fünften Phase werden wiederum Ziel- und Attributdimensionen simultan dargestellt, die Darstellung der Zielkategorien aus Phase vier wird beibehalten, die Items sollen wieder so schnell und korrekt wie möglich zugeordnet werden.

Abbildung 5: IAT Bildschirm in Block 5

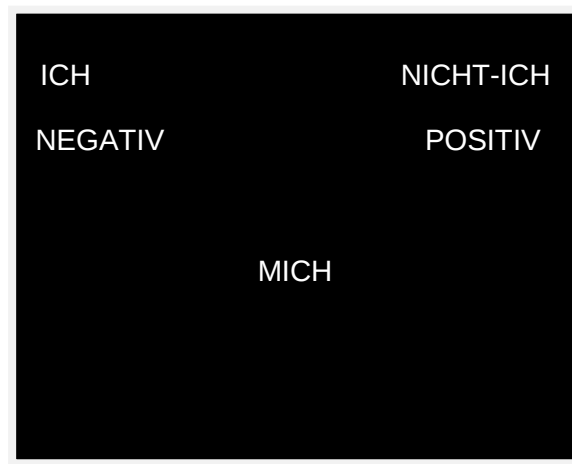


Tabelle 2: Übersichtsdarstellung der Blöcke des IAT, Items und Trialanzahl

Block	Trials	Beschreibung	Kategorien links	Items	Kategorien rechts
1	24	Übungsblock Zielkategorien	ICH • • • • • •	MEIN MICH MEINE ICH SELBST MIR EUCH EURE IHR IHNEN ANDERE IHRE	NICHT-ICH • • • • • •
2	24	Übungsblock Attributkategorien	POSITIV • • • • • •	LÄCHELN FRIEDEN GLÜCK FREUDE GESUNDHEIT URLAUB GESTANK KRIEG UNGLÜCK LEID KRANKHEIT QUAL	NEGATIV • • • • • •
3	(96) 24 72	Kompatibler Block Übungsphase Erhebungsphase	ICH POSITIV • •	MEIN IHR LÄCHELN KRIEG	NICHT-ICH NEGATIV • •
4	24	Übungsblock Attributkategorien	NEGATIV •	KRIEG GLÜCK	POSITIV •
5	(96) 24 72	Inkompatibler Block Übungsphase Erhebungsphase	ICH NEGATIV • •	ICH IHNEN FREUDE UNGLÜCK	NICHT-ICH POSITIV • •

Note: Die korrekte Antwort ist jeweils durch einen schwarzen Punkt gekennzeichnet.

In der Tabelle werden in Block 1 und 2 alle Items, die in diesem Selbstwert-IAT verwendet wurden, aufgelistet.

7.1.2. Verrechnungsmodus

In dieser Arbeit wurde der verbesserte Verrechnungsalgorithmus („improved scoring algorithm“) nach Greenwald, Nosek und Banaji (2003) verwendet. Er zeichnet sich durch eine eindeutigere Erfassung des Konstrukts aus, ist weniger empfindlich gegenüber externen Einflüssen und bildet somit Assoziationsstärken besser ab.

Extreme *D*-Werte, die durch extrem langsame Antworten entstehen, werden ausgeschlossen. Der neue Algorithmus soll laut den Autoren insofern validere Korrelationen des IAT-Wertes und anderer interessierender Variablen, wie Kapazität des Arbeitsgedächtnisses oder des Alters aufweisen. Auch vor dem Artefakt der Vorerfahrung mit impliziten Assoziationsverfahren soll der neue Algorithmus schützen, obwohl auch dieser Berechnungsmodus den Effekt von Vorerfahrungen nicht komplett ausmerzen konnte.

Der verbesserte IAT Wert wird in mehreren Schritten berechnet. Zunächst werden sowohl die Daten aus den Übungs- als auch Erhebungsphasen der Blöcke drei und fünf verwendet. Im zweiten Schritt werden alle Trials der Blöcke drei und fünf mit Latenzzeiten von über 10 000 Millisekunden entfernt, die Daten von TeilnehmerInnen, bei welchen mehr als 10 Prozent der Trials Latenzzeiten von weniger als 300 Millisekunden aufweisen, werden komplett ausgeschlossen. Anschließend wird der Mittelwert für die Latenzzeiten aller korrekten Antworten jeweils getrennt für die Übungs- und Erhebungsphasen von Block fünf und drei berechnet und die gepoolte Standardabweichung für alle Trials (sowohl korrekte als auch inkorrekte) der Erhebungsphasen aus Block fünf und drei sowie für die Übungsphase aus Block fünf und drei jeweils gemeinsam gebildet. Im nächsten Schritt wird für die Latenzzeit jeder fehlerhaften Antwort der für den jeweiligen Block errechnete Mittelwert plus 600 Millisekunden eingesetzt. Der aus diesen Änderungen neu entstandene Mittelwert wird nun für alle Trials jeweils getrennt für die Übungs- und Erhebungsphase der beiden Blöcke berechnet. Schließlich werden aus den im letzten Schritt resultierenden Mittelwerten die Differenzen der Erhebungs- und Übungsphasen gebildet, indem der Mittelwert der Erhebungsphase von Block drei von der Erhebungsphase von Block fünf und der Mittelwert der Übungsphase von Block drei von der Übungsphase von Block fünf subtrahiert wird. Die resultierenden Werte werden durch die zuvor berechnete gepoolte Standardabweichung dividiert und die beiden Quotienten jeweils entsprechend der Anzahl der Trials in Übungs- und Erhebungsphase gewichtet und der Mittelwert gebildet. Der daraus resultierende Wert ist der *D*-Wert und Maß der Assoziationsstärke.

7.2.ILICA

ILICA ist, wie bereits mehrfach erwähnt, ein Computersimulationsverfahren, Ziel ist es mit Hilfe eines möglichst realitätsnahen Szenarios Planungskompetenz, im Sinne einer Persönlichkeitseigenschaft zu erfassen. Selbstverwaltungskompetenz fasst Kubinger (persönliche Mitteilung, März 1990; zit. n. Möseneder, 1992; S. 11) folgendermaßen zusammen: „die Fähigkeit, mittel- und längerfristige Ziele erfolgreich zu erreichen und damit konkurrierende Ablenkungen von außen – soweit nötig - widerstehen zu können.“

Aufgabe der TeilnehmerInnen ist es, einen freien Tag vor dem Antritt eines Urlaubes zu planen. Es gibt einige Erledigungen, die vom Programm vorgegeben werden, die Durchführung dieser Tätigkeiten soll bevorzugt werden. Neben diesen Aufgaben gibt es allerdings noch genug Zeit für frei wählbare Aktivitäten.

Nach einer kurzen Einleitung in das Programm, in welcher auch alle zur Verfügung stehenden Hilfsmittel vorgestellt werden, Uhr, Kalender, Telefon und Notizblock, beginnt die Simulation. Ein Urlaubsziel kann frei ausgesucht werden, was den persönlichen Anreiz der Testsituation erhöhen soll. Man kann aus 40 Aktivitäten frei wählen, wobei zu beachten ist, dass einige der vorgegebenen Aufgaben an bestimmte Uhrzeiten gebunden sind, drei dieser Aufgaben sind auch bereits in den Kalender eingetragen. Während der Bearbeitung des Tests treten je nach gewählten Aktivitäten auch immer wieder unvorhergesehene Ereignisse auf. Auf diese Ablenkungen muss die Testperson reagieren, es stehen jeweils drei Verhaltensmöglichkeiten zur Auswahl. Die Simulation ist beendet, wenn 18:00 Uhr erreicht ist, während der Testbearbeitung besteht kein Zeitdruck.

7.2.1. Skalen

Der Testperson werden während der Bearbeitung des Tests mehrere Entscheidungen abverlangt, die jeweils zur Verfügung stehenden drei Antwortmöglichkeiten werden dazu verwendet, eine Typenzuordnung vorzunehmen. Es wird zwischen impulsivem,

flexiblem und reflexivem Typ unterschieden. Der relative Anteil der gewählten Antworten in Bezug auf den jeweiligen Typ wird angegeben.

Die Typen werden folgendermaßen beschrieben:

- Der **impulsive Typ** ist ablenkbar, er neigt dazu auch Dinge zu tun, die in Konflikt mit seinen ursprünglichen Zielen stehen.
- Der **reflexive Typ** handelt überlegt und zielstrebig, er ist nicht kompromissbereit.
- Der **flexible Typ** ist kompromissbereit, hat sein Ziel vor Augen, ist aber bereit bei auftretenden Störungen diese auch gegenüber anderen Bedürfnissen abzuwägen.

Neben dieser Typzuordnung liefert ILICA noch eine umfassende Liste an Kennwerten bezüglich der erhobenen Selbstverwaltungskompetenz. Im Testprotokoll werden folgende Werte angeführt:

- Plan: Diese dichotome Variable gibt an, ob die Person sich dafür entschieden hat mit Tagesplanung zu arbeiten, das heißt vorab eine Liste der Tätigkeiten zu erstellen, die sie im Laufe des Tages durchführen will, oder ob sie ohne Tagesplanung gearbeitet hat.
- Kalender, Uhr und Telefon: Diese drei Werte zeigen an, wie häufig das jeweilige Hilfsmittel verwendet worden ist.
- Zielanzahl: Dieser Kennwert wird nur angegeben, wenn die Person mit Tagesplanung gearbeitet hat, und erfasst, wie viele Tätigkeiten die Person vorab für den Tagesverlauf geplant hat.
- Aktivitätenanzahl: Dieser Wert gibt die Anzahl der tatsächlich durchgeführten Aktivitäten wieder.
- Planerfüllung: Diese dichotome Variable wird nur angegeben wenn mit Tagesplanung gearbeitet wurde, sie gibt an, ob alle Aktivitäten, die zuerst geplant wurden, auch tatsächlich durchgeführt wurden.
- Vorgabeerfüllung: Dieser Kennwert zeigt, wie viele der vom Programm im Vorhinein festgelegten Tätigkeiten auch tatsächlich erledigt wurden.
- Urlaubsaktivitäten, Freizeitaktivitäten, Alltagsaktivitäten und Sozialaktivitäten: Diese Variablen geben jeweils den relativen Anteil von Urlaubsaktivitäten,

Freizeitaktivitäten, Alltagsaktivitäten und Sozialaktivitäten an der Gesamtzahl der Aktivitäten wieder.

- Plananzahl: Die Variable Plananzahl gibt die Menge jener Aktivitäten wieder, die geplant und dann auch tatsächlich durchgeführt wurden. Dieser Kennwert wird nur angegeben, wenn die Strategie mit Tagesplanung gewählt wurde.
- Ungeplant: Der Wert Ungeplant gibt an, wie viele ungeplante Aktivitäten durchgeführt wurden. Auch diese Variable wird nur angegeben, wenn mit Tagesplanung gearbeitet wurde.

7.3.AKZEPT!-L

Akzeptanz stellt laut Kersting (2008) „die Summe der auf das individuelle Erleben und Bewerten diagnostischer Verfahren gerichteten Aspekte dar“ (S. 422). Die Messung von Akzeptanz erscheint nur sinnvoll wenn folgende Kriterien erfüllt sind: Die Befragten müssen das bezüglich der Akzeptanz zu untersuchende Verfahren tatsächlich ausgeführt haben, das in Frage stehende Verfahren sollte psychometrischen Standards genügen, theoretisch fundiert und mehrdimensional sein, Umgebungs- und situationsabhängige Faktoren sollten genauso beachtet werden wie individuelle Unterschiede, wie zum Beispiel persönliches Interesse oder Kompetenz und wenn es um den Vergleich zweier unterschiedlicher Verfahren geht, sollen diese unter möglichst denselben situativen Bedingungen erhoben werden.

Kersting entwickelte auf Basis dieser Postulate einen Fragebogen, dessen Ziel es ist, auch für unterschiedliche Erhebungsinstrumente anwendbar und anpassbar zu sein und gleichzeitig Akzeptanz in möglichst gut vergleichbarer Art und Weise zu erheben. Schließlich wurden drei unterschiedliche Fragebogen entwickelt, der AKZEPT!-L zur Erhebung eines standardisierten Akzeptanzwertes bei Leistungstests, der AKZEPT!-P zur Anwendung bei Persönlichkeitsfragebogen und der AKZEPT!- AC für den Einsatz bei Assessment-Centern.

In dieser Arbeit kommt der AKZEPT!-L zum Einsatz. Da es sich um Objektive Persönlichkeitstests handelt und beide Tests eine Leistungskomponente beinhalten,

war dieser Fragebogen der mit der besten Passung. Für Objektive Persönlichkeitstests im Speziellen gibt es bisher noch keinen Fragebogen.

7.3.1. Skalen

Der AKZEPT!-L besteht aus vier Skalen, pro Skala gibt es vier Items, zu den einzelnen Items steht jeweils eine sechsstufige Bewertungsskala zur Verfügung, welche von „trifft nicht zu“ bis zu „trifft genau zu“ reicht. Eine besonders negative Beurteilung resultiert in einem niedrigen Wert (1) eine besonders positive Bewertung entspricht einem hohen Wert (6). Verneinend umschriebene Items werden anschließend bei der Auswertung umgewandelt, so dass insgesamt ein hoher Wert hoher Akzeptanz entspricht.¹

- Die Skala Messqualität soll die persönliche Einschätzung des Teilnehmers erfassen, inwiefern das Verfahren Leistungsunterschiede zwischen verschiedenen Testpersonen auch tatsächlich abbildet (Zimmerhofer, 2008).

Beispielitem:

„Der Test ermöglicht es, die zwischen verschiedenen Menschen bestehenden Leistungsunterschiede in der vom Test gemessenen Fähigkeit exakt zu messen.“

- Augenscheinvalidität bedeutet in diesem Zusammenhang die Übereinstimmung des von der Testperson vermuteten Ziels der Messung mit dem tatsächlichen Ziel der Messung oder dem genannten Ziel der Messung (Zimmerhofer, 2008).

Beispielitem:

„Dass man mit dem Test geeignete Personen für einen Job herausfinden kann, ist zu bezweifeln.“

¹ Der AKZEPT!-L wurde kostenlos zur Verfügung gestellt, eine Weitergabe allerdings ausgeschlossen, insofern muss hier auf eine vollständige Darstellung des Fragebogens verzichtet werden, die hier angeführten Items entsprechen den unter: <http://www.kersting-internet.de/akzept.html> [05.01.2010] frei verfügbaren Beispielitems.

- Kontrollierbarkeit erfasst das Verständnis der Aufgabenstellung, also ob die Testperson klar verstanden hat worin ihre Aufgabe besteht, also auch ob Instruktionen verständlich formuliert sind (Zimmerhofer, 2008).

Beispielitem:

„Bei der Bearbeitung der Testaufgaben wusste ich jederzeit, was ich tun muss.“

- Mit der Skala Belastungsfreiheit soll erfasst werden inwieweit die Testperson von den gestellten Aufgaben gefordert wurde, ob eine Unter- oder Überforderung vorliegt (Zimmerhofer, 2008).

Beispielitem:

„Die Bearbeitung der Testaufgaben ist belastend.“

Zusätzlich wird mit zwei Items die globale Bewertung des Verfahrens erhoben. Einerseits wird die Gesamtbewertung des Verfahrens an sich erhoben:

„Welche Schulnote würden Sie dem soeben bearbeiteten Test geben?“

Das zweite Item soll Aufschluss darüber geben, wie TeilnehmerInnen ihre eigene Leistung in dem bearbeiteten Test einschätzen:

„Im Vergleich mit anderen Personen meiner Altersgruppe (mit gleicher Schulbildung) denke ich, dass ich im Test...abgeschnitten habe.“

Bei der hier vorliegenden Arbeit wurden alle Items der Originalfassung beibehalten, auch wenn einige Items aufgrund der Besonderheit Objektiver Persönlichkeitstests nicht vollkommen den vorgegebenen Verfahren entsprachen. Der Fragebogen wurde jedoch unter der Bedingung der Beibehaltung aller Items zur Verfügung gestellt.

Der AKZEPT!-L war handschriftlich zu beantworten, jeweils nach dem ersten Testdurchgang von ILICA und Selbstwert-IAT wurden die Testteilnehmer per Instruktion dazu aufgefordert aufzuzeigen, der Fragebogen wurde vorgelegt.

8. Durchführung der Untersuchung

Die Testungen im April und Mai 2010 in den drei unterschiedlichen Schulen in Niederösterreich bedurften einer genauen Planung und insbesondere einer gezielt für die speziellen Anforderungen, der Testung an unterschiedliche Orten, und der unterschiedlichen Voraussetzungen der Testungsprogramme, ausgelegten und überlegten Implementierung und Aufbereitung der Programme. In den Monaten März und April 2010 wurde der Selbstwert-IAT mit Hilfe von Inquisit (Version 3.0.4.0; Millisecond Software LLC, 2010) programmiert, sowie die gesamte Testbatterie erstellt. Anschließend wurde an einigen PsychologiestudentInnen eine Probetestung durchgeführt (diese Daten sind nicht in die empirische Auswertung mit eingeflossen).

8.1. Beschreibung der Datenerhebung

Die Anforderungen an die Testbatterie waren umfassend, so sollte die Testung an vielen unterschiedlichen Orten stattfinden, die Voraussetzungen in den Schulen und der dort zur Verfügung gestellten PCs und EDV Systeme waren unterschiedlich. Teilweise gab es eingeschränkte Berechtigungen auf den Schul-PCs, die Installation von Software war unerwünscht. Ziel war es, einen durchgängigen Testablauf zu ermöglichen, der von den Testteilnehmern nicht beeinflusst werden muss, damit die Bearbeitung der unterschiedlichen Verfahren ungestört erfolgen kann.

Angedachte Möglichkeit war beispielsweise eine Onlinetestung, leider wäre die Umsetzung mit erheblichem Programmieraufwand einer Onlineversion von ILICA verbunden gewesen. Eine weitere Möglichkeit wäre die Ausstattung von Testnotebooks mit den Testprogrammen gewesen, da allerdings das Ausleihen einer entsprechenden Anzahl von Notebooks über einen langen Zeitraum nicht möglich war, wurde auch diese Möglichkeit ausgeschlossen.

Einzig realisierbare Variante der Implementierung blieb die Erstellung einer virtuellen Maschine, in welcher alle Testprogramme und Instruktionen fertig konfiguriert werden konnten, weiterer Vorteil war die Möglichkeit der beliebig häufigen Vervielfältigung. Nach Evaluierung mehrerer Virtualisierungslösungen kristallisierte sich QEMU

(Version 0.12.3, Fabrice Bellard, 2010) als einzige Möglichkeit heraus. QEMU ist eine frei verfügbare Virtualisierungssoftware, die ohne Administratorenrechte und ohne Installation funktioniert. Als Basis für die virtuelle Maschine diente Windows-XP (Windows NT Version 5.1, Microsoft Corporation, 2001). Die Größe der Festplatte wurde auf 4 Gigabyte beschränkt, um auch Testungen direkt via USB-Stick an den unterschiedlichen Schul-PCs durchführen zu können.

Mit Hilfe von Batch-Skripten und einigen Hilfsprogrammen wurden die Testabläufe modelliert. Alle erforderlichen Verfahren wurden nach Eingabe der zufällig zugeteilten Teilnehmernummer entsprechend der Zugehörigkeit zu den Untersuchungsgruppen der Reihe nach gestartet. Der komplette Testablauf inklusive den Instruktionsanweisungen stellte einen abgeschlossenen Ablauf dar. Für die TeilnehmerInnen waren, abgesehen vom Eingeben der Probandennummer, der Bearbeitung der Aufgaben innerhalb der beiden Verfahren und dem Aufzeigen vor Vorgabe des AKZEPT!-L keine Aktionen notwendig.

Vor der Durchführung wurden die virtuellen Maschinen jeweils von den TestleiterInnen gestartet und nach Beendigung wieder heruntergefahren. Anschließend konnten die Testergebnisse aus den virtuellen Maschinen abgerufen und in SPSS überspielt werden.

8.2.Instruktionen

Um sicherzustellen, dass die Personen das Testergebnis und damit den Ausgang der Studie bei der zweiten Vorgabe nicht willentlich beeinflussen, wurden die Testpersonen über den wahren Grund der Untersuchung (die Erhebung der Stabilität von ILICA und Selbstwert-IAT) im Unklaren gelassen. Als Grund für die zweite Vorgabe wurde angegeben, dass dadurch ein höherer Genauigkeitsgrad in den Ergebnissen erreicht werden könne. Die Einleitungsinstruktion lautete folgendermaßen:

Instruktion Einleitung:

„Ich begrüße Sie sehr herzlich und freue mich, dass Sie an dieser Untersuchung teilnehmen.

Sie werden nun zwei unterschiedliche Tests bearbeiten. Um möglichst genaue Ergebnisse zu erhalten, wird jeder Test zweimal vorgegeben. Nach der ersten Vorgabe werden Sie jeweils handschriftlich um die Meinung zu dem vorangegangenen Test gebeten werden.“

Wie bereits erwähnt, wurde das Computersimulationsverfahren ILICA allen Testpersonen zweimal vorgegeben, etwa eine Hälfte der TeilnehmerInnen erhielt bei der zweiten Vorgabe eine Instruktion, die sie zur Anwendung neuer Strategien aufforderte. Diese Manipulation von Verhaltensvielfalt soll Aufschluss darüber geben, ob und wie stark die Testpersonen ohne Aufforderung im Vergleich zu jenen mit Aufforderung ihr Verhalten verändern und insofern ihr Testergebnis beeinflussen.

Vor Beginn des zweiten Testdurchganges wurde jenen Testpersonen folgende Instruktion präsentiert:

Instruktion Manipulation von Verhaltensvielfalt:

„Es folgt nun ein zweiter Testdurchgang der Urlaubsplanung.
Um ein möglichst breites Spektrum Ihrer Arbeitsweise zu erfassen, versuchen Sie bitte neue Strategien bei der Durchführung des Tests anzuwenden und Ihre Arbeitsweise zu variieren, aber Ihre Tagesziele trotzdem zu verfolgen!“

Die andere Hälfte der Testpersonen erhielt nur die Information, dass nun ein zweiter Testdurchgang folgt:

Instruktion keine Manipulation von Verhaltensvielfalt:

„ Es folgt nun ein zweiter Testdurchgang der Urlaubsplanung.“

9. Beschreibung der Stichprobe

Die Testungen zu dieser Arbeit fanden im Zeitraum von April bis Mai 2010 am Bundesrealgymnasium Gröhrmühlgasse in Wiener Neustadt, im Bundesoberstufenrealgymnasium Ternitz, und am Gymnasium und Realgymnasium Sachsenbrunn in Kirchberg am Wechsel statt.

An der Testung nahmen insgesamt 163 Schüler der siebenten und achten Schulstufe teil. Insgesamt wurden 54 SchülerInnen des Gymnasiums und Realgymnasiums Sachsenbrunn, 86 SchülerInnen des Bundesrealgymnasiums Wiener Neustadt und 23 SchülerInnen des Oberstufenrealgymnasiums Ternitz getestet. Nach Ausschluss nicht verwertbarer Daten (aufgrund von Hardwareproblemen oder Testungsabbruch) blieben 157 TeilnehmerInnen. Das Alter der SchülerInnen lag zwischen 16 und 20 Jahren (Mittelwert: 17,23). Die Stichprobe setzte sich insgesamt aus 75 Männern (47,8 Prozent) und 82 Frauen (52,2 Prozent) zusammen.

Die Testungen fanden in den Computersälen der jeweiligen Schule statt, teilweise mit Einsatz des mobilen Testplatzsystems des Arbeitsbereichs Psychologische Diagnostik. Die Teilnahme war freiwillig und die Erziehungsberechtigten der Schüler mussten vorher ihr Einverständnis für die Teilnahme ihres Kindes an der Testung geben (siehe Anhang).

Vorab wurde das Einverständnis des Landesschulrates für Niederösterreich eingeholt (siehe Anhang).

9.1. Bestimmung des Stichprobenumfangs

Der optimale Stichprobenumfang wurde mit Hilfe des Programmes Cademo light (Version 3.20, siehe www.biomath.de) berechnet. Es wurde ein hoher Korrelationskoeffizient, von $\rho = .70$, zwischen den beiden Testzeitpunkten für ILICA und Selbstwert-IAT angenommen, da ein niedrigerer Zusammenhang nicht für eine entsprechende Stabilität der Verfahren spricht. Es wurden folgende Werte festgesetzt: $\alpha = .05$, $\beta = .20$ und das Mindestmaß des Korrelationskoeffizient $\lambda = .7$, mit einem

Intervall von $\delta = .10$. Daraus ergibt sich, dass ein $\rho > \lambda + \delta = .80$ mit einer Wahrscheinlichkeit von mindestens $1 - \beta = .80$ registriert wird. Es ergibt sich ein notwendiger Stichprobenumfang von insgesamt 119 Personen.

Während der Durchführung der Testung ergab sich ein etwas größerer Stichprobenumfang ($n = 157$).

9.2. Verteilung der TeilnehmerInnen auf die Untersuchungsbedingungen

Die Zuteilung der einzelnen Teilnehmer zu den einzelnen Versuchsgruppen erfolgte per Los, dadurch konnte zwar keine vollkommene Gleichverteilung bezüglich des Geschlechts erreicht werden (besonders in Gruppe 1), dennoch sollten durch die zufällige Aufteilung Verzerrungen verhindert werden. Die exakte Verteilung der TeilnehmerInnen kann Tabelle drei entnommen werden.

Tabelle 3: Verteilung der männlichen und weiblichen TeilnehmerInnen auf die einzelnen Untersuchungsbedingungen

Versuchsgruppe	Anzahl männliche Teilnehmer	Anzahl weibliche Teilnehmer	Gesamt
Gruppe 1 (Nr. 100-199)	15	25	40
Gruppe 2 (Nr. 200- 299)	19	19	38
Gruppe 3 (Nr. 300-399)	19	21	40
Gruppe 4 (Nr. 400-499)	22	17	39
Gesamt	75	82	157

10. Ergebnisse

Im folgenden Abschnitt werden die Ergebnisse der statistischen Auswertung dargestellt, die Berechnungen wurden mit dem Programm SPSS 17.0.2 durchgeführt.

10.1. Mittelwerte und Standardabweichungen der Untersuchungsbedingungen

Um einen besseren Überblick zu gewährleisten werden im Folgenden alle Mittelwerte und Standardabweichungen der einzelnen Untersuchungsbedingungen zu beiden Messzeitpunkten dargestellt.

Tabelle vier sind die Mittelwerte und Standardabweichungen des Impliziten Assoziations Task zu Messzeitpunkt eins und zwei zu entnehmen.

Tabelle 4: Mittelwerte und Standardabweichungen des Selbstwert-IAT für die einzelnen Untersuchungsgruppen zu beiden Messzeitpunkten

Untersuchungsgruppe	IAT 1	IAT 2
1 (n=40)	.658 (.256)	.424 (.292)
2 (n = 38)	.618 (.257)	.456 (.289)
3 (n = 40)	.688 (.275)	.462 (.286)
4 (n= 39)	.679 (.279)	.557 (.317)
Gesamt (n = 157)	.661 (.266)	.474 (.298)

Note: Mittelwerte und die Standardabweichungen (in Klammern) des D-Werts des IAT, jeweils für die einzelnen Untersuchungsbedingungen sowie für die Gesamtstichprobe

In Tabelle Nummer fünf sind die Mittelwerte und Standardabweichungen der einzelnen Skalen von ILICA zu Messzeitpunkt eins und zwei, jeweils für die einzelnen Untersuchungsbedingungen, dargestellt.

Tabelle 5: Mittelwerte und Standardabweichungen von ILICA für die einzelnen Untersuchungsgruppen zu beiden Messzeitpunkten.

Skala (Messzeitpunkt)	Gruppe 1 (n = 40)	Gruppe 2 (n = 38)	Gruppe 3 (n = 40)	Gruppe 4 (n = 39)	Insgesamt (n = 157)
Kalender (1)	2.73 (2.28)	2.55 (2.54)	2.23 (2.09)	1.64 (1.97)	2.29 (2.25)
Kalender (2)	1.73 (1.96)	1.16 (1.79)	1.32 (1.69)	1.38 (1.79)	1.40 (1.80)
Uhr (1)	8.58 (5.41)	8.53 (5.97)	7.75 (4.56)	7.90 (5.41)	8.18 (5.31)
Uhr (2)	8.15 (3.66)	7.03 (4.61)	7.87 (3.95)	7.15 (4.95)	7.51 (4.30)
Telefon (1)	1.75 (2.04)	1.45 (1.88)	1.38 (1.92)	.85 (1.04)	1.36 (1.78)
Telefon (2)	.60 (.93)	.47 (.80)	.73 (1.18)	.56 (1.17)	.59 (1.03)
Urlabsaktivitäten (1)	52.45 (13.21)	50.47 (12.99)	55.00 (12.07)	52.15 (13.50)	52.55 (12.93)
Urlabsaktivitäten (2)	51.80 (12.03)	45.32 (15.78)	51.43 (12.87)	52.33 (12.77)	50.27 (13.58)
Freizeitaktivitäten (1)	13.83 (10.65)	14.05 (10.88)	14.23 (10.24)	15.13 (11.27)	14.31 (10.67)
Freizeitaktivitäten (2)	12.40 (10.90)	13.79 (12.24)	13.12 (11.41)	14.59 (12.15)	13.46 (11.59)
Alltagsaktivitäten (1)	19.13 (8.57)	18.87 (11.49)	17.03 (9.54)	16.72 (10.04)	17.93 (9.91)
Alltagsaktivitäten (2)	21.32 (7.98)	24.32 (10.27)	20.45 (9.90)	16.44 (9.06)	20.61 (9.66)
Sozialaktivitäten (1)	14.63 (8.10)	16.37 (8.69)	13.70 (6.93)	15.85 (9.71)	15.11 (8.39)
Sozialaktivitäten (2)	14.38 (7.86)	16.45 (8.37)	14.95 (7.21)	16.59 (5.63)	15.57 (7.33)
Flexibler Typ (1)	35.48 (13.65)	34.37 (16.35)	32.07 (17.57)	41.38 (15.82)	35.81 (16.13)
Flexibler Typ (2)	35.77 (12.23)	32.55 (14.66)	29.60 (17.79)	38.87 (13.78)	34.19 (15.03)
Impulsiver Typ (1)	34.78 (12.80)	38.24 (15.98)	37.40 (14.98)	32.67 (12.94)	35.76 (14.25)
Impulsiver Typ (2)	36.05 (13.69)	43.50 (16.92)	37.43 (15.63)	32.67 (15.67)	37.36 (15.83)
Reflexiver Typ (1)	29.30 (11.99)	27.34 (16.48)	30.35 (13.85)	25.82 (14.57)	28.23 (14.25)
Reflexiver Typ (2)	27.90 (12.04)	23.58 (12.66)	32.88 (15.42)	28.26 (14.30)	28.21 (13.94)

Note: Mittelwerte und Standardabweichungen (in Klammern) für die Skalen Kalender, Uhr, Telefon, Urlabsaktivitäten, Alltagsaktivitäten, Sozialaktivitäten, Freizeitaktivitäten, Impulsiver Typ, Reflexiver Typ und Flexibler Typ, des ILICA, jeweils für die einzelnen Untersuchungsbedingungen sowie für die Gesamtstichprobe zu Messzeitpunkt eins und zwei. Untersuchungsbedingungen mit Manipulation von Verhaltensvielfalt wurden grau unterlegt.

10.2. Retest-Reliabilität IAT

Alternativhypothese Ia: Der Korrelationskoeffizient zwischen erstem und zweitem Testergebnis des Selbstwert-IAT ist signifikant und höher als .70.

$$H(Ia)0: 0 < \rho \leq .70$$

$$H(Ia)1: \rho > .70$$

Entsprechend der Formulierung der Hypothese wurde auch die Signifikanz der Korrelationskoeffizienten entsprechend der von Kubinger, Rasch und Šimeckova (2007) vorgeschlagenen Vorgehensweise berechnet.¹ Die H0 wird verworfen wenn $u > u(1 - \alpha)$, wobei $(1 - \alpha)$ dem Quantil der einseitigen Standardnormalverteilung entspricht.

Für die Testwerte des IAT war Normalverteilung gegeben, es wurde eine Korrelation nach Pearson gerechnet.

Tabelle 6: Korrelationskoeffizient Selbstwert-IAT

		d – Wert IAT 1	d - Wert IAT 2
d - Wert IAT 1	Korrelation nach Pearson	1	,377
	N	157	157
d - Wert IAT 2	Korrelation nach Pearson	,377	1
	N	157	157

Der Korrelationskoeffizient zwischen dem IAT-Effekt zu Messzeitpunkt eins und dem IAT-Effekt zu Messzeitpunkt zwei beträgt .377 und ist nicht signifikant. Die Nullhypothese muss dementsprechend beibehalten werden, da der Zusammenhang nicht dem im Vorhinein festgesetzten Kriterium, eines Korrelationskoeffizienten von .70 entspricht. Dieses Ergebnis entspricht nicht der geforderten Stabilität.

¹ Zur Berechnung der Signifikanz liegt eine SPSS Syntax vor, diese kann dem Artikel von Kubinger, Rasch und Šimeckova (2007) zur Testung von Korrelationskoeffizienten entnommen werden.

10.3. Retest-Reliabilität ILICA

Alternativhypothese Ib: Der Korrelationskoeffizient zwischen erstem und zweitem Testergebnis von ILICA ist signifikant und höher als .70.

$$H(Ib)0: 0 < \rho \leq .70$$

$$H(Ib)1: \rho > .70$$

Bei der Berechnung der Korrelation wurden in diesem Fall die TeilnehmerInnen mit Verfälschungsinstruktion von der Analyse ausgeschlossen, es resultiert eine Stichprobengröße von 80 Personen. Auch hier wurden die Signifikanzberechnungen entsprechend der von Kubinger, Rasch und Šimeckova (2007) empfohlenen Formel berechnet.

Da bei ILICA die Testergebnisse in einzelnen Skalen ausgegeben werden, wurden die Korrelationskoeffizienten jeweils getrennt für die einzelnen Skalen berechnet. Die Korrelationskoeffizienten können Tabelle sieben entnommen werden.

Nicht für alle Skalen war Normalverteilung gegeben, es wurde dennoch eine Produkt-Moment Korrelation nach Pearson berechnet. Laut Havlicek und Peterson (1977) ist der Pearsons Korrelationskoeffizient robust gegenüber starken Verletzungen der Voraussetzung der Normalverteilung.

Für die Berechnung der Korrelation einiger Skalen konnte nur eine eingeschränkte Stichprobe herangezogen werden, da Kennwerte nur angegeben werden, sofern die Testpersonen mit einer Tagesplanung gearbeitet haben, es wurden nur jene Fälle ausgewählt, bei welchen sowohl bei Testzeitpunkt eins als auch zwei mit Tagesplanung gearbeitet wurde. Es ergibt sich ein Stichprobenumfang von $n = 47$, die Korrelationskoeffizienten können Tabelle acht entnommen werden.

Die Korrelationskoeffizienten liegen zwischen .298 und .803, keiner der gefundenen Korrelationskoeffizienten ist signifikant größer als .70.

Tabelle 7: Korrelationskoeffizienten der Skalen von ILICA (n = 80)

SKALA	r
Kalender	.298
Uhr	.498
Telefon	.354
Aktivitätenanzahl	.539
Vorgabeerfüllung	.592
Urlaubsaktivitäten	.693
Freizeitaktivitäten	.645
Alltagsaktivitäten	.672
Sozialaktivitäten	.576
Flexibler Typ	.477
Impulsiver Typ	.552
Reflexiver Typ	.695

Tabelle 8: Korrelationskoeffizienten der Skalen von ILICA, eingeschränkte Stichprobe (n = 47)

SKALA	r
Zielanzahl	.803
Plananzahl	.801
Ungeplant	.728

Die Nullhypothese muss für die Skalen Kalender, Uhr, Telefon, Aktivitätenanzahl, Vorgabeerfüllung, Urlaubsaktivitäten, Freizeitaktivitäten, Alltagsaktivitäten, Sozialaktivitäten, Flexibler Typ, Impulsiver Typ und Reflexiver Typ (s. Tabelle 7) beibehalten werden, keines der gefundenen Ergebnisse erreicht einen Korrelationskoeffizienten von .70.

Auch für die Skalen Zielanzahl, Plananzahl und Ungeplant (s. Tabelle 8) muss die Nullhypothese beibehalten werden, keines der gefundenen Ergebnisse für diese Skalen ist signifikant größer als .70.

Insgesamt entsprechen die Skalen nicht dem vorgegebenen Kriterium der Reliabilität des Verfahrens.

10.4. Manipulation von Verhaltensvielfalt ILICA

Hypothese II: Die Gruppe, die eine Manipulation von Verhaltensvielfalt erhalten hat, unterscheidet sich in ihren Testergebnissen zu Messzeitpunkt zwei nicht signifikant von der Gruppe, die keine Manipulation von Verhaltensvielfalt erhalten hat.

$$H(II)0: \mu_1 = \mu_2$$

$$H(II)1: \mu_1 \neq \mu_2$$

Zur Prüfung wurde ein nicht parametrisches Verfahren gewählt, der Mann-Whitney U-Test. Aufgrund der Vielzahl an Skalen die verglichen werden, steigt die Gefahr einer Kumulierung des Fehlers erster Art, es wurde eine Bonferroni Korrektur vorgenommen um diesem Fehler vorzubeugen.

Es werden die sieben Skalen, Urlaubsaktivitäten, Freizeitaktivitäten, Sozialaktivitäten, Alltagsaktivitäten, impulsiver Typ, reflexiver Typ und flexibler Typ der Gruppe ohne Manipulation von Verhaltensvielfalt ($n = 80$) und der Gruppe mit Manipulation von Verhaltensvielfalt ($n = 77$) verglichen. Die Ergebnisse der Analyse können Tabelle neun entnommen werden.

Laut Bonferroni Korrektur müsste ein $p < 0.0071$ vorliegen. Keines der Ergebnisse ist insofern signifikant, aber auch wenn keine dementsprechende Korrektur vorgenommen werden würde, wären die Ergebnisse nicht signifikant (s. Tabelle 9). Die Nullhypothese muss beibehalten werden, es konnten keine signifikanten Unterschiede zwischen der Gruppe mit Manipulation von Verhaltensvielfalt und der Gruppe ohne Manipulation festgestellt werden.

Tabelle 9: Manipulation von Verhaltensvielfalt Mann-Whitney U-Test

Skala	Keine Manipulation n = 80		Manipulation n = 77		U	p
	Mdn	Mittlerer Rang	Mdn	Mittlerer Rang		
Urlaubsaktivitäten	53	82.33	50	75.55	2814	.351
Freizeitaktivitäten	10	76.68	11	81.42	2894	.513
Alltagsaktivitäten	21	80.44	20	77.51	2965	.688
Sozialaktivitäten	15	72.28	15	85.47	2582	.080
Flexibler Typ	33	73.13	38	85.10	2610.5	.099
Impulsiver typ	38	78.94	36	79.06	3075	.987
Reflexiver Typ	30	85.93	23	71.80	2525.5	.051

Note: Es wird der Median (Mdn) für jede Skala berichtet, da diese Statistik im Fall eines nicht parametrischen Tests angemessener erscheint, zusätzlich werden jeweils der mittlere Rang, sowie die Mann-Whitney U-Statistik und p, der Signifikanzwert, angegeben.

10.4.1. Vergleich der Korrelationskoeffizienten

Um auch die Ergebnisse des ersten Messzeitpunktes mit einzubeziehen, wird zusätzlich ein Vergleich der Korrelationskoeffizienten zwischen der Gruppe ohne Manipulation von Verhaltensvielfalt und der Gruppe mit Manipulation von Verhaltensvielfalt durchgeführt.

Nullhypothese II ($K_{orr.}$): Der Korrelationskoeffizient der Gruppe ohne Manipulation von Verhaltensvielfalt unterscheidet sich nicht von dem Korrelationskoeffizienten der Gruppe mit Manipulation von Verhaltensvielfalt.

Alternativhypothese II ($K_{orr.}$): Der Korrelationskoeffizient der Gruppe ohne Manipulation von Verhaltensvielfalt unterscheidet sich von dem Korrelationskoeffizienten der Gruppe mit Manipulation von Verhaltensvielfalt.

$$H(II_{Korr.})_0: \rho_1 = \rho_2$$

$$H(II_{Korr.})_1: \rho_1 \neq \rho_2$$

Tabelle 10: Manipulation von Verhaltensvielfalt, Vergleich der Korrelationskoeffizienten

Skala	Keine Manipulation n = 80		Manipulation n = 77		z	p
	r	Z _F	r	Z _F		
Urlaubsaktivitäten	.693	.85	.518	.57	1.72	.085
Freizeitaktivitäten	.645	.77	.585	.67	.59	.553
Alltagsaktivitäten	.672	.81	.370	.39	2.62	.009
Sozialaktivitäten	.576	.66	.326	.34	1.95	.051
Flexibler Typ	.477	.52	.388	.41	.67	.501
Impulsiver Typ	.552	.62	.561	.63	.08	.936
Reflexiver Typ	.695	.86	.554	.62	1.43	.152

Note: Es werden jeweils die Korrelationskoeffizienten (r), Fisher's Z-Werte (Z_F) und Testgröße (z) berichtet. Signifikante Ergebnisse sind fett gedruckt.

Die Korrelationskoeffizienten sind in allen Skalen, außer der Skala impulsiver Typ, in der Bedingung mit Verfälschungsinstruktion niedriger als in der Bedingung ohne Verfälschungsinstruktion. Es ist festzuhalten, dass nur für die Skala Alltagsaktivitäten die Nullhypothese verworfen wird, für alle anderen Skalen muss die Nullhypothese beibehalten werden, da die Ergebnisse nicht signifikant sind.

10.5. Akzeptanz

Hypothese III: Es gibt keinen signifikanten Unterschied zwischen ILICA und IAT in den Ergebnissen der Skalen „Kontrollierbarkeit“, „Messqualität“, „Augenscheinvalidität“, „Belastungsfreiheit“, „Gesamtbeurteilung“ und „Verfahrensbezogene Selbsteinschätzung“.

$$H(II)0: \mu_1 = \mu_2$$

$$H(II)1: \mu_1 \neq \mu_2$$

Zur Prüfung dieser Hypothese wurde an der gesamten Stichprobe ein Wilcoxon Signed-Rank Test durchgeführt. Da es sich um Werte handelt, die von derselben Stichprobe erhoben wurden und die Daten nur Ordinalskalenniveau aufweisen, wurde dieses nicht parametrische Verfahren gewählt.

Es werden die Bewertungen aller TeilnehmerInnen (n = 157) von ILICA und IAT bezüglich der einzelnen Skalen des AKZEPT!-L verglichen.

Tabelle 11: Median der Skalen des AKZEPT!-L für ILICA und IAT

Skala	Median
Schulnote ILICA	2,00
Vergleich ILICA	3,00
Kontrollierbarkeit.ILICA	5,25
Messqualität.ILICA	3,75
Augenscheinvalidität.ILICA	3,75
Belastungsfreiheit.ILICA	5,25
Schulnote IAT	2,00
Vergleich IAT	3,00
Kontrollierbarkeit.IAT	6,00
Messqualität.IAT	4,00
Augenscheinvalidität.IAT	3,0000
Belastungsfreiheit.IAT	5,0000

Den Tabellen zwölf und dreizehn können die Häufigkeiten und dazugehörige Prozentwerte der vergebenen Schulnoten für die beiden Verfahren entnommen werden, wobei eine Eins einer besonders positiven, eine Sechs einer besonders negativen Bewertung entspricht.

Tabelle 12: Globale Bewertung ILICA

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig sehr gut	13	8,3	8,3	8,3
gut	70	44,6	44,6	52,9
befriedigend	59	37,6	37,6	90,4
ausreichend	12	7,6	7,6	98,1
mangelhaft	2	1,3	1,3	99,4
ungenügend	1	,6	,6	100,0
Gesamt	157	100,0	100,0	

Tabelle 13: Globale Bewertung IAT

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig sehr gut	19	12,1	12,1	12,1
gut	65	41,4	41,4	53,5
befriedigend	56	35,7	35,7	89,2
ausreichend	12	7,6	7,6	96,8
mangelhaft	4	2,5	2,5	99,4
ungenügend	1	,6	,6	100,0
Gesamt	157	100,0	100,0	

Es wurde eine Bonferroni Korrektur vorgenommen, um einer Kumulierung des Fehlers erster Art vorzubeugen. Das Signifikanzkriterium liegt daher bei .008.

Tabelle vierzehn können die Ränge, Tabelle fünfzehn die Teststatistiken des Vergleichs der beiden Konditionen entnommen werden.

Der Selbstwert-IAT (MD = 6) wurde bezüglich der Skala Kontrollierbarkeit signifikant höher eingeschätzt als ILICA (MD = 5.25), $T = 1303$, $p < .008$.

ILICA (MD = 3.75) wurde bezüglich der Augenscheinvalidität signifikant besser bewertet als der Selbstwert-IAT (MD = 3), $T = 1734$, $p < .008$. Auch innerhalb der Skala Belastungsfreiheit wurde ILICA (MD = 5.25) signifikant besser bewertet als der IAT (MD = 5), $T = 3014,5$, $p < .008$. Für die Skalen Messqualität, Bewertung des Verfahrens (Schulnote) und Vergleich der eigenen Leistung mit jener anderer Personen derselben Altersklasse und Schulbildung (Vergleich) muss die Nullhypothese beibehalten werden.

Tabelle 14: Ränge AKZEPT!-L

		N	Mittlerer Rang	Rangsumme
Kontrollierbarkeit.IAT - Kontrollierbarkeit.ILICA	Negative Ränge	25	52,12	1303,00
	Positive Ränge	111	72,19	8013,00
	Bindungen	21		
Messqualität.IAT - Messqualität.ILICA	Negative Ränge	59	70,37	4152,00
	Positive Ränge	86	74,80	6433,00
	Bindungen	12		
Augenscheinvalidität.IAT - Augenscheinvalidität.ILICA	Negative Ränge	106	80,77	8562,00
	Positive Ränge	37	46,86	1734,00
	Bindungen	14		
Belastungsfreiheit.IAT - Belastungsfreiheit.ILICA	Negative Ränge	90	61,12	5500,50
	Positive Ränge	40	75,36	3014,50
	Bindungen	27		
Schulnote IAT - Schulnote ILICA	Negative Ränge	56	52,88	2961,00
	Positive Ränge	51	55,24	2817,00
	Bindungen	50		
Vergleich IAT - Vergleich ILICA	Negative Ränge	31	37,97	1177,00
	Positive Ränge	45	38,87	1749,00
	Bindungen	81		

Note: Negative Ränge Skala IAT < Skala ILICA; Positive Ränge: Skala IA T> Skala ILICA; Bindungen: Skala IAT = Skala ILICA

Tabelle 15: Wilcoxon-Signed Rank Test AKZEPT!-L

	Kontrollierbar.	Messqualität	Augenscheinv.	Belastungsfrei.	Schulnote	Vergleich
Z	-7,312 ^a	-2,259 ^a	-6,893 ^b	-2,896 ^b	-,238 ^b	-1,597 ^a
p	,000	,024	,000	,004	,812	,110

a. Basiert auf negativen Rängen

b. Basiert auf positiven Rängen

Note: Kontrollierbar. = Kontrollierbarkeit, Augenscheinv. = Augenscheinvalidität, Belastungsfrei. = Belastungsfreiheit

Zusätzlich wird die Effektstärke für die Skalen mit signifikanten Ergebnissen nach der Formel von Rosenthal (1991, S. 19) berechnet:

$$r = \frac{Z}{\sqrt{N}}$$

Wobei N hier für die Anzahl an Beobachtungen steht, es ergibt sich ein $N = 314$.

Es ergeben sich folgende Effektgrößen:

Tabelle 16: Effektgrößen AKZEPT!-L ILICA und IAT

Skala	Z	r
Kontrollierbarkeit	-7.312	-.413
Augenscheinvalidität	-6.893	-.389
Belastungsfreiheit	-2.896	-.163

Die Effektstärken der Skala Kontrollierbarkeit, $r = -.413$, und der Skala Augenscheinvalidität, $r = -.389$, liegen zwischen den von Cohen (1992) als mittlerer ($r = .30$) und als großer ($r = .50$) Effekt bezeichneten Größen. Im Fall der Skala Kontrollierbarkeit erklärt der Effekt 17 Prozent der gesamten Varianz ($r^2 = .171$), für die Skala Augenscheinvalidität beträgt der Anteil erklärter Varianz 15 Prozent ($r^2 = .151$). Die Skala Belastungsfreiheit weist allerdings eine niedrigere Effektgröße, $r = -.163$ ($r^2 = .026$), auf (laut Cohen entspricht ein $r = .10$ einem kleinen Effekt), was für einen geringeren Unterschied der beiden Verfahren in dieser Skala spricht.

11. Interpretation und Diskussion

Die Ergebnisse der Retest-Reliabilitätsuntersuchung des Selbstwert-IAT zeigen einen verhältnismäßig niedrigen Zusammenhang der Testergebnisse zu den beiden Messzeitpunkten. Eine geforderte Stabilität, also ein Korrelationskoeffizient von .70, konnte bei weitem nicht erreicht werden. Vergleicht man das hier vorliegende Ergebnis mit Resultaten bisheriger Untersuchungen zum Selbstwert-IAT (Rudolph et al., 2008), wie sie auch in Kapitel 2.2. beschrieben werden, bleibt der hier gefundene Zusammenhang hinter jenen Ergebnissen zurück. Ein Grund könnte möglicherweise der geringe Zeitabstand sein, da die meisten anderen Untersuchungen doch mindestens ein Intervall von einer Woche zwischen den beiden Messzeitpunkten aufweisen. Bei Betrachtung der Mittelwerte der verschiedenen Untersuchungsgruppen in dieser Untersuchung fällt ins Auge, dass in allen Bedingungen der *D*-Wert, also der IAT-Effekt zu Messzeitpunkt zwei, abgenommen hat. Aufgrund des äußerst geringen Zeitraumes zwischen den beiden Testungen (das Intervall lag je nach Geschwindigkeit der Bearbeitung der Testbatterie bei ungefähr 40 Minuten), könnte es zu verstärkten Übungseffekten gekommen sein, die das Ergebnis beeinflusst haben könnten. Situative Einflussfaktoren, die von einigen Autoren als Ursache für schlechte Stabilitätswerte des IAT angenommen werden, können hier allerdings weitgehend ausgeschlossen werden, da aufgrund der unmittelbaren aufeinanderfolgenden Testungen kaum derartige Einwirkungen auf das Testergebnis möglich waren.

Die Korrelationskoeffizienten von ILICA zu den beiden Messzeitpunkten entsprechen nicht den entsprechenden Forderungen nach Stabilität. Verglichen mit den Ergebnissen von Herle (1999) sind die hier gefundenen Stabilitätsmaße etwas höher. Innerhalb der Skalen Zielanzahl, Anzahl ungeplant durchgeführter Aktivitäten und Anzahl der geplanten und auch tatsächlich durchgeführten Aktivitäten konnte ein besonders hoher Zusammenhang gefunden werden. Betrachtet man die Mittelwerte der vier Untersuchungsbedingungen, so zeigt sich auch hier, wie in der Untersuchung von Herle, eine Abnahme der Verwendung von Uhr, Kalender und Telefon zu Messzeitpunkt zwei. Insgesamt muss für das Verfahren ILICA insofern festgestellt werden, dass es das geforderte Mindestmaß an Stabilität nicht erreicht wird.

Vergleicht man die Ergebnisse der beiden indirekten Verfahren, so zeigt sich, dass ILICA insgesamt eine höhere Stabilität aufweist, liegen doch die meisten Ergebnisse der einzelnen Skalen immer noch über dem Korrelationskoeffizienten des Selbstwert-IAT.

Ein Teil der Stichprobe erhielt bei der Bearbeitung von ILICA, bei der zweiten Vorgabe, die Aufforderung das eigene Verhalten bei Durchführung des Verfahrens zu variieren. Ziel war es, eine mögliche Erklärung für die bereits von Herle (1999) attestierte, teilweise eingeschränkte Stabilität der Skalenwerte zu finden. Während Herle von situativen Einflüssen als Ursache ausging, soll dieser Ansatz untersuchen, ob Personen ihr Verhalten bei zweiter Testvorgabe variieren. In der hier vorliegenden Untersuchung konnten keine signifikanten Unterschiede in den betreffenden Skalen zu Messzeitpunkt zwei, zwischen den beiden Untersuchungsbedingungen festgestellt werden. Zusätzlich wurde ein Vergleich der Korrelationskoeffizienten beider Untersuchungsgruppen durchgeführt, hier zeigte sich nur ein signifikanter Unterschied innerhalb der Skala Alltagsaktivitäten. Dennoch konnte beobachtet werden, dass die Korrelationskoeffizienten, auch wenn die Unterschiede nicht signifikant sind, in der Gruppe mit Manipulation von Verhaltensvielfalt niedriger sind, als in der Bedingung ohne Manipulation. Da sich die beiden Untersuchungsgruppen kaum unterscheiden, spricht dies dafür, dass auch diejenigen Personen, die nicht dazu aufgefordert wurden, ihr Verhalten bei der zweiten Vorgabe des Verfahrens verändern. Worauf diese Verhaltensänderung genau zurückzuführen sein könnte, kann aufgrund der hier vorliegenden Daten nicht beantwortet werden und sollte insofern Gegenstand zukünftiger Untersuchungen sein. Eine mögliche, plausible Erklärung ist, dass aufgrund der Bekanntheit des Verfahrens neue Strategien angewendet werden und unterschiedliche Verhaltensweisen ausprobiert werden, auch wenn keine konkrete Aufforderung dazu erfolgt.

Sowohl ILICA als auch der Selbstwert-IAT wurden von den TestteilnehmerInnen relativ positiv bewertet. So wurden bei ILICA die Noten „mangelhaft“ und „nicht genügend“ nur dreimal vergeben. Der IAT wurde nur fünfmal „mangelhaft“ oder „nicht genügend“ bewertet.

Der Selbstwert-IAT wurde bezüglich der Kontrollierbarkeit sehr hoch eingeschätzt, was bedeutet, dass die Testpersonen die Aufgabenstellung allgemein sehr gut verstanden haben. Die Augenscheinvalidität, die Übereinstimmung mit dem vom Teilnehmer vermuteten Ziel der Testung und dem von Testleiter angegebenen, und die Messqualität, also die Einschätzung der TeilnehmerInnen ob (Leistungs-)Unterschiede vom Selbstwert-IAT adäquat abgebildet werden, wurden eher mittelmäßig bewertet. Bezüglich der Belastungsfreiheit beurteilten die TeilnehmerInnen den Selbstwert - IAT im oberen Bereich.

Das Computersimulationsverfahren ILICA wurde ähnlich bewertet. Kontrollierbarkeit wurde eher hoch eingeschätzt, Belastungsfreiheit, also die Beurteilung ob eine Unter- oder Überforderung vorliegt, wurde durchschnittlich positiv bewertet, Augenscheinvalidität und Messqualität liegen im mittleren Bereich.

Vergleicht man nun die beiden Verfahren, zeigen sich in den Skalen Belastungsfreiheit, Augenscheinvalidität und Kontrollierbarkeit signifikante Unterschiede. Der Selbstwert-IAT wurde bezüglich der Kontrollierbarkeit besser als ILICA bewertet. Die Aufgabenstellung und Instruktion wurde von den TeilnehmerInnen beim IAT besser verstanden. Dies kann möglicherweise an der unterschiedlichen Aufgabenstellung und auch Aufbereitung der Verfahren liegen, sind doch beim IAT die Anweisungen sehr konkret vorgegeben (man soll schnell und möglichst genau arbeiten), auch die Aufgabe ist relativ klar umrissen, besteht sie doch nur darin Wörter zuzuordnen. Das Verfahren ILICA hingegen zeichnet sich dadurch aus, den Teilnehmern viele Handlungsfreiheiten zu lassen und ist insofern auch komplizierter (Handhabung der Uhr, des Telefons, Vielzahl an Aktivitäten aus welchen man frei wählen kann). Bezüglich der Augenscheinvalidität wurde das Verfahren ILICA signifikant besser bewertet. Die TeilnehmerInnen beurteilten also die Übereinstimmung des vermuteten Zieles der Testung mit dem angegebenen Ziel bei ILICA als höher. Das indirekte Verfahren ILICA zeichnet sich durch eine hohe Realitätsnähe aus, während der Selbstwert-IAT als Verfahren zur Messung des impliziten Selbstwerts relativ abstrakt erscheint, der gefundene Unterschied bezüglich der Einschätzung der beiden Verfahren liegt also relativ nahe. Auch bezüglich des Kriteriums Belastungsfreiheit wurde ILICA besser eingeschätzt, wobei der Unterschied hier eher gering ist. Die Vielzahl an Durchgängen des IAT, die für

TestteilnehmerInnen höchst wahrscheinlich etwas ermüdend ist, könnte Ursache dieser Bewertung sein, ILICA hingegen ist eher abwechslungsreich gestaltet.

Intention dieser Arbeit war es festzustellen, ob sich der IAT zur Fallbehandlung eignet. Zu diesem Zweck wurde die Retest-Reliabilität ermittelt und Akzeptanz, mit Hilfe des Fragebogens AKZEPT!-L erhoben. Der Selbstwert-IAT sollte mit dem Computersimulationsverfahren ILICA verglichen werden. Zusätzlich wurde eine Manipulation von Verhaltensvielfalt bei ILICA vorgenommen.

Zusammenfassend kann gesagt werden, dass sich der Selbstwert-IAT zur Einzelfallbehandlung nicht eignet. Die hier vorliegenden Ergebnisse zur Stabilität dieses Verfahrens sprechen dagegen, differenzierte Aussagen auf Ergebnissen dieses Verfahrens aufzubauen. Kritisch zu betrachten ist dahingehend auch, dass nach wie vor nicht klar ist, welche Prozesse den IAT-Effekt eigentlich beeinflussen. Welche Faktoren für die niedrige Retest-Reliabilität verantwortlich sind, sollten zukünftige Untersuchungen klären. Für diese Untersuchung kann insofern festgestellt werden, dass situative Einflüsse weitestgehend ausgeschlossen werden können, da es sich um eine unmittelbare Testwiederholung handelte, vielmehr scheinen Übungs- und Lerneffekte den niedrigen Korrelationskoeffizienten mitbedingt zu haben. Die diesbezüglichen Ergebnisse von ILICA waren eindeutig besser, dennoch konnten auch von diesem Verfahren die angestrebten Stabilitätswerte nicht in vollkommen erreicht werden. Auch ILICA wird insofern nur als Screeninginstrument empfohlen.

Die Ergebnisse beider Verfahren bezüglich der Akzeptanz waren durchaus positiv, Befürchtungen, dass aufgrund der Undurchschaubarkeit des Messprinzips die Akzeptanz dieser Verfahren beeinträchtigt wäre, konnten entkräftet werden.

12. Zusammenfassung

Die hier vorliegende Arbeit hat sich zum Ziel gesetzt einen Beitrag zu der Frage zu leisten, ob sich Implizite Assoziationstasks zum Einsatz in der Fallbehandlung eignen. Der Selbstwert-IAT wurde mit dem Computersimulationsverfahren ILICA bezüglich der Retest-Reliabilität und der Akzeptanz verglichen. Beide hier zum Einsatz gekommenen Instrumente sind indirekte Verfahren. Zusätzlich wurde eine Manipulation von Verhaltensvielfalt bei ILICA vorgenommen.

Der erste Teil dieser Arbeit widmet sich der theoretischen Auseinandersetzung mit der Qualität indirekter Verfahren, Objektiver Persönlichkeitstests, impliziter Messmethoden und Simulationsverfahren. Gemeinsam ist indirekten Verfahren, dass sie Persönlichkeitseigenschaften nicht über direktes Nachfragen, wie beispielsweise herkömmliche Fragebogen, erfassen, sondern indirekt, theoriegeleitet, aufgrund experimentalpsychologischer Verhaltenserschassung. Charakterisiert sind derartige Verfahren durch die Undurchschaubarkeit des Messprinzips, die eine Verfälschung durch TeilnehmerInnen verhindern soll.

Bisherige Untersuchungen verglichen meist indirekte Verfahren mit expliziten Techniken. Die hier vorliegende Arbeit hingegen widmete sich dem Vergleich zweier indirekter Verfahren, die unterschiedliche Techniken verwenden und auch verschiedenen Untergruppen indirekter Verfahren zuzuordnen sind.

Der Implizite Assoziationstask ist ein nicht standardisiertes, implizites Verfahren, Ziel ist es die Stärke von Assoziationen zwischen Einstellungsobjekten und Bewertungen im Gedächtnis zu prüfen. In dieser Arbeit kam ein Selbstwert-IAT zum Einsatz, er soll ein Maß impliziter Selbstwertschätzung liefern.

ILICA ist ein Computersimulationsverfahren, ein Objektiver Persönlichkeitstest im engeren Sinne. Mit Hilfe dieses Verfahrens soll Selbstverwaltungs-kompetenz erfasst werden.

Um die Akzeptanz dieser beiden Verfahren zu erfassen kam ein Fragebogen, der AKZEPT!-L zum Einsatz. Der Akzeptanzfragebogen erfasst das individuelle Erleben und die Bewertung des in Frage stehenden Verfahrens durch die TeilnehmerInnen.

Der empirische Teil umfasst die Beantwortung der drei Hypothesen. Untersucht wurde die Retest-Reliabilität von ILICA und IAT, der Einfluss einer Manipulation von Verhaltensvielfalt auf die Ergebnisse von ILICA, sowie die Akzeptanz der beiden Verfahren.

Die bisherigen Ergebnisse bezüglich der Retest-Reliabilität beider Verfahren lagen unterhalb festgelegter Standards. Für den Selbstwert-IAT lagen die Ergebnisse zwischen $r = .54$ (Rudolph et al., 2008) und $r = .69$ (Bosson et al., 2000), für die Skalen von ILICA zwischen $r = .28$ und $r = .66$ (Herle, 1999). Durch die unmittelbare Vorgabe sollten situative Einflüsse möglichst gering gehalten werden, da von verschiedenen Autoren spekuliert wurde, sie könnten Grund für die geringen Stabilitätswerte des IAT sein (Bosson et al., 2000, Gawronski, 2006).

Bezüglich der Akzeptanz der beiden Verfahren gab es noch keine Ergebnisse. Fraglich war, ob die Undurchschaubarkeit des Messprinzips die Akzeptanz negativ beeinflussen könnte (Ortner et al., 2007).

Für die Manipulation von Verhaltensvielfalt für den Test ILICA, lagen noch keine Ergebnisse aus der Literatur vor. Es sollte geklärt werden, ob neben situativen Einflüssen und Lerneffekten auch Verhaltensvariationen ein Grund für niedrigere Retest-Reliabilitäten sein könnten.

Die hier untersuchte Stichprobe setzte sich aus SchülerInnen der siebenten und achten Klassen (elfte und zwölfte Schulstufe) aus drei Gymnasien in Niederösterreich zusammen. In die Analyse gingen insgesamt die Ergebnisse von 157 SchülerInnen ein. Die Ergebnisse zeigten schlechte Retest-Reliabilitäten des Selbstwert-IAT ($r = .377$). Die diesbezüglichen Ergebnisse von ILICA waren zwar etwas besser ($r = .298 - .803$), konnten allerdings das im Vorhinein festgesetzte Kriterium eines Korrelationskoeffizienten von $.70$ auch nicht erreichen.

Die Manipulation von Verhaltensvielfalt bei ILICA zeigte kaum Wirkung, es konnte nur innerhalb einer Skala (Alltagsaktivitäten) ein signifikanter Unterschied der Korrelationskoeffizienten festgestellt werden, der Mittelwertsvergleich lieferte kein signifikantes Ergebnis.

Die Ergebnisse des Akzeptanzfragebogens zeigten, dass sowohl ILICA als auch IAT durchaus positiv von den TeilnehmerInnen bewertet wurden. Es wurden signifikante Unterschiede in der Bewertung der beiden Verfahren in den Skalen Augenscheinvalidität, Kontrollierbarkeit und Belastungsfreiheit gefunden. ILICA wurde bezüglich Belastungsfreiheit und Augenscheinvalidität besser bewertet. In der Skala Kontrollierbarkeit schnitt der Selbstwert-IAT besser als das Computersimulationsverfahren ab.

Zusammenfassend kann gesagt werden, dass der Selbstwert-IAT sich, trotz seiner guten Ergebnisse bezüglich der Akzeptanz, nicht zum Einsatz in der Fallbehandlung eignet. Der Stabilitätswert war, trotz einer unmittelbaren Wiederholung, sehr niedrig und spricht insofern nicht dafür, dass durch dieses Verfahren ein situationsüberdauerndes Merkmal gemessen wird.

13. Literaturverzeichnis

Amelang, M. & Bartussek, D. (1990). *Differentielle Psychologie und Persönlichkeitsforschung*. Stuttgart: W. Kohlhammer.

Asendorpf, J. B., Banse, R. & Mücke, D. (2002). Double dissociation between implicit and explicit personality self-concept: The case of shy behavior. *Journal of personality and social psychology*, 83, 380-393.

Banse, R., Seise, J. & Zerbes, N. (2001). Implicit attitudes towards homosexuality: Reliability, validity, and controllability of the IAT. *Zeitschrift für experimentelle Psychologie*, 48, 145-160.

Bortz, J. & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler*. Berlin-Heidelberg: Springer.

Bosson, J. K., Swenn, W. B. & Pennebaker, J. B. (2000). Stalking the perfect measure of implicit Self-Esteem: The blind men and the elephant revisited? *Journal of Psychology*, 79, 631-643.

Brendl, C. M., Markman, A. B. & Messner, C. (2001). How do indirect measures of evaluation work? Evaluating the inference of prejudice in the Implicit Association Test. *Journal of Personality and Social Psychology*, 81, 760-773.

Cademo 3.20 [Computer Software]. (2009). Rostock: BioMath GmbH.

Cattell, R. B. (1957). *Personality and motivation structure and measurement*. Yonkers-on-Hudson, New York: World Book Company.

Cohen, J. (1992): A power primer. *Psychological Bulletin*, 112, 155-159.

De Houwer, J. (2001). A structural and process analysis of the implicit association test. *Journal of Experimental Social Psychology*, 37, 443-451.

De Houwer, J. (2003). The extrinsic affective Simon task. *Experimental Psychology*, 50, 77–85.

De Houwer, J., & Eelen, P. (1998). An affective variant of the Simon paradigm. *Cognition and Emotion*, 12, 45–61.

Degner, J. Wentura, D. & Rothermund, K. (2006). Indirect Assessment of Attitudes with Response-Time-Based Measures Chances and Problems. *Zeitschrift für Sozialpsychologie*, 37, 131-139.

Egloff, B. & Schmukle, S. C. (2002). Predictive validity of an implicit association test for assessing anxiety. *Journal of personality and social psychology*, 83, 1441 -1455.

Eysenck, H. J. (1960). *Structure of human personality*. London: Methuen.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, 50, 229–238.

Fiedler, K. & Blümke, M. (2005). Faking the IAT: Aided and unaided response control on the Implicit Association Tests. *Basic and applied social psychology*, 27, 307 -316.

Funke, J. (1999). ILICA. Ein Simulationstest zur Erfassung des Entscheidungsverhaltens. In E. Fay (Hrsg.), *Tests unter der Lupe II* (S. 61-75). Lengerich Pabst Science Publishers.

Gawronski, B. (2006). Die Technik des Impliziten Assoziationstests als Grundlage für Objektive Persönlichkeitstests. In T. M. Ortner, R. Proyer & K. D. Kubinger (Hrsg.), *Theorie und Praxis Objektiver Persönlichkeitstests* (S. 53-69) Bern: Huber.

Gawronski, B. & Conrey, F. R. (2004). Der Implizite Assoziationstest als Maß automatisch aktivierter Assoziationen: Reichweite und Grenzen. *Psychologische Rundschau*, 55, 118–126.

Greenwald, A. G. & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self esteem, and stereotypes. *Psychological Review*, 102, 4-27.

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A. & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and selfconcept. *Psychological Review*, 109, 3-25.

Greenwald, A. G. & Farnham, S. D. (2000). Using the Implicit Association Test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology*, 79, 1022-1038.

Greenwald, A. G., McGhee, D. E. & Schwartz, J. K. L. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, 74, 1464-1480.

Greenwald, A. G., Nosek, B. A. & Banaji, M. R. (2003). Understanding and Using the Implicit Association Test: I. An Improved Scoring Algorithm. *Journal of Personality and Social Psychology*, 85, 197–216.

Havlicek, L. L. & Peterson, N. L. (1977). Effect of the violation of assumptions upon significance levels of the pearson r. *Psychological Bulletin*, 84, 373-377.

Herle, M. (1999). *Zur Reliabilität von Psychologischen Testergebnissen in Simulationsverfahren am Beispiel von ILICA und Mailbox '90*. Unveröffentlichte Diplomarbeit, Universität Wien.

Hofman, W., Gawronski, B., Gschwendner, T., Le, H. & Schmitt, M. (2005). A meta analysis on the correlation between the implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin*, 31, 1369-1385.

Inquisit 3.0.4.0 [Computer Software]. (2010). Seattle, WA: Millisecond Software LCC.

Kersting, M. (2008). Zur Akzeptanz von Intelligenz- und Leistungstests. *Report Psychologie*, 33, 420-433.

Kersting, M. (2009). *AKZEPT! Fragebogen zur Messung der Akzeptanz diagnostischer Verfahren*. Verfügbar unter: <http://www.kersting-internet.de/akzept.html> [05.01.2010]

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

Kubinger, K. D. (2003). Objektiver Persönlichkeitstest. In: K. D. Kubinger & R. S. Jäger (Hrsg.). *Schlüsselbegriffe der Psychologie* (S. 304-309). Weinheim: Beltz.

Kubinger, K. D. (2006). Ein Update der Definition von Objektiven Persönlichkeitstests: experimentalpsychologische Verhaltensdiagnostik. In T. M. Ortner, R. Proyer & K. D. Kubinger (Hrsg.), *Theorie und Praxis Objektiver Persönlichkeitstests* (S. 38-52). Bern: Huber.

Kubinger, K. D. (2009). *Psychologische Diagnostik – Theorie und Praxis psychologischen Diagnostizierens*. Göttingen: Hogrefe.

Kubinger, K. D. & Rasch, D. (2006). *Statistik für das Psychologiestudium*. München: Elsevier GmbH.

Kubinger, K. D., Rasch, D. & Šimeckova, M. (2007). Testing a correlation coefficient's significance: Using $H_0: 0 < \rho \leq \lambda$ is preferable to $H_0: \rho = 0$. *Psychology Science*, 49, 74-87.

Maslo, R. (1994). *Validierung von ILICA – Eine Computersimulation zur Erfassung des Konzepts „Selbstverwaltung“*. Unveröffentlichte Diplomarbeit, Wien, Universität Wien.

Mierke, J. (2004). *Kognitive Prozesse bei der indirekten Messung von Einstellungen mit dem Impliziten Assoziationstest*. Unveröffentlichte Dissertation, Universität: Freiburg.

Mierke, J. & Klauer, K. C. (2001). Implicit association measurement with the IAT: Evidence for effects of executive control processes. *Zeitschrift für Experimentelle Psychologie*, 48, 107-122.

Möseneder, D. (1992). *Entwicklung eines Tests zur Erfassung der Selbstverwaltungsfähigkeit*. Unveröffentlichte Diplomarbeit, Wien, Universität Wien.

Möseneder, D. & Ebenhö, J. (1996). *ILICA: Ein Simulationstest zur Erfassung des Entscheidungsverhaltens*. Frankfurt: Swets Test Services.

Murray, H. A. (1971). *Thematic apperception test*. Cambridge Massachusetts: Harvard University Press.

Neumann, R., Hulsenbeck, K. & Seibt, B. (2004). Attitudes towards people with AIDS and avoidance behavior: Automatic and reflective bases of behavior. *Journal of experimental social psychology*, 40, 543 -550.

Nosek, B. A. & Banaji, M. R. (2001). The Go/No-go Association Task. *Social Cognition*, 19, 625-666.

Ortner, T. M., Horn, M., Kersting, M., Kubinger, K. D., Schmidt-Atzert, L., Proyer, R. T., Schuhfried, G., Schütz, A., Wagner-Menghin, M. & Westhoff, K. (2007). Standortbestimmung und Zukunft Objektiver Persönlichkeitstests. *Report Psychologie*, 32, 60-69.

Ortner, T. M. & Janous, G. (2006). Zur Nützlichkeit des Einsatzes eines Objektiven Persönlichkeitstests für die Auswahl von Fachhochschulstudierenden. In T. M. Ortner, R. T. Proyer & K. D. Kubinger (Hrsg.), *Theorie und Praxis Objektiver Persönlichkeitstests* (S. 88-101). Bern: Huber.

Ortner, T. M., Kubinger, K. D., Schrott, A., Radinger, R. & Litzenberger, M. (2006). *Belastbarkeits-Assessment: computerisierte Objektive Persönlichkeits- Testbatterie - Deutsch (BAcO-D)*. Frankfurt/M.: Harcourt: Test Services.

Paulhus, D. L. (1984). Two component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46, 598-609.

Pawlik, K. (2006). Objektive Tests in der Persönlichkeitsforschung. In T. M. Ortner, R. Proyer & K. D. Kubinger (Hrsg.), *Theorie und Praxis Objektiver Persönlichkeitstests* (S. 16-23). Bern: Huber.

QEMU 0.12.3 [Computer Software]. (2010). Fabrice Bellard. Verfügbar unter: <http://wiki.qemu.org/Download>

Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Newbury Park, CA: Sage.

Rothermund, K. & Wentura, D. (2001). Figure-ground asymmetries in the Implicit Association Test (IAT). *Zeitschrift für Experimentelle Psychologie*, 48, 94-106.

Rudolph, A. (2009). *Verfahren zur Erfassung impliziter Selbstwertschätzung. Psychometrische Eigenschaften und die Möglichkeit der Vorhersage von ängstlichem, selbstsicherem und defensive Verhalten*. Dissertation, Technische Universität: Chemnitz. Verfügbar unter: <http://archiv.tu-chemnitz.de/pub/2009/0151>

Rudolph, A., Schröder-Abé, M., Riketta, M. & Schütz, A. (2010) Easier when done than said! Implicit self-esteem predicts observed on spontaneous behaviour, but not self reported or controlled behaviour. *Zeitschrift für Psychologie*, 218, 12-19.

Rudolph, A., Schröder-Abé, M. & Schütz, A. (2006). Ein impliziter Assoziationstest zur Erfassung von Selbstwertschätzung. In Ortner T. M., Proyer R. T. & Kubinger K. D. (Hrsg.), *Theorie und Praxis Objektiver Persönlichkeitstests* (S. 153-163). Bern: Huber.

Rudolph, A., Schröder-Abé, M. & Schütz, A. (2008). Reaktionszeitbasierte Testverfahren in der Persönlichkeitsdiagnostik – ein impliziter Assoziationstest zur Erfassung von Selbstwertschätzung. In: W. Sarges & D. Scheffer (Hrsg.), *Innovative Ansätze für die Eignungsdiagnostik* (S. 197-206). Göttingen: Hogrefe.

Rudolph, A., Schröder-Abé, M., Schütz, A., Gregg, A. & Sedikides, C. (2008). Through a glass less darkly? Reassessing convergent and discriminant validity in measures of implicit self-esteem. *European Journal of Psychological Assessment*, 24, 273-281.

Schlenker, B. R. & Weigold, M. F. (1992). Interpersonal Processes involving impression regulation and management. *Annual Review of Psychology*, 43, 133 -168.

Schmidt, L. R. (2006). Objektive Persönlichkeitstests in der Tradition Cattells: Forschungslinien und Relativierungen. In T. M. Ortner, R. Proyer & K. D. Kubinger (Hrsg.), *Theorie und Praxis Objektiver Persönlichkeitstests* (S. 24-37). Bern: Huber.

Seiwald, B. B. (2002). Replicability and generalizability of Kubinger's results: Some more studies on faking personality inventories. *Psychologische Beiträge*, 44, 17-23.

Steffens, M. C. (2004). Is the Implicit Association Test Immune to Faking? *Experimental Psychology*, 51, 165-179.

Stieger, S. (2006). *Testgütekriterien des impliziten Assoziationstest (IAT) und des single-target impliziten Assoziationstest (STIAT): Einflüsse auf den IAT Effekt sowie die Retest-Reliabilität*. Unveröffentlichte Dissertation, Universität Wien.

Windows-XP, Windows NT Version 5.1 [Computersoftware] (2001). Redmond, WA: Microsoft Corporation One Microsoft Way.

Winter, D. G. & Stewart, A. J. (1977). Power Motive Reliability as a Function of Retest Instructions. *Journal of Consulting and Clinical Psychology*, 45, 436-440.

Zimmerhofer, A. (2008). *Studienberatung im deutschen Hochschulsystem auf der Basis psychologischer Tests: Studienfachprofile, Vorhersagevalidität und Akzeptanz*. Dissertation, Aachen, Rheinisch-Westfälische Technische Hochschule. Verfügbar unter: http://darwin.bth.rwth-aachen.de/opus3/volltexte/2008/2418/pdf/Zimmerhofer_Alexander.pdf [23.10.2010]

14. Abbildungsverzeichnis

Abbildung 1: IAT Bildschirm in Block 1	43
Abbildung 2: IAT Bildschirm in Block 2	44
Abbildung 3: IAT Bildschirm in Block 3	44
Abbildung 4: IAT Bildschirm in Block 4	45
Abbildung 5: IAT Bildschirm in Block 5	45

15. Tabellenverzeichnis

Tabelle 1: Einteilung der Gruppen	41
Tabelle 2: Übersichtsdarstellung der Blöcke des IAT, Items und Trialanzahl	46
Tabelle 3: Verteilung der männlichen und weiblichen Teilnehmer auf die einzelnen Untersuchungsbedingungen	57
Tabelle 4: Mittelwerte und Standardabweichungen des Selbstwert-IAT für die einzelnen Untersuchungsgruppen zu beiden Messzeitpunkten	58
Tabelle 5: Mittelwerte und Standardabweichungen von ILICA für die einzelnen Untersuchungsgruppen zu beiden Messzeitpunkten	59
Tabelle 6: Korrelationskoeffizient Selbstwert-IAT	60
Tabelle 7: Korrelationskoeffizienten der Skalen von ILICA (n = 80)	62
Tabelle 8: Korrelationskoeffizienten der Skalen von ILICA, eingeschränkte Stichprobe (n = 47)	62
Tabelle 9: Manipulation von Verhaltensvielfalt Mann-Whitney U-Test	64
Tabelle 10: Manipulation von Verhaltensvielfalt, Vergleich der Korrelationskoeffizienten	65
Tabelle 11: Median der Skalen des AKZEPT!-L für ILICA und IAT	66
Tabelle 12: Globale Bewertung ILICA	66
Tabelle 13: Globale Bewertung IAT	67
Tabelle 14: Ränge AKZEPT!-L	68
Tabelle 15: Wilcoxon-Signed Rank Test AKZEPT!-L	68
Tabelle 16: Effektgrößen AKZEPT!-L ILICA und IAT	69

16. Anhang

16.1. Schreiben Landesschulrat

Elena Simon

Paulusplatz 9/1/16

1030 Wien

Tel.nr.: 0676/965 65 35

E – Mail: elena_simon@yahoo.de

An den Landesschulrat für Niederösterreich

Rennbahnstraße 29

3109 St. Pölten

Betreff: Genehmigung der Durchführung einer empirischen Untersuchung im
Rahmen der Diplomarbeit.

Wien, am 01.03.2010

Sehr geehrte Damen und Herren,

Ich bin Studentin der Psychologie an der Universität Wien. Derzeit schreibe ich meine
Diplomarbeit im Arbeitsbereich psychologische Diagnostik bei Herrn Univ.- Prof. Kubinger.
Das Thema meiner Arbeit ist:

„Zur Praktikabilität des IAT in der Fallbehandlung –
Machbarkeit, Akzeptanz und wiederholte Retest-Reliabilität im Vergleich zu ILICA“.

Bei den Tests IAT (Impliziter Assoziationstest) und ILICA handelt es sich um Computertests.

Der IAT erfasst den impliziten Selbstwert mittels der Prüfung automatischer Assoziationen.

ILICA ist ein interaktives Computersimulationsverfahren. Ziel dieses Tests ist es, die so genannte Selbstverwaltungskompetenz, die Fähigkeit mittel- und längerfristige Ziele erfolgreich zu erreichen und damit konkurrierende Ablenkungen von außen widerstehen zu können, zu erfassen.

Beide Verfahren sind indirekte Verfahren, diese Tests zeichnen sich durch eine Undurchschaubarkeit des Messprinzips aus, persönliche Stilmerkmale werden aus beobachtbarem Verhalten bei bestimmten Anforderungen erschlossen.

Akzeptanz wird mittels des Fragebogens AKZEPT! –L erfasst.

Die Vorgabe der Tests erfolgt unmittelbar hintereinander, wobei jedes Verfahren zweimal vorgegeben wird, da es sich um eine Untersuchung zur Retest-Reliabilität handelt. Nach der ersten Vorgabe jedes Tests soll jeweils einmal der Fragebogen AKZEPT! vorgegeben werden.

Ziel der Untersuchung ist es festzustellen ob sich der IAT auch zum Einsatz als Individualverfahren eignet, da er bisher nur als Screeningverfahren eingesetzt wird. Von zusätzlichem Interesse ist die Resonanz einer jungen Klientel auf den Modus der Objektiven Persönlichkeitstests.

Die Testdurchführung würde in etwa 120 Minuten dauern (inklusive Rückmeldung). Alle Testergebnisse sind anonymisiert und nicht zurück verfolgbar.

Im Rahmen der Untersuchung wird den Schülern im Psychologie und Philosophie Unterricht ein Einblick in psychologische Forschung und Verfahren gewährt, sofern dies von den Schulen gewünscht wird.

Weiters gibt es auf Wunsch die Möglichkeit einer mündlichen Rückmeldung der persönlichen Ergebnisse.

Die Untersuchung wird an SchülerInnen der 7. Und 8. Klasse (11. und 12. Schulstufe) voraussichtlich an folgenden Schulen durchgeführt:

Bundesrealgymnasium Wiener Neustadt, Gröhrmühlgasse
Gröhrmühlgasse 27
A-2700 Wiener Neustadt
Telefonnr.: 2622 23115

Bundesgymnasium Babenbergerring

Babenbergerring 10
A-2700 Wiener Neustadt
Telefonnr.: 02622 / 22380

Gymnasium und Realgymnasium Sachsenbrunn
Schulstiftung der Erzdiözese Wien
2880 Kirchberg am Wechsel
Telefonnr.: 02641 2202

Klemens Maria Hofbauer Gymnasium
Eichbühler Straße 97
2801 Katzelsdorf / Leitha
Telefonnr.: 02622-78402

Oberstufenrealgymnasium Ternitz
F. Lichtenwörther-Gasse 1
2630 Ternitz
Telefonnr.: 0664 / 5145336

Umfassende Beschreibungen der Verfahren IAT und ILICA, den AKZEPT! – L Fragebogen, das Informationsblatt für Eltern und die Einverständniserklärung sowie die schriftliche Bestätigung des Instituts zur Diplomarbeit entnehmen Sie bitte dem Anhang.

Ich versichere, dass die Daten der Testung anonymisiert werden und die Tests keinerlei die Privatsphäre der Testpersonen verletzende Fragen enthalten.

Für weitere Informationen stehe ich natürlich gerne zur Verfügung.

Ich hoffe auf eine rasche, positive Antwort.

Mit freundlichen Grüßen

Elena Simon

16.2. Elternbrief

Sehr geehrte Damen und Herren, Liebe Eltern!

Ich bin Studentin der Psychologie an der Universität Wien. Derzeit schreibe ich meine Diplomarbeit im Arbeitsbereich psychologische Diagnostik. Das Thema meiner Arbeit ist: „Zur Praktikabilität des IAT in der Fallbehandlung – Machbarkeit, Akzeptanz und wiederholte Retest-Reliabilität im Vergleich zu ILICA“.

Die Untersuchung zu diesem Thema wird an SchülerInnen der 7. Und 8. Klasse durchgeführt.

Bei den Tests IAT (Impliziter Assoziationstest) und ILICA handelt es sich um Computertests.

Beide Verfahren sind indirekte Persönlichkeitsverfahren, sie zeichnen sich durch eine Undurchschaubarkeit des Messprinzips aus, persönliche Stilmerkmale werden aus beobachtbarem Verhalten bei bestimmten Leistungsanforderungen erschlossen.

Ein Einblick in psychologische Forschung und aktuelle Verfahren wird den SchülerInnen gewährt.

Weiters gibt es natürlich die Möglichkeit einer Rückmeldung der persönlichen Ergebnisse.

Die Testdurchführung würde in etwa 100 bis 120 Minuten dauern. Alle Testergebnisse sind anonymisiert und nicht zurück verfolgbar.

Der genaue Termin steht noch nicht fest und wird noch bekanntgegeben.

Für weitere Informationen stehe ich unter der E-Mail Adresse:

elena_simon@yahoo.de, gerne zu Ihrer Verfügung.

Mit freundlichen Grüßen

Elena Simon

Hiermit bestätige ich, dass mein Sohn / meine Tochter _____
an der psychologischen Untersuchung teilnehmen darf.

Unterschrift

Abstract – Deutsch

Indirekte Verfahren zur Messungen von Persönlichkeitseigenschaften erfreuen sich wachsender Beliebtheit, häufig werden diese Instrumente mit expliziten Fragebogentechniken verglichen. Diese Arbeit hingegen widmet sich dem Vergleich zweier indirekter Verfahren, einem Impliziten Assoziationstest und einem Computersimulationsverfahren.

Diese Arbeit hat sich zum Ziel gesetzt einen Beitrag zu der Frage zu leisten, ob sich der Implizite Assoziationstest (IAT) zum Einsatz in der Fallbehandlung eignet. Zu diesem Zweck wurden Retest-Reliabilität und Akzeptanz eines Selbstwert-IAT und des Computersimulationsverfahren ILICA (Phantasiename eines Urlaubszieles) erhoben und anschließend verglichen. Zusätzlich wurde eine Manipulation von Verhaltensvielfalt für den Test ILICA bei einem Teil der Stichprobe vorgenommen, um festzustellen ob Variationen im Verhalten beim zweiten Testdurchgang geringe Stabilitätswerte mitverursachen.

Die Daten wurden an einer Stichprobe von 157 SchülerInnen der 7. und 8. Klasse, im Alter von 16 bis 20 Jahren, erhoben. Sowohl ILICA als auch IAT wurden zweimal unmittelbar hintereinander vorgegeben, um situative Einflüsse auszuschließen. Nach dem ersten Durchgang wurde jeweils mit Hilfe eines Fragebogens die Akzeptanz der beiden Verfahren erhoben. Es konnte belegt werden, dass beide indirekten Verfahren den Stabilitätsanforderungen nicht vollkommen genügten. Insgesamt waren die Ergebnisse von ILICA besser. Die Verhaltensmanipulation führte zu keinen signifikanten Unterschieden in den Ergebnissen zu Messzeitpunkt zwei. Die Akzeptanz beider Verfahren war trotz der Undurchschaubarkeit des Messprinzips hoch. ILICA wurde bezüglich der Augenscheinvalidität und Belastungsfreiheit besser als der Selbstwert-IAT eingeschätzt, der IAT hingegen wurde in Bezug auf die Kontrollierbarkeit positiver bewertet.

Insgesamt kann der Selbstwert-IAT nicht zum Einsatz in der Fallbehandlung empfohlen werden, da der niedrige Korrelationskoeffizient der Testergebnisse zu den beiden Messzeitpunkten eine zu niedrige Stabilität attestiert. Verbesserungen des Verfahrens, eine genaue Abklärung, was den IAT – Wert verursacht und insbesondere

welche Ursache die geringe Stabilität dieses Verfahrens hat, sind notwendig, bevor man diese Messmethode zum Einsatz in der Fallbehandlung empfehlen kann.

Schlüsselwörter:

Impliziter Assoziationstest, ILICA, Retest-Reliabilität, Akzeptanz, Manipulation von Verhaltensvielfalt, indirekte Verfahren

Abstract – Englisch

Indirect methods for measuring attributes of personality increase in popularity. Those instruments are often compared with explicit techniques of questionnaires. On the contrary this paper is dedicated to the comparison of two indirect techniques, an implicit association test and a computer simulation technique.

The objective of this work is to make a contribution to the question whether the implicit association test (IAT) is qualified for the application in individual treatment. For this purpose, the retest-reliability and acceptance of a self-esteem IAT and of the computer simulation test ILICA (fancy name of a holiday destination) were collected and subsequently compared. In addition, a manipulation of varieties of behavior was conducted for a part of the sample for ILICA to prove whether varieties in behavior in the second testing cause low stability values.

The data were collected via a sample of 157 pupils of the 7th and 8th grade, in the age of 16 to 20 years. ILICA as well as the IAT were immediately conducted two times in succession to foreclose situational ascendancies. After the first testing of each procedure, the acceptance of the two tests was conducted with a questionnaire. It was indicated that both indirect methods do not entirely meet the stability requirements. Altogether the results of ILICA were better. The manipulation of varieties of behavior didn't entail any significant differences in the results at the second inquiry. The acceptance of both instruments was, in spite of the intransparency of the measurement principle, high. ILICA was assessed better for face validity and freedom of stress than the self – esteem – IAT, the IAT on the other hand was judged more positive for controllability.

All in all, the self-esteem IAT cannot be recommended for use in individual treatment, the low correlation coefficient of the test results of the two times of measurement certifies too low stability. Improvement of the instrument, an exact clarification of what causes the IAT effect and in particular which reason the low stability of the technique has, is necessary before this measuring method can be advised for individual treatment.

Key words:

Implicit associationtest, ILICA, retest-reliability, acceptance, manipulation of varieties in behavior, indirect methods

Lebenslauf

CURRICULUM VITAE

Elena Simon

26.11.1984

österreichische Staatsbürgerin

AUSBILDUNG

1991 – 1993	Volksschule Maria Regina in Wien
1993 – 1995	Volksschule Kirchberg am Wechsel
1995 – 2003	Gymnasium Sachsenbrunn Abschluss mit Matura
seit Herbst 2003	Studium der Psychologie an der Universität Wien

SPRACHKENNTNISSE

- Muttersprache: Deutsch
- Englisch in Wort und Schrift
- Französisch in Wort und Schrift

ARBEITSPRAXIS

August 2000 – August 2010	Ferialarbeit beim internationalen Ludwig Wittgenstein Symposium Tätigkeiten: administrative und organisatorische Aufgaben
Oktober 2004 – Februar 2005 & November 2005 – Februar 2006	Angestellt in der Boutique Majolie in Wien als kaufmännische Angestellte
Juni & Juli 2006	Angestellt als Servierkraft beim Stiegenwirt in Kirchberg am Wechsel
Oktober 2006 - Dezember 2007	Angestellt beim Roten Kreuz in Wien im „Haus Jupiter“ Tätigkeiten: Betreuung von AsylwerberInnen, Ansprechperson bei Problemen, Krisenintervention, unter psychologischer Supervision
März 2008 – September 2010	Angestellt in der Kanzlei „Steuer & Beratung“ Tätigkeiten: Sekretariat, Unterstützung der Lohnverrechnung, Buchhaltung, Bilanzierung, Erstellung von Steuererklärungen in Excel
August 2009 – September 2009	Praktikum in der Justizanstalt Favoriten Tätigkeiten: Leitung des Arbeitsintegrationstrainings
seit September 2010	Jugendbetreuerin im Jugendzentrum KOMET in Traiskirchen Tätigkeiten: Betreuung von Jugendlichen, Ansprechperson bei Problemen, Krisenintervention, Planung und Mitwirkung bei Projekten