



universität
wien

DIPLOMARBEIT

Titel der Diplomarbeit

Modernisierung Objektiver Persönlichkeitstests:
Nutzen von virtuellen Umgebungen zur Erfassung des
Big Five Konstrukts Gewissenhaftigkeit
- eine Machbarkeitsstudie

Verfasser

Michael Mühlegger

Angestrebter akademischer Grad

Magister der Naturwissenschaften (Mag. rer. nat.)

Wien, Mai 2015

Studienkennzahl lt. Studienblatt: A 298

Studienrichtung lt. Studienblatt: Psychologie

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

I hear you say "Why?"

Always "Why?"

You see things; and you say "Why?"

But I dream things that never were; and I say "Why not?"

(George Bernhard Shaw)

Danksagung

Danken möchte ich in erster Linie meiner Betreuerin, Frau Mag. Dr. Khorramdel-Ameri, dass sie mir die Chance gegeben hat, eine Thema zu bearbeiten dass mich leidenschaftlich interessiert und sie mir zudem alle nötigen Freiheiten gab, um meine Ideen in die Tat umzusetzen. Des Weiteren möchte ich mich für ihre konstruktive Kritik und Geduld bedanken. Vielen Dank für die Zeit und Mühen, die Sie in meine Arbeit investiert haben. Auch möchte ich mich herzlich bei den Mitarbeitern der Test- und Beratungsstelle der Fakultät für Psychologie an der Universität Wien für ihre tatkräftige Mithilfe bei den Erhebungen bedanken. Ebenso möchte ich der Firma Schuhfried meinen Dank aussprechen, für die freundliche Bereitstellung zweier Wiener Testsysteme zum Zweck der Datenerhebung.

Besonders möchte ich mich bei meiner Mutter Heidi Mühlegger bedanken, da sie mich Zeit meines Lebens mit all ihrer Kraft unterstützt hat und mir stets ein Vorbild in Zuversicht und Beharrlichkeit war und noch immer ist. Auch möchte ich mich bei Frau Dr. Elisabeth Staus-Rausch und Frau Dr. Eva-Maria Fuchs-Rausch und ihren Lieben bedanken, die mir über die letzten Jahre meines Studium hinweg zur Familie geworden sind.

Zum Abschluss möchte ich noch einem besonderen Menschen meinen Dank aussprechen, meiner Partnerin Victoria Staus. Sie ist nicht nur Namenspatin des in dieser Arbeit entstandenen Verfahrens, sondern musste auch während seiner Entstehungsphase diesen Test als Qualitätskontrolleurin so oft wie kein anderer durchführen. Danke für die Kraft und Motivation die Du mir seit dem wir uns kennen zukommen hast lassen.

Inhaltsverzeichnis

I	Einleitung	12
II	Theoretischer Teil	14
1	Das Big Five Modell der Persönlichkeit	15
1.1	Das Konstrukt "Gewissenhaftigkeit"	16
1.2	Facetten der Gewissenhaftigkeit	19
1.3	Schwierigkeiten bei der Erfassung von Gewissenhaftigkeit	21
1.3.1	Vorteile der Fragebogenmethode	21
1.3.2	Grenzen der Fragebogenmethode	22
2	Experimentalpsychologische Verhaltens- beobachtung	27
2.1	Objektive Persönlichkeitstests	27
2.1.1	Objektivität bei Objektiven Persönlichkeitstests	29
2.1.2	Kritik an experimentalpsychologischen Verhaltensdiagnostik	31
3	Computersimulationen	33
3.1	Virtuelle Realität	36
3.1.1	Telepresence und Immersion	36
3.1.2	Anwendung virtueller Umgebung in der Psychologie . . .	37
III	Empirischer Teil	41
4	Zielsetzung der empirischen Untersuchung	41

5	Testbeschreibung und -Konstruktion (VICI)	42
5.1	Entstehungsgeschichte	42
5.2	Phase 1 (Einleitung)	43
5.2.1	Datenblatt	43
5.2.2	Szenario	44
5.3	Screening Auge-Hand Koordination (AHKS)	45
5.4	Test VICI - Phase 2 (Training)	47
5.4.1	Trainingsphase	47
5.4.2	Visuelle Orientierung	47
5.4.3	Navigation im Raum	49
5.4.4	Interaktion mit Objekten	50
5.4.5	Arbeit am Fließband	52
5.4.6	Abholbereich	56
5.4.7	Das Büro	57
5.5	Test VICI - Phase 3 (Testphase)	57
5.5.1	Aufgabe 1	57
5.5.2	Aufgabe 2	60
5.5.3	Aufgabe 3	62
6	Testkennwerte (VICI)	64
6.1	AHK-Screening (AHKS)	64
6.2	Kennwerte der Aufgabe 1 (Aufräumen)	64
6.3	Kennwerte der Aufgabe 2 (Fließband)	65
6.4	Kennwerte der Aufgabe 3 (Bestellung)	68
7	Weitere verwendete Untersuchungsinstrumente	69
7.1	NEO- Persönlichkeitsinventar (NEO PI-R)	69
7.2	Räumliches Orientierungsvermögen (3D)	70

7.3 Arbeitshaltungen (AHA)	72
8 Fragestellungen	76
8.1 Hypothesen	76
9 Statistische Auswertungsmethoden	79
10 Beschreibung der Stichprobe	80
10.1 Beschreibung der Stichprobe bezüglich Geschlechts- & Alters- verteilung	80
10.2 Beschreibung der Stichprobe bezüglich des Bildungsgrads	83
11 Ausgewählte Ergebnisse des Tests (VICI)	85
11.1 Ergebnisse Auge-Hand Koordination (AHKS)	85
11.2 Ergebnisse Aufgabe 1	86
11.2.1 Ordnungssinn (Abfall)	86
11.2.2 Ordnungssinn (Kartons)	87
11.3 Ergebnisse Aufgabe 2	87
11.3.1 Quantitative Arbeitsleistung (Produktionsleistung)	87
11.3.2 Entscheidungsfreudigkeit	88
11.3.3 Qualitative Arbeitsleistung (Exaktheit)	89
11.3.4 Arbeitsstil (Impulsivität / Reflexivität)	90
11.4 Ergebnisse Aufgabe 3	91
11.4.1 Aufgabenerfüllung	91
12 Hypothesenprüfung	92
12.1 Prüfung der Fragestellung 1	92
12.2 Prüfung der Fragestellung 2	95
12.3 Prüfung der Fragestellung 3	98

12.4 Prüfung der Fragestellung 4	101
13 Konstruktvalidierung	104
13.1 Prüfung der Fragestellung 5	104
13.2 Prüfung der Fragestellung 6	104
14 Faktorenanalyse	106
15 Zusammenfassung der Ergebnisse	110
16 Interpretation und Diskussion der Ergebnisse	113
IV Anhang	125
Literatur	136

Aus Gründen der besseren Lesbarkeit wird auf die gleichzeitige Verwendung männlicher und weiblicher Sprachformen verzichtet. Sämtliche Personenbezeichnungen gelten gleichwohl für beiderlei Geschlecht.

Abstract - Deutsch

Die vorliegende Arbeit möchte einen Beitrag dazu leisten Objektive Persönlichkeitstests (experimental- psychologische Verhaltensbeobachtungen) moderner und realitätsnäher (und dadurch vielleicht für die Testperson motivierender) zu gestalten, indem mit dreidimensionalen virtuellen Umgebungen gearbeitet wird, was heutzutage technisch leichter möglich ist. Dies bietet mehr Möglichkeiten, aber auch neue Schwierigkeiten, vor allem für Testpersonen die vielleicht weniger geübt im Umgang mit dem Computer sind. Die vorliegende Arbeit ist daher als eine Machbarkeitsstudie zu verstehen, die testen und evaluieren möchte, inwiefern neueste Technik für den Einsatz experimental- psychologischer Verhaltensbeobachtung genutzt werden kann. Im Fokus liegt dabei das Konstrukt Gewissenhaftigkeit. Es wurde eigens ein Instrument (VICI) für die vorliegende Studie entwickelt und an insgesamt 85 Personen erprobt, die größtenteils Studierende der Psychologie waren und über die Testerfahrung der Fakultät für Psychologie an der Uni Wien und im privaten Umfeld des Autors rekrutiert wurden. Mittels t-Tests wurde geprüft, ob das Geschlecht, räumliche Orientierungsfähigkeit, Augen-Hand Koordination oder 3D-Computerspiel-Erfahrung die Ergebnisse des Tests VICI beeinflussen. Zur Überprüfung der Dimensionalität wurde eine Hauptkomponentenanalyse mit schiefwinkliger Oblimin Rotation mit Kaiser-Normalisierung unter Verwendung von 16 Variablen des Test VICI durchgeführt. Die Konstruktvalidierung erfolgte anhand ausgewählter Untertests des objektiven Persönlichkeitsverfahrens Arbeitshaltungen (AHA) und des Persönlichkeitsfragebogens NEO PI-R. Die Ergebnisse zeigen, dass weder Geschlecht, räumliche Orientierungsfähigkeit, Augen-Hand Koordination oder 3D-Computerspiel-Erfahrung einen signifikanten Einfluss auf die Ergebnisse des Tests VICI haben. Die Hauptkomponentenanalyse identifizierte fünf Faktoren, die jeweils einer der

Aufgaben des Test VICI zugeordnet werden konnten. Die korrelative Überprüfung im Rahmen der Konstruktvalidierung ergab, dass es keine signifikanten Zusammenhänge mit NEO PI-R und dem AHA gibt. Diese erste Machbarkeitsstudien zeigt vielversprechende Ergebnisse für den Einsatz virtueller Umgebungen für experimentalpsychologische Verhaltensbeobachtung, wobei zukünftige Studien die Konstruktvalidierung weiter untersuchen sollten.

Abstract - English

The present study aims to modernize objective personality tests to make them more modern and more realistic (and thus more motivating to the subjects) by working with three-dimensional virtual environments which is nowadays technically easier. This provides more opportunities but also new difficulties, particularly for subjects with lower computer proficiency. The present study is a feasibility study, which aims to test and evaluate the usefulness of the latest technology for the development and construction of modern objective personality tests. The study thereby focusses on the Big Five construct conscientiousness. An instrument (VICI) was specifically developed for this study and tested on a total of 85 persons, mostly psychology students of the Faculty of Psychology at the University of Vienna. Using t-tests, it was examined whether gender, spatial orientation ability, eye-hand coordination, or 3D computer game experience influences the results of the test VICI. To evaluate the dimensionality of the instrument an exploratory factor analysis with oblique oblimin rotation with Kaiser normalization was performed on 16 variables of the test VICI. The construct validity was conducted based on chosen subtests of objective personality test "Arbeitshaltungen" (AHA) and the personality questionnaire NEO PI-R. It was found that gender, spatial orientation ability, eye-hand coordination and 3D computer game experience have no significant impact on the results of the test VICI. The explanatory factor analysis identified five factors that could each be assigned to one of the tasks of the test VICI. With regard to the construct validation results showed that there are no significant correlations with NEO PI-R and AHA. The current feasibility study shows promising results for the use of virtual environments for experimental psychological behavior observation, but future studies should further examine the construct validity.

Teil I

Einleitung

Der Bereich der Psychologie welcher sich mit den individuellen Differenzen zwischen Menschen befasst, wird als Differentielle Psychologie bezeichnet. Um individuelle Differenzen überhaupt feststellen zu können, bedarf es zunächst geeigneter diagnostischer Instrumente. Die Aufgabe der psychologischen Diagnostik ist es diese Instrumente theoretisch fundiert und evaluiert zu konstruieren. Grob können diese Verfahren in die Bereiche Leistungs- und Persönlichkeitsdiagnostik unterteilt werden. (Kubinger, 2009)

Die Anwendungsgebiete der entwickelten Verfahren ist vielfältig und erstreckt sich von der Grundlagenforschung über Beratungssituationen bis hin zur Personalselektion, um nur einige Beispiele zu nennen. Besonders in der Personalselektion kann die Anwendung psychologisch- diagnostischer Verfahren einen erheblichen Nutzen bieten. Bei der Auswahl eines ungeeigneten Bewerbers können für ein Unternehmen, höhere Kosten sowie Verluste an Zeit und Ressourcen entstehen, die es möglichst zu vermeiden gilt (Ruf, 2010).

Daher finden immer mehr psychologisch- diagnostische Instrumente eine Anwendung in der Selektionsdiagnostik. Aus der Sicht eines Bewerbers wird, verständlicherweise, ein möglichst gutes Abschneiden innerhalb des Auswahlprozesses angestrebt. Bei der Erhebung von Konstrukten des Leistungsbereichs, beispielsweise bei der Fähigkeit zum logisch schlussfolgernden Denken, ist es der Testperson nur schwer, sich besser darzustellen, als es dem eigenen Fähigkeitsniveau entspricht. Anders stellt sich dies bei der Persönlichkeitsdiagnostik

dar, hier ist es auf Grund der häufig verwendeten Persönlichkeits-Fragebögen leicht ersichtlich, was das Ziel der Befragung ist und welche Antworten als sozial Erwünscht gelten und damit positiv gewertet werden. Die Möglichkeit einer bewussten Verzerrung der Ergebnisse bei Persönlichkeitsfragebögen stellt besonders innerhalb einer Beratungs- oder Auswahl-situation ein Problem dar. (Kubinger, 2009)

Direkte, im Sinne einer systematischen Verhaltensbeobachtung ist zeit- und kostenintensiv und daher bei einer größeren Stichprobe nicht ohne Weiteres durchführbar. Außerdem ist die Objektivität seitens des Beobachters nicht immer gegeben oder nur durch zeitintensive Schulungen zu erreichen. Ein Versuch der erhöhten Verfälschbarkeit von Fragebögen entgegen zu wirken, stellt die Entwicklung Objektiver Persönlichkeitstests dar, welche versuchen Persönlichkeitsmerkmale anhand von beobachtbarem Verhalten zu erschließen. Publierte Verfahren sind ähnlich einem Computerspiel aufgebaut und werden in einem zweidimensionalen Raum dargestellt. Die Möglichkeit die Ergebnisse nach Belieben zu verzerren wird durch eine Verschleierung der Testintention erreicht. Wenn der Testperson nicht klar ist, was gemessen wird, oder wie es verrechnet wird, ist eine Verzerrung zwar immer noch nicht auszuschließen, jedoch erheblich erschwert (L. Schmidt, 2006). Leider hat dieser Ansatz nur zu einer geringen Anzahl, auf diese Weise entwickelter Verfahren geführt. Zum einen ist die Entwicklung solcher Objektiven Persönlichkeitsverfahren bedeutend aufwendiger, verglichen mit herkömmlichen Persönlichkeitsfragebögen, und darüber hinaus mangelt es noch an einer ausreichenden Zahl überzeugender Validierungsstudien zu dieser Verfahrensklasse (Amelang, Bartussek, Stemmler & Hagemann, 2011). Die überschaubare Zahl an entwickelten Objektiven Persönlichkeitsverfahren nutzt den Computer zur Vorgabe und Auswertung, allerdings

ist der Großteil dieser Verfahren veraltet und nutzen die über die letzten Jahrzehnte gewachsenen Möglichkeiten dieses Mediums kaum aus.

Mit dem Einbezug des computergestützten Testens konnte eine Vielzahl möglicher Fehlerquellen von traditionellen Fragebögen, wie beispielsweise Testleiter-Effekten oder Auswertefehlern gemindert werden, beziehungsweise das Testgütekriterium Objektivität verbessert werden. Doch obwohl die Möglichkeiten bei Weitem noch nicht ausgeschöpft sind und die technische Entwicklung immer rasanter voranschreitet, hat sich am klassischen Stimulusmaterial der meisten Verfahren kaum etwas verändert. Nach wie vor besteht der Großteil aller Verfahren aus einer passiven Vorgabe von Fragen im Textformat, einzelnen Bildern oder Videos, auf die seitens der Testperson geantwortet werden soll. Auch die zuvor genannten objektiven Persönlichkeitsverfahren arbeiten bisher fast ausschließlich mit zweidimensionalen Stimulusmaterial.

Ziel dieser Studie ist es, die Machbarkeit und generelle Anwendbarkeit eines neuen Ansatzes für die Vorgabe von Stimulusmaterial zu erproben. Beeinflusst von klassischen objektiven Persönlichkeitsverfahren, wird in dieser Arbeit versucht, Persönlichkeitsmerkmale anhand von beobachtbarem Verhalten innerhalb einer Computersimulation zu erfassen. Dies soll am Beispiel eines eher verzerrungssensiblen Konstrukts wie jenem der „Gewissenhaftigkeit“ aus den Big-Five (vgl. Abschnitt 1.3.2) veranschaulicht werden. Die interaktive Gestaltung, sowie die aktive Bewegung einer Testperson im dreidimensionalen Raum, liefert hierbei eine zusätzliche Informationsquelle und bietet zugleich mehr Möglichkeiten die Testsituation einer realen Situation anzunähern.

Teil II

Theoretischer Teil

1 Das Big Five Modell der Persönlichkeit

Wenn ein so komplexes Konstrukt, wie die Persönlichkeit eines Menschen beschrieben werden soll, so verwendet man hierzu meist Eigenschaftswörter, welche dazu dienen, die unterschiedlichen Ausprägungen der Persönlichkeit genauer zu charakterisieren. Eigenschaftswörter wie „pflichtbewusst“, oder „genau“ dienen nicht nur der reinen Beschreibung, sie können auch dazu verwendet werden, um Unterschiede zwischen einzelnen Personen festzustellen. Die Annahme, dass sich alle Aspekte individueller Differenzen, welche bedeutsam, interessant oder nützlich sind oder waren, Eingang in die Sprache gefunden haben, begründet in der Folge die sogenannte *Sedimentationshypothese* (Kubinger, 2009). Aufbauend auf dieser Hypothese und der Vielzahl an unterschiedlichen Eigenschaftswörtern, entstand der psycholexikalische Ansatz, welcher sich mit der Ähnlichkeit bzw. Unähnlichkeit solcher Begriffe befasst (Amelang, 2003). Mittels Faktorenanalyse konnten fünf stabile Dimensionen gefunden werden, welche heute das Fünf-Faktoren-Modell der Persönlichkeit (FFM) begründen (Kubinger, 2009). Zu diesen, heute gemeinhin als „Big Five“ bezeichneten Dimensionen, zählen Costa und McCrae (1992) zufolge *Neurotizismus*, *Extraversion*, *Offenheit für Erfahrung*, *Verträglichkeit* und *Gewissenhaftigkeit*. Die auf dem Big Five-Persönlichkeitsmodell hervorgegangenen diagnostischen Verfahren, wie zum Beispiel das NEO-PI-R (Costa & McCrae, 1992) bilden derzeit die Basis psychologischen Diagnostizierens im Persönlichkeitsbereich und finden breite Anwendung (Kubinger, 2009). Man könnte sogar sagen, dass dieses

Modell innerhalb der Persönlichkeitsforschung eine Hegemonialstellung eingenommen hat (Schuler & Höft, 2001).

1.1 Das Konstrukt “Gewissenhaftigkeit”

Diese Studie konzentriert sich hauptsächlich auf das Konstrukt Gewissenhaftigkeit und seine Subdimensionen beziehungsweise Facetten. Dem deutschen Philosophen Wilhelm T. Krug (1819) nach entspringt die Tugend der gewissenhaften Pflichterfüllung. Also wird Gewissenhaftigkeit nicht nur als Grundlage für Tugendhaftigkeit angesehen, sondern wird immer auch mit einer Handlung und der Art und Weise der Ausführung in Verbindung gesetzt. Borkenau und Ostendorf (2008) zufolge unterscheidet diese Persönlichkeitsdimension ordentliche, zuverlässige, anspruchsvolle, pünktliche, disziplinierte sowie systematisch vorgehende Menschen von nachlässigen, gleichgültigen, undisziplinierten und inkompetenten Personen. Andere Persönlichkeitstheorien zählen zusätzlich zur Planungsfähigkeit, Strukturiertheit und Zielstrebigkeit einer Person noch weitere Konzepte zum Konstrukt der Gewissenhaftigkeit, wie beispielsweise impulsbeziehungsweise Selbstkontrolle, oder die Fähigkeit, Belohnungen aufschieben zu können (vgl. Corker, Oswald & Donnellan, 2012; DeYoung et al., 2010; Lievens & Ones, 2009).

Je nach Stärke der Zielsetzung und dem Antrieb zur Zielerreichung unterscheiden sich Personen in der Art und Weise, wie sie eine Aufgabenstellung bearbeiten und welches Maß an Energie sie dafür bereit sind aufzubringen (Lord, 2011). Gewissenhaftigkeit sei also jenes Merkmal, welches die Differenz zwischen dem Leistungspotential und der realisierten Leistung einer Person beschreibt und somit auch die Wahrscheinlichkeit angibt, die Ergebnisse zu erreichen, welche zuvor als Ziel gesteckt wurden (Lord, 2011). Menschen mit hoher

Ausprägung in Gewissenhaftigkeit „... besitzen einen Drang zur Vollendung von Aufgaben, obwohl sie sich möglicherweise zu lange mit Projekten aufhalten“ (Lord, 2011, S.25).

Gewissenhaftigkeit und Eignungsdiagnostik

Besonders in der Eignungs- und Selektionsdiagnostik ist eine positive Ausprägung in der Skala Gewissenhaft nicht nur wünschenswert, sondern oft eine Grundvoraussetzung für einen erfolgreichen Ausgang einer Bewerbung. Wenn die Wahl zwischen zwei Bewerbern getroffen werden soll, würde kaum ein Arbeitgeber denjenigen Bewerber bevorzugen, welche sich als unpünktlich oder ungenau beschreiben würde. Bezogen auf zukünftiges Arbeitsverhalten ist die Gewissenhaftigkeit der wohl am meisten beforschte Persönlichkeitsfaktor und auch jener mit der stärksten Vorhersagekraft (vgl. Barrick, Mount & Judge, 2001; Minbashian, Wood & Beckmann, 2010; Sutherland, De Bruin & Crous, 2007).

Gewissenhaftigkeit und Gesundheit

Doch der Faktor Gewissenhaftigkeit beschränkt sich nicht allein auf das Verhalten im beruflichen Kontext. In einer Längsschnittstudie von Terracciano, Löckenhoff, Zonderman, Ferrucci und Costa (2008) untersuchten die Autoren eine repräsentative US-Bevölkerungstichprobe (N = 2.359), über mehrere Jahre hinweg und erhoben wiederholt neben dem Gesundheitszustand auch mehrere Persönlichkeitsmaße. Die Autoren kamen zu dem Ergebnis, dass eine hoch ausgeprägte Gewissenhaftigkeit mit einer geringeren Sterblichkeit einhergeht. Diese Ergebnisse führten sie darauf zurück, dass gewissenhafte Personen eher einen gesundheitsförderlichen Lebensstil (z.B. gesunde Ernährung, Sport) hätten und auf gesundheitsschädliche Verhaltensweisen wie z.B. Rauchen eher

verzichten würden. Zudem wären diese Personen auch konsequenter bei gesundheitsförderliche Praktiken (z.B. regelmäßige Medikamenteneinnahme). Gewissenhaftigkeit ist als ein Teil der Big Five ein Persönlichkeitsfaktor mit Auswirkungen, die über den beruflichen Kontext hinausgehen und sogar über Gesundheit und Lebenserwartung entscheiden können (Roberts, Walton & Bogg, 2005).

1.2 Facetten der Gewissenhaftigkeit

Nahezu für jedes Verfahren, welches Gewissenhaftigkeit misst, gibt es eine eigene Konzeption darüber, was die Anzahl, Benennung oder Art der Interpretation der Facetten dieses Konstrukts angeht. Zu den wichtigsten gehören NEO PI-R (Ostendorf & Angleitner, 2004), NEO-FFI (Borkenau & Ostendorf, 2008), BIP (Hossiep & Paschen, 2003), FPI (Fahrenberg, Hampel & Selg, 2010) oder PRF (Jackson, 1974) um nur einige weitverbreitete Verfahren zu nennen. Die Vielzahl an Verfahren und deren unterschiedliche Definition des selben Konstrukts macht es schwierig eine allgemeingültige Aussage darüber zu treffen, was nun eigentlich als Gewissenhaft zu verstehen ist. In einem regelbasierten Ansatz stellten Westhoff, Steinborn und Schurz (2013) aus allen aktuellen Fragebögen ein Kategoriensystem zur Operationalisierung und besseren Interpretierbarkeit des Konstrukts Gewissenhaftigkeit her. Die Autoren bedienten sich der Strategie des Vergleichens nach (Klauer, 2011 zitiert nach Westhoff et al., 2013) und erstellten einen Datenpool bestehend aus all jenen Facetten, welche entweder in allen, oder zumindest in den meisten Konzeptionen enthalten waren. Nachdem übereinstimmende Faktoren unter einer gemeinsamen Bezeichnung zusammengefasst wurden und redundante Begriffe entfernt wurden, blieben zehn unterschiedliche Facetten übrig. Im Anschluss wurde versucht diese entsprechend einer positiven Verhaltensbeschreibung zu operationalisieren. In der nachfolgenden Tabelle werden die einzelnen, durch Westhoff et al. (2013) zusammengefügte Facetten, in einem Kategoriensystem samt jeweiliger Operationalisierung veranschaulicht.

Tabelle 1: Das Kategoriensystem der Gewissenhaftigkeit nach Westhoff (2013),
Tabelle 1 (S. 206)

Facette	Operationalisierung
Beharrlichkeit	Hält konsequent und über eine lange Zeit an der korrekten Erfüllung einer Aufgabe fest. Investiert so viel Zeit in die Arbeitsaufgabe, wie nötig ist, um das gewünschte Ergebnis zu erzielen. Neigt auch bei auftauchenden Schwierigkeiten nicht zum Aufgeben, sondern ist in der Regel motiviert, Lösungen zu finden. Verfolgt diese Lösungen geduldig und unermüdlich, auch wenn sie eine Verlängerung der Arbeitszeit an der Aufgabe und eine Aufwandserhöhung mit sich bringen. Interessiert sich für Aufgaben, die viel Ausdauer erfordern.
Besonnene Arbeitsweise	Überlegt vor Beginn einer Arbeit gründlich. Entwickelt einen Arbeitsplan, der alle Teilschritte in der richtigen Reihenfolge enthält. Verschafft sich Klarheit über die benötigten Mittel (zum Beispiel Informationen, Finanzmittel). Bedenkt alle Umstände, die für die Konsequenzen (Gelingen oder Misslingen) bedeutsam werden könnten. Schätzt auch den zeitlichen Aufwand für Gesamtaufgabe sowie für die Teilaufgaben ab. Arbeitet nach dem Arbeitsplan und passt ihn an veränderte Gegebenheiten an.
Impulskontrolle	Neigt nicht dazu, Gefühlen und Wünschen spontan nachzugeben. Kann widerstehen, etwas anderes zu tun als das eigentlich Begonnene. Hat also die Fähigkeit, Dinge auch tatsächlich zu lassen, die nicht beabsichtigt waren.
Leistungsbereitschaft	Interessiert sich für berufliche Tätigkeiten, die hohe Leistungen erfordern. Stellt sich herausfordernden beruflichen Anforderungen und möchte diesen gerecht werden.
Ordentlichkeit	Organisiert den eigenen Arbeitsplatz nach einer nachvollziehbaren Ordnung. Alles hat seinen Platz und befindet sich in der Regel an diesem. Verwaltet Arbeitsmaterialien so, dass sie auf Anhieb zu finden sind, auch von eingewiesenen Personen.
Pflichtbewusstsein	Kennt die mit Arbeitsaufgaben und dem Umgang mit anderen verbundenen Pflichten und erfüllt diese. Kommt diesen Pflichten auch in schwierigen Situationen nach. Erledigt in der Regel zuerst alle wichtigen, unbedingt zu erledigenden Aufgaben und wendet sich erst danach anderem zu.
Regelbewusstsein	Hält sich auch gegen innere und äußere Widerstände an geschriebene und ungeschriebene Regeln des beruflichen Alltages (zum Beispiel im Umgang mit Mitarbeitern, Kollegen, Vorgesetzten und Kunden).
Selbstdisziplin	Kann sich selbst zum Arbeiten motivieren, unabhängig von der Lust zum Arbeiten. Erledigt diejenigen Aufgaben, die auch erledigt werden sollten. Führt begonnene Aufgaben trotz störender Einflüsse (zum Beispiel Langeweile, Ablenkungen) zu Ende.
Verantwortungsbewusstsein	Arbeitet zuverlässig auch ohne Kontrolle von außen. Übernimmt die Verantwortung für eigene Fehler und versucht nicht, sie auf andere abzuschieben.
Zuverlässigkeit	Führt Aufgaben pünktlich und in der erforderlichen Qualität aus.

1.3 Schwierigkeiten bei der Erfassung von Gewissenhaftigkeit

Wie im vorangegangenen Abschnitt beschrieben, beziehen sich der Großteil aller Facetten der Gewissenhaftigkeit auf das Verhalten und die Art und Weise, wie eine Aufgabe bearbeitet wird. Die Beobachtung von realem Verhalten ist jedoch sehr aufwendig und in vielen Situationen auch nur schwer möglich. Die klassische Methode Gewissenhaftigkeit zu erfassen stellt daher das Fragen dar. Es kann hier zwischen zwei grundsätzlichen Methoden unterschieden werden. Zum einen können Fragen direkt, also mündlich während eines psychologisch, diagnostischen Gesprächs gestellt werden, oder in schriftlicher Form als Fragebogen vorgegeben werden. Bei der Fragebogenmethode kann generell zwischen Selbst- und Fremdeinschätzung unterschieden werden (Becker, 2003). Hierbei werden einer Testperson schriftlich vorformulierte Aussagen, Fragen, oder Eigenschaftswörter vorgegeben, zu welchen diese im vorgegebenen Antwortformat ein Urteil abgibt. Nach Becker (2003) kann ein Persönlichkeitsfragebogen definiert werden als „ ... standardisiertes Instrument zur Erhebung von Selbst- oder Fremdberichtsdaten, aus denen unter Anwendung testtheoretisch begründeter Auswertungsprinzipien, Testwerte abgeleitet werden, die als Indikatoren für den individuellen Ausprägungsgrad von Persönlichkeitseigenschaften dienen“ (S.332).

1.3.1 Vorteile der Fragebogenmethode

Die klassische Fragebogenmethode zu Konstrukten der Persönlichkeit hat zahlreiche Vorteile, wobei hier nur einige genannt werden sollen. Zum einen ist es sicher der direkteste Weg Informationen über Empfindungen, Emotionen, Erlebnisse und Motive einer Person zu gewinnen, diese danach zu fragen. Zum

anderen liegen die Vorteile des direkten Zugangs zu Merkmalen der Persönlichkeit, in einer hohen Objektivität in Durchführung, Auswertung und Interpretation. Ebenso ist es mit dieser Methode möglich, auf äußerst ökonomischem Weg sehr viele Informationen in kurzer Zeit zu erhalten. Doch so stark die genannten Vorteile auch sind, so hat die Fragebogentechnik unter speziellen Bedingungen einige Einschränkungen.

1.3.2 Grenzen der Fragebogenmethode

Jede Person ist in gewisser Weise der einzige wahre Experte, wenn es darum geht, Fragen zu beantworten, die sich auf das eigene Erleben beziehen. Bei Fragen mit geringer Komplexität, welche sich auf Fakten beziehen, zum Beispiel das Alter, fällt die Beantwortung leicht und es kann davon ausgegangen werden, dass es bei der Beantwortung der Frage nur einen geringen Spielraum gibt. Bezieht sich die Fragestellung auf subjektive Inhalte, wie beispielsweise den momentanen Grad des Wohlbefindens, steigt die Komplexität der Aufgabe und die Antwort muss erst in Beziehung zu Momenten mit gleichen oder ähnlich erlebten Empfindungen des Wohlbefindens gesetzt werden. Dies erfordert nicht nur eine gute Beobachtungsgabe bezogen auf innere Vorgänge der eigenen Person, sondern auch die Fähigkeit frühere Empfindungen korrekt zu erinnern und mit der aktuellen Situation zu vergleichen. Untersuchungen zum autobiographischen Gedächtnis haben gezeigt, dass es sich beim Erinnern nicht lediglich um die Aktivierung von Gespeichertem handelt, sondern eher um eine Rekonstruktion vergangener Ereignisse mit heuristischen Strategien (Amelang, Bartussek, Stemmler & Hagemann, 2006). Sobald also die

benötigte Information mittels Nachdenken gefunden werden muss, ist eine Verzerrung des Ergebnisses durch fehlerhafte Erinnerung nicht auszuschließen. Beispielsweise werden bei der so genannten Verfügbarkeitsheuristik eher jene Ereignisse als wahrscheinlicher eingeschätzt, welche aktuell einfacher aus dem Gedächtnis abrufbar sind (Gordon, 1996 zitiert nach Kirchler, 2008). Hat die befragte Person also erst kürzlich eine ähnliche Situation erlebt, wird die eigene Selbsteinschätzung durch dieses noch frische Erlebnis beeinflusst. Nach Jungermann, Pfister und Fischer (1998) sind gerade jene Informationen, welche erst kürzlich auftraten lebhafter, bildhafter und somit auch leichter für Erinnerungen abrufbar. Ebenso werden angenehme Ereignisse eher überschätzt und unangenehme hingegen eher unterschätzt (Kirchler, 2008). Obwohl fehlerhafte Erinnerungen auftreten können basiert der Großteil, aller mit Persönlichkeitsfragebögen erhobenen Information, lediglich auf Selbstberichtsdaten und liefert so hauptsächlich Erkenntnisse darüber, wie sich eine Testperson selbst sieht und entspricht damit eher dem Selbstbild der Person (Becker, 2003). Nach Mummendey und Grau (2008) lässt sich die Beantwortung eines Fragebogens in einem gewissen Ausmaß auch als eine Form der Selbstdarstellung auffassen.

Generelle Verfälschbarkeit

Schwierigkeiten bei der Technik des Fragens sind nicht nur die zum Teil verzerrten Erinnerungen und Selbsteinschätzungen, sondern auch eine Grundvoraussetzung, welcher nach, die befragten Personen sich selbst überhaupt kennen und zu beobachten im Stande sein müssen. Es ist jedoch auch denkbar, dass es einer Person generell an der Fähigkeit zur Selbstbeschreibung mangelt. In diesem Zusammenhang spricht man vom Phänomen der *Alexithymie*, oder dem Unvermögen einer Person Gefühle angemessen wahrnehmen und beschreiben zu können (Kubinger, 2009). Ob eine Testperson über eine ausreichende Fä-

higkeit zur Selbstbeschreibung verfügt oder nicht, wird in den meisten Fällen vor der Verwendung eines Persönlichkeitsfragebogens nicht ermittelt und stellt somit ein Problem hinsichtlich der Interpretation der gewonnenen Ergebnisse dar. Selbst wenn eine befragte Person über ausreichende Selbstbeschreibungsfähigkeiten verfügt, so muss diese auch bereit sein, die gefragte Information zur Verfügung zu stellen (Amelang et al., 2006). Da im Regelfall bei den meisten Fragebögen die Messintention klar ersichtlich ist, besteht für die befragte Person auch die Möglichkeit ihr Antwortverhalten in gewünschter Richtung zu manipulieren und damit das Endergebnis zu verzerren. Kubinger (2002) zeigte in einer Studie zur Verfälschbarkeit von Persönlichkeitsfragebögen, dass selbst ohne zu erwartende negative Konsequenzen und unabhängig vom vorgegeben Antwortformat sogenanntes „faking-good“ betrieben wird.

In einer weiteren Studie zur Verfälschbarkeit von Ziegler, Danay, Schölmerich und Bühner (2010) wurden studentische Versuchspersonen in eine Experimental- und eine Kontrollgruppe aufgeteilt. Die Kontrollgruppe wurde instruiert ehrlich zu antworten, während die Testpersonen der Experimentalgruppe sich vorstellen sollten, Bewerber eines Auswahlverfahrens zu sein, bei dem es um die Vergabe von begrenzten Studienplätzen gehe. Die Teilnehmer der Experimentalgruppe sollten dabei einen möglichst guten Eindruck erzeugen, aber gleichzeitig so unauffällig vorgehen, dass ein Experte den Manipulationsversuch nicht bemerke. (Ziegler et al., 2010)

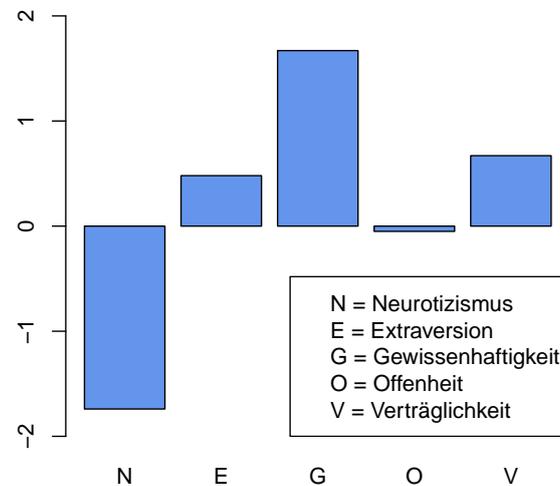


Abbildung 1: Verfälschung eines Persönlichkeitsinventars in einer imaginierten Auswahl-situation nach Ziegler et al. (2010), Tab.1 ; Abgebildet sind hier die Abweichungen der neutralen Situation in Standardabweichungen von der Kontrollbedingung.

In Abbildung 1 sind die Ergebnisse der Studie von Ziegler et al. (2010) dargestellt. Die Abweichungen der Experimentalgruppe von der neutralen Kontrollgruppe sind hier klar veranschaulicht. Diejenigen Testpersonen, welche sich bei der Beantwortung des Fragebogens eine Auswahl-situation vorstellen sollten, beschrieben sich im Vergleich zur Kontrollgruppe als wesentlich gewissenhafter und weniger neurotisch. Dies lässt den Schluss zu, dass die Testpersonen davon ausgingen, dass es vorteilhafter erschien gewissenhafte Charakterzüge systematisch zu übertreiben und gleichzeitig bei neurotische Eigenschaften zu untertreiben (Ziegler et al., 2010).

Nach Kubinger (2009) erfüllt ein Test das Gütekriterium der Unverfälschbarkeit dann, wenn die getestete Person ihr Testergebnis nicht, oder nur unwesentlich nach eigenem Belieben beeinflussen kann. Während es in allen Leistungstests zwar möglich, aber kaum wahrscheinlich ist, ein besseres Ergebnis zu erzielen als es dem eigenen Fähigkeitsniveau entspricht, so ist dies bei Persönlichkeitsfragebögen aufgrund der hohen Augenscheinvalidität nicht nur möglich, sondern in der Regel sogar sehr wahrscheinlich, dass eine Testperson zu ihrem persönlichen Vorteil antwortet (Kubinger, 2009). Die hohe Transparenz von Fragebögen führt dazu, dass für die Testperson Schlüsse über jenes Konstrukt möglich werden, auf welches die einzelnen Fragen abzielen und eröffnen somit auch die Möglichkeit ihre Antworten in eine positive Richtung zu verfälschen (Khorramdel & Kubinger, 2006).

Als Möglichkeiten zur Kontrolle von Antworttendenzen, also zum Beispiel von Verfälschbarkeit, nennt Seiwald (2003) die adäquate Wahl des Antwortformats, umgepolte Items, Kontrollitems oder Kontrollskalen sowie eine vor Verfälschung warnende Instruktion. Laut Vogel (2011) ist es trotz solcher Methoden nicht möglich Verfälschbarkeit komplett auszuschalten, denn die hohe Augenscheinvalidität von Persönlichkeitsfragebögen ist trotzdem weiterhin gegeben. In Hinblick auf die Verfälschbarkeit von Persönlichkeitsfragebogen stellen Objektive Persönlichkeitsverfahren laut Kubinger (2009) eine Lösung für dieses Problem dar. Des Weiteren gibt Kubinger (2009) zu bedenken, dass trotz der unveränderlichen Problematik der Verfälschbarkeit weiterhin mehr Persönlichkeitsfragebögen als Objektive Persönlichkeitstests erstellt werden.

2 Experimentalpsychologische Verhaltensbeobachtung

Innerhalb der Persönlichkeitsforschung gibt es kaum ein Gebiet, welches nicht maßgeblich durch das Wirken von Raymond B. Cattell beeinflusst wurde. Nicht nur seine Einflüsse in Bezug auf sein faktorenanalytisch begründetes Modell der 16 bipolaren, situationsunabhängigen Grundeigenschaften der Persönlichkeit waren richtungsweisend, sondern auch sein multimodaler Ansatz der Persönlichkeitsmessung verdient besondere Beachtung (Amelang et al., 2006).

In seiner Theorie der Persönlichkeit geht Cattell von drei zentralen Messmodalitäten aus: Q-Daten (*Fragebögen*), L-Daten (*Verhaltensbeobachtungen und Lebensdaten*) sowie T-Daten (*objektive Tests*) zur Erfassung der Persönlichkeit, wobei versucht wird für letztere das Verhalten eines Individuums in einer standardisierten Situation möglichst ohne Selbstbeurteilung zu erfassen (Ortner, Proyer & Kubinger, 2006). Wenn auch nicht als Erster oder Einziger überhaupt, doch sicher am konsequentesten versuchte Cattell mit diesen Mitteln die Schwachstellen in Fragebogenverfahren auszugleichen (Amelang et al., 2006).

2.1 Objektive Persönlichkeitstests

Laut Pawlik (2006) existiert die Grundlage der Erforschung der Objektive Persönlichkeitstests schon seit Ende der 1920er Jahre, jedoch entwickelten erst Cattell und Eysenck in den 1960er Jahren diesen Ansatz in systematischer Art und Weise weiter. Der Begriff „Objektive Persönlichkeitstests“ geht ursprünglich

auf Cattell zurück und bezieht sich auf eine alternative Erhebungstechnik zur Erfassung der Persönlichkeit, welche anders als klassische Persönlichkeitsfragebögen nur schwer oder überhaupt nicht verfälschbar sein sollen.

L. R. Schmidt (1975) definiert Objektive Persönlichkeitstests als:

Objektive Tests (T-Daten) zur Messung der Persönlichkeit und Motivation sind Verfahren, die unmittelbar das Verhalten eines Individuums in einer standardisierten Situation erfassen, ohne dass dieses sich in der Regel selbst beurteilen muss. Die Verfahren sollen für den Probanden keine mit der Messintention übereinstimmende Augenscheinvalidität haben. Das kann durch die Aufgabenauswahl oder bestimmte Auswertungsmethoden erreicht werden. Um als Test zu gelten, müssen auch die Objektiven Verfahren den üblichen Gütekriterien psychologischer Tests genügen. (S.292)

Laut Kubinger (2009) sei das Ziel von Objektiven Persönlichkeitstests eine Registrierung und Aufzeichnung des Arbeits- und Prozessverhaltens bei der Bearbeitung von theoriegeleiteten (psychologischen) Aufgabenstellungen und entsprechen daher so gesehen eher dem Konzept einer systematischen Verhaltensbeobachtung unter experimentell variierten Bedingungen. Er schlägt daher den Begriff „Experimentalpsychologische Verhaltensbeobachtung“ mit folgender Definition vor:

„Die experimentalpsychologische Verhaltensdiagnostik als eine (psychologische) „Technologie“ bezieht sich auf Verfahren, die aus dem beobachtbaren Verhalten bei experimentell variierten Leistungsanforderungen persönliche Stilmerkmale erschließen, wobei die Regis-

trierung der Art und Weise der Problembearbeitung der Computer übernimmt“ (Kubinger, 2009, S.260).

Unter Art und Weise der Problembearbeitung können nach Kubinger (2009) Variablen wie Reaktionsgeschwindigkeit oder der Bearbeitungsverlauf verstanden werden.

2.1.1 Objektivität bei Objektiven Persönlichkeitstests

Erhebungsobjektivität

Objektivität in Bezug auf die Experimentalpsychologische Verhaltensbeobachtung lässt sich laut Pawlik (2006) in Erhebungsobjektivität, Durchführungsobjektivität und Auswertungsobjektivität unterteilen. Erhebungsobjektivität bezeichnet das Ausmaß, wie sehr das Ergebnis eines diagnostischen Verfahrens von äußeren Rahmenbedingungen, wie Setting, Umgebungsfaktoren oder Testleiter unabhängig sind (Pawlik, 2006). Im Gegensatz zu den ersten Objektiven Persönlichkeitstests, welche noch als Paper-Pencil Verfahren durchgeführt wurden, bauen neuere Verfahren dieser Klasse auf die computerisierte Testvorgabe. Diese garantiert ein Maximum an Standardisierung bezogen auf die Erhebungsobjektivität.

Erfassungsobjektivität

Die Erfassungsobjektivität kann auch gleichgesetzt werden mit dem Gütekriterium der Unverfälschbarkeit. Ein Test entspricht diesem Kriterium, wenn es der Testperson nicht möglich ist, ihre Ergebnisse nach eigenem Belieben zu beeinflussen (Pawlik, 2006). Ob Objektive Persönlichkeitstests wirklich den Ansprüchen der Unverfälschbarkeit gerecht werden können, wurde in einer Studie

von Kubinger und Hofmann (2001) zur Validität der Objektiven Persönlichkeitsbatterie Arbeitshaltungen (Kubinger & Ebenhöf, 2002) überprüft. Bei einem Vergleich, mittels sog. Multitrait-Multimethod Matrix (Campbell & Fiske, 1959) wurden Vertreter unterschiedlicher Erhebungsmethoden einander gegenübergestellt und verglichen. Zum Einsatz kamen die beiden objektiven Persönlichkeitstestbatterien, Arbeitshaltungen (Kubinger & Ebenhöf, 2002), und das objektive Belastbarkeits-Assessment BACO (Ortner, Kubinger, Schrott & Litzenberger, 2006). Diese beiden Verfahren wurde jeweils mit dem NEO-FFI Borkenau und Ostendorf (1993) als klassischen Vertreter der Fragebogenmethode, sowie einem semiprojektiven Verfahren verglichen. Die beobachteten Daten ließen sich jedoch eher durch die Erfassungsmethode erklären, als durch ein gemeinsames Eigenschaftsmodell. Folglich schlossen die Autoren, dass es sich bei den von objektive Persönlichkeitsverfahren erfassten Aspekten der Persönlichkeit, um andere handele als es bei klassischen Fragebögen der Fall wäre. Wobei zu erwarten war, dass Ergebnisse einer Verhaltensbeobachtung von jenen einer Selbstbeschreibung abweichen. Ein weiteres Ergebnis war, dass ca. 25% der Probanden das Messprinzip des Objektiven Persönlichkeitstests durchschauten, die Testergebnisse dadurch aber unbeeinflusst blieben. Im Vergleich dazu durchschauten nahezu 90% der Testpersonen das Messprinzip des Persönlichkeitsfragebogens und erreichten in diesem dadurch auch bessere Ergebnisse (Kubinger & Hofmann, 2001) .

Auswertungsobjektivität

Unter Auswertungsobjektivität versteht man das Ausmaß, in dem das Ergebnis einer Testperson stets auf die gleiche Weise ausgewertet wird und unabhängig von der Person ist, die dieses Ergebnis erstellt Pawlik (2006). Ebenso wie bei der Erhebungsobjektivität, führt die Vorgabe eines objektiven Persönlichkeits-

tests als computerisiertes Verfahren zu dem Vorteil einer maximalen Auswertungsobjektivität und Verrechnungssicherheit.

2.1.2 Kritik an experimentalpsychologischen Verhaltensdiagnostik

Die Grundlage dieser Verfahrensklasse ist im Prinzip eine Täuschung der Testperson. Indem die zu bearbeitende Aufgabe primär als Leistungsaufgabe und nicht als Teil eines Persönlichkeitsverfahrens aufgefasst wird, bleibt die Messintention für die Testperson verschleiert. Eine Verzerrung der Ergebnisse seitens der Testperson wird dadurch deutlich erschwert (Kubinger & Hofmann, 2001).

Ein Vorenthalten der Messintention ist aber streng genommen ethisch nicht korrekt und steht damit in Konflikt mit der Zumutbarkeit dieser Verfahren für die Testperson (Kubinger, 2009). Neuere Vertreter dieser Verfahrensklasse versuchen durch eine möglichst hohe ökologische Validität, im weitesten Sinne eine Arbeitsprobe herzustellen. Die Messintention wird dabei zwar nicht wirklich verheimlicht, aber ein Erkennen dieser durch die Testperson bleibt weiterhin erschwert. Weitere Probleme ergeben sich durch einen nicht unbeträchtlichen Mehraufwand bei der Entwicklung solcher Verfahren. Zudem besteht derzeit noch ein Mangel an überzeugenden Validierungsstudien zu dieser Verfahrensklasse (Schmidt-Atzert & Amelang, 2012).

Ein weiterer Kritikpunkt betrifft weniger eine inhaltliche Gültigkeit dieser Verfahren, als mehr die Generalisierbarkeit ihrer erzielten Beobachtungen (Schmidt-Atzert & Amelang, 2012). Hinsichtlich einer prognostischen Validität konnte Frebort (2002) nachweisen, dass einzelne Testkennwerte der *Arbeitshaltungen* (Kubinger & Ebenhöf, 2002) bei der Auswahl von TierpflegeschülerInnen signifikant zwischen erfolgreichen und weniger erfolgreichen Schülern unterscheiden

können. Die zuweilen noch sehr überschaubare Anzahl an Studien zur prognostischen Validität zeigt noch deutlichen Forschungsbedarf auf. Haupteinsatzgebiete dieser Verfahren liegen derzeit innerhalb der Selektionsdiagnostik, beispielsweise bei der verkehrspsychologischen Fahreignungsdiagnostik. Da die Ergebnisse und somit der Ausgang einer Testung in diesem Anwendungsgebiet für die Testperson besonders relevant sind, erhöht sich bei herkömmlichen Fragebögen auch das Auftreten vorsätzlich verzerrter Ergebnisse (Kubinger, 2009). Obwohl die Technik Objektiver Persönlichkeitstests den geforderten Ansprüchen an das Gütekriterium der Objektivität weitaus besser entsprechen als andere Verfahren, wurde dieser vielversprechende Ansatz nur von einigen Wenigen weiter verfolgt. Aufgrund des Mangels an Forschung und Neuentwicklungen in diesem Gebiet, wirken die bisher entwickelten Objektiver Persönlichkeitsverfahren meist veraltet. Als Beispiel kann der Test ILICA (Möseneder & Ebenhöf, 1996) genannt werden, welcher aus einer Fülle von Text besteht und daher eher umständlich für die Testperson zu bearbeiten ist. Die neu gewonnenen technischen Möglichkeiten des Mediums Computer, werden bisher kaum genutzt und bieten daher noch großes Potential für Neuentwicklungen.

3 Computersimulationen

Aufgabe der psychologischen Diagnostik ist nicht nur die Erfassung eines aktuellen „Ist-Stand“ von Leistungs- und Persönlichkeitsparameter zu erheben, sondern darüber hinaus auch Prognosen für zukünftiges Verhalten zu erstellen (Kubinger, 2009). Die gängigsten Instrumente innerhalb der Persönlichkeitsdiagnostik, welche derzeit Verwendung finden, sind textbasierte Verfahren (van der Linden & Glas, 2000). Diese Form der Erhebung ist hauptsächlich an Aussagen über vergangenes Verhalten gebunden und muss darauf vertrauen, dass die gemachten Angaben der Realität entsprechen. Die Durchführung eines Quasiexperiments mit einer Testperson unter kontrollierten Bedingungen, wäre die ökologisch validere Alternative zu textbasierten Erhebungsinstrumenten. Reales Verhalten könnte so in einer möglichst realitätsnahen Situation beobachtet werden und durch zusätzliche Informationen auch eine präzisere Prognose über zukünftiges Verhalten erstellt werden. Eine solche Art der Untersuchung bringt allerdings auch Nachteile mit sich. Der Mehraufwand bezüglich Planung, Durchführung und Auswertung, sowie die höheren Kosten übersteigen in den meisten Fällen den dadurch gewonnen Zusatznutzen. Ebenso ist die Kontrolle der Störvariablen, wie beispielsweise Versuchsleitereffekte, deutlich erschwert. Dennoch fand dieser Ansatz in abgemilderter Weise innerhalb der Eignungsdiagnostik vielfach Anwendung in Form von Arbeitsproben (Nerdinger, Blickle & Schaper, 2008). Eine Arbeitsprobe ist ein diagnostisches Verfahren, welches aus einer realitätsnahen Simulation wichtiger Arbeitsaufgaben besteht (Höft & Funke, 2001). Die Bearbeitung einer sogenannten „Postkorbaufgabe“ stellt eine typischen Vertreter einer Arbeitsprobe dar und wird häufig als Teil eines Assessmentcenters durchgeführt. Hierbei werden, anders als bei einem Probearbeitstag, nur grundlegende Fähigkeiten zur Bewältigung einer späteren Aufgabe

überprüft und nicht die spätere Aufgabe selbst. Die Art der Aufgabenbewältigung wird von geschulten Beobachtern registriert und notiert. Der Vorteil liegt in einer standardisierten Aufgabenstellung und Interpretation (Nerdinger et al., 2008).

Schon in den Anfängen des 20. Jahrhunderts, zu Zeiten der „industriellen Psychotechnik“ wurde eine Apparatur zur Tauglichkeitsuntersuchung von Straßenbahnfahrern entwickelt, die einer frühen Form eines Fahrsimulators ähnelte. Die Bewerber mussten auf Reize, welche auf einer vor ihnen aufgebauten Anzeigetafel präsentiert wurden, möglichst schnell und mit dem korrekten Verhalten reagieren (Kirchler, 2008). Die psychologische Diagnostik bediente sich also schon sehr früh, ihrer jeweiligen Epoche entsprechend, technischer Hilfsmittel und wurde auch auf den Einsatz von Computern als Hilfsmittel für den diagnostischen Prozess aufmerksam. Mit der Entwicklung immer leistungsstärkerer Computer veränderten sich auch dessen Anwendungsgebiete.

Die Anwendung von Computern erstreckt sich heute über die gesamte Spannweite des diagnostischen Prozesses, von der Item-Vorgabe, über die Aufzeichnung, bis hin zur Verrechnung und Auswertung (van der Linden & Glas, 2000). Da Computer multimedial nutzbar sind, bietet diese Technologie ein Vielfaches an Möglichkeiten zur Vorgabe von Stimulusmaterial. Von der Vorgabe einfacher Texte, über Bilder und Videos, bis hin zu aufwendigen Computersimulationen (van der Linden & Glas, 2000). Computersimulationen versuchen „...dynamische und vernetzte Aspekte der Umwelt, die das Verhalten, die Einstellung oder Emotionen beeinflussen, realitätsnah abzubilden. An die Stelle von Schilderungen in textlicher Form, treten wirklichkeitsnahe Stimuli“ (Funke & Reuschenbach, 2011, S.602).

Der Unterschied einer Computersimulation zu einem Objektiven Persönlichkeitsverfahren kann am Grad der angestrebten Realitätsnähe festgemacht werden. Objektive Persönlichkeitsverfahren bestehen zumeist aus eher abstrakten, oder sehr vereinfachten Aufgabenstellungen mit geringer Realitätsnähe, was die Bestimmung sinnvoller Maßeinheiten für die Erfassung psychologischer Konstrukte vereinfacht. Als typischer Vertreter der experimentalpsychologischen Verhaltensbeobachtung kann hier beispielsweise der Subtest "Symbole Kodieren" im Test *Arbeitshaltungen* (Kubinger & Ebenhöf, 2002) genannt werden. Computersimulation hingegen streben eine möglichst hohe Realitätsnähe an. Was nicht bedeuten soll, dass sich die genannten Ansätze ausschließen. Der Subtest „Ungünstige Arbeitsbedingungen“ der Objektiven Persönlichkeitsbatterie BACO (Ortner & Kubinger, 2006) versucht, mit Hilfe einer möglichst Alltagsnahen Simulation eines störanfälligen Computers die Arbeitsweise einer Testperson bei ungünstigen Arbeitsbedingungen am PC zu erfassen.

Nach Schiff, Arnone und Cross (1994, zitiert nach Funke & Reuschenbach, 2011) hängt die Realitätsnähe einer Computersimulation zum einen von der Ähnlichkeit der präsentierten Reize mit der Umwelt ab (*Display realism*) und zum anderen, von der Art, wie in die Simulation eingegriffen werden kann (*Response system realism*). Damit ist gemeint, dass bei einer Simulation nicht nur die Objekte möglichst wirklichkeitsgetreu dargestellt werden sollten, sondern auch deren physikalisches Verhalten eine Rolle spielt. Computersimulationen, welche sich komplexer 3D Technologie bedienen, sind heutzutage in der Lage ein computergeneriertes Abbild der Realität zu schaffen, welches als Virtuelle Realität (*Virtual Reality, VR*) bezeichnet wird und der Realität zum Verwechseln ähnlich ist.

3.1 Virtuelle Realität

Nach Kuhlen (2014) bezeichnet man als Virtuelle Realität die Darstellung und gleichzeitige Wahrnehmung der Wirklichkeit und ihrer physikalischen Eigenschaften in einer in Echtzeit interaktiven, computergenerierten Welt. Im Gegensatz zu einer reinen Animation steht bei der Virtuellen Realität die Möglichkeit der Interaktion im Vordergrund. Das bedeutet, der Benutzer kann in der virtuellen Welt frei navigieren und virtuelle Objekte manipulieren. Der Benutzer einer VR erlebt diese nicht nur als passiver Zuschauer, sondern als aktiver Akteur. Des Weiteren ist die Interaktion in der VR multimodal und bezieht möglichst viele menschliche Sinne mit ein. Je nachdem, welche technischen Hilfsmittel Verwendung finden, kann beim Benutzer das Empfinden entstehen, vollständig von der virtuellen Szene umgeben zu sein, oder in diese förmlich einzutauchen (Kuhlen, 2014). In diesem Zusammenhang spricht man auch von den beiden Begriffen *Telepresence* und *Immersion* (Minsky, 1980 zitiert nach Ijssels-teijn, Harper, Group et al., 2000).

3.1.1 Telepresence und Immersion

Sowohl Telepresence als auch Immersion befassen sich mit dem Gefühl des Eintauchens in eine virtuelle Welt, jedoch aus verschiedenen Perspektiven. Telepresence (abgekürzt *Presence*) bezieht sich auf den subjektive Bewusstseinszustand einer Person als *sense of being in the virtual environment* (Bente, Krämer & Petersen, 2002). Durch eine Fokussierung der Aufmerksamkeit grenzt sich die in der VR befindliche Person von äußeren Reizen ab und fühlt sich mehr in der VR anwesend, als in der physischen Welt. Der Begriff Presence beschreibt dieses Gefühl auf einer rein subjektiven Ebene. Die Subjektivität dieser Empfindung bedeutet auch, dass dieses Gefühl interindividuell stark variieren

kann. Die Einflussfaktoren auf das Presence-Empfinden können nach Witmer und Singer (1998) in drei Klassen unterteilt werden: *Control factors*, beschreiben das Ausmaß an Kontrolle und die Modifizierbarkeit der VR seitens des Nutzers. *Sensory factors* beschreiben, die Vielzahl verwendeter Sinnesmodalitäten und deren Detailtreue. *Realism factors* umschreiben den Grad an Realitätsnähe zu realen Objekten und deren Bedeutsamkeit in der VR für den Anwender. Ebenfalls müssen auch *distraction factors* berücksichtigt werden, welche sich auf den Grad der Ablenkung von der VR, sowie die Qualität der Bearbeitung dieser beziehen (Witmer & Singer, 1998).

Die individuelle Tendenz, sich von einer VR involvieren zu lassen (*involvement*) und in diese einzutauchen, kann nach Witmer und Singer (1998) als Immersion beschrieben werden. Nach Schubert, Friedmann und Regenbrecht (2001) ist Immersion folgendermaßen von Presence abzugrenzen:

„While immersion is objectively quantifiable, [...] the sense of presence, is a subjective experience and only quantifiable by the users experiencing it“ (S.3). Das Gefühl von Immersion hängt also eher von „objektiven Stimulibedingungen“, wie den verwendeten technischen Geräten zur VR-Erzeugung ab, während das Gefühl der Presence einen durch Immersion erzeugten, Bewusstseinszustand darstellt (Schubert et al., 2001).

3.1.2 Anwendung virtueller Umgebung in der Psychologie

Schon Loomis, Blascovich und Beall (1999) heben die möglichen Anwendungsgebiete von virtuellen Umgebungen für die psychologische Forschung hervor. Jedoch ist die Nutzung virtueller Umgebungen innerhalb der Psychologie noch größtenteils auf die Grundlagenforschung beschränkt. Eine Ursache dafür ist sicherlich der große Entwicklungsaufwand, eine solche virtuelle Umgebung um-

zusetzen, sowie die hohen technischen Anforderungen die solch ein System benötigt. Als weiterer Nachteil erweist sich das Problem der *Cybersickness* (vgl. LaViola, 2000), welche symptomatisch der See- oder Reisekrankheit ähnelt. Eine Differenz zwischen visuellem und vestibulären Sinneswahrnehmungen, wie sie in einer VR vorkommt, kann zu Übelkeit und Schwindelgefühlen führen, was die Anwendbarkeit virtueller Systeme erheblich einschränken kann (LaViola, 2000).

Dennoch, die Möglichkeit eine realitätsnahe Umgebung zu schaffen, welche zugleich streng kontrollierbar bleibt, findet derzeit schon innerhalb der Verhaltenstherapie Anwendung. Im Bereichen spezifischer Phobien, kann so mittels VR eine Expositionsbehandlung, mit schrittweisem Annähern an das angstausslösende Objekt erzielt werden. Die hohe Kontrollierbarkeit, ermöglicht es sowohl dem Therapeuten als auch dem Patienten jeder Zeit die optimale Dosis des Reizes zu regulieren. So konnten bei Studien zu Höhenangst (Emmelkamp et al., 2002), Spinnenphobie (Emmelkamp et al., 2002; Opdyke, Williford & North, 1995), Sozialen Ängsten (Anderson, Rothbaum & Hodges, 2003) oder auch bei Flugangst (Rothbaum, Hodges, Anderson, Price & Smith, 2002) positive Behandlungserfolge erzielt werden. In Bereich der neuropsychologischen Forschung konnte gezeigt werden, dass auch die Nutzung von nicht immersiven virtuellen Umgebungen, bei vereinfachter Ausführung, sogar im geriatrischen Bereich Anwendung finden kann (vgl. Allain et al., 2014).

In einer Studie von Hartig, Frey und Ketzler (2003) versuchten die Autoren zum einen die generelle Anwendbarkeit virtueller Umgebungen für die experimentalpsychologische Forschung zu erproben und zum anderen dafür bereits bestehende Computerspiele zu modifizieren, um so den Aufwand für die Neuentwick-

lung virtueller Umgebungen zu reduzieren. Für diesen Zweck verwendeten die Autoren das 3D-Spiel Quake III Arena, welches der Kategorie der First-Person-Shooter angehört. Es ist bei dieser Art Videospiel möglich, sich frei in einer am PC Bildschirm präsentierten Welt zu bewegen. Als weitere Option können bei diesem Spiel eigene Umgebungen gestaltet und genutzt werden. Die Autoren fixierten den Blickwinkel und beschränkten die Steuerung lediglich auf die Pfeiltasten der Tastatur, womit die Steuerung erheblich vereinfacht wurde. Darüber hinaus wurden alle, für dieses Spielgenre typische Elemente wie z. B. Waffen entfernt, so dass eine neutrale virtuelle Umgebung übrig blieb. An einer Stichprobe von 85 Personen wurden fünf abgeschlossene virtuelle Umgebungen unterschiedlicher Komplexität untersucht. Dabei sollten die Testpersonen mithilfe der Pfeiltasten durch die ihnen präsentierten Räume navigieren und dabei möglichst schnell ans Ziel kommen. Hartig et al. (2003) kamen zu dem Schluss, dass sich insbesondere virtuelle Umgebungen mit niedriger Komplexität für psychologische experimentelle Forschung eignen und vielfältige Möglichkeiten bieten können.

Diese Aussage stützt auch die Aussage von Washburn (2003), wonach sich Computerspiele als Werkzeug experimentell-psychologischer Forschung durch ökologisch valide Verhaltensanforderungen, standardisierte Untersuchungen sowie einem höheren Wohlbefinden von Versuchspersonen auszeichnen würden. Gleichwohl stellt sich die Frage, wie man aus der Verhaltensvielfalt von Computerspielen sinnvolle Maßeinheiten für die Erfassung psychologischer Konstrukte extrahieren kann. Der Vorteil objektiver Persönlichkeitstests im Vergleich zu Computerspielen ist eine wesentlich schlichtere Gestaltung des Testmaterials, was die Möglichkeit bietet das beobachtete Verhalten zu Testkennwerten zusammenzufassen und zu interpretieren. Der Nachteil einer simpleren Gestaltung

bleibt jedoch die geringere Realitätsnähe und Generalisierbarkeit. Die vorliegende Studie versucht hier eine Brücke zwischen gesteigerter Realitätsnähe und gleichzeitiger Interpretierbarkeit der gewonnenen Daten zu bilden.

Teil III

Empirischer Teil

4 Zielsetzung der empirischen Untersuchung

In den bisherigen Ausführungen konnte gezeigt werden, dass die Methoden der psychologischen Diagnostik sich größtenteils der Technik des Fragens bedienen. Persönlichkeitsfragebögen haben jedoch besonders in Auswahl-situation aufgrund ihrer hohen Augenscheinvalidität den Nachteil besonders anfällig für eine Verfälschung der Ergebnisse zu sein. Die direkte Beobachtung von Verhalten, im Rahmen von systematischen Verhaltensbeobachtungen in einem sogenannten Assessmentcenter, anstelle des reinen Erfragens, ist jedoch häufig zu kostenintensiv und bei einer größeren Anzahl von Bewerbern schlicht nicht mehr durchführbar. Als zweiter Kritikpunkt einer direkten Beobachtung ist die Objektivität seitens des Beobachters nicht immer gegeben, oder nur durch zeitintensive Schulungen zu erreichen. Obwohl objektive Persönlichkeitsverfahren eine hier eine sinnvolle Alternative bieten, so stellt sich doch die Frage, was eigentlich genau dabei gemessen wird, da es immer noch an ausreichenden Validierungsstudien mangelt.

Daher ist das Ziel dieser Studie eine objektive, computerisierte Alternative für diese Probleme zu erproben. Unter Berücksichtigung der Erkenntnisse experimentell-psychologischer Verhaltensbeobachtung und virtueller Computersimulationen wird versucht, anhand eines bedeutsamen und verzerrungssensiblen Merkmals der Persönlichkeit eine multidimensionale Alternative zu eindimensionalen Fragebögen zu schaffen. So soll diese Arbeit einen Beitrag dazu leisten experimentell-

psychologische Verhaltensbeobachtungen moderner und realitätsnäher (und dadurch vielleicht für die Testperson motivierender) zu gestalten. Zu diesem Zweck wurde ein neuartiger Test entwickelt, der auf das Konstrukt Gewissenhaftigkeit bezogenes Verhalten in einer virtueller Umgebung möglichst objektiv erfassen soll. Gleichweg werden so auch viele neue Fragen aufgeworfen. Beispielsweise, ob tatsächlich mehr erfasst wird als mit herkömmlichen Objektiven Persönlichkeitsverfahren und somit der Mehraufwand gerechtfertigt ist, oder inwieweit virtuelles Verhalten realem Verhalten entspricht. Auf Grund mangelnder Vergleichsstudien handelt es sich daher in der vorliegenden Arbeit um eine reine Machbarkeitsstudie.

5 Testbeschreibung und -Konstruktion (VICI)

5.1 Entstehungsgeschichte

Der Test VICI (Vienna Conscientiousness Inventory) wurde ausschließlich unter Verwendung von lizenzfreier Software erstellt. Die 3-dimensionalen Objekte für die virtuelle Umgebung wurden mit dem Programm Blender 3D erstellt und alle dafür nötigen Texturen und Grafiken mit dem Programm Gimp 2.8. Blender 3D ist eine freie unter der GPL lizenzierte Software und dient zur Bearbeitung von dreidimensionalen Objekten. Diese können modelliert, texturiert, animiert sowie gerendert werden (Wikipedia, 2014a). Standardmäßig ist in Blender auch eine 3D Game-Engine enthalten. Eine Game-Engine ist ein spezielles Framework für Computerspiele, welche den Spielverlauf steuert und für die visuelle Darstellung sowie Simulation physikalischer Gesetze innerhalb des Spielablaufs verantwortlich ist (Wikipedia, 2014b). So ist es möglich eine Computersimulation von der Erstellung aller 3D-Modelle, über die Texturierung, Animation bis hin

zu Kontrolle aller erdenklichen physikalischen Parameter zu erstellen und über die objektorientierte Programmiersprache Python zu kontrollieren. Die Erstellung und Programmierung des Tests VICI erstreckte sich über einen Zeitraum von zwei Jahren und wurde in allen Schritten vom Autor selbst ausgeführt.

5.2 Phase 1 (Einleitung)

Der Test VICI kann in drei unterschiedliche Phasen unterteilt werden. Die erste Phase umfasst ein Datenblatt zur Erfassung soziodemografischer Daten, einem Fragebogen, sowie ein Screening zur Augen-Hand Koordination.

5.2.1 Datenblatt

Im Datenblatt werden Informationen über Geschlecht, Alter, Bildungsgrad sowie Grad der Computererfahrung im Allgemeinen, als auch die spezielle Erfahrung im Bereich der 3D-Computerspiele erfasst. Bei letzteren beiden kann die jeweilige Erfahrung aus drei Abstufungen (*keine*, *wenig* und *viel Erfahrung*) gewählt werden. Um die anonymisierten Ergebnisse den einzelnen Testpersonen zuordnen zu können, wird den Testpersonen im Datenblatt ein spezifischer Probandencode zugewiesen. Im Anschluss an das Datenblatt folgt die Vorgabe der 48 Items zur Gewissenhaftigkeit des NEO PI-R nach Borkenau und Ostendorf (2008). Bei der Testkonstruktion des VICI wurde versucht, ein Setting zu konstruieren indem sich einige Facetten der Gewissenhaftigkeit beobachten lassen, welche auch im NEO PI-R vorkommen. Der NEO PI-R wurde ausschließlich für die Validierungsstudie in das Computerprogramm integriert und stellt keinen festen Bestandteil des Test VICI dar. Es geht lediglich um den Vergleich von Selbstbeschreibung und beobachteten Verhalten.



Abbildung 2: Datenblatt des Test VICI sowie die Einbettung der Facette Gewissenhaftigkeit des NEOPI- R (Borkenau & Ostendorf, 2008)

5.2.2 Szenario

Nach der Beantwortung des Fragebogens wird die Testperson in einem kurzen Text über das virtuelle Szenario aufgeklärt. Die Testperson erfährt, dass sie zu einem Probearbeitstag in eine Schokoladenfabrik eingeladen wurde und bei erfolgreichem Abschneiden eine lohnende Anstellung erwarten könne. Als Aufgabe müsse die Testperson alleine die Schokoladenproduktion in einer Halle beaufsichtigen und für einen reibungslosen Ablauf sorgen. Es wurde also eine virtuelle Bewerbungssituation geschaffen, um das Anspruchsniveau der Testperson von Anfang an zu steigern.



Abbildung 3: Produktionshalle im Überblick

Das Szenario einer Schokoladenfabrik, als industrielle Produktionsstätte, wurde gewählt um eine möglichst breite Akzeptanz gewährleisten können. Darüber hinaus wurde versucht, das zu produzierende Produkt und die zu bewältigenden Aufgaben geschlechtsneutral zu gestalten, so dass systematische Verzerrungen reduziert werden. Die Wiener Schokoladenfabrik *Heindl* diente als Inspiration für die Gestaltung der virtuellen Umgebung im Test VICI. Wichtige Elemente wie Förderbänder, Kartons oder ein Büro wurden zwar nachgebildet, jedoch nur in stark vereinfachter Art und Weise. Des Weiteren wurde die Produktionshalle mit großen Fenstern versehen um offen und freundlich zu wirken. Die Halle besteht aus nur einem leicht zu überblickenden Raum, der es erlaubt, von jeder Perspektive den Überblick über die räumliche Orientierung zu behalten.

5.3 Screening Auge-Hand Koordination (AHKS)

Nachdem die Testperson grob über das virtuelle Setting informiert wurde, folgt ein Screening der Auge-Hand Koordination mittels PC-Maus. Das Screening

wurde in den Kontext des Probearbeitstages eingebettet, indem die Testperson mit einem Auto zunächst zu ihrer Fabrikhalle fahren muss. Da, dieser Ablauf plausibel erscheint, wird das Screening nicht als separater Test wahrgenommen und unterstützt so das virtuelle Szenario.



Abbildung 4: Screening zur Auge-Hand Koordination

Beim Screening selbst hat die Testperson die Aufgabe den Zeiger der PC-Maus auf ein dargestelltes Auto zu richten, um dieses in Bewegung zu setzen. Solange sich der Mauszeiger auf dem Auto befindet, fährt dieses. Bei einem Abrutschen des Mauszeigers stoppt das Auto. Die Aufgabe ist es nun, möglichst ohne Stopp, in die Arbeit zu gelangen. Dabei wird gemessen, wie häufig die Testperson mit dem Mauszeiger vom Auto abrutscht und anhält. Annahme ist, je häufiger das Auto stoppen muss, desto schlechter ist die Fähigkeit zur Kontrolle des Mauszeigers und damit auch zur Auge-Hand Koordination.

5.4 Test VICI - Phase 2 (Training)

5.4.1 Trainingsphase

Nach Abschluss der Vortestungen beginnt die Trainingsphase für den Test VICI. Von diesem Zeitpunkt an befindet sich die Testperson in einer dreidimensionalen, virtuellen Fabrikhalle und sieht diese aus der Ego-Perspektive.

5.4.2 Visuelle Orientierung

Bei modernen Videospiele wird meist die Steuerung aus der Ego-Perspektive gewählt. Als Ego-Perspektive oder auch *first-person point-of-view* wird eine Kameraeinstellung in Videospiele bezeichnet, bei welcher der Spieler die virtuelle Umgebung durch die Augen eines dort befindlichen Avatars sieht. Diese Perspektive wird genutzt um den Grad an Immersion zu steigern und somit ein stärkeres Gefühl der Präsenz zu erzeugen (Slater & Usoh, 1993). Der Spieler kann mit Hilfe eines Eingabegerätes (z.B. PC-Maus oder Joystick) den Blick nach Belieben verändern und sich so frei umsehen, als wäre er selbst vor Ort.

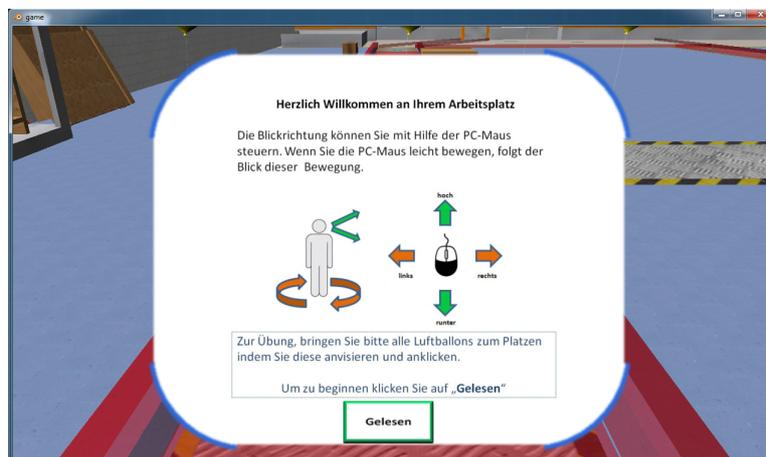


Abbildung 5: Instruktion "Visuelle Orientierung"



Abbildung 6: Training "Visuelle Orientierung"

Da das Wissen um die Steuerung von klassischen 3D-Spielen aus der Ego-Perspektive nicht bei allen Testpersonen vorausgesetzt werden kann, beginnt der Test VICI mit einem Training der grundlegenden Navigation. Die Blicksteuerung, stellt die wichtigste Fähigkeit für die Orientierung im virtuellen Raum dar und wird daher als erstes erläutert und trainiert. Die Übungsphase beginnt in der virtuellen Schokoladenfabrik mit einer Instruktion der Blicksteuerung durch die PC-Maus. Eine Bewegung der Maus nach rechts und links lenkt auch den Blick in der virtuellen Welt in die jeweilige Richtung. Analog dazu, entspricht eine vor oder zurück Bewegung der Maus, dem Blick nach oben oder unten. Nachdem die erste Instruktion gelesen und bestätigt wurde, trainiert die Testperson dies, indem sie Luftballons durch Anklicken zum Platzen bringt. Die so erlernte Blicksteuerung bietet der Testperson die Möglichkeit ihre neue Umgebung frei zu explorieren. Um den Entscheidungsspielraum noch zu erhöhen, kann unter Verwendung der Computertastatur auch die Position im Raum auf der X und Y-Achse frei variiert werden.

5.4.3 Navigation im Raum

Nach Beenden des Trainings zur visuellen Orientierung, folgt das Training zur Navigation im Raum. In klassischen 3D-Videospielen werden die Tasten W-A-S-D der Tastatur genutzt, um sich innerhalb der virtuellen Umgebung vor, zurück oder seitwärts zu bewegen. Diese Tastenbelegung wurde beim Test VICI übernommen und zur besseren Erkennbarkeit für die Testpersonen farblich markiert.

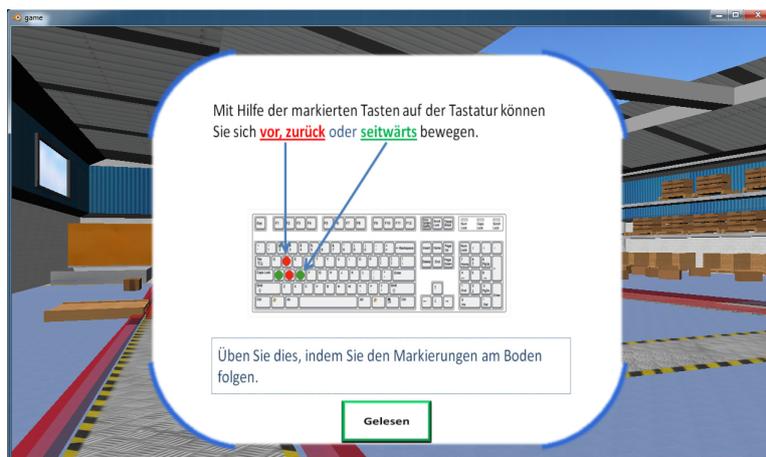


Abbildung 7: Instruktion "Navigation im Raum"



Abbildung 8: Training "Visuelle Orientierung"

Nachdem die Testperson mittels W-A-S-D-Tasten erfolgreich einige Schritte getan hat, ist dieser Teil der Instruktion abgeschlossen. Die Kombination aus freier Navigation im Raum und freier Steuerung der Blickrichtung, ermöglicht es der Testperson sich ab diesem Zeitpunkt selbständig durch den Raum zu bewegen. Um zu vermeiden, dass einzelne Segmente des Trainings ausgelassen werden, ist der Bewegungsspielraum während der Trainingsphase auf einen, durch rote Balken gekennzeichnet Trainingspfad beschränkt. Blinkenden Pfeile auf diesem Weg zeigen die Richtung der nächsten, zu absolvierenden, Aufgabe an. Nach Beenden der Übungsphase, sind alle Beschränkungen aufgehoben und es hängt rein von den individuellen Entscheidungen der Testperson ab, wo sie sich im Raum aufhält und welche Aktionen sie als nächstes setzt.

5.4.4 Interaktion mit Objekten

Um den Grad an Freiheit weiter zu steigern, ist es der Testperson möglich, mit nahezu allen Objekten innerhalb der virtuellen Umgebung zu interagieren. Wie schon beim Training der Blicksteuerung erlernt, können Objekte in der Fabrikhalle angeklickt werden und reagieren darauf. Eine Tür öffnet sich wenn diese anvisiert und angeklickt wird oder bewegliche Objekte können durch die Testperson bewegt werden.

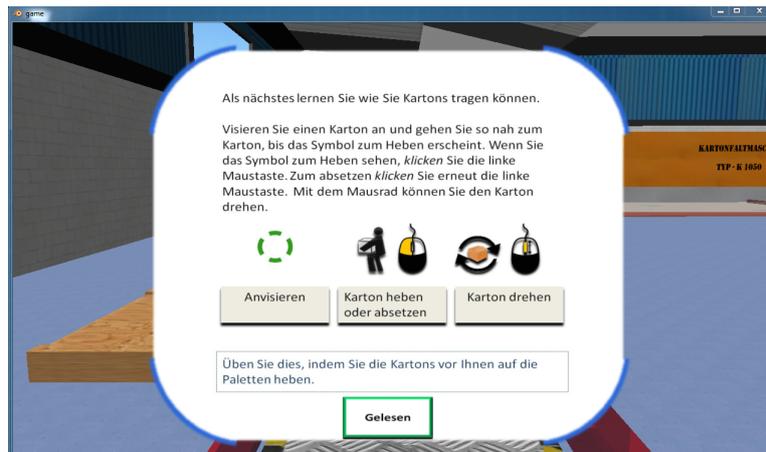


Abbildung 9: Instruktion “Interaktion mit Objekten”

Die dritte Stufe der Trainingsphase erläutert, wie Objekte hochgehoben und wieder abgesetzt werden können. Um ein bewegliches Objekt hochzuheben, muss die Testperson das gewünschte Objekt anvisieren und sich diesem so lange nähern, bis ein Symbol eingeblendet wird, welches anzeigt dass es sich in greifbarer Nähe befindet. Aus ausreichender Distanz zum gewünschten Objekt wird mittels Klick auf die linke Maustaste das Objekt hochgehoben. Ein hochgehobenes Objekt bleibt bis zu einem weiteren Klick der linken Maustaste im Blickfeld der Testperson fixiert. Während ein Objekt gehoben wird, kann es auch mittels Mausrad in vertikaler Achse gedreht werden.



Abbildung 10: Training "Interaktion mit Objekten"

Die Übungsaufgabe in diesem Fall stellen zwei Kartons dar, welche auf ihre jeweiligen Paletten gehoben werden sollen. Sobald beide Kartons auf ihrer Zielpalette abgesetzt wurden, ist dieser Teil des Trainings abgeschlossen. Die Manipulation von Objekten lässt zum einen mehr Verhalten beobachten und zum anderen steigert es das Interesse und die Motivation der Testperson, sich mit der virtuellen Umgebung zu beschäftigen. In diesem Fall können die Kartons innerhalb der Fabrik hochgehoben, gedreht und sogar geworfen werden. Es ist also möglich zu erfassen, welcher Karton wie lange, in welche Richtung getragen oder sogar geworfen wurde. Der individuelle Stil, eine Aufgabe zu bearbeiten, kann erfasst werden und gibt Aufschluss darüber, wie eine Person eine ähnliche Aufgabe in der Realität am wahrscheinlichsten lösen würde.

5.4.5 Arbeit am Fließband

Die Hauptaufgabe innerhalb des simulierten Probetages ist die Schokoladenproduktion am Fließband. In der vierten Station der Trainingsphase hat die Testperson die Möglichkeit diese Arbeit zu kennenzulernen und zu üben.



Abbildung 11: Ansicht auf das Fließband und die dazugehörigen Bedienelemente

Die Testperson kann mittels klick eines blinkenden Symbols, sichtbar am Computerterminal vor dem Fließband, in die Fließbandaufgabe wechseln. Einmal angeklickt, ändert sich die Perspektive von 3D-Egoperspektive, in eine stabilisierte Vogelperspektive mit Blick auf das Fließband. Bei einer nicht stabilisierten Perspektive wäre eine Vergleichbarkeit der Testpersonen nicht gegeben, da der Blickwinkel bei der Fehlersuche am Fließband ausschlaggebend sein kann. Die stabilisierte Vogelperspektive hat den Vorteil, dass jede Testperson bei der Arbeit am Fließband den gleichen Blickwinkel auf die Schokoladentafeln hat und dadurch auch die gleiche Chance hat, eine fehlerhafte Tafel zu entdecken. Der Fließbandbildschirm besteht im Grunde aus zwei Bereichen. Zum einen aus dem Fließband selbst und zum anderen aus den dazugehörigen Bedienelementen.

Schaltflächen zum Sortieren

Durch Klick auf die Schaltfläche „weiter“ wird eine Schokoladentafel produziert. Diese bewegt sich von der linken Seite des Fließbandes zur Mitte des Förder-

bandes, wo die produzierte Tafel überprüft werden kann. Auf der rechten Seite des Bildschirms ist eine Mustertafel zu sehen, mit welcher alle Tafeln abgeglichen werden können. Falls die produzierte Tafel nicht dem Muster entspricht, soll diese vom Fließband entfernt werden und die weitere Produktionskette verlassen. Dies kann mittels Schaltfläche „aussortieren“ erreicht werden. Ist die Tafel fehlerfrei, kann mit einem erneuten Klick auf die Schaltfläche „weiter“ diese Schokoladentafel in Kartons befördert werden.

Schaltfläche „Vergrößerung“

Eine dritte Schaltfläche mit der Aufschrift „vergrößern“ kann bei Unsicherheit seitens der Testperson genutzt werden. Diese Lupenfunktion zeigt im aktiven Zustand, auf einem dargestellten Bildschirm, ein vergrößertes Bild der aktuellen Tafel an. Im inaktiven Zustand zeigt der dieser Bildschirm lediglich die aktuelle Uhrzeit an. Es ist so möglich, eine Tafel genau auf Fehler zu prüfen, jedoch kostet diese genauere Prüfung auch mehr Zeit.



Abbildung 12: Aktive „Vergrößerung“

Für jede Entscheidung bleiben der Testperson fünf Sekunden Zeit. Wenn die Entscheidung vor Ablauf der Zeit erfolgt, spart diese Zeit und die Testperson kann mehr Tafeln in kürzerer Zeit produzieren. Trifft die Testperson keine Entscheidung innerhalb des Zeitlimits, wird die aktuelle Tafel auf dem Fließband als fehlerfrei angesehen und automatisch weitergeleitet. Falls die Schaltfläche zur „Vergrößerung“ gewählt wurde, ist es zwar möglich vor Ablauf der verbleibenden Zeit die Tafel auszusortieren, allerdings ist es dann nicht möglich die „weiter“ Schaltfläche zu verwenden. Die genauere Untersuchung einer Tafel mittels „Vergrößerung“ beinhaltet also immer einen Zeitverlust, erhöht jedoch die Wahrscheinlichkeit auch sehr kleine Fehler zu finden.

Arbeitsleistung

Auf der linken Seite des Fließbandbildschirms (siehe Abb. 12 auf S.53) ist eine Grafik zu sehen, welche der Testperson als Orientierung für die eigene Leistungseinschätzung dienen soll. Hier kann die Testperson sehen, welche Leistung als unter-, über- oder etwa durchschnittlich anzusehen ist. Zusätzlich sind unterhalb dieser Grafik, aktuelle Eigenleistung an produzierten Tafeln sowie die Anzahl der aussortierten Tafeln zu sehen. Während des Trainings sortiert die Testperson 20 Schokoladentafeln und erhält für jede getroffene Entscheidung eine Rückmeldung, über die Korrektheit der getroffenen Entscheidung. Die ersten zehn Tafeln sind ohne zeitliche Begrenzung, als auch ohne Vergrößerungsfunktion. In dieser Phase kann sich die Testpersonen, an die Position und Handhabung der einzelnen Schaltflächen gewöhnen. Ab der zehnten Tafel, ist es möglich, die Schaltfläche zur Vergrößerung zu aktivieren und die zeitliche Begrenzung tritt in Kraft. Nach der zwanzigsten Tafel ist dieser Teil der Trainingsphase erfüllt und die Testperson wechselt automatisch zurück in die Ego-Perspektive, um mit dem Training fortzufahren.

5.4.6 Abholbereich

Nachdem das Training am Fließband erfolgreich absolviert wurde, wird die Testperson angehalten, weiter den grünen Markierungen auf dem Weg zu folgen, welcher durch den Rest der Produktionshalle in das Büro führt.



Abbildung 13: Instruktion für den Abholbereich

Auf dem Weg in das Büro passiert die Testperson den sogenannten Abholbereich. In diesem Bereich werden Kundebestellungen zur Abholung bereitgestellt. Im Laufe des Probearbeitstages wird der Proband den Auftrag erhalten, eine Bestellung für einen Kunden vorzubereiten und die bestellten Kartons in diesem Bereich zu platzieren. Im Abholbereich befindet sich ein Monitor, welcher die aktuelle Uhrzeit sowie eintreffende Bestellungen anzeigt. Beim Passieren des Abholbereichs bestätigt die Testperson wieder die gelesenen Informationen und geht anschließend weiter in das Büro.

5.4.7 Das Büro

Das Büro dient als Ausgangspunkt für alle Aufgaben, die von der Testperson erledigt werden sollen. Ein Computer, der sich im Büro befindet, kann von der Testperson angeklickt werden und wechselt die Perspektive von der Ego-Perspektive in eine 2D-Perspektive, welche den Bildschirm des Computers darstellt.



Abbildung 14: virtueller PC im Büro”

Am Computer erfährt der Proband alle Aufgaben für den Tag und soll diese dort auch als erledigt quittieren. Drei Aufgaben sollen an diesem Tag erledigt werden. Von diesen drei Aufgaben muss zunächst die erste Aufgabe erledigt und quittiert werden, bevor die beiden weiteren Aufgaben bearbeitet werden können.

5.5 Test VICI - Phase 3 (Testphase)

5.5.1 Aufgabe 1

Diese Aufgabe bezieht sich auf die Facette Ordnungsliebe der Dimension Gewissenhaftigkeit. Laut Costa und McCrae (1992) beschreiben sich Personen mit

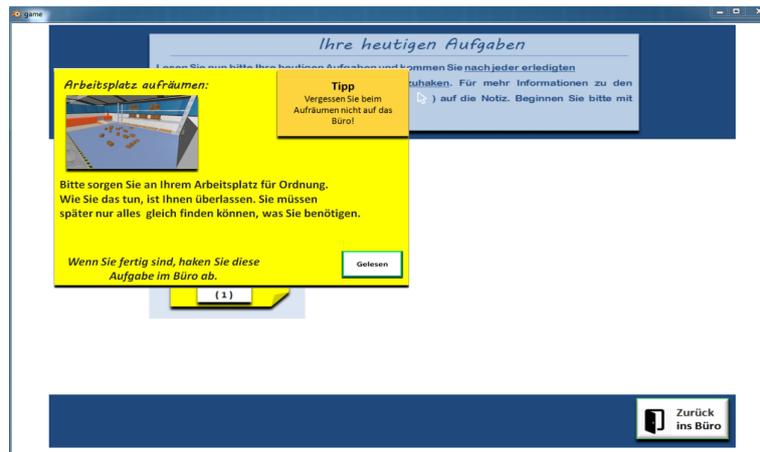


Abbildung 15: Instruktion zur ersten Aufgabe

einer hohen Ausprägung dieser Facette als ordentlich und sehr systematisch, sie bewahren Dinge dort auf, wo sie auch hingehören. Personen mit einer niedrigen Ausprägung sind unfähig, sich die Arbeit einzuteilen und beschreiben sich als wenig planvoll in ihrem Leben (Costa & McCrae, 1992). Nach Westhoff et al. (2013) organisieren ordnungsliebende Personen den eigenen Arbeitsplatz nach einer nachvollziehbaren Ordnung, bei dem alles seinen Platz hat und auf Anhieb zu finden ist. Die erste Aufgabe besteht darin, den gesamten Arbeitsbereich in Ordnung zu bringen. Dabei ist es der Testperson selbst überlassen, was diese unter Ordnung versteht und ab wann sie denkt, diese Aufgabe erfüllt zu haben. Da es schwer zu definieren ist, was Ordnung genau ist und es von Mensch zu Mensch unterschiedlich ist, was als Ordnung anzusehen ist, wird in dieser Aufgabe kein genaues System vorgegeben. Es gibt nur grobe Ziele, welche erreicht werden sollen. Die Instruktion der ersten Aufgabe deutet an, dass es wichtig ist die Kartons innerhalb der Produktionshalle zu sortieren, um diese später schnell wieder finden zu können.



Abbildung 16: Kartons, die verstreut in der Halle liegen

Neben dem Fließband liegen 20 Kartons unsystematisch verstreut. Es können drei Sorten unterschieden werden: Alpenmilch, Erdbeere und Haselnuss. Jede Sorte ist von außen an einem spezifischen Symbol zu erkennen, jedoch ist dieses Symbol unauffällig angebracht. Das macht es dem Probanden schwer, von weitem erkennen zu können, welcher Inhalt sich in den Kartons befindet. Weiter unterscheiden sich die jeweiligen Kartons einer Sorte in ihrer Füllmenge. Es gibt Füllmengen zu 50, 100 und 150 Tafeln. Die erste Aufgabe besteht im Wesentlichen darin, die ungeordneten Kartons in der Produktionshalle nach einem frei wählbaren System zu ordnen. Es sollte, nachdem die Testperson aufgeräumt hat, möglich sein einen Karton gezielt zu finden. Da jeder Mensch eine eigene Vorstellung von Ordnung hat, ist es hier interessanter kein bestimmtes Schema vorzuschreiben und stattdessen zu beobachten, ob das gewählte System zum Ziel führt. Ein erfolgreiches Ordnungssystem führt dazu, dass der Proband immer genau weiß, wo sich etwas befindet und nicht erst danach suchen muss. Bei einer eingehenden Bestellung werden also nur jene Kartons gehoben, die auf der Bestellung aufgelistet sind, wenn das System erfolgreich war.



Abbildung 17: Abfall, neben statt im Abfalleimer

Während des Durchschreitens der virtuellen Halle sollte der Testperson auch auffallen, dass an einigen Stellen Abfall herumliegt. Dieser Abfall sollte erkannt und angeklickt werden. Wenn dieser angeklickt wird, verschwindet der Abfall. Wenn der Proband der Meinung ist, diese Aufgabe erfüllt zu haben, hakt sie diese im Büro ab und kann anschließend die nächsten beiden Aufgaben lesen.

5.5.2 Aufgabe 2

Die zweite Aufgabe bezieht sich auf das Produzieren von Schokolade. Es soll so viel Schokolade wie möglich produziert werden. Als Mindestanzahl werden 200 Tafeln vorgeschlagen, ob diese Zahl eingehalten wird oder nicht bleibt der Testperson überlassen. Die Bearbeitung dieser Aufgabe erfolgt, wie im Abschnitt 5.4.5 erläutert, und kann so oft unterbrochen werden, wie es die Testperson für nötig hält.

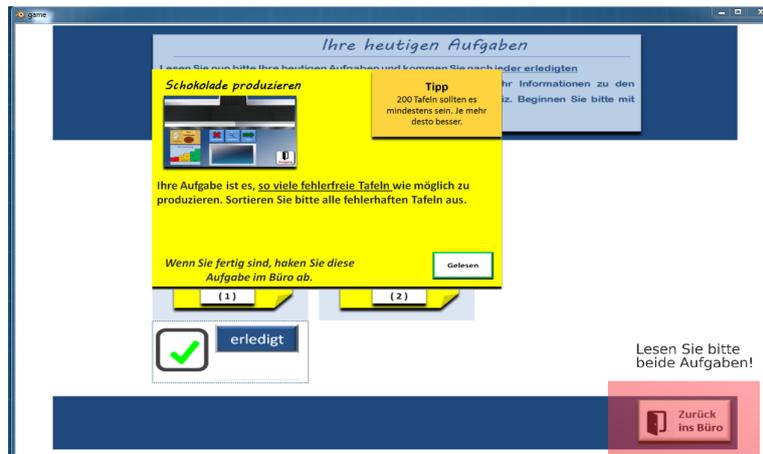


Abbildung 18: Instruktion zur zweite Aufgabe

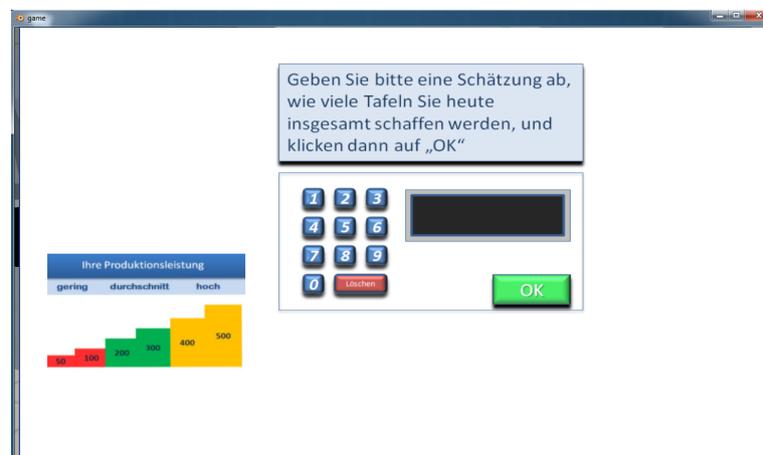


Abbildung 19: Schätzung der Tagesproduktion

Bei erstmaliger Arbeit am Fließband nach der Trainingsphase, wird die Testperson gebeten, eine Schätzung über die persönliche Tagesleistung abzugeben. Neben einem Eingabefeld indem eine beliebige Zahl angegeben werden kann, befindet sich eine Grafik welche als Bezugsrahmen für die persönliche Einschätzung dient. Es bleibt dem Probanden selbst überlassen, ob sie die eigene Einschätzung erfüllt oder nicht. Aus einer vorangegangenen Usability-Studien ist bekannt, dass die Anstrengung dieser Aufgabe nach dem Training

meist unterschätzt wird, was dazu führt, dass es bei gleichzeitigem Anspruch sehr genau zu arbeiten, durchaus reizvoll sein kann das selbst gesteckte Ziel nicht zu erfüllen. Diese Aufgabe nimmt Bezug auf den Arbeitsstil der Testperson (Impulsiv vs. Reflexiv), welcher am ehesten mit den Facetten, Besonnenheit, Leistungsstreben der Dimension Gewissenhaftigkeit im Fragebogen NEO PI-R (Costa & McCrae, 1992) in Verbindung gebracht werden kann

5.5.3 Aufgabe 3

Inhalt der dritten und letzte Aufgabe ist eine Kundenbestellung, welche durch die Testperson bearbeitet werden soll. Die Testperson erfährt, dass um eine bestimmte Uhrzeit eine Bestellung eintreffen wird, welche zu einem zweiten Termin durch ein Transportunternehmen abgeholt wird. Da noch keine Bestellung eingetroffen ist, weiß die Testperson auch noch nicht welche Kartons benötigt werden.



Abbildung 20: Instruktion zur dritte Aufgabe

Die Testperson hat also die Wahl auf die Bestellung zu warten, oder schon mit der Fließbandaufgabe zu beginnen. Während der gesamten Zeit besteht die Möglichkeit, die Uhrzeit im Blick zu behalten. In der Halle wird die aktuelle Uhrzeit auf einem Bildschirm über dem Abholbereich angezeigt und am Fließband ist die Uhrzeit auf dem Bildschirm für Vergrößerungen zu sehen. Wenn die Bestellung eingetroffen ist, erscheint auf dem Bildschirm über dem Abholbereich eine Liste mit den bestellten Kartons. Diese sollten anschließend im dafür vorgesehenen Abholbereich abgestellt werden. Nachdem die Bestellung im Abholbereich abgeliefert wurde, sollte diese Aufgabe am Computer im Büro quittiert werden. Die Testperson kann selbst entscheiden, ob Sie nach dieser Aufgabe noch weiter Schokolade produzieren, oder den Test beenden möchte. Um den Test zu beenden, muss die Testperson durch die Tür in der Haupthalle mit der Aufschrift „Ausgang“ gehen. Der Test kann jeder Zeit auf diese Weise beendet werden.

6 Testkennwerte (VICI)

6.1 AHK-Screening (AHKS)

Der Wert für die Variable AHKS besteht aus der Anzahl an Fehler im Auge-Hand Koordination Screening des Test VICI. Für die Auswertung wird angenommen, dass eine niedrige Fehlerzahl für gute Fähigkeit zur Auge-Hand Koordination steht.

6.2 Kennwerte der Aufgabe 1 (Aufräumen)

Summe gefundenen Abfalls (Trash-Score)

Der Trash-Score wird gebildet aus dem in der Halle gefundenen Abfall. Pro gefundenem Item wird ein Punkt vergeben. Insgesamt können maximal 6.5 Punkte erreicht werden. Die volle Punktzahl ist auch vom Zeitpunkt des Findens abhängig. Wenn das Item während der laufenden Aufgabe 1 aufgehoben wird, zählt dies als ein Punkt. Für Items, welche nach der ersten Aufgabe aufgehoben wurden, wird ein halber Punkt vergeben. Eine detailliertere Aufschlüsselung befindet sich hierzu im Anhang ab Seite 128.

BoxSystem-Score

Die Art der Gruppierung wurde anhand eines Kategoriensystem nach vier Typen unterteilt, „*Gruppirt nach Stückzahl*“, „*Gruppirt nach Sorte*“, „*Gruppirt nach Stückzahl & Sorte*“ sowie „*Unsortiert*“. Unsortiert wurde mit null Punkten gewertet und jede Art der Sortierung erhielt einen jeweils einen Punkt. Welches Art der Sortierung gewählt wurde, wird in der Variable *System* festgehalten. Zudem wurden weitere Punkte für die Güte des Ordnungssystems verge-

ben. Für eine saubere (klare Abgrenzung zu anderen Gruppen) und fehlerfreie Gruppierung wurden zwei Punkte vergeben für eine klare, jedoch nicht fehlerfreie Gruppierung ein Punkt. Wenn eine klares System zu erkennen war, wurde ein weitere Punkt vergeben. Wenn beispielsweise, die Kartons nach Sorte und Stückzahl sortiert waren, jedoch die Kartongruppen nicht nachvollziehbar in der Halle verteilt waren, wurde kein extra Punkt für das System gezählt. Im Gegenzug, wenn neben der reinen Gruppenbildung eine Art zusätzliches System zu erkennen war (Stückzahl aufsteigen, eine wiederkehrende Sortenabfolge, oder lineare Anordnung), so wurde ein weiterer Punkt vergeben.

Ordnungsliebe-Score

Der Ordnungsliebe-Score besteht zum einen aus der Summe gefunden Abfalls (Trash-Score) und zum anderen aus der Art und Weise wie die Kartons in der Halle sortiert wurden zusammen. (BoxSystem-Score). Die Kombination dieser beide Scores wird im folgenden als Ordnungsliebe-Score bezeichnet.

Kartonbewegung zum Zeitpunkt T1

Die Anzahl an bewegten Kartons während der Aufgabe 1.

6.3 Kennwerte der Aufgabe 2 (Fließband)

Geschätzt

Als Wert für die Variable „Geschätzt“ wird die geschätzte Tagesleistung an produzierten Schokoladentafeln bezeichnet.

Anspruchsniveau

Das Anspruchsniveau bildet eine Kategorisierung der getroffenen Selbsteinschätzung aus dem Wert „Geschätzt“.

Produziert

Die Summe aller als fehlerfrei deklarierten Tafeln bildet den Wert für die Variable „Produziert“

Zieldiskrepanz

Als Zieldiskrepanz wird die absolute Differenz zwischen der geschätzten Tagesleistung am Fließband und der tatsächlich erbrachten Leistung bezeichnet.

Prozentuale Mehrarbeit (MehrPR)

Der Wert „MehrPR“ bezeichnet die prozentuale Mehrarbeit am Fließband, bezogen auf die geschätzte Produktionsleistung.

Entscheidungsfreude

Für den Wert „Entscheidungsfreude“ wurde die Anzahl an getroffenen Entscheidungen am Fließband innerhalb der ersten 60 Sekunden verwendet.

Exaktheit (Exaktheit100)

Die Variable „Exaktheit 100“ bildet die Anzahl richtig getroffenen Entscheidungen während der ersten 100 Tafeln, dividiert durch 100.

Exaktheit (ExaktheitAll)

Die Variable „ExaktheitAll“ bildet die Anzahl richtig getroffenen Entscheidungen insgesamt dividiert durch die Anzahl an getroffenen Entscheidungen.

Exaktheit-Score

Der „Exaktheits-Score“ stellt eine gewichtete Form der Variable „Exaktheit100“ dar. Die durchschnittlich benötigte Betrachtungsdauer einer Tafel, bis zu einer korrekten Aussortierung dient als Schätzung der Schwierigkeit den jeweiligen Fehler zu identifizieren und korrekt zu handeln. Als Gewicht wurde die Anzahl an Sekunden bis zu einer korrekten Reaktion verwendet. Summe richtiger Antworten der ersten 100 Tafeln nach Schwierigkeit gewichtet bildet somit den „Exaktheit-Score“.

Arbeitsgeschwindigkeit (Time100)

Die durchschnittlich benötigte Zeit bis zu einer getroffene Entscheidungen während der ersten 100 Tafeln bildet das Maß für die Arbeitsgeschwindigkeit. Dabei ist unerheblich, ob die getroffene Entscheidungen richtig oder falsch war.

Arbeitsgeschwindigkeit (TimeAll)

Die durchschnittlich benötigte Zeit bis zu einer getroffene Entscheidungen über alle Tafeln hinweg bildet das Maß für die Arbeitsgeschwindigkeit. Dabei ist unerheblich, ob die getroffene Entscheidungen richtig oder falsch war.

Impulsivität/Reflexivität (IR-Score)

Impulsivität bzw. Reflexivität berechnet sich aus dem Grad an Exaktheit und der Arbeitsgeschwindigkeit einer Person bezogen auf die ersten 100 Tafeln. Hierfür wurden die Prozenträge für Exaktheit und Arbeitsgeschwindigkeit gebildet um eine Vergleichbarkeit der Variablen zu gewährleisten. Hohe Prozenträge der Variable Time100 bedeuteten, dass die Testperson eher viel Zeit benötigt hat um eine Entscheidung zu treffen und niedrige Werte bedeuten, dass relativ rasch eine Entscheidung getroffen wurde. Hohe Prozenträge der Variable Exaktheit100 bedeuten, dass diese Person eher wenig Fehler bei den Entscheidungen gemacht hat, wobei niedrige Prozenträge bedeuten, dass eher viele Fehler am Fließband gemacht wurden. Diese wurden im Anschluss z-transformiert um den Mittelwert der Beiden Variablen auf Null zu setzen. Eine Addition dieser Werte bildet den Score für Impulsivität/Reflexivität. Positive Werte können als Tendenz für einen reflexiven und negative Werte als Tendenz für einen impulsiven Arbeitsstil betrachtet werden. Eine detailliertere Aufschlüsselung befindet sich hierzu im Anhang dieser Arbeit.

Lupe

Wie oft die Lupenfunktion innerhalb der ersten 100 Tafeln Verwendung fand, bildet stellt die Variable „Lupe100“ dar. „LupeAll“ stellt die Häufigkeit der Lupenfunktion über alle getroffene Entscheidungen am Fließband dar.

6.4 Kennwerte der Aufgabe 3 (Bestellung)

Kartonbewegung zum Zeitpunkt T2

Die Anzahl an bewegten Kartons während der Aufgabe 3.

Bestellung richtig

Anzahl an richtig bereitgestellten Kartons

Bestellung falsch

Anzahl an falsch bereitgestellten Kartons

Pünktlichkeit

Gibt an, ob die Bestellten Kartons innerhalb der vorgegeben Zeitspanne in den Abholbereich abgelegt wurden oder nicht.

7 Weitere verwendete Untersuchungsinstrumente

7.1 NEO- Persönlichkeitsinventar (NEO PI-R)

Im Rahmen der Konstruktvalidierung des VICI wurden 48 Items der Dimension Gewissenhaftigkeit, in revidierter deutschsprachige Fassung, des NEO-Persönlichkeitsinventar (NEO PI-R) nach Costa und McCrae von Ostendorf und Angleitner (2004) vorgegeben. Der NEO PI-R ist eine Persönlichkeits- Fragebogenbatterie zur Erfassung der Dimensionen Neurotizismus, Extraversion, Offenheit für Erfahrungen, Verträglichkeit und Gewissenhaftigkeit (Berth, Goldschmidt, Ostendorf & Angleitner, 2006). Die Dimension Gewissenhaftigkeit besteht aus den sechs Facetten Kompetenz, Ordnungsliebe, Pflichtbewusstsein, Leistungsstreben, Selbstdisziplin und Besonnenheit und wird von Berth et al. (2006) definiert als dem „...aktiven Prozess des Planens, der Organisation und Ausführung von Aufgaben. Personen mit hohen Merkmalsausprägungen sind demnach zielstrebig und willensstark, leistungs- und pflichtbewusst, genau,

pünktlich und zuverlässig“ (S.95). Die vorgegebenen 48 Items sind auf einer fünf-stufigen Likertskala zu beantworten, welche sich zwischen starker Ablehnung und starker Zustimmung erstreckt. Im Manual des NEO PI-R werden zur besseren Interpretierbarkeit für jede Facette typische Adjektive angeführt. Die Reliabilität der Dimension Gewissenhaftigkeit wird mit Cronbach Alpha $\alpha = .90$ angegeben (Berth et al., 2006). In der voliergenden Stichprobe konnte für die Dimension Gewissenhaftigkeit ein Cronbach Alpha von $\alpha = .85$ erzielt werden.

7.2 Räumliches Orientierungsvermögen (3D)

Der Test 3D (Schuhfried, 2004) ist ein sprachfreier, computerisierter Leistungstest und dient der Erfassung des räumlichen Orientierungsvermögens. Dieses Konstrukt kommt am ehesten dem Faktor „Space“ in Thurstone’s Generalfaktortheorie gleich und könnte im Deutschen als Fähigkeit zur „räumlicher Wahrnehmung“ bezeichnet werden. Das räumliche Wahrnehmungsvermögen kann grundsätzlich in zwei Bereich unterteilt werden. Zum einen in das räumliche Vorstellungsvermögen und zum anderen in das räumliche Orientierungsvermögen. Als räumliches Vorstellungsvermögen wird die Fähigkeit, zwei und dreidimensionale Formen und Figuren auffassen und manipulieren zu können verstanden. Das räumliche Orientierungsvermögen hingegen beschreibt die Fähigkeit, die Umwelt von einer neuen Position aus und einer neuen Perspektive auffassen zu können (Schuhfried, 2004). Beim Test 3D werden der Testperson aus Bausteinen zusammengefügte Figuren präsentiert, bei denen sich der Proband vorstellen soll, wie diese Anordnung aus einer anderen Perspektive aussehen würde. Der gewünschte Blickwinkel wird hierbei mit einem Pfeil gekennzeichnet und variiert bei den unterschiedlichen Anordnungen.

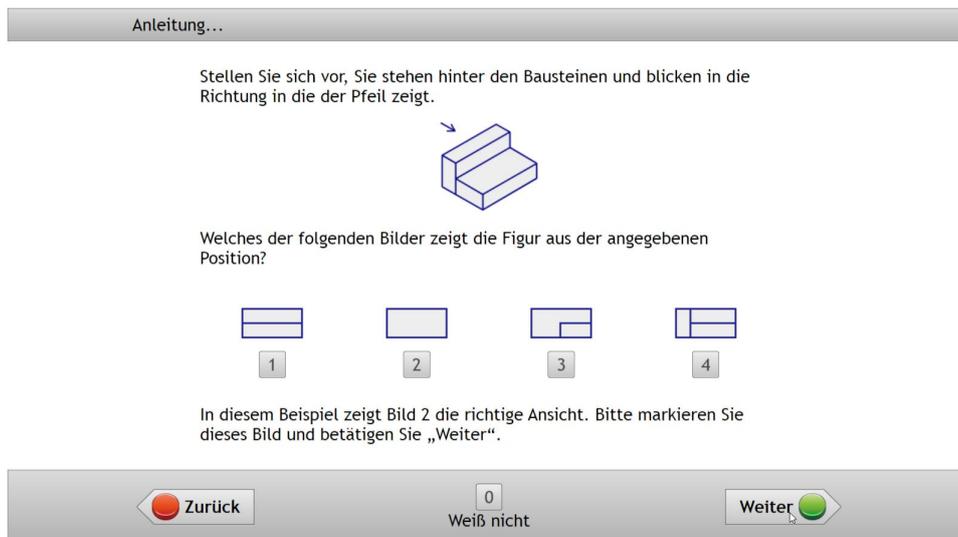


Abbildung 21: Test 3D (Schuhfried, 2004)

Die Testperson soll nun aus vier verschiedenen Antwortalternativen, jenes Bild auswählen, dass zu der vorgestellten Perspektive am ehesten darstellt. Insgesamt hat der Proband drei Minuten Zeit und kann maximal 30 unterschiedlichen Anordnungen bearbeiten. Für die hier durchgeführte empirische Studie wurde die Fähigkeit zum räumlichen Vorstellungsvermögen als besonders wichtig erachtet, da diese eine Grundlage darstellt, um sich in einer virtuellen Umwelt zurecht finden zu können (Schuhfried, 2004). Die Split-Half-Reliabilität der verwendeten Testform S1 mit Zeitlimit beträgt $.87$ (Schuhfried, 2004). Aufgrund unvollständiger Exporte aus dem Wiener Testsystem 6 konnten keine Reliabilitäten für die erhobene Stichprobe berechnet werden.

7.3 Arbeitshaltungen (AHA)

Die Persönlichkeitstestbatterie Arbeitshaltungen (Kubinger & Ebenhöf, 2002) wurde ist ein objektiver Persönlichkeitstest sensu R.B.Cattell und besteht aus drei Untertests, die zum einen den kognitiven Stil Impulsivität/Reflexivität, als auch motivationspsychologischen Konstrukte wie Anspruchsniveau, Leistungsmotivation und Frustrationstoleranz erfassen (Kubinger & Ebenhöf, 2002). Die Berechnung der Reliabilität ist, den Autoren folgend, bei den vorhandenen Variablen der einzelnen Untertests nicht möglich bzw. nicht sinnvoll (Schuhfried, 2014). Im Folgenden werden die drei Untertests inhaltlich kurz vorgestellt.

Untertest „Flächengrößen vergleichen“

Dieser Untertest besteht aus 20 Items auf welchen jeweils zwei Bilder mit unterschiedlichen Flächen zu sehen sind. Der Proband hat die Aufgabe, immer die größere der beiden Flächen anzuklicken und innerhalb von 30 Sekunden so viele Bildpaare wie möglich zu bewerten. Kann der Proband sich für keine Fläche entscheiden, gibt es die Möglichkeit dies mittels Klick auf einen Button mit der Aufschrift „Keine Entscheidung“ zu kennzeichnen.

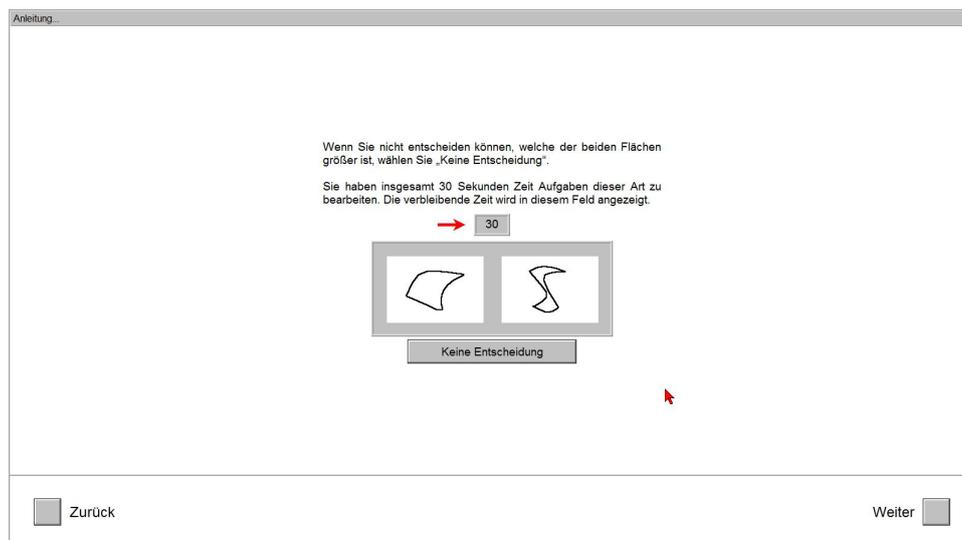


Abbildung 22: Untertest “Flächengrößen vergleichen” (Kubinger & Ebenhö, 2002)

Erfasst wird in diesem Untertest das kognitive Konstrukt Impulsivität/ Reflexivität. Es kann also zwischen Personen unterschieden werden, die „in Problemsituationen entweder langsam und fehlerarm oder schnell und fehlerreich [...] arbeiten“ (Kubinger & Ebenhö, 2002, S. 5), wobei dies nicht als Wertung zu verstehen ist. Zusätzlich werden noch die Testkennwerte Exaktheit und Entschlussfreudigkeit erhoben. Die Anzahl richtiger Antworten kann hierbei als Maß für Exaktheit verstanden werden und die Anzahl gegeben Antworten gibt Aufschluss über die Entschlussfreudigkeit eines Probanden. (Kubinger & Ebenhö, 2002)

Untertest „Symbole Kodieren”

Die Aufgabe des Probanden bei diesem Untertest ist es, abstrakte, farbige Symbole den ebenfalls abstrakten, schwarz-weißen Figuren durch Mausklick zuzuordnen. Die Aufgabe wird in fünf Abschnitte unterteilt, wobei dem Probanden 50 Sekunden pro Durchgang zur Verfügung stehen. Am Ende der Abschnit-

te eins bis vier erscheint jeweils eine Rückmeldung über die Anzahl der richtig kodierten Symbole. Gleichzeitig soll eine Prognose über die im nächsten Durchgang erwartete Leistung abgegeben werden. Ab dem dritten Durchgang wird der Proband zudem noch einer Frustrationsbedingung ausgesetzt, indem fälschlich suggeriert wird, dass andere Personen an diesem Punkt des Tests durchschnittlich höhere Leistung erzielen. Dieser Untertest misst die motivationspsychologischen Konstrukte Anspruchsniveau und Frustrationstoleranz. Das Anspruchsniveau wird vom Probanden selbst gesetzt und über die einzelnen Durchgänge hinweg mit den real erbrachten Leistungen verglichen. Dadurch erhält man Auskunft über das erfolgsaufsuchende bzw. misserfolgsvermeidende Motiv eines Probanden.

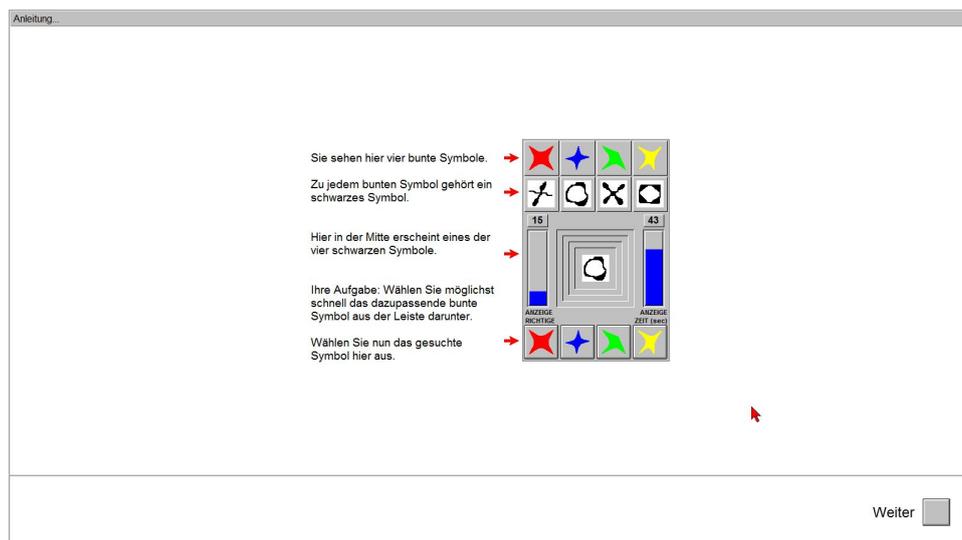


Abbildung 23: Untertest “Symbole kodieren” (Kubinger & Ebenhöf, 2002)

Das Konstrukt Frustrationstoleranz wird über die Differenz zwischen erster und letzter Prognose und der erbrachten Leistung ermittelt. Interessant ist hierbei die Art und Weise, wie sich die ständigen negativen Rückmeldungen auf den Probanden auswirken. Darüber hinaus werden zusätzlich die Testkennwerte Leistungsniveau, Zieldiskrepanz und der Zeitpunkt des Leistungsmaximums erhoben. Das Leistungsniveau wird durch die Anzahl richtiger Kodierungen im zweiten Durchgang wiedergegeben. Die Zieldiskrepanz ergibt sich aus Differenz zwischen real erbrachter und geschätzter Leistung im Kodieren. Als Zeitpunkt des Leistungsmaximums wird jenen Durchgang gewertet, in welchem der Proband die höchste Anzahl richtiger Kodierungen erzielte. (Kubinger & Ebenhöf, 2002)

Untertest „Figuren unterscheiden“

Der dritte Untertest „Figuren unterscheiden“ dient der Erfassung der Leistungsmotivation. Die Testperson soll hierbei aus jeweils vier sehr einfachen geometrischen Figuren diejenige zu identifizieren, welche sich gegenüber den übrigen drei Figuren deutlich unterscheidet. Nach insgesamt 170 bearbeiteten Items erhält die Testperson eine Rückmeldung, darüber, dass sich zwar „recht erfolgreich“ sei und sie jeder Zeit den Test abbrechen könne, jedoch andere Personen an dieser Stelle noch weiter gearbeitet hätten. Nach 425 Items wird die Testperson noch einmal daran erinnert, dass sie den Test jeder Zeit abbrechen könne. Gleichzeitig wird ihr aber auch mitgeteilt, dass es immer Personen gibt, die noch weiter arbeiten. Falls die Testperson mindestens 150 Items bearbeitet und die Dauer zwischen zwei Bearbeitungsschritten größer ist, als der Mittelwert plus dreimal der Standardabweichung aller vorhergehenden Bearbeitungszeiten folgt eine Meldung, welche der Testperson einen Konzentrationsabfall signalisiert und diese daran erinnert, dass sie den Test beenden könne, aber wieder

mit der Information, andere Personen hätten an dieser Stelle noch weitergearbeitet. Die Testvorgabe erfolgt so lange, bis die Testperson die Bearbeitung von sich aus beendet, falls diese dies nicht tut ist die Testdauer jedoch insgesamt auf 30 Minuten begrenzt. Dieser Untertest zielt darauf ab, die Leistungsmotivation dadurch zu erfassen, wie ausdauernd eine Person an einer Aufgabe festhält und diese intrinsisch motiviert weiterführt. (Kubinger & Ebenhöf, 2002)

8 Fragestellungen

8.1 Hypothesen

1. Fragestellung: Unterscheiden sich die Ergebnisse des Tests VICI (Ordnungsliebe, Exaktheit, Impulsivität/Reflexivität und Trash-Score) aufgrund der 3D-Gaming Erfahrung der Teilnehmer?
 - 1.1. H1: Die Werte der Dimension Ordnungsliebe im VICI der Teilnehmer mit 3D-Gaming Erfahrung unterscheiden sich von den jenen Teilnehmern, die keine 3D-Spiele Erfahrung haben.
 - 1.2. H1: Die Werte der Dimension Exaktheit im VICI der Teilnehmer mit 3D-Gaming Erfahrung unterscheiden sich von den jenen Teilnehmern, die keine 3D-Spiele Erfahrung haben.
 - 1.3. H1: Die Werte der Dimension Impulsivität/Reflexivität im VICI der Teilnehmer mit 3D-Gaming Erfahrung unterscheiden sich von den jenen Teilnehmern, die keine 3D-Spiele Erfahrung haben.
 - 1.4. H1: Die Werte der Dimension Trash-Score im VICI der Teilnehmer mit 3D-Gaming Erfahrung unterscheiden sich von den jenen Teilnehmern, die keine 3D-Spiele Erfahrung haben.

2. Fragestellung: Unterscheiden sich die Ergebnisse des Tests VICI (Ordnungsliebe, Exaktheit, Impulsivität/Reflexivität und Trash-Score) aufgrund der Augen-Hand-Koordination (AHKS) der Teilnehmer?
 - 2.1. H1: Die Werte der Dimension Ordnungsliebe im VICI der Teilnehmer mit wenigen Fehlern im AHKS unterscheiden sich von den jenen Teilnehmern, die mehrere Fehler im AHKS gemacht haben.
 - 2.2. H1: Die Werte der Dimension Exaktheit im VICI der Teilnehmer mit wenigen Fehlern im AHKS unterscheiden sich von den jenen Teilnehmern, die mehrere Fehler im AHKS gemacht haben.
 - 2.3. H1: Die Werte der Dimension Impulsivität/Reflexivität im VICI der Teilnehmer mit wenigen Fehlern im AHKS unterscheiden sich von den jenen Teilnehmern, die mehrere Fehler im AHK-Test gemacht haben.
 - 2.4. H1: Die Werte der Dimension Trash-Score im VICI der Teilnehmer mit wenigen Fehlern im AHKS unterscheiden sich von den jenen Teilnehmern, die mehrere Fehler im AHKS gemacht haben.

3. Fragestellung: Unterscheiden sich die Ergebnisse des Tests VICI (Ordnungsliebe, Exaktheit, Impulsivität/Reflexivität und Trash-Score) aufgrund des Geschlechts der Teilnehmer?
 - 3.1. H1: Die Werte der Dimension Ordnungsliebe im VICI unterscheiden sich zwischen Männern und Frauen.
 - 3.2. H1: Die Werte der Dimension Exaktheit im VICI unterscheiden sich zwischen Männern und Frauen.
 - 3.3. H1: Die Werte der Dimension Impulsivität/Reflexivität im VICI unterscheiden sich zwischen Männern und Frauen.

- 3.4. H1: Die Werte der Dimension Trash-Score im VICI unterscheiden sich zwischen Männern und Frauen
4. Fragestellung: Unterscheiden sich die Ergebnisse des Tests VICI (Ordnungsliebe, Exaktheit, Impulsivität/Reflexivität und Trash-Score) aufgrund der Fähigkeit zur räumlichen Orientierungsfähigkeit der Teilnehmer?
 - 4.1. H1: Die Werte der Dimension Ordnungsliebe im VICI der Teilnehmer mit wenigen richtigen Antworten im 3D-Test unterscheiden sich von den jenen Teilnehmern, die mehrere richtigen Antworten im 3D-Test gemacht haben.
 - 4.2. H1: Die Werte der Dimension Exaktheit im VICI der Teilnehmer mit wenigen richtigen Antworten im 3D-Test unterscheiden sich von den jenen Teilnehmern, die mehrere richtigen Antworten im 3D-Test gemacht haben.
 - 4.3. H1: Die Werte der Dimension Impulsivität/Reflexivität im VICI der Teilnehmer mit wenigen richtigen Antworten im 3D-Test unterscheiden sich von den jenen Teilnehmern, die mehrere richtigen Antworten im 3D-Test gemacht haben.
 - 4.4. H1: Die Werte der Dimension Trash-Score im VICI der Teilnehmer mit wenigen richtigen Antworten im 3D-Test unterscheiden sich von den jenen Teilnehmern, die mehrere richtigen Antworten im 3D-Test gemacht haben.
5. Fragestellung: Gibt es einen signifikanten Zusammenhang zwischen den Dimensionen des AHA und des VICI?
 - 5.1. H1: Es gibt einen signifikanten Zusammenhang zwischen den Dimensionen des AHA und des VICI.

6. Fragestellung: Gibt es einen signifikanten Zusammenhang zwischen den Dimensionen des NEO PI-R und des VICI?

6.1. H1: Es gibt einen signifikanten Zusammenhang zwischen den Dimensionen des NEO PI-R und des VICI.

9 Statistische Auswertungsmethoden

Vergleich zweier Gruppenmittelwerte

Um die Fragestellungen 1-3 statistisch prüfen zu können, wurde der t-Test für unabhängige Stichproben ausgewählt. Der t-Test für unabhängige Stichproben hat zum Ziel die Mittelwerte zweier Gruppen zu vergleichen. Um dieses Verfahren einsetzen zu können, müssen bestimmte Voraussetzungen erfüllt werden:

- Die Stichproben müssen voneinander unabhängig sein.
- Die Variablen müssen zumindest ein Intervallskalenniveau aufweisen.
- Beide Gruppen müssen normalverteilt sein.
- Die Homogenität der Varianzen muss gegeben sein.

Sollten eine Normalverteilung innerhalb der Gruppen nicht gegeben sein, wird als Ersatzverfahren der nichtparametrische Mann-Whitney U-Test eingesetzt. Dabei werden nicht wie beim t-Test die Mittelwerte, sondern die Rangreihen der beiden Gruppen verglichen. Sollte eine Homogenität der Varianzen über die Gruppe nicht gegeben sein, wird als Ersatzverfahren ein Welch-Test als nichtparametrische Alternative gerechnet.

Prüfung auf Normalverteilung

Die Normalverteilung der Gruppen wird mit Hilfe des Shapiro-Wilk-Tests geprüft. Fällt der Shapiro-Wilk-Tests nicht signifikant aus ($p > .05$) kann die Normalverteilung der Stichprobe angenommen werden.

Prüfung auf Homogenität der Varianzen

Die Homogenität der Varianzen wird anhand des Levene-Tests überprüft. Im Rahmen dieses Levene-Tests wird für jede abhängige Variable getestet, ob die beiden Gruppen aus einer gemeinsamen Grundgesamtheit mit den gleichen Varianzen stammen. Fällt der Levene-Test nicht signifikant aus ($p > .05$), kann die Homogenität der Varianzen angenommen werden.

10 Beschreibung der Stichprobe

10.1 Beschreibung der Stichprobe bezüglich Geschlechts- & Altersverteilung

Die Stichprobe besteht aus insgesamt 85 Personen, die über die Testerfahrung der Fakultät für Psychologie und im privaten Umfeld rekrutiert wurden. Die Standorte, an denen die Testungen durchgeführt wurden, lagen zum einen in der Computerdiagnostik der Fakultät für Psychologie an der Uni Wien und zum anderen in den Wohnung des Autors. Alle Personen haben freiwillig an der Untersuchung teilgenommen und haben im Anschluss an ihre Teilnahme eine Einverständniserklärung für Weiterverarbeitung ihrer Testergebnisse für Forschungszwecke zugestimmt. Von den insgesamt 85 Datensätzen mussten 2 Datensätze ausgeschlossen werden, da diese auf Grund eines Testabbruchs

nicht auswertbar waren. Von den ursprünglichen Datensätzen konnten 83 zu weiteren Berechnung verwendet werden. Die Geschlechterverteilung lag bei 53 weiblichen und 30 männlichen Versuchspersonen. In Tabelle 2 wird die Stichprobe anhand von Kenngrößen beschrieben.

	N	Mittelwert	Median	Range	Minimum	Maximum	Quantile 25%	Quantile 50%	Quantile 75%
Alter	83	24.2	24	38	16	54	21	24	25

Tabelle 2: Stichprobenbeschreibung bezüglich der Altersverteilung

In Tabelle 2 sind Alterskennwerte in Bezug auf das Geschlecht dargestellt. Das Durchschnittsalter der Männer liegt bei 24 Jahren ($SD = 6.22$). Das Durchschnittsalter bei den Frauen beträgt 24 Jahre ($SD = 5.11$). Der Median liegt bei 23 Jahren. Das Box-Plot Diagramm (siehe Abb. 24) stellt die Altersverteilung grafisch in Bezug auf das Geschlecht dar.

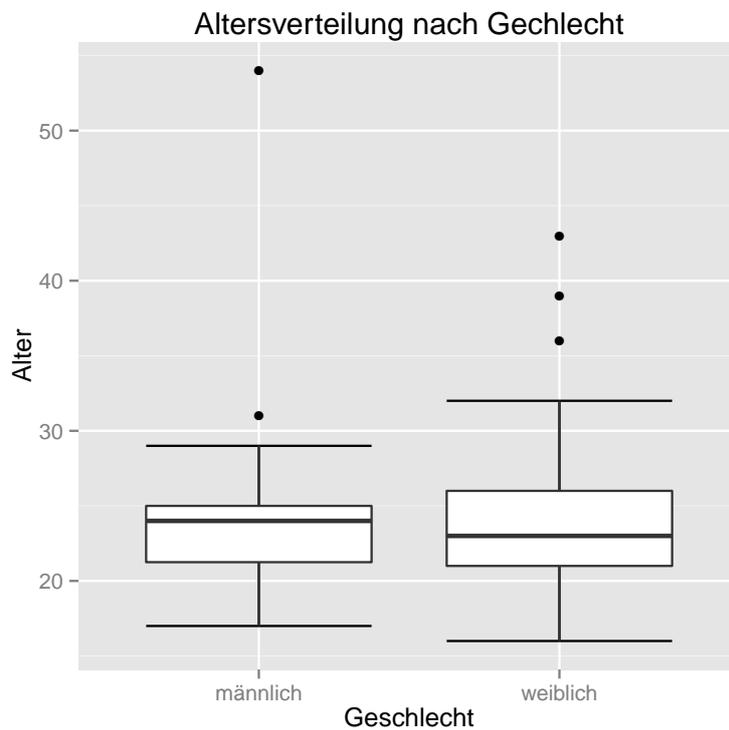


Abbildung 24: Box-Plot: Altersverteilung in Jahren in Bezug auf das Geschlecht

Aus dem Box-Plot-Diagramm ist ersichtlich, dass bei den Männern der Altersbereich zwischen 21.25 und 25 Jahren stark vertreten ist. Bei den Frauen liegt der stark repräsentierte Bereich zwischen 21 und 26 Jahren. In beiden Gruppen sind demnach junge Testpersonen überrepräsentiert.

10.2 Beschreibung der Stichprobe bezüglich des Bildungsgrads

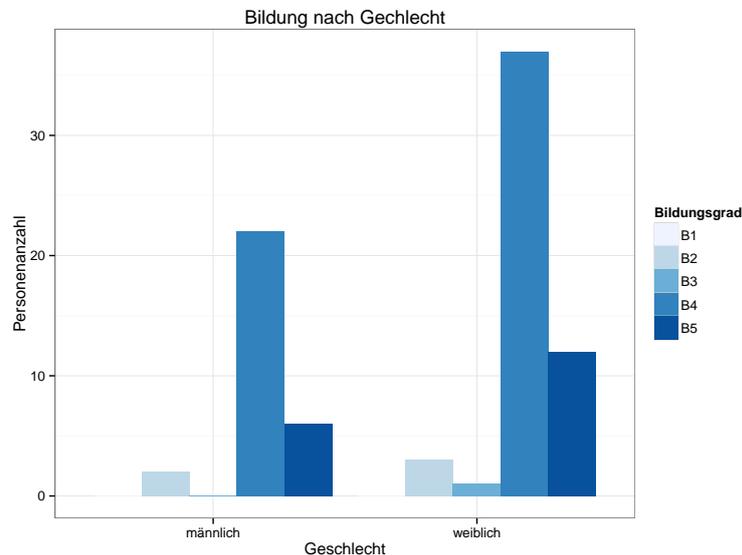


Abbildung 25: Stichprobenbeschreibung bezüglich des Bildungsgrads

Die Stichprobenverteilung bezüglich des Bildungsstatus geht aus der Abbildung 25 hervor. Durch die Unterscheidung in fünf Kategorien sollte das gesamte österreichische Ausbildungsangebot abgedeckt werden. Die Verteilung wird für beide Geschlechter separat dargestellt. Folgender Kodierschlüssel liegt der Abbildung 25 zugrunde:

- (B1) Bildungsgrad 1: kein Schulabschluss oder Sonderschule
- (B2) Bildungsgrad 2: Pflichtschule oder Realschule abgeschlossen
- (B3) Bildungsgrad 3: Fachschule bzw. Berufsausbildung abgeschlossen
- (B4) Bildungsgrad 4: Höhere Schule mit Abitur (Matura) abgeschlossen
- (B5) Bildungsgrad 5: Universitäts- bzw. Hochschulabschluss

Die Zuordnung der Personen zu einer Kategorie erfolgte nur dann, wenn sie bereits über die in der Kategorie dargestellte Ausbildung verfügten. Personen, die zum Zeitpunkt der Testung an der Universität studierten und ihr Studium noch nicht beendet hatten, wurden mit dem Bildungsgrad „Höhere Schule mit Abitur (Matura) abgeschlossen“ kodiert.

Insgesamt 21.7% der Gesamtstichprobe verfügen über einen Universitäts bzw. Hochschulabschluss. Dabei gaben 71% der Gesamtstichprobe als höchste abgeschlossene Ausbildung einen Abiturabschluss an. 8% der Stichprobe verfügen über einem niedrigeren Schulabschluss als Abitur. Aus dieser Stichprobenverteilung wird ersichtlich, dass Personen, die über einen höheren Schulabschluss (Bildungsgrad 4 & 5; siehe Abbildung 25) verfügen, mit 92.8% am stärksten vertreten sind.

11 Ausgewählte Ergebnisse des Tests (VICI)

Bei der Erhebung sind besonders umfangreiche Mengen an Daten zustande gekommen. Um den Rahmen dieser Diplomarbeit nicht zu sprengen, werden im Anschluss nur einige ausgewählte Ergebnisse näher dargestellt.

11.1 Ergebnisse Auge-Hand Koordination (AHKS)

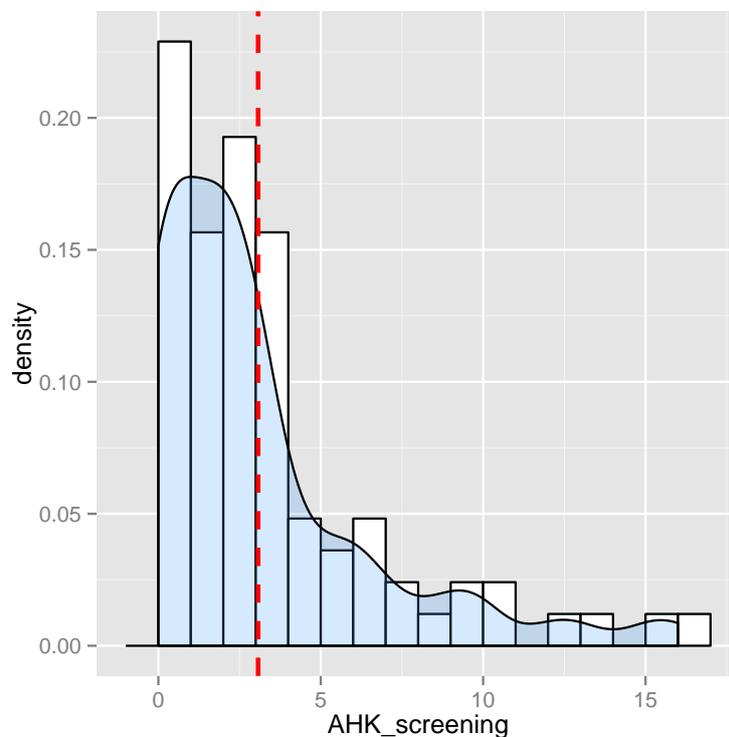


Abbildung 26: Histogramm zur Auge-Hand Koordination

Bei dem in Abschnitt 5.3 beschriebenen Screening zur Auge-Hand Koordination (AHKS) zeigte sich eine durchschnittliche Fehlerhäufigkeit von 3.07 ($Median = 2$, $Min = 0$, $Max = 16$). Wie in Abbildung 26 weiter unten zu sehen ist, handelt es sich um eine links schiefe Verteilung, bei welcher in etwa 58% der Testperso-

nen weniger als 3 Fehler gemacht haben.

11.2 Ergebnisse Aufgabe 1

11.2.1 Ordnungssinn (Abfall)

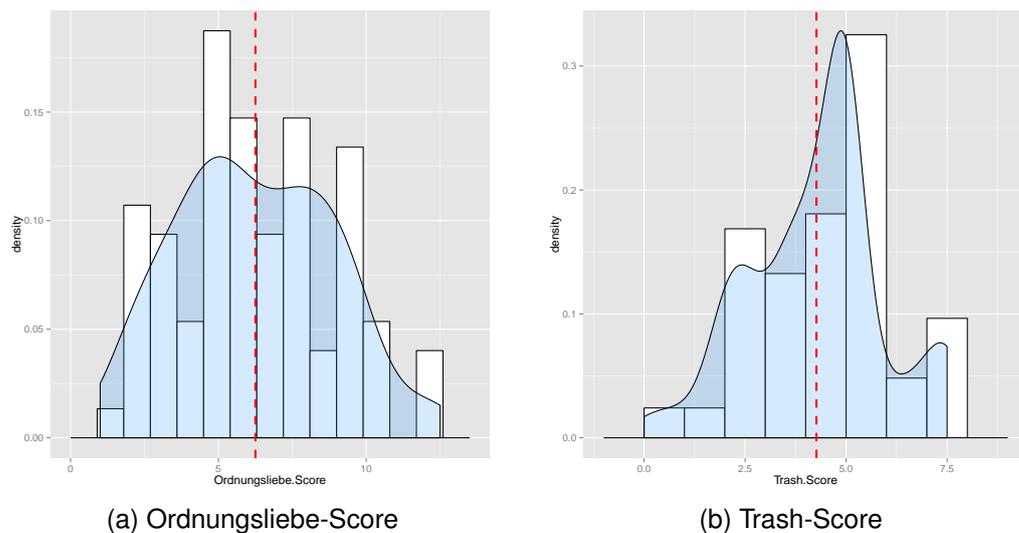


Abbildung 27: Verteilung der wichtigsten Kenngrößen in Aufgabe 1

Wie in Abschnitt 5.5.1 beschrieben, lag die erste Aufgabe für die Testperson darin, für Ordnung am Arbeitsplatz zu sorgen. Nahezu alle Testpersonen 98% bemerkten den Abfall neben dem Abfalleimer und hoben diesen auf. Ebenso verhielt es sich mit der Plastikflasche unter dem Tisch im Büro, welche von 94% der Teilnehmer bemerkt wurde. Etwa 16% der Testpersonen öffneten den Büroschrank und entdeckten den dort liegenden Abfall. Innerhalb der Produktionshalle variieren die Prozentzahlen der zu findenden Objekte von 78 bis 88%. Zusammenfassend wurden, abgesehen von dem im Büroschrank abgelegtem Abfall, von ca. 80% der Teilnehmer alle Objekte gefunden und weggeräumt. Im Durchschnitt wurde so ein Trash-Score von 4.27 ($SD = 1.65$) erreicht.

11.2.2 Ordnungssinn (Kartons)

Bei der Kartonanordnung sortierten 39% (32 Personen) die Kartons ausschließlich nach ihrer Sorte, nur zwei Testpersonen sortierten die Kartons ausschließlich nach ihrer Stückzahl. Eine Kombination aus Sorte und Stückzahl als Ordnungssystem wählten 25% (21 Personen) der Probanden. Annähernd 34% (28 Personen) nahmen keinerlei Gruppierung bei den Kartons vor. Somit kann gesagt werden, dass mit 66%, zwei Drittel der Testpersonen ein Ordnungssystem erstellten und ein Drittel die Kartons unangetastet ließen. Im Durchschnitt wurde so ein Ordnungsliebe-Score von 6.25 ($SD = 2.62$) erreicht.

11.3 Ergebnisse Aufgabe 2

11.3.1 Quantitative Arbeitsleistung (Produktionsleistung)

Von den insgesamt 83 Teilnehmern schätzten 60% (50 Personen) ihre Tagesproduktion vor Beginn der Arbeit im durchschnittlichen Bereich zwischen 100 und 400 Tafeln ein. Anhand des Mittelwerts von 227 ($SD = 110.23$) Tafeln ist zu sehen, dass die in Aufgabe 2 erwünschte Marke von 200 Tafeln vom Großteil der Teilnehmer erfüllt wurde. Etwa 20% (17 Personen) schätzten ihre Produktionsleistung zu Beginn der Arbeit mit 300 oder mehr Tafeln, und somit als überdurchschnittlich ein. Von den Probanden schätzten sich 19% (16 Personen) mit einem Mittelwert von 57.44 ($SD = 43.56$) als unterdurchschnittlich ein. Jedoch produzierten Personen dieser Selbsteinschätzung im Durchschnitt 311 ($SD = 141.18$) Tafeln. Teilnehmer, die ihre Leistung als überdurchschnittlich einschätzten ($M = 373$, $SD = 29.42$) kamen am Ende auf eine reale Produktionsleistung von durchschnittlich 571 ($SD = 224.86$) produzierten Schokoladentafeln. Im mittleren Bereich lagen die Schätzungen bei 235 ($SD = 44.07$) Tafeln und die realisierte Leistung bei durchschnittlich 449 ($SD = 169.26$) Tafeln.

11.3.2 Entscheidungsfreudigkeit

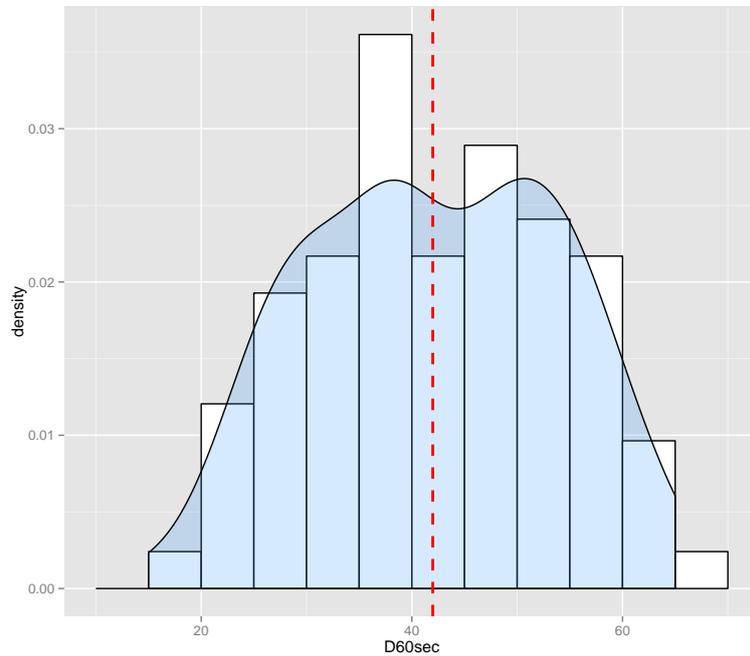


Abbildung 28: Verteilung der Entscheidungsfreudigkeit

Wie in Abschnitt 6.3 beschrieben, stellt die Entscheidungsfreudigkeit die Summe getroffener Entscheidungen innerhalb der ersten sechzig Sekunden dar. Im Durchschnitt wurden 42 ($SD = 11.63$) Tafeln in einer Minute bearbeitet, ($Min = 15$, $Max = 65$). Die Verteilung der Variable Entscheidungsfreudigkeit ist in Abbildung 28 dargestellt.

11.3.3 Qualitative Arbeitsleistung (Exaktheit)

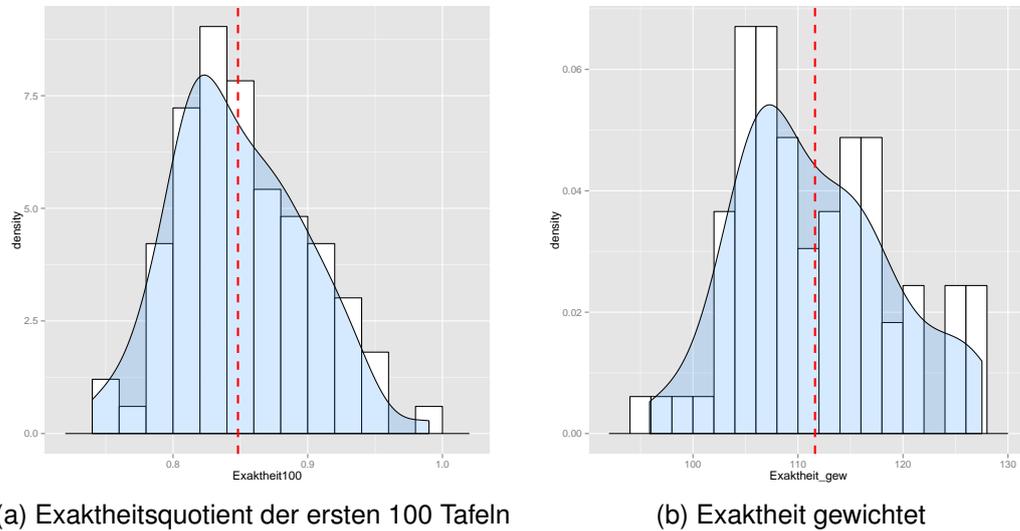


Abbildung 29: Verteilung Aufgabe 2

Die in Abbildung 29a dargestellte Verteilung des Exaktheitsquotienten für die ersten 100 Tafeln zeigt eine eingipflige Verteilung mit einem Mittelwert von 0.85 ($SD = 0.05$). Verglichen mit dem Exaktheitsquotienten über die gesamte Produktion gerechnet, bleibt der Mittelwert konstant bei 0.85 ($SD = 0.06$). Die in Abbildung 29b dargestellte Verteilung des gewichteten Exaktheitsquotienten zeigt eine zweigipflige Verteilung mit einem Mittelwert von 111.62 ($SD = 7.38$). Es zeigte sich, dass der Grad an Exaktheit bei der Aussortierung am Fließband durchweg sehr hoch war, selbst der niedrigste Wert lag noch immer bei 0.74.

11.3.4 Arbeitsstil (Impulsivität / Reflexivität)

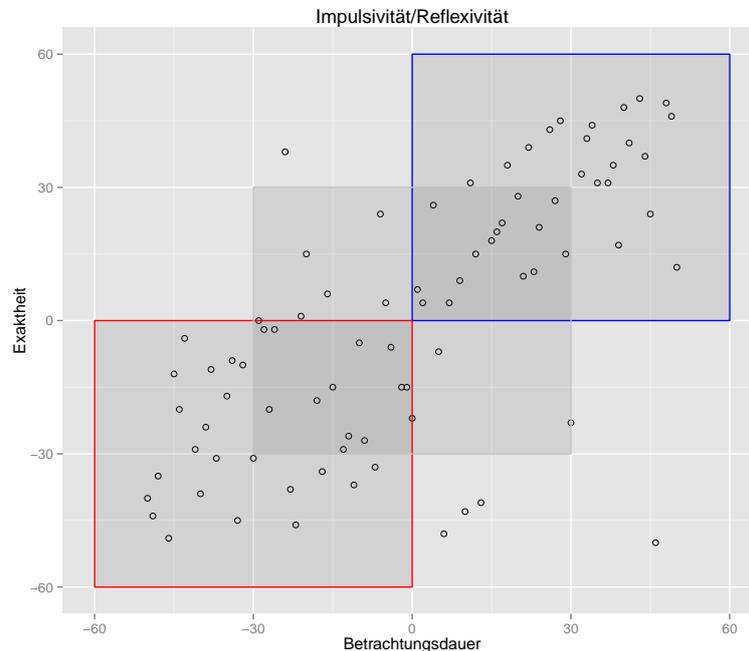


Abbildung 30: Streudiagramm zum Score "Impulsivität-Reflexivität"

Die Abbildung 30 zeigt eine hohe, positive Korrelation nach Pearson, von $r = .73$, $p < .0001$ zwischen der benötigten Betrachtungsdauer (Time100) bis zu einer korrekten Entscheidung am Fließband und dem Exaktheitsquotienten (Exaktheit100), bezogen auf die ersten 100 produzierten Tafeln. Um ein Maß für den Arbeitsstil im Sinne von Impulsivität und Reflexivität zu erhalten, wurden zunächst Prozentränge für jede der Variablen gebildet und diese anschließend z-transformiert. Für die genauere Beschreibung der Transformation der Daten, siehe Abschnitt 6.3. Der hohe Zusammenhang zwischen Betrachtungsdauer und Exaktheit verdeutlicht, dass eine höhere Betrachtungsdauer gleichzeitig auch die Genauigkeit am Fließband steigert. Dieser Arbeitsstil ist in Abbildung 30 als blaues Quadrat in der oberen rechten Ecke gekennzeichnet und kann als reflexiver Stil bezeichnet werden. Bei niedriger Betrachtungsdauer der Tafeln

werden Fehler häufiger übersehen, was auf einen impulsiven Arbeitsstil hindeutet. In der Abbildung 30 ist dieser Bereich mit einem rot umrandeten Quadrat im unteren linken Quadranten gekennzeichnet. Der Übergangsbereich in der Mitte lässt sich keinem der beiden Arbeitsstile zuordnen, es handelt sich hierbei um einen Mischtyp. Des Weiteren können auch Ausreißer nicht eindeutig zugeordnet werden, hohe Werte in der Exaktheit und niedrige Werte in der Betrachtungsdauer können auf eine effiziente Arbeitsweise hindeuten. Personen, deren Betrachtungsdauer hoch und deren Exaktheit niedrig ausfällt würden einem ineffizientem Arbeitsstil zugerechnet werden. Mischtypen und Ausreißer sollten anhand ihrer einzelnen Kennwerte in Exaktheit und Betrachtungsdauer bewertet werden. Sechzehn Teilnehmer hatten zumindest eine teilweise korrekte Bestellung im Abholbereich deponiert.

11.4 Ergebnisse Aufgabe 3

11.4.1 Aufgabenerfüllung

Von den insgesamt 83 Teilnehmern lieferten 47% (39 Personen) drei richtige von drei bestellten Kartons im Abholbereich ab. Mit ca. 19% gaben insgesamt 16 Personen überhaupt keine Kartons ab. Von denjenigen, die ausschließlich die bestellten Kartons ablieferten, hatten 7 Personen zum Zeitpunkt von Aufgabe 1 keine erkennbares Ordnungssystem für die Kartons, 15 Personen hatten die Schachteln zuvor nach den Sorten und 2 Teilnehmer nach der Stückzahl geordnet. Der Großteil derjenigen, die die drei korrekten Boxen in den Abholbereich brachten, sortierte zuvor sowohl nach Stückzahl, als auch nach Sorte.

12 Hypothesenprüfung

12.1 Prüfung der Fragestellung 1

Unter den 83 Teilnehmern gab es keine Person, die angab, über keinerlei Computererfahrung zu verfügen, daher wurde auf eine Auswertung bezüglich der Computererfahrung verzichtet. In Abbildung 31 (siehe S.93) ist die Erfahrung in 3D-Computerspielen nach Geschlecht dargestellt. Von den insgesamt 83 Teilnehmern gaben etwa 48% an, keinerlei Erfahrung mit 3D-Computerspielen zu haben, wovon 85% weiblich waren. Etwa 29% der getesteten Personen gaben an, über ein wenig 3D-Spiele Erfahrung zu verfügen und 23% hatten schon viel Erfahrung bei diesem Typ von Videospiele. Von jene Testpersonen mit viel Erfahrung bestand der Großteil mit 79% aus männlichen Teilnehmern. Für die Überprüfung der Fragestellungen 1 wurde die Stichprobe anhand der Selbstbeschreibung zur 3D-Spiele-Erfahrung in zwei annähernd gleich große Gruppen geteilt. Eine Gruppe mit "wenig und viel" Gaming-Erfahrung (Gruppe 1, $N = 43$) und eine Gruppe ohne jegliche Erfahrung mit 3D-Computerspielen (Gruppe 2, $N = 40$).

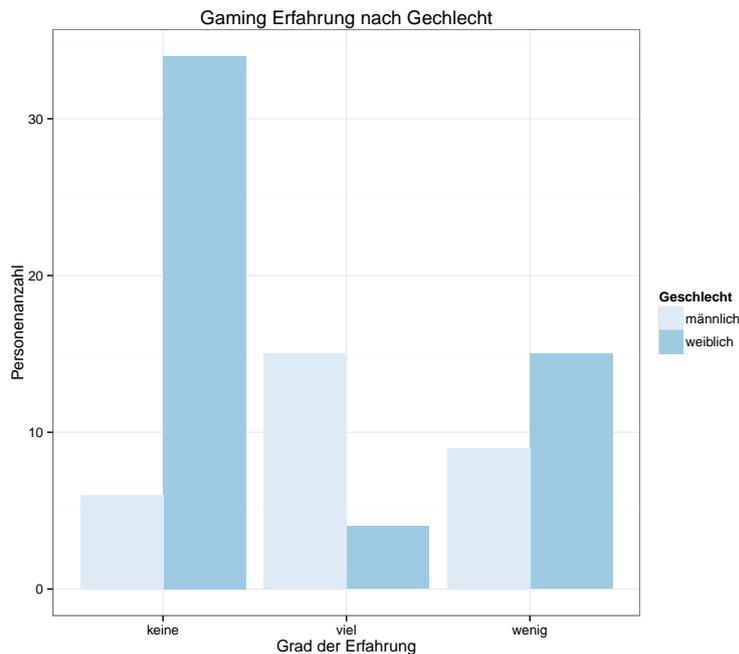


Abbildung 31: Balkendiagramm zum Grad der 3D Spiele- Erfahrung

Ergebnis H1.1 (Game-Ordnungsliebe)

Die Ergebnisse für den Shapiro-Wilk Test zur Überprüfung der Normalverteilung zeigen, sowohl für Gruppe 1 ($W(43) = 0.97, p = .30$), als auch für Gruppe 2 ($W(40) = 0.973, p = .46$), ein nicht signifikantes Ergebnis. Es kann daher für beide Gruppen eine Normalverteilung angenommen werden. Der Levene-Test, $df = 81, F = 0.263, p = .609$, kam ebenfalls zu keinem signifikanten Unterschied bezüglich der Varianzen. Aufgrund aller erfüllten Voraussetzungen wurde ein t-Test für unabhängig Stichproben gerechnet. Die Ergebnisse zeigen, dass es keinen signifikanten Unterschied, $t(83) = -0.549, df = 78.9, p = .585$, zwischen Teilnehmern mit 3D-Gaming Erfahrung und Teilnehmern ohne 3D-Gaming Erfahrung in Bezug auf die Werte der Ordnungsliebe im Test VICI gibt. Es kann daher angenommen werden, dass die Erfahrung in 3D-Computerspielen keine

Auswirkungen auf die Ergebnisse des Test VICI in der Dimension Ordnungsliebe hat.

Ergebnis H1.2 (Game-Exaktheit)

Die Ergebnisse für Shapiro-Wilk Test zur Überprüfung der Normalverteilung zeigen, dass die Verteilung in Gruppe 1 ($W(43) = 0.957, p = .10$) angenommen werden konnte, jedoch in Gruppe 2 konnte die Normalverteilung nicht angenommen werden, ($W(40) = 0.859, p = .00$). Der Levene-Test, $df = 81, F = 0.455, p = .502$, kam zu keinem signifikanten Unterschied zwischen den Varianzen. Aufgrund der Verletzung der Voraussetzung der Normalverteilung in den Gruppen wurde ein Mann-Whitney U-Test durchgeführt. Die Ergebnisse zeigen, dass es keinen signifikanten Unterschied ($W(83) = 685.5, p = .113$) zwischen Teilnehmern mit 3D-Gaming Erfahrung und Teilnehmern ohne 3D-Gaming Erfahrung in Bezug auf die Werte der Exaktheit im Test VICI gibt. Es kann daher angenommen werden, dass die Erfahrung in 3D-Computerspielen keine Auswirkungen auf die Ergebnisse des Test Vici in der Dimension Exaktheit hat.

Ergebnis H1.3 (Game-IR)

Die Ergebnisse für Shapiro-Wilk Test zur Überprüfung der Normalverteilung zeigen, dass die Verteilung in Gruppe 1 ($W(43) = 0.957, p = .10$) angenommen werden konnte, jedoch in Gruppe 2 konnte die Normalverteilung nicht angenommen werden, ($W(40) = 0.859, p = .00$). Der Levene-Test $df = 81, F = 0.455, p = .502$ kam zu keinem signifikanten Unterschied zwischen den Varianzen. Aufgrund der Verletzung der Voraussetzung der Normalverteilung in den Gruppen wurde ein Mann-Whitney U-Test durchgeführt. Die Ergebnisse zeigen, dass es keinen signifikanten Unterschied ($W(83) = 685.5, p = .113$), zwischen Teilnehmern mit 3D-Gaming Erfahrung und Teilnehmern ohne 3D-Gaming Erfah-

rung in Bezug auf die Werte der Impulsivität/Reflexivität im Test VICI gibt. Es kann daher angenommen werden, dass die Erfahrung in 3D-Computerspielen keine Auswirkungen auf die Ergebnisse des Test Vici in der Dimension Impulsivität/Reflexivität hat.

Ergebnis H1.4 (Game-Trash)

Die Ergebnisse für Shapiro-Wilk Test zur Überprüfung der Normalverteilung zeigen, dass zwar in Gruppe 1, ($W(43) = 0.965, p = .20$) die Normalverteilung angenommen werden konnte, jedoch nicht in Gruppe 2, ($W(40) = 0.908, p = .00$). Der Levene-Test fiel nicht signifikant aus, $df = 81, F = 0.336, p = .564$, diesen Voraussetzungen folgend wurde ein Mann-Whitney-U-Test durchgeführt. Die Ergebnisse zeigen, dass es keinen signifikanten Unterschied $W(83) = 1046.5, p = .084$, zwischen Teilnehmern mit 3D-Gaming Erfahrung und Teilnehmern ohne 3D-Gaming Erfahrung in Bezug auf die Werte des Trash-Score im Test VICI gibt. Es kann daher angenommen werden, dass die Erfahrung in 3D-Computerspielen keine Auswirkungen auf die Ergebnisse des Test VICI in der Dimension Trash-Score hat.

Beantwortung der Fragestellungen 1

Ausgehend von den Ergebnissen der einzelnen Unterhypothese kann die Haupthypothese 1 somit verworfen werden, es gibt keine Auswirkungen von 3D-Spiele Erfahrung auf die Ergebnisse im Test VICI.

12.2 Prüfung der Fragestellung 2

Für die folgenden Berechnungen wurde die Stichprobe anhand des vorab durchgeführten Augen-Hand-Koordinationstests in zwei Gruppen geteilt. Im Schnitt

machten die Teilnehmer 3.07 Fehler ($Min = 0$, $Max = 16$). Von den insgesamt 83 Teilnehmern machten 23% (19 Personen) keine Fehler. Die Stichprobe wurden anhand des Median geteilt (siehe Abschnitt 11.1), in Gruppe 1 befanden sich Personen, die keine bis maximal 2 Fehler beim AHKS machten, in Gruppe 2 waren Personen, die 3-16 Fehler hatten. Auch bei Fragestellung 2 wurde der t-Test als Verfahren der Wahl eingesetzt, wenn die Prüfung der Voraussetzungen dies ermöglichte. War die Voraussetzung der Normalverteilung der Gruppen nicht gegeben wurde erneut der Mann-Whitney-U-Test eingesetzt.

Ergebnis H2.1 (AHKS-Ordnungsliebe)

Die Ergebnisse für Shapiro-Wilk Test zur Überprüfung der Normalverteilung zeigen, dass sowohl in Gruppe 1, $W(48) = 0.977$, $p = .44$ als auch in Gruppe 1, ($W(35) = 0.977$, $p = .65$), die Normalverteilung angenommen werden konnte. Der Levene-Test $df = 81$, $F = 0.068$, $p = .795$ kam zu keinem signifikanten Unterschied zwischen den Varianzen. Aufgrund aller erfüllten Voraussetzungen wurde ein t-Test für unabhängig Stichproben gerechnet. Die Ergebnisse zeigen, dass es keine signifikanten Unterschiede, $t(81) = -0.824$, $df = 71.09$, $p = .413$, zwischen Teilnehmern mit wenigen Fehlern und Teilnehmern mit mehreren Fehlern im AHKS in Bezug auf die Werte der Dimension Ordnungsliebe im Test VICI gibt. Es kann daher angenommen werden, dass die Augen-Hand Koordination keine Auswirkungen auf die Ergebnisse des Test VICI in der Dimension Ordnungsliebe hat.

Ergebnis H2.2 (AHKS-Exaktheit)

Die Ergebnisse für Shapiro-Wilk Test zur Überprüfung der Normalverteilung zeigen, dass sowohl in Gruppe 1, $W(48) = 0.878$, $p = .00$ als auch in Gruppe 1, ($W(35) = 0.927$, $p = .02$), die Normalverteilung nicht angenommen werden

konnte. Der Levene-Test $df = 81$, $F = 0.961$, $p = .33$ kam zu keinem signifikanten Unterschied zwischen den Varianzen.. Aufgrund der Verletzung der Voraussetzung der Normalverteilung in den Gruppen wurde ein Mann-Whitney U-Test durchgeführt. Die Ergebnisse zeigen, dass es keinen signifikanten Unterschiede, $W(83) = 802.5$, $p = .733$, zwischen Teilnehmern mit wenigen Fehlern und Teilnehmern mit mehreren Fehlern im AHKS in Bezug auf die Werte der Dimension Exaktheit im Test VICI gibt. Es kann daher angenommen werden, dass die Augen-Hand Koordination keine Auswirkungen auf die Ergebnisse des Test VICI in der Dimension Exaktheit hat.

Ergebnis H2.3 (AHKS-IR)

Die Ergebnisse für Shapiro-Wilk Test zur Überprüfung der Normalverteilung zeigen, dass sowohl in Gruppe 1, $W(48) = 0.959$, $p = .09$ als auch in Gruppe 1, ($W(35) = 0.933$, $p = .03$), die Normalverteilung nicht angenommen werden konnte. Der Levene-Test kam zu keinem signifikanten Unterschied zwischen den Varianzen $df = 81$, $F = 0.113$, $p = .737$. Aufgrund der Verletzung der Voraussetzung der Normalverteilung in den Gruppen wurde ein Mann-Whitney U-Test durchgeführt. Die Ergebnisse zeigen, dass es keinen signifikanten Unterschiede, $W(83) = 850$, $p = .93$, zwischen Teilnehmern mit wenigen Fehlern und Teilnehmern mit mehreren Fehlern im AHKS in Bezug auf die Werte der Dimension Impulsivität/Reflexivität im Test VICI gibt. Es kann daher angenommen werden, dass die Augen-Hand Koordination keine Auswirkungen auf die Ergebnisse des Test VICI in der Dimension Impulsivität/Reflexivität hat.

Ergebnis H2.4 (AHKS-Trash)

Die Ergebnisse für Shapiro-Wilk Test zur Überprüfung der Normalverteilung zeigen, dass sowohl in Gruppe 1, $W(48) = 0.959$, $p = .09$ als auch in Gruppe

1, ($W(35) = 0.934$, $p = .03$), die Normalverteilung nicht angenommen werden konnte. Der Levene-Test $df = 81$, $F = 1.011$, $p = .318$ kam zu keinem signifikanten Unterschied zwischen den Varianzen. Aufgrund der Verletzung der Voraussetzung der Normalverteilung in den Gruppen wurde ein Mann-Whitney U-Test durchgeführt. Die Ergebnisse zeigen, dass es keinen signifikanten Unterschiede, $W(83) = 976$, $p = .203$, zwischen Teilnehmern mit wenigen Fehlern und Teilnehmern mit mehreren Fehlern im AHKS in Bezug auf die Werte der Dimension Trash-Score im Test VICI gibt. Es kann daher angenommen werden, dass die Augen-Hand Koordination keine Auswirkungen auf die Ergebnisse des Test VICI in der Dimension Trash-Score hat.

Beantwortung der Fragestellungen 2

Ausgehend von den Ergebnissen der einzelnen Unterhypothese kann die Haupthypothese 2 somit verworfen werden, es gibt keine Auswirkungen von Augen-Hand Koordination auf die Ergebnisse im Test VICI.

12.3 Prüfung der Fragestellung 3

Für die folgenden Berechnungen wurde die Stichprobe nach dem Geschlecht der Teilnehmer in zwei Gruppen geteilt. Bei Fragestellung 3 wurde der t-Test mit unabhängigen Stichproben für die statistische Auswertung eingesetzt, wenn die Prüfung der Voraussetzungen dies ermöglichte. War die Voraussetzung der Normalverteilung der Gruppen nicht gegeben wurde der Mann-Whitney-U-Test eingesetzt.

Ergebnis H3.1 (Geschlecht-Ordnungsliebe)

Die Ergebnisse für den Shapiro-Wilk Test zur Überprüfung der Normalverteilung zeigen, sowohl für Gruppe 1 ($W(30) = 0.952, p = .18$), als auch für Gruppe 2 ($W(53) = 0.958, p = .06$), ein nicht signifikantes Ergebnis. Es kann daher für beide Gruppen eine Normalverteilung angenommen werden. Der Levene-Test $df = 81, F = 0.269, p = .606$ kam ebenfalls zu keinem signifikanten Unterschied zwischen den Varianzen. Aufgrund aller erfüllten Voraussetzungen wurde ein t-Test für unabhängig Stichproben gerechnet. Die Ergebnisse zeigen, dass es keinen signifikanten Unterschied, $t(83) = -1.67, df = 56.69, p = .1$, zwischen Männern und Frauen in Bezug auf die Werte der Dimension Ordnungsliebe im Test VICI gibt. Es kann daher angenommen werden, dass das Geschlecht keine Auswirkungen auf die Ergebnisse des Test VICI in der Dimension Ordnungsliebe hat.

Ergebnis H3.2 (Geschlecht-Exaktheit)

Die Ergebnisse für den Shapiro-Wilk Test zur Überprüfung der Normalverteilung zeigen, für Gruppe 1 ($W(30) = 0.94, p = .09$), ein nicht signifikantes Ergebnis. Für Gruppe 2 ($W(53) = 0.873, p = .00$), resultiert ein signifikantes Ergebnis. Es kann daher nicht für beide Gruppen eine Normalverteilung angenommen werden. Der Levene-Test $df = 81, F = 0.269, p = .726$ kam zu keinem signifikanten Unterschied zwischen den Varianzen. Aufgrund der Verletzung der Voraussetzung der Normalverteilung in den Gruppen wurde ein Mann-Whitney U-Test durchgeführt. Die Ergebnisse zeigen, dass es keinen signifikanten Unterschied, $W(83) = 613.5, p = .086$, zwischen Männern und Frauen in Bezug auf die Werte der Dimension Exaktheit im Test VICI gibt. Es kann daher angenommen werden, dass das Geschlecht keine Auswirkungen auf die Ergebnisse des

Test VICI in der Dimension Exaktheit hat.

Ergebnis H3.3 (Geschlecht-IR)

Die Ergebnisse für den Shapiro-Wilk Test zur Überprüfung der Normalverteilung zeigen, für Gruppe 1 ($W(30) = 0.965, p = .40$), ein nicht signifikantes Ergebnis. Für Gruppe 2 ($W(53) = 0.943, p = .01$), resultiert ein signifikantes Ergebnis. Es kann daher nicht für beide Gruppen eine Normalverteilung angenommen werden. Der Levene-Test $df = 81, F = 0.479, p = .491$ kam zu keinem signifikanten Unterschied zwischen den Varianzen. Aufgrund der Verletzung der Voraussetzung der Normalverteilung in den Gruppen wurde ein Mann-Whitney U-Test durchgeführt. Die Ergebnisse zeigen, dass es keinen signifikanten Unterschied, $W(83) = 679, p = .274$, zwischen Männern und Frauen in Bezug auf die Werte der Dimension Impulsivität/Reflexivität im Test VICI gibt. Es kann daher angenommen werden, dass das Geschlecht keine Auswirkungen auf die Ergebnisse des Test VICI in der Dimension Impulsivität/Reflexivität hat.

Ergebnis H3.4 (Geschlecht-Trash)

Die Ergebnisse für den Shapiro-Wilk Test zur Überprüfung der Normalverteilung zeigen, für Gruppe 1 ($W(30) = 0.939, p = .08$), ein nicht signifikantes Ergebnis. Für Gruppe 2 ($W(53) = 0.946, p = .01$), resultiert ein signifikantes Ergebnis. Es kann daher nicht für beide Gruppen eine Normalverteilung angenommen werden. Der Levene-Test $df = 81, F = 0.813, p = .37$ kam zu keinem signifikanten Unterschied zwischen den Varianzen. Aufgrund der Verletzung der Voraussetzung der Normalverteilung in den Gruppen wurde ein Mann-Whitney U-Test durchgeführt. Die Ergebnisse zeigen, dass es keinen signifikanten Unterschied, $W(83) = 888.5, p = .369$, zwischen Männern und Frauen in Bezug auf die Werte der Dimension Trash-Score im Test VICI gibt. Es kann daher angenommen

werden, dass das Geschlecht keine Auswirkungen auf die Ergebnisse des Test VICI in der Dimension Trash-Score hat.

12.4 Prüfung der Fragestellung 4

Für die folgenden Berechnungen wurde die Stichprobe nach der Fähigkeit zum räumlichen Orientierungsvermögen in zwei extremgruppen Gruppen geteilt. Dafür wurden die Korrekten Antworten der Testpersonen im Test 3D in Prozenträge unterteilt. All jene Personen mit einem Prozentrang (≤ 25) bildeten die Extremgruppe mit eher niedriger Fähigkeit ($N = 26$) in räumlichen Orientierungsvermögens und all jene mit einem Prozentrang von ≥ 75 wurden jener Gruppe mit hoher Fähigkeit ($N = 24$) in räumlichen Orientierungsvermögen zugeteilt. Bei Fragestellung 4 wurde der t-Test mit unabhängigen Stichproben für die statistische Auswertung eingesetzt, wenn die Prüfung der Voraussetzungen dies ermöglichte. Im Falle ungleicher Standardabweichungen in beiden Gruppen, wurde der Welch-Test eingesetzt als Alternative für den t-Test verwendet. War die Voraussetzung der Normalverteilung der Gruppen nicht gegeben wurde der Mann-Whitney-U-Test eingesetzt.

Ergebnis H4.1 (3D-Ordnungsliebe)

Die Ergebnisse für den Shapiro-Wilk Test zur Überprüfung der Normalverteilung zeigen sowohl für Gruppe 1 ($W(26) = 0.927, p = .06$), also auch für Gruppe 2 ($W(24) = 0.939, p = .15$) keine signifikanten Ergebnisse. Es kann daher für beide Gruppen eine Normalverteilung angenommen werden. Der Levene-Test $df = 64, F = 2.475, p = .004$ kam zu einem signifikanten Unterschied zwischen den Varianzen. Aufgrund der fehlenden Homogenität der Varianzen, wurde ein Welch-Test für unabhängig Stichproben gerechnet. Die Ergebnisse

zeigen, dass es keinen signifikanten Unterschied, $t(50) = -0.728$, $df = 45.63$, $p = .471$, zwischen Personen mit niedrigen und Personen mit hohen Fähigkeiten in räumlichen Orientierungsvermögen bezogen auf die Werte der Dimension Ordnungsliebe im Test VICI gibt. Es kann daher angenommen werden, dass die Fähigkeit im räumlichen Orientierungsvermögen keine Auswirkungen auf die Ergebnisse des Test VICI in der Dimension Ordnungsliebe hat.

Ergebnis H4.2 (3D-Exaktheit)

Die Ergebnisse für den Shapiro-Wilk Test zur Überprüfung der Normalverteilung zeigen ein nicht signifikantes Ergebnis für Gruppe 1 ($W(26) = 0.931$, $p = .08$), jedoch nicht signifikantes Ergebnis für Gruppe 2 ($W(24) = 0.787$, $p = .00$). Es kann daher nicht für beide Gruppen eine Normalverteilung angenommen werden. Der Levene-Test $df = 64$, $F = 4.195$, $p = .00$ kam zu einem signifikanten Unterschied zwischen den Varianzen. Aufgrund der Verletzung der Voraussetzung der Normalverteilung als auch der fehlenden Homogenität der Varianzen in den Gruppen wurde ein Mann-Whitney U-Test durchgeführt. Die Ergebnisse zeigen, dass es keinen signifikanten Unterschied, $W(50) = 352.5$, $p = .437$, zwischen Personen mit niedrigen und Personen mit hohen Fähigkeiten in räumlichen Orientierungsvermögen in bezogen auf die Werte der Dimension Exaktheit im Test VICI gibt. Es kann daher angenommen werden, dass die Fähigkeit in räumlichen Orientierungsvermögens keine Auswirkungen auf die Ergebnisse des Test VICI in der Dimension Exaktheit hat.

Ergebnis H4.3 (3D-IR)

Die Ergebnisse für den Shapiro-Wilk Test zur Überprüfung der Normalverteilung zeigen sowohl für Gruppe 1 ($W(26) = 0.932$, $p = .08$), also auch für Gruppe 2 ($W(24) = 0.938$, $p = .14$) keine signifikanten Ergebnisse. Es kann daher

für beide Gruppen eine Normalverteilung angenommen werden. Der Levene-Test $df = 64$, $F = 1.923$, $p = .004$ kam zu einem signifikanten Unterschied zwischen den Varianzen. Aufgrund der fehlenden Homogenität der Varianzen, wurde ein Welch-Test für unabhängig Stichproben gerechnet. Die Ergebnisse zeigen, dass es keinen signifikanten Unterschied, $t(50) = 1.612$, $df = 47.88$, $p = .471$, zwischen Personen mit niedrigen und Personen mit hohen Fähigkeiten in räumlichen Orientierungsvermögen in bezogen auf die Werte der Dimension Impulsivität/Reflexivität im Test VICI gibt. Es kann daher angenommen werden, dass die Fähigkeit in räumlichen Orientierungsvermögens keine Auswirkungen auf die Ergebnisse des Test VICI in der Dimension Impulsivität/Reflexivität hat.

Ergebnis H4.4 (3D-Trash)

Die Ergebnisse für den Shapiro-Wilk Test zur Überprüfung der Normalverteilung zeigen sowohl für Gruppe 1 ($W(26) = 0.942$, $p = .14$), also auch für Gruppe 2 ($W(24) = 0.923$, $p = .06$) keine signifikanten Ergebnisse. Es kann daher für beide Gruppen eine Normalverteilung angenommen werden. Der Levene-Test $df = 64$, $F = 0.654$, $p = .841$ kam zu einem signifikanten Unterschied zwischen den Varianzen. Aufgrund der fehlenden Homogenität der Varianzen, wurde ein Welch-Test für unabhängig Stichproben gerechnet. Die Ergebnisse zeigen, dass es keinen signifikanten Unterschied, $t(50) = -0.565$, $df = 47.66$, $p = .575$, zwischen Personen mit niedrigen und Personen mit hohen Fähigkeiten in räumlichen Orientierungsvermögen in bezogen auf die Werte der Dimension Trash-Score im Test VICI gibt. Es kann daher angenommen werden, dass die Fähigkeit in räumlichen Orientierungsvermögens keine Auswirkungen auf die Ergebnisse des Test VICI in der Dimension Trash-Score hat.

13 Konstruktvalidierung

Anhand einer Korrelationsmatrix wurde versucht eine Konstruktvalidierung des Test VICI mittels AHA und NEO PI-R durchzuführen. Aus Gründen des Umfangs werden diese Tabellen (siehe Seite 125) im Anhang zur Verfügung gestellt.

13.1 Prüfung der Fragestellung 5

H5

Für die Überprüfung der Konstruktvalidierung von VICI und AHA wurde eine Pearson Produkt-Moment-Korrelationsmatrix über die 5 Dimensionen des AHA (Impulsivität/ Reflexivität, Exaktheit, Entscheidungsfreudigkeit, Leistungsniveau Anspruchsniveau) sowie ausgewählten Dimensionen des VICI (Ordnungsliebe Score, Trash Score, Kartonbewegung zu T1, Entscheidungsfreude, Exaktheit Score, Arbeitsgeschwindigkeit, Impulsivität/ Reflexivität Score, Produktivität, Zieldiskrepanz, Mehr PR, Anspruchsniveau, Geschätzte Produktivität, Kartonbewegungen zu T2, Bestellung Richtig, Bestellung Falsch und dem System) gerechnet.

13.2 Prüfung der Fragestellung 6

H6

Für die Überprüfung der Konstruktvalidierung von VICI und NEO PI-R wurde eine Pearson Produkt-Moment-Korrelationsmatrix über die 5 Dimensionen der Gewissenhaftigkeit sowie den Gesamtwert der Gewissenhaftigkeit des NEO PI-R, sowie ausgewählten Dimensionen des VICI (Ordnungsliebe Score, Trash Score, Kartonbewegung zu T1, Entscheidungsfreude, Exaktheit Score, Arbeits-

geschwindigkeit, Impulsivität/Reflexivität, Produktivität, Zieldiskrepanz, Mehr PR, Anspruchsniveau, Geschätzte Produktivität, Kartonbewegungen zu T2, Bestellung Richtig, Bestellung Falsch und dem System) gerechnet.

Ergebnisse der Konstruktvalidierung

Die Korrelationsmatrizen (siehe Anhang, Seite 125) zeigen, dass es neben der Korrelation von Kartonbewegungen zu T1 und Selbstdisziplin ($r = -.306$, $p = .005$), Kartonbewegungen zu T1 und dem Gesamtscore des NEO PI-R-Gewissenhaftigkeit ($r = -.253$, $p = .021$) und Mehr PR und Besonnenheit ($r = -.229$, $p = .037$) keine signifikante Korrelationen zwischen dem Persönlichkeits-Fragebogen (NEO PI-R) und der Computersimulation (VICI) gab, bei den genannten Korrelationen handelt es sich um niedrige bis moderate negative Zusammenhänge. Ähnlich verhielt es sich mit dem Versuch der Konstruktvalidierung einiger Kennwerte aus dem AHA mit der Computersimulation VICI. Die Korrelationsmatrizen zeigen, dass es neben der Korrelation von Mehr PR und Anspruchsniveau ($r = -.246$, $p = .030$) und Bestellung Richtig und Leistungsniveau ($r = .253$, $p = .022$) keine signifikanten Zusammenhänge zwischen dem AHA und dem VICI gab. Eine klassische Validierung anhand anderer konstrukt-naher Objektive Persönlichkeitstests ist schwerlich durchzuführen, da es bisher nur wenige solcher Verfahren gibt (Ortner et al., 2007). Daraus kann geschlossen werden, dass entweder unterschiedliche Methoden, unterschiedliche Konstrukte oder unterschiedliche Aspekte des selben Konstrukts erfasst werden (vgl. Ifcic, 2012; Kubinger & Litzenberger, 2003; Skinner & Howarth, 1973).

14 Faktorenanalyse

Eine Hauptkomponentenanalyse mit der schiefwinkligen Oblimin Rotation mit Kaiser-Normalisierung wurde mit 16 Items des Test VICI durchgeführt. Das Kaiser-Meyer-Olkin Kriterium der Stichprobeneignung gibt an, ob ein Datensatz für die Faktorenanalyse geeignet ist, mit einem KMO = .557 ist diese Stichprobe mittelmäßig gut für die Durchführung der Faktorenanalyse geeignet (Field, 2009). Der Bartlett-Test für Sphärizität $\chi^2(120) = 1292.68, p < .001$, zeigt, dass die Korrelationen zwischen den einzelnen Items ausreichend groß waren um eine Hauptkomponentenanalyse durchzuführen. Eine erste Analyse wurde ausgeführt um die Eigenwerte für jede Komponente der Daten zu erlangen. Fünf Komponenten wiesen einen Eigenwert > 1 (Kaiser Kriterium) auf, in Kombination erklärten diese Komponenten 79% der Varianz.

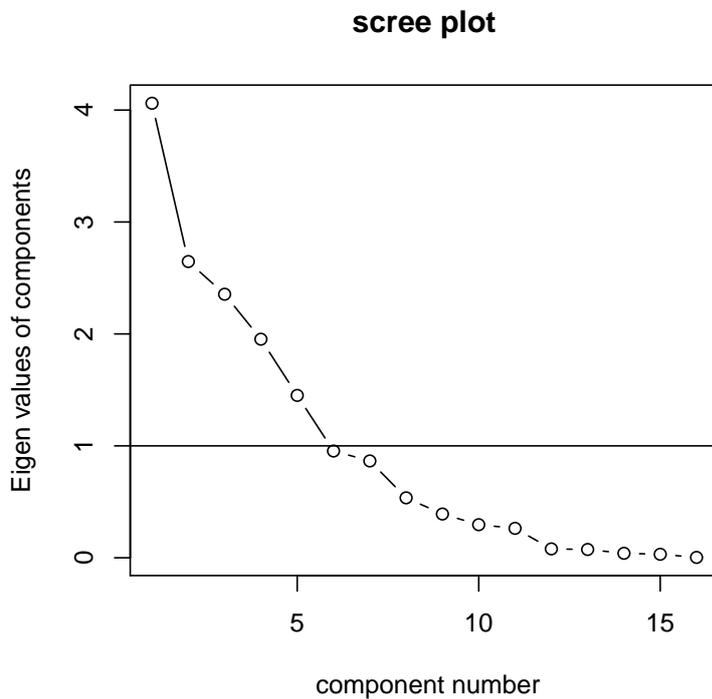


Abbildung 32: Screeplot

Der Screeplot (Abbildung 32) war etwas zweideutig und zeigte Flexionen, die die Beibehaltung von einem oder fünf Faktoren rechtfertigen würde. Es wurden, in Übereinstimmung mit dem Eigenwert-Kriterium und dem Screeplot, fünf Faktoren für die finale Analyse beibehalten. Abbildung 33 zeigt die Faktorladungen nach der Rotation. Aufgrund der Faktorladungen auf den einzelnen Komponenten wurden diese Faktoren inhaltlich interpretiert.

```
FAC.output<-fa(FAC,nfactors=5,SMC=TRUE,fm="pa",rotate="oblimin")
print(FAC.output$loadings,cutoff=.4)
```

Loadings:

	PA1	PA2	PA4	PA5	PA3
Ordnungsliebe.Score				1.007	
Trash.Score				0.747	
Kartonbewegung zu T1				0.407	
Entscheidungsfreude	-0.927				
Exaktheit.Score	0.549				
Arbeitsgeschwindigkeit	0.931				
IR.Score	0.977				
Produktivität			0.924		
Zieldiskrepanz			0.969		
MehrPR		-0.649			
Anspruchsniveau		0.908			
Geschätzte Produktivität		0.995			
Kartonbewegung zu T2					-0.515
Bestellung Richtig					0.565
Bestellung Falsch					-0.783
System					0.567
	PA1	PA2	PA4	PA5	PA3
SS loadings	3.031	2.430	2.020	1.923	1.705
Proportion Var	0.189	0.152	0.126	0.120	0.107
Cumulative Var	0.189	0.341	0.468	0.588	0.694

Abbildung 33: Faktorladungen

Faktor 1

Faktor 1 setzt sich aus vier Items der Aufgabe 2 zusammen, die den Arbeitsstil, also wie am Fließband gearbeitet wird, beschreibt. Zu den Items, die auf Faktor 1 laden, zählen demnach die *Entscheidungsfreude*, die *Exaktheit*, *Arbeitsgeschwindigkeit* (durchschnittlich benötigte Entscheidungszeit, während der ersten 100 Schokoladentafeln) sowie der Score *Impulsivität/Reflexivität* (IR.Score).

Faktor 2

Faktor 2 besteht aus drei Items der Aufgabe 2, die die Selbsteinschätzung in Bezug auf die quantitative Produktionsleistung darstellen. Zu den Items, die auf Faktor 2 laden, zählen der Wert für die *geschätzte Anzahl an Schokoladentafel* (Geschätzte Produktivität), die von der Person produziert werden, das *Anspruchsniveau* sowie das Maß für die Abweichung vom geschätzten Wert in Prozent (MehrPR).

Faktor 3

Faktor 3 umfassen die Items, die für Aufgabe 3, der Absolvierung einer Bestellung, wichtig sind. Dazu zählen die Items *Bestellung Richtig* (die Anzahl der richtig abgelieferten Kartons), *Bestellung Falsch* (Anzahl an falsch abgelieferten Kartons), *Anzahl der Kartonbewegungen zu T2* sowie die *Art des Systems*, das von der Person im Rahmen des Aufräumens in Aufgabe 1 erstellt wurde (System).

Faktor 4

Faktor 4 ist, wie bereits Faktor 1 und 2, als eine Untereinheit von Aufgabe 2 zu verstehen, bei dem es sich um die Leistungsbereitschaft der Person handelt. Die Items *Anzahl an produzierten Tafeln* (Produktivität) und die *Zieldiskrepanz* laden auf diesen Faktor.

Faktor 5

Faktor 5 umfasst die Items, die zu Aufgabe 1, dem Ordnungsschaffen in der Fabrik, zugehörig sind. Dazu zählen das System, das von der Person erstellt

wurde, der *Ordnungsliebe-Score*, der *Trash-Score* sowie die *Anzahl der Kartonbewegungen zu T1*.

15 Zusammenfassung der Ergebnisse

Im folgenden Abschnitt sollen die Ergebnisse dieser Studie kurz zusammengefasst werden. Die Stichprobe umfasste 85 Teilnehmer, von denen 83 Datensätze für die Auswertung verwendet werden konnten. Die Stichprobe bestand zum größten Teil aus Studierenden der Psychologie, die im Rahmen der Testerfahrung den Test VICI absolvierten. Um die Stichprobe zu beschreiben, werden ausgewählte Ergebnisse und Kennwerte des Test VICI präsentiert. Bezüglich der Computer-Erfahrung gab keine der Testpersonen an, über "keine Erfahrung" zu verfügen, daher wurde diese Variable aus der Auswertung entfernt. Von den Teilnehmern gab etwa die Hälfte an, keinerlei Erfahrung mit 3D-Computerspielen zu haben. Etwa 29% der getesteten Personen gaben an, über ein wenig 3D-Spiele Erfahrung zu verfügen und 23% hatten schon viel Erfahrung bei diesem Typ von Videospiele.

Beim Screening zur Augen-Hand-Koordination konnte festgestellt werden, dass die Testpersonen im Schnitt 3 Fehler machten.

Bei der ersten Aufgabe des Tests wurden von ca. 80% der Teilnehmer alle Objekte gefunden und weggeräumt (*Trash-Score*). Beim *Ordnungsliebe-Score* stellte sich heraus, dass zwei Drittel der Testpersonen ein Ordnungssystem erstellten und ein Drittel die Kartons unangetastet ließen.

Zu Aufgabe 2 kann gesagt werden, dass die erwünschte Marke von 200 Tafeln vom Großteil der Teilnehmer erfüllt wurde. In Bezug auf die Entscheidungsfreude wurden im Durchschnitt 42 Tafeln in einer Minute bearbeitet. Es zeigte sich, dass der Grad an Exaktheit bei der Aussortierung am Fließband durchweg sehr hoch war. Der hohe Zusammenhang zwischen Betrachtungsdauer und Exaktheit verdeutlicht, dass eine höhere Betrachtungsdauer gleichzeitig auch die Genauigkeit am Fließband steigert (Impulsivität/ Reflexivität).

Bei Aufgabe 3 lieferten von den insgesamt 83 Teilnehmern in etwa die Hälfte die drei richtigen bestellten Kartons im Abholbereich ab. Der Großteil derjenigen, die die drei korrekten Boxen in den Abholbereich brachten, sortierte zuvor sowohl nach Stückzahl, als auch nach Sorte.

Bei der Auswertung von Fragestellung 1-4 wurde untersucht, ob der Einsatz virtueller Umgebungen bezogen auf Handhabbarkeit und Orientierungsfähigkeit generell eine Möglichkeit für die psychologische Diagnostik darstellt. Dazu wurde in den Hypothesen 1.1-1.4 geprüft, ob sich die Ergebnisse des Tests VICI (*Ordnungsliebe, Exaktheit, Impulsivität/Reflexivität* und *Trash-Score*) aufgrund der 3D-Gaming Erfahrung der Teilnehmer unterscheiden. Die Ergebnisse der t-Tests oder U-Tests zeigten, dass es keinen signifikanten Unterschied zwischen Teilnehmern mit 3D-Gaming Erfahrung und Teilnehmern ohne 3D-Gaming Erfahrung in Bezug auf die Werte des Test VICI gibt. Es kann daher angenommen werden, dass die Erfahrung in 3D-Computerspielen keine Auswirkungen auf die Ergebnisse des Test VICI hat.

Die Hypothesen 2.1-2.4 testeten, ob sich die Ergebnisse des Tests VICI (*Ordnungsliebe, Exaktheit, Impulsivität/Reflexivität* und *Trash-Score*) aufgrund der

Augen-Hand-Koordination der Teilnehmer unterscheiden. Die Ergebnisse zeigen, dass es keine signifikanten Unterschiede zwischen Teilnehmern mit wenigen Fehlern und Teilnehmern mit mehreren Fehlern im AHKS in Bezug auf die Werte des Test VICI gibt. Es kann daher angenommen werden, dass die Augen-Hand Koordination keine Auswirkungen auf die Ergebnisse des Test VICI hat.

Die Hypothesen 3.1-3.4 testeten, ob sich die Ergebnisse des Tests VICI (*Ordnungsliebe, Exaktheit, Impulsivität/Reflexivität* und *Trash-Score*) aufgrund des Geschlechts der Teilnehmer unterscheiden. Die Ergebnisse zeigen, dass es keine signifikanten Unterschiede zwischen weiblichen und männlichen Teilnehmern in Bezug auf die Werte des Test VICI gibt. Es kann daher angenommen werden, dass das Geschlecht keine Auswirkungen auf die Ergebnisse des Test VICI hat.

Die Hypothesen 4.1-4.4 testeten, ob sich die Ergebnisse des Tests VICI (*Ordnungsliebe, Exaktheit, Impulsivität/Reflexivität* und *Trash-Score*) aufgrund der Fähigkeit zur räumlichen Orientierungsfähigkeit der Teilnehmer unterscheiden. Die Ergebnisse zeigen, dass es keine signifikanten Unterschiede zwischen Personen mit niedrigen und Personen mit hohen Fähigkeiten in räumlichen Orientierungsvermögen in bezogen auf die Werte im Test VICI gibt. Es kann daher angenommen werden, dass das räumliche Orientierungsvermögen keine Auswirkungen auf die Ergebnisse des Test VICI hat. Aufgrund der vorliegenden Ergebnisse kann gesagt werden, dass individuelle Ausprägungen in 3D-Spiele-Erfahrung, Augen-Hand Koordination, Geschlecht oder räumliches Orientierungsvermögen keinen Einfluss auf die Ergebnisse im Test VICI haben.

Im Rahmen der Hypothesen 5 und 6 wurde versucht eine Konstruktvalidierung durchzuführen. Die Korrelationen zwischen dem Objektiven Persönlichkeitsverfahren AHA und der Computersimulation VICI beziehungsweise zwischen dem Persönlichkeits-Fragebogen NEO PI-R und der Computersimulation VICI fielen zum größten Teil nicht signifikant aus. Die Schlussfolgerung lautet, dass die Kennwerte aus dem AHA wie auch aus dem NEO PI-R nicht geeignet sind für die Konstruktvalidierung der hier vorliegenden Computersimulation. Es dürften also entweder unterschiedliche Methoden, unterschiedliche Konstrukte oder unterschiedliche Aspekte des selben Konstrukts bei AHA, NEO PI-R und VICI erfasst werden.

Eine abschließende Faktorenanalyse prüfte die Dimensionalität des VICI. Die exploratorische Faktorenanalyse ergab, dass die einzelnen Variablen des Test VICI auf 5 Faktoren laden, welche jeweils einer der 3 Aufgaben des Test VICI zugeordnet werden konnten.

16 Interpretation und Diskussion der Ergebnisse

In diesem Kapitel sollen die Ergebnisse in Bezug auf eventuelle Fragestellung interpretiert und diskutiert werden.

Fragestellungen 1-4

Bezüglich Fragestellung 1, ob sich die Ergebnisse des Tests VICI (*Ordnungsliebe, Exaktheit, Impulsivität/Reflexivität* und *Trash-Score*) aufgrund der 3D-Gaming Erfahrung der Teilnehmer unterscheiden, ist zu sagen, dass die vier Unterhypothesen nicht signifikant ausfielen, das bedeutet, dass die Nullhypothese beibe-

halten werden kann und es keine Auswirkungen von 3D-Spiele Erfahrung auf die Ergebnisse im Test VICI gibt. Es kann daraus geschlossen werden, dass Personen aufgrund ihrer 3D-Spiel Erfahrung keinen Vorteil bei der Bearbeitung eines solchen Verfahrens haben, welcher gegen das Gütekriterium der Fairness verstoßen würde. In diesem Zusammenhang soll auf die Einführungsphase des Tests VICI verwiesen werden (siehe Abschnitt 5.4). Ziel der Instruktionsphase war es, möglichst simple und präzise Hilfestellungen für die Navigation und Interaktion zu geben, um die Probanden möglichst einfach an die Steuerung zu gewöhnen. Nicht zu vernachlässigen ist jedoch, dass auf eine Auswertung in Bezug auf die Computer-Erfahrung der Probanden verzichtet werden musste, da keine der Personen angegeben hatte, dass sie über keine Computer-Erfahrung verfüge. Wäre die Stichprobe in Bezug auf dieses Merkmal heterogener, würde dies die Ergebnisse eventuell erheblich beeinflussen.

Bezüglich Fragestellung 2, ob sich die Ergebnisse des Tests VICI (*Ordnungs-
liebe, Exaktheit, Impulsivität/Reflexivität* und *Trash-Score*) aufgrund der Augen-Hand Koordination der Teilnehmer unterscheiden, zeigte sich, dass auch diesmal die vier Unterhypothesen nicht signifikant ausfielen, daraus kann gefolgert werden, dass die Nullhypothese beibehalten werden kann und es keine Auswirkung der Augen-Hand Koordination auf die Ergebnisse im Test VICI gibt. Dieses Ergebnis ist sehr wesentlich in Bezug auf den Einsatz multidimensionaler virtueller Umgebungen, da die Navigation und die Bearbeitung damit zusammenhängender Aufgaben nicht eingeschränkt wird. Auch für die Augen-Hand Koordination gilt der gleiche Einwand in Bezug auf die Computer-Erfahrung der Testpersonen, da das Screening für Augen-Hand Koordination mit der Computermaus bearbeitet wird, ist das Verständnis im Umgang mit dieser eine Voraussetzung für die Ergebnisse dieser Hypothesen.

Bezüglich Fragestellung 3, ob sich die Ergebnisse des Tests VICI (*Ordnungs-
liebe, Exaktheit, Impulsivität/Reflexivität* und *Trash-Score*) aufgrund des Ge-
schlechts der Teilnehmer unterscheiden, ist zu sagen, dass die vier Unterhy-
pothesen nicht signifikant ausfielen, das bedeutet, dass die Nullhypothese bei-
behalten werden kann und es keine Auswirkungen des Geschlechts auf die
Ergebnisse im Test VICI gibt. Dazu sei auf die Testkonstruktion in Abschnitt 5
verwiesen, bei der versucht wurde, geschlechtsneutrale Stimuli einzusetzen, um
motivationale Vorteile bzw. Nachteile oder systematische Verzerrungen zu redu-
zieren. Das Szenario einer Schokoladenfabrik als industrielle Produktionsstätte
wurde gewählt um eine möglichst breite Akzeptanz gewährleisten zu können.

Bezüglich Fragestellung 4, ob sich die Ergebnisse des Tests VICI (*Ordnungs-
liebe, Exaktheit, Impulsivität/Reflexivität* und *Trash-Score*) aufgrund des räum-
lichen Orientierungsvermögen der Teilnehmer unterscheiden, zeigte sich, dass
die vier Unterhypothesen nicht signifikant ausfielen, das bedeutet, dass die Null-
hypothese beibehalten werden kann und es keine Auswirkungen des räumli-
chen Orientierungsvermögen auf die Ergebnisse im Test VICI gibt. Die räumli-
che Orientierung ist ein wichtiger Faktor in einer virtuellen Umgebung, da zu-
sätzlich zur Bewegung auch gleichzeitig der Blick gesteuert werden muss. Diese
beiden Steuerungen auch noch mit der Interaktion mit der Umgebung und der
Orientierung im Raum zu kombinieren stellt eine äußerst komplexe Angelegen-
heit dar. In der Konstruktion des Tests VICI wurde für die benutzerfreundliche
Gestaltung der Produktionshalle viel Zeit investiert. Zu erwähnen wären hierbei
die große Halle, welche von jedem Standpunkt aus eine uneingeschränkte Sicht
auf den Innenraum erlaubt um einen Verlust der räumlichen Orientierung vor-
zubeugen. Sowie Fenster in den Wänden und in der Decke der Halle, um eine

bedrückende Enge zu vermeiden. Ein Verlust der räumlichen Orientierungsfähigkeit hätte erheblichen Einfluss auf die Testergebnisse, welche als Facetten einer Persönlichkeitseigenschaft unabhängig von kognitiven Fähigkeiten bleiben sollten.

Konstruktvalidierung/ Übereinstimmungsvalidität

Beim Versuch einer Konstruktvalidierung (Fragestellungen 5 und 6) anhand des NEO-PI-R und des AHA konnten nur sehr wenige signifikante Zusammenhänge dieser Verfahren mit dem VICI gefunden werden. Auch in anderen Studien (vgl. Hauser, 2011; Ifcic, 2012) sind Zusammenhänge zwischen Objektiven Persönlichkeitsverfahren und Persönlichkeitsfragebögen ausgeblieben. Aus dieser Tatsache kann geschlossen werden, dass davon ausgegangen werden muss, dass Persönlichkeitsfragebogen, in diesem Fall der NEO PI-R, und objektive Persönlichkeitsverfahren, in dieser Studie der VICI, entweder unterschiedliche Konstrukte oder aber unterschiedliche Aspekte eines Konstruktes erfassen (vgl. Hauser, 2011; Ifcic, 2012; Kubinger & Litzenberger, 2003; Skinner & Howarth, 1973). Überraschend ist jedoch, dass besonders jene Kennwerte, welche namensgleich im AHA sowie im VICI vorkommen (*Impulsivität/Reflexivität, Exaktheit* oder *Entscheidungsfreude*) ebenfalls nicht signifikant korrelieren. Jene wenigen signifikanten Korrelationen (wie zwischen *Bestellung Richtig* und *Leistungsniveau*) weisen keinen besonders starken Zusammenhang auf. Weshalb beim VICI und AHA keine erfolgreiche Konstruktvalidierung durchgeführt werden konnte, kann an verschiedenen Gründen liegen. Zum einen verfügen diese beiden Objektiven Persönlichkeitsverfahren über vollkommen unterschiedliches Stimulusmaterial, andererseits ist auch die Herangehensweise divers. Der Test AHA verwendet abstraktes und uneindeutiges Stimulusmaterial, der Test VICI hingegen versucht als dreidimensionale Computersimulation eine explorative

Verhaltensbeobachtung sowie inhaltliche als auch ökologische Validität zu vereinen. Eine klassische Validierung anhand anderer konstruktnaher objektiver Persönlichkeitstests ist außerdem deshalb schwerlich durchzuführen, da es bisher nur wenige solcher Verfahren gibt (Ortner et al., 2007). Es kann geschlossen werden, dass entweder unterschiedliche Methoden, unterschiedliche Konstrukte oder unterschiedliche Aspekte des selben Konstrukts erfasst werden (vgl. Ifcic, 2012; Kubinger & Litzenberger, 2003; Skinner & Howarth, 1973). Für zukünftige Studien wäre eine Konstruktvalidierung mittels Multitrait-Multimethod-Ansatz zu überlegen.

Faktorenanalyse

Die explorative Faktorenanalyse bestätigt, dass die einzelnen Aufgaben des Test VICI jeweils auf eigene Faktoren laden. Aufgabe 2 ist als mehrdimensional anzusehen. Die drei Faktoren Arbeitsstil, Leistungsbereitschaft und Selbsteinschätzung können unterschieden werden. Ein weiterer interessanter Punkt ist, dass das in Aufgabe 1 erstellte System auf den Faktor der Aufgabe 3 lädt. Dieses interessante Ergebnis zeigt, dass diese Aufgaben nicht voneinander unabhängig sind, sondern aufeinander aufbauen, es liegt die Vermutung nahe, dass dies ein in der Realität häufiges Phänomen darstellt.

AHKS

Das Screening zur Augen-Hand Koordination konnte eingesetzt werden, um zwischen Personen bezüglich ihrer Leistungsfähigkeit in diesem Bereich zu differenzieren. Es zeigte sich, dass knapp die Hälfte der Probanden keine oder nur sehr wenige Fehler machten und die andere Hälfte der Personen eine große Spannweite aufwies. Als Intro für den Test VICI wurde das Screening gut ange-

nommen. Ein wichtiges Kriterium für einen reibungslosen Testablauf bei computergestützten psychologisch-diagnostischen Verfahren, welche häufig nicht beachtet werden, stellen die Mauseinstellungen sowie unterschiedliche Bildschirmgröße dar. Da in dieser Studie ein Großteil der Testpersonen in den Testräumen der Computerdiagnostik teilnahmen, konnten Maus,- und Bildschirmeinstellungen für diese Personen standardisiert werden. Jene Testpersonen, welche nicht in der Universität getestet wurden, wurden an einem Laptop mit den selben Einstellungen getestet.

Aufgabe 1

Kartons sortieren

Die Vermutung, die der Möglichkeit zu freien Bearbeitung dieser Aufgabe zugrunde lag, scheint sich bestätigt zu haben: Ordnung ist als etwas sehr subjektives anzusehen. Bei der Bearbeitung der Aufgabe wurden viele verschiedene Arten von Systemen zur Sortierung der Kartons verwendet. Es gab klar strukturierte Systeme, die nach Sorte und Stückzahl differenzierten und weniger strukturierte Systeme, die jeweils nur einer dieser Eigenschaften umfassten. Es gab allerdings auch Ordnungen, in denen kein System erkennbar war, sondern die Kartons zum Beispiel an den Rand gerückt, oder auf einen Haufen gelegt wurden. Für die Bearbeitung der Aufgabe wurde die Halle großräumig genutzt, die Kartons konnten von den Testpersonen bewegt und gestapelt werden. Durch das erstellte System war es den Probanden möglich, später bestimmte Kartons ohne viel Aufwand wiederzufinden. In der Hauptkomponentenanalyse stellte sich heraus, dass das Item *System* das einzige war, welches in einer anderen Aufgabe (auf einem anderen Faktor) höher lud, als in jenem, zu dem es eigentlich gehörte. Das Item *System* lud auf Faktor 3, welcher die Aufga-

be 3 darstellte. Das System stellte sich als elementar für die Bearbeitung der letzten Aufgabe heraus. Zu kritisieren ist jedoch, dass die Instruktion teilweise nicht eindeutig genug waren, gewisse Informationen (zum Beispiel, dass es verschiedene Sorten und Stückzahlen der Kartons gibt), hätte klarer ausgedrückt werden können. Leider wurde auch der Abholbereich von einigen Testpersonen falsch interpretiert. Diese Personen hielten den Abholbereich für jenen Ort, an dem alle Kartons aufzustellen wären und nicht als den Ort, an dem ein Lieferant die Bestellung abholen könne. Genauere Instruktion würde womöglich mehr Testpersonen auffordern, beim Ordnungschaffen ein System zu erstellen. Teilweise wurde diese Aufgabe mit viel Ehrgeiz bearbeitet. Die Kartons genau zu stapeln, ist aber auch sehr zeitintensiv und ohne Zeitdruck mit einer stressigen Arbeitssituationen in der Realität nur wenig vergleichbar. Hier könnten in Zukunft zeitliche Begrenzungen eingeführt werden.

Abfall

Retrospektiv kann gesagt werden, dass die einzelnen Items relativ einfach zu finden waren, außer jene, welche unrealistischerweise im Kasten des Büros versteckt waren und somit viel seltener gefunden wurden. Generell kann die Aufgabe, den Abfall zu finden und aufzuheben als eine gute Möglichkeit zur Kontrolle von Explorationsverhalten angesehen werden. So kann festgestellt werden, wenn zum Beispiel eine Aufgabe nicht bearbeitet wurde, ob die Person sich überhaupt an dem Ort aufgehalten hatte und die Möglichkeit hatte diese Aufgabe zu sehen, oder ob die Person überhaupt nicht in die Nähe der Aufgabe gekommen war und deshalb keine Chance hatte, die Aufgabe zu erledigen. Die Abfall-Items im Kasten des Büros wurden vor allem als ein Indikator für das Interaktionsverhalten mit der virtuellen Umwelt genutzt. Es zeigten sich hierbei keinen Unterschied zwischen den Teilnehmern in Bezug auf die

3D-Computerspiel Erfahrung. Des Weiteren wurde der Abfall auch als ein Indiz verwendet, ob die Aufgabe, in der Fabrik Ordnung zu schaffen, von den Testpersonen verstanden wurde. Für zukünftige Weiterentwicklungen des Tests könnten, die Arten von Abfällen und die Orte, an denen diese zu finden sind, sowie die Schwierigkeit der Items größer variieren.

Aufgabe 2

Die Arbeit am Fließband wurde von den Testpersonen weitestgehend erwartungsgemäß bearbeitet. Die Genauigkeit der Fehler-Aussortierung nahm mit der Dauer der Aufgabe aufgrund von Übungseffekten zu. Relativ schnell wurde das jeweilige Niveau der Testperson erreicht, welches keinen Leistungseinbruch mit der Dauer der Arbeit am Fließband verzeichnete. Die Fehlertypen konnten, wie bei der Testkonstruktion angenommen, dem vorgefertigten Schwierigkeitschema zugeordnet werden. Die Selbsteinschätzung zu Beginn der Aufgabe wurde zumeist eher konservativ abgegeben, der Großteil der Teilnehmer produzierte um einiges mehr, als sie geschätzt hatten, vielleicht wurde die Aufgabe zu Beginn als deutlich schwieriger eingeschätzt, als sie den Testpersonen schlussendlich gefallen war. Obwohl die monotone Arbeit am Fließband seitens der Teilnehmer im allgemeinen als ermüdend angesehen wurde. Die Grafik zur Selbsteinschätzung wurde laut einiger Probanden als Entscheidungshilfe übersehen. Die Lupenfunktion wurde selten bis gar nicht eingesetzt. Da die Verwendung der Lupe mit einem garantierten Zeitverlust einherging, könnte der Gewinn an Exaktheit in Anbetracht der Menge der zu bearbeitenden Tafeln als unverhältnismäßig betrachtet worden sein. Vielleicht hätte eine spezifischere Instruktion die Verwendung der Lupe gefördert.

Der *Exaktheits-Score* als Grad der Genauigkeit mit welcher die richtige Entscheidung getroffen wurde, zeigt keinen Zusammenhang mit der Variable Exaktheit aus dem AHA. Die verwendeten Bilder beim Subtest *Flächengrößenvergleichen* im AHA bestehen aus geometrischen Formen bei denen es auch ohne Zeitdruck nicht mit Sicherheit eine richtige Wahl getroffen werden kann. Ein erraten der Antwort kann nicht nur nicht ausgeschlossen werden, sondern wird so eher noch forciert. Im Test VICI ist eine klare Entscheidung und Identifizierung der richtigen Antwort aufgrund der Realitätsnähe leichter zu treffen. Auch wenn Raten auch beim Test VICI nicht ausgeschlossen werden kann, so ist der Wert für Exaktheit leichter nachvollziehbar als jener im Test AHA. Im Rahmen der Konstruktvalidierung fiel die Produkt-Moment-Korrelation des *Exaktheit-Score* des VICI mit dem des AHA nicht signifikant aus.

Der Wert für die Impulsivität und Reflexivität ist inhaltlich valide und seine Berechnung transparent und nachvollziehbar. Wie in Grafik 30 veranschaulicht, differenziert dieser Score zwischen Personen eines impulsiven Arbeitsstils und jenen eines reflexiven Arbeitsstils. Jedoch sollten Ausreißer, die sich nicht auf einem Kontinuum zwischen diesen beiden Stilen einordnen lassen eher anhand ihrer Einzelwerte in *Arbeitsgeschwindigkeit* und *Exaktheit* betrachtet werden. Im Rahmen der Konstruktvalidierung fiel die Produkt-Moment-Korrelation des *Impulsivität/Reflexivität Score* des VICI mit dem des AHA nicht signifikant aus.

Aufgabe 3

Aufgabe 3 wurde von den meisten Teilnehmern gut gelöst, das lag unter anderem am System, welches in Aufgabe 1 erstellt werden konnte. Die Aufgabenstellung war sehr klar formuliert und ohne größere Probleme umsetzbar. Die Dimension *Pünktlichkeit* ist mit einem einmaligen Abprüfen nicht ausreichend er-

hoben worden, es bräuchte mehr Möglichkeiten das Verhalten in Bezug auf die Pünktlichkeit, aber zum Beispiel auch auf die Selbstverwaltungsfähigkeit oder bei weiteren Aufgabenkollisionen zu beobachten. Für zukünftige Revisionen des Tests wurde angedacht, öfter Bestellungen durchzuführen, diese komplexer zu gestalten sowie nicht angekündigte oder ungeplante Unterbrechungen einzubinden.

Allgemeine Anmerkungen

Der Test VICI wurde im Feedback durch die Testleiter von einem Großteil der Testpersonen als interessanter Ansatz für die Diagnostik, „innovatives Instrument“ oder als „nette Abwechslung“ beschrieben, Generell kann gesagt werden, dass der Test positiv bei den Probanden ankam. Besonders die freien Explorationsmöglichkeiten wurden seitens einiger Testpersonen als positiv beschrieben. Einige Interaktionsmöglichkeiten wurden den Aufzeichnungen des Programms zufolge kaum oder nie genutzt, genau diese sogenannten *Display-* und *Response system realism* (Schiff, Arnone und Cross; 1994, zitiert nach Funke & Reuschenbach, 2011) sind jedoch für eine virtuelle Realität von besonderer Bedeutung. Der innovative Ansatz dieses Tests liegt in der offenen Bearbeitung der Aufgaben, nur wenn keine Grenzen gesetzt werden, kann wahres Verhalten beobachtet werden. Sobald man, wie zum Beispiel bei Fragebögen, eine Leitlinie oder Auswahlmöglichkeiten vorgibt, ist das wahre Verhalten bereits überdeckt. Bei test-erprobten Versuchspersonen, wie Psychologie-Studierende es sind, ist ein grenzenloser virtueller Freiraum mit einer großen Menge an Interaktionsmöglichkeiten möglicherweise auch etwas verwirrend. Wer es gewöhnt ist, sehr strikt abgesteckten Grenzen bei psychologisch-diagnostischen Tests zu folgen sowie Anweisungen und Zeitbeschränkungen zu bekommen, ist möglicherweise schnell verloren, wenn nach Hinweisen, wie zum Beispiel der zu absolvie-

renden Bestellung, gesucht werden muss. Ein anderer Grund für leichte Ratlosigkeit kann auch in der Menge an zu verarbeitender Information liegen. Beim Test VICI wurde versucht möglichst kurze und prägnante Instruktionen zu geben. Dabei kann es, wie bei Fragebögen und anderen klassischen Instrumenten dazu kommen, dass Anweisungen nicht oder nicht vollständig gelesen werden oder schnell wieder vergessen werden.

Kritik und Ausblick für die Zukunft

In dieser Arbeit konnte gezeigt werden, dass eine Anwendung virtueller Umgebungen für zukünftige Entwicklungen innerhalb der psychologischen Diagnostik durchaus als praktikabler Ansatz mit vielen Vorteilen, wie der ökologischen Validität des Verfahrens oder der Möglichkeit zur unbeeinträchtigenden Monitoring des Verhaltens betrachtet werden kann. Es bleibt jedoch immer noch ungeklärt, inwieweit Verhalten in einer virtuellen Umgebung auf reales Verhalten übertragbar ist. Ebenfalls mangelt es derzeit, ähnlich wie bei den bisher bereits entwickelten Objektiven Persönlichkeitstests, noch an geeigneten Validierungsstudien (Amelang et al., 2011). Auch bleibt ungeklärt, inwieweit sich die gewonnenen Erkenntnisse auch auf eine weniger homogene Stichprobe übertragen lassen. Anzumerken ist, dass diese kleine Stichprobe für eine Konstruktvalidierung nicht groß genug ist. Des Weiteren können die gefundenen Ergebnisse nicht verallgemeinert werden, da die Stichprobe zum Großteil aus Studierenden bestand und die Computerkenntnisse bei allen Teilnehmern gegeben waren. Von einer generellen Machbarkeit kann ausgehend von den in dieser Arbeit gewonnenen Erkenntnissen ausgegangen werden. In Zukunft wird sich zeigen, inwieweit dieser Ansatz optimiert werden kann und wo sich Anwendungsbereiche finden lassen. Für zukünftige Revisionen des Test VICI soll besondere Aufmerksamkeit auf eine kurze, aber genaue Instruktion gelegt werden, um die wichtigen

Informationen klar zu kommunizieren, ohne dabei den Rahmen zu sprengen. Wie bereits erwähnt, ist die Dimension *Pünktlichkeit* mit einem einmaligen Abprüfen nicht valide zu erheben, es bräuchte mehr Möglichkeiten das Verhalten in Bezug auf die Pünktlichkeit, aber zum Beispiel auch auf die Selbstverwaltungsfähigkeit oder bei weiteren Aufgabenkollisionen zu beobachten. Außerdem wurde angedacht, öfter Bestellungen durchzuführen, diese komplexer zu gestalten, sowie nicht angekündigte oder ungeplante Unterbrechungen einzubinden. Auch in Bezug auf die Schwierigkeit der Abfall-Items oder der zu sortierenden Items ist eine variierende Gestaltung anzudenken, um zu leichte Items und die damit in Verbindung stehenden Deckeneffekte zukünftig zu vermeiden.

Teil IV

Anhang

Korrelationsmatrix zur Konstruktvalidierung

```
----- Korrelationskoeffizienten NEO Pi-R und VICI -----  
// ----- //  
Kurzezeichnung der verwendeten Variablen des Tests NEO Pi-R:  
NeoC1 = Ordnungsliebe  
NeoC2 = Pflichtbewusstesein  
NeoC3 = Leistungsstreben  
NeoC4 = Selbstdisziplin  
NeoC5 = Besonnenheit  
SumNEO = Gesamtscore  
//-----//
```

	NeoC1	NeoC2	NeoC3	NeoC4	NeoC5	SumNEO
Ordnungsliebe.Score	0.074	0.011	-0.023	-0.069	-0.001	-0.032
Trash.Score	-0.026	-0.028	-0.038	-0.057	-0.060	-0.069
Kartonbewegung.zu.T1	-0.055	-0.116	-0.120	-0.306	-0.128	-0.253
Entscheidungsfreude	0.000	-0.131	-0.040	-0.032	-0.027	-0.058
Exaktheit.Score	0.037	-0.034	-0.060	0.066	-0.046	-0.030
Arbeitsgeschwindigkeit	-0.083	0.025	-0.057	-0.028	-0.071	-0.065
IR.Score	0.003	0.052	-0.010	0.049	-0.001	0.015
Produktivität	0.193	0.141	0.093	-0.057	0.008	0.092
Zieldiskrepanz	0.098	0.029	0.058	-0.005	-0.054	0.015
MehrPR	-0.107	-0.196	-0.051	-0.064	-0.229	-0.210

Anspruchsniveau	0.195	0.208	0.106	-0.072	0.136	0.173
Geschätzte.Produktivität	0.181	0.213	0.051	-0.080	0.121	0.150
Kartonbewegung.zu.T2	-0.139	0.213	0.023	0.059	-0.115	0.018
Bestellung.Richtig	0.091	0.111	0.065	-0.137	0.157	0.061
Bestellung.Falsch	0.000	0.019	0.008	0.205	-0.117	0.038
System	0.068	-0.028	-0.047	-0.106	0.025	-0.059

----- Signifikanzlevel (p < .05) NEO Pi-R und VICI -----
// ----- //

	NeoC1	NeoC2	NeoC3	NeoC4	NeoC5	SumNEO
Ordnungsliebe.Score	0.504	0.921	0.833	0.535	0.996	0.776
Trash.Score	0.818	0.803	0.732	0.610	0.592	0.536
Kartonbewegung.zu.T1	0.619	0.295	0.279	0.005	0.251	0.021
Entscheidungsfreude	0.997	0.237	0.722	0.775	0.807	0.601
Exaktheit.Score	0.737	0.760	0.592	0.554	0.681	0.786
Arbeitsgeschwindigkeit	0.456	0.822	0.610	0.798	0.525	0.562
IR.Score	0.981	0.642	0.927	0.657	0.993	0.894
Produktivität	0.080	0.203	0.404	0.610	0.942	0.409
Zieldiskrepanz	0.378	0.796	0.605	0.961	0.631	0.896
MehrPR	0.335	0.076	0.650	0.566	0.037	0.056
Anspruchsniveau	0.078	0.060	0.339	0.519	0.219	0.119
Geschätzte.Produktivität	0.102	0.054	0.649	0.472	0.277	0.176
Kartonbewegung.zu.T2	0.225	0.061	0.841	0.609	0.317	0.876
Bestellung.Richtig	0.414	0.320	0.561	0.218	0.155	0.581
Bestellung.Falsch	0.999	0.865	0.940	0.063	0.292	0.733
System	0.539	0.805	0.670	0.341	0.824	0.594

```

----- Korrelationskoeffizienten AHA und VICI -----
// ----- //
Kurzezeichnung der verwendeten Variablen des Tests AHA:
IR = Impulsivität / Reflexivität
EX = Exaktheit
EN = Entschlussfreudigkeit
LN = Leistungsniveau
AN = Anspruchsniveau
//-----//

```

	IR	EX	EN	LN	AN
Ordnungsliebe.Score	0.113	0.089	0.028	0.122	0.031
Trash.Score	-0.027	0.070	0.135	-0.004	0.064
Kartonbewegung.zu.T1	0.033	0.003	-0.101	-0.115	-0.084
Entscheidungsfreude	-0.054	0.112	0.184	0.088	-0.046
Exaktheit.Score	0.114	-0.016	-0.210	-0.041	0.102
Arbeitsgeschwindigkeit	0.095	-0.097	-0.208	-0.062	-0.033
IR.Score	0.088	-0.036	-0.201	-0.044	0.090
Produktivität	0.153	0.223	-0.036	0.132	0.072
Zieldiskrepanz	0.192	0.193	-0.118	0.115	-0.010
MehrPR	0.103	0.156	0.008	0.218	-0.246
Anspruchsniveau	-0.061	0.120	0.157	0.035	0.152
Geschätzte.Produktivität	-0.060	0.057	0.137	0.043	0.152
Kartonbewegung.zu.T2	0.029	0.134	0.063	0.046	0.025
Bestellung.Richtig	0.137	0.018	-0.072	0.253	-0.181
Bestellung.Falsch	-0.055	0.080	0.036	-0.170	0.062
System	0.141	0.007	-0.152	0.128	-0.005

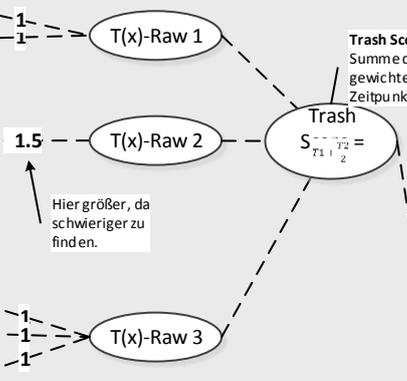
----- Signifikanzlevel(p < .05) AHA und VICI -----					
// ----- //					
	IR	EX	EN	LN	AN
Ordnungsliebe.Score	0.308	0.424	0.801	0.275	0.787
Trash.Score	0.808	0.531	0.225	0.972	0.579
Kartonbewegung.zu.T1	0.769	0.976	0.363	0.302	0.465
Entscheidungsfreude	0.631	0.311	0.095	0.433	0.688
Exaktheit.Score	0.304	0.883	0.056	0.717	0.376
Arbeitsgeschwindigkeit	0.392	0.383	0.060	0.582	0.774
IR.Score	0.430	0.747	0.069	0.692	0.433
Produktivität	0.168	0.043	0.745	0.238	0.529
Zieldiskrepanz	0.083	0.081	0.287	0.304	0.930
MehrPR	0.356	0.160	0.942	0.050	0.030
Anspruchsniveau	0.581	0.279	0.156	0.756	0.183
Geschätzte.Produktivität	0.590	0.611	0.216	0.699	0.185
Kartonbewegung.zu.T2	0.803	0.241	0.581	0.691	0.837
Bestellung.Richtig	0.219	0.872	0.519	0.022	0.113
Bestellung.Falsch	0.618	0.475	0.749	0.126	0.592
System	0.204	0.952	0.171	0.251	0.962

Zusammensetzung der Testkennwerte im Test VICI

Bildung des Trash.Scores sowie des Ordnungsliebe.Scores im Test VICI

Abfall in der Produktionshalle und im Büro

Plastikflasche unter dem Tisch im Büro	Mist Typ 7
Abfall neben dem Abfalleimer	Mist Typ 1
Zerkümmertes Papier im Kasten	Mist Typ 4&5
Plastikflasche im Kasten	
Zerkümmertes Papier im Abholbereich	Mist Typ 3
Dose und Plastikflasche neben dem Fließband	Mist Typ 6
Leere Dose neben dem Ausgang	Mist Typ 2



Trash Score:
Summe des entdeckten Abfalls gewichtet, abhängig vom Zeitpunkt des Aufhebens.

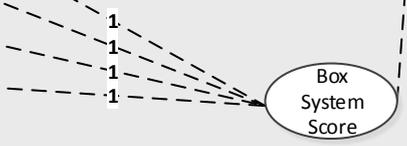
$$S_{Trash} = \sum_{i=1}^n T(x)_{i,2}$$

Hier größer, da schwieriger zu finden.

Ordnungsliebe Score

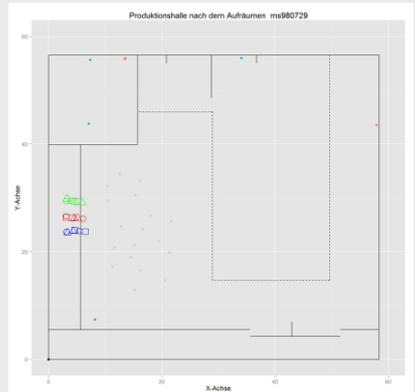
Sortierung der Kartons

Kartons gruppiert nach Größe [Ja/Nein]	Box_grouped by_size
Kartons gruppiert nach Sorte [Ja/Nein]	Box_grouped by_type
Gibt es ein erkennbares System [Ja/Nein]	Box_system
Qualität der Gruppierung	Sauber. gruppiert
Gibt es eine Art des Systems	System.Art
Summe bewegter Kartons zu T1	BoxT1
Summe gedrehte Kartons T1	BoxTwistT1
Summe bewegter Kartons zu T2	BoxT2
Summe gedrehte Kartons T2	BoxTwistT2

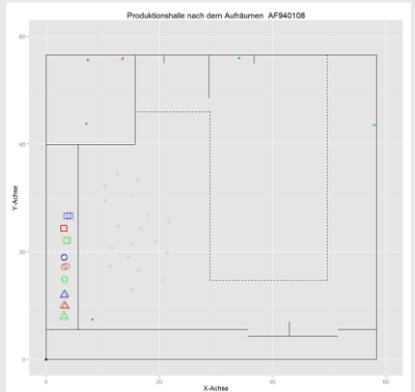


Legende

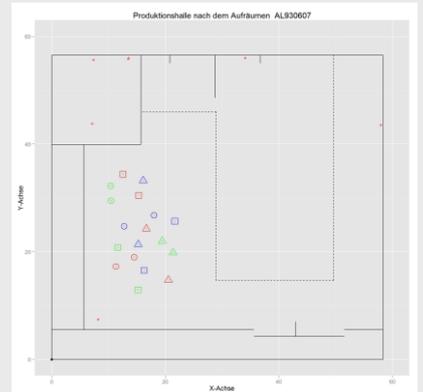
Sorten	Größen
Erdbeere	150
Haselnuss	100
Alpenmilch	50



Gruppirt nach Sorte



Gruppirt nach Stückzahl und Sorte



Unsortiert

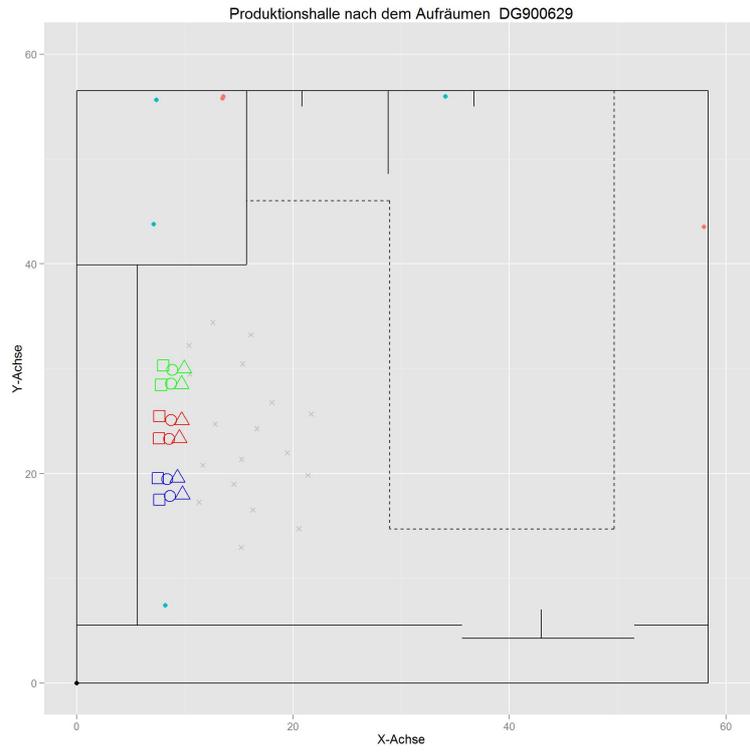


Abbildung 34: Beispiel eines Ordnungssystems

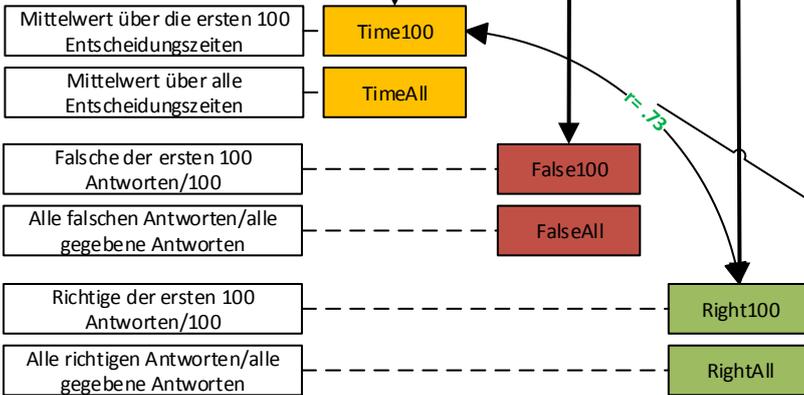
Schokoladentafeln

Original	Entscheidungszeiten	Fehler	Richtig
Anzahl [Typ 1]	Time [Typ 1]	Fehler [Typ 1]	Richtig [Typ 1]
Leicht zu erkennen	Time [Typ 2]	Fehler [Typ 2]	Richtig [Typ 2]
	Time [Typ 3]	Fehler [Typ 3]	Richtig [Typ 3]
	Time [Typ 4]	Fehler [Typ 4]	Richtig [Typ 4]
	Time [Typ 5]	Fehler [Typ 5]	Richtig [Typ 5]
Schwer zu erkennen	Time [Typ 6]	Fehler [Typ 6]	Richtig [Typ 6]
	Time [Typ 7]	Fehler [Typ 7]	Richtig [Typ 7]
	Time [Typ 8]	Fehler [Typ 8]	Richtig [Typ 8]
	Time [Typ 9]	Fehler [Typ 9]	Richtig [Typ 9]
	Time [Typ 10]	Fehler [Typ 10]	Richtig [Typ 10]

Gewichte:
Die durchschnittliche Betrachtungsdauer bis zu einer richtigen Entscheidung dienen als Gewichtungen für die Schwierigkeit der einzelnen Tafeln.

Gewichtung	
	2.02 Sek. Typ 10
	1.88 Sek. Typ 9
	1.66 Sek. Typ 6
	1.41 Sek. Typ 7
	1.38 Sek. Typ 8
	1.36 Sek. Typ 5
	1.34 Sek. Typ 4
	1.29 Sek. Typ 1
	1.24 Sek. Typ 3
	1.16 Sek. Typ 2

Schwierigkeit den Fehler zu entdecken

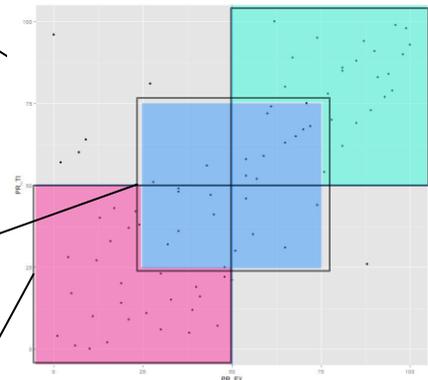


Gemischter Arbeitstyp:
Mittlere Betrachtungszeiten bei gleichzeitig mittlerer Exaktheit

Impulsiver Arbeitstyp:
Kurze Betrachtungszeiten bei gleichzeitig geringer Exaktheit

Reflexiver Arbeitstyp:
Längere Betrachtungszeiten bei gleichzeitig hoher Exaktheit

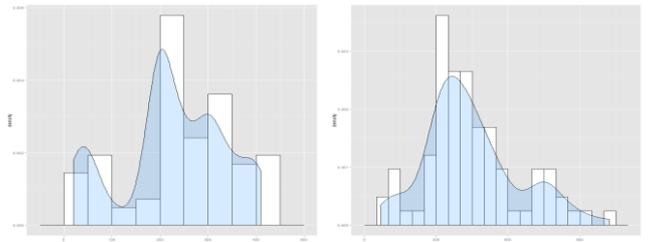
Jelänger die Betrachtungsdauer, desto weniger Fehler wurden gemacht. Hier in Prozenträngen angegeben.



Positiver Zusammenhang zwischen Betrachtungsdauer und richtiger Antwort

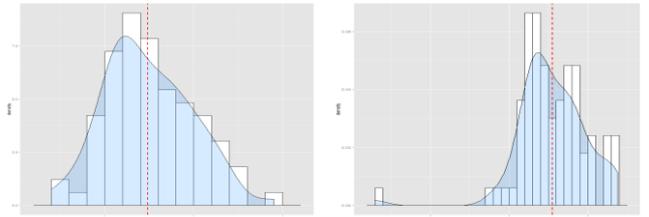
Anspruchsniveau und Beharrlichkeit (Geschätzt & Realisiert)

Zieldiskrepanz	DiffProEst
Summe produzierte Tafeln	Produced
Summe geschätzt	Estimate
Mehrproduktion in Prozent	MehrPR



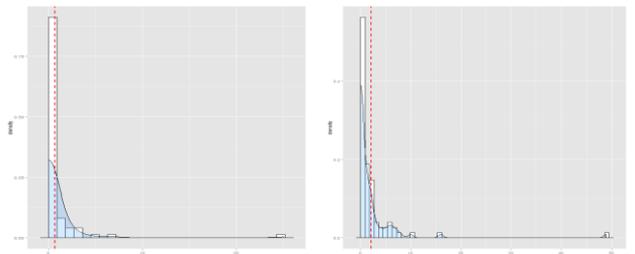
Exaktheit

Richtige Antworten der ersten 100 Tafeln	Exaktheit100
Richtige Antworten gesamt Produzierte Tafeln	ExaktheitAll
Summe richtiger Antworten der ersten 100 Tafeln nach Schwierigkeit gewichtet	Exaktheit_gew



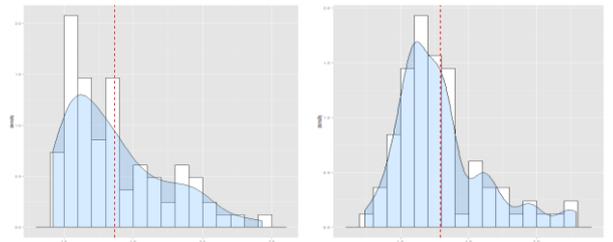
Verwendung der Lupe

Summe Lupe verwendet (Sample)	Lupe100
Summe Lupe verwendet	LupeAll



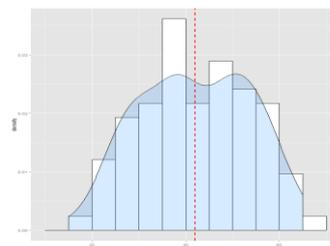
Durchschnittliche Geschwindigkeit

Mittelwert der ersten 100 getroffenen Entscheidungen	Time100
Mittelwert aller getroffenen Entscheidungen	TimeAll



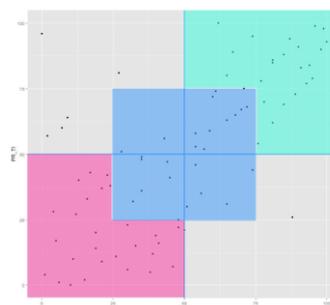
Entschlussfreudigkeit

Anzahl produzierter Tafeln innerhalb der ersten 60 Entscheidungssekunden	D60sec
---	--------



Impulsivität vs. Reflexivität

Z-transformierte Prozentränge des Exaktheit.Scores	EX_Z
Z-transformierte Prozentränge des Time100	TI_Z
EX_Z+TI_Z	IR_Z



Hinweis:
Positives Vorzeichen = Reflexiver Typ
Negatives Vorzeichen = Impulsiver Typ
Je höher der Wert, desto stärker ist der Stil ausgeprägt.

Michael Mühlegger

Curriculum Vitae

Staatsangehörigkeit: Österreich

Geburtsdatum: 11. September 1984

Ausbildung

2005–2015 **Studium der Psychologie**, *Universität*, Wien, Österreich.

Erfahrung

2012– **Angestellter**, *Schuhfried GmbH*, Mödling, Österreich.

2013–2014 **Praktikum**, *Datenwerk innovationsagentur GmbH*, Wien, Österreich.

IT - Kompetenzen

OS Linux, Unix, Windows
Programming Python 2.7, Python 3, XML
Scientific R -Statistics, SPSS, LyX
Graphics Photoshop, Gimp 2.8
3D design (Blender 3D)

Modeling, Texturing, Game
Engine, Rendering, Anima-
tions, Logics, Scripting

Originalitätserklärung

Ich, Michael Mühlegger, versichere, dass ich die Diplomarbeit selbständig verfasst habe, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und mich auch sonst keiner unerlaubten Hilfe bedient habe.

Des Weiteren habe ich diese Diplomarbeit bisher weder im In- noch im Ausland in irgendeiner Form als Prüfungsarbeit vorgelegt.

Wien, am 05.Mai 2015

Unterschrift

Abbildungsverzeichnis

1	Verfälschung eines Persönlichkeitsinventars in einer imaginierten Auswahl-situation nach Ziegler et al. (2010),Tab.1 ; Abgebildet sind hier die Abweichungen der neutralen Situation in Standardabweichungen von der Kontrollbedingung.	25
2	Datenblatt des Test VICI sowie die Einbettung der Facette Gewissenhaftigkeit des NEOPI- R (Borkenau & Ostendorf, 2008) . . .	44
3	Produktionshalle im Überblick	45
4	Screening zur Auge-Hand Koordination	46
5	Instruktion "Visuelle Orientierung"	47
6	Training "Visuelle Orientierung"	48
7	Instruktion "Navigation im Raum"	49
8	Training "Visuelle Orientierung"	49
9	Instruktion "Interaktion mit Objekten"	51
10	Training "Interaktion mit Objekten"	52
11	Ansicht auf das Fließband und die dazugehörigen Bedienelemente	53
12	Aktive "Vergrößerung"	54
13	Instruktion für den Abholbereich	56
14	virtueller PC im Büro"	57
15	Instruktion zur ersten Aufgabe	58
16	Kartons, die verstreut in der Halle liegen	59
17	Abfall, neben statt im Abfalleimer	60
18	Instruktion zur zweite Aufgabe	61
19	Schätzung der Tagesproduktion	61
20	Instruktion zur dritte Aufgabe	62
21	Test 3D (Schuhfried, 2004)	71

22	Untertest “Flächengrößen vergleichen” (Kubinger & Ebenhöf, 2002)	73
23	Untertest “Symbole kodieren” (Kubinger & Ebenhöf, 2002)	74
24	Box-Plot: Altersverteilung in Jahren in Bezug auf das Geschlecht	82
25	Stichprobenbeschreibung bezüglich des Bildungsgrads	83
26	Histogramm zur Auge-Hand Koordination	85
27	Verteilung der wichtigsten Kenngrößen in Aufgabe 1	86
28	Verteilung der Entscheidungsfreudigkeit	88
29	Verteilung Aufgabe 2	89
30	Streudiagramm zum Score “Impulsivität-Reflexivität”	90
31	Balkendiagramm zum Grad der 3D Spiele- Erfahrung	93
32	Screeplot	107
33	Faktorladungen	108
34	Beispiel eines Ordnungssystems	130

Tabellenverzeichnis

1	Das Kategoriensystem der Gewissenhaftigkeit nach Westhoff (2013), Tabelle 1 (S. 206)	20
2	Stichprobenbeschreibung bezüglich der Altersverteilung	81

Literatur

Allain, P., Foloppe, D. A., Besnard, J., Yamaguchi, T., Etcharry-Bouyx, F., Le Gall, D., ... Richard, P. (2014). Detecting everyday action deficits in alzheimers disease using a nonimmersive virtual reality kitchen. *Journal of the International Neuropsychological Society*, 20 (05), 468–477.

- Amelang, M. (2003). Differentielle Psychologie. In K. D. Kubinger & R. S. Jäger (Hrsg.), *Schlüsselbegriffe der psychologischen Diagnostik* (S. 96-102). Beltz.
- Amelang, M., Bartussek, D., Stemmler, G. & Hagemann, D. (2006). *Differentielle Psychologie und Persönlichkeitsforschung*. W. Kohlhammer Verlag.
- Amelang, M., Bartussek, D., Stemmler, G. & Hagemann, D. (2011). *Differentielle Psychologie und Persönlichkeitspsychologie*, 7. Kohlhammer.
- Anderson, P., Rothbaum, B. O. & Hodges, L. F. (2003). Virtual reality exposure in the treatment of social anxiety. *Cognitive and Behavioral Practice*, 10 (3), 240–247.
- Barrick, M. R., Mount, M. K. & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? In (Bd. 9, S. 9-30). Wiley Online Library.
- Becker, P. (2003). Persönlichkeitsfragebogen. In K. D. Kubinger & R. S. Jäger (Hrsg.), *Schlüsselbegriffe der psychologischen Diagnostik* (S. 332-337). Beltz.
- Bente, G., Krämer, N. C. & Petersen, A. (2002). Virtuelle Realität als Gegenstand und Methode in der Psychologie. In G. Bente (Hrsg.), *Virtuelle Realitäten* (S. 1–31). Hogrefe Verlag.
- Berth, H., Goldschmidt, S., Ostendorf, F. & Angleitner, A. (2006). NEO-PI-R. NEO-Persönlichkeitsinventar nach Costa und McCrae. Revidierte Fassung. *Diagnostica*, 52 (2), 95–99.
- Borkenau, P. & Ostendorf, F. (1993). NEO-Fünf-Faktoren Inventar (NEO-FFI) [Software-Handbuch]. Göttingen.
- Borkenau, P. & Ostendorf, F. (2008). NEO-Fünf-Faktoren-Inventar (NEO-FFI) nach Costa & McCrae (Bd. 2) [Software-Handbuch]. Göttingen: Hogrefe.
- Campbell, D. T. & Fiske, D. W. (1959). Convergent and discriminant validation

- by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.
- Corker, K. S., Oswald, F. L. & Donnellan, M. B. (2012). Conscientiousness in the classroom: A process explanation. *Journal of personality*, 80 (4), 995–1028.
- Costa, P. & McCrae, R. (1992). Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor-Inventory (NEO-FFI) Professional Manual. [Software-Handbuch]. Odessa.
- DeYoung, C. G., Hirsh, J. B., Shane, M. S., Papademetris, X., Rajeevan, N. & Gray, J. R. (2010). Testing predictions from personality neuroscience: Brain structure and the big five. *Psychological Science*, 21 (6), 820-828.
- Emmelkamp, P., Krijn, M., Hulsbosch, A., De Vries, S., Schuemie, M. & Van der Mast, C. (2002). Virtual reality treatment versus exposure in vivo: a comparative evaluation in acrophobia. *Behaviour Research and Therapy*, 40 (5), 509–516.
- Fahrenberg, J., Hampel, R. & Selg, H. (2010). *FPI-R: Freiburger Persönlichkeitssinventar; Manual*. Hogrefe.
- Field, A. P. (2009). *Discovering statistics using spss; (and sex and drugs and rock 'n' roll)* (3. ed. Aufl.). Los Angeles, Calif. [u.a.]: Sage. (Hier auch später erschienene, unveränderte Nachdrucke (2010, 2011))
- Frebort, M. (2002). *Evaluation einer Batterie Objektiver Persönlichkeitstests zur Auswahl von TierpflegeschülerInnen* (Unveröffentlichte Diplomarbeit). Universität Wien.
- Funke, J. & Reuschenbach, B. (2011). Einsatz technischer Mittel in der psychologischen Diagnostik. In L. Hornke, M. Amelang & M. Kersting (Hrsg.), *Leistungs-, Intelligenz- und Verhaltensdiagnostik* (S. 595-667). Göttingen: Hogrefe.
- Hartig, J., Frey, A. & Ketzler, A. (2003). Modifikation des Computerspiels Quake

- III Arena zur Durchführung psychologischer Experimente in einer virtuellen 3D-Umgebung. *Zeitschrift für Medienpsychologie*, 15 (4), 149–154.
- Hauser, M. (2011). *Extremgruppenvalidierung von LAMBDA-neu* (Unveröff. Dipl. Arbeit,). Universität Wien.
- Höft, S. & Funke, U. (2001). Simulationsorientierte Verfahren der Personalauswahl. In *Lehrbuch der Personalpsychologie* (S. 135-173). Göttingen: Hogrefe.
- Hossiep, R. & Paschen, M. (2003). *BIP: das Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung*. Hogrefe.
- Ifcic, K. (2012). *Ein Beitrag zur Konstruktvalidierung einer Kurzversion des WSP aus dem Wiener Self-Assessment Psychologie* (Unveröff. Dipl. Arbeit,). Universität Wien.
- Ijsselstein, W., Harper, B., Group, T. et al. (2000). Virtually there? a vision on presence research. *Proceedings of Presence 2002, Oct., 9, 11*, 245–259.
- Jackson, D. N. (1974). *Personality research form manual*. Research Psychologists Press.
- Jungermann, H., Pfister, H. & Fischer, K. (1998). *Die Psychologie der Entscheidung*. Heidelberg: Spektrum.
- Khorrandel, L. & Kubinger, K. D. (2006). The effect of speediness on personality questionnaires: an experiment on applicants within a job recruiting procedure. *Psychology Science*, 48 (3), 378-397.
- Kirchler, E. (2008). *Arbeits- und Organisationspsychologie* (2., korr. Aufl. Aufl.). Wien: Facultas WUV.
- Krug, W. T. (1819). *System der praktischen philosophie* (Bd. 3). Autor.
- Kubinger, K. D. (2002). On faking personality inventories. *Psychologische Beiträge*, 44 (1), 10.
- Kubinger, K. D. (2009). *Psychologische Diagnostik Theorie und Praxis psycho-*

- logischen Diagnostizierens*. Göttingen: Hogrefe.
- Kubinger, K. D. & Ebenhöf, J. (2002). *Arbeitshaltungen Kurze Testbatterie* (Version 26.00. Test: Software und Manual (Verf.dieses Manuals: T. Karner & M. Sommer). Aufl.). Mödling: Schuhfried.
- Kubinger, K. D. & Hofmann, K. (2001). Herkömmliche Persönlichkeitsfragebogen und Objektive Persönlichkeitstests im Wettstreit um (Un-) Verfälschbarkeit. *Report Psychologie*, 26, 298-304.
- Kubinger, K. D. & Litzenberger, M. (2003). Zur Validität der Objektiven Persönlichkeits-Test-Batterie <Arbeitshaltungen>. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 24 (2), 119-133.
- Kuhlen, T. (2014). Virtuelle Realität als Gegenstand und Werkzeug der Wissenschaft. In S. Jeschke, L. Kobbelt & A. Dröge (Hrsg.), *Exploring virtuality* (S. 133-147). Springer.
- LaViola, J. (2000). A discussion of cybersickness in virtual environments. *SIGCHI Bull.*, 32 (1), 47–56. Zugriff auf <http://doi.acm.org/10.1145/333329.333344> doi: 10.1145/333329.333344
- Lievens, F. & Ones, S., D. S. and Dilchert. (2009). Personality scale validities increase throughout medical school. *Journal of Applied Psychology*, 94 (6), 1514-1535.
- Loomis, J. M., Blascovich, J. J. & Beall, A. C. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behavior Research Methods, Instruments, & Computers*, 31 (4), 557–564.
- Lord, W. (2011). *Das NEO-Persönlichkeitsinventar in der berufsbezogenen Anwendung: Interpretation und Feedback*. Hogrefe Verlag.
- Minbashian, A., Wood, R. E. & Beckmann, N. (2010). Task-contingent conscientiousness as a unit of personality at work. *Journal of Applied Psychology*, 95 (5), 793.

- Möseneder, D. & Ebenhöf, J. (1996). ILICA, Ein Simulationstest zur Erfassung des Entscheidungsverhaltens. *Swets, Frankfurt am Main*.
- Mummendey, H. & Grau, I. (2008). *Die Fragebogenmethode* (5. Aufl.). Hogrefe (Göttingen).
- Nerdinger, F., Blickle, G. & Schaper, N. (2008). *Arbeits-und Organisationspsychologie: mit 32 Tabellen*. Springer.
- Opdyke, D., Williford, J. S. & North, M. (1995). Effectiveness of computer-generated (virtual reality) graded exposure in the treatment of acrophobia. *Am J Psychiatry*, 1 (152), 626–28.
- Ortner, T., Horn, R., Kersting, M., Krumm, S., Kubinger, K. D., Proyer, R. T., . . . others (2007). Standortbestimmung und Zukunft Objektiver Persönlichkeitstests. *Report Psychologie*, 32 (2), 60–69.
- Ortner, T., Kubinger, K., Schrott, R., A. and Radinger & Litzemberger, M. (2006). Belastbarkeits-Assessment: computerisierte Objektive Persönlichkeits-Testbatterie (BAcO) [Software-Handbuch]. Mödling.
- Ortner, T. & Kubinger, K. D. (2006). BAcO-D: Eine Objektive Persönlichkeitstest-batterie zur Erfassung von Belastbarkeit. In T. M. Ortner, R. T. Proyer & K. D. Kubinger (Hrsg.), *Theorie und Praxis Objektiver Persönlichkeitstests* (S. 133-142). Bern: Huber.
- Ortner, T., Proyer, R. T. & Kubinger, K. D. (Hrsg.). (2006). *Theorie und Praxis Objektiver Persönlichkeitstests*. Bern: Huber.
- Ostendorf, F. & Angleitner, A. (2004). NEO- Persönlichkeitsinventar nach Costa und McCrae Revidierte Fassung [Software-Handbuch]. Göttingen.
- 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: Huber.
- Roberts, B. W., Walton, K. E. & Bogg, T. (2005). Conscientiousness and health

- across the life course. *Review of General Psychology*, 9 (2), 156-168.
- Rothbaum, B. O., Hodges, L., Anderson, P. L., Price, L. & Smith, S. (2002). Twelve-month follow-up of virtual reality and standard exposure therapies for the fear of flying. *Journal of consulting and clinical psychology*, 70 (2), 428.
- Ruf, A. (2010). *Die Bedeutung des Fünf-Faktoren-Modells als umfassendes Persönlichkeitsmodell für die Eignungsdiagnostik*. GRIN Verlag.
- Schmidt, L. (2006). Objektive Persönlichkeitstests in der Tradition Cattells: Forschungslinien und Relativierungen. In T. M. Ortner, R. T. Proyer & K. D. Kubinger (Hrsg.), *Theorie und Praxis Objektiver Persönlichkeitstests* (S. 24-37). Bern: Huber.
- Schmidt, L. R. (1975). *Objektive Persönlichkeitsmessung in diagnostischer und klinischer Psychologie*. Weinheim [u.a.]: Beltz.
- Schmidt-Atzert, L. & Amelang, M. (2012). *Psychologische Diagnostik* (5., vollständig überarbeitete und erweiterte Auflage Aufl.). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Schubert, T., Friedmann, F. & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence*, 10 (3), 266–281.
- Schuhfried. (2004). Räumliches Orientierungsvermögen (kurzbezeichnung 3D) [Software-Handbuch]. Mödling.
- Schuhfried. (2014). Arbeitshaltungen (Kurzbezeichnung AHA) [Software-Handbuch]. Mödling.
- Schuler, H. & Höft, S. (2001). Konstruktorientierte verfahren in der personalauswahl. In H. Schuler (Hrsg.), *Lehrbuch der personalpsychologie* (S. 94-133). Hogrefe.
- Seiwald, B. (2003). Antworttendenzen (response set). In K. Kubinger & R. Jäger (Hrsg.), *Schlüsselbegriffe der psychologischen diagnostik*. Weinheim:

Beltz.

- Skinner, N. S. F. & Howarth, E. (1973). Cross-media independence of questionnaire and objective-test personality factors. *Multivariate Behavioral Research*, 8 (1), 23-40.
- Slater, M. & Usoh, M. (1993). Representations systems, perceptual position, and presence in immersive virtual environments. *Presence: Teleoperators and virtual environments*, 2 (3), 221-233.
- Sutherland, R., De Bruin, G. & Crous, F. (2007). The relation between conscientiousness, empowerment and performance. *Journal of Human Resource Management*, 5 (2), 60-67.
- Terracciano, A., Löckenhoff, C. E., Zonderman, A. B., Ferrucci, L. & Costa, P. T. (2008). Personality predictors of longevity: activity, emotional stability, and conscientiousness. *Psychosomatic medicine*, 70 (6), 621–627.
- van der Linden, W. J. & Glas, C. A. (2000). *Computerized adaptive testing: Theory and practice*. Springer.
- Vogel, S. (2011). *Zum Unterschied von traditionellen psychologisch- diagnostischen Verfahren und Objektiven Persönlichkeitstests* (Unveröffentlichte Diplomarbeit). Universität Wien.
- Washburn, D. A. (2003). The games psychologists play (and the data they provide). *Behavior Research Methods, Instruments, & Computers*, 35 (2), 185–193.
- Westhoff, K., Steinborn, M. & Schurz, A. (2013). Was versteht man unter dem Konstrukt Gewissenhaftigkeit? *Report Psychologie*, 38 (5), 200-208.
- Wikipedia. (2014a, September). *Blender (software)*. Zugriff auf http://de.wikipedia.org/wiki/Blender_%28Software%29/
- Wikipedia. (2014b, September). *Spiel-engine (software)*. Zugriff auf <http://de.wikipedia.org/wiki/Spiel-Engine>

- Witmer, B. G. & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and virtual environments*, 7 (3), 225–240.
- Ziegler, M., Danay, E., Schölmerich, F. & Bühner, M. (2010). Predicting academic success with the big 5 rated from different points of view: Self-rated, other rated and faked. *European Journal of Personality*, 24 (4), 341–355.