



universität
wien

DIPLOMARBEIT

Titel der Diplomarbeit

Everest

Experimentalpsychologisches Verfahren zur Erfassung der
Selbstverwaltungsfähigkeit
Neukonstruktion

Verfasser

Tobias Griepenstroh

Angestrebter akademischer Grad

Magister der Naturwissenschaften (Mag. rer. nat.)

Wien, September 2014

Studienkennzahl lt. Studienblatt: A 298

Studienrichtung lt. Studienblatt: Psychologie

Betreuerin / Betreuer: Mag. Dr. Lale Khorramdel-Ameri

DANKSAGUNG

An erster Stelle möchte ich meinen Eltern danken, die mich stets in all meinen Vorhaben unterstützt haben und ohne die mein Studium nicht möglich gewesen wäre.

Danken möchte ich ebenfalls Frau Mag. Dr. Khorramdel-Ameri, welche mir ermöglicht hat meine eigenen Ideen umzusetzen und mir stets mit Anregung, Kritik und Hilfe zur Seite stand.

Für die Gestaltung der Charaktere möchte ich an dieser Stelle Frau Isabella Kailing Tu danken, die all meine Vorstellungen erfolgreich zu Papier gebracht hat und mich auch bei der Coloration der Charaktere tatkräftig unterstützt hat.

Ganz besonderer Dank gebührt meiner Lebensgefährtin Neele, deren Unterstützung und Fürsorge den Abschluss dieser Arbeit überhaupt erst ermöglicht haben.

Darüber hinaus danke ich den vielen Testpersonen für ihre Teilnahme sowie meinen Freunden für ihren Rückhalt.

INHALTSVERZEICHNIS

EINLEITUNG	9
THEORETISCHER TEIL	11
1 Entscheidungsfindungsprozesse	11
2 Zwei-Prozess-Modelle	14
3 Objektive Persönlichkeitstests	17
4 Computersimulationen	19
5 Testbeschreibungen und Testkonstruktion Everest	22
5.1 ILICA	22
5.1.1 Testbeschreibung ILICA	22
5.1.2 Typisierungen	23
5.1.3 Kritik	24
5.2 Neuerungen im Vergleich Everest – ILICA	25
5.3 Testkonstruktion des Everest	27
5.3.1 Entstehungsgeschichte	27
5.3.2 Bedienung des Everest	29
5.4 Testdurchführung des Everest	30
5.4.1 Trainings- und Instruktionsphase	30
5.4.2 Testitems	36
5.4.3 Testumgebungen und Personen	38
5.4.4 Typisierung des Everest	45
5.5 Testkennwerte	48

EMPIRISCHER TEIL	50
6 Stichprobenbeschreibung	50
7 Itemanalyse	51
8 Skalenanalyse	56
9 Clusteranalyse	59
9.1 Deskriptive Darstellung der Cluster	60
9.2 Weiterführende deskriptive Darstellung	65
9.3 Inhaltliche Beschreibung der Cluster	69
10 Gütekriterien	71
10.1 Skalierung	71
10.2 Reliabilität	71
10.3 Validität	71
10.4 Objektivität	72
10.5 Ökonomie	72
10.6 Eichung	72
10.7 Fairness	73
10.8 Zumutbarkeit	73
10.9 Unverfälschbarkeit	73
11 Diskussion und Ausblicke	74
12 Zusammenfassung	76
13 Literaturverzeichnis	79
14 Anhang	82
14.1 Anhang A – Abstract	82
14.2 Anhang B – Einverständniserklärung	83
14.3 Anhang C – Curriculum Vitae	84

ABBILDUNGSVERZEICHNIS

Abb. 1: Schaltflächen Hilfe und Uhr	31
Abb. 2: Notizblock	32
Abb. 3: Veränderter Mauszeiger bei Interaktionsmöglichkeit	32
Abb. 4: Diensthandy	33
Abb. 5: Schaltfläche „Aufgabenstellung“ am Beispiel Powerpoint-Präsentation	34
Abb. 6: Virtueller Desktop der Testperson während der Testbearbeitung	35
Abb. 7: Büro der Testperson	38
Abb. 8: Tic Tac Toe Spielfeld	39
Abb. 9: Außenansicht Bürogebäude	39
Abb. 10: Gang im unteren Stockwerk	40
Abb. 11: Büro des Vorgesetzten	40
Abb. 12: Kantine	41
Abb. 13: Büro der Kolleginnen	41
Abb. 14: Gang im oberen Stockwerk	42
Abb. 15: Konferenzraum	42
Abb. 16: Kopierraum	43
Abb. 17: Männlicher und weiblicher Avatar	43
Abb. 18: Vorgesetzter Herr Lehmann und Haustechniker Pröll	44
Abb. 19: Kollegium	44
Abb. 20: Screeplot für Skala Impulsiv	57
Abb. 21: Screeplot für Skala Flexibel	57
Abb. 22: Screeplot für Skala Reflexiv	58
Abb. 23: Screeplot für alle 3 Skalen	58
Abb. 24: Bestimmung der Clusteranzahl	60

TABELLENVERZEICHNIS

Tab. 1: Alter	50
Tab. 2: Geschlechterverteilung	50
Tab. 3: Ernsthaftigkeit und Belastung	51
Tab. 4: Antworthäufigkeiten	53
Tab. 5: Itemschwierigkeiten	54
Tab. 6: Itemtrennschärfen	54
Tab. 7: Skalenreliabilität	55
Tab. 8: Skalenreliabilität nach Itemausschluss	55
Tab. 9: Itemtrennschärfen nach Itemausschluss	56
Tab. 10: Leven's Test für die clusterbildenden Variablen	61
Tab. 11: Mittelwerte der Cluster-bildenden Variablen	61
Tab. 12: ANOVA – Impulsiv	63
Tab. 13: Post-Hoc-Test Impulsiv	63
Tab. 14: Welch-Test Ergebnisse Reflexiv, Flexibel, Arbeitsscore	64
Tab. 15: Post-Hoc-Tests Reflexiv, Flexibel, Arbeitsscore	64
Tab. 16: Leven's Test nicht clusterbildende Variablen	66
Tab. 17: ANOVA Ergebnisse nicht clusterbildende Variablen	67
Tab. 18: Post-Hoc-Tests Freizeit und Bearbeitungszeit	68

EINLEITUNG

Die Zielsetzung dieser Arbeit war die Entwicklung eines Objektiven Persönlichkeitstest zur Erfassung der Selbstverwaltungsfähigkeit nach dem Vorbild des Verfahrens ILICA (Möseneder & Ebenhöf, 1996). Der so entstandene Test „Everest“ basiert auf der Grundidee des ILICA, beinhaltet jedoch Items sowie eine neue Umgebung für die Computersimulation. Der Test wurde außerdem an den Arbeitsalltag angepasst, indem statt eines freien Tages ein Arbeitstag in einer Büroumgebung simuliert wird. In der vorliegenden Arbeit wird die Art und Weise in welcher eine Neuprogrammierung, Optimierung sowie Weiterentwicklung des ILICA vorgenommen wurde beschrieben. Weiters werden die Ergebnisse einer Ersterprobung sowie die Gütekriterien des Everest vorgestellt.

Everest steht für „experimentalpsychologische Verfahren zur Erfassung der Selbstverwaltungsfähigkeit“. Die grundlegenden Neuerungen des Everest gegenüber dem ILICA sind die gesteigerte Realitätsnähe sowie das Einbeziehen arbeitsrelevanter Aufgabenstellungen. Es wurde dabei versucht, das eher (v.a. programmiertechnisch) veraltete Design und Konzept des ILICA neu zu gestalten und an die derzeit gegebenen technologischen Möglichkeiten anzupassen. Dabei sollte v.a. die Computersimulation realitätsnäher gestaltet werden und der (Instruktions)Text gekürzt werden.

Ziel ist es, der Testperson eine mittels Avatars frei begehbare, zweidimensionale Umgebung zu bieten in der sie sich frei bewegen und die an sie gestellten Aufgaben in der von ihr präferierten Art und Weise bearbeiten kann. Die Testperson kann sich damit (über den Avatar) durch Raum und Simulation bewegen anstatt, wie im ILICA, nur davon zu lesen. Darüber hinaus wird versucht die drei verschiedenen Entscheidungstypen die bereits im ILICA postuliert wurden über neue Items zu erfassen und mittels Clusteranalyse sowie Exploratorischer Faktorenanalyse zu replizieren und im Hinblick auf ihre Dimensionalität zu analysieren.

Im Folgenden sollen folgende Bereiche ausführlich dargestellt werden:

- Überblick über theoretischer Hintergrund von Entscheidungsmodellen
- Beschreibung des ILICA von Möseneder und Ebenhöf (1996)
- Ausführliche Darstellung des Everest
- Erste testtheoretische Analysen des Everest
- Betrachtung des Everest hinsichtlich der Testgütekriterien.

Theoretischer Teil

1 Entscheidungsfindungsprozesse

„Gegenstand der Psychologie sind Verhalten, Erleben und Bewusstsein des Menschen, deren Entwicklung über die Lebensspanne und deren innere [...] und äußere [...] Bedingungen und Ursachen [...]“ (Zimbardo & Gerrig, 1999, S.2)

Menschliches Verhalten ist hierbei geprägt von Entscheidungen. Im Verlauf eines Tages ist der Mensch mit einer Vielzahl an Situationen konfrontiert, welche ihn vor mehr oder minder schwierige Entscheidungen stellen, die oftmals auch sehr kurzfristig getroffen werden müssen.

Die Bandbreite an Entschlüssen reicht hierbei von trivialen Problemstellungen, wie der Wahl des passenden Kleidungsstückes bis hin zu gravierenden Entscheidungen, die den gesamten weiteren Lebensverlauf eines Individuums beeinflussen können und oftmals auch direkt oder indirekt Auswirkungen auf das engere oder weitere soziale Umfeld des Entscheidungsträgers bzw. der Entscheidungsträgerin haben.

Die menschliche Entscheidungsfindung zu verstehen, zu erklären und in weiterer Folge vorherzusagen, stellte somit stets ein Kerngebiet der psychologischen Forschung dar, welches im Rahmen des Forschungsfeldes „Judgement and Decision Making“ genauer untersucht wird. Dadurch entstand im Laufe der Zeit eine Sammlung an vielfältigen Theorien und Modellen zum Prozess der Entscheidungsfindung.

Vielfach wurde postuliert, dass Menschen sich in ihrem Verhalten generell stets nach objektiven und rationalen Verhaltens- und Entscheidungsmustern richten würden.

Diese Ansicht ruht auf der Annahme, dass Menschen in der Regel Entscheidungen bevorzugen, welche für sie den größtmöglichen Vorteil versprechen und auch die Wahl des Verhaltens stets auf die individuelle Gewinnoptimierung ausgerichtet erfolgt.

Ein solches Modell an Verhaltens- und Entscheidungsmustern erachtet den Menschen somit als rational getriebenes Wesen, welches über die nötigen Kapazitäten verfügt um innerhalb sehr kurzer Zeitintervalle eine Kosten-Nutzen-Analyse hinsichtlich der Konsequenzen seiner gewählten Handlungen aufzustellen.

Diese Annahmen finden sich auch im Menschenbild des Homo oeconomicus. Nach Eduard Spranger (1950) ist der Homo oeconomicus ein Mensch, der stets in allen Lebenssituationen als höchste Priorität den Nützlichkeitswert sieht.

Entscheidend für das Konzept des Homo oeconomicus ist die Grundannahme, dass Menschen in einer Art und Weise Entscheidungen treffen, welche der normativen Nutzentheorie von Von Neumann und Morgenstern (1953) entspricht. Demzufolge treffen Menschen grundsätzlich rationale Entscheidungen auf der Basis des zugrundeliegenden vollständigen Wissens über Alternativen, Werte und Wahrscheinlichkeiten. Der gesamte Entscheidungsprozess fußt somit auf komplexen kognitiven Operationen, welche auf klar definierten mathematischen Regeln basieren.

Beim Treffen von Entscheidungen existiert eine klare Rangfolge der Alternativen, sodass nach Abwägen des Für und Wider stets die Möglichkeit präferiert wird, welche den höchsten persönlichen Nutzen verspricht. Dieses Modell beschreibt somit einen Menschen, der rein wirtschaftlich denkt und handelt und demzufolge zu rein rationalem Verhalten befähigt ist, welches nicht von Emotionen oder sozialen Kontexten beeinflusst wird.

Eine Entscheidungsfindung nach der oben beschriebenen Art und Weise beansprucht jedoch ein Höchstmaß an kognitiven Ressourcen, da von einer fehlerfreien Verarbeitung der Informationen ausgegangen wird. Zudem muss auch die Bearbeitungszeit bei einem Maximum liegen, da das Modell unendliche Anpassungsgeschwindigkeit annimmt, mit der sich der Homo oeconomicus unmittelbar auf strukturelle Veränderungen in seiner Umwelt einstellt.

Eine anschauliche Umschreibung der Grenzen dieser rein rationalen Entscheidungsfindung findet sich bei H. A. Simon in seinem Konzept der „bounded rationality“ (1997).

Demzufolge handelt es sich bei der Entscheidungsfindung keineswegs um einen uneingeschränkt rationalen Prozess. Vielmehr ist davon auszugehen, dass Entscheidungsprozesse immer nur in begrenztem Maß rational getroffen werden können, da der Mensch nicht nur in seinen kognitiven Fähigkeiten und seiner Informationsverarbeitungskapazität beschränkt ist, sodass oftmals nicht alle relevanten Eigenschaften einer Entscheidung mit in die Entscheidungsfindung einfließen. Zudem ist auch das lückenlose Vorhandensein aller nötigen Informationen im Rahmen einer Entscheidungssituation nicht realistisch.

Oftmals erfordern Situationen eine unmittelbare Entscheidung ohne, dass ausreichend Zeit vorhanden ist um die Alternativen gegen einander abzuwägen.

Zusammenfassend führen all diese Komponenten zu einer Erklärung dafür, dass Entscheidungen oftmals anders getroffen werden, als dies unter Idealbedingungen möglich wäre.

2 Zwei-Prozess-Modelle

Basierend auf diesen Erkenntnissen die zeigten, dass Verhalten und Entscheidungen nicht nur auf einem rein rationalen sondern auch auf einem intuitiven Konzept basieren, brachte die jüngste Forschung innerhalb der Psychologie eine Vielzahl an sogenannten Zwei-Prozess-Modellen hervor. Als bekannteste Vertreter gelten das Elaboration-Likelihood-Modell von Petty und Cacioppo und das Heuristisch-Systematische-Modell von Chaiken.

Allen vorhandenen Zwei-Prozess-Modellen liegt die Annahme zugrunde, dass Verhalten multimodal bestimmt ist. Obwohl es viele verschiedene Modelle diesbezüglich gibt, zeigen sie alle inhaltliche Konsistenz durch die Unterscheidung zwei verschiedener Modi an Informationsverarbeitung.

Smith und DeCoster (1999) nennen drei grundlegende Annahmen, die sich in allen bislang publizierten Ansätzen zu Zwei-Prozess-Modellen widerspiegeln: Zum einen erfolgt die Informationsverarbeitung nach einem auf einfachen Regeln basierenden Prozess mit möglichst geringem kognitivem Aufwand. Auf diese Art und Weise erfolgen die meisten Entscheidungsprozesse. Mitunter kommt es jedoch zu speziellen Konstellationen, die einen anderen Weg der Informationsverarbeitung verlangen, bei dem aktiv alle zur Verfügung stehenden Informationen in den Entscheidungsfindungsprozess mit einbezogen werden. Entscheidungen, die so getroffen werden führen zu stabiler verankertem Verhalten.

Die dritte Annahme ist die Abhängigkeit des Informationsverarbeitungsumfangs sowohl von ausreichend kognitiven und sonstigen Kapazitäten als auch von der vorhandenen Eigenmotivation der betroffenen Person, sich mit der vorgegebenen Problemstellung auch in adäquater Weise auseinander setzen zu wollen.

Das Zwei-Prozess-Modell nach Sloman (1996) beschreibt als grundlegende Prozesse den assoziativen und den regelbasierten Informationsverarbeitungsprozess. Bei der assoziativen Verarbeitung kommt es in kurzer Zeit unter geringem kognitivem Aufwand und eher intuitiv zu Entscheidungen. Grundlage der Entscheidungsfindung sind Konzepte, welche zuvor durch wiederholte Erfahrung erlernte Assoziationen mit Schlüsselreizen in der Problemstellung verbinden und so zu einer automatischen Entscheidung auf Grund von Ähnlichkeiten und Mustern führen.

Anders verhält es sich bei der regelbasierten Informationsverarbeitung in deren Verlauf

es unter Anwendung und Befolgung abstrakter, theoriegeleiteter Regeln aus den Bereichen der Logik und Kausalität zu einer Entscheidungsfindung kommt.

Unterschiede in den Modellen zeigen sich vor allem hinsichtlich der Frage, ob die beiden postulierten Informationsverarbeitungsprozesse simultan, sequentiell oder alternativ ablaufen und ob die Wahl des Informationsverarbeitungsprozesses eher von den motivationalen oder den kapazitätsbedingten Gegebenheiten abhängt.

Sloman (1996) betrachtet die beiden Informationsverarbeitungsprozesse als Ergänzungen und weniger als Alternativen. Demzufolge kann ein Mensch während seines Entscheidungsfindungsprozesses zu mehreren möglichen Entscheidungen gelangen, wenn assoziative und regelbasierte Prozesse unterschiedliche Lösungswege bereitstellen.

Strack und Deutsch (2004) beschreiben in Form von 10 Thesen ein Zwei-Prozess-Modell in dem die Interaktion eines reflexiven und eines impulsiven Informationsverarbeitungssystems eine tragende Rolle für soziales Verhalten darstellt. Diese beiden Prozesse arbeiten parallel. Im reflexiven System das Verhalten als Konsequenz eines vorangehenden Entscheidungsprozesses zu sehen ist, bei dem Wahrscheinlichkeiten der Zielerreichung, Konsequenzen und Werte mit einbezogen werden. Dieser Prozess erfordert eine hohe kognitive Kapazität und kommt somit nicht in Situationen zum Tragen, in denen akut und schnell gehandelt werden muss. Hier führt das impulsive System das Verhalten herbei indem auch unter weniger guten Bedingungen rasch und anhand assoziativer Verbindungen Schemata aktiviert werden die zu einem in erster Linie motivational beeinflussten Verhalten führen. Die impulsiven Informationsverarbeitungsprozesse verbrauchen weniger kognitive Kapazitäten jedoch sind sie nicht so wohl überlegt und gut durchdacht wie die reflexiven Prozesse, die auf Schlussfolgerungen und Intentionen beruhen.

Payne, Bettman und Johnson (1993) stellen die Strategiewahl bei multiattributiven Entscheidungen als das Resultat eines Metakalküls dar. Hierbei entscheidet eine Kosten-Nutzen-Analyse unter Einbezug der Struktur der Umwelt und unter Berücksichtigung der Belange des Entscheiders über die Wahl der Strategie. Ein Beispiel für einen Einflussfaktor, der der Umweltstruktur entstammt ist etwa die Anzahl der zur Verfügung stehenden Alternativen während in der Person verankerte Einflussfaktoren etwa die kognitive Kapazität oder der notwendige Aufwand darstellen. Nach Betsch (2004) erfolgt die Wahl des jeweiligen Entscheidungsmodus in Anlehnung

an das Prozessmodell von Strack und Deutsch (2004) sowohl in Abhängigkeit von situativen Gegebenheiten wie der Präsenz von Wahrscheinlichkeiten oder dem Vorhandensein physiologischer Bedürfnisse, als auch aufgrund eines individuellen Personenfaktors, der Ausdruck für eine über vielfache Entscheidungssituationen hinweg erprobte Strategiepräferenz ist. In weiterer Folge steht somit nicht mehr die Wahl einzelner Strategien oder Entscheidungen im Zentrum des Forschungsinteresses, sondern die Anwendung eines Entscheidungsstils aufgrund dessen sich weitere Vorhersagen treffen lassen.

Die vorliegende Arbeit orientiert sich an der Annahme eines individuell präferierten Entscheidungsstils nach Betsch (2004) , sowie am Zwei-Prozess-Modell von Strack und Deutsch (2004), in dem von einem reflexiven und einem impulsiven Informationsverarbeitungsprozess ausgegangen wird. In Anlehnung an den Test ILICA wird neben einem reflexiven und einem impulsiven auch ein flexibler Stil angenommen.

3 Objektive Persönlichkeitstests

Konzeptuell begründet sind Objektive Persönlichkeitstests auf den Ausführungen von Raymond B. Cattell (Pawlik, 2006). Der hierbei postulierte Ansatz bot erstmals eine systematische Herangehensweise an die Messung von Persönlichkeit. Die wesentliche Neuerung stellte die Verbindung verschiedener Quellen der Datenerhebung zur besseren Differenzierung der Persönlichkeit dar.

Als Datenquellen nutzte Cattell Q-Daten, L-Daten und T-Daten. L-Daten lieferten Persönlichkeitskennwerte anhand von Fremdbeurteilungen oder aus dem lebensgeschichtlichen Kontext entnommenen Daten. Q-Daten sind Selbstbeschreibungen der Testperson, die in Form von Persönlichkeitsfragebögen oder Interviewdaten vorhanden sind. Deutlich anders als die Q-Daten liefern die T-Daten, welche durch Verhaltensbeobachtungen oder mittels Leistungstests erhoben werden, Persönlichkeitskennwerte die gegenüber Verzerrungstendenzen stabil sind.

Die Bezeichnung „objektiver Test“ führte Cattell (1958) ein, um eine klare Abgrenzung zwischen Q-Daten und T-Daten zu schaffen. Diese inhaltlich begründete Differenzierung zwischen durch Fragen erhobenen Q-Daten und auf Auswertung erbrachter Leistung basierender T-Daten fand jedoch keinen Einzug in die psychologische Praxis. Nach Cattell (1958) gilt ein Verfahren dann als „objektiver Test“, wenn es sich um eine unter standardisierten Bedingungen durchgeführte und grundsätzlich übertragbare Stimulus-Situation handelt, in der das Verhalten der Testperson gemessen wird ohne, dass diese mit der eigentlichen Messintention vertraut ist sodass es nur in geringem Ausmaß zu Verfälschungstendenzen wie sozial erwünschtem Antwortverhalten oder auch zu systematischen Messfehlern kommen kann.

Obwohl das Konzept der Objektiven Persönlichkeitstests durchaus positiv in der Fachgemeinschaft aufgenommen wurde, ließen neue Forschungsansätze und Neuerungen in diesem Bereich gut drei Jahrzehnte auf sich warten. Gerade der erhöhte Aufwand im Bereich der Testkonstruktion gegenüber herkömmlichen Fragebögen mag die weitere Verbreitung dieser Erhebungstechnik eingeschränkt haben. Erst mit dem Aufkommen neuer Technologien, speziell dem computergestützten Testen, erlebte das Konzept der Objektiven Persönlichkeitstests einen erneuten Aufschwung.

Aufgrund der Möglichkeit Situationen und Handlungsabläufe virtuell zu simulieren wurde es wissenschaftlich wieder adäquat T-Daten zu erheben.

Bei Kubinger (2006) findet sich eine zeitgemäße Definition Objektiver Persönlichkeitstests. „Objektive Persönlichkeitstests versuchen, persönliche Stilmerkmale aus dem beobachtbaren Verhalten bei bestimmten (Leistungs-) Anforderungen zu erschließen, wobei die Registrierung der Art und Weise der Problembearbeitung der Computer übernimmt.“ (Kubinger, 2006, S. 256)

Im Sinne dieser aktuelleren Definition der Objektiven Persönlichkeitstests bildet die Erfassung des Arbeits- und Prozessverhaltens während der Durchführung hypothesengeleiteter Aufgaben den zentralen Forschungspunkt. Nach Kubinger (2009) bietet sich mit der Bezeichnung „Experimentalpsychologische Verhaltensdiagnostik“ eine adäquate Möglichkeit, um den essentiellen Punkt der „systematischen Verhaltensbeobachtung unter experimentell variierten Bedingungen“ im Rahmen solcher Verfahren zu verdeutlichen. (Kubinger, 2009, S. 260)

4 Computersimulationen

Seit Beginn der psychologischen Diagnostik stellte dieser Teilbereich der Psychologie einen Zweig dar, der eng an die technologische Entwicklung gebunden war.

Nach Funke und Reuschenbach (2011) nutzte Schulte bereits 1921 in den Zeiten der „Industriellen Psychotechnik“ zum Zweck der Tauglichkeitsprüfung angehender Damenfriseurinnen eine Apparatur, welche die Zugstärke des Bürstenvorgangs messbar machte. Seit diesen Anfängen hat die Leistungsdiagnostik eine stete Entwicklung durchgemacht bis hin zu aktuellen Verfahren, die sich der momentan verfügbaren Technologien bedienen. Die Schnelllebigkeit des technologischen Wandels sorgt zwar für eine Vielzahl an potenziellen, neuen Verwendungsmöglichkeiten jedoch halten diese nicht immer auch in gleichem Tempo Einzug in die psychologische Diagnostik.

Die Computersimulation stellt einen der vergleichsweise neueren Zweige dieser Entwicklung dar. Erst mit Aufkommen einer gewissen Rechenleistung der Prozessoren neuerer Computer wurden komplexere, interaktive Simulationen überhaupt möglich. Eine Definition des Begriffs Computersimulation geben Funke und Reuschenbach (2011) wonach Computersimulationen versuchen „...dynamische und vernetzte Aspekte der Umwelt, die das Verhalten, die Einstellung oder Emotionen beeinflussen, realitätsnah abzubilden.“ (Funke und Reuschenbach, 2011, S. 602)

Bislang genutzte Testanweisungen in Textform werden durch wirklichkeitsnahe Stimuli ersetzt, sodass ein Kompromiss zwischen Messungen entsteht, welche in situ durchgeführt werden und solchen, die unter der Kontrollierbarkeit einer Laborsituation erfolgen.

Schiff, Arnone und Cross (1994) zitiert nach Funke und Reuschenbach (2011) verstehen unter Realitätsnähe der Computersimulation einerseits die Item-Realität-Passung (display realism), sodass die im Simulationsverfahren verwendeten Items möglichst optimal die reale Umwelt abbilden. Zudem können realistische Bedingungen hergestellt werden indem das Eingreifen seitens der Testperson möglichst realen Abläufen folgt (response-system realism).

Der Vorteil von Computersimulationen liegt in der Komplexität der Anforderungen hinsichtlich derer Reaktionen und Einstellungen analysiert werden. Dies ergibt sich aus dem großen Umfang zusammenspielender akustischer und visueller Stimuli, die für die Testperson nicht zur Gänze in ihrer Messintention transparent sind.

Abzugrenzen sind Computersimulationen von Arbeitsproben, die ebenfalls die Simulation einer Tätigkeit zum Ziel haben. Da hierbei neben den Anforderungen an die Testperson jedoch auch die von ihr geforderten Handlungen realistischer sind, sind Fehlentscheidungen der Testperson nicht nur für diese selbst sondern auch für die durchführenden Personen gravierender, da sie reale Konsequenzen nach sich ziehen.

Im Rahmen einer Computersimulation erfolgt die Auswahl der entsprechenden Reaktion seitens der Testperson hingegen auf Basis vordefinierter Antwortmuster, welche zu vordefinierten Konsequenzen führen, die sich wiederum aus der Natur der Programmierung ergeben. Dies führt zwar zu einer höheren Ökonomie der Computersimulation gegenüber den Arbeitsproben, welche sich jedoch durch die Rigidität der definierten Antwortmuster erkaufte und somit die Lebensnähe des Verfahrens mitunter beschneidet. Die Verwendung des Mediums Computer ermöglicht hierbei jedoch zugleich evidente Vorteile gegenüber anderen Verfahren.

Eine Beschreibung der Anforderungen an eine Computersimulation um eine möglichst realitätsnahe Aufgabenstellung zu gewährleisten, findet sich bei Funke und Reuschenbach (2011). So muss ein gewisses Maß an Komplexität gewährleistet sein, das heißt die Systemvariablen müssen in ausreichend hoher Zahl vorhanden sein um Konsequenzen für die Verarbeitungskapazität zu haben, sodass Maßnahmen zur Reduktion der Komplexität seitens der Testperson notwendig werden. Eine weitere zu erfüllende Anforderung liegt in der Vernetztheit der Variablen. Die multiplen Wirkungsrelationen zwischen den Variablen führen zu netzwerkartiger Effektausbreitung und in weiterer Konsequenz zu konfundierten Wirkungen, Nebenwirkungen und Fernwirkungen.

Da im Rahmen von Computersimulationsverfahren nicht nur durch eingreifende Handlungen der Testperson der Systemverlauf bestimmt wird, sondern sich aufgrund der Programmierung auch eine Eigendynamik ergibt, welche die Testperson unter Zeit- und Handlungsdruck setzt, entstehen hier besondere Anforderungen an die Testperson, die gefordert ist, den Ablauf in der Computersimulation abzuschätzen.

Im Rahmen der Testkonstruktion bietet sich die Möglichkeit der Zeitmanipulation, da Ereignisse in der Computersimulation sowohl in Zeitraffer als auch in Zeitlupe ablaufen können.

Nach Funke und Reuschenbach (2011) ist eine in diagnostischer Hinsicht relevante Anforderung an die Computersimulation die Intransparenz. Gemeint ist damit die Unzugänglichkeit von Informationen über die Systemvariablen, die Variablenverknüpfungen oder die Konsequenzen von Eingriffen entweder aufgrund zeitlicher Einschränkungen oder aufgrund struktureller Intransparenz. Dies erfordert von der Testperson ein Abschätzen der Notwendigkeit von Informationsbeschaffung und Informationsbewertung. Die Intransparenz der Computersimulation erfüllt somit auch die zuvor erwähnte Anforderung Catells (1958) an einen objektiven Test, wonach die eigentliche Messintention für die Testperson unzugänglich ist.

Die letzte Anforderung an eine Computersimulation stellt nach Funke und Reuschenbach (2011) die Polytellie dar, die durch gleichzeitig oder aufeinanderfolgend gegebene, multiple und eventuell gegensätzliche Teilzeile bei der Testperson einen Interessenskonflikt herbeiführt sodass es zu einem Entscheidungsprozess und zu Kompromissbildung kommt.

Kritisch zu betrachten ist nach Funke und Reuschenbach (2011) vor allem der hohe Ressourcenaufwand im Rahmen der Testkonstruktion. Auch muss eine Anforderungsanalyse die Passung zwischen Umwelt und Systemanforderungen sichern, wenn ein möglichst hohes Maß an Realitätsnähe im Vordergrund steht. Zudem ist es wichtig, vorab das Computerwissen der Testpersonen zu kontrollieren, um möglichen Generationseffekten vorzubeugen.

Einen klaren Kritikpunkt in Bezug auf ein adaptives Computersimulationsverfahren sehen Funke und Reuschenbach (2011) in der stets personenspezifischen Bearbeitungsweise, welche den Vergleich einzelner Testwerte erschwert.

„Bereits mit den ersten Eingriffen erzeugt jede einzelne Versuchsperson ihre eigene Umwelt und ist damit nur noch bedingt mit anderen Personen vergleichbar, die das gleiche Szenario bearbeiten.“ (Funke und Reuschenbach, 2011, S.609)

5 Testbeschreibungen und Testkonstruktion des Everest

5.1 ILICA

Bei dem mit dem Phantasienamen ILICA bezeichneten psychologisch-diagnostischen Verfahren handelt es sich um einen computergestützten Simulationstest von Möseneder und Ebenhöf aus dem Jahre 1996.

5.1.1 Testbeschreibung ILICA

Inhaltlich versucht der Test realistische Bedingungen des alltäglichen Lebens bestmöglich virtuell abzubilden und bietet als Rahmenhandlung einen gesamten Tagesablauf, der von der Testperson individuell durch Festsetzen von Freizeitaktivitäten, aber auch durch zuvor festgelegte und während des Testverlaufs aufkommende Aufgabenstellungen gestaltet wird. Dabei erfolgt die Simulation ausschließlich auf der Basis von Textmaterial.

Der ILICA dient der Erfassung der Planungsstrategien der Testperson bei organisatorisch geringfügiger anspruchsvollen Aufgabenstellungen und liefert einen Kennwert für das Ausmaß, in dem es zu einer Ablenkung der Testperson beim Erreichen der angestrebten Zielvorgaben kam.

Das zu Grunde liegende Konstrukt, welches hierbei erfasst wird, ist die sogenannte „Selbstverwaltungsfähigkeit“. Damit ist nach Möseneder und Ebenhöf (1996) „der persönliche Stil einer Testperson, mit Ablenkungen während einer angestrebten Zielerreichung sowie mit aktualisierten Problemen und Frustrationen umzugehen“ gemeint (S. 4). Das Konstrukt der Selbstverwaltungsfähigkeit erscheint aufgrund seiner Komplexität als ungeeignet für die Erhebung mittels herkömmlicher psychologisch-diagnostischer Verfahren, sodass der Einsatz eines computergestützten Verfahrens angebracht ist.

Möseneder und Ebenhöf (1996) geben zu bedenken, dass der Begriff Selbstverwaltungsfähigkeit zwar ein direkt leistungsbezogenes Verhalten charakterisiert, jedoch in erster Linie eine Persönlichkeitsvariable dargestellt wird, die den individuellen Stil der Testperson dadurch beschreibt, wie die Handhabung von Unterbrechungen während einer intendierten Zielerreichung, die kurzfristige Aktualisierung von Problemstellungen und der Umgang mit Frustration erfolgt.

„Selbstverwaltungsfähigkeit“ beschreibt somit in erster Linie nicht einen effizienten Arbeitsstil, sondern vielmehr die Fähigkeit zum abwägenden Verhalten. Das Kernstück des Verfahrens stellt somit die Differenzierung hinsichtlich des kognitiven Stils der „Reflexivität versus Impulsivität“ (R-I) dar.

In Anlehnung an diesen bei Amelang und Bartussek (1990) angeführten R-I-Stil repräsentiert „Selbstverwaltungsfähigkeit“ im Rahmen des ILICA Personen, die bei kognitiv weniger anspruchsvollen Aufgabenstellungen entweder langsam und fehlerarm oder schnell und fehlerreich arbeiten. Um der inhärenten Komplexität der sozialen Bedingungsfaktoren der im ILICA genutzten Aufgabenstellungen gerecht zu werden, erweiterten Möseneder und Ebenhöf diese Typisierung um eine weitere Stilrepräsentation, wodurch sich letztlich folgende drei Typen klassifizieren lassen:

5.1.2 Typisierungen

„Reflexiver Typ

Personen dieses Typs gehen keine Kompromisse ein. Im Mittelpunkt ihres Denkens und Handelns stehen die eigenen Ziele, die aktiv und unbeeinflussbar verfolgt werden; ihr Verhalten ist also zielorientiert, konsequent und wenig durch Emotion gekennzeichnet.“ (Mösender und Ebenhöf, 1996, S. 6)

„Impulsiver Typ

Personen dieses Typs lassen sich leicht zu Dingen überreden bzw. verleiten, die nicht nur nicht geplant, sondern auch im Widerspruch zu bestimmten Zielen stehen; ihre Reaktionen erfolgen rasch, instinktiv und wenig überlegt.“ (Mösender und Ebenhöf, 1996, S. 6)

„Flexibler Typ

Personen dieses Typs wägen einander widersprechende Ziele und Bedürfnisse gegeneinander ab; sie zeigen Kompromißbereitschaft und lassen sich auch unter emotionalen Druck setzen.“ (Mösender und Ebenhöf, 1996, S. 6)

Abgesehen von diesen Typisierungen erfasst der Test ILICA eine Vielzahl weiterer formaler Kriterien wie z.B. die Häufigkeit der Nutzung diverser Hilfsmittel und kommt somit auf insgesamt 17 ermittelte Testkennwerte.

5.1.3 Kritik

Aufgrund der Tatsache, dass der Test seit seiner Veröffentlichung 1996 nicht weiter aktualisiert wurde, fallen die Unzulänglichkeiten der Testmethodik (im Vergleich zu den aktuellen Möglichkeiten der Computertechnik und des computerbasierten Testens) zum Teil gravierend auf. Die angestrebte Realitätsnähe kann objektiv betrachtet nicht erreicht werden. Limitierender Faktor hierbei ist die gravierende Diskrepanz zwischen den aktuellen technologischen Möglichkeiten sowie der Testumsetzung. Bereits für das Jahr 1996 bediente sich der ursprüngliche Test nicht den aktuellen Möglichkeiten der Technik. Die Simulation eines Tages nur auf Grundlage von Textmaterial erscheint für den angestrebten Nutzen des Tests unzulänglich.

Eine umfangreiche Kritik der Methodik des ILICA findet sich bei Funke 1999. Er merkte unter anderem an, dass die freie Gestaltung des Tages rein illusionärer Natur sei. Die zu bearbeitenden Aufgaben verfügen über ein bestimmtes Maß an festgelegter Zeit welches sie verbrauchen. Sie in unterschiedlichem Ausmaß zu unterschiedlichen Zeiten zu bearbeiten ist für die Testperson nicht möglich. So „kostet“ das Item „Hund füttern“ beispielsweise 15 Minuten Zeit und kann zudem beliebig oft wiederholt werden. Das Item „Staubsaugen“ verbucht 25 Minuten an virtueller Zeit und kann nur am Stück bearbeitet werden. Eventuell freie Zeit vor einem festen eingeplanten Termin mit beispielsweise 5 Minuten Staubsaugen zu überbrücken wäre zwar realistisch, jedoch wurde eine derartige Realitätsnähe im ILICA nicht erfolgreich umgesetzt.

5.2 Neuerungen im Vergleich Everest-ILICA

Die Entscheidung zur Neukonzipierung des ILICA und somit der Grundstein zur Entstehung des Everest entstanden aus dem Bedürfnis, die zuvor angeführten Kritikpunkte adäquat zu überarbeiten und bisherige Limitierungen aufzuheben. Zugleich wurde mit der Verlagerung des Schwerpunkts hin zum Arbeitskontext und weg von der freien Tagesgestaltung die Möglichkeit zum Einsatz des Everest für arbeitsrelevante Personalentscheidungen geschaffen. Aufgrund der zahlreichen Implementierungen des Everest wurde auch namentlich eine deutliche Abgrenzung vom Test ILICA geschaffen.

Der Name Everest stellt ein Akronym dar und steht in seiner Abkürzung für „Experimentalpsychologisches Verfahren zur Erfassung der Selbstverwaltungsfähigkeit“. So erfolgte auch eine Annäherung an die von Kubinger (2009) vorgeschlagene, aktuelle Bezeichnung der Experimentalpsychologischen Verhaltensdiagnostik.

Die Aufgaben benötigen noch immer ein bestimmtes Ausmaß an Zeit um erledigt zu werden, aus Gründen der Fairness erscheint es unsinnig für verschiedene Testperson verschiedene Aufgabenzeiten zu definieren. So sollte beispielsweise das Kopieren von Unterlagen für jede Testperson gleich lange dauern, ungeachtet der Tatsache ob es sich bei der Person um eine Person mit Kopiererfahrung handelt oder nicht. Was jedoch implementiert wurde, ist die Möglichkeit von Seiten der Testperson festzulegen wie viel ihrer Zeit sie auf die aktuelle Aufgabe investieren möchte. Eine Aufgabe die beispielsweise 60 Minuten Zeit benötigt um erledigt zu werden muss nun nicht mehr am Stück erledigt werden sondern erlaubt der Testperson die Bearbeitungszeit individuell aufzuteilen wodurch eine größere Realitätsnähe gewährleistet wird, da umfangreichere Aufgaben nun stückweise während möglicher Leerlaufzeiten erledigt werden können.

Eine der wichtigsten Neuerungen ist die implementierte Sprachausgabe. Aufgrund der Tatsache, dass ein gesamter Arbeitstag simuliert wird, der sowohl soziale als auch arbeitsrelevante Distraktoren enthält, ist die Menge an relevantem Textmaterial umfangreich.

Zur Erleichterung der Bedienung beziehungsweise des Verständnisses wurden sämtliche relevanten, an die Person gerichteten Texte auditiv unterlegt. Zugleich sollte

auf diese Art und Weise der angestrebte Immersionseffekt verstärkt werden, der beim ILICA aufgrund des umfassenden Textmaterials nicht in vergleichbarem Maße erreicht werden konnte.

Nach Kryspin-Exner (2011) versteht man unter Immersionseffekt „ein Versinken des Spielers in den Inhalt des Spiels und sein Eintauchen in die Rahmenhandlung“. (S. 175). Um dieses Gefühl während der Bearbeitung des Everest für die Testperson möglich durchgehend aufrecht zu erhalten wurde die Benutzeroberfläche des ILICA zu grafisch ansprechenden, virtuellen Umgebungen umgewandelt. Zudem wird die Testperson im Everest durch einen sogenannten Avatar repräsentiert, was zu einer zusätzlichen Identifikation der Testperson mit dem virtuell agierenden Charakter beiträgt. Auch die Aufgabenstellungen im Everest werden von virtuellen Charakteren vorgegeben. Zusammen mit der oben beschriebenen Audioausgabe soll so eine möglichst realistische, virtuelle Umgebung für die Testperson geschaffen werden.

Ein weiterer Punkt der bei der Neuauflage des Tests besonderes Augenmerk erhielt, war die Gestaltung der Testinstruktion. Gerade für wenig erfahrene Benutzer stellt die Einführung in die Steuerung des Tests sowie dessen Hilfsmittel einen wichtigen Punkt dar. So wurde darauf Wert gelegt, dass alle Aspekte der Testbedienung, seien es einfache Mausklicks oder die Verwendung testinterner Hilfsmittel, von der Testperson im Rahmen der Instruktionsphase zwangsweise selbst durchgeführt werden müssen. Eine Überforderung der Testperson sollte somit verhindert werden.

5.3 Testkonstruktion des Everest

5.3.1 Entstehungsgeschichte

Unstandardisierte Experteninterviews bilden die Grundlage für die computergestützte Simulation des Arbeitstages. Auf diese Art und Weise wurde ein prototypischer Tag in einem Büro ermittelt. Hierbei galt es, Geschehnisse zu ermitteln, die durch die Testperson entweder beeinflussbar oder nicht beeinflussbar waren.

Dazu wurden eine ehemalige Personalverrechnerin der Coca Cola HBC Austria GmbH sowie eine ehemalige Mitarbeiterin der Card Complete Service Bank AG eingehend befragt, um anhand der so gewonnenen Informationen größtmögliche Realitätsnähe nicht nur in Bezug auf die Aufgabenstellungen sondern auch hinsichtlich der Distraktoren zu erzeugen. Basierend auf den ermittelten Aussagen der beiden Expertinnen konnte das zu Grunde liegende Konzept der Computersimulation wie folgt festgelegt werden.

Bereits die Instruktionsphase des Tests stimmt die Testperson auf den für den Test gewählten Arbeitskontext ein. Zu diesem Zweck dient die sowohl schriftlich als auch auditiv an die Testperson herangetragene Instruktion, die nicht nur als Testphase für die Bedienung des Tests fungiert, sondern auch eine Rahmenhandlung schafft durch die der Testperson ein Eintauchen in die virtuelle Welt ermöglicht wird.

Besonderes Augenmerk liegt hierbei in der Formulierung der Einleitung, die bereits einen hohen Aufforderungscharakter hat (Siehe Punkt 5.4.1).

Die individuelle Gestaltung des Arbeitstages wird durch festgelegte Termine zeitlich limitiert sodass realistische Bedingungen geschaffen werden. Zudem wird der Tagesablauf – wie bereits im ILICA – durch für die Testperson nicht vorherzusehende, ungeplante Störungen unterbrochen. Dies dient zum einem der Erhöhung des Drucks aufgrund dessen es seitens der Testperson zu Entscheidungsfindungsprozessen kommt, sowie zum anderen der Erfassung des zugrundeliegenden Entscheidungsstils.

Sowohl die festgelegten Aufgabenstellungen des Arbeitsalltags als auch die im Test auftauchenden Distraktoren wurden in den Experteninterviews eingehend besprochen um Ereignisse zu wählen, die auch im realen Arbeitsalltag eintreffen. Den ExpertInnen wurden exemplarisch einige Situationen genannt, welche hinsichtlich ihrer

Realitätsnähe eingeschätzt wurden. Danach erfolgte eine Ergänzung der Situationsmöglichkeiten seitens der befragten ExpertInnen. Die so entstandene Liste an Aufgabenstellungen und Distraktoren wurde im Anschluss von den ExpertInnen hinsichtlich ihres realistischen Auftretens im Arbeitsalltag gereiht und in den Test übernommen, sodass der Immersionseffekt für die Testperson während der Testbearbeitung weiter verstärkt wird.

Programmierung und grafische Gestaltung

Der Test Everest wurde in der integrierten Entwicklungsumgebung (IDE) „Adventure Game Studio“ Version 3.2 programmiert und umfasst ungefähr 8000 Zeilen Code. Selbige IDE ist auf die Entwicklung sogenannter „point and click“ Adventure Games spezialisiert.

Die Hintergründe der einzelnen Szenen des Everest wurden mittels „Sweet Home 3D“ erstellt. Die hierbei verwendeten Objekte innerhalb der Szenen wurden aus frei verfügbaren und nicht lizenzierten Objekten zusammengestellt, welche zuvor mittels AutoCad 3D texturiert wurden. Nachdem Lichtquellen gesetzt wurden, wurde die jeweilige Szene finalisiert um schließlich ein einzelnes Hintergrundbild zu erhalten.

Für die verwendeten Menügrafiken wurden ebenfalls frei verfügbare Grafiken verwendet beziehungsweise wurden sie mittels des Bildbearbeitungsprogramms GIMP selbst erstellt.

Die Vertonung des Tests erfolgte mittels der Software Audacity.

Die Avatare und Charaktere des Tests wurden professionell seitens einer Grafikdesignerin erstellt. Die Animation der Figuren erfolgte mithilfe des Programms Anime Studio Pro, die dazu notwendige Vorarbeit hierzu, sprich das Layering und Colorieren erfolgte wiederum mittels GIMP.

5.3.2 Bedienung des Everest

„Point and Click“ als Bedienkonzept für computergestützte Programme entwickelte sich erstmals im Jahre 1988 mit dem Spiel „Manhunter“ von Sierra. Bevor dieses Konzept etabliert wurde, bediente man sich überwiegend der Texteingabe zur Steuerung interaktiver Spielesoftware. Die Idee des Konzepts ist gleichsam elegant wie simpel: durch einen einfachen Klick auf eine grafische Abbildung des gewünschten Zielobjektes mit der Maus oder anderen Eingabegeräten, setzt die bedienende Person ihre Aktionen in der virtuellen Welt.

Dieses Konzept war bereits durch grafische Benutzeroberflächen für Computer bekannt. Im Jahre 1973 war es erstmals möglich einen Computer, genauer gesagt den Xerox Alto, in einer Art und Weise zu bedienen wie sie uns heute aus dem alltäglichen Gebrauch bekannt ist. Microsoft Windows stellt die wohl am weitesten verbreitete und bekannteste grafische Benutzeroberfläche dar und obgleich es in den Jahren seiner Entstehung 1983 bis heute 2014 einen rasanten technologischen Wandel durchlebt hat, gestaltet sich die Bedienung noch immer gleich. „Point and click“ stellt somit eine bekannte Größe dar. Durch seine Verbreitung ist davon auszugehen, dass der überwiegende Teil der Endnutzer mit der Art der Steuerung nicht überfordert sein sollte.

Um dies weiter zu illustrieren sei an dieser Stelle eine moderne und sehr populäre Art und Weise der „point and click“ Steuerung angeführt. So genannte „Tower defence“ Spiele erfreuen sich im Internet besonderer Beliebtheit und sind auf Seiten wie z.B. www.armorgames.com kostenlos und direkt im Web-Browser nutzbar. Analog zur Steuerung des „Everest“ nutzen sie eine zwei dimensionale Umgebung mit welcher mittels „point and click“ interagiert werden kann. Das von Ironhidegames entwickelte Spiel „Kingdom Rush“ ist hierbei eines der populärsten Exemplare. Veröffentlicht wurde es auf der angeführten Seite am 29.07.2011, seit dieser Zeit gab es 55.816.203 Zugriffe auf dieses Spiel.

5.4 Testdurchführung des Everest

Vor Beginn des eigentlichen Tests wird die Testperson mittels eines animierten Bildes, sowie eines kurzen Instruktionstextes dazu aufgefordert die bereitgestellten Kopfhörer aufzusetzen. Durch Bestätigen dieses Fensters ertönt ein Testton, welchen die Testperson bestätigen muss, sofern er vernommen wird. Dies dient der Überprüfung der technischen Gegebenheiten und stellt einen reibungslosen Testablauf sicher. Eine Kontrolle der Audioausgabe ist aufgrund der zentralen Bedeutung für die Testbearbeitung unabdingbar.

Anschließend wird die Testperson dazu aufgefordert ihre demographischen Merkmale, Beruf, Alter, Geschlecht und Schulbildung anzugeben. Sollten die Angaben unvollständig sein, so wird die Testperson darauf hingewiesen.

5.4.1 Trainings- und Instruktionsphase

Basierend auf den Angaben der Testperson bezüglich ihres Geschlechts, findet sich nun ihr entsprechend weiblicher oder männlicher Avatar vor dem Gebäude der Firma Everest wieder. Zu diesem Zeitpunkt ist es der Testperson noch nicht möglich mit der virtuellen Umwelt zu interagieren. Diese Möglichkeit wird infolge Stück für Stück initiiert, während die entsprechenden Trainingseinheiten durchlaufen werden.

Zu Beginn wird die Testperson über das virtuelle Setting des kommenden Arbeitstages informiert. Sie erfährt, dass sie bereits lange für die Firma Everest arbeitet, heute jedoch das erste Mal in einer neuen Abteilung zum Einsatz kommen wird.

Hiermit sollte der Testperson vermittelt werden, dass es in Folge nicht darum geht neue Dinge zu erlernen, sondern einen „gewöhnlichen“ Arbeitstag zu absolvieren, der auch ohne besondere Vorkenntnisse spezifischer Aufgabenbereiche zu bewältigen ist.

Nachdem die Rahmenbedingungen dargelegt wurden um die Testperson auch emotional an das Setting heranzuführen, erfolgt eine Einführung in die Steuerung sowie eine Erläuterung und der zur Verfügung stehenden Hilfsmitteln während der Testbearbeitung.

Nach einer kurzen Erläuterung des Prinzips des „point and click“, wird die Testperson dazu aufgefordert dieses erlernte Wissen sogleich anzuwenden. Der Bewegungsfreiraum der Maus wird hierbei auf einen kleinen Teilbereich des Bildschirms beschränkt und die Testperson dazu aufgefordert durch Klicken mit der Maus ihren Avatar zu eben jener Stelle zu bewegen. Aufgrund der Restriktion des Handlungsfreiraumes ist eine Überforderung gleich zu Beginn des Tests auszuschließen. Erst nach erfolgreicher Absolvierung der vorgegebenen Aufgabe fährt der Test fort.

Nun werden die der Testperson zur Verfügung stehenden Hilfsmittel, die Uhr und der Notizblock, näher erklärt (Abbildung 1). Es wird explizit darauf hingewiesen, dass die Zeit innerhalb des Tests nicht mit normaler Geschwindigkeit, sondern beschleunigt abläuft. Um welchen Faktor der Umrechnung es sich hierbei genau handelt wird der Testperson nicht mitgeteilt. Allerdings kann diese durch Beobachten der Uhr ein Gefühl hierfür erlangen. Die Testperson wird dazu aufgefordert die Uhr probeweise durch einen Klick auf die Schaltfläche „Hilfe“ zu öffnen. Die Maus wird hierbei auf diese Schaltfläche fixiert, sodass sichergestellt wird, dass diese Instruktionsaufgabe erfüllt wird.

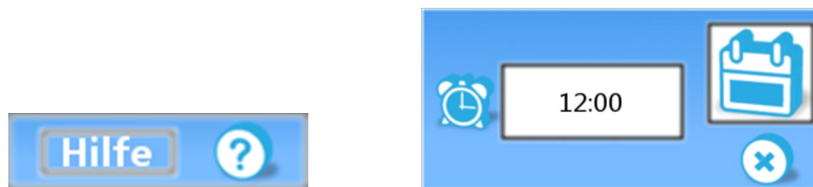


Abb. 1: Schaltflächen „Hilfe“ und „Uhr“

Nachdem die Funktionsweise der Uhr erklärt wurde, wird die Testperson dazu instruiert den Notizblock zu öffnen (Abbildung 2). Auch hier wird der Aktionsradius der Maus wieder beschränkt.

Durch einen Klick öffnet sich der Notizblock der Trainingsphase. Dieser unterscheidet sich lediglich durch die aufgeführten Aufgaben von dem späteren Notizblock.

So wird verhindert, dass sich die Testperson bereits im Vorfeld einen Überblick über die während der Testbearbeitung anstehenden Aufgaben verschaffen kann. Die Testperson wird darauf hingewiesen, dass erledigte Aufgaben abgehakt werden sofern sie zur Gänze erledigt wurden. Anschließend wird sie dazu aufgefordert den Notizblock wieder zu schließen.

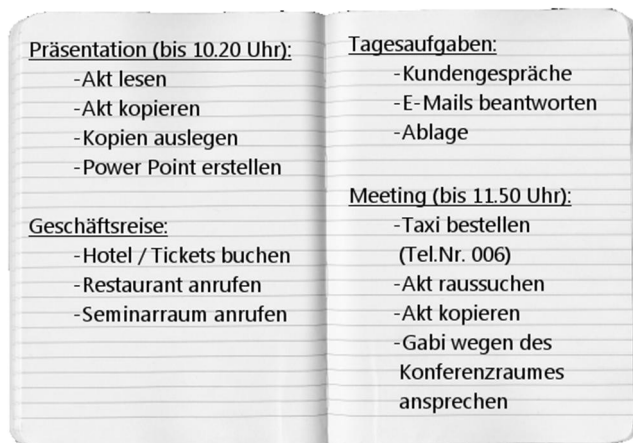


Abb. 2: Notizblock

Zum Abschluss der Trainingsphase wird der Testperson erklärt, wie sie Dinge in der Umgebung hervorheben kann mit denen eine Interaktion möglich ist.

Aufgrund der realistischen Hintergründe bietet sich der Testperson eine Fülle an optischen Informationen. Obgleich die Hintergründe nach Möglichkeit zweckdienlich gestaltet wurden, wurden im Sinne der Realitätsnähe Objekte eingefügt, mit denen eine Interaktion nicht erfolgen kann. Dies sind beispielsweise Objekte wie Aktenordner oder Flipcharts.

Damit die Testperson stets einen Überblick darüber behält, was Teil des Tests und was Dekoration ist, erhält sie die Möglichkeit mittels der Leertaste jedes interaktive Objekt hervorheben zu lassen. Sollte der Mauszeiger sich über einem Objekt befinden, mit dem die Testperson auch tatsächlich interagieren kann, so verändert sich der Mauszeiger entsprechend und die Testperson erhält einen Hinweis (Abbildung 3).

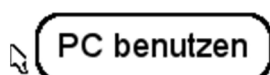


Abb. 3: Veränderter Mauszeiger bei Interaktionsmöglichkeit am Beispiel „PC benutzen“

Nachfolgend wird der Testperson erläutert, dass Dialoge in dem Test so gestaltet sind, dass stets zwischen drei Antwortalternativen ausgewählt werden kann. Um dies selbst auszuprobieren wird sie dazu angewiesen den abgebildeten männlichen Charakter vor den Türen des Bürogebäudes anzuklicken. Der eigentliche Test, sowie die virtuelle Arbeitszeit der Testperson beginnen erst nach Erledigung dieser Aufgabe. Ohne diese Aufgabe zu erledigen ist es der Testperson nicht möglich das Bürogebäude zu betreten. Eine entsprechend formulierte Nachricht würde die Testperson in diesem Falle darauf aufmerksam machen, dass zuvor die Probeaufgabe zu absolvieren ist.

Im Gespräch mit der Person vor dem Gebäude erhält die Testperson anschließend unter anderem Zugriff auf ihr Diensthandy (Abbildung 4). Es wird kurz erklärt, wie die Bedienung desselben verläuft, bevor nun der eigentliche Arbeitstag beginnt. Die Erläuterung des Handys ist hierbei absichtlich kurz gehalten, da die Bedienung überaus intuitiv erfolgt, da sich besonders hier eine große Realitätsnähe erkennen lässt und die Bedienung mit der eines realen Mobiltelefons vergleichbar ist. Mittels Druck auf die Tasten wird die entsprechende Nummer eingegeben, durch Klicken des grünen Buttons erfolgt der Anruf, ein Klick auf den roten Button löscht die Eingabe.



Abb. 4: Diensthandy

Im Inneren des Bürogebäudes wird die Testperson von ihrem Vorgesetzten Herrn Lehmann, sowie ihrer neuen Kollegin Frau Schmidt begrüßt. Letztere begleitet die Testperson in ihr eigenes Büro. Dort angekommen wird die Testperson mit der Bedienung der Arbeitsaufgaben vertraut gemacht.

Das Prinzip gestaltet sich hierbei über jede Aufgabe hinweg gleich, lediglich optisch unterscheiden sich die entsprechenden Fenster durch eine kleine Grafik im oberen linken Eck, welches die Art der Tätigkeit beschreibt.

Abbildung 5 zeigt die Aufgabenstellung der PowerPoint Vorbereitung, welche durch das Symbol des kleinen Präsentationsschirmes untermalt wird. Der Testperson wird vermittelt, dass eine Aufgabe stets ein gewisses Maß an Arbeitszeit zur erfolgreichen Erfüllung benötigt. In diesem Falle wären es beispielsweise 20 Minuten. Mittels des Schiebereglers kann die Testperson ein beliebiges Ausmaß an Zeit (hier: von 1 – 20 Minuten) für diese Aufgabe in diesem Moment aufwenden.

Verschiebt sie hierbei den Regler, so aktualisiert sich die Anzeige der verbleibenden Minuten bis zur Zielerfüllung entsprechend. Ein Klick auf den Haken bestätigt die entsprechende Auswahl, wohingegen ein Klick auf das X das aktuelle Fenster schließt und die Auswahl abbricht.



Abb. 5: Schaltfläche „Aufgabenstellung“ am Beispiel PowerPoint-Präsentation

Weiteres wird die Testperson darauf hingewiesen, dass Sie durch einen Klick auf den PC Bildschirm auf ihrem Arbeitsplatz ein entsprechendes Menü öffnet, über das sich eine Reihe an Aufgaben aufrufen lässt wie beispielsweise die Erstellung der PowerPoint-Präsentation (Abbildung 6).



Abb. 6: Virtueller Desktop der Testperson während der Testbearbeitung

Nachdem die Bedienung des Tests vermittelt wurde, wird der Testperson von Seiten ihres Vorgesetzten Herrn Lehmann der Tagesablauf erklärt. Sie erfährt welche Aufgaben es zu erledigen gilt und welche Deadlines einzuhalten sind. Diese Informationen werden in dem zur Verfügung stehenden und aus der Instruktionsphase bereits bekannten Notizblock automatisch vermerkt. Es wird der Testperson jedoch die Möglichkeit geboten die Instruktion bei Unklarheiten erneut zu erfragen.

5.4.2 Testitems

Der Test umfasst mehrere Kategorien von Aufgaben, die die Testperson erledigen muss. Zum einen wird die Testperson mit Aufgabenstellungen konfrontiert die an ein Zeitlimit gebunden sind, zum anderen erhält die Testperson Aufgabenstellungen die sie selbstständig über den Tag verteilt erledigen kann.

Arbeitsaufgaben mit Zeitlimit

Die Testperson soll eine Präsentation bis 10.20 Uhr vorbereiten, hierzu müssen folgende Aufgaben erledigt werden:

- Akt lesen
- Akt kopieren
- Kopien auslegen
- Powerpoint-Präsentation erstellen

Zudem erteilt der Vorgesetzte der Testperson im Verlauf der Testbearbeitung eine Anzahl weiterer Aufgaben, da unerwartet ein wichtiger Kunde eintrifft und kurzfristig ein zusätzliches Meeting stattfindet. Hierfür gilt es folgende Aufgaben zu erledigen:

- ein Taxi für den Kunden bestellen
- Akt raussuchen
- Akt kopieren
- Kollegin bitten den Konferenzraum vorzubereiten
- kopierten Akt an Chef übermitteln

Arbeitsaufgaben ohne Zeitlimit

Der Testperson wird mitgeteilt, dass sie in der nächsten Woche eine Geschäftsreise unternehmen muss. Hierfür gilt es folgende Aufgaben zu erledigen:

- Hotel/ Tickets buchen
- Restaurant anrufen
- Seminarraum anrufen

Darüber hinaus erhält die Testperson generelle Büroaufgaben die sie im Laufe des Tages erledigen sollte:

- Kundengespräche am Telefon führen
- E-Mails beantworten
- Ablage erledigen

Letztere Aufgabe besteht aus zwei Teilen: der Aufbereitung der Ablage, sowie dem Ablegen der Ablage am entsprechendem Ort.

Die Kundengespräche kann die Testperson nicht aktiv beeinflussen, selbige treten über den Tag verteilt auf und zählen somit gleichzeitig zu der folgenden Kategorie, den Distraktoren.

Distraktoren

Die oben angeführten Aufgaben stellen keinen wirklichen organisatorischen Anspruch an die Testperson, da sie innerhalb des simulierten Zeitfensters leicht zu absolvieren sind.

Um die zeitlichen Ressourcen zu verkürzen und somit die Aufgaben in Bezug auf das Ressourcenmanagement schwieriger zu gestalten, treten über den Tag verteilt folgende Distraktoren auf, die laufende Aufgaben unterbrechen können:

- drei Kundenanrufe
- eine Kollegin bietet einen Rundgang an
- ein alter Freund und Kollege möchte einen Kaffee trinken
- eine Kollegin hat Probleme mit dem Drucker
- das Kind der Kollegin muss ins Krankenhaus, die Testperson soll die Ablage übernehmen

Auf diese sieben Distraktoren, welche sowohl arbeits- als auch freizeitbezogen gestaltet sind, reagiert die Testperson jedes Mal durch Auswahl aus drei verschiedenen Antwortalternativen. Jede Antwortalternative ist einem der drei zugrundeliegenden Entscheidungstypen inhaltlich zugeordnet und verändert in bestimmter Art und Weise durch unterschiedliche zeitliche Ressourcenanforderungen den weiteren Testverlauf.

5.4.3 Testumgebungen und Personen

Aufgrund der Tatsache, dass es sich bei dem Test Everest um die Simulation eines Bürotages handelt, bietet der Testablauf der Testperson eine Vielzahl an Interaktionsmöglichkeiten. Darüber hinaus kann sich die Testperson innerhalb des virtuellen Bürogebäudes frei bewegen.

Im Folgenden sollen die zehn im Everest enthaltenen Testumgebungen, sowie die jeweils enthaltenen Interaktionsmöglichkeiten kurz vorgestellt werden.

Büro der Testperson (Abbildung 7)

In diesem Raum verbringt die Testperson den Großteil ihres virtuellen Arbeitstages.

Es bieten sich Interaktionsmöglichkeiten mit dem Computer, der Ablage, dem Akt, dem Telefon und der Tür.



Abb. 7: Büro der Testperson

Mittels des Computers in ihrem Büro ist es der Testperson möglich ihre Zeit mit einem Spiel zu verbringen (Abbildung 8). Sollte die Testperson sich für diese Option entscheiden, sieht sie das oben abgebildete Fenster. Als Interaktionsmöglichkeiten stehen das Spielfeld sowie der Beenden-Button in der unteren rechten Ecke zur Verfügung. Das Tic Tac Toe Spiel ist mit ungefähr 30 Sekunden realen Zeitaufwandes nicht sonderlich langwierig, verbraucht jedoch 10 Minuten der virtuellen Testzeit.

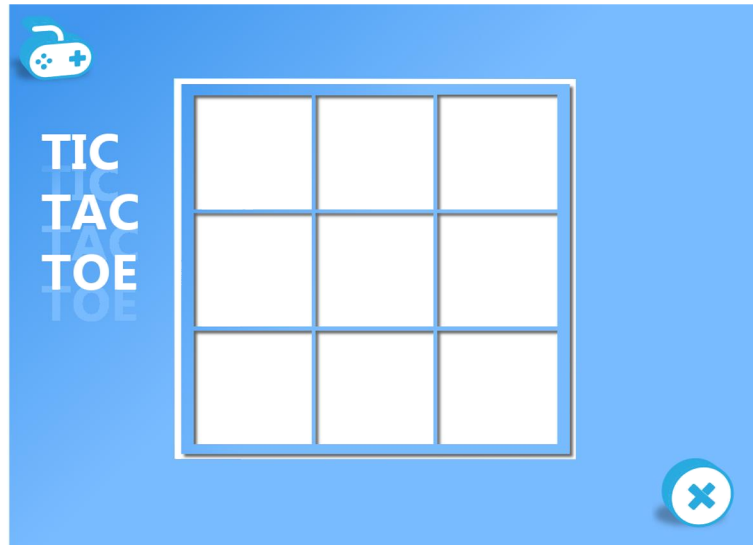


Abb. 8: Tic Tac Toe Spielfeld

Außenbereich (Abbildung 9)

In dieser Umgebung beginnt die Testperson ihren Arbeitsalltag.

Es bieten sich Interaktionsmöglichkeiten mit der Tür sowie mit dem Hausmeister, welcher in dieser Umgebung zu finden ist. Der Hausmeister bietet der Testperson darüber hinaus die Möglichkeit eine Pause von fünf Minuten einzulegen.



Abb. 9: Außenansicht Bürogebäude

Flur Untergeschoss (Abbildung 10)

Dieser Gang verbindet das Büro der Testperson mit allen anderen Räumen des Bürogebäudes. Darüber hinaus ist es der Testperson möglich über die Treppe in den zweiten Stock des Gebäudes zu wechseln. Neben den Türen bietet sich der Kaffeeautomat als Interaktionsmöglichkeit an. Selbiger ermöglicht das Einlegen von Pausen, deren Länge von der Testperson frei gewählt werden kann.

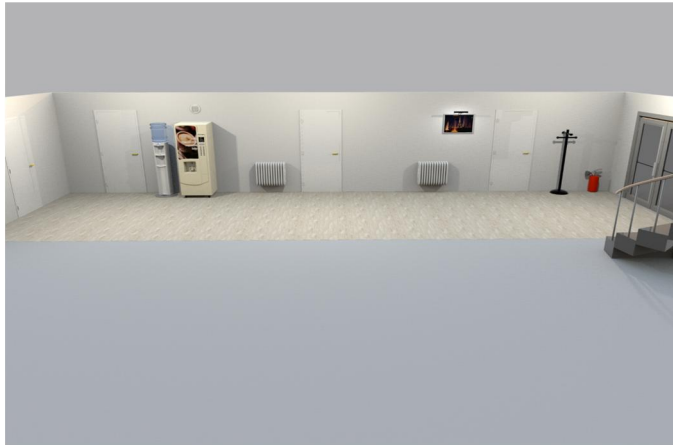


Abb. 10: Gang im unteren Stockwerk

Büro des Vorgesetzten (Abbildung 11)

Dieser Raum befindet sich im unteren Stockwerk. Die Testperson kann hier ihren Vorgesetzten antreffen, um ihm im Rahmen einer zu erledigenden Aufgabe einige kopierte Unterlagen zu übergeben. Als Interaktionsmöglichkeiten bieten sich lediglich die Tür sowie der Vorgesetzte.

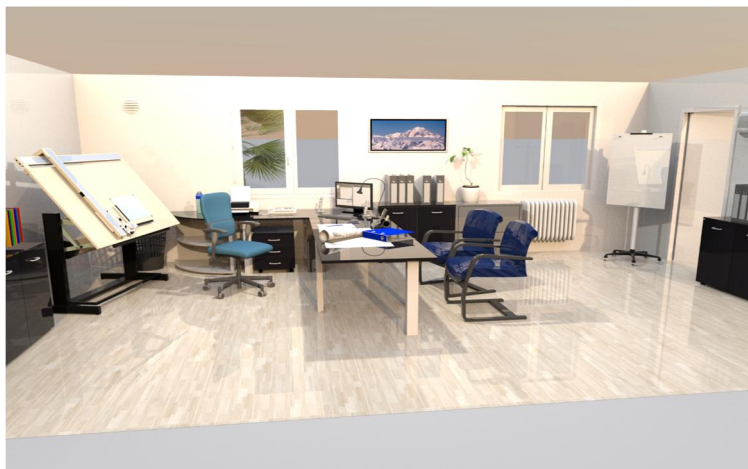


Abb. 11: Büro des Vorgesetzten

Kantine (Abbildung 12)

Ebenfalls im unteren Stockwerk befindet sich die Kantine. Hier verbringt die Testperson Zeit mit einem ihrer Kollegen, sofern sie sich für eine entsprechende Antwortalternative entscheidet. Abgesehen von besagtem Kollegen bietet sich als Interaktionsmöglichkeit lediglich die Tür.



Abb. 12: Kantine

Büro der Kolleginnen (Abbildung 13)

Dieser Raum befindet sich ebenfalls im unteren Stockwerk. Die Testperson kann hier ihre beiden Kolleginnen antreffen und mit ihnen in Dialog treten. Abgesehen davon bietet sich nur der untere Bildschirmrand als Interaktionsmöglichkeit um den Raum zu verlassen.

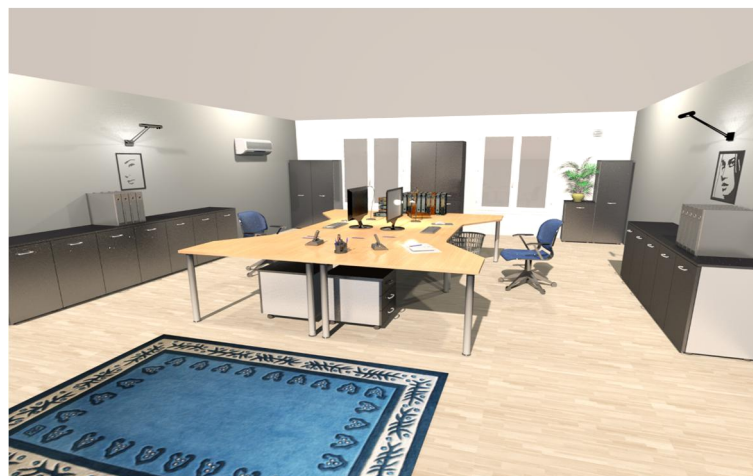


Abb. 13: Büro der Kolleginnen

Flur Obergeschoss (Abbildung 14)

Wählt die Testperson im unteren Gang die Treppe, so wechselt sie in diesen Raum. Hier kann sie wiederum entscheiden ob sie in den Kopierraum oder in den Konferenzraum wechseln möchte. Die dritte Tür bietet die Möglichkeit zurück in den unteren Gang zu wechseln.

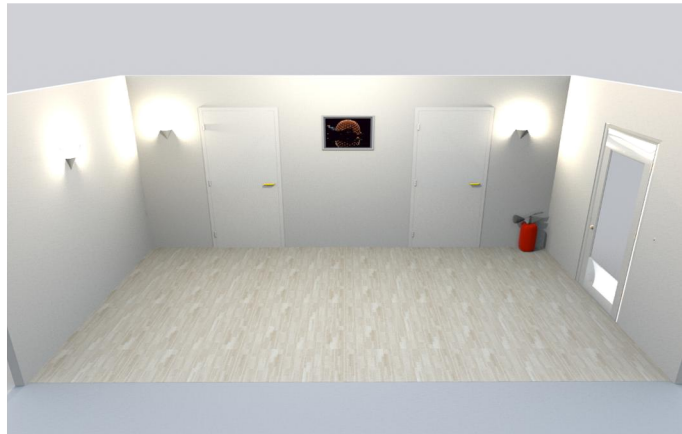


Abb. 14: Gang im oberen Stockwerk

Konferenzraum (Abbildung 15)

In diesem Raum findet die Präsentation statt, die die Testperson als Aufgabe vorbereiten soll. Je nachdem, ob alle Aufgabenstellungen die mit diesem Überpunkt zusammenhängen erledigt wurden oder nicht, hält entweder der Avatar der Testperson den Vortrag, oder ihr Vorgesetzter. Sollten nicht alle Aufgaben erledigt worden sein, so weist der Vorgesetzte die Testperson an dieser Stelle darauf hin und erwartet in Zukunft bessere Leistungen. Als Interaktionsmöglichkeiten bieten sich der Tisch zur Bereitstellung der Unterlagen, sowie die Tür.

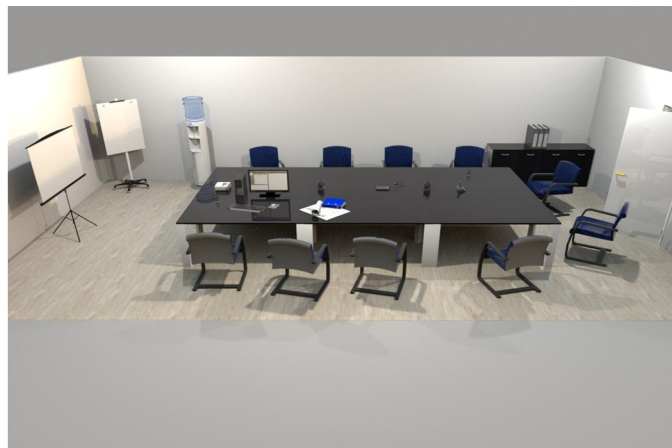


Abb. 15: Konferenzraum

Kopierraum (Abbildung 16)

In diesem Raum bieten sich der Testperson Interaktionsmöglichkeiten mit der Tür, dem Drucker, sowie dem Aktenschrank für die Ablage. Der Drucker fällt bei erstmaliger Benutzung aus und erschwert somit die Einhaltung der zeitlichen Vorgaben. Die Testperson muss sich infolge an den Haustechniker wenden, um den Drucker reparieren zu lassen.



Abb. 16: Kopierraum

Avatare (Abbildung 17)

Basierend auf der Angabe des Geschlechts der Testperson im Zuge der Erfassung der demographischen Daten wird selbiger, wie bereits eingangs erwähnt, ein entsprechend männlicher oder weiblicher Avatar zugewiesen.



Abb. 17: Männlicher und weiblicher Avatar

Abgesehen davon besteht für die Testperson die Möglichkeit mit folgenden Personen im virtuellen Arbeitsumfeld zu interagieren.

Vorgesetzter Herr Lehmann und Haustechniker (Abbildung 18)



Abb. 18: Vorgesetzter Herr Lehmann und Haustechniker Pröll

Kollegium (Abbildung 19)

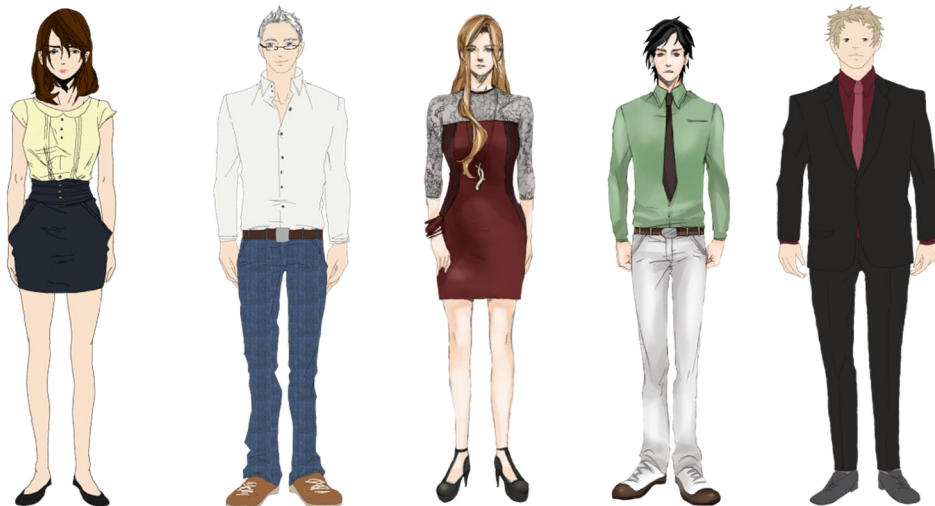


Abb. 19: Kollegium

5.4.4 Typisierung

Analog zu der Typisierung des ILICA erfolgt im Everest eine Einteilung der Entscheidungsmuster in folgende Klassen:

Reflexiver Typ

„Personen dieses Typs gehen keine Kompromisse ein. Im Mittelpunkt ihres Denkens und Handelns stehen die eigenen Ziele, die aktiv und unbeeinflussbar verfolgt werden; ihr Verhalten ist also zielorientiert, konsequent und wenig durch Emotion gekennzeichnet.“ (Mösender und Ebenhö, 1996, S. 6).

Sie gehen weniger auf die Bedürfnisse anderer ein und stellen Ihre Interessen über die Belange anderer um eine optimale Zielerreichung zu sichern.

Ermittelt wird dieses innerhalb des Everest dadurch, dass in der Antwortalternative klar ersichtlich ist welche Konsequenzen sich aus der Wahl der Antwortalternative ergeben, da ein direkter Bezug zur zeitlichen Dauer der Aufgabe gegeben ist. Die Testperson kann auf diese Weise Wahrscheinlichkeiten dafür berechnen ob das Verhalten das aus der Antwortalternative resultiert entgegen ihrer angestrebten Zielerreichung ausgerichtet ist oder nicht. Die Testperson muss also entscheiden, ob genügend zeitliche Ressourcen zur Verfügung stehen um die gegebenen Optionen ausführen zu können. Hierfür ist es erforderlich, einen Überblick über die erledigten, zu erledigenden und aktuellen Aufgaben sowie die aktuelle Zeit bzw. Fristen zu haben. Dies stellt einen höheren Anspruch an die kognitiven Ressourcen der Testperson dar. Wie bei Strack und Deutsch (1999) dargelegt, werden durch die benötigte höhere kognitive Kapazität und durch den Einbezug von Konsequenzen und Wahrscheinlichkeiten in den Entscheidungsfindungsprozess reflexive Informationsverarbeitungsprozesse in Gang gesetzt, woraus sich die Bezeichnung des Typs ergibt.

Impulsiver Typ

„Personen dieses Typs lassen sich leicht zu Dingen überreden bzw. verleiten, die nicht nur nicht geplant, sondern auch im Widerspruch zu bestimmten Zielen stehen; ihre Reaktionen erfolgen rasch, instinktiv und wenig überlegt.“ (Mösender & Ebenhöf, 1996, S. 6).

Sie gehen vermehrt auf die Bedürfnisse anderer ein, sind sozial bemüht und vernachlässigen die angestrebte Zielerreichung aufgrund der Belange anderer.

Ermittelt wird dies innerhalb des Everest dadurch, dass Wahrscheinlichkeiten im Gegensatz zu den anderen Entscheidungsalternativen nicht gegeben sind.

Für die Testperson ist somit nicht direkt ersichtlich wie lange bzw. wie viel Zeit sie für die gewählte Antwortalternative investieren muss. Die Testperson wägt somit nicht in erster Linie über das Für und Wider aufgrund zeitlicher Ressourcen ab, sondern lässt sich bei ihrer Entscheidungsfindung primär von den Erwartungen der Gesprächspartner beeinflussen. Strack und Deutsch (1999) zufolge führen impulsive Prozesse Verhalten auch unter suboptimalen Gegebenheiten sehr rasch herbei. Für die Testperson sind durch die fehlenden Zeitangaben sowie den emotionalen Druck aufgrund sozialer Erwartungen in den entsprechenden Situationen keine idealen Bedingungen geschaffen. Da in die Entscheidungsfindung weder Konsequenzen noch Wahrscheinlichkeiten mit einbezogen werden können beruht diese auf motivationalen Komponenten. Auch dies spricht für impulsive Informationsverarbeitungsprozesse, woraus sich wiederum die Bezeichnung des Typs ergibt.

Flexibler Typ

„Personen dieses Typs wägen einander widersprechende Ziele und Bedürfnisse gegeneinander ab; sie zeigen Kompromißbereitschaft und lassen sich auch unter emotionalen Druck setzen.“ (Mösender & Ebenhö, 1996, S. 6).

Sie gehen auf die Bedürfnisse anderer ein, versuchen dabei jedoch zwischen eigener Zielerreichung und den Belangen anderer zu vermitteln.

Ermittelt wird dies innerhalb des Everest dadurch, dass sich die Testperson für eine Alternative entscheidet, welche im Gegensatz zu der impulsiven Variante über eine Zeitangabe verfügt. Dennoch steht diese Aufgabe, anders als die reflexive Antwortalternative, im Gegensatz zur angestrebten Zielerreichung. Die Testperson muss folglich entscheiden, ob die zeitlichen Ressourcen es ihr erlauben die entsprechende Option zu wählen.

Zwar sind auch hier die Bedingungen für die Testperson nicht ideal, jedoch ist es ihr möglich Wahrscheinlichkeiten und Konsequenzen in den Entscheidungsfindungsprozess mit einzubeziehen. Die Bezeichnung des Typs erfolgte in Anlehnung an die Typenbezeichnung bei Möseneder und Ebenhö (1996).

5.5 Testkennwerte des Everest

In dem folgendem Abschnitt sollen die als relevant erachteten Testkennwerte näher erläutert werden. Abgesehen von den aufgeführten Kennwerten erfasst der Everest eine Vielzahl an weiteren Variablen, wie beispielsweise die demographischen Daten, oder die Häufigkeit bestimmter Aktionen, auf die in den folgenden Ausführungen jedoch nicht weiter eingegangen wird, da sie für die Auswertung des Verfahrens keine tragende Rolle spielen.

- **Summe der reflexiven Entscheidungen**
- **Summe der flexiblen Entscheidungen**
- **Summe der impulsiven Entscheidungen**

Die drei angeführten Summenscores bilden sich aus den Antworten der Testperson auf die unter Punkt 5.4.2 angeführten Distraktoren.

- **Summe der erreichten Arbeitsaufgaben**

Stellt einen Summenscore über alle Arbeitsaufgaben dar, die unter Punkt 5.4.2 aufgeführt wurden.

- **Summe der durchgeführten Freizeitaktivitäten**

Stellt einen Summenscore über die Aktivitäten Kaffeepause, Pause sowie Spielen dar.

- **Verwendung der Uhr**

Erfasst wie häufig die Testperson auf die Uhr zugreift.

- **Verwendung des Notizblock**

Erfasst wie häufig die Testperson auf den Notizblock zugreift.

- **Bearbeitungszeit**

Erfasst wird hierbei die Zeit in Minuten die die Testperson für die Bearbeitung des Tests benötigt.

Anhand eines angefügten Akzeptanzfragebogens ließen sich folgende weiterführende Variablen erheben:

- **Die erlebte Belastung durch den Test**

- **Ernsthaftigkeit bei Bearbeitung des Tests**

Beide angeführten Testkennwerte werden auf einer Skala von 0 bis 100 erfasst, wobei 0 für eine geringe Ausprägung und 100 für eine maximale Ausprägung steht.

Empirischer Teil

6 Stichprobenbeschreibung

An der Testung zur Skalierung und Validierung des Everest nahmen im Zeitraum vom 12.05.2014 bis zum 20.08.2014 insgesamt 65 Personen teil. Das in Tabelle 1 ersichtliche Alter der Testpersonen liegt hierbei zwischen 19 und 65 Jahren (Mittelwert: 35,48, SD: 13,16).

Tab. 1: Alter

	N	Minimum	Maximum	Mittelwert	Standardabweichung
Alter	65	19,00	65,00	35,4769	13,16020

Die Stichprobe setzt sich aus 40 Männern (61,5 Prozent) und 25 Frauen (38,5 Prozent) zusammen (Tabelle 2).

Tab. 2: Geschlechterverteilung

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Männlich	40	61,5	61,5	61,5
Weiblich	25	38,5	38,5	100,0
Gesamt	65	100,0	100,0	

Auf einer Skala von 0 bis 100 wurde der Test mit einem Mittelwert von 40,18 als wenig belastend empfunden. Mit einem Mittelwert von 80,49 wurde der Test hierbei überaus ernsthaft bearbeitet (Tabelle 3).

Tab. 3: Ernsthaftigkeit und Belastung

Statistiken

	Ernsthaftigkeit	Belastung
N Gültig	65	65
Mittelwert	80,4923	40,1846
Standardabweichung	17,06250	16,47146

Die eigentliche Testung wurde mittels Laptop durchgeführt, wodurch ortsungebunden Testungen ermöglicht wurden. Ein Großteil der Testpersonen wurde auf diese Weise in Deutschland, Nordrhein-Westfalen, akquiriert.

Darüber hinaus wurde ein Teil der Testungen in den Räumen der Computerdiagnostik an der Fakultät für Psychologie an der Universität Wien durchgeführt.

Die Teilnahme an der Testung erfolgte freiwillig und nach schriftlicher Einverständniserklärung (siehe Anhang).

7 Itemanalyse

Der nachfolgende Abschnitt beschreibt die sieben relevanten Items, welche zur Berechnung der Skalen Reflexiv, Impulsiv und Flexibel herangezogen werden. Die Antworthäufigkeiten sind hierbei in Tabelle 4 ersichtlich, die Itemtrennschärfe sowie Itemschwierigkeiten sind den Tabellen 5 und 6 zu entnehmen.

- **Item1 : IRund**

Es zeigte sich, dass sich 52% der Testpersonen bei Item 1 für die reflexive Option, 22% für die impulsive, sowie 26% für die flexible entschieden. Das Item zeigte hierbei eine Itemtrennschärfe von -0,42 bei einer Itemschwierigkeit von 65,4.

- **Item 2: IAnruf1**

52 % der Testpersonen entschieden sich bei Item 2 für die flexible Option, lediglich 9% entschieden sich hierbei für die reflexive Option, 39% entschieden sich für die impulsive Option. Das Item zeigte hierbei eine Itemtrennschärfe von 0,57 bei einer Itemschwierigkeit von 35,4.

- **Item 3 : IFreund**

60% der Testpersonen entschieden sich bei Item 3 für die reflexive Option, nur 3% wählten die impulsive Antwortalternative. 37% entschieden sich für die flexible Option. Das Item zeigte hierbei eine Itemtrennschärfe von 0,25 bei einer Itemschwierigkeit von 78,5

- **Item 4 : IDruck**

69% der Testpersonen entschieden sich bei Item 4 für die flexible Option, die reflexive Antwortalternative fiel mit 9% gering aus. 22% entschieden sich für die impulsive Option. Das Item zeigte hierbei eine Itemtrennschärfe von 0,65 bei einer Itemschwierigkeit von 43,9.

- **Item 5 : IAnruf2**

48% der Testpersonen entschieden sich bei Item 5 für die flexible Option, 28% wählten die impulsive Option. 24% entschieden sich für die reflexive Option. Das Item zeigte hierbei eine Itemtrennschärfe von 0,58 bei einer Itemschwierigkeit von 48,5.

- **Item 6 : IKind**

48% der Testpersonen entschieden sich bei Item 6 für die flexible Option, nur 18% wählten die reflexive Alternative. 34% wählten die impulsive Option. Das Item zeigte hierbei eine Itemtrennschärfe von 0,42 bei einer Itemschwierigkeit von 42,3.

- **Item 7 : IAnruf3**

45% der Testpersonen entschieden sich bei Item 7 für die impulsive Option, lediglich 21% wählten hierbei die reflexive Alternative. 34% entschieden sich für die flexible Option. Das Item zeigte hierbei eine Itemtrennschärfe von 0,13 bei einer Itemschwierigkeit von 35,4.

Tab. 4: Antworthäufigkeiten

	N	Mittelwert
	Statistik	Statistik
IImp1	65	,2154
IFlex1	65	,2615
IRef1	65	,5231
IImp2	65	,3846
IFlex2	65	,5231
IRef2	65	,0923
IImp3	65	,0308
IFlex3	65	,3692
IRef3	65	,6000
IImp4	65	,2154
IFlex4	65	,6923
IRef4	65	,0923
IImp5	65	,2769
IFlex5	65	,4769
IRef5	65	,2462
IImp6	65	,3385
IFlex6	65	,4769
IRef6	65	,1846
IImp7	65	,4462
IFlex7	65	,3385
IRef7	65	,2154
Gültige Werte	65	

Tab. 5: Itemschwierigkeiten

	N	Mittelwert
	Statistik	Statistik
IRund	65	1,3077
IAnruf1	65	,7077
IFreund	65	1,5692
IDruck	65	,8769
IAnruf2	65	,9692
IKind	65	,8462
IAnruf3	65	,7077
Gültige Werte	65	

Tab. 6: Itemtrennschärfen

Item-Skala-Statistiken

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala- Korrelation	Cronbachs Alpha, wenn Item weggelassen
IRund	5,6769	6,472	-,042	,713
IAnruf1	6,2769	4,953	,571	,501
IFreund	5,4154	5,965	,248	,602
IDruck	6,1077	5,066	,648	,493
IAnruf2	6,0154	4,609	,580	,483
IKind	6,1385	5,090	,422	,546
IAnruf3	6,2769	5,891	,128	,650

Die Skalenreliabilität unter Berücksichtigung aller sieben Items liegt mit 0,62 unter der allgemein gültigen Anforderung von 0,7 – 0,8 für brauchbare Skalen (Tabelle 7).

Tab. 7: Skalenreliabilität

Reliabilitätsstatistiken

Cronbachs Alpha	Anzahl der Items
,616	7

Bezüglich der Itemtrennschärfe weist insbesondere Item 1 einen sehr geringen Wert auf, welcher durch das negative Vorzeichen darauf schließen lässt, dass dieses Item nicht entsprechend den anderen Items in dieselbe Richtung korreliert. Es ist anzunehmen, dass durch die gegebene Situation zu Beginn des Tests dieses Item nicht aufgrund der zugrundeliegenden Entscheidungsmuster beantwortet wird, sondern für die Entscheidung vielmehr unmittelbar zeitliche Faktoren eine Rolle spielen.

In Folge wurde das Item 1 aufgrund einer zu geringen Itemtrennschärfe ausgeschlossen, wodurch sich eine angepasste Skalenreliabilität von 0,71 ergibt (Tabelle 8). Die angepassten Itemtrennschärfen sind in Tabelle 9 ersichtlich.

Tab. 8: Skalenreliabilität nach Itemausschluss

Reliabilitätsstatistiken

Cronbachs Alpha	Anzahl der Items
,713	6

Tab. 9: Itemtrennschärfen nach Itemausschluss

Item-Skala-Statistiken

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item- Skala-Korrelation	Cronbachs Alpha, wenn Item weggelassen
IAnruf1	4,9692	4,530	,575	,634
IFreund	4,1077	5,504	,251	,724
IDruck	4,8000	4,663	,643	,625
IAnruf2	4,7077	4,210	,579	,627
IKind	4,8308	4,580	,454	,671
IAnruf3	4,9692	5,030	,250	,742

8 Skalenanalyse

Im Folgenden wird die Dimensionalität der Skalen Reflexiv, Flexibel und Impulsiv überprüft. Hierzu wurden die entsprechenden Items jeweils einer explorativen Faktorenanalyse unterzogen, um zu eruieren ob hinter den jeweiligen Skalen lediglich ein Faktor steht. Abschließend wurde über alle drei Skalen eine weitere explorative Faktorenanalyse berechnet, um die angenommene Eindimensionalität bezüglich aller drei Skalen zu bestätigen.

Aus diesem Grunde wurde eine nicht rotierte Faktorenanalyse nach der Hauptkomponentenmethode durchgeführt.

Aufgrund der zuvor angeführten Schwächen im Bereich der Trennschärfe des Items 1 wurde selbiges von der Untersuchung ausgeschlossen. Zur Bestimmung der Faktorenanzahl wurde eine graphische Kontrolle anhand des Screeplots vorgenommen. Bortz (1999) legt fest, dass im Zuge des „Scree-Test“ all diejenigen Faktoren als relevant angesehen werden, deren Eigenwerte sich im Screeplot vor dem „Knick“ befinden.

Es zeigte sich, dass die Eindimensionalität der Skalen als gegeben angesehen werden kann. Im Folgenden die Ergebnisse der Faktorenanalyse der einzelnen Skalen.

Ergebnisse für die Skala Impulsiv

In Abbildung 20 ist ein klar ersichtlicher „Knick“ in der graphisch dargestellten Faktorenlösung zu erkennen, welcher für eine einfaktorielle Lösung spricht.

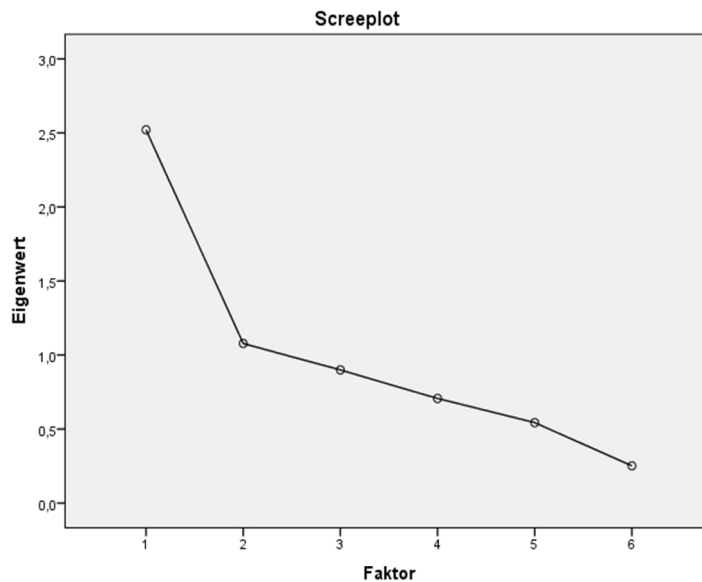


Abb. 20: Screeplot für Skala Impulsiv

Ergebnisse für die Skala Flexibel

Abbildung 21 zeigt einen klar ersichtlichen „Knick“ in der graphisch dargestellten Faktorenlösung, welcher für eine einfaktorielle Lösung spricht.

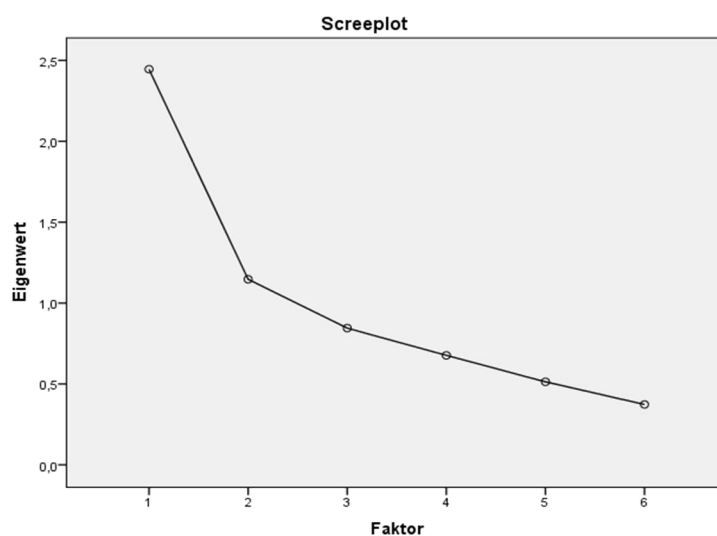


Abb. 21: Screeplot für Skala Flexibel

Ergebnisse für die Skala Reflexiv

In Abbildung 22 ist ein klar ersichtlicher „Knick“ in der graphisch dargestellten Faktorenlösung über dem Eigenwertkriterium von 1 zu erkennen, welcher für eine einfaktorielle Lösung spricht.

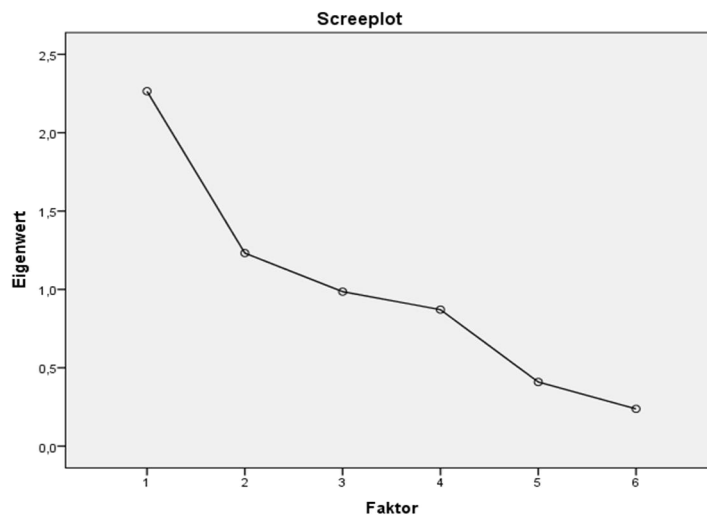


Abb. 22: Screeplot für Skala Reflexiv

Ergebnisse für die drei Skalen

In Abbildung 23 ist ein klar ersichtlicher „Knick“ in der graphisch dargestellten Faktorenlösung über dem Eigenwertkriterium von 1 zu erkennen, welcher für eine dreifaktorielle Lösung spricht. Es kann somit angenommen werden, dass allen drei Skalen jeweils ein von den anderen unabhängiger Faktor zugrunde liegt.

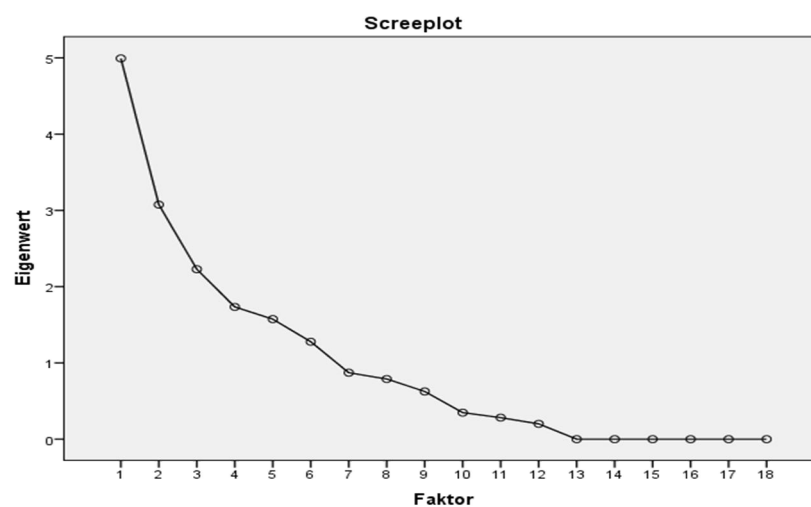


Abb. 23: Screeplot für alle 3 Skalen

9 Clusteranalyse

Kernstück des ILICA war die Typisierung der Testpersonen in drei verschiedene Stilrichtungen der Entscheidungsfindung.

Da der Everest auf denselben theoretischen Konzepten beruht und im Hinblick auf die drei Typisierungen des ILICA konzipiert wurde, wurde eine Clusteranalyse berechnet. Ziel dieses Verfahrens ist es, eine Klassifikation der vorliegenden Daten hinsichtlich der Ähnlichkeit der einzelnen Objekte vorzunehmen. Es wird dabei versucht, eine bestimmte Anzahl an Clustern zu extrapolieren für welche gilt, dass die einzelnen Objekte der jeweiligen Cluster untereinander eine möglichst hohe Ähnlichkeit aufweisen, sich gegenüber Objekten der anderen Cluster jedoch maximal unterscheiden. Im Detail findet sich eine Beschreibung dieses Verfahrens beispielsweise bei Kubinger, Rasch und Yanagida (2011).

Angewandt wurde im Folgenden eine hierarchische Clusteranalyse. Als Distanzmaß wurde die quadrierte euklidische Distanz herangezogen und die Clusterzuordnung wurde mittels der Methode nach Ward berechnet. Aufgrund der unterschiedlichen Skalenniveaus der berücksichtigten Variablen, wurden die entsprechenden Testkennwerte im Vorfeld z-standardisiert.

Folgende clusterbildende Variablen wurden zur Berechnung herangezogen:

- ReflexivNeu (Summenscore)
- ImpulsivNeu (Summenscore)
- FlexibelNeu (Summenscore)
- Arbeitsscore (Summenscore)

Der Arbeitsscore setzt sich hierbei aus all denjenigen Variablen zusammen, die eine Aufgabe beschreiben, welche die Testperson im Arbeitskontext über den Tag verteilt erledigen sollte. Zur Berechnung der Cluster konnten alle 65 Testpersonen herangezogen werden, da keine fehlenden Werte vorlagen.

Abbildung 24 zeigt anhand der „Rest“-Komponenten (siehe Abbildung 2) eine 3-Cluster-Lösung, deutlich zu erkennen an dem „Knick“ im Übergang der 3-Cluster-Lösung zur 2-Cluster-Lösung.

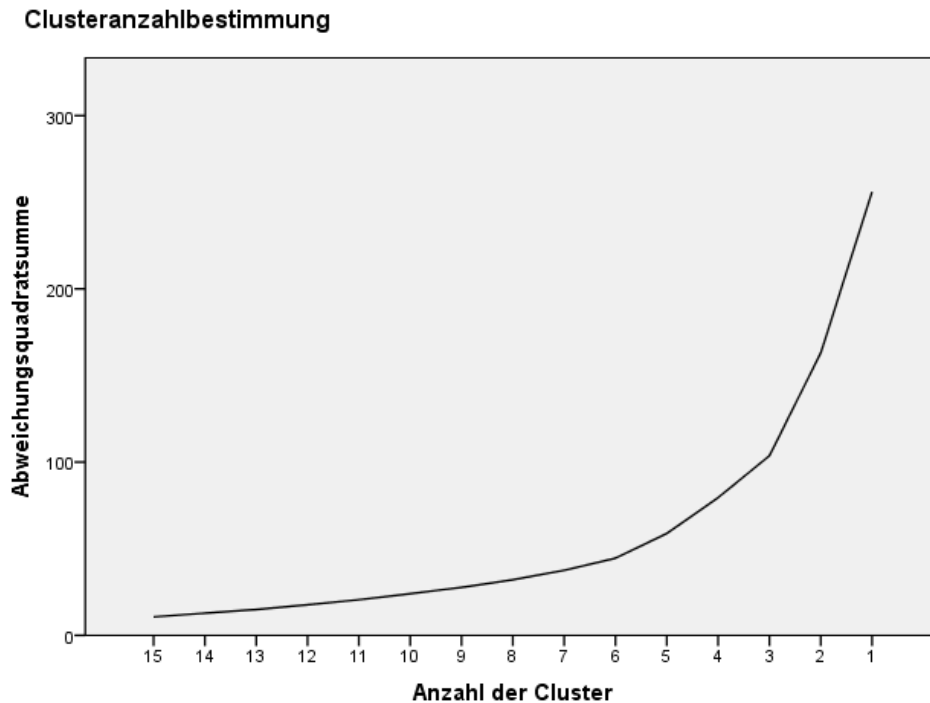


Abb. 24 : Bestimmung der Clusteranzahl

9.1 Deskriptive Darstellung der Cluster

Um die Unterschiede der vorliegenden drei Cluster inhaltlich einzuordnen und die Cluster infolge charakterisieren zu können, wurde eine einfaktorielle Varianzanalyse durchgeführt, um zu überprüfen ob zwischen den drei Cluster-Lösungen ($N = 65$) signifikante Unterschiede bestehen.

Die Annahme der Normalverteilung erscheint aufgrund der vorliegenden Histogramme verletzt. Allerdings beschreibt unter anderem Bortz (1999), dass der F-Test der Varianzanalyse auch bei kleineren Stichproben wie den hier gegebenen, robust gegenüber Nicht-Normalität ist.

Zur Überprüfung der Annahme der Homogenität der Varianzen wurde ein Levene's Test (siehe Tabelle 10) durchgeführt, welcher mit $F(2,62) = 7,52$, $p = 0,001$ für die Variable Arbeitsscore, und für die Variable Flexibel mit $F(2,62) = 3,78$, $p = 0,028$, sowie für die Variable Reflexiv mit $F(2,62) = 7,77$, $p = 0,001$ als signifikant einzustufen ist. Infolge ist somit die zugrundeliegende Annahme der Homogenität der Varianzen als verletzt einzustufen. Aufgrund dessen wurde auf diese Variablen der Welch-Test angewendet, um auf signifikante Unterschiede innerhalb der Gruppen zu testen.

Für die übrige Variable Impulsiv $F(2,62) = 3,04$, $p = 0,055$ gilt die Annahme der Homogenität der Varianzen als nicht verletzt.

Tab. 10: Levene's Test für die clusterbildenden Variablen

Test der Homogenität der Varianzen				
	Levene-Statistik	df1	df2	Signifikanz
ArbeitsScore	7,524	2	62	,001
RefNeu	7,768	2	62	,001
FlexNeu	3,780	2	62	,028
ImpNeu	3,036	2	62	,055

Tabelle 11 bietet eine Übersicht der Mittelwerte der clusterbildenden Variablen auf die im Folgenden Bezug genommen wird. Den Tabellen 12 bis 15 können die im Folgenden diskutierten Werte in tabellarischer Form entnommen werden.

Tab. 11: Mittelwerte der Cluster-bildenden Variablen

Mittelwerte

Cluster	ArbeitsScore	RefNeu	FlexNeu	ImpNeu
1	15,2500	2,8333	2,2917	,8750
2	12,3750	,2500	1,6250	4,1250
3	13,4000	,8400	4,2000	,9600
Insgesamt	13,8308	1,4308	2,8615	1,7077

Reflexiv

Bezüglich der Ausprägung im Bereich Reflexiv zeigte sich ein signifikanter Effekt der Clusterzugehörigkeit bei einem $p < ,05$ Wert für die drei Bedingungen [Welch's $F(2, 34,67) = 47,74, p = 0,001$].

Zum Vergleich der Mittelwerte wurde in diesem Fall ein Post-Hoc-Test nach Tamhane durchgeführt. Die Ergebnisse verdeutlichen, dass sich die Cluster 1 ($M = 2,83, SD = 1,17$), Cluster 2 ($M = 0,25, SD = 0,45$) und Cluster 3 ($M = 0,84, SD = 1,35$) signifikant voneinander unterscheiden.

Flexibel

Für die Variable Flexibel zeigte sich ein signifikanter Effekt der Clusterzugehörigkeit bei einem $p < ,05$ Wert für die drei Bedingungen [Welch's $F(2, 40,86) = 62,46, p = 0,001$].

Zum Vergleich der Mittelwerte wurde in diesem Fall ein Post-Hoc-Test nach Tamhane durchgeführt. Die Ergebnisse verdeutlichen, dass Cluster 3 ($M = 4,20, SD = 0,96$) sich signifikant von Cluster 2 ($M = 1,63, SD = 0,50$) und Cluster 1 ($M = 2,29, SD = 1,16$) unterscheidet. Zwischen den Clustern 2 und 1 finden sich hingegen keine signifikanten Unterschiede.

Impulsiv

Im Bereich der Variable Impulsiv zeigte sich ein signifikanter Effekt der Clusterzugehörigkeit bei einem $p < ,05$ Wert für die drei Bedingungen [$F(2, 62) = 103,05, p = 0,001$].

Ein Post-Hoc-Test Vergleich nach der Scheffé-Prozedur deutet darauf hin, dass sich die Mittelwerte der Cluster 1 ($M = 0,88, SD = 0,74$) und Cluster 2 ($M = 4,13, SD = 0,62$) signifikant voneinander unterscheiden. Cluster 3 ($M = 0,96, SD = 0,89$) unterscheidet sich zwar signifikant von Cluster 2, jedoch nicht signifikant von Cluster 1.

Arbeitsscore

Es zeigte sich ein signifikanter Effekt der Clusterzugehörigkeit auf das Ausmaß des Merkmals Arbeitsscore bei einem $p < ,05$ Wert für die drei Bedingungen [Welch's $F(2, 38,32) = 15,96, p = 0,001$].

Zum Vergleich der Mittelwerte wurde auch in diesem Fall ein Post-Hoc-Test nach Tamhane durchgeführt. Die Ergebnisse verdeutlichen, dass sich die Cluster 1 ($M = 15,25, SD = 2,17$), Cluster 2 ($M = 12,38, SD = 1,02$) und Cluster 3 ($M = 13,40, SD = 1,12$) signifikant voneinander unterscheiden.

Zusammenfassend lässt sich somit sagen, dass Cluster 1 durch hohe Werte in der Variable Reflexiv charakterisiert wird. Gleichzeitig erreichen Personen dieses Clusters die höchsten Werte im Bereich des Arbeitsscores. Cluster 2 hingegen wird durch die Variable Impulsiv charakterisiert. Personen dieses Clusters erreichen zudem den niedrigsten Arbeitsscore. Cluster 3 schließlich wird charakterisiert durch die Variable Flexibilität, hier weist dieser Cluster die höchsten erreichten Ausprägungen auf. Der erreichte Arbeitsscore von Personen dieses Clusters liegt hierbei im Mittelfeld.

Tab. 12: ANOVA - Impulsiv

Einfaktorielle ANOVA

ImpNeu					
	Quadratsumme	df	Mittel der Quadratsumme	F	Signifikanz
Zwischen den Gruppen	124,111	2	62,056	103,052	,000
Innerhalb der Gruppen	37,335	62	,602		
Gesamt	161,446	64			

Tab. 13: Post-Hoc-Test Impulsiv

Mehrfachvergleiche

Abhängige Variable: ImpNeu

Scheffé-Prozedur

(I) Ward Method	(J) Ward Method	Mittlere Differenz (I-J)	Standardfehler	Signifikanz
1	2	-3,25000*	,25045	,000
	3	-,08500	,22176	,929
2	1	3,25000*	,25045	,000
	3	3,16500*	,24844	,000
3	1	,08500	,22176	,929
	2	-3,16500*	,24844	,000

*. Die Differenz der Mittelwerte ist auf dem Niveau 0.05 signifikant.

Tab. 14: Welch-Test Ergebnisse Reflexiv, Flexibel, Arbeitsscore

Robuste Testverfahren zur Prüfung auf Gleichheit der Mittelwerte					
		Statistik ^a	df1	df2	Sig.
RefNeu	Welch-Test	47,741	2	34,653	,000
FlexNeu	Welch-Test	62,460	2	40,862	,000
ArbeitsScore	Welch-Test	15,956	2	38,321	,000

a. Asymptotisch F-verteilt

Tab. 15: Post-Hoc-Tests Reflexiv, Flexibel, Arbeitsscore

Mehrfachvergleiche					
Tamhane					
Abhängige Variable	(I) Ward Method	(J) Ward Method	Mittlere Differenz (I-J)	Standardfehler	Signifikanz
RefNeu	1	2	2,58333*	,26318	,000
		3	1,99333*	,24973	,000
	2	1	-2,58333*	,26318	,000
		3	-,59000*	,13454	,000
	3	1	-1,99333*	,24973	,000
		2	,59000*	,13454	,000
FlexNeu	1	2	,66667	,26779	,053
		3	-1,90833*	,30455	,000
	2	1	-,66667	,26779	,053
		3	-2,57500*	,22867	,000
	3	1	1,90833*	,30455	,000
		2	2,57500*	,22867	,000
ArbeitsScore	1	2	2,87500*	,51204	,000
		3	1,85000*	,49655	,002
	2	1	-2,87500*	,51204	,000
		3	-1,02500*	,34004	,014
	3	1	-1,85000*	,49655	,002
		2	1,02500*	,34004	,014

*. Die Differenz der Mittelwerte ist auf dem Niveau 0.05 signifikant.

9.2 Weiterführende deskriptive Darstellung der Cluster

Im Folgenden wurde der Datensatz anhand der gefundenen Cluster-Lösung hinsichtlich relevanter Testkennwerte analysiert. Es wurden hierzu weitere einfaktorielle Varianzanalysen berechnet, welche die Cluster als Gruppierungsvariable verwenden, um wiederum zu überprüfen ob signifikante Unterschiede zwischen den drei Cluster-Lösungen ($N = 65$) existieren.

Ziel dieser Analysen ist es, die gefundenen drei Cluster inhaltlich weitergehend zu beschreiben um letztendlich eine fundierte Typisierung zu ermöglichen.

Zur weiteren Analyse wurden die Testbearbeitungszeit in Minuten, die Verwendung des Notizblocks, das Aufrufen der Uhr, die Ernsthaftigkeit, die Belastung sowie das Ausmaß an Freizeitaktivitäten herangezogen.

Der letztgenannte Wert stellt einen Summenscore über die Variablen „Spielen“, „Pause“ sowie „Kaffee“ dar und gibt somit an, wie oft die Testperson Aktivitäten ausgeführt hat, die in keinerlei Kontext zum Arbeitsalltag stehen. Diese Aktivitäten umfassen das Spielen von Tic-Tac-Toe, das selbstständige Einlegen einer Kaffeepause sowie das Pausieren außerhalb des Bürogebäudes.

Die Annahme der Normalverteilung wird aufgrund der zuvor erwähnten Aspekte bei Bortz (1999) im Folgenden außer Acht gelassen.

Zur Überprüfung der Annahme der Homogenität der Varianzen wurde ein Leven's Test durchgeführt (siehe Tabelle 16), welcher mit $F(2,62) = 1,63$, $p = 0,205$ für die Variable Freizeit, mit $F(2,62) = 2,44$, $p = 0,096$ für die Variable Ernsthaftigkeit, mit $F(2,62) = 2,20$, $p = 0,119$ für die Variable Belastung, mit $F(2,62) = 1,25$, $p = 0,295$ für die Variable Uhr, mit $F(2,62) = 1,74$, $p = 0,183$ für die Variable Notizblock, sowie mit $F(2,62) = 2,45$, $p = 0,094$ für die Variable Minuten ein nicht signifikantes Ergebnis lieferte. Für diese Variablen kann infolge die Homogenität der Varianzen als gegeben angesehen werden. Den Tabellen 17 und 18 können die im Folgenden diskutierten Werte in tabellarischer Form entnommen werden.

Tab. 16: Levene's Test nicht clusterbildende Variablen

Test der Homogenität der Varianzen				
	Levene-Statistik	df1	df2	Signifikanz
Ernst	1,247	2	62	,295
Belstung	2,204	2	62	,119
Notizblock	1,743	2	62	,183
Hilfe	,575	2	62	,566
Freizeit	1,627	2	62	,205
Minuten	2,451	2	62	,094

Bearbeitungszeit

Es zeigte sich ein signifikanter Effekt der Clusterzugehörigkeit auf das Ausmaß der Bearbeitungszeit bei einem $p < ,05$ Wert für die drei Bedingungen [$F(2,62) = 20,68$, $p = 0,001$].

Ein Post-Hoc-Test Vergleich nach der Scheffé-Prozedur deutet darauf hin, dass der Mittelwert des Cluster 2 ($M = 22,63$, $SD = 2,33$) sich signifikant von dem des Cluster 1 ($M = 28,08$, $SD = 1,98$), sowie des Cluster 3 ($M = 27,24$, $SD = 3,53$) unterscheidet. Allerdings zeigte sich kein signifikanter Effekt zwischen Cluster 1 und Cluster 3.

Freizeit

Hierbei zeigte sich ein signifikanter Effekt der Clusterzugehörigkeit auf das Ausmaß der Freizeitaktivitäten bei einem $p < ,05$ Wert für die drei Bedingungen [$F(2, 62) = 7,08$, $p = 0,002$].

Zum Vergleich der Mittelwerte wurde ein Post-Hoc-Test nach der Scheffé-Prozedur durchgeführt. Es zeigt sich, dass Cluster 1 ($M = 1,83$, $SD = 1,37$) sich sowohl von Cluster 2 ($M = 0,88$, $SD = 0,96$) als auch Cluster 3 ($M = 0,64$, $SD = 1,04$) signifikant unterscheidet. Cluster 2 und Cluster 3 unterscheiden sich hingegen nicht signifikant voneinander.

Verwendung des Notizblocks

In der Häufigkeit der Verwendung des Notizblocks als Hilfestellung zeigten die gefundenen drei Cluster keine signifikanten Unterschiede [$F(2, 62) = 1,22, p = 0,302$].

Verwendung der Uhr

Es zeigte sich kein signifikanter Unterschied zwischen den gefundenen Cluster bezüglich der Verwendung der Uhr als Hilfestellung [$F(2, 62) = 1,89, p = 0,159$].

Ernsthaftigkeit

Im Hinblick auf die angegebene Ernsthaftigkeit der Testpersonen ließ sich über die drei Gruppen der Clusterzugehörigkeit bei einem $p > ,05$ Wert kein signifikanter Effekt finden [$F(2, 62) = 0,23, p = 0,793$].

Belastung

Bezüglich der angegebenen Belastung der Testpersonen zeigte sich kein signifikanter Effekt der Clusterzugehörigkeit bei einem $p > ,05$ Wert für die drei Bedingungen [$F(2, 62) = 1,43, p = 0,248$].

Tab. 17: ANOVA Ergebnisse nicht clusterbildende Variablen

Einfaktorielle ANOVA						
		Quadratsumme	df	Mittel der Quadrate	F	Signifikanz
Ernst	Zwischen Gruppen	138,809	2	69,404	,233	,793
	Innerhalb Gruppen	18493,438	62	298,281		
	Gesamt	18632,246	64			
Belstung	Zwischen Gruppen	764,629	2	382,314	1,428	,248
	Innerhalb Gruppen	16599,156	62	267,728		
	Gesamt	17363,785	64			
Notizblock	Zwischen Gruppen	51,511	2	25,756	1,221	,302
	Innerhalb Gruppen	1308,335	62	21,102		
	Gesamt	1359,846	64			
Hilfe	Zwischen Gruppen	91,795	2	45,898	1,894	,159
	Innerhalb Gruppen	1502,143	62	24,228		
	Gesamt	1593,938	64			
Freizeit	Zwischen Gruppen	18,911	2	9,455	7,076	,002
	Innerhalb Gruppen	82,843	62	1,336		
	Gesamt	101,754	64			
Minuten	Zwischen Gruppen	313,641	2	156,821	20,681	,000
	Innerhalb Gruppen	470,143	62	7,583		
	Gesamt	783,785	64			

Tab. 18: Post-Hoc-Tests Freizeit und Bearbeitungszeit

Mehrfachvergleiche

Scheffé-Prozedur

Abhängige Variable	(I) Ward Method	(J) Ward Method	Mittlere Differenz (I-J)	Standardfehler	Signifikanz
Freizeit	1	2	,95833 [*]	,37308	,043
		3	1,19333 [*]	,33034	,003
	2	1	-,95833 [*]	,37308	,043
		3	,23500	,37008	,818
	3	1	-1,19333 [*]	,33034	,003
		2	-,23500	,37008	,818
Minuten	1	2	5,45833 [*]	,88876	,000
		3	,84333	,78694	,566
	2	1	-5,45833 [*]	,88876	,000
		3	-4,61500 [*]	,88162	,000
	3	1	-,84333	,78694	,566
		2	4,61500 [*]	,88162	,000

*. Die Differenz der Mittelwerte ist auf dem Niveau 0.05 signifikant.

Zusammenfassend lässt sich sagen, dass sich die gefundenen Cluster bezüglich der Variablen Bearbeitungszeit sowie Freizeit signifikant voneinander unterscheiden. Cluster 2 (Impulsiv) setzt sich aus Personen zusammen, welche den Test signifikant schneller bearbeitet haben als Personen der anderen Cluster. Cluster 1 (Reflexiv) setzt sich aufgrund der vorliegenden Ergebnisse aus Personen zusammen, welche signifikant mehr nicht relevante Arbeitstätigkeiten ausführen.

Bezüglich der Verwendung von Hilfsmitteln wie der Uhr oder des Notizblocks besteht zwischen den Clustern kein Unterschied. Obwohl sich in den Variablen Ernsthaftigkeit und Belastung aufgrund der Mittelwerte ein Trend erkennen lässt, fielen die vorliegenden Unterschiede der Mittelwerte jedoch nicht signifikant aus, wodurch davon ausgegangen werden kann, dass sich die Cluster diesbezüglich nicht unterscheiden.

9.3 Inhaltliche Beschreibung der Cluster

In Anlehnung an die von Möseneder und Ebenhöf (1996) vorgeschlagene Typisierung, lassen sich die gefundenen drei Cluster den entsprechenden Entscheidungstypen zuordnen. Im Folgenden sollen die gefundenen drei Cluster aufgrund der zuvor aufgeführten Ergebnisse inhaltlich beschrieben werden.

Cluster 1 – Der reflexive Entscheidungstyp

Cluster 1 setzt sich aus 24 Personen zusammen und zeichnet sich durch einen überdurchschnittlich hohen Wert im reflexiven Bereich aus.

Personen dieses Typs erreichen zudem überdurchschnittliche Werte im Bereich des Arbeitsscores, sprich sie erledigen signifikant mehr der ihnen gestellten Aufgaben als alle anderen Entscheidungstypen. Die Zielerreichung scheint somit wie in der Definition der Entscheidungstypen eingangs festgelegt einen zentralen Stellenwert einzunehmen.

Charakteristisch ist zudem die unterdurchschnittliche Ausprägung in den Bereichen Impulsiv sowie Flexibel. Zur Erreichung ihrer Leistungen benötigen sie hierbei signifikant mehr Zeit als Personen des Typs 2 – des impulsiven Entscheidungstyps. Die überdurchschnittliche Bearbeitungszeit scheint hierbei durch das erhöhte Ausmaß an Freizeitaktivitäten bedingt zu sein. Personen dieses Typs erledigen die ihnen gestellten Aufgaben vor Ablauf des virtuellen Testtages, wodurch sich ihnen die Möglichkeit bietet nicht geplante Aktivitäten, wie Spielen oder Pausen in den Tagesablauf zu integrieren.

Cluster 2 – Der impulsive Entscheidungstyp

Cluster 2 setzt sich aus 16 Personen zusammen. Bezeichnend für diesen Cluster ist hierbei der überdurchschnittlich hohe Wert im Bereich des Merkmals Impulsiv, sowie der unterdurchschnittliche Wert in dem Merkmal Reflexiv. Einhergehend mit der impulsiven Arbeitsweise zeigen Personen dieses Clusters zudem eine signifikant geringere Bearbeitungszeit als Personen der anderen Cluster. Durch den fehlenden kognitiven Aufwand beim Abwägen verschiedener Entscheidungsoptionen, erfolgt die Entscheidungsfindung schneller wodurch die Bearbeitungszeit sinkt. Aufgrund der impulsiven Arbeitsweise erreichen Personen dieses Typs den geringsten Arbeitsscore der drei verschiedenen Typen. Personen dieses Typs lassen sich leicht zu Dingen überreden beziehungsweise verleiten, die nicht geplant sind und im Widerspruch zur Zielerreichung stehen, wodurch es aufgrund zeitlicher Ressourcenknappheit nicht möglich ist alle der ihnen gestellten Aufgaben zu erledigen

Cluster 3 – Der flexible Entscheidungstyp

Cluster 3 stellt mit 25 Personen den größten der drei Cluster dar. Gekennzeichnet sind Personen dieses Typs durch einen überdurchschnittlichen Wert im Bereich des Merkmals Flexibel, sowie eine mittlere Ausprägung im Merkmal Reflexiv.

Ihre Zielerreichung liegt mit einem durchschnittlichen Wert im Bereich der erbrachten Arbeitsleistungen zwischen denen des reflexiven und des impulsiven Entscheidungstyps.

Auch im Bereich der Bearbeitungszeit zeigt sich ein durchschnittliches Bild. So liegt die Bearbeitungszeit dieses Typs zwischen denen des reflexiven und impulsiven Entscheidungstyps, unterscheidet sich hierbei jedoch nicht signifikant von der Zeit des Reflexiven Typs. In Anbetracht dieser Tatsache erscheint der nicht signifikant erhöhte Freizeit-Wert ein Indiz dafür zu sein, dass Personen des flexiblen Entscheidungstyps ihre Entscheidungen gegeneinander abwägen um so die Zielerreichung in zufriedenstellendem Ausmaß sicherzustellen.

10 Gütekriterien

10.1 Skalierung

„Ein Test erfüllt das Gütekriterium der Skalierung, wenn die laut Verrechnungsvorschriften resultierenden Testwerte die empirischen Verhaltensrelationen adäquat abbilden.“ (Kubinger, 2009, S. 82).

Aufgrund der geringen Anzahl an Testpersonen in der erhobenen Stichprobe ließ sich die Anforderung der Skalierung nicht mittels eines Item-Response-Theorie-Modells überprüfen. Weiterführende Analysen diesbezüglich sind in kommenden Studien durchzuführen.

10.2 Reliabilität

„Die Reliabilität eines Tests beschreibt den Grad der Genauigkeit, mit dem er ein bestimmtes psychisches Merkmal misst, gleichgültig, ob er dieses Merkmal auch zu messen beansprucht.“ (Kubinger, 2009, S. 49).

Bezüglich der Skalenreabilität erscheint der Everest eine zufriedenstellende innere Konsistenz aufzuweisen mit einem Cronbach-alpha von 0,71.

Bei der Erfassung der Bearbeitungszeit kann aufgrund der computerbasierten Vorgabe von maximaler Messgenauigkeit ausgegangen werden.

Bezüglich der arbeitsrelevanten Aufgaben kann zurzeit keine Aussage in Hinblick auf die Reliabilität getroffen werden.

10.3 Validität

„Unter Validität eines Tests ist zu verstehen, dass er tatsächlich jenes psychische Merkmal misst, welches er zu messen behauptet.“ (Kubinger, 2009, S. 55).

Zu diesem Punkt der Gütekriterien lassen sich bislang keine Aussagen treffen, da entsprechende Validitätstudien noch ausständig sind.

10.4 Objektivität

„Unter Objektivität eines Tests ist zu verstehen, dass die mit ihm gewonnenen Ergebnisse unabhängig vom Untersucher sind.“ (Kubinger, 2009, S. 38).

Da es sich beim Everest um ein computergestütztes Verfahren handelt, wodurch die Vorgabe und Auswertung automatisiert abläuft, ist die Testleiterunabhängigkeit sowie Verrechnungssicherheit als gesichert anzusehen. Die Interpretationseindeutigkeit ist zum jetzigen Zeitpunkt aufgrund fehlender Normierungen noch nicht gegeben.

10.5 Ökonomie

„Ein Test erfüllt das Gütekriterium Ökonomie, wenn er, gemessen am diagnostischen Informationsgewinn, relativ wenig Ressourcen (Zeit und Geld) beansprucht.“ (Kubinger, 2009, S. 98).

Mit einer durchschnittlichen Bearbeitungszeit von 26 Minuten und einer maximalen Bearbeitungszeit von 34 Minuten lässt sich der Everest als ökonomisch einstufen. Aufgrund der automatisierten Instruktionen gestaltet er sich auch im Hinblick auf die vom Testleiter zu investierende Zeit als ökonomisch. Darüber hinaus sorgen die geringen Systemanforderungen an die verfügbare Hardware dafür, dass der Test auch auf älteren beziehungsweise günstigeren Computern problemlos lauffähig ist, wodurch Finanzielle Ressourcen geschont werden.

10.6 Eichung

„Ein Test erfüllt das Gütekriterium Eichung, wenn für sein Bezugssystem zur Relativierung des individuellen Testergebnisses (die sog. „Eichtabellen“) folgende Bedingungen gegeben sind: Die Eichtabellen sind gültig, d.h. nicht veraltet, die Population, für die die Eichtabellen gelten, ist definiert, die für die Erstellung der Eichtabellen herangezogene Stichprobe ist repräsentativ.“ (Kubinger, 2009, S. 68).

Aufgrund der nicht repräsentativen Stichprobe welche lediglich über einen Umfang von 65 Personen verfügt, mangelt es dem Test an den Voraussetzungen einer gültigen Normierung.

10.7 Fairness

„Ein Test erfüllt das Gütekriterium Fairness, wenn die resultierenden Testwerte zu keiner systematischen Diskriminierung bestimmter Testpersonen zum Beispiel aufgrund ihrer ethnischen, soziokulturellen oder geschlechtsspezifischen Gruppenzugehörigkeit führen.“ (Kubinger, 2009, S. 123).

Keine der Testpersonen äußerte Schwierigkeiten bei der Bearbeitung des Tests. Die vorliegenden empirischen Ergebnisse lassen nicht darauf schließen, dass eine systematische Beeinträchtigung einer bestimmten Gruppe an Testpersonen existiert. Weiterführende Untersuchungen diesbezüglich sollten in zukünftigen Studien durchgeführt werden um den Umstand der aktuell zu kleinen Stichprobe zu begegnen. Aufgrund der Vielzahl verbaler Testinhalte empfiehlt sich eine Überprüfung, ob eine Beeinträchtigung für Personen mit nicht deutscher Muttersprache vorliegt.

10.8 Zumutbarkeit

„Ein Test erfüllt das Gütekriterium Zumutbarkeit, wenn er die Testperson absolut und relativ zu dem aus seiner Anwendung resultierenden Nutzen in zeitlicher, psychischer (insbesondere energetisch-motivationaler und emotionaler) sowie körperlicher Hinsicht schont.“ (Kubinger, 2009, S. 116).

Aufgrund der vorliegenden Ergebnisse ist ersichtlich, dass der Test als nicht übermäßig belastend empfunden wurde. Die spielerische Natur der gestellten Situation sorgt im Vergleich zu anderen Verfahren zu einer gesteigerten Bereitschaft zur Teilnahme an einer psychologischen Testung.

10.9 Unverfälschbarkeit

„Ein Test erfüllt das Gütekriterium der Unverfälschbarkeit, wenn die getestete Person ihr Testergebnis nicht oder nur unwesentlich nach eigenem Belieben beeinflussen kann.“ (Kubinger, 2009, S. 120).

Aufgrund der Tatsache, dass die eigentliche Messintention des Tests der Testperson im Vorfeld nicht klar ersichtlich ist, ist davon auszugehen, dass ein gezieltes Manipulieren des Testverhaltens in eine gewisse Richtung nicht möglich ist. Bezüglich der Erreichung der gestellten arbeitsrelevanten Aufgaben mag die Augenscheinvalidität als hoch einzustufen sein, jedoch messen diese ein Leistungsverhalten, welches von der Testperson nicht besser gestaltet werden kann als es sich tatsächlich darstellt. Im Zuge dessen ist dem Test somit grundlegend Unverfälschbarkeit zu attestieren. Studien hierzu sehen jedoch noch aus.

11 Diskussion und Ausblicke

Eine Neuprogrammierung des Everest erscheint aufgrund der Limitierungen, die die Entwicklungsumgebung „Adventure Game Studio“ mit sich bringt als angebracht. Auf diese Weise würde sich die Möglichkeit zur Erfassung weiterer Kennwerte eröffnen. Denkbar wäre beispielsweise die Erfassung der Bearbeitungszeit pro Item. Selbiges wurde in der aktuellen Auflage nicht erhoben, da die Entwicklungsumgebung zu diesem Zweck nicht konzipiert wurde und ein adäquates Abbilden im Millisekundenbereich nicht möglich ist.

Um den Immersionseffekt zu verstärken wäre es denkbar, Abstand von den vorgefertigten Avataren zu nehmen. Hilfreich wäre es hierbei der Testperson die Möglichkeit zu bieten sich mittels weniger Mausklicks ihren eigenen Avatar zu erstellen um sich bestmöglich mit diesem identifizieren zu können. Um die Testpersonen bei dieser Aufgabenstellung nicht zu überfordern könnte sich die Individualisierung auf einzelne Aspekte wie Hautfarbe, Haarfarbe und Kleidung beschränken. Dies wäre auch angebracht, um die Testbearbeitungszeit nach wie vor möglichst gering zu halten.

Einhergehend mit einer Neuauflage wäre auch die grafische Gestaltung der Testumgebungen zu überarbeiten. Die aktuelle Auflage nutzt ein rein zweidimensionales Design. Möglich wäre hierbei jedoch auch ein sogenanntes zweieinhalbdimensionales Design. Hierbei werden dreidimensionale Hintergründe mit zweidimensionalen Avataren verbunden. Das könnte die Realitätsnähe des Testsettings erhöhen, ohne einen höheren Aufwand hinsichtlich der Bedienung für die Testperson zu bedeuten.

Abgesehen von den inhaltlichen Aspekten des Tests sollte in zukünftigen Studien darauf Wert gelegt werden, die Geltung der Testgütekriterien wie im obigen Abschnitt angesprochen zu überprüfen. Gerade die Limitierungen aufgrund des geringen Stichprobenumfangs sollten in weiteren Studien adressiert werden. Aktuelle Aussagen bezüglich der gefundenen Typen sind wegen der geringen Anzahl an Personen pro Cluster nicht generalisierbar. Cluster 2 setzt sich beispielsweise aus nur 16 Personen zusammen, aber auch die anderen Cluster liegen unter $n = 30$.

Auf Basis der gewonnenen Informationen ließe sich infolge eine Auswertung der Testkennwerte am Ende des Testdurchlaufes anfügen. Hierbei sollte dem Testleiter auch in grafisch anschaulicher Weise adäquat dargestellt werden wie die Ergebnisse der Testperson einzuordnen sind, um diese bestmöglich an die Testperson rückzumelden.

Auch bietet sich zur weiteren Überprüfung eine genauere Betrachtung der gefunden Typen an. Hierbei wäre es ebenfalls denkbar, dass der bislang nicht direkt erhobene Aspekt der sozialen Kompetenz einen direkten Einfluss auf den Entscheidungsprozess hat. Bislang leistet der Everest eine Differenzierung der drei Typen hinsichtlich der Arbeitseffizienz. Es werden der impulsive Typ, welcher die geringste Arbeitseffizienz zeigt, der flexible Typ, welcher eine moderate Arbeitseffizienz zeigt sowie der reflexive Typ, welcher die höchste Arbeitseffizienz zeigt, unterschieden. Die Items die zur Typisierung herangezogen werden ließen sich unterschieden in Aufgaben mit sozialen Anforderung und Aufgaben mit rein organisatorischen Anforderungen. Aufgrund dessen wäre es denkbar, dass in zukünftigen Studien überprüft wird, inwiefern die soziale Kompetenz einer Testperson Auswirkungen auf ihr Antwortverhalten zeigt. Es könnte zum Beispiel erwartet werden, dass hierbei Personen des impulsiven Typs zwar die geringste Arbeitseffizienz zeigen, hierbei jedoch die höchste soziale Kompetenz beweisen würden. Bezüglich des flexiblen Typs wäre eine moderate soziale Kompetenz bei gleichzeitiger moderater Arbeitseffizienz zu erwarten. Der reflexive Typ würde sich durch eine sehr hohe Arbeitseffizienz auszeichnen, würde jedoch eine geringe soziale Kompetenz aufweisen.

Aufgrund der vorliegenden Daten dieser Ersterprobung lassen sich die oben geschilderten Schlüsse jedoch leider nicht ziehen und bieten somit mögliche Ansätze für weitere Studien.

12 Zusammenfassung

Das zugrundeliegende Thema der vorliegenden Arbeit war die Entwicklung eines neuen objektiven Persönlichkeitstest zur Erfassung von im Arbeitskontext basierend auf dem Konzept des ILICA.

Hierbei wurde versucht die Programmierung auf Basis der aktuellen technologischen Möglichkeiten realitätsnäher zu gestalten als es im ILICA der Fall war, sowie die Testinhalte auf den Arbeitsalltag anzupassen. Von besonderer Bedeutung war hierbei die Kompensation der Kritikpunkte und Schwächen des ILICA .

Gerade in Bezug auf die Realitätsnähe des Verfahrens wurde versucht den Kritikpunkten des ILICA zu begegnen.

Aufgrund der aktualisierten Definition objektiver Persönlichkeitstest durch Kubinger (2006) wurde das neu entwickelte Verfahren als Verfahren der experimentalpsychologischen Verhaltensdiagnostik definiert, wodurch sich der Name experimentalpsychologisches Verfahren zur Erfassung der Selbstverwaltungsfähigkeit kurz „Everest“ ergab.

Der Everest simuliert den Arbeitsalltag in einem Büro und stellt die Testperson vor organisatorische Aufgaben. Die Testperson kann sich hierbei vollkommen frei innerhalb des simulierten Arbeitsumfeldes bewegen und die Aufgabenstellungen in der von ihr präferierten Art und Weise bewältigen.

Der Umgang mit auftretenden Distraktoren sowie der erreichte Arbeitsscore werden herangezogen um zwischen den zugrundeliegenden Entscheidungstypen zu differenzieren (Der reflexive Entscheidungstyp, der flexible Entscheidungstyp und der impulsive Entscheidungstyp).

Im Zuge der Neukonstruktion des Everest wurden die Möglichkeiten des Computers als Testmediums besser genutzt, als dies ILICA der Fall ist. Die Realitätsnähe wird durch eine interaktive zweidimensionale Umgebung im Vergleich zur textbasierten Umgebung des ILICA erhöht. Zusätzlich erleichtert eine komplette Sprachausgabe der Instruktion und des Textes den Testablauf und verringert den für die Testperson durch umfangreiches Textmaterial entstehenden kognitiven Aufwand. Anzumerken ist, dass der Everest wesentlich weniger Text beinhaltet als der ILICA.

Es wurde darauf Wert gelegt, die gesteigerten Anforderungen aufgrund der Bedienung mittels ausführlicher Trainings- und Instruktionsphasen zu kompensieren.

Insgesamt nahmen 65 Personen an der vorliegenden Studie teil, wobei das Verfahren durchweg auf hohe Akzeptanz gestoßen ist. Die Ergebnisse zeigten, dass der Test durchgängig ernsthaft bearbeitet wurde, zugleich jedoch als nicht belastend empfunden wurde.

Über Exploratorische Faktorenanalysen und Clusteranalysen konnten die drei (theoretisch) angestrebten Entscheidungstypen neuerlich klassifizieren werden und die Dimensionalität der Messung der drei Typen konnte zufriedenstellend geprüft werden. Über Mittelwertsvergleiche konnte gezeigt werden, dass die erzielten Testwerte und das Verhalten während der Testbearbeitung erwartungsgemäß den theoretischen Hintergründen entsprechen. Aufgrund der geringen Stichprobengröße, sollten die hier vorliegenden Ergebnisse nochmals repliziert und abgesichert werden.

Der Everest weist folgende relevante Testkennwerte aus: Reflexiv Score, Impulsiv Score, Flexibel Score, Arbeitsscore, Bearbeitungszeit, Aufrufen von Hilfen.

Die Konstruktion eines experimentalpsychologischen Verfahrens zur Erfassung von Selbstverwaltungsfähigkeit und sozialer Kompetenz im Arbeitskontext kann aufgrund der vorliegenden Ergebnisse als erfolgreich angesehen werden. Weiterführende Untersuchungen hinsichtlich der Gütekriterien müssen noch vorgenommen werden.

Literaturverzeichnis

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

Armor Games Inc. *Homepage*. Abgerufen am 20.09.2014 von <http://armorgames.com/play/12141/kingdom-rush>

Betsch, C. (2004). Präferenz für Intuition und Deliberation (PID). *Zeitschrift für Differentielle und Diagnostische Psychologie*, 25 (4), 179-197

Bortz, J. (1999). *Statistik für Human- und Sozialwissenschaftler*. (5. Aufl.). Heidelberg: Springer Medizin Verlag

Cattell, R. B. (1958). What Is „Objective“ in „Objective Personality Tests“?. *Journal of Counseling Psychology*, 5 (4), 285-289

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

Funke, J., & Reuschenbach, B. (2011). Einsatz technischer Mittel in der psychologischen Diagnostik. In L. F. Hornke, M. Amelang, & M. Kersting (Hrsg.), *Leistungs-, Intelligenz- und Verhaltensdiagnostik* (S.595-667). Göttingen: Hogrefe

Kryspin-Exner, I., Felnhöfer, A., & Kothgassner, O. D. (2011). Die digitale Büchse der Pandora: Psychische Störungen im Cyberspace. *Neuropsychiatrie*, 25 (4), 172-182

Kubinger, K. D. (2006) Ein Update der Definition von Objektiven Persönlichkeitstests. In T. M. Ortner, R. T. Proyer, & K. D. Kubinger (Hrsg.), *Theorie und Praxis Objektiver Persönlichkeitstests* (S. 38-51). Bern: Verlag Hans Huber

Kubinger, K. D. (2009). *Psychologische Diagnostik. Theorie und Praxis psychologischen Diagnostizierens*. (2., überarb. u. erw. Aufl.). Göttingen: Hogrefe

Kubinger, K. D., Rasch, D., Yanagida, T. (2011). *Statistik in der Psychologie*. Göttingen: Hogrefe

Möseneder, D., & Ebenhöf, J. (1996). *ILICA – Ein Simulationstest zur Erfassung des Entscheidungs-verhaltens*. Test: Software und Manual. Frankfurt/M.: Swets Test Services

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

Payne, J. W., Bettman, J.R., & Johnson, E.J. (1993). *The adaptive decision maker*. Cambridge: University Press

Simon, H. A. (1997). *Models Of Bounded Rationality. Emperically Grounded Economic Reason* (3. Aufl.) MIT Press

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, 119, 3-22

Smith, E.R., & DeCoster, J. (1999) Associative and Rule-Based Processing – A Connectionist Interpretation of Dual-Process Models. In S. Chaiken, & Y. Trope (Hrsg.), *Dual-Process Theories in Social Psychology* (S. 323-324). New York: Guilford Press

Spranger, E. (1950). *Lebensformen. Geisteswissenschaftliche Psychologie und Ethik der Persönlichkeit* (8. Aufl.) Tübingen: Neomarius Verlag

Strack, F., & Deutsch, R. (2004). Reflective and Impulsive Determinants of Social Behavior. *Personality and Social Psychology Review*, 8 (3), 220-247

Von Neumann, J., & Morgenstern, O. (1953). *Theory Of Games And Economic Behavior*. Abgerufen von <http://ia600301.us.archive.org/29/items/theoryofgamesand030098mbp/theoryofgamesand030098mbp.pdf>

Zimbardo, P.G., & Gerrig, R.J. (1999). *Psychologie* (7., neu übers. u. bearb. Aufl.).
Berlin: Springer

Anhang

Anhang A - Abstract

In der vorliegenden Arbeit wird die Entwicklung und Ersterprobung eines neuen Objektiven Persönlichkeitstests (experimentalpsychologische Verhaltensdiagnostik) zur Erfassung der Selbstverwaltungsfähigkeit, kurz „Everest“ basierend auf der Idee des ILICA (Möseneder & Ebenhöf, 1996) vorgestellt. Das Verfahren wurde 65 Personen im Alter von 19 bis 65 Jahren vorgegeben ($M = 35.48$, $SD = 13.16$). Es gelang mittels Clusteranalyse drei verschiedene Entscheidungstypen zu extrapolieren welche sich hinsichtlich der Ausprägung ihrer Testkennwerte voneinander unterschieden. Analysen mittels Exploratorischer Faktorenanalysen ergaben zufriedenstellende Ergebnisse bezüglich der Dimensionalität der Items (3-faktorielle Lösung), mit welchen diese drei Typen erfasst werden. Mittelwertsvergleiche zeigen außerdem, dass die drei gefunden Entscheidungstypen reflexiv, flexibel und impulsiv sich zufriedenstellend mit der zugrundeliegenden Theorie begründen lassen.

Die innere Konsistenz der relevanten Skalen konnte nach Eliminierung eines nicht Trennscharfen Items als zufriedenstellend eingestuft werden. Insgesamt kann die Konstruktion und Ersterprobung des Everest als erfolgreich angesehen werden.

Versuchspersonen-Information und Einverständniserklärung

Wir laden Sie ein, freiwillig als Versuchsperson, an folgender Testung teilzunehmen:

Skalierung und Validierung des Everest

Bei diesem Test handelt es sich um ein in der Entwicklung befindliches Verfahren, dessen Ergebnisse noch nicht verlässlich für den Einzelfall verwendet werden können. Ziel dieser Testung ist es Daten zu erheben anhand derer zukünftige Testergebnisse in Relation gesetzt werden können.

Ihre Rechte:

Selbstverständlich können Sie vor und jederzeit während der Studie weitere Informationen über Zweck, Ablauf, etc. der Studie von den studiendurchführenden Personen erfragen. Gerne werden Sie auch nach Ende der Studie über die Ergebnisse der Untersuchung informiert. Sie können die Untersuchung jederzeit, auch ohne Angabe von Gründen, von sich aus abbrechen.

Datenschutz:

Sämtliche Ihre Person betreffenden Daten werden getrennt von den erhobenen Daten aufbewahrt, sodass Ihre Anonymität gewahrt bleibt. Es ist geplant, die im Rahmen der Untersuchung erhobenen Daten in einer wissenschaftlichen Zeitschrift zu veröffentlichen.

Einverständniserklärung:

Durch Ihre Unterschrift bestätigen Sie, dass Sie die Versuchspersoneninformation gelesen und verstanden haben. Sie erklären sich mit der Teilnahme an der Testung zur Skalierung und Validierung des Everest, sowie der Analyse Ihrer Daten durch befugte Personen einverstanden.

Name: _____

Geboren am: _____ in: _____

Datum: _____ Unterschrift: _____

Anhang B - Einverständniserklärung

Persönliche Daten

Name: Griepenstroh
Vorname: Tobias
Adresse: Engerthstraße 193/3/38, 1020 Wien
Telefon: 0681 10857016
E-Mail: tobias.griepenstroh@web.de
Nationalität: Deutsch

Praktika und berufliche Erfahrungen

Seit 04/2012 Verein BALANCE
Besuchsdienst

06/2013 – 08/2013 ORS Service GmbH
Praktikum als Psychologischer Betreuer

Ausbildung

Seit 10/2007 Studium der Psychologie an der Universität Wien

1997 – 06/2006 Gymnasium der Stadt Rahden

Wehrdienst/Zivildienst

11/2006 – 06/2007 Zivildienst im Jugendwartbezirk Stemwede

10/2006 – 11/2006 Wehrdienst in Rotenburg/ Wümme

