



universität
wien

EUcO 2.0

**Entscheiden unter Unsicherheit computerisierter
Objektiver Persönlichkeitstest
Neukonstruktion**

Diplomarbeit

zur Erlangung des Magistergrades der Naturwissenschaften
an der Fakultät für Psychologie
der Universität Wien

eingereicht von
Josef Past

Magister der Naturwissenschaften (Mag. rer. nat.)

Wien, im Jänner 2011

Studienkennzahl:	298
Studienrichtung:	Psychologie
Betreuer:	Professor K. Kubinger

Danksagung

Ich möchte mich an dieser Stelle bei Professor Klaus Kubinger bedanken, dass er mir dieses Diplomarbeitsthema, den EUcO zu verbessern und anschließend neu zu konstruieren, angeboten hat.

Meinen Studienkollegen möchte ich meinen Dank aussprechen für ihre konstruktive Kritik an den ersten Verbesserungsvorschlägen. Insbesondere Elisabeth Quendler, sie hat die entscheidenden Überlegungen für die Umsetzung bei mir wachgerufen.

Dr. Brugger vom Heerespsychologischen Dienst möchte ich danken, dass ich an der Stellungskommission des Österreichischen Bundesheers in Wien Stellungspflichtige Personen testen durfte. Der verantwortlichen Psychologin an der Stellungskommission Mag. Christina Pollak danke ich herzlich für die geschaffenen Rahmenbedingungen, sowie jene Personen die sie und schlussendlich mich dabei unterstützten.

Für die Möglichkeit FH-Studenten im Rahmen einer Lehreinheit zu testen, danke ich Dr. Petra Wagner von der Fachhochschule Linz.

Meinen Dank möchte ich auch den vielen Testpersonen, die freiwillig an der Testung teilnahmen, aussprechen.

Ebenso möchte ich mich bei Clemens Stadler bedanken, der durch die Programmierung des Tests erst die Grundlage dafür geschaffen hat, dass diese Arbeit an Personen vorgegeben werden konnte. Für das Korrekturlesen danke ich herzlich Andreas Krainz, obwohl er sehr beschäftigt war, nahm er sich dafür Zeit.

Meiner Familie gebührt ein ganz besonderer Dank. Sie zeigte immerzu Verständnis, wenn meine Aufmerksamkeit auf das Studium gerichtet war und ich weniger Zeit für sie übrig hatte.

Schließlich möchte ich mich bei meinen Freunden Rüdiger Brandstetter und Thomas Kammerer bedanken, die mich immer wieder ermutigten und motivierten diese Arbeit zu beenden.

Inhaltsverzeichnis

1 Einleitung	7
2 Theoretischer Teil	10
2.1 Objektive Persönlichkeitstest	10
2.1.1 Begriffsbestimmung	10
2.1.2 Entwicklung	10
2.1.3 Unverfälschbarkeit als Zielsetzung	11
2.1.4 Möglichkeiten und Probleme Objektiver Persönlichkeitstests	13
2.1.4.1 Testtheoretische Problemfelder	13
2.1.4.2 Konstruktvalidität	14
2.1.4.3 Möglichkeiten von Objektiven Persönlichkeitstests	15
2.1.5 Computerisierte Vorgabe	17
2.2 Testinstruktion	20
2.2.1 Formale Testinstruktion	20
2.2.1.1 Instruktionsform bezüglich Testdurchführung	21
2.2.1.2 Regeln zur Instruktions- und Itemform	22
2.2.1.3 Testlänge, Testzeit und Zielgruppe	28
2.2.1.4 Antwortform	29
2.2.2 „Verbale“ Verständlichkeit der Instruktion	30
2.3 Methoden der Entscheidung	32
2.3.1 Grundbegriffe	32
2.3.2 Merkmale einer Entscheidung	34
2.3.3 Komponenten einer Entscheidung	38
2.3.3.1 Persönlichkeitsspezifische Komponenten	38
2.3.3.2 Inhaltliche Komponenten	39
2.3.4 Entscheidungsprobleme	41
2.3.5 Entscheiden unter Unsicherheit	42
2.3.5.1 Unsicherheit	42
2.3.5.2 Entscheidungsmodelle	47
2.3.5.3 Entscheidung und Informationsverarbeitung am Bildschirm	49
2.3.6 Wahrscheinlichkeit	50

2.3.6.1 Messung der subjektiven Wahrscheinlichkeit.....	52
2.3.6.2 Das Theorem von Bayes.....	53
2.3.7 Entscheidungshilfen.....	58
2.3.8 Attribution.....	60
2.3.9 Rationalität.....	62
2.3.9.1 Grundbegriff.....	62
2.3.9.2 Rationalität des Menschen.....	64
2.3.10 Zufall, Glück und Geschick.....	65
2.3.11 Spieltheorie.....	66
2.3.12 Zeitdruck.....	68
2.4 Beschreibung EUcO 2.0.....	69
2.4.1 Kritik am EUcO.....	70
2.4.2 Instruktion.....	72
2.4.3 Beispielaufgaben.....	74
2.4.4 Aufgabenmaterial EUcO 2.0.....	79

3 Empirischer Teil.....85

3.1 Fragestellungen und Hypothesen.....	85
3.2 Datenerhebung, Stichprobenbeschreibung und Testdurchführung.....	86
3.3 Ergebnisdarstellung.....	90
3.3.1 Auswertung der Testkennwerte des EUcO 2.0.....	90
3.3.1.1 Rohscore.....	90
3.3.1.2 Entscheidungsgüte.....	93
3.3.1.3 Konsistenz.....	95
3.3.1.4 Verwendung von Hilfen.....	97
3.3.1.5 Attributionsstil einer Person.....	99
3.3.1.6 Einfluss von Zeitdruck.....	100
3.3.1.7 Entscheidungsgeschwindigkeit.....	101
3.3.2 Untersuchung der Mittelwerte in den beiden Teilstichproben.....	104
3.3.3 EUcO 2.0 Ergebnisse im Vergleich zur alten Version.....	113
3.3.4 Beantwortung der Fragestellungen und Hypothesenüberprüfung.....	115
3.4 Ausblick.....	116

4 Zusammenfassung.....	119
5 Literaturverzeichnis.....	122

1. Einleitung

Der Zweck der vorliegenden Arbeit ist die Verbesserung, Korrektur und Weiterentwicklung des EUcO, ein Test zur Erfassung des individuellen Entscheidungsverhaltens, welcher von Radinger (2002) im Rahmen ihrer Diplomarbeit entwickelt und evaluiert wurde. Das Kürzel EUcO steht für „Entscheiden unter Unsicherheit computerisierter Objektiver Persönlichkeitstest“. Im Folgenden wird zwischen EUcO und der überarbeiteten Form EUcO 2.0 unterschieden.

Das inhärente Problem der meisten Persönlichkeitsfragebogen liegt in ihrer Durchschaubarkeit für die Testperson und somit in der bewussten bzw. unbewussten Verfälschung der Ergebnisse. In herkömmlichen Verfahren finden so genannte Lügenskalen Verwendung, welche die Ehrlichkeit der Testperson messen sollen. Solche Skalen sind aber äußerst umstritten, da es Testpersonen sehr leicht fällt, das Bestreben dieser Fragen zu erkennen und eine Tendenz zu bestimmten Antworten entwickeln. Auf Grund der hohen Augenscheinvalidität solcher Skalen wird ihre Sinnhaftigkeit von vielen Testautoren bezweifelt. Unter „Objektiven Persönlichkeitstest“ wird heute mehrheitlich ein eigenes Testgenre verstanden, welches überwiegend auf frühere Überlegungen und Konzepte von Raymond B. Cattell zurückgeht (Ortner, Horn, Kersting, Krumm, Kubinger, Proyer, Schmidt-Atzert, Schuhfried, Schütz, Wagner-Menghin, & Westhoff, 2007). Bei Objektiven Persönlichkeitstests können aus dem Verhalten Rückschlüsse auf die Persönlichkeit abgeleitet werden. In Anlehnung daran haben die Testautoren diesen Weg für die Messung des Entscheidungsverhaltens gewählt, um Verfälschungen und sozial erwünschte Antworttendenzen zu vermeiden.

Zu den wichtigsten Anwendungsbereichen Psychologischer Diagnostik zählen die klinisch-psychologische Diagnostik sowie das Gebiet der Personalselektion.

In vielen Berufssituationen müssen Entscheidungen getroffen werden. Ob die Entscheidung richtig war, kann oft nicht unmittelbar festgestellt werden. Da Entscheidungen in komplexen Umwelten meist einer bestimmten Unsicherheit

unterliegen, können deren Konsequenzen verzögert eintreten und schwerwiegende Folgen aufweisen. Der EUcO 2.0 kann als aufschlussreiches Instrument in der Personalselektion Verwendung finden, speziell zur Auswahl in Berufssparten, in denen Unsicherheiten im Entscheidungsstadium existieren.

In Entscheidungssituationen hat der Entscheider mehrere Möglichkeiten, mindestens aber zwei Alternativen zur Auswahl. Ist nun die Wahl auf eine Alternative gefallen, muss dem Entscheider bewusst sein, dass die gewählte Option nicht „hundertprozentig“ richtig sein kann und die Auswirkungen der nicht gewählten Alternativen zu berücksichtigen sind. Diese Einschätzung ist im medizinischen, wirtschaftlichen, militärischen usw. Umfeld von enormer Bedeutung.

Die Testung wird mittels Computer durchgeführt, zumal mit Hilfe der computerisierten Vorgabe die Anforderungen an die Gütekriterien für einen „guten“ Test nach Lienert und Raatz (1998) optimal erfüllt werden können

Verschiedenen Verfahren, welche angeben Entscheidungsverhalten zu messen bzw. beabsichtigen dieses zu messen, wurden von Radinger (2002) diskutiert und sind dort nachzulesen.

Eine Teilaufgabe beim EUcO besteht darin, dass die Testperson Wahrscheinlichkeiten schätzen muss, davon abhängig, wie sich die verschieden farbigen Kugeln in den beiden Behältern zusammensetzen. Diese wurde außergewöhnlich oft überschätzt und als wesentliche Schwäche kritisiert.

Ausgangspunkt von Überschätzungen sind Heuristiken, welche aus vereinfachten Annahmen bestehen, um damit ein Problem zu lösen, ohne dass alle Informationen vorhanden sind und ausreichend Zeit verfügbar ist. In vielen Experimenten konnten Heuristiken als Problemlösungsansätze nachgewiesen werden, speziell Kahneman, Slovic und Tversky (1982) und Gigerenzer (2001) widmeten sich diesem Thema.

Als weitere Schwäche gilt die äußerst lange und zum Teil sehr unverständliche Instruktion, welche „von den Teilnehmern mit höherer Bildung nicht immer auf Anhieb erfasst wird“ (Radinger, 2002, S. 110).

Einige Kritikpunkte, die sich im Laufe der wissenschaftlichen Beschäftigung der Testautoren mit der Anwendung in der Praxis als bedeutend herausstellten, waren:

- fehlende Einführungsaufgaben,
- die hohe Anzahl an Hilfen,
- statische Testeigenschaft und der
- Wahrscheinlichkeitsbegriff.

In dieser Arbeit wird auf die angeführten Kritikpunkte eingegangen, die aufgetretenen Schwächen sollen korrigiert werden. Weiters wird der Versuch unternommen, die angebotenen Hilfen in einen wissenschaftlichen Kontext zu setzen und darüber hinaus weitere Persönlichkeitsmerkmale zu extrahieren.

Der theoretische Teil dieser Arbeit beschäftigt sich mit Untersuchungen zum Entscheidungsverhalten im Kontext von Unsicherheit und der Zuschreibung dieses Verhaltens der Ursache und Wirkung von Handlungen. Zudem wird auf grundsätzliche Testmerkmale eingegangen. Der empirische Teil enthält die konkrete Handhabung bei der Durchführung, Auswertung und Interpretation der Ergebnisse.

2. Theoretischer Teil

2.1 Objektive Persönlichkeitstests

2.1.1 Begriffsbestimmung

Eine Begriffsdefinition für Objektive Persönlichkeitstests lässt sich in den einschlägigen psychologischen Lexika noch nicht finden. Häcker und Stapf (2004) setzen in ihrem Standardbuch „Dorsch Psychologisches Wörterbuch“ Objektiver Persönlichkeitstest gleich einem objektiver Test, bei dem es durch die Konzeption der Testvorlage und Antwortgestaltung zu keiner Verfälschung auf Testpersonenseite kommen kann. Die Bezeichnung „objektiver Test“ wurde ursprünglich eingeführt, um zwischen Fragebogen und Test – bei Tests findet eine tatsächliche Handlung statt, während bei Fragebogen lediglich ein Verhalten beschrieben wird – zu unterscheiden (Cattell, 1973). Für Schmidt (1975) wird bei objektiven Persönlichkeitstests in standardisierten Situationen das Verhalten eines Individuums erfasst, das sich nicht selbst beurteilen muss. Eine treffendere Definition gibt Kubinger (1996, S.7; zitiert nach Kubinger, 2003a, S. 304) ab, „Objektive Persönlichkeitstests verfolgen das Prinzip, persönliche Stilmerkmale aus dem *beobachtbaren* Verhalten bei bestimmten (Leistungs-) Anforderungen zu erschließen“.

2.1.2 Entwicklung

Obwohl das Gebiet der Persönlichkeitsforschung ein sehr altes ist, wenn es in Bezug zur gesamten wissenschaftlichen Psychologie gesetzt wird, wurden auch objektive Persönlichkeitstests bereits früh behandelt. Die Anfänge gehen hier auf Raymond R. Cattell zurück. Überlegungen zur objektiven Messung der Persönlichkeit hatte er schon relativ früh in den 1930er Jahren des letzten Jahrhunderts gegeben. Durch seine systematischen Untersuchungen und die

daraus gewonnenen faktorenanalytischen Ergebnisse publizierte er 1975 eine ganze Testbatterie: OA-TB75 (Häcker, 1975), die auch in deutschsprachiger Version erhältlich ist. Zu dieser Zeit gab es schon eine Vielzahl an Objektiven Persönlichkeitstests, die jedoch dem Gütekriterium Unverfälschbarkeit (im Sinne von Durchschaubarkeit) und damit verbunden der Kontrollposition des Antwortverhaltens von Seiten der Testperson nicht ausreichend genügten.

Die Forschungen u.a. von Cattell, Eysenck oder Fahrenberg und die Weiterentwicklung von objektiven Persönlichkeitstests wurden über einen langen Zeitraum weitgehend ignoriert. Obwohl sich eine Reihe von Spezialisten mit dem Thema beschäftigte, fanden die Erkenntnisse nicht die nötige Würdigung und folglich auch nicht die Aufnahme in die psychologisch-diagnostische Praxis. Ein entscheidender Grund für die geringe Beachtung objektiver Persönlichkeitstests war nach Schmidt (1975) die mangelnde Begriffsbestimmung und den daraus abgeleiteten Persönlichkeitskonstrukten.

In den 1990er Jahren wurde der Fokus auf Objektive Persönlichkeitstests wieder verstärkt. Speziell durch die Verwendung des Computers wurden Vorgabe und Datenerhebung revolutioniert.

Im deutschsprachigen Raum existieren bereits eine Reihe von Objektiven Persönlichkeitstests wie beispielsweise die Testbatterie „Arbeitshaltungen“ von Kubinger und Ebenhöf (1996), welche unter anderem Impulsivität/Reflexivität erfasst oder die Testbatterie „BAcO“ (Kubinger, Schrott & Ortner, 2000), welche verschiedene Facetten von emotionaler Belastbarkeit misst.

2.1.3 Unverfälschbarkeit als Zielsetzung

Das größte Manko von Persönlichkeitsfragebogen liegt in ihrer Verfälschbarkeit. D.h. sie können wissentlich oder unbeabsichtigt manipuliert werden. Immerhin wird bei der Konstruktion von Fragebögen durch so genannte Lügen- oder Kontrollskalen der Versuch unternommen, diese Tendenzen zu kontrollieren.

Während die PRF (Deutsche Personality Research Form) von Stumpf (1984) durch die so genannte „Infrequenzskala“ die Offenheit einer Testperson noch zu messen versucht, wird beim TPF (Trierer Persönlichkeitsfragebogen) von Becker, (1989) die Testperson nur mehr instruiert nicht lange für die Beantwortung zu überlegen um damit unverfälschte Informationen zu bekommen. Es stellt sich die Frage, wie das Ergebnis zu interpretieren ist, wenn die Auswertung solcher Kontrollskalen einen „unauffälligen“ Wert zeigt.

Fisseni (1997) nennt drei Tendenzen, an welchen Maßstäben sich eine Testperson bei der Beantwortung orientieren kann:

- von sich ein sozial erwünschtes Bild zu bieten
- bestimmte Merkmale hervorzukehren oder aber zu verschweigen
- einem vorgegebenen Antwortmuster zu folgen

Diese Selbstpräsentation bringen Mummendey und Grau (2008, S. 14) auf den Punkt: „Eine Person ist stets bestrebt, das Bild mitzugestalten, das andere Personen von ihr „haben“ (sollen). (...), ständig wie ein Schauspieler, der sich mehr und weniger bewußt an seine Rolle anpaßt“.

In Zusammenhang mit der Unverfälschbarkeit eines Tests ist der Begriff der Durchschaubarkeit zu erwähnen. Während es bei Leistungstests im hohen Maße beabsichtigt wird, das Messprinzip von Testpersonenseite für eine Test konforme Beantwortung zu verstehen, um eine bestmögliche Leistungserbringung zu gewähren, ist bei Persönlichkeitsfragebogen die Durchschaubarkeit nicht erwünscht, weil dadurch Rückschlüsse auf die Messintention gezogen werden und folglich verfälscht werden können (Kubinger, 2006a).

An dieser Stelle sei ein Beispiel von Cattell aus seiner Testbatterie angeführt: Testpersonen mussten Items bearbeiten wie z.B. „Wenn jemand seinen Geldbeutel auf dem Tisch eines Restaurants vergisst, wird er in 95 von 100 Fällen zurückgegeben“. Über die Verrechnung von mehreren ähnlichen Items (Antwortmöglichkeiten waren a) stimme zu oder b) stimme nicht zu) wurden die Eigenschaften „Pessimismus“ (Verhältnis pessimistischer Antworten zu allen bearbeiteten Items) und „Zustimmungstendenz“ (Quotient aus den bejahten und

verneinten Items) gemessen. Eine kritische Testperson könnte sich nun darüber Gedanken machen, ob die Situation in ein schlichtes oder gehobenes Restaurant projiziert wird; davon abhängig ist das Antwortverhalten und in weiterer Folge kommt es zur Verfälschung. Dieses Beispiel soll zeigen, wie schwierig Verfälschungstendenzen zu kontrollieren sind.

In zahlreichen Experimenten wurde bereits nachgewiesen, dass Personen zu ihrem persönlichen Vorteil verfälschen, wenn damit bestimmte Konsequenzen verbunden sind (Ortner et al., 2007).

Die Unverfälschbarkeit kann laut Schmidt (1975) durch eine durchdachte Itemwahl oder spezielle Auswertungsmethoden erreicht werden. Nach Kubinger (2003b) ist bei einem psychologischen Test die Unverfälschbarkeit hoch, wenn einer Testperson keine Kontrolle über Art und Inhalt der Informationen geliefert werden, sie das Messprinzip nicht durchschauen kann.

Nach den Qualitätsstandards der DIN 33430 (DIN - Deutsche Institut für Normung, 2004) sollen im Zuge berufsbezogener Eignungsbeurteilung Kandidaten ihre Ergebnisse möglichst nicht verfälschen. Genau diese Anforderung erfüllen Objektive Persönlichkeitstests. Testpersonen können aufgrund mangelnder Durchschaubarkeit des Testprinzips Ergebnisse nicht in eine gewünschte Richtung beeinflussen.

2.1.4 Möglichkeiten und Probleme Objektiver Persönlichkeitstests

2.1.4.1 Testtheoretische Problemfelder

In der sozialwissenschaftlichen Methodenlehre gibt es zwei Begriffe von Testtheorie. Der erste bezeichnet eine Theorie über Konklusionen, die man anhand von Stichproben bezüglich gewisser Eigenschaften der Population zieht.

Solche Konklusionen werden als Test bezeichnet, weil eine Hypothese getestet wird.

Der zweite Begriff bezeichnet die Theorie über psychologische Tests, im Sinne von klassischer und probabilistischer Testtheorie. Testtheorien befassen sich mit der Frage, in welche Bestandteile sich Messwerte aufgliedern oder mit dem Zusammenhang von Testverhalten und dem zu erfassenden Merkmal. In dieser Beziehung muss die Verrechnung von Testkennwerten eine verlässliche Zuordnung garantieren, um die Leistungsfähigkeit aller Testpersonen in eine faire Relation setzen zu können (Kubinger, 2006a).

Um den testtheoretischen Problemen entgegenzutreten und höheren messtheoretischen Anforderungen zu genügen, können Tests konstruiert werden, die dem Rasch-Modell standhalten. Radinger (2002) zeigte in ihrer Diplomarbeit, dass mit dem Rohscore des EUcO mehrere Dimensionen erfasst werden und somit die Gültigkeit des Rasch-Modells verworfen werden musste. Für eine detailliertere Untersuchung müsse eine größere Stichprobe gezogen werden. Bei der Teilung nach hoher vs. niedriger Score konnte in der Teilgruppe „Wertpapierhändler“ das Rasch-Modell jedoch bestätigt werden.

2.1.4.2 Konstruktvalidität

„Die Validität oder Gültigkeit eines Tests gibt den Grad der Genauigkeit an, mit dem dieser Test dasjenige Persönlichkeitsmerkmal (...), tatsächlich misst oder vorhersagt“ (Lienert und Raatz, 1998, S. 10). Verschiedene psychologische Verfahren weisen unterschiedliche Validitätswerte auf, wobei aber nur jene zum Einsatz kommen sollen dessen Validitätswert am Höchsten ist. Der Validitätsbegriff lässt sich in drei Gruppen unterteilen: Inhaltliche Validität, kriterienbezogene Validität und Konstruktvalidität.

Näher eingegangen wird auf die Konstruktvalidität und auf die Frage, ob ein Test das zu interessierende Konstrukt überhaupt erfasst. Während Leistungstests, die z.B. die rechnerische Fähigkeit durch Rechenaufgaben messen „relativ

offensichtlich“ konstruktvalide sind, ist das Gütekriterium der Konstruktvalidität bei Objektiven Persönlichkeitstest schwieriger zu bestimmen.

Der Gestaltwahrnehmungstest von Hergovich und Hörndler (1994) misst den kognitiven Stil der Feldabhängigkeit. Feldabhängige Personen orientieren sich stärker an anderen Personen, halten einen geringeren Körperabstand und sind emotional offener. Ihnen wird eine höhere soziale Kompetenz zugesprochen. In empirischen Studien konnte die Validität des Tests bestätigt werden, u. a. (Stift, 1995, zitiert nach Hergovich, 2006) dass sich Feldabhängige bei Arbeiten stärker von aktivierten Gefühlen beeinflussen lassen. Allerdings wird durch die Ergebnisse in neueren Untersuchungen eine Abkehr vom Konstrukt der Feldabhängigkeit gemacht. Der Gestaltwahrnehmungstest ermittelt demnach eine Fähigkeit und müsste als Leistungstest kategorisiert werden.

Anastasi (1954) spricht noch von „factorial validity“ inzwischen gibt es bereits mehrere Ansätze für die Überprüfung der Konstruktvalidität. Neben der Faktorenanalyse sind noch der Multitrait-Multimethod Ansatz von Campbell und Fiske (1959) sowie nicht-korrelative Ansätze zu nennen (vgl. Kubinger, 1996 oder Bühner, 2004).

2.1.4.3 Möglichkeiten von Objektiven Persönlichkeitstests

In Objektiven Persönlichkeitstests wird gezielt die Augenscheinvalidität (face validity) verringert, um damit die Verfälschbarkeit zu verhindern. Auf der anderen Seite erhöht sich die Akzeptanz eines Verfahrens, wenn Augenscheinvalidität gegeben ist. Kersting (2003) spricht in dieser Beziehung von einem Akzeptanz-Verfälschbarkeits-Dilemma.

Auf Grund der geringen Augenscheinvalidität und der auf diese Weise einhergehenden Täuschung der Testperson müssen ethische Prinzipien diskutiert werden. Aus ethischer Sicht ist es unbedingt notwendig, Testpersonen in entsprechender Form zu informieren. Nach Ortner et al. (2007) stellt dies aber keinen Widerspruch dar: demnach kann eine angemessene Aufklärung gegeben

werden, ohne der Testperson die genaue Verrechnung ihres Verhaltens zu einem Testwert preiszugeben.

Ein weiterer Punkt ist die starke Sprachgebundenheit und deren Problematik bei Persönlichkeitsfragebogen, welche unter anderem auch bei der Übersetzung zu Problemen führt (vgl. Horn, 2003). Bei Objektiven Persönlichkeitstest wird man um die verbale Formulierung der Instruktion nicht umhin kommen, hingegen kann bei den Aufgaben der Verbalteil auf ein Minimum reduziert werden (vgl. Ortner et al, 2007).

Viele Persönlichkeitsfragebögen erfassen die Persönlichkeit bzw. Dimensionen der Persönlichkeit, meist aus zuvor faktorenanalytischen Untersuchungen, um damit ein theoriegeleitetes Abbild der Persönlichkeit zu schaffen (vgl. Amelang & Bartussek, 2001). Faktoren sind hier primäre Einflüsse, die das menschliche Verhalten in fundamentaler Weise bestimmen. Diesen Standpunkt vertreten mehrere Forscher u. a. auch Cattell und Thurstone. Eine entgegengesetzte Position nehmen jene Forscher ein, für welche es eben nicht möglich ist, mittels Faktorenanalyse solche Einflüsse ausfindig zu machen. Diese diametralen Ansichten bringen Lord und Novick (1965, zitiert nach Piaggio, 1973, S. 11) auf den Punkt: „Thought much of psychological theory is based on a trait orientation, it is not necessary that these traits exist in any physical or physiological sense. It is sufficient for most purposes of psychology that a person behaves (...)”.

Bei der Konzeption von Objektiven Persönlichkeitstests können bereits bekannte Konstrukte valide und unverfälscht erfasst werden; darüber hinaus auch neue Konstrukte, die bisher nicht mit den gewohnten Methoden erfassbar waren (Ortner et al, 2007).

Kubinger (2006b, S. 42) diskutierte in diesem Zusammenhang die Messung von theoriekonformen Persönlichkeitsdimensionen und Objektiven Persönlichkeitstests und schlussfolgerte, dass sie „den persönlichen Stil eines Menschen, nämlich sich in typischen Situationen in ziemlich typischer Weise zu verhalten“ erfassen.

2.1.5 Computerisierte Vorgabe

Der Computer als Vorgabemedium hat sich inzwischen zum integralen Bestandteil des diagnostischen Prozesses entwickelt. Er bietet in den Teilbereichen Testdurchführung und -auswertung, Testentwicklung, Testevaluation und schließlich bei der diagnostischen Urteilsbildung mannigfaltige Unterstützung.

Einer Reihe von Problemen kann bei herkömmlichen Papier-Bleistift-Verfahren durch die computerisierte Vorgabe begegnet werden (Fisseni, 1997):

Testinstruktion

Die Anfänge von computerisierten, damals noch computergestützten, Instruktionen reichen weit in die siebziger und achtziger Jahre des letzten Jahrhunderts zurück. Zu dieser Zeit wurden hauptsächlich computergestützte Lernprogramme in Großunternehmen eingesetzt (Niegemann, 2008).

Die (computerisierte) Vorgabe der Instruktion (siehe Kapitel 2.2.) ermöglicht es der Testperson ein Verständnis für die Testintention – Stichwort Augenscheinvalidität – aufzubauen, viel wichtiger ist jedoch, dass der Informationszeitbedarf individuell angepasst werden kann. Sidiropoulou (1997) spricht von „maßgeschneiderten Instruktionen“. Dabei „handelt es sich um lineare Lernprogramme, die ein stufenweises Erlernen der Aufgabenstellung in aufsteigender Komplexität beinhalten“ (Bukasa, Kisser & Wenniger, 1990, zitiert nach Volz-Sidiropoulou, 2004, S. 386). Zwei Anliegen werden damit verbunden: erstens soll eine Verständniserleichterung für die Testperson erfolgen und zweitens die Informationszeit minimiert werden. Volz-Sidiropoulou (2004) berichtet von einer Gesamttestzeitverkürzung bei solchen Instruktionsprogrammen im Vergleich zu herkömmlichen Tests.

Die Multimedia-Eigenschaft des Computers erlaubt es die Instruktion auditiv darzubieten. Damit kann parallel zur visuellen Erklärung die Instruktion akustisch vermittelt werden. Auf diese Weise kommt es zu keiner Benachteiligung von Personen, welche Inhalte besser über den auditiven Zugang aufnehmen können.

Mittels der Computerrealisierung kann die Testperson über das Abspielen von Video- und Audiosequenzen selbst entscheiden (Parshall, Spray, Kalohn & Davey, 2002).

Beim EUcO 2.0 wird in der Testsituation die Möglichkeit des wiederholten Instruktionsabrufs geboten.

Exemplarisch sei an dieser Stelle die Papier-Bleistift-Version des dreidimensionalen Würfeltests von Gittler (1989) angeführt, der verschiedene Interaktionszenarien während der Instruktionsphase vorsieht und die Durchführungsobjektivität beeinflussen könnte. Bei der Computervorgabe erhöht sich die Durchführungsobjektivität, und die bekannten Testleitereffekte werden ausgeschlossen (Volz-Sidiropoulou, 2004).

Testübungsphase

Wie die Instruktion kann die Testübungsphase dem individuellen Arbeitsstil der Person angepasst werden.

Itempräsentation

Durch die Grafik- und Soundfähigkeit des Computers können Objekte beweglich und farblich dargestellt oder mit akustischen Signalen unterlegt werden. Klassische Antwortformate wie z.B. Multiple-Choice können angereichert und variiert werden. Nach Parshall et al. (2002) können bei computerbasierten Tests die Antwortformate komplexer konstruiert vorgegeben werden.

Die Möglichkeiten des Computers bei der Testentwicklung, -evaluation und -auswertung seien hier nur am Rande angemerkt (vgl. Jäger & Krieger, 1994).

Weitere Nutzungsmöglichkeiten durch den Computer sind laut Kubinger (2006a) unter anderem:

- mehr oder weniger komplexe Reizeanordnung, variiert über die Zeit
- Kontrolle der Speed-Komponente bei Gruppenverfahren
- Reduktion der a-priori Ratewahrscheinlichkeit

- beliebige Nuancierung der Antworten (Analogskala), egal wie die Antworten verrechnet werden
- Interaktion zwischen Testperson und Computer
- experimentelle Manipulation der Testperson

Die Computervorgabe erlaubt eine genaue und zusätzliche Registrierung von Daten, die in herkömmlichen Papier-Bleistift-Verfahren nicht aufgezeichnet werden, wie etwa vorgenommene Korrekturen und deren Häufigkeit, Bearbeitungszeit oder Latenzzeiten. Gemäß Klieme und Stumpf (1990, zit. nach Sidriopoulou, 1997, S. 387) wird als Latenzzeit das „Intervall vom Beginn der Darbietung eines Items auf dem Bildschirm bis zum Moment der Registrierung einer Reaktion durch das Gerät“ verstanden.

Daneben können Items (ohne „zurückblättern“) wiederholt vorgegeben werden.

Die Frage nach einer möglichen Benachteiligung von Personen, die keine oder nur geringe Computererfahrung haben, wurde in der Arbeit von Klinck (2002) untersucht. Sie stellte eine positive Korrelation zwischen Computererfahrung und Testwerten in computerisierten Testverfahren fest. Allerdings ist dies kein Beleg für die Existenz eines derartigen Bias, weil auch bei Papierversionen dieselben Korrelationshöhen aufgefunden wurden. Dies ist darauf zurückzuführen, dass kognitiv leistungsfähigere Personen eher den Computer als Arbeitsmittel benutzen und somit computererfahrener sind als weniger leistungsfähige Personen.

Computerisierte Verfahren bieten nicht nur inhaltliche und methodische Vorteile. Die Durchführung der Testvorgabe muss nicht unbedingt durch Spezialisten erfolgen, sondern kann durch ein geschultes Personal administriert werden. Dadurch werden die verantwortlichen Psychologen entlastet und können sich tätigkeitsspezifischen Aufgaben widmen (Krumm & Schmidt-Atzert, 2009).

Die kritiklose Verwendung des Computers wird von Booth (1994) diskutiert, der neben dem Nutzen auch Nachteile wie die Gefahr der Datenüberflutung oder der missbräuchlichen Verwendung der Daten sieht. Insgesamt überwiegen aber die

Vorteile der Computertechnologie: die Verantwortung für den Anwender wird jedoch höher.

2.2 Testinstruktion

Bevor detaillierter auf die Instruktion eingegangen wird, muss zunächst darauf eingegangen werden, welche Punkte bei der Konstruktion von psychologischen Tests von besonderer Bedeutung sind.

Bei der Testkonstruktion müssen Verfahrensvorschriften erarbeitet werden, die zur Sammlung und Auswertung der psychologischen Daten dienlich sind (Vukovich, 1964). In der Testsituation werden diagnostische Informationen gewonnen, die durch Testmaterial, Instruktion und Durchführungsbedingungen bestimmt werden. Bühner (2004) spricht in diesem Zusammenhang von „Stufen der Testentwicklung“, die ein Test innerhalb des Konstruktionsprozesses durchlaufen muss. Eine Stufe ist die „Revision eines Tests“, wenn neue Informationen gewonnen wurden, die eine Überarbeitung rechtfertigen.

Bei der Konstruktion oder Weiterentwicklung eines Verfahrens muss für die DIN 33430 (2004) hinreichend begründet werden, warum eine (Neu)-Konstruktion erforderlich ist. Die zwei wichtigsten Gründe sind, dass es für ein interessierendes psychologisches Merkmal tatsächlich kein Verfahren gibt oder dass ein bestehendes Verfahren offensichtliche Schwächen hat.

2.2.1 Formale Testinstruktion

Die Erklärung oder Einführung der Testpersonen für die Bearbeitung ist in der Literatur begriffsdefinitorisch nicht einheitlich festgelegt. In den verschiedensten Arbeiten wird zwischen Anweisung, Anleitung und /oder Instruktion unterschieden. Eine zufrieden stellende Differenzierung nehmen Guthke und Sprung (1990) vor, wenn sie im Hinblick auf die Bedingungen der Durchführung und Auswertung die Anweisung auf Seite des Testleiters platzieren. Die Instruktion stellt für die

Testperson jene Informationen bereit, die über den Zweck der Testung Aufschluss geben und darüber, welche Leistungen erwartet werden.

„Die Testinstruktion ist sozusagen die Eintrittskarte zu einem Test“ (Moosbrugger & Kelava, 2007, S. 68), sie soll die Testpersonen zur Teilnahme anregen und klare Handlungsanweisungen geben. Nach der DIN 33430 (2004) stellt die Instruktionsphase bei eignungsdiagnostischen Verfahren jene Informationen über die Rahmenbedingungen dar, in denen erklärt wird, wie das Verfahren zu bearbeiten ist.

Wiewohl Standardwerke über Testkonstruktionen (vgl. Lienert & Raatz, 1998 oder Guthke & Sprung, 1990) die Notwendigkeit einer Testanweisung deutlich machen, so kommt bei Objektiven Persönlichkeitstest der Instruktion eine besondere Bedeutung zu. Da die Testperson einerseits über die „richtige“ Bearbeitung informiert werden muss, andererseits auch auf Grund der Spezifität dieser Tests. Weil bei Objektiven Persönlichkeitstests die Augenscheinvalidität so gering wie möglich gehalten werden soll bzw. die Messintention überhaupt im Verborgenen bleibt, ist die instruktionskonforme Bearbeitung für die Interpretation der resultierenden Testwerte von entscheidender Wichtigkeit.

2.2.1.1 Instruktionsform bezüglich Testdurchführung

Lienert und Raatz (1998) unterscheiden verschiedene Instruktionsformen. Abhängig davon wie ein Test beschaffen ist, sprechen sie von genereller, spezieller oder separater Instruktion.

Bei der generellen Instruktion werden alle relevanten Informationen (Ablauf und Bearbeitung) über den gesamten Test vor Beginn der Durchführung gegeben. Die Instruktion gilt gleichermaßen für die Bearbeitung aller Items (Guthke & Sprung, 1990). Spezielle Testinstruktionen finden meistens bei Testbatterien Verwendung, die aus mehreren Untertests zusammengesetzt und deren einzelne Items homogen sind. Der Intelligenz-Struktur-Test 2000 R (Amthauer, Brocke, Liepmann und Beauducel, 2001) ist ein solches Verfahren: jeder Untertest verlangt eine aufgabenadäquate Instruktion.

Das Charakteristikum der separaten Instruktion hängt mit der extremen Heterogenität des Aufgabenmaterials zusammen. Jede Aufgabe wird einzeln instruiert, was häufig bei Entwicklungstest und Tests für Kleinkinder vorkommt.

„Die Formulierung der Instruktion muss standardisiert vorgegeben werden, da sie sonst auf Kosten der Durchführungsobjektivität geht und dadurch die Ergebnisse verfälscht werden“ (Bellmann & Sarges, 1980, zit. nach Lienert & Raatz, 1998). Die Instruktion soll (Vukovich, 1964) keine interindividuell unterschiedlichen Auslegungen zulassen, die später nicht berücksichtigt werden können.

2.2.1.2 Regeln zur Instruktions- und Itemform

Wie bereits weiter oben erwähnt wird auf eine ausführliche Regelung bei der Instruktionsgestaltung weitgehend verzichtet. Meist lassen sich wichtige Aspekte (so kurz wie möglich und so ausführlich wie nötig) in kurzen Sätzen formulieren. Hingegen werden für die Itemgestaltung ganze Kapitel ausgefüllt (vgl. Bühner, 2004 oder Rost, 1996).

Da die Entwicklung von Objektiven Persönlichkeitstests auf die Probleme und Schwächen von Fragebögen zurückgeht, können grundsätzliche Regelungen und Anforderung für die Itemgestaltung übernommen werden. Wegen der besonderen Eigenschaft von Objektiven Persönlichkeitstests ist es nur konsequent diese Bestimmungen auch auf die Instruktionsgestaltung, im Besonderen der Sprache betreffen siehe Kapitel 2.2.2, zu übernehmen.

Itemform

Das Item lässt sich nach Rütter (1978) in zwei Teile, Stimuluskomponente und Reaktionskomponente, untergliedern, wobei der Stimulusteil nochmals in Informationsfeld und Fragefeld aufgesplittet werden kann. Das Item stellt die allgemeine Situation dar, in der die Testperson ihr Verhalten zeigt und kann aus einer Frage, einer Aussage, einer Geschichte, einer visuellen Stimulation oder kognitiven Anforderung bestehen. Es stellt die kleinste Beobachtungseinheit in einem Test dar.

Items können für ein besseres Verständnis in Teilaufgaben zerlegt werden. Lienert und Raatz (1998) erwähnen ein Beispiel mit Drahtbiegeproben (wobei einmal Länge und einmal Richtung bewertet werden, gemessen an der Drahtbiegevorlage). Durch die Aufgabenteilung ist das Item für die Testperson verstehbar und wirkt weniger komplex.

Die Reihenfolge der Items hängt für Lienert und Raatz (1998) eng mit der Durchführungstechnik zusammen. Reine Power-Tests verlangen keine Einhaltung einer bestimmten Reihfolge. Hingegen müssen bei Speed and Power-Tests die Items nach aufsteigendem Schwierigkeitsgrad geordnet werden. Zu Beginn jeder Testung sollten ein oder zwei Items als so genannte Eisbrecher vorgegeben werden, um die Motivation zu erhöhen.

Jedes Item sollte so konkret wie möglich gestaltet werden. Angleitner und Wiggins (1986) fordern einen konkreten Gesamteindruck, da ein abstraktes Item für Verwirrung sorgt und ein Antwortverhalten auslöst, das von der intendierten Testabsicht abweicht. Nisbett und Ross (1980, zit. nach Angleitner et. al., 1986, S.34) erbrachten den Beweis, dass „konkrete und anschauliche Informationen leichter abgerufen werden“. Bei einigen Tests bzw. Items kann dieser Forderung nicht nachgekommen werden, da die formale Gestaltung dies einfach nicht zulässt. Man denke nur an die Raven-Tests, die trotzdem von großer diagnostischer Relevanz sind (Guthke & Sprung, 1990).

Itemwiederholungen sollten laut Kreutz und Titscher (1974) nur in so großen Abständen erfolgen, dass die Antworten stochastisch voneinander unabhängig sind. Sie warnen vor einer möglichen Motivationsreduzierung, wenn die Wiederholung von der Testperson bemerkt wird.

Für Guthke und Sprung (1990) sollen die verwendeten Items inhaltlich unabhängig sein, um damit Reihenfolgeeffekte zu vermeiden – ein vorausgegangenes Item soll keinen Hinweis zur Lösung eines nachfolgenden geben. Durch Pufferitems können in einigen Fällen Reihenfolgeeffekte vermindert werden (Rost, 1996).

Instruktionsform

Damit die Testperson zur aktiven Teilnahme animiert wird und motiviert an die Testbearbeitung herantritt, braucht sie klare Handlungsanweisungen. Eine Hauptregel für die Instruktion lautet (Guthke & Sprung, 1990, S. 227), „So kurz wie möglich, aber auch so eindeutig und allgemein verständlich wie notwendig“.

Die Instruktion soll nur klare Begriffe enthalten. Nach Tränkle (1993, zit. nach Fisseni, 2004, S. 206) „eignen sich Begriffe in dem Maße, in dem sie denotativ eindeutig und konnotativ arm sind“. Komplexe Formulierungen können zu Missverständnissen führen, die eine fehlerbehaftete Itembearbeitung hervorrufen, welche sich schließlich auf die Validität auswirkt (Jäger & Petermann, 1992).

Die Verwendung von Häufigkeiten (z.B. oft, selten, usw.) sollte nicht in Worten angegeben werden, sondern in Zahlen oder in Beispielen (Fisseni, 2004). Diese Regel muss aus der Itemformulierung in die Instruktionsgestaltung übernommen werden, um etwaige Uneindeutigkeiten vorwegzunehmen. Weiters sind suggestive Satzformulierungen, wie sie manchmal gewollt oder ungewollt in Fragebögen vorkommen zu unterlassen.

Wiederholungen einiger Textpassagen lassen sich zwangsweise, aufgrund von unterschiedlichen Bildungsniveaus, nie vermeiden, durch eine computerisierte Vorgabe kann dem jedoch jede Testperson individuell begegnen.

„Nur am Rande spielen so genannte Testroutiniers eine Rolle, da Objektive Persönlichkeitstest per se keine bzw. keine hohe Augenscheinvalidität haben sollten“ (Anastasi, 1976, zitiert. nach Grubitzsch, 1991, S.151). Jedoch sollten sie bei der Instruktionsgestaltung Berücksichtigung finden.

Die Instruktion ist umso besser gelungen, je weniger Rückfragen während der Anleitungsphase bzw. bei der eigentlichen Testung auftreten. Für Sader (1959) muss die Instruktion in Vorversuchen auf ihre Verständlichkeit hin überprüft werden. Bei vielen psychologischen Verfahren werden Übungsbeispiele zur

Veranschaulichung für richtiges Lösungsverhalten gegeben. Guthke und Sprung (1990) verlangen zu Beginn eines Tests einfache Übungsbeispiele, damit unerfahrene Personen mit dem konfrontierten Material vertraut werden. Verzerrungen der eigentlichen Leistungsfähigkeit werden minimiert und die Testperson lernt den formell-technischen Aufbau der Items kennen.

Übungsbeispiele

Die Vorgabe von Übungsbeispielen ist unerlässlich für die Einführung - die Testperson kann ihr Verständnis für die Aufgaben und der von ihr später geforderten Leistung überprüfen, nachdem sie zuvor die dafür nötigen theoretischen Informationen erhalten und Übungsbeispiele unter kontrollierten Bedingungen mit Rückmeldungscharakter bearbeitet hat. Die Testperson kann quasi selbst kontrollieren, ob alle relevanten Informationen tatsächlich verstanden wurden. Sollten Unklarheiten zu Tage treten, so kann die Testperson den Testleiter kontaktieren oder im Falle einer computerisierten Vorgabe die Instruktion nochmals abrufen.

Die Testperson sollte durch Übungsbeispiele den Umgang mit den Aufgaben lernen.

Rost (1996) geht davon aus, dass es bei der eigentlichen Testung bzw. bei der Bearbeitung der Items zu Lerneffekten kommen kann, die das Ergebnis beeinflussen.

Die meisten Lernvorgänge sind reaktionskontingent (mit der Reaktion zusammenhängend) und gefährden die lokale stochastische Unabhängigkeit. „Findet dagegen lediglich Lernen im Sinne von Üben statt, was im wesentlichen von der Anzahl und Qualität der Aufgaben (...) abhängt, (...), die mit der Annahme der stochastischen Unabhängigkeit vereinbar ist“ (Rost, 1996, S. 75).

Um alle Testpersonen auf den gleichen Level oder viel mehr auf ein gleiches Basisniveau zu bringen empfehlen sich lineare Lernprogramme. Ein noch besseres Vorgehen bei der Übungsphase wird nach Volz-Sidiropoulou (2004) durch die adaptive Vorgabe ermöglicht. Die Testperson steuert durch ihr Antwortverhalten die Übungsphase, wodurch individueller auf das Verständnis der

Testintention eingegangen wird. Dies erfordert allerdings eine aufwendigere Konstruktion und eine komplexere Programmierung.

Grafische Gestaltung

Die typographische und farbliche Gestaltung und das Layout von psychologischen Verfahren wurden bisher nur von einigen wenigen Testautoren beachtet. (Hartley, Lindsey & Burnhill, 1977), obgleich dieses Thema bei den Testentwicklern Berücksichtigung findet. Verfolgt man nur die Veränderungen des Layouts über die letzten Jahrzehnte, so gibt es überraschenderweise diesbezüglich keine allgemeine Regelung oder Norm. Lienert und Raatz (1998) fordern zwar eine übersichtliche Anordnung, gute räumliche Trennung durch Umranden oder Unterstreichen eine Hervorhebung des Textes: auf detailliertere Angaben wird aber verzichtet.

Karmasin und Karmasin (1977) fordern für das Layout eine eindeutige optische Unterscheidung durch verschiedene Schrifttypen oder Typengrößen. Einzelne Fragen – in konsequenter Weise auch einzelne Sätze – sollten optisch nicht dominieren, sie sollten zu einem Block gegliedert werden. Die Größe dieser Blocks sollte aber begrenzt bleiben und durch Überschriften getrennt werden, um kognitive Erschöpfung zu vermeiden (Richter, 1969).

In Untersuchungen konnte kein besonderer Nutzen bei der Verwendung von Farben festgestellt werden (vgl. Tränkle, 1983). Um Ermüdungs-, Sättigungs- oder Monotonieerlebnisse entgegenzuwirken, sie sind aber allemal hilfreich, das Testmaterial optisch aufzufrischen.

Eine übertriebene Verwendung von Farbe in psychologischen Verfahren sollte trotzdem vermieden werden, da nach Häcker und Stapf (2004) etwa 5 % der Männer und 0,25 % der Frauen an einer partiellen Farbblindheit leiden.

Bei der graphischen Gestaltung muss auch die Menge der gegebenen Information Beachtung finden, ansonsten kann es leicht zur Überlastung kommen. „Eine besonders schnelle und gedanklich bequeme Informationsaufnahme kann durch

Bilder erfolgen“ (Kroeber-Riel & Esch, 2000, S. 15). Bei einem Bild von mittlerer Komplexität sind cirka 2 Sekunden für die Aufnahme nötig, in der gleichen Zeit können auch zehn Wörter aufgenommen werden. Dennoch wird über ein Bild wesentlich mehr Botschaft vermittelt.

Ein nicht unwesentlicher Unterschied zwischen visueller und verbaler Verarbeitung liegt darin, dass Bildinformationen automatisch und mit geringerer Anstrengung verarbeitet werden können. Weitere Nutzen sind die stärkere Aktivierung und die leichtere Abrufbarkeit von Bildinformationen (Kroeber-Riel, 1993).

Durch Kontraste in der Helligkeit kann die Aufmerksamkeit von verschiedenen Objekten gesteuert werden. „Kontrasteffekte vergrößern die Reaktion auf einen konstanten Reiz entsprechend dem vergrößerten Intensitätsunterschied zwischen Reiz und Reizhintergrund“ (Zimbardo, 1995, S. 178). Mittels Helligkeitskontrast werden zwischen Oberflächen Grenzen schärfer wahrgenommen.

Ein Problem beim EUcO (Radinger, 2002) war für die Testpersonen die Zuordnung der gezogenen Kugel im Auswurf. Überlegungen, ob die Kugel zum richtigen Behälter noch hinzugezählt werden muss oder nicht, stiftete Verwirrung. Gegenstände die aus dem Kontext gerissen werden, wirken erschwerend auf die Klassifikation, und die Informationsverarbeitung dauert länger an. Eine bedeutende Quelle für die Einordnung von Objekten liefert der räumliche und zeitliche Kontext.

Nach Zimbardo (1995) hängt die Klassifikation von Wahrnehmungen von den Erwartungen und von den Eigenschaften der Gegenstände ab. Diese Wahrnehmung wird vom räumlichen und zeitlichen Kontext gelenkt. Die Interpretation einer Person ist davon abhängig, was sie weiß und was sie um sich sehen kann. Steht dies im Widerspruch zueinander, so gelingt die Interpretation nicht. Die Kommunikation zwischen „bottom-up“ und „top-down“ Prozessen ist gestört.

2.2.1.3 Testlänge, Testzeit und Zielgruppe

Die Frage, wie lange ein Test sein darf, sein soll bzw. sein muss, hängt unweigerlich mit der Frage der Testökonomie und der Messgenauigkeit zusammen. Testökonomie heißt, „keinen unnötigen Aufwand bei der psychologischen Testung zu betreiben“ (Woike, 2003, S. 405). Dies impliziert, die Kosten (Testlänge) aller Beteiligten so gering wie möglich zu halten. Allerdings können Testökonomie und Messgenauigkeit nicht getrennt voneinander betrachtet werden, weil eine Testverkürzung meistens mit einer geringeren Reliabilität einhergeht. Hier sind die Testautoren gefordert, dass einerseits den Gütekriterien Rechnung getragen wird und andererseits, dass es zu keinen Erschöpfungserscheinungen oder Motivationsminderungen kommt.

Unter Testlänge verstehen Moosbrugger und Kelava (2007) zweierlei. Erstens die Summe der vorgegeben Items und zweitens die Zeit für die Bearbeitung der Aufgaben, wobei letzteres von Lienert und Raatz (1998) als Testzeit und die Itemanzahl als eigentliche Testlänge verstanden wird. Für die Testlänge ist weiters noch entscheidend, ob der Test ein (homogen) oder mehrere (heterogen) Konstrukte erfasst. Bei einem homogen ausgeprägten Merkmalsbereich werden weniger Items benötigt, und die Testlänge ist kürzer. Gewöhnlich enthalten einfache Leistungstests 20 bis 40 Einzelaufgaben. Die Testzeit ist außerdem davon abhängig, ob Schnelligkeit (Speed) oder Niveau (Power) verlangt sind. Bei Niveautests wird allgemein eine großzügige Zeitvorgabe für die Bearbeitung gegeben, und das Ergebnis wird nicht aufgrund von Stresskomponenten verzerrt.

Das „Einsatzgebiet“ eines Tests ist immer beschränkt auf einen gewissen Personenkreis, welcher aber bereits bei der Testplanung festgelegt werden muss. So unterscheiden sich die Zielgruppen Büroangestellte, Geschäftsführer und Militärstrategen wesentlich voneinander, und die Vorgabe eines Tests zum räumlichen Vorstellungsvermögen in einem Auswahlverfahren ist bei der einen Zielgruppe sinnvoller als bei der anderen.

Die Anforderungen an ein psychologisches Verfahren, das für einen weiten Personenkreis konzipiert ist sind viel höher als für eine relativ enge Zielgruppe. Bei

der Planung eines Tests für eine breite Zielgruppe muss auch der Schwierigkeitsbereich breiter streuen (Moosbrugger & Kelava, 2007).

2.2.1.4 Antwortform

Antwortformate können in drei Gruppen gegliedert werden (Moosbrugger & Kelava, 2007):

- a) freies Antwortformat
- b) atypisches Antwortformat
- c) gebundenes Antwortformat

Bei freien Antwortformaten werden keine Antwortalternativen geboten: die Person muss die Antwort selbst formulieren. Atypische Formate setzen sich oft aus Kombinationen von gebundenen Antwortformaten zusammen, der Zahlen-Verbindungs-Test (Oswald & Roth, 1987) wäre ein Beispiel für ein solches Format. Gebundene Antwortformate bieten der Testperson für die Beantwortung mehrere vom Testautor festgelegte Antwortmöglichkeiten (Lienert & Raatz, 1998).

Gebundene Antwortformate haben eine höhere Objektivität, weil die Auswertungsobjektivität sehr hoch ist. Jedoch kann es durch die vorgegebenen Antwortmöglichkeiten zu Rateeffekten kommen. Die Ratewahrscheinlichkeit bei einem dichotomen Antwortformat liegt bei 50 %. Wie bereits im EUcO (Radinger, 2002) wird dieses Format auch bei der Revision beibehalten, weil in Diskussionen festgestellt wurde, dass bei Vorgabe mehrerer Behälter diese von der Testperson immer wieder auf zwei tatsächlich Infragekommende reduziert werden. Das eigentliche Testergebnis wird in der zweiten Teilaufgabe, der Wahrscheinlichkeitsschätzung, abgelesen.

Auf so genannten Ratingskalen - sie zählen auch zu den gebundenen Formaten - können Antworten mehrkategorial abgestuft werden. Sie haben den Vorteil, dass sie von der Testperson sehr präzise und differenziert beantwortet werden können.

2.2.2 „Verbale“ Verständlichkeit der Instruktion

Um für die Testperson die Chance zu wahren, ihr bestmöglichstes Ergebnis zu erbringen, ist das sprachliche Verständnis der Instruktion von großer Bedeutung. Angleitner und Wiggins (1986, S. 86) meinen dazu „comprehensibility is of obvious importance for the response process“. Die Instruktion sollte bereits nach einmaligem Durchlesen verständlich sein. Ist dies nicht der Fall, so besteht die Gefahr von Fehlinterpretationen, woraus Verzerrungen bei der Itembeantwortung resultieren können. Solche Verzerrungen gehen auf Kosten der Gütekriterien und die Testbearbeitung wird ad absurdum geführt. Sader (1959, zitiert nach Vukovich, 1964, S.121) forderte bereits Ende der 50er Jahre, dass „eine logische Analyse und grammatikalisch richtige Formulierung keine Garantien für eine eindeutige Testinstruktion sind und die Verständlichkeit in Voruntersuchungen empirisch erprobt werden muss“.

Nach Rogers (1974) müssen nachstehende Fragen bei der Gestaltung gestellt werden:

- Ist die Instruktion einfach oder schwer zu verstehen?
- Sind Sätze eindeutig oder enthalten sie mehrere Bedeutungsinhalte?
- Wurde die gesamte nötige Informationen geben?
- Sorgt zusätzliche Information für Verwirrung oder trägt sie zum Verständnis bei

Verallgemeinerungen jeglicher Art, besonders aber jene, die nur unter bestimmten Bedingungen Gültigkeit haben, sollen unterlassen werden. Um die Testlänge nicht unnötig zu verlängern, muss eine umständliche Länge vermieden werden. Auf der anderen Seite führen telegrafische Kürzel häufig zu Missverständnissen (Lienert & Raatz, 1998).

Damit die sprachliche Verständlichkeit gewährleistet ist, verlangen Moosbrugger und Kelava (2007), dass folgende Aspekte Berücksichtigung finden:

- positive Formulierung und Vermeidung von Negationen

- einfache Satzformulierung
- Zielgruppenorientierte Sprache
- Häufigkeitsangaben von Zahlen, anstatt von verbalen Umschreibungen

Grundsätzlich ist es für Karmasin und Karmasin (1977) ausgeschlossen ein absolutes Optimum der Bedeutungsäquivalenz der sprachlichen Formulierung bei allen Testpersonen zu erreichen. Sie schlagen deshalb eine Orientierung an der Alltagssprache (Umgangssprache) eines „durchschnittlichen“ Mitglieds der Zielgruppe vor, um interindividuelle Sprachbegabungen zu kompensieren. Für die Bewerkstellung dieser Anforderung muss sich der Testautor die Kennzeichen dieser Sprache vergegenwärtigen.

Die Formulierung von Sätzen soll sich auf die gebräuchlichsten Wörter reduzieren. Noch wichtiger ist die Gebräuchlichkeit eines Wortes in einem gegebenen Zusammenhang. Fremdwörter und abstrakte Begriffe sollten so weit wie möglich vermieden werden (Friedrichs, 1997).

Die grammatikalische Satzkonstruktion der Alltagssprache ist durch eine geringe „Satztiefe“ geprägt. Kreutz und Titscher (1974) verlangen sogar eine Beschränkung auf Hauptsätze. Dem schließen sich Karmasin und Karmasin (1977) nicht an: jedoch sollten komplizierte Nebensatzformulierungen und adverbiale Konstruktionen vermieden werden.

Vielfach wird in der Literatur ein kurzer Satzbau empfohlen (vgl. Karmasin & Karmasin, 1977 oder Kreutz & Titscher, 1974). Oppenheim (1966) setzt die Obergrenze für die durchschnittliche Satzlänge bei 20 Wörtern an. In Analysen wurden Korrelationen von 0.54 bzw. 0.71 zwischen Satzlänge und freier Reproduktion festgestellt.

Die syntaktischen Merkmale bei der sprachlichen Formulierung wurden im Zusammenhang mit der Reliabilität eines psychologischen Verfahrens untersucht. Die Komplexität wurde neben einigen anderen Eigenschaften besonders genau geprüft. Es zeigte sich, dass der Komplexitätsgrad der sprachlichen Formulierung

die Reliabilität beeinflusst. Für Mummendey und Grau (2008) wäre aber die Reduzierung auf eine äußerst simple Form der falsche Schluss. Denn die syntaktische Komplexität könnte durchaus auf ein semantisches Niveau der Aussage verweisen.

Mummendey und Grau (2008) verstehen den Prozess des Verständnisses als einen zeitabhängigen, bei dem es darum geht, einen Übergang von der Oberflächenstruktur zur Tiefenstruktur zu vollziehen. Verschiedene grammatische Merkmale können den Abstand zwischen dem Kern eines Satzes und der Oberflächenstruktur, in der diese Aussage dem Leser erscheint, erhöhen. Eine Minimierung des Abstandes zwischen Oberfläche und Tiefe ist deshalb nur konsequent.

2.3 Methoden der Entscheidung

2.3.1 Grundbegriffe

Der Entscheidungsbegriff wird durch Feger (1968, S. 9) folgendermaßen dargestellt: Es „soll unter einer Entscheidungssituation der Abschnitt auf dem biographischen Kontinuum eines Individuums verstanden werden, der in dem Augenblick beginnt, in dem sich für ein Individuum mindestens zwei Möglichkeiten des Sich-Verhaltens eröffnen und in der im Augenblick einen (durchaus nicht endgültigen und vollständigen) Abschluß findet, in dem sich das Individuum entschließt, einer der sich bietenden Alternativen den Vorzug zu geben“.

Eine allgemeine Definition geben Häcker und Stapf (2004), wobei die Entscheidung als Wahl einer Handlungs- oder Reaktionsmöglichkeit in einer Situation, in der mehrere Möglichkeiten bestehen, verstanden wird.

Entscheidungen sind keine isolierten Aktionen sondern stehen immer im Kontext zu Freiheit, Zeitdruck und Verantwortung. Dieser Kontext beeinflusst die Wahrscheinlichkeit, eine optimale Entscheidungsstrategie zu wählen.

Die Entscheidungsforschung im eigentlichen Sinne ist nur einige Jahrzehnte alt, sie hat aber ältere Wurzeln in der Philosophie, Ökonomie und Mathematik. Die philosophischen Ansätze gehen auf Jeremy Bentham (1748 – 1832) zurück, der Handlungen nur von deren Konsequenzen her bewertete. Aus ökonomischer Sicht beschäftigte sich Adam Smith (1723 – 1790). Er sah den Menschen als egoistisch handelnden, der gerade deshalb zum allgemeinen Wohl betrug. Die mathematischen Wurzeln hängen unweigerlich mit Jacob Bernoulli (1654 – 1705) und Pierre Simon de Laplace (1749 – 1827) zusammen, die sich mit Wahrscheinlichkeiten im Glücksspiel auseinander setzten.

Die Etablierung im Wissenschaftsbereich erfolgte durch das 1947 erschienene Buch „Spieltheorie und wirtschaftliches Verhalten“ von John Neumann und Oskar Morgenstern (Jungermann, Pfister & Fischer, 2010).

Nach Berg (1972) ist die Spieltheorie sowohl für die Beschreibung und Analyse realer Verhaltensweisen (deskriptives Element) geeignet, als auch für die operationale Verwendung in realen Entscheidungssituationen (normatives Element).

Der Mittelpunkt der Entscheidungsforschung ist heute ohne Zweifel in der Ökonomie zu finden. In der Betriebswirtschaftslehre steht dabei das Interesse im Vordergrund, Grundlagen und Verfahren für die Optimierung von wirtschaftlichen Entscheidungsprozessen zu entwickeln und zur Verfügung zu stellen (Eisenführ & Weber, 2003).

Bei der Untersuchung von Entscheidungsverhalten werden zwei Ansätze, die deskriptive und die präskriptive Entscheidungstheorie, verfolgt. Erstere widmet sich der Erklärung und Beschreibung des tatsächlichen Entscheidungsverhaltens von Menschen, wobei eine Vielzahl systematischer Fehler aufgedeckt wurde, die zu schlechten intuitiven Entscheidungen führen können. Der zweite, auch normative Entscheidungstheorie genannte Ansatz, unterstellt dem entscheidenden Menschen, dass sich dieser an Maximen rationalen Handelns orientiert. Präskriptiv wird Entscheidungsverhalten dadurch erklärt, indem analysiert wird, wie sich eine Person bei einer gegebenen Entscheidungssituation und einem vorhandenen Zielsystem entscheiden sollte (Manz, Dahmen & Hoffmann, 2000).

Ausgangspunkt der präskriptiven Theorie ist die Tatsache, dass der Mensch Schwierigkeiten mit neuartigen, nicht routinierten Entscheidungssituationen hat. Nach Eisenführ und Weber (2003) liegen diese Schwierigkeiten in der Unsicherheit der Zukunft, der verschiedenartigen Konsequenzen bei den unterschiedlichen Handlungen und Anzahl der zur Verfügung stehenden Alternativen. Weiters wird die Situation umso komplexer, je mehr unsichere Einflüsse auf das Ergebnis einwirken und je mehr Ziele zu beachten sind. Dieser Grad der Komplexität in Entscheidungssituationen kann zu Schwierigkeiten führen.

Die Auswirkungen der Entscheidung lassen sich in der Regel nie mit Sicherheit voraussagen. Je größer die Unsicherheit ist, umso schwieriger wird im Allgemeinen die Behandlung der Entscheidungssituationen (Bühlmann, 1969).

Voraussetzungen für eine Entscheidung sind Wissen und Motivation. Der Entscheider muss Wissen aktivieren, um infrage kommende Optionen zu sammeln, Erfahrungen von ähnlichen Situationen auf Neue zu transferieren und die Konsequenzen der jeweiligen Optionen zu erkennen. Außerdem muss er eine Lösung selbst und nach seinen eigenen Vorstellungen herbeiführen wollen. Entscheidungen setzen also Wissensaktivierung und Motivationsdynamik voraus. Oft sind sie durch einem emotionalen Kontext eingefärbt (Jungermann et al., 2010). Debus (1977) differenziert die Emotionen nach ihrer Intensität, der Qualität und dem zeitlichen Verlauf. Sie lösen beim Entscheider einen unterschiedlichen Grad der Aktivierung aus: z.B. im motorischen Verhalten verändert sich Stärke und Schnelligkeit der Reaktionen.

2.3.2 Merkmale einer Entscheidung

Entscheidungen können von einer Person (individuelle Entscheidungssituation) oder in Interaktion mit einem Partner bzw. einer Gruppe (soziale Entscheidungssituation) getroffen werden. Individuelle Entscheidungen sind

Gegenstand der klassischen Ökonomie und Statistik, soziale Entscheidungen werden in der Spieltheorie und Sozialpsychologie behandelt (Jungermann, 2005).

Nach Bühlmann (1969) werden vier Arten von Entscheidungssituationen unterschieden:

- a) Risikosituation
- b) Ungewissheitssituation oder Unsicherheit
- c) Sicherheitssituation
- d) Spielsituation

Zu a) In Risikosituationen ist der Entscheider in der Lage, die einzelnen Umweltzustände nach subjektiven und objektiven Eintrittswahrscheinlichkeiten zu ordnen. Damit keine Informationen vernachlässigt werden, müssen bei der Auswahl der optimalen Handlungsalternative die unterschiedlichen Eintrittswahrscheinlichkeiten Berücksichtigung finden (Manz et al., 2000).

Zu b) Von Unsicherheitssituationen spricht Bitz (1981), wenn ein Zustand durch zufällige Ereignisse eintritt (z.B. den Lauf einer Roulette-Kugel). Die relevanten Umweltbedingungen sind dem Entscheider nicht bekannt und er hat keine Möglichkeit, die Ereignisse in Form bekannter Wahrscheinlichkeitsverteilungen wahrzunehmen (Berg, 1972).

Zu c) Entscheidungen bei Sicherheit sind dadurch gekennzeichnet, dass der Entscheider Konsequenzen seiner Handlungen sowie die eintretenden Umweltzustände mit Sicherheit kennt. Manz et al., 2000).

Zu d) In Spielsituationen wird die Entscheidung durch die Wahl von Strategien rational handelnder Gegenspieler bestimmt (Bitz, 1981).

Optionsmenge

Eine weitere Unterscheidung kann in Hinblick darauf getroffen werden, ob zu Beginn eines Entscheidungsprozesses die Menge der Optionen, zwischen denen eine Wahl getroffen werden kann, bereits vorliegt oder nicht. So steht die Anzahl der Optionen bei einem Autokauf im Vorhinein fest (z.B. Anschaffungskosten,

Verbrauch, Farbe, Aufbau, usw.). Dagegen müssen bei der Konstruktion einer Maschine die Optionen erst generiert werden (Jungermann et al., 2010).

Der Zweck des Entscheidungsprozesses ist, die beste erreichbare Alternative zu wählen. Damit verbunden ist die Suche nach den vorhandenen Wahlmöglichkeiten. Nach Eisenführ und Weber (2003) unterbleibt in der Praxis häufig die systematische Suche nach Alternativen. Voreilig legen sich die involvierten Personen auf etwaige Alternativen fest, wobei neue Vorschläge später eher als störend empfunden werden.

Entscheidungsstufen

Nach Jungermann (2005) können Entscheidungen einstufig oder mehrstufig ablaufen. Bei Entscheidungen, die in mehreren Schritten (Szenarios) vollzogen werden, hängt jeder weitere Schritt vom vorherigen Ergebnis ab.

Entscheidungshäufigkeit

Ein weiteres Merkmal von Entscheidungen ist, ob sie ein einziges Mal oder wiederholt getroffen werden. Für Svenson (1996) ist der Prozess der Entscheidungsfindung in bestimmten Situationen von der Ein- oder Mehrmaligkeit der Entscheidung abhängig.

Kognitiver Aufwand

Die Beschäftigung mit Entscheidungsproblemen hängt mit dem damit verbundenen kognitiven Aufwand zusammen, wie Informationen repräsentiert sind oder ob überhaupt mentale Repräsentationen vorliegen. Ablaufende Entscheidungen fordern bei der Informationssuche und –verarbeitung kognitive Anstrengungen, je nach dem, wie die Nutzung der kognitiven Ressourcen stattfindet (Jungermann et al., 2010).

Der kognitive Aufwand wird in vier Ebenen unterteilt:

- *routinisierte Entscheidungen*: Diese Entscheidungen fordern den geringsten kognitiven Aufwand. Die gegebene Situation wird mit gespeicherten Situationen verglichen (Matching).

- *stereotype Entscheidungen*: Entscheidungen werden einerseits durch die Art der möglichen Entscheidungsoptionen ausgelöst, andererseits kommt es zu einem minimalen Bewertungsprozess (trade-offs).
- *reflektierte Entscheidung*: Der Entscheider sucht nach Informationen in seinem Gedächtnis und / oder in seiner Umwelt. Dabei werden diese Informationen nach Präferenzen geordnet.
- *Konstruktive Entscheidung*: Den höchsten kognitiven Aufwand fordern diese Entscheidungen, da erstens die Optionen nicht vorgegeben oder nicht hinreichend definiert sind und zweitens die relevanten Werte unklar sind oder erst generiert werden müssen.

Entscheidungskriterium

Die Entscheidungsregel soll die Lösung des Entscheidungsproblems ermöglichen. Im Gegensatz dazu führt das Entscheidungsprinzip grundsätzlich nicht zur Lösung des Problems. Das Entscheidungsprinzip legt die Präferenzen einer Entscheidung nicht eindeutig fest, sie gibt nur die Richtlinie für die Entscheidungsregel vor. Entscheidungskriterium wird als Oberbegriff für Entscheidungsprinzip und Entscheidungsregel verwendet (Laux, 2003).

Jungermann (2005) nimmt noch eine Unterteilung der Entscheidungsregeln in non-kompensatorische und kompensatorische vor. Bei der non-kompensatorischen Regel wird eine Option nicht gewählt, wenn sie nicht bestimmte Merkmale besitzt. Die Anwendung der kompensatorischen Regel erfolgt, wenn eine schlechte Ausprägung auf einem Attribut durch ein Attribut mit guter Ausprägung kompensiert werden kann.

Für Bronner (1989) ist die Wahrscheinlichkeit, dass Entscheidungen erkannt, aufgegriffen und bewältigt werden, umso geringer

- je stärker Vorurteile oder Gewohnheiten das Verhalten prägen,
- je weniger sich der Entscheider mit dem Alternativen-Konflikt auseinandersetzt
- je geringer die Bereitschaft des Entscheiders für die Risikoübernahme bei den verschiedenen Alternativen ist,

- je negativer die Sanktionen der Entscheidung sind,
- je geringer die persönliche Betroffenheit ist,
- je höher die Ungewissheit durch unzureichende oder widersprüchliche Informationen ist.

2.3.3 Komponenten einer Entscheidung

2.3.3.1 Persönlichkeitsspezifische Komponenten

Interferenzneigung

Hörmann (1960) hatte nachgewiesen, dass einige Aspekte des Entscheidungsverhaltens von der Interferenzneigung beeinflusst werden. Interferenz liegt bei der Wahrnehmung vor, wenn über die tatsächlich gegebene Information hinaus, auf weitere nicht beobachtbare Eigenschaften des Wahrnehmungsobjekts, Schlüsse gezogen werden. Personen mit hoher oder niedriger Interferenzneigung unterscheiden sich im Entscheidungsverhalten in der Menge der Information, die bis zur Entscheidung verlangt wurde. Ebenso konnten Unterschiede bei den Entscheidungszeiten festgestellt werden je ähnlicher die Alternativen wurden. Schließlich zeigten sich noch Differenzen im Selbstvertrauen der Richtigkeit des eigenen Urteils. In Murrelversuchen mussten Versuchspersonen eine Murrel aus einem Beutel mit zwei verschiedenfarbigen Murrel ziehen und die Wahrscheinlichkeiten dafür schätzen, wobei sich herausstellte, dass Niedrig-Interferenzler die subjektive Wahrscheinlichkeit mehr zu den Extrempunkten verschoben (Goldsmith, 1967).

Neurotische Tendenz

Goldsmith (1967) hat Entscheidungsverhalten auf neurotische Tendenzen hin untersucht. Neurotische Personen neigen auf Grund ihrer emotionalen Labilität bei der Auswahl zur Intransitivität. Die neurotische Tendenz wurde mit einem psychologischen Verfahren von Brengelmann (1960) gemessen.

2.3.3.2 Inhaltliche Komponenten

Optionen

Optionen sind nach Jungermann et al. (2010) Gegenstände, Handlungen, Regeln oder Strategien, zwischen denen gewählt werden kann. Sie müssen vom Entscheider gesucht oder entwickelt werden. Manchmal sind Optionen aber bereits vorhanden. Damit nicht unnötig Ressourcen (Zeit und Geld) in die Suche nach Optionen verschwendet werden, erfolgt zuvor eine Abklärung des Zielsystems (Eisenführ & Weber, 2003). Kauf- und Standortentscheidungen sind typisch für Situationen, bei denen eine Wahl stattfindet. Handlungen sind meist auf ein konkretes Ziel gerichtet, welches durch die Handlung verwirklicht werden soll. Der Entscheider muss dafür die Fähigkeiten für die Ausführung besitzen. Eine spezielle Handlungsoption wird bei der Überprüfung des Status quo verlangt, wenn der momentane Zustand fortgesetzt oder abgebrochen werden soll.

Strategische Optionen sehen keine konkreten Handlungspläne vor, sondern werden durch allgemeine Wertvorstellungen des Entscheiders dominiert. Eine eher einfache Art Entscheidungen zu treffen, wird durch Regeln bewältigt (Jungermann et al., 2010).

Ereignisse

Ereignisse sind Vorkommnisse und Sachverhalte, die den Ausgang einer Entscheidung beeinflussen können, auf die der Entscheider aber keinen Einfluss nehmen kann, obwohl er oft Erwartungen bezüglich der Art und Wahrscheinlichkeit des Ereignisses hat (Jungermann, et al., 2010). Ereignisse die in der natürlichen und sozialen Umwelt des Entscheiders stattfinden, werden als externe Ereignisse bezeichnet. Dagegen gelten Ereignisse, welche in der Person selbst erfolgen, als interne Ereignisse.

In Entscheidungssituationen rufen Ereignisse Unsicherheit hervor, wenn die Folgen einer Wahl die Option beeinflussen können. Diese Unsicherheit liegt den meisten Entscheidungsproblemen zugrunde, weil die Folgen einer Entscheidung nachgeordnet sind.

Konsequenzen

Konsequenzen sind Zustände, die sich als Folge der Wahl einer Option ergeben. Optionen werden meist wegen ihrer Konsequenzen gewählt. Dies ist eine Grundannahme der Entscheidungstheorien. Implizit oder explizit werden diese Konsequenzen bewertet und auf ihren subjektiven Nutzen (Wert) geprüft. Kahneman und Snell (1992) unterscheiden zwischen vorhergesagtem und erlebtem Nutzen, denn der vorgestellte Nutzen zum Zeitpunkt der Entscheidung stimmt nicht immer mit dem später tatsächlich erlebten Nutzen überein.

Der Nutzen einer Entscheidung ist im Gegensatz zur gewählten Alternative nicht beobachtbar. Die Beurteilung des Nutzens von Konsequenzen stellt die eigentliche Grundlage für die Entscheidung dar (Jungermann et al. 2010). Wenn die Konsequenzen sicher sind, so basiert die Wahl allein durch die Bewertung des Nutzens; sind die Konsequenzen unsicher, so erfolgt die Wahl auf Grund von Wahrscheinlichkeiten. Der subjektive Wert einer Konsequenz wird als Nutzen bezeichnet.

Manchmal werden aber auch Entscheidungen gesucht, deren Basis einer ethischen Qualität zugrunde liegt und die nicht durch Konsequenzen bestimmt werden (Jungermann et al., 2010).

In einigen Entscheidungssituationen erscheinen die unmittelbaren Konsequenzen sehr klar und einfach. Typischerweise handelt es sich meist um monetäre Beträge. Hingegen sind in anderen Entscheidungssituationen die Konsequenzen multidimensional: d.h. das Resultat der Entscheidung hat viele Attribute. Welche Attribute als relevant gehalten werden, hängt von den Zielen ab.

Ziele

Die Menge an Optionen und Konsequenzen schränken die Ziele ein. Der Entscheider schränkt durch seine Ziele die in Frage kommenden Möglichkeiten aus dem „Möglichkeitsuniversum“ ein (Jungermann, et al., 2010). Umgekehrt hängen die vom Entscheider in Betracht gezogenen Alternativen unter anderem davon ab, an welchen Zielen er sich orientiert (Laux, 2003).

Gründe

Gründe spielen für den Entscheider in zweifacher Hinsicht eine Rolle: Sie können die Entscheidung in eine Richtung lenken, die aus den vorher erwähnten Komponenten nicht ableitbar ist. Entscheidungen können davon abhängen, wie sie anderen Personen gegenüber begründet werden können, z.B. durch überzeugende Argumente.

Entscheidungsfeld

Die Gesamtheit der erfassten Alternativen, Umweltzustände bzw. Eintrittswahrscheinlichkeiten sowie alternative Ergebnisse werden als Entscheidungsfeld bezeichnet (Bitz, 1981).

2.3.4 Entscheidungsprobleme

Die Entscheidungstheorie behandelt die Probleme der Entscheidungsfindung theorieabhängig. D.h. in der deskriptiven Entscheidungstheorie werden Hypothesen über das Verhalten von Individuen im Entscheidungsprozess generiert, mit deren Hilfe Entscheidungen prognostiziert werden. Sie versucht zu erklären, wie Personen tatsächlich in der Realität entscheiden. Deskriptive Theorien sind Aussagesysteme, die in Untersuchungen erarbeitet werden. Probleme ergeben sich hinsichtlich der realisierbaren Alternativen in Entscheidungssituationen und zu welchen Konsequenzen sie führen (Laux, 2003).

In der präskriptiven Entscheidungstheorie können rationale Entscheidungen nur dann getroffen werden, wenn Zielvorstellungen für die Beurteilung der Alternativen existieren (Laux, 2003). Präskriptive Theorien beschreiben nicht die Realität, sondern geben Empfehlungen für alternative Entscheidungssituationen.

Nach Eisenführ und Weber (2003) können die Strukturen eines Entscheidungsproblems in den Alternativen, Umwelteinflüssen, Konsequenzen sowie in den Zielen und Präferenzen des Entscheiders beschrieben werden.

Die Suche nach Alternativen wirft stets die Frage auf, wann der Prozess beendet und die Entscheidung getroffen werden soll. Durch Zeit- oder Budgetvorgaben wird der Suchprozess begrenzt.

2.3.5 Entscheiden unter Unsicherheit

2.3.5.1 Unsicherheit

In den meisten Entscheidungssituationen herrscht das Problem vor, dass man nicht mit Sicherheit sagen kann, wie die Zukunft aussehen wird. Nach Jungermann (2005) finden alle Entscheidungen unter Unsicherheit statt, weil mögliche Konsequenzen in der Zukunft liegen und von Faktoren abhängen, die der Entscheider nicht kontrollieren kann.

Unsicherheit liegt dann vor, wenn mindestens zwei Umweltzustände für möglich gehalten werden. Der Entscheider kennt zwar, welche entscheidungsrelevanten Umweltzustände existieren, nicht aber mit welcher Wahrscheinlichkeit sie eintreten werden.

Ein wichtiges Thema der Entscheidungsforschung ist die Bildung, Revidierung und die Berücksichtigung der subjektiven Wahrscheinlichkeit des Menschen. Jeder Mensch ist tagtäglich mit Entscheidungen konfrontiert, bei denen er eine gewisse Anzahl von Alternativen – verbunden mit den dazugehörigen Konsequenzen – hat. Angefangen bei der Wahl der Bekleidung, wenn die Wettervorhersage ungenau ist oder der Kauf eines Gebrauchtwagens, dessen tatsächliche Substanz unbekannt ist. Die Unsicherheit ist ungleich größer, wenn die Konsequenzen weit reichende Folgen haben können und die Entscheidung nicht mehr rückgängig gemacht werden kann wie bei Operationen, Investitionen oder etwa strafrechtlichen und politischen Entscheidungen. Bei Entscheidungen die irreversibel sind, muss ein gewisses Maß an Risiko eingeplant werden, welches bei Entscheidungen zu Unsicherheit führt. In vielen Entscheidungssituationen ist die Unsicherheit das

Hauptproblem. Wird die Unsicherheit explizit berücksichtigt, so muss sie in einem Modell abgebildet werden. Modelle sind durch Tatbestände gekennzeichnet, denen Ereignisse oder Zustände zu Grunde liegen (Eisenführ & Weber, 2003). Ein Beispiel für einen unsicheren Tatbestand wäre: Wie wird das Wetter morgen? Die dazugehörigen Zustände sind sonnig, bewölkt oder regnerisch usw.

Oft kommen Zustände nicht isoliert vor, sondern sind Teil einer komplex zusammenwirkenden Umwelt. Die Entscheidung für die Einführung eines neuen Produktes hängt zum Beispiel von der Konkurrenz oder vom Währungskurs des Landes ab. Eisenführ und Weber (2003) nennen die aus Einzeleinflüssen zusammengesetzten Ereignisse „Szenarien“. Eine große Anzahl von Einflüssen ist bei komplexen Entscheidungen zu beobachten: die Darstellung der Unsicherheit gelingt dann mittels „Szenario-Analyse“.

Russo und Schoemaker (1989) geben eine Reihe von Methoden an, die zu einer verbesserten Modellierung der Unsicherheit beitragen:

- Auflistung aller Pro- und Contra-Argumente zu Wahrscheinlichkeitsaussagen
- Erstellung von Ursachenbäumen
- Kreieren optimistischer und pessimistischer Szenarien
- Antizipation eines zukünftigen Fehlschlags und rückblickender Erklärungen, woran es gelegen haben könnte.

Unsicherheit wird durch verbale Äußerungen oder numerischen Feststellungen ausgedrückt. Je nach Kontext werden Wörter (vielleicht, möglicherweise, sicher, usw.) verwendet, die sich auf Häufigkeiten beziehen, mit denen Ereignissen auftreten oder mit denen sie erwartet werden. Präziser wird Unsicherheit durch numerische Feststellungen ausgedrückt, z.B. 5%, die Wahrscheinlichkeit ist die wichtigste Möglichkeit Unsicherheit auszudrücken (Jungermann et al., 2010).

Die Beziehung von numerischen zu verbalen Ausdrücken wurde von Fischer und Jungermann (1996) untersucht. Erstens interpretieren Menschen die gleichen Ausdrücke höchst unterschiedlich, zweitens decken Worte eine sehr

unterschiedliche Wahrscheinlichkeitsbreite ab und drittens spielt der Kontext des verbalen Ausdrucks für die Interpretation eine große Rolle.

Koriat, Lichtenstein und Fischhoff (1980) haben ein Prozessmodell entwickelt, das zur Entstehung von subjektiver Sicherheit beiträgt. Zuerst wird beim Lösungsprozess mit der Suche nach relevanten Informationen begonnen. Daraus entsteht ein subjektives Sicherheitsgefühl, welches zum Schluss in eine numerische Antwort übersetzt wird.

Unsicherheit ist als verbaler Ausdruck anzusehen. In der Wahrscheinlichkeitstheorie wird Unsicherheit mathematisch formalisiert und in einen quantitativen Ausdruck transferiert (Jungermann, et al., 2010).

Menschen sind es jedoch nicht gewohnt und sträuben sich sogar dagegen, Wahrscheinlichkeiten präzise zu quantifizieren. Für die Unterstützung schwerwiegender Entscheidungen unter Unsicherheit sollten aber Wahrscheinlichkeiten numerisch bestimmt werden (Eisenführ & Weber, 2003).

Wahrscheinlichkeiten nach dem klassischen Wahrscheinlichkeitskonzept sind intersubjektiv überprüfbar und werden als objektive Wahrscheinlichkeit bezeichnet. Solche Wahrscheinlichkeiten sind in der Realität selten gegeben.

In vielen Entscheidungssituationen bestehen über das Eintreten bestimmter ungewisser Ereignisse Vorstellungen, die keiner objektiven Wahrscheinlichkeit entsprechen. Diese gelten als subjektive Wahrscheinlichkeiten. Subjektive Vorstellungen und Überzeugungen können zu subjektiven Wahrscheinlichkeiten quantifiziert werden (Laux, 2003).

Neben dem Konzept der Wahrscheinlichkeit gibt es weitere Möglichkeiten der Formalisierung der Unsicherheit, gekennzeichnet dadurch, dass sie weniger schweren Restriktionen unterliegen als die Wahrscheinlichkeit. Solche neuartigen Konzepte sind nach Jungermann et al. (2010), das Konzept der „Gewissheitsfaktoren“, die „Evidenz-Theorie“ und die „Möglichkeits-Theorie“.

Diese Konzepte werden kurz vorgestellt, vertieft können sie bei Jungermann et al. (2010) nachgelesen werden.

Gewissheits-Faktoren

Gewissheitsfaktoren sind Expertenurteile über die subjektive Gewissheit, mit der eine Hypothese (DANN-Teil) angenommen wird, sofern eine bestimmte Evidenz (WENN-Teil) vorliegt.

Evidenz-Theorie

In der Evidenz-Theorie wird die Unsicherheit durch ein so genanntes Plausibilitätsintervall beschrieben. Evidenz beinhaltet alles Wissen und alle Vermutungen, die für und gegen ein Ereignis sprechen. Drei Repräsentationen von Unsicherheit werden zugelassen.

Möglichkeits-Theorie

Sprachliche Ausdrücke (unwahrscheinlich, möglicherweise, usw.) für Unsicherheit können als eine unscharfe Menge einer Verteilung verstanden werden. Die Verteilung wird in einer Zugehörigkeitsfunktion abgebildet, wobei das Intervall zwischen 0 und 1 liegt.

Wenn Menschen in Entscheidungssituationen Unsicherheit erleben, ohne dass dabei auf bestimmte Wahrscheinlichkeitsangaben zur Berechnung zurückgegriffen werden kann, erfolgt die Entscheidung „intuitiv“ (Jungermann et al., 2010). Teigen (1994) aber auch Jungermann (1997) haben eine Differenzierung solcher Intuitionen, also Varianten von Unsicherheit (subjektive Wahrscheinlichkeit) vorgeschlagen. Eine erste wichtige Unterscheidung bezieht sich darauf, ob die Ursache für Unsicherheit auf externe oder interne Faktoren zurückzuführen ist. Bei einer Lottoziehung wird die Unsicherheit extern attribuiert. Die Zahlenggebung basiert allein auf der Apparatur. Wenn wir sagen „Kabul ist die Hauptstadt der Mongolei“ (Anm.: tatsächlich Ulan Bator), so beruht Unwissenheit auf unserem fehlenden Wissen. In diesem Fall wird Unsicherheit internen Ursachen zugeschrieben.

Die externe Unsicherheit kann wiederum in frequentistische und tendenzielle Unsicherheit gegliedert werden. Bei ersteren sind die relativen Häufigkeiten

bekannt oder können geschätzt werden und die verschiedenen Ereignisse hängen von externen Faktoren ab. Das ist z.B. beim (Glücks-)Spiel der Fall.

Tendenzielle Unsicherheit ist ein Spezialfall und tritt nur bei einmaligen Ereignissen auf. Die Beurteilung einer Wahrscheinlichkeit ist in etwa die Einschätzung der Tendenz eines Ereignisses, tatsächlich einzutreten. Nach Jungermann et al., (2010) bringen Formulierungen, wie „der Tumor ist wahrscheinlich tödlich“, Tendenzen zu Ausdruck.

Bei internen Unsicherheiten kann Unsicherheit direkt, als unmittelbarer Ausdruck eines Gefühls, entstehen. Die Unsicherheit ergibt sich aus dem Gefühl der Stärke einer gespeicherten assoziativen Verknüpfung. Die subjektive Sicherheit (zu) einer (genaueren) Aussage wird durch Introspektion erhöht.

Unsicherheit kann aber auch aus Argumenten abgeleitet werden. Schwache oder starke Argumente veranlassen uns, an Aussagen zu glauben. In abgeänderter Form wäre nun „Kabul ist die Hauptstadt der Mongolei, ich bin mir aber nicht sicher“ ein Beispiel für abgeleitete Unsicherheit. Die Urteile werden kognitiv argumentativ abgeleitet, d.h. Wissen wird überprüft und aus diesem ein Schluss gezogen.

Zwei weitere Unterscheidungen von Unsicherheit, welche den erwähnten Konzepten nicht zuordenbar sind, wurden von Teigen (1994) vorgenommen. In der ersten Variante hängt die Wahrnehmung der Wahrscheinlichkeit von externen Ereignissen, von der subjektiven Kontrollvorstellung der Ereignisse ab. Die zweite Variante wird durch die wahrgenommene Plausibilität des Entscheiders geformt. Das Verhalten von Hauptdarstellern in Actionfilmen wird oft, weil unplausibel, als äußerst unwahrscheinlich empfunden. Der Grad der Plausibilität einer Geschichte oder Behauptung bestimmt die (Un-)Sicherheit.

2.3.5.2 Entscheidungsmodelle

SEU-Modell

Das SEU-Modell (Subjective Expected Utility) wurde von Edwards (1954) genannt. In ihm werden die Faktoren Unsicherheit und Nutzen als subjektive Größen behandelt. Nach diesem Modell wählt ein Entscheider jene Option, die den höchsten subjektiv erwarteten Nutzen hat. Der Wert einer Option ist die Summe der Nutzenwerte der einzelnen möglichen Konsequenzen, gewichtet mit den Wahrscheinlichkeiten ihres Eintretens (Jungermann et al., 2010). Es wird angenommen, dass die Wahrscheinlichkeiten der möglichen Konsequenzen der zur Verfügung stehenden Optionen in einer bestimmten Weise kognitiv integriert werden und dass die Entscheidung von Nutzen und Wahrscheinlichkeit der einzelnen Konsequenzen abgeleitet wird.

Für die Anwendung des SEU-Modells müssen gewisse Voraussetzungen durch den Entscheider erfüllt werden. Savage (1954) legt vier Axiome vor, die das SEU-Modell fundieren:

- Vergleichbarkeit
- Transitivität
- Unabhängigkeit
- Kontinuität

Ist eine der genannten Bedingungen verletzt, so entscheidet eine Person nicht nach dem SEU-Modell, es findet keine Maximierung des Nutzens statt. Die Überprüfung der Axiome führt zu einer Einschränkung des Modells, da die Axiome keine allgemeine Gültigkeit beanspruchen können. Insofern kann das SEU-Modell nur unter spezifischen Bedingungen Entscheidungsverhalten erklären.

Prospect-Theorie

Mit der Prospect-Theorie wurde von Kahneman und Tversky (1979) eine Revision der SEU-Theorie vorgenommen. Sie gehen auch davon aus, dass die Konsequenzen einer Option und deren Wahrscheinlichkeiten die Entscheidung

determinieren, jedoch macht sie spezifische Annahmen zur Transformation der Konsequenzen und der Wahrscheinlichkeiten. Die Entscheidung läuft in zwei Prozessen ab. In der ersten Phase wird das Problem nach bestimmten Regeln enkodiert, transformiert und schließlich mental repräsentiert. In der zweiten Phase erfolgt dann eine Bewertung über den subjektiven Wert jeder Option. Die Option mit dem höchsten subjektiven Wert wird gewählt.

Aus den mentalen Repräsentationen werden Erklärungen für das nicht adäquate Entscheidungsverhalten von Personen nach der klassischen SEU-Theorie abgeleitet. Nach Kahneman und Tversky (1979) führt die Präsentation eines Entscheidungsproblems zu einer mentalen Repräsentation dieses Problems. Dieses bezeichnen sie als „decision frame“ und verstehen es als interne Repräsentation der relevanten Komponenten des spezifischen Entscheidungsproblems. Durch die Verarbeitung von Informationen und Wissensaktivierung wird vom Entscheider der „decision frame“ aufgebaut.

Die Theorie ist wie das SEU-Modell axiomatisch begründet, nur das Axiom der Vergleichbarkeit wurde durch das Axiom der Vollständigkeit ausgetauscht. In Entscheidungssituationen kann die Prospect-Theorie nur Optionen mit bis zu drei Konsequenzen Entscheidungsverhalten erklären.

Weitere Theorien

In der Disappointment-Theorie wird dem Ansatz Rechnung getragen, dass nicht nur der spezifische Nutzen einer Konsequenz relevant ist, sondern dass der Nutzen auch von Erwartungen abhängt. Bei positive Erwartungen wird Freude und bei negativen Enttäuschung empfunden.

Bei der Regret-Theorie werden nicht nur Konsequenzen jeder Option durch den Entscheider antizipiert, sondern er assoziiert auch Emotionen, die aus dem Vergleich mit verpassten Konsequenzen resultieren. Positive Emotionen empfindet (Freude) der Entscheider bei verpassten schlechten Konsequenzen und negative (Bedauern) bei besseren Konsequenzen als die realisierten.

Bei beiden Theorien werden emotionale Komponenten miteinbezogen.

In rangabhängigen Nutzenmodellen werden die Wahrscheinlichkeiten von Konsequenzen nach dem Wert bzw. Nutzen der möglichen Ereignisse gewichtet. Es wird angenommen, dass die Wahrscheinlichkeit über- oder untergewichtet werden, abhängig davon, ob die assoziierte Konsequenz eher erwünscht oder unerwünscht ist (Jungermann et al., 2010).

Camerer (1992) untersuchte die vorgestellten Modelle und schlussfolgerte:

- dass keine Theorie eine universelle Geltung beanspruchen kann
- bei extremen Wahrscheinlichkeiten (nahe 0 oder nahe 1) ist die Prospect-Theorie besser anwendbar; bei mittleren Wahrscheinlichkeiten ist die SEU-Theorie besser
- bei negativen Konsequenzen kommt es zu systematischer Asymmetrie (Risiko-Freude bei Verlusten, Risiko-Aversion bei Gewinnen), nur Theorien mit einem Referenzpunkt können zwischen Gewinn und Verlust differenzieren (z.B. Prospect-Theorie)

2.3.5.3 Entscheidung und Informationsverarbeitung am Bildschirm

Gertzen (1990) untersuchte die Informationsverarbeitung bei Entscheidungsprozessen am Bildschirm. Er stellte fest, dass nur ein Teil der verfügbaren Information tatsächlich bearbeitet wird - selektive Informationsverarbeitung. Erhöhen sich die Alternativmöglichkeiten, so nimmt die prozentuale Informationsnutzung ab. Nach Russo und Doshier (1983) gelingt zwar die Betrachtung von Informationen am Bildschirm gut, die Ausführung von expliziten Abrufoperationen misslingt aber sehr oft.

Bei großer Informationsmenge sinkt die Qualität der Entscheidung und die verfügbaren Informationen werden nur mehr selektiv genützt.

Bei sehr komplexen Aufgaben kommt es zu einer überproportionalen Verkürzung der Bearbeitungszeit. Onken, Hastie und Revelle (1985) interpretieren dies als

Hinweis darauf, dass auf die hohe kognitive Belastung und den hohen Komplexitätsgrad mit vereinfachenden Strategien reagiert wird.

Personen versuchen, ihren kognitiven Aufwand bei der Entscheidungsfindung gering zu halten. Deshalb werden statt normativer oft heuristische Regeln eingesetzt. Informationen werden dadurch nicht vollständig genützt.

In Hinblick auf die kognitive Anforderung in Entscheidungssituationen soll die Informationsverarbeitung bei komplexen Aufgaben seriell erfolgen (Huber, 1982). Gertzen (1990) liefert Argumente, warum am Bildschirm eine sequenzielle Darbietung von Informationen erfolgen soll:

- Personen sind es gewohnt, Informationen seriell aufzunehmen.
- Technische Systeme werden von alternativenweisen Sequenzierung dominiert.
- Alternativen können hinsichtlich ihrer Stimmigkeit im Gesamteindruck überprüft werden.

2.3.6 Wahrscheinlichkeit

Unsicherheit wird in der Wahrscheinlichkeitstheorie durch Zahlen formalisiert, völlig unabhängig von der interpretativen Bedeutung dieser Zahlen (Jungermann et al., 2010). Meistens kann davon ausgegangen werden, dass das Auftreten von Ereignissen nicht gleich wahrscheinlich ist.

Es wurden drei Interpretationen des Wahrscheinlichkeitsbegriffs vorgeschlagen:

1. Der Begriff der Wahrscheinlichkeit eines Ereignisses ist nach der *objektivistischen* Auffassung, dass ein objektives Merkmal von materiellen Prozessen unabhängig vom Beobachter besteht. In Übereinstimmung mit der Realität kann ein Wahrscheinlichkeitsurteil einem Wahrnehmungsurteil ähnlich sein.
2. Bei der *frequentistischen* Interpretation wird durch Beobachtungen im Sinne von relativen Häufigkeiten eines Ereignisses die Wahrscheinlichkeit bezüglich einer Referenzmenge, beschrieben. „Die relative Häufigkeit kann

nur dann sinnvoll als Wahrscheinlichkeit interpretiert werden, wenn die Menge hinreichend groß ist (...)“ (Laux, 2003).

3. Nach der subjektivistischen Interpretation ist eine Angabe über die Wahrscheinlichkeit der Ausdruck einer rein subjektiven Größe. Der Grad an Gewissheit determiniert das Wahrscheinlichkeitsurteil, welches weder richtig noch falsch sein kann. Die Wahrscheinlichkeit ist abhängig vom Wissensstand und von der Informationsverarbeitung einer Person, d.h. demselben Ereignis können verschiedene Personen unterschiedliche Wahrscheinlichkeiten zuordnen. Aus subjektiver Sicht ist es unerheblich, ob das Ereignis in der Vergangenheit oder in der Zukunft liegt (Eisenführ & Weber, 2003).

Unbewusst bilden wir ständig Erwartungen über Ereignisse und verwerten diese bei der Entscheidungsfindung. Erwartungen über unsichere Ereignisse werden oft in verbaler Form ausgedrückt (dass es heute regnet, ist unwahrscheinlich). Im täglichen Leben kommt es auch meistens nicht auf eine sehr hohe Genauigkeit der Wahrscheinlichkeitsangabe an. Bei wichtigen Entscheidungen ist eine quantitative Angabe der Wahrscheinlichkeit aber sehr empfehlenswert. Sie biete einerseits die Vorteile des Vergleiches und der Diskussion, andererseits können quantitative Wahrscheinlichkeiten mathematisch verarbeitet werden. Eisenführ und Weber (2003) stellen gerade in professionellen Bereichen, wo Erwartungen über unsichere Ereignisse von entscheidender Bedeutung sind, eine Abneigung beim Verwenden von numerischen Wahrscheinlichkeiten fest. Sie nennen Beispiele aus der Medizin und der Justiz, diesbezügliche Untersuchungen wurden von Fischer und Jungermann (1996) durchgeführt.

Eisenführ und Weber (2003) bestehen nicht auf punktgenaue Wahrscheinlichkeitsangaben, da es meist schwierig ist, die exakte Wahrscheinlichkeit zu bestimmen. Aber alleine schon die Angabe von Wahrscheinlichkeitsintervallen kann bei vorliegenden Entscheidungsproblemen weiterhelfen.

2.3.6.1 Messung der subjektiven Wahrscheinlichkeit

In der Regel ist es unmöglich, numerische Wahrscheinlichkeiten abzufragen. Die meisten Menschen sind im Umgang mit Wahrscheinlichkeiten ungeübt und unerfahren. Gewöhnlich sind die gesuchten Wahrscheinlichkeiten zu Beginn nur schemenhaft vorhanden (Eisenführ & Weber, 2003).

Menges (1970) erörtert einige Einwände gegen die Messung von subjektiven Wahrscheinlichkeiten:

- A) Die persönliche Wahrscheinlichkeitsschätzung weicht systematisch von den objektiven ab.
- B) Die interpersonale Kommunizierbarkeit aber auch intellektuelle und emotionale Faktoren spielen bei der subjektiven Wahrscheinlichkeitsmessung eine gewichtige Rolle.
- C) Personen sind nicht in der Lage, zwischen annähernd gleichen Wahrscheinlichkeiten zu unterscheiden.
- D) Menschen sind nicht in der Lage, ihren „degree of belief“ ohne weiteres zu messen, zumindest nicht ohne Hilfe.

Morgan und Henrion (1990) entwickelten Befragungstechniken, mit denen Personen ihre Wahrscheinlichkeitsvorstellungen präzisieren und ausdrücken können. Sie unterscheiden zwischen direkter und indirekter Methode. Bei der direkten Meßmethode antwortet der Befragte mit einer Zahl für die unsichere Variable für eine vorgegebene Wahrscheinlichkeit. Für die Abfrage von direkten Wahrscheinlichkeitsangaben ist es hilfreich, die Ereignisse zunächst nach ihren Wahrscheinlichkeiten zu ordnen, bevor es zu einer Quantifizierung kommt. Für eine bessere Anschaulichkeit eignet sich dabei die grafische Darstellung der Wahrscheinlichkeiten auf einer Verteilungsfunktion.

Bei der indirekten Messmethode stellt der Befragte Vergleiche zu einfachen Analogsituationen an und schließt so auf die Wahrscheinlichkeit. Da der Wahrscheinlichkeitsbegriff für den naiven Befragten kein vertrautes Konzept ist, kann mit Hilfe der Analogsituation die Vorstellung darüber verbessert werden. In

indirekten Wahrscheinlichkeitsabfragen muss der Befragte keine Wahrscheinlichkeiten nennen oder sie verstehen.

2.3.6.2 Das Theorem von Bayes

Das Theorem von Bayes (1763) ist eine Formel zur Berechnung von Wahrscheinlichkeiten aus anderen Wahrscheinlichkeiten. Das Besondere liegt darin, dass es dazu benutzt werden kann, eine zunächst a priori angenommene Wahrscheinlichkeit in Bezug auf neue Daten zu revidieren (Eisenführ & Weber, 2003). Bayes analysierte Zusammenhänge von Beobachtungen mit der „wahren“ Umwelt. Anhand von Beobachtungen wird eine Meinung (belief) gebildet, und auf Grund von neuen Beobachtungen wird das Urteil korrigiert (Jungermann et al., 2010).

Die a-priori-Wahrscheinlichkeit ist diejenige Wahrscheinlichkeit, mit der eine Hypothese zutrifft, bevor irgendwelche Anhaltspunkte berücksichtigt werden. Die a-posteriori-Wahrscheinlichkeit ist jene Wahrscheinlichkeit, die nach Berücksichtigung eines eingetretenen Ereignisses zutrifft. Mit Hilfe des Bayes-Theorems kann die a-posteriori-Wahrscheinlichkeit einer Hypothese, wenn die a-priori-Wahrscheinlichkeit und die entsprechenden bedingten Wahrscheinlichkeiten gegeben sind, exakt berechnet werden (Anderson, 2001).

Die Wahl einer Alternative beeinflusst die Wahrscheinlichkeiten der möglichen Folgen der Entscheidung. Diese Folgen sollen so berücksichtigt werden, dass Fehlentscheidungen, die schwerwiegende und kostspielige Folgen haben können weniger wahrscheinlich gemacht werden, dagegen belanglosere Fehlentscheidungen eher in Kauf genommen werden, und richtige Entscheidungen möglichst hohe Wahrscheinlichkeiten erhalten (Wendt, 1983).

Darstellung der Formel für die Berechnung der a-posteriori-Wahrscheinlichkeit:

$$P(H|E) = \frac{P(E|H) \cdot P(H)}{P(E|H) \cdot P(H) + P(E|-H) \cdot P(-H)}$$

$P(H|E)$ a-posteriori Wahrscheinlichkeit, dass die Hypothese H auftritt, unter der Bedingung, dass Ereignis E bereits eingetreten ist

$P(H)$ a-priori Wahrscheinlichkeit, dass Hypothese H (ohne Kenntnis von E) eintritt

$P(E|H)$ Wahrscheinlichkeit des Eintretens von E unter der Voraussetzung, dass H zutrifft

$P(-H)$ Wahrscheinlichkeit, dass Hypothese nicht eintritt

$P(E|-H)$ Wahrscheinlichkeit, dass E eintritt, unter der Voraussetzung, dass H nicht zutrifft

Die Gewinnung der bedingten Wahrscheinlichkeiten (likelihoods) lassen sich nach Eisenführ und Weber (2003) auf Basis empirischer relativer Häufigkeiten bei sich wiederholenden Ereignissen gewinnen, z.B. durch das Sammeln von Daten über Krankheitssymptome oder auf Basis einer vorausgesetzten theoretischen (z.B. binomial) Verteilung. Eine dritte Möglichkeit wäre eine rein intuitive Gewinnung.

Abweichungen von Wahrscheinlichkeitsschätzungen

Es wurde zunächst angenommen, dass sich durch das Bayes Theorem die Wahrscheinlichkeitsurteile von Personen beim Auftreten neuer Daten angemessen beschreiben lassen – das Urteilsverhalten der Personen der wahrscheinlichkeitstheoretischen Norm entspreche. Empirisch mussten jedoch systematische Abweichungen festgestellt werden (Jungermann et al., 2010).

Basisraten-Fehler

Personen vernachlässigen oft die a-priori Wahrscheinlichkeit (Basisrate), durch eine besonders auffällige neue Information, die nicht im gerechtfertigten Ausmaß der theoretischen Wahrscheinlichkeit steht, wird die Ausgangswahrscheinlichkeit ignoriert. Kahneman und Tversky (1973) konnten den Basisraten-Fehler in zahlreichen Experimenten nachweisen.

Unzulänglichkeiten bei der Berücksichtigung der Basisrate könnten zu ungerechtfertigten Schlüssen führen.

Die Basisrate wird weniger oft vernachlässigt, wenn Ereignisse in absoluter und nicht in relativer Häufigkeit angegeben werden (Gigerenzer & Hoffrage, 1995).

Konservatismus

Ein gegenläufiges Phänomen wird durch den Konservatismus beschrieben. Hat sich jemand einmal ein erstes Urteil über die Wahrscheinlichkeit eines Ereignisses gebildet, so wird die Revidierung bei Vorlage valider neuer Information nur unzureichend vorgenommen (Jungermann et al., 2010). An der Anfangsmeinung wird entgegen der gerechtfertigten theoretischen Wahrscheinlichkeit festgehalten. Ein konservatives Verhalten beobachtet man vor allem, wenn man Probanden sukzessiv viele einzelne neue Informationen gibt und sich anschaut, wie sehr bzw. schnell sich ihre Meinung dann ändert (Lee, 1977).

„Bei der Einschätzung von Wahrscheinlichkeiten unterschätzen Menschen oft die kumulative Kraft von eingetretenen Ereignissen“ (Anderson, 2001).

Zwar sind Menschen nicht in der Lage die korrekte Wahrscheinlichkeit nach Bayes anzugeben, es gibt aber Hinweise, dass viele Aspekte ihres Verhaltens mit den Bayesianischen Prinzipien im Einklang stehen. Personen besitzen oft ein impliziertes Wissen über diese Prinzipien, auch wenn sie kein diesbezügliches explizites Wissen zur Anwendung bringen, wenn sie zu expliziten Einschätzungen aufgefordert werden. Gluck und Bower (1988) führten dahingehend Experimente durch und stellten fest, dass sich Personen den wahren Wahrscheinlichkeiten sehr

gut annäherten, ohne Informationen über Basisrate und bedingte Wahrscheinlichkeit. Das Verhalten, bei dem man zwischen Alternativen auswählt und sich an den objektiven Erfolgsanteil dieser Alternativen orientiert, wird als Wahrscheinlichkeitsabgleich bezeichnet.

Versuchspersonen sind bei proportionalen Einschätzungen ziemlich exakt, wenn sie dabei nicht auf ihr Gedächtnis zurückgreifen müssen (Robinson, 1964).

Hoffrage (1993, S. 91) weist darauf hin, "daß Häufigkeitsschätzungen genauer, realistischer und konsistenter sind als Wahrscheinlichkeitsschätzungen".

Kognitive Heuristiken

Eine Heuristik besteht aus vereinfachten Annahmen, mit deren Hilfe ein Problem schneller gelöst werden kann als ohne Vereinfachung. Sie werden oft bei Informationsmangel, in komplexen Situationen oder wenn der Zeitrahmen begrenzt ist, angewendet. Der Nachteil solcher Schnellverfahren liegt in ihren systematischen Verzerrungen.

Nach Teigen (1994) geben Personen unter bestimmten Bedingungen Urteile nach den Regeln der Wahrscheinlichkeitstheorie ab, im Normalfall erfolgen die Urteile aber „intuitiv“.

Kahneman und Tversky (1982) haben viele Forschungen bezüglich Heuristiken durchgeführt und in einem Sammelband „heuristics and biases“ publiziert.

Repräsentativitätsheuristik

Die subjektive Wahrscheinlichkeit für ein Ereignis ist umso größer, je repräsentativer das Ereignis für die Population ist, aus der es kommt (Jungermann et al., 2010). Die Heuristik ist meistens nützlich und auch erfolgreich, sie kann aber zu systematischen Verzerrungen führen.

Häufige Fehler sind:

- Insensitivität gegenüber der Basisrate
- kleine Stichproben werden unterschätzt (Gesetz der kleinen Zahl)
- Reihenfolgen abhängige Fehler
- extreme Ausgänge werden hinsichtlich ihrer Wahrscheinlichkeit überschätzt (Regression zur Mitte)
- Überschätzung der Wahrscheinlichkeit von Konjunktionen

Verfügbarkeitsheuristik

Je leichter Beispiele für ein Ereignis abgerufen oder erinnert werden, desto größer ist die subjektive Wahrscheinlichkeit eines Ereignisses. Das legendäre „Buchstabenhäufigkeits“-Experiment von Kahneman und Tversky (1973) sei hier angeführt, Anderson (2001) begründet dies gedächtnispsychologisch. Wörter werden stärker mit ihren Anfangsbuchstaben als mit ihrem zweit platzierten oder später platzierten Buchstaben assoziiert.

Zu Verzerrungen kommt es durch:

- die Lebhaftigkeit der Darstellung
- Präsenz des Ereignisses
- Verknüpfung der Ereignisse

Anker- und Anpassungsheuristik

Die Heuristik setzt dann ein, wenn ein Anker (Ausgangswahrscheinlichkeit) gesetzt wird. Nach diesem Anker erfolgt der Schätzwert nach oben oder nach unten. Für die Urteilsbildung wird oft herangezogen wenn wenig Information zur Verfügung steht, folgende Fehler können auftreten:

- numerische Größen werden falsch eingeschätzt
- Erinnerungsverzerrungen
- fehlerhafte Vorstellungen

Diese drei Heuristiken sind jene, die am häufigsten in der Fachliteratur zu finden sind und durch eine Vielzahl an Experimenten am besten erforscht wurden. Bei Autoren, die sich diesem Thema widmeten, können ein Fülle von weiteren

Heuristiken, Verzerrungen und Fehler vertieft nachgelesen werden (vgl. Jungermann et al., 2010; Anderson, 2001 oder Eisenführ & Weber, 2003).

Overconfidence

Ein weiterer Bias muss an dieser Stelle angeführt werden, da er bereits bei der ersten EUcO-Version aufgetreten ist. Es handelt sich dabei um das Overconfidence-Phänomen. Damit ist gemeint, dass sich Menschen bezüglich der Richtigkeit ihrer Antworten allzu sicher sind. Die Sicherheit der Richtigkeit einer gegebenen Antwort ist also höher als die relative Häufigkeit der richtigen Antwort (Jungermann et al., 2010). Der Bias ist auf die Tendenz rückführbar, dass Personen nach bestätigender Information suchen. Zuerst wählt man eine Alternative aus und im Anschluss wird nach weiterer Information gesucht, die die gewählte Alternative bestätigt. Es wird aber nicht nach Informationen gesucht, die gegen die gewählte Alternative sprechen. Diese selektive Informationssuche führt zu einer überhöhten Sicherheit in die eigene Antwort.

2.3.7 Entscheidungshilfen

Eine Recherche in der Suchmaschine Google brachte am 7. August 2010 nach dem Schlagwort „Entscheidungshilfe“ 1.89 Millionen Einträge; es ist daher anzunehmen, dass die Unterstützung bei Entscheidungen von großer Bedeutung sein muss, bzw. dass es zumindest ein riesiges Angebot an Vereinen, Unternehmen oder Personen gibt, die Hilfestellungen bei Entscheidungen anbieten. Auffallend bei der Recherche war die Unterteilung in Hilfe und Unterstützung, Hilfe wird meistens Personen oder Unternehmen angeboten und Unterstützung durch technische Systeme (z.B. Software) gegeben.

Nach Schäfer (1976) lassen sich Entscheidungshilfen nach verschiedenen Gesichtspunkten gestalten. Eine wesentliche Punkt ist dabei, den Entscheidungsprozess in Teilabschnitte zu gliedern und in (Fluss)-Diagrammen grafisch dazustellen. Entscheidungshilfen sind davon abhängig, ob sich die

Informationen auf Unsicherheit oder Nutzen beziehen. Ebenso können die weiter oben erwähnten Heuristiken als Hilfen in Entscheidungssituationen angewendet werden, wenn keine anderen Modelle verfügbar sind. Schließlich ist die technische Hilfe durch den Computer wichtig. Er ermöglicht große Datenmengen zu speichern und zu verarbeiten.

Jungermann et al. (2010) referieren eine Reihe von Unterstützungen, wobei hier nicht nur technische Systeme gemeint sind:

Formen und Bereiche von Beratung

Der Nutzen von entscheidungstheoretischen Konzepten reicht von stark formalisierten quantitativen Verfahren der Entscheidungsanalyse (decision analysis) (Eisenführ & Weber, 2003) bis hin zu verbalen qualitativen Verfahren der Entscheidungsberatung (decision counseling). Diese Verfahren kommen im ökonomischen, medizinischen und individuellen Bereich (z.B. Beruf, Familie, usw.) zum Einsatz.

Beratung in dyadischen Situationen

Entscheidungsprobleme, bei denen ein Experte um Rat gebeten wird, werden als „Judge-Advisor-Systeme“ bezeichnet. Der „judge“ trifft eine Entscheidung und holt sich vom „advisor“ Informationen, Meinungen oder Empfehlungen ein. In Experimenten wurde untersucht, welche Informationen gesucht werden, abhängig davon, ob die Entscheidungen für sich selbst oder für andere Personen getroffen werden. Jungermann und Fischer (2005) untersuchte zum Beispiel welchen Einfluss der Expertenstatus auf die Informationsverarbeitung hat.

Entscheidungsunterstützungssysteme

Das sind Techniken und Systeme, die bei der Wahrnehmung und Strukturierung von Problemen, bei der Informationsverarbeitung, bei der Bewertung und Auswahl sowie bei der Steuerung von Prozessen Unterstützung leisten. Diese Systeme basieren auf entscheidungstheoretischen Modellen. Entscheidungshilfen werden

oft abgelehnt, weil sich nicht die Wünsche oder Bedürfnisse der potentiellen Nutzer befriedigen (Yates, Veinott und Patalano, 2003).

Computer-Unterstützung

Damit sind spezielle Softwareprogramme gemeint, die sowohl im individuellen wie auch unternehmerischen Bereich angewendet werden.

Szenario-Technik

In Szenarien werden mögliche „Zukünfte“ in Abhängigkeit von Aktionen und Ereignissen dargestellt. In vereinfachter Form handelt es sich dabei um Pfade in einem Entscheidungsbaum. In den politischen und wirtschaftlichen Bereichen werden sie zur Suche und Bewertung von Handlungsoptionen entwickelt.

Anleitung zur Verbesserung des Entscheidungsverhaltens

Darunter versteht man eine Reihe von mehr oder weniger streng an entscheidungstheoretischen Konzepten orientierten Anleitungen, wie Präferenzen geordnet und wie Urteilsfehler vermieden werden und wie schließlich bessere Entscheidungen getroffen werden. Ein aktuelles Beispiel, ein Trainingsprogramm für die Beratung bei der Berufswahl, wurde von Teuschner (2003) entwickelt und mit Erfolg, eingesetzt.

2.3.8 Attribution

Der Attributionsbegriff stammt aus dem lateinischen und bezeichnet in der Psychologie die Ursachenzuschreibung von Handlungen und die daraus resultierenden Konsequenzen für das Erleben und Verhalten von Menschen. „Gegenstand der Attributionsforschung sind Meinungen über Kausalzusammenhänge, sowie das Zustandekommen und die Auswirkungen solcher Meinungen“ (Herkner, 1980, S. 11).

Die Attributionspsychologie liefert nicht nur Erklärungen von Kausalwahrnehmungen: sie kann auch Vorhersagen über zukünftige Ereignisse machen.

Kelley (1967) formulierte die umfassendste aller Attributionstheorien. Sie lässt sich auf die Personen- und Selbstwahrnehmung wie auch auf Verhaltenskonsequenzen und andere Ereignisse anwenden. Zentrum seiner Theorie ist das Kovariationsprinzip, welches besagt, dass eine Person ein Ereignis auf Ursachen zurückführt, mit der es kovariiert. Damit ein Beobachter Kovariationen feststellen kann, benötigt er Informationen über das Ereignis und den Bedingungen für das Auftreten des Ereignisses. Informationen über die Person, die Objekte (Stimuli, Reiz, usw.) und die Zeit sind dabei besonders wichtig. In der Regel ist der Ausgangspunkt die Beobachtung des Verhaltens einer Person gegenüber einem bestimmten Objekt. Nach Kelley (1967) erhält man bei der Verhaltensbeobachtung Informationen über *Konsensus*, *Distinktheit* und *Konsistenz*. Der Konsensus ist hoch, wenn sich viele Personen gegenüber dem Objekt verhalten wie die Referenzperson. Eine hohe Distinktheit ergibt sich, wenn sich die Vergleichsperson nur gegenüber wenigen Objekten so verhält, wie gegenüber dem Ausgangsobjekt. Und schließlich ist die Konsistenz hoch, wenn sich verschiedene Personen zu vielen Zeitpunkten in gleicher Weise verhalten wie gegenüber dem Ausgangsreiz.

Kelley formulierte dazu drei Hypothesen:

1. Ein Ereignis wird dem Objekt (Reiz) zugeschrieben, wenn hoher Konsensus, hohe Distinktheit und hohe Konsistenz besteht. Man spricht von Stimulusattribution.
2. Ein Ereignis wird der Person zugeschrieben, wenn der Konsensus gering, die Distinktheit gering und die Konsistenz hoch ist (Personenattribution).
3. Ein Ereignis wird den Umständen zugeschrieben, wenn geringer Konsensus, hohe Distinktheit und geringe Konsistenz besteht (Situationsattribution).

Viele Experimente bestätigten Kelleys Hypothesen. Es stellte sich allerdings heraus, dass mehr Personenattributionen als Stimulusattributionen vorgenommen wurden. Weiters hatten Konsensusinformationen insgesamt schwächere Effekte als Distinktheits- oder Konsistenzinformationen (Herkner, 1980).

Eine ausführliche Erklärung zu Kelleys Attributionstheorie kann bei Herkner (1980) nachgelesen werden.

2.3.9 Rationalität

2.3.9.1 Grundbegriff

Der Begriff Rationalität stammt aus dem lateinischen Wort ratio, was Vernunft bedeutet. Unter „rationalitas“ wird das Denkvermögen verstanden. Rationalität bzw. rational wird als Erklärungsmuster für Handlungen verwendet. Eine in diesem Sinne „rationale Erklärung“ einer bestimmten Handlung versucht diese – durch Angabe der jeweils (handlungs-)situationsspezifischen Bedingungen – intersubjektiv verständlich, also für andere Personen nachvollziehbar zu machen.

Der Rationalitätsbegriff findet wissenschaftlich interdisziplinär Verwendung, etwa in der Philosophie, Mathematik oder Ökonomie. Nach Kirchler (2005) handelt der „homo oeconomicus“, ein Menschbild in der Wirtschaftspsychologie das einen fiktiven Durchschnittsmenschen beschreibt, völlig zweckrational. „Mögliche Entscheidungsalternativen werden vor der Entscheidung für eine Alternative gesichtet und entsprechend den individuellen Vorteilen oder Präferenzen gereiht“ (Kirchler, 2005, S. 28).

Die präskriptive Entscheidungstheorie will Entscheidern helfen, möglichst rationale Entscheidungen zu treffen. In der Entscheidungstheorie lässt sich aber schwer feststellen, wie gut oder schlecht eine Entscheidung ist. Darum kann man auch

nicht von einer rationalen oder irrationalen Entscheidung sprechen, sondern nur von „mehr oder weniger rational“ (Eisenführ & Weber, 2003). Auch der spätere Erfolg oder Misserfolg ist keine sichere Messlatte für Rationalität, da man zwischen einer rationalen und einer erfolgreichen Entscheidung unterscheiden muss. Entscheidungen, die sich wiederholen oder keine Unsicherheit aufweisen, können leichter aufgrund von Ergebnissen beurteilt werden als einmalige, schwerwiegende Entscheidungen unter großer Unsicherheit.

Das Rationalitätskonzept einer Entscheidung enthält keine Vorschriften über die Inhalte. Erwartungen und Präferenzen sind grundsätzlich subjektiv. Sie müssen aber mit den Rationalitätspostulaten konsistent und begründbar sein. Die Rationalität von Entscheidungen in Bedeutung von bewussten, begründbaren und verstehbaren Entscheidungen sind vom Wertesystem und den abrufbaren Informationen des Entscheiders abhängig (Lee, 1977).

Eisenführ und Weber (2003) nennen zwei wichtige Kriterien für die Umsetzung der Rationalität in Entscheidungsprozesse. Das erste ist die prozedurale Rationalität, d.h. die Beachtung bestimmter als rational akzeptierter Anforderungen an das Denken.

Die prozedurale Rationalität wird bestimmt:

- Entscheider muss sich überlegen, ob er das richtige Problem löst (jede Entscheidung ist nur ein Ausschnitt aus dem Problemuniversum).
- Die Rationalität verlangt eine angemessen sorgfältige und systematische, aber nicht maximale Entscheidungsvorbereitung. Reduzierung des Aufwandes.
- Bei der Erwartungsbildung sollen relevante objektive Daten in Betracht gezogen werden.
- Eigene Ziele und Präferenzen sollen gründlich überlegt werden, damit Selbsttäuschung vermieden wird.

Das zweite Kriterium ist die Konsistenz, dass die bei einer Entscheidung eingehenden Prämissen untereinander kongruent sind.

Anforderungen an das Konsistenzkriterium sind:

- Zukunftsorientierung. Die Grundlage für Handlungsalternativen basiert auf den jeweiligen Konsequenzen.
- Transitivität
- Invarianz. Die möglichen Präferenzen sollen nicht von der Darstellung des Entscheidungsproblems abhängen.
- Unabhängigkeit gegenüber irrelevanten Alternativen.

Nach Bronner (1989) erhöht sich das Rationalitätsniveau von Entscheidungen, wenn Rationalitäts-Barrieren erkannt und bewusst gemacht werden.

2.3.9.2 Rationalität des Menschen

Es gibt drei Quellen (Zimbardo, 1995), warum rationales Denken verloren geht. Eine Erklärung ist der Einfluss der Masse. Durch sie wird der Verstand untergraben und man unterwirft sich der Gruppe. Eine zweite Quelle ist die Erfüllung der eigenen Bedürfnisse: sie behindern das rationale Denken. Die dritte Quelle sind Vorurteile. Diese müssen sich nicht nur auf Personen beziehen, sondern können auch gegenüber Handlungsalternativen bestehen.

Quellen nicht rational handelnder Menschen sind intrapsychische Konflikte, die sich in irrationalen Überzeugungen, Ängsten und Handlungen manifestieren. Stabile kognitive Täuschungen runden das Bild des nicht rational handelnden Menschen ab (Gigerenzer, 1993).

Rational fundiertes Entscheiden gelingt, wenn alle relevanten Informationen gesammelt, abgewogen und verrechnet werden (Gigerenzer, 2007). Dass dem nicht so ist, Menschen also oft intuitiv handeln, liegt an der begrenzten Fähigkeit, verfügbare Informationen zu transzendieren (überschreiten der Wahrnehmung).

Linda-Problem

Personen wird folgende Aufgabe gestellt:

Linda ist sehr intelligent und äußert offen ihre Meinung. Sie ist studierte Philosophin. Als Studentin hat sie sich für Gleichberechtigung und soziale Gerechtigkeit engagiert, weiters hat sie gegen Atomkraftwerke demonstriert. Personen müssen nach der Textpräsentation eine Alternative auswählen, die für sie am wahrscheinlichsten ist.

A1: Linda ist Bankangestellte.

A2: Linda ist Bankangestellte und in der Frauenbewegung aktiv.

Die meisten Personen wählten Alternative 2 und unterlagen dadurch einen Fehlschluss, weil sie sich von der Beschreibung beeinflussen ließen. Würde man nach mathematischen Wahrscheinlichkeiten antworten, wären beide Alternativen gleich wahrscheinlich.

Das „Linda-Problem“ dient als Beweis, dass Menschen oft nicht rational handeln, sondern sich von intuitiven Gefühlen leiten lassen. Sie interpretieren verfügbare Informationen falsch und verstoßen gegen die Gesetze der Logik. So kommt es vor, dass manchmal die Teilmenge größer als die Gesamtmenge ist und Beziehungselemente falsch ins Verhältnis gesetzt werden.

Gigerenzer (2007) spricht vom Konjunktionsfehler, wenn die Semantik von Wörtern trügerische Verwendung findet. Dem Wort „und“ zum Beispiel wird dabei die Bedeutungen einer logischen Verknüpfung zugewiesen, die Einfluss auf die Alternativenwahl nimmt.

2.3.10 Zufall, Glück und Geschick

Menschen geben selbst bei numerisch gleich großen Wahrscheinlichkeiten unterschiedliche Bedeutungen ab, je nachdem, ob das Ereignis auf das eigene Geschick oder den Zufall zurückzuführen ist. Während beim Roulette der Zufall die richtige Zahl bestimmt, ist beim Dart-Spiel das Geschick für die höchstmögliche Punktsomme verantwortlich.

Menschen wollen die Geschehnisse ihrer Umwelt verstehen und lenkend eingreifen. Der Beobachtung kann mittels Zufall eine Bedeutungseinheit gegeben werden. Der Zufall stellt eine Erklärung für alles dar, wofür wir keine Ursachen entdecken können. Auch Glück wird zur Erklärung verwendet, wenn Erfolg nicht auf Fähigkeiten wie Fleiß oder Intelligenz beruht (Lee, 1977).

„Was man unter „Zufall“ versteht und welche Beweggründe man für oder gegen das Ernstnehmen des Zufalls hat, ist jeweils an die Begriffssprache gebunden, in der die Auseinandersetzungen geführt werden“ (Gigerenzer, Szijntink, Porter, Daston, Beatty & Krüger, 1999).

Das Reich des Zufalls hat mit den Versuchen und Experimenten im Glückspiel begonnen. Seit der statistischen Wahrscheinlichkeitstheorie wird dem Begriff des Zufalls in den verschiedensten Wissenschaften (Biologie, Rechtswesen, Ökonomie, usw.) begegnet.

2.3.11 Spieltheorie

In den Anfängen der mathematischen Spieltheorie wurde den Gesellschaftsspielen wie zum Beispiel, Schach oder Mühle große Aufmerksamkeit gewidmet, daher der Begriff Spieltheorie.

Die Spieltheorie ist ein Teilgebiet der Mathematik, das sich damit befasst, Systeme mit mehreren Darstellern (Spielern) zu modellieren. Sie versucht das rationale Entscheidungsverhalten in sozialen Konfliktsituationen abzuleiten. Entscheidungssituationen, in denen der Erfolg eines Spielers nicht nur vom eigenen Handeln, sondern auch von den Handlungen anderer abhängt, werden durch die Spielsituation beschrieben.

Die Spieltheorie handelt also von der Strategie der beteiligten Spieler und dem Gewinn als Konsequenz der gewählten Strategie (Bühlmann, 1969). Die Methode der Spieltheorie wird zur Lösung von Entscheidungsproblemen eingesetzt. Daher

ist sie auch ein grundlegender Bestandteil der entscheidungstheoretischen Literatur.

Gegenstand der Spieltheorie ist nach Holler und Illing (1996) die Analyse von strategischen Entscheidungssituationen, in denen

1. das Ergebnis von den Entscheidungen anderer abhängt, so dass ein einzelner das Ergebnis nicht unabhängig von der Wahl der anderen bestimmen kann,
2. jeder Entscheidungsträger sich dieser Interdependenz bewusst ist,
3. jeder Entscheidungsträger davon ausgeht, dass alle anderen sich ebenfalls der Interdependenz bewusst sind,
4. jeder bei seinen Entscheidungen 1. – 3. berücksichtigt.

Der Informationsstand der Spieler ist für die Darstellung und Lösung entscheidend. Dabei wird zwischen vollständiger und vollkommener Information, abhängig davon, ob der Spieler über die Spielregeln, die eigenen Züge und die Züge der anderen Spieler informiert ist, unterschieden. (Bamberg & Coenenberg, 2006). Wenn der Spieler einen Zug ausgeführt hat und er über den bisherigen Verlauf des Spiels völlig informiert ist, so spricht man von einem Spiel mit vollkommener Information. Bei einem Spiel mit vollständiger Information sind alle Spieler über die Charakteristika (Strategienmenge und Auszahlungsfunktionen) der restlichen Spieler vollständig informiert.

In den Spielsituationen beeinflussen sich die Aktionen von verschiedenen Entscheidungssubjekten die zu Konfliktsituationen führen in ihren Ergebnissen wechselseitig. Die Lösung dieser Konfliktsituation gelingt dann, wenn für alle beteiligten Spieler ein optimaler Kompromiss gefunden wird. Jeder Spieler hat einerseits das egoistische Motiv, die beste verfügbare Alternative zu wählen, andererseits soll der Nutzen aller Beteiligten maximal sein (Herkner, 2001). Das Gefangenendilemma ist ein analoges Beispiel, wie eine Nutzenmaximierung erfolgen kann:

Zwei Personen werden des gleichen Deliktes verdächtigt und werden getrennt voneinander, ohne die Möglichkeit zur Absprache in der Untersuchungshaft, verhört. Gesteht einer der Beiden, so kommt der jeweils andere für lange Zeit ins Gefängnis und der Geständige frei. Bekennen sich beide schuldig, müssen sie eine mittlere Gefängnisstrafe absitzen und wenn beide leugnen, wird nur eine leichte Strafe verhängt. Die beste Alternative, um den größtmöglichen Nutzen (in diesem Beispiel die geringst mögliche Haftstrafe) zu erhalten, wäre die Option zu gestehen.

2.3.12 Zeitdruck

Die Dringlichkeit von Entscheidungen wird durch die verfügbare Zeit bestimmt (Bronner, 1989). Ist der Zeitrahmen für eine Entscheidung sehr knapp bemessen, führt dies beim Entscheider zu Stresssymptomen. Entscheidungen, die unter Zeitdruck gefällt werden, gehen mit einer Leistungseinschränkung einher, die zu erheblichen Konsequenzen in der Entscheidungsqualität führen kann.

Die verfügbare Zeit wird von Personen interindividuell unterschiedlich wahrgenommen, der empfundene Zeitdruck hängt folglich nicht nur von der tatsächlichen Zeitbegrenzung, sondern ebenso vom Individuum und wie dieses den Zeitrahmen bewertet, ab.

Bei routinierten Entscheidungen spielt die Zeitkomponente nur mehr eine sekundäre Rolle, hingegen führt Zeitdruck bei neuartigen Entscheidungen nach Bronner (1989) dazu, dass sich die Informationssuche reduziert, die Entscheidungsaktivitäten eingegrenzt und Effizienz mindernde Effekte ausgelöst werden.

Unter Zeitdruck wird die Informationsverarbeitung durch die Anwendung non-kompensatorischer Entscheidungsregeln beschleunigt und vereinfacht. Optionen werden aufgrund ihrer wichtigsten Attribute ausgewählt. Dies erfordert keine Verrechnung der Vor- und Nachteile.

Entscheidungsträger konzentrieren sich unter Zeitdruck eher auf relevante Informationen, ignorieren unwichtige und bevorzugen weniger riskante Optionen (Payne, Bettman & Johnson, 1993).

2.4 Beschreibung EUcO 2.0

Der EUcO 2.0 basiert auf der Version von Radinger (2002). Beim EUcO wurden Testpersonen mit mathematischen Wahrscheinlichkeitsaufgaben konfrontiert, anhand deren verschiedene Aspekte des Entscheidungsverhaltens erfasst werden.

Testkennwerte des EUcO:

- Summe der richtigen Entscheidungen (Rohscore)
- Entscheidungsgenauigkeit (geschätzte Wahrscheinlichkeit)
- Entscheidungskonsistenz
- Entscheidungsgeschwindigkeit
- Verwendung von Hilfen
- Zeitdruck

Aufgrund der computerisierten Vorgabe konnten weitere Testkennwerte wie Reaktionszeit oder die sukzessive Verkürzung der Bearbeitungszeit (kontinuierliche Erhöhung des Zeitdrucks) realisiert werden. Die Itembearbeitung erfolgt ausschließlich mit der Maus.

Der Test sollte Entscheidungsverhalten möglichst unabhängig von situativen Rahmenbedingungen, individuellen Vorerfahrungen und emotionaler Involviertheit erfassen.

Die Aufgaben wurden so konstruiert, dass gewisse objektive Informationen vorgegeben wurden, allerdings der Zufall die Behälterwahl und somit die Entscheidung beeinflusste. Bei Entscheidungen unter Unsicherheit spielt die

objektive Wahrscheinlichkeit der möglichen Alternativen eine gewisse Rolle: anhand des EUcO's können Personen, welche diese Wahrscheinlichkeiten hinreichend genau berücksichtigen, diagnostiziert werden.

Die Eigenschaft eines Objektiven Persönlichkeitstests erhält der EUcO dadurch, dass die Testperson aufgefordert wird, Wahrscheinlichkeiten mittels verschiedener optionaler Hilfsmöglichkeiten zu schätzen, ohne die wahre Testintention durchschauen zu können.

2.4.1 Kritik am EUcO

Als Hauptkritikpunkt kristallisierte sich bei der Verwendung des Tests seine komplexe Verständlichkeit heraus, einerseits durch eine lange und schwer verständliche Einführung, andererseits aufgrund der Wahrscheinlichkeit. Bevor zukünftige Testpersonen den Test überhaupt bearbeiten konnten, mussten sie bereits hohe kognitive Leistungen beim Instruktionsverständnis erbringen, um die an sie gestellten Anforderungen zu verarbeiten.

Von den insgesamt 17 Hilfen im EUcO wurde die Hilfe „Letzte Entscheidung richtig“ am häufigsten gewählt. Diese Hilfe gibt Aufschluss darüber, ob die Testperson das zuvor bearbeitete Item richtig gelöst hat. Daraus lässt sich der Wunsch nach Feedback für die richtige Behälterwahl schlussfolgern. Ebenso kann aber auch die Testperson eine Antwort danach suchen, ob die jeweils verwendete Strategie instruktionskonform ist und somit im Sinne eines fehlerfreien Verständnisses erfasst wurde.

Für die Revision des EUcO gab der Autor den Test mehreren Personen vor, wobei sich herausstellte, dass es bei der Abgabe der Wahrscheinlichkeitsschätzung einigen Testpersonen Probleme bereitete, die Frage „Wie sicher sind sie sich dabei“ wahrscheinlichkeitsmäßig zu antworten. Bei manchen Personen bezog sich die Wahrscheinlichkeitsschätzung auf die subjektive Sicherheit ihrer gewählten Behälter. Einige Testpersonen wollten die Wahrscheinlichkeitsschätzung sehr genau am Prozentstreifen markieren und benötigten dadurch etwas mehr Zeit.

Einige Testpersonen taten sich bezüglich der Zuordnung der bereits im Auswurf befindlichen Kugel schwer. Es erforderte immer wieder Überlegungen, ob diese Kugel zum favorisierenden Behälter hinzugezählt oder abgezogen werden musste.

Bei vielen psychologischen Verfahren sind Beispielaufgaben für ein leichteres Verständnis Usus, welche beim EUcO jedoch nicht realisiert worden sind. Die Vorgabe von Beispielaufgaben trägt nicht nur zum Verständnis bei, sondern erleichtert den Umgang mit den Wahrscheinlichkeitsaufgaben, weil interindividuelle Niveauunterschiede in der Kenntnis des Wahrscheinlichkeitsprinzips anzunehmen sind. Durch Beispielaufgaben kann von einem zumindest gleichen „Basiswissen“ ausgegangen werden. Die Berechnung der genauen Wahrscheinlichkeit nach Bayes kann ohnehin nicht ohne technische Hilfsmittel erfolgen.

Die hohe Anzahl an Hilfen sorgt für eine zusätzliche Verwirrung, wobei gewisse Hilfen relativ wenig genutzt wurden. Die Frage, ob Entscheidungen getroffen oder aufgeschoben werden, mag entscheidungspsychologisch eine wichtige sein, jedoch aufgrund der geringen Inanspruchnahme tragen sie eher zur Irritation als zur Unterstützung bei.

Nach Abschluss jeder Aufgaben haben die Testpersonen die Möglichkeit, ihre Entscheidung rückgängig zu machen. Hier stellt sich die Frage, ob es sinnvoll ist, die getätigte kognitive Anstrengung zu hinterfragen und die Entscheidung zu revidieren. Für die Wahl eines Behälters müssen alle Überlegungen bis spätestens vor der Entscheidung abgeschlossen sein. In der Regel können Personen in der Berufswelt Entscheidungen nicht ständig revidieren.

Beim EUcO konnten Testpersonen bei der Hilfe „Expertenrat“ zwischen zwei Optionen (Mathematikprofessor oder Wirtschaftsforscher) wählen; die Sinnhaftigkeit einer Aufsplittung dieser Hilfe ist zu bezweifeln und führt an der eigentlichen Testintention vorbei.

Als letzter Kritikpunkt ist anzuführen, dass die Möglichkeiten der Testgestaltung z.B. Optik oder Animation durch die computerisierte Vorgabe unausgeschöpft blieben. Die statische Eigenschaft tut ihr übriges zur komplexen Wahrnehmung.

2.4.2 Instruktion

Die gesamte Instruktion des EUcO 2.0 wurde auf 12 Bildschirmseiten dargestellt. Auf Grund der Komplexität des Aufgabenmaterials wurden die Seiten in zwei Teile (Abbildung 1 und 2) gegliedert, die in unterschiedlichen Schwarz- und Graunuancierungen angeordnet sind. Die Testperson sollte dadurch mit der Informationsdarbietung nicht überfordert werden. Sie kann durch Anklicken des „Weiter“ Buttons selbst entscheiden, wann die nächste Seite bzw. der nächste Teilbereich der Instruktion erscheinen soll. Mit Hilfe des „Zurück“ Buttons wird der vorher dargestellte Bereich angezeigt.

Die Instruktion kann von der Testperson während der Einführung aber auch bei der eigentlichen Testung beliebig oft abgerufen werden.

Für ein besseres Verständnis werden in der Instruktionsphase fünf Beispielaufgaben gegeben.

Abbildung 1: oberer Informationsteilbereich



Abbildung 2: unterer Informationsteilbereich



2.4.3 Beispielaufgaben

Die Darstellung der Beispielaufgaben erfolgte nach dem „Lotto-Prinzip“ (Abbildung 3). Der Testperson werden zuerst die beiden Behälter mit den verschiedenfarbigen (schwarzen und weißen) Kugeln präsentiert. Im Anschluss werden die Kugeln herumgewirbelt (animiert) bis eine Kugel in den Auswurf fällt. Damit sollte eine analogische Verbindung zur Zahlenlottoziehung hergestellt werden. Die explizite Illustration des so genannten Schließmechanismus (Abbildung 4) sollte der Testperson verdeutlichen, dass nur eine einzige Kugel in den Auswurf fallen kann und alle weiteren Kugeln sofort zurückgehalten werden.

Die fünf Beispielaufgaben sind mit ansteigender Schwierigkeit (Abbildung 5 und 6, Abbildung 5a und 6a mit Wahrscheinlichkeiten) angeordnet. Die Veränderung der Kugelanzahl als auch deren Farbzusammensetzung in den Behältern führt zu einer Veränderung der a posteriori Wahrscheinlichkeit: dadurch soll die Testperson den Umgang mit Wahrscheinlichkeiten erlernen.

Nachdem eine Kugel in den Auswurf gefallen ist, wird die Testperson aufgefordert, jenen Behälter auszuwählen, aus dem die Kugel höchstwahrscheinlich stammt. Wählt die Testperson einen Behälter aus, erhält sie eine Rückmeldung über die Richtigkeit ihrer Wahl (Abbildung 7). Wird der falsche Behälter gewählt (Abbildung 8), muss die Person das Übungsbeispiel, mit dem Hinweis es sich genauer anzusehen, nochmals durchführen.

Wird der richtige Behälter gewählt, erfolgt die Information für die zweite Teilaufgabe und die richtige Markierung am Zahlenband. Damit sollen Assoziationen zur späteren Wahrscheinlichkeitsschätzung in der Testphase hergestellt werden (Abbildung 9).

Nach der ersten dargebotenen Ziehung wird betont, dass das Bild eine nicht logische Situation abbildet, da eine Kugel nach der Ziehung in einem Behälter

fehlen müsste (Abbildung 10). Die Testperson muss sich dieser Fiktion immer bewusst sein.

Die Testperson wird am Ende der Instruktion darauf hingewiesen, dass nach der Wahl eines Behälters die geschätzte Wahrscheinlichkeit nicht unter 50 % liegen kann, da dies ein Widerspruch wäre.

Abbildung 3: Lotto-Prinzip

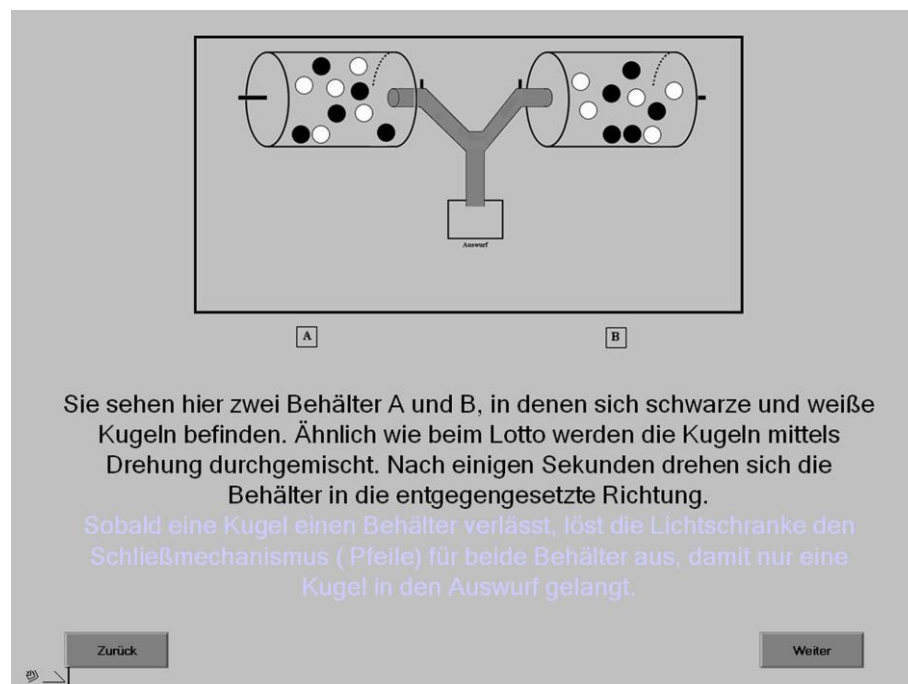


Abbildung 4: Schließmechanismus

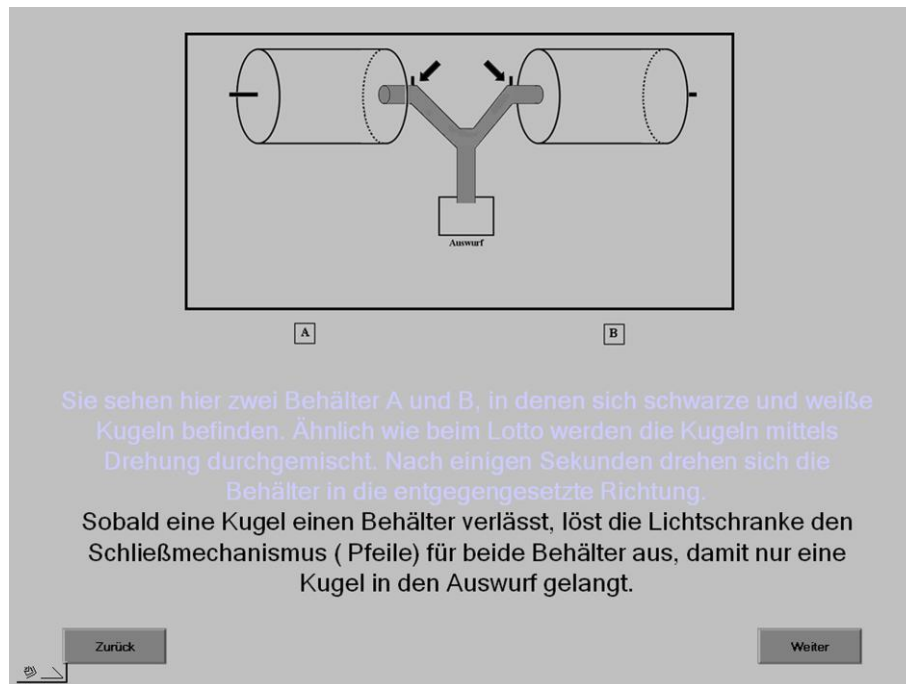


Abbildung 5: erste Beispielaufgabe

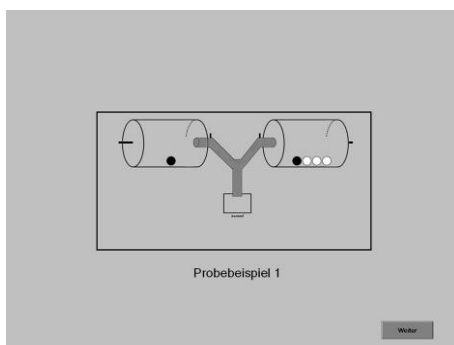


Abbildung 6: zweite Beispielaufgabe

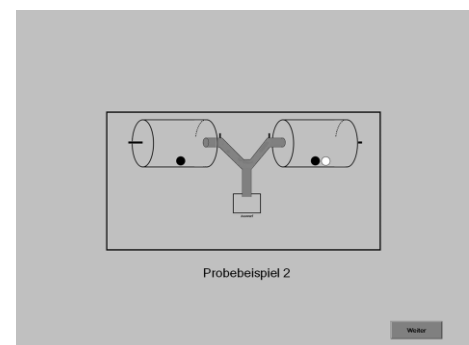


Abbildung 5a: Wahrscheinlichkeit 80%

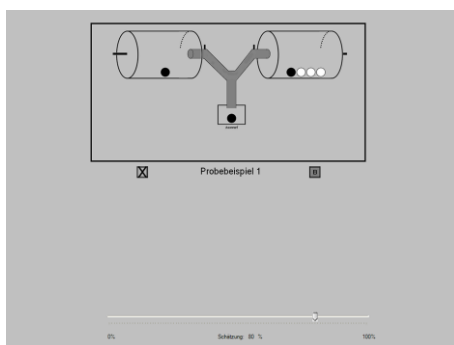


Abbildung 6a: Wahrscheinlichkeit 67%

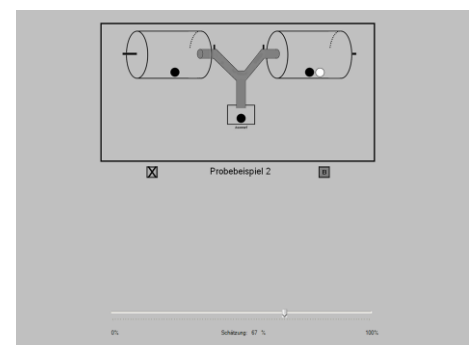
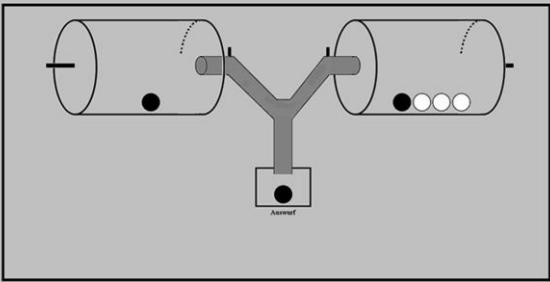


Abbildung 7: positive Rückmeldung bei richtiger Behälterwahl



☒ Probebeispiel 1 ☐ B

Der Behälter A ist richtig!

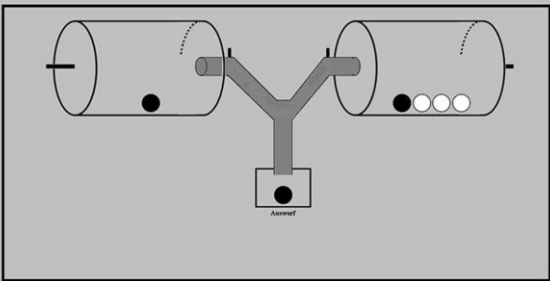
Nun sollten Sie den zweiten Teil der Aufgabe bearbeiten, die Markierung am Zahlenband. Bei den Probebeispielen erfolgt dies automatisch.

Bei einem Versuch mit 100 Durchgängen, wobei nur Durchgänge mit einer schwarzen Kugel im Auswurf zählten, kam die schwarze Kugel 80 Mal aus Behälter A und 20 Mal aus Behälter B.

Die richtige Markierung am Zahlenband liegt folglich bei 80

0% Schätzung: 80 % 100%

Abbildung 8: negative Rückmeldung bei falscher Behälterwahl



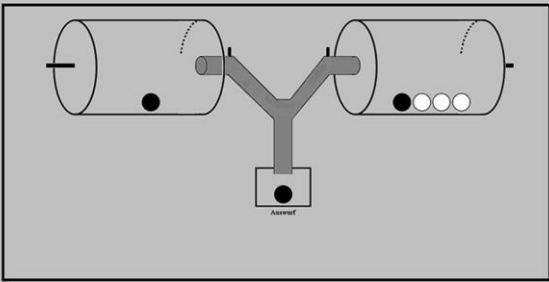
☐ A ☒ B

Probebeispiel 1

Der Behälter B ist leider falsch!

Klicken Sie auf den Button „Zurück“ und sehen Sie sich das Probebeispiel noch einmal genau an!

Abbildung 9: Herstellung der Assoziation zur Bayes Wahrscheinlichkeit



Probebeispiel 1

Der Behälter A ist richtig!

Nun sollten Sie den zweiten Teil der Aufgabe bearbeiten, die Markierung am Zahlenband. Bei den Probebeispielen erfolgt dies automatisch.

Bei einem Versuch mit 100 Durchgängen, wobei nur Durchgänge mit einer schwarzen Kugel im Auswurf zählten, kam die schwarze Kugel 80 Mal aus Behälter A und 20 Mal aus Behälter B.

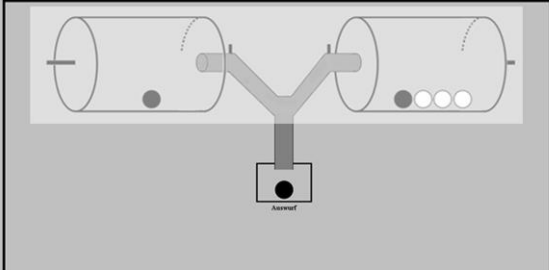
Die richtige Markierung am Zahlenband liegt folglich bei 80

0% Schätzung: 80 % 100%

Zurück Weiter

Abbildung 10: Hinweis auf die nicht logische Bildsituation

Beachten Sie !



Probebeispiel 1

In der hell unterlegten Fläche sehen Sie die Ausgangssituation, noch **bevor** eine Kugel gezogen wurde. Im Auswurf sehen Sie eine schwarze Kugel, also **nach** der Ziehung. Hierbei handelt es sich um eine **unlogische Situation!** Die Kugel müsste eigentlich in einem der Behälter fehlen, aber aufgrund der Aufgabenstellung wird Ihnen die korrekte Abbildung vorenthalten.

Weiter

2.4.4 Aufgabenmaterial EUcO 2.0

Der EUcO 2.0 besteht aus 30 Items mit jeweils zwei Teilaufgaben. Bei der ersten Teilaufgabe werden der Testperson zwei Behälter angeboten, in denen sich eine unterschiedliche Anzahl schwarzer und weißer Kugeln befinden. Eine weitere Kugel (schwarz oder weiß) befindet sich in einem Auswurfbehälter (Itembeispiel¹ Abbildung 11). Der Testperson wird verdeutlicht, dass die Abbildung der beiden Behälter den Zeitpunkt vor der Ziehung zeigt, also bevor eine Kugel in den Auswurf fiel. Sie wird anschließend gefragt, aus welchem Behälter die in Auswurf befindliche Kugel höchstwahrscheinlich stammt. Nach der Wahl des Behälters erfolgt die Bearbeitung der zweiten Teilaufgabe, in dem die Testperson eine Schätzung der Wahrscheinlichkeit auf einem Zahlenband (von 0 % bis 100 %) vornehmen muss (Abbildung 13). Bei der Schätzung wird der Bezug von 100 Ziehungen (Durchgängen) hergestellt, in denen ebenfalls nur Kugeln mit der gleichen Farbe zählten, wie jene sich im Auswurf Befindende. Grundsätzlich kann die exakte statistische Wahrscheinlichkeit nicht ad hoc berechnet werden.

Von den 30 Items werden 10 unter Zeitdruck vorgegeben. Die verfügbare Zeit wird bei jedem Item reduziert, beim ersten Item haben die Testpersonen 25 Sekunden und beim letzten Item nur mehr 10 Sekunden Zeit für eine erfolgreiche Bearbeitung. Die verbleibende Zeit wird mittels eines roten Countdown eingeblendet (Abbildung 14). Die Testperson erhält Informationen zur Zeitlimitierung zu Beginn des Itemblocks „Zeitdruck“ und wenn das Zeitlimit überschritten wird. Ist die Zeit abgelaufen, wird die Testperson darauf hingewiesen, dass sie zu langsam war und die Aufgabe als falsch gewertet wird. Mit Bestätigung der Zeitablaufinformation wird das nächste Item präsentiert (Abbildung 17 weist auf die bevorstehende Zeitlimitierung hin, Abbildung 18 auf die Beachtung der Zeitvorgabe).

¹ Itembeispiel frei erfunden.

Der Testperson stehen bei der Bearbeitung der Items sechs verschiedene Hilfen zu Verfügung (siehe Abbildung 11; Abbildung 15 und 16 zeigen die Hilfen „Entscheidung der letzten drei Personen“ und „Farbe nächste Kugel selber Behälter“). Sie hat die Möglichkeit, bis zu zwei Hilfen in Anspruch zu nehmen, wobei die Hilfen bis zum Anklicken des Buttons „Jetzt Entscheiden“ abrufbar sind. Nach dem ersten Item werden nur mehr fünf Hilfen variiert angeboten, damit die Testperson verschiedene Hilfen verwendet.

Die Einteilung der Hilfen erfolgt in sinnvolle Hilfen (Expertenrat einholen, Computersimulation und Farbe nächste Kugel selber Behälter) und in sinnlose Hilfen (Entscheidung der letzten drei Personen, Computer soll entscheiden und Farbe vorherige Kugel selber Behälter). Die Bezeichnung „sinnvolle Hilfe“ basiert aufgrund logischer Prämissen: Tatsächlich können sich Experten manchmal irren und bei Simulationen zu geringe Daten verwendet worden sein. Die Hilfe „Farbe nächste Kugel selber Behälter“ ist insofern zu den sinnvollen Hilfen zu zählen, weil sich die a-priori Wahrscheinlichkeit verändert, während die Hilfe „Farbe vorherige Kugel selber Behälter“ ein vergangenes und unabhängiges Ereignis darstellt.

Die sinnvollen Hilfen werden aufgrund ihrer inhaltlichen Überlegungen als solche bezeichnet. Faktisch bekommen Testpersonen nur eine bedingt richtige Hilfestellung, so dass zum Beispiel der Experte nicht immer einen „richtigen“ Rat, bezogen auf die mathematische Wahrscheinlichkeit, gibt.

Die Itembearbeitung erfolgt in drei Stufen: In der ersten bekommt die Testperson die Aufgabe inklusive Hilfsmöglichkeiten präsentiert. In Stufe zwei muss sie den favorisierten Behälter auswählen (Abbildung 12) und in der letzten Stufe erfolgt die Wahrscheinlichkeitsschätzung. Sollte die Testperson während der Schätzung Bedenken bezüglich ihrer Behälterwahl haben, so kann sie in Stufe zwei zurückkehren und den Behälter ändern.

Abbildung 11: Itembeispiel, (Stufe 1)

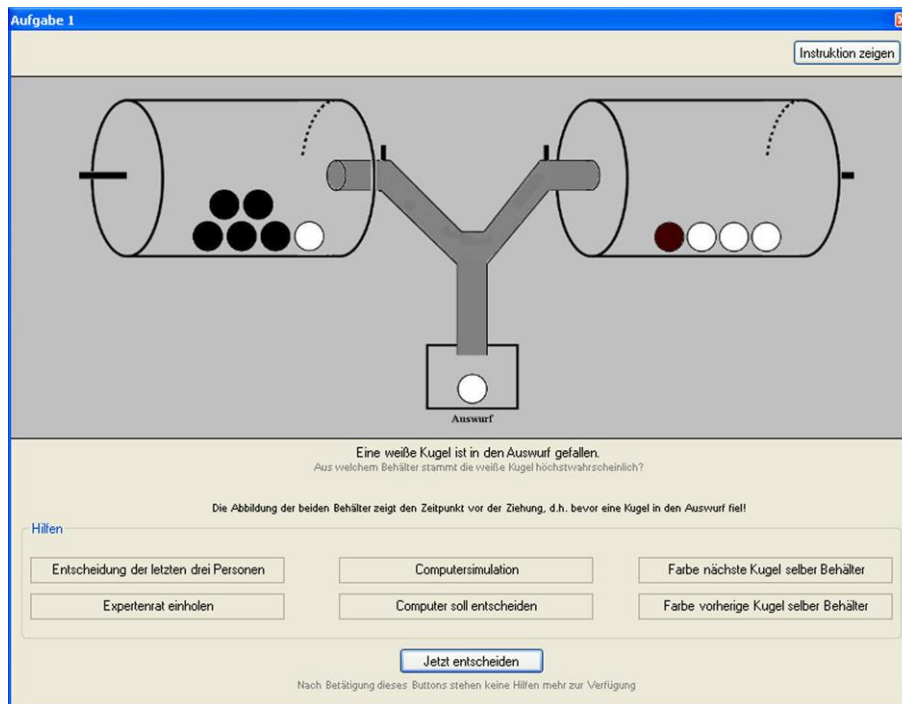


Abbildung 12: Behälterwahl (Stufe 2)

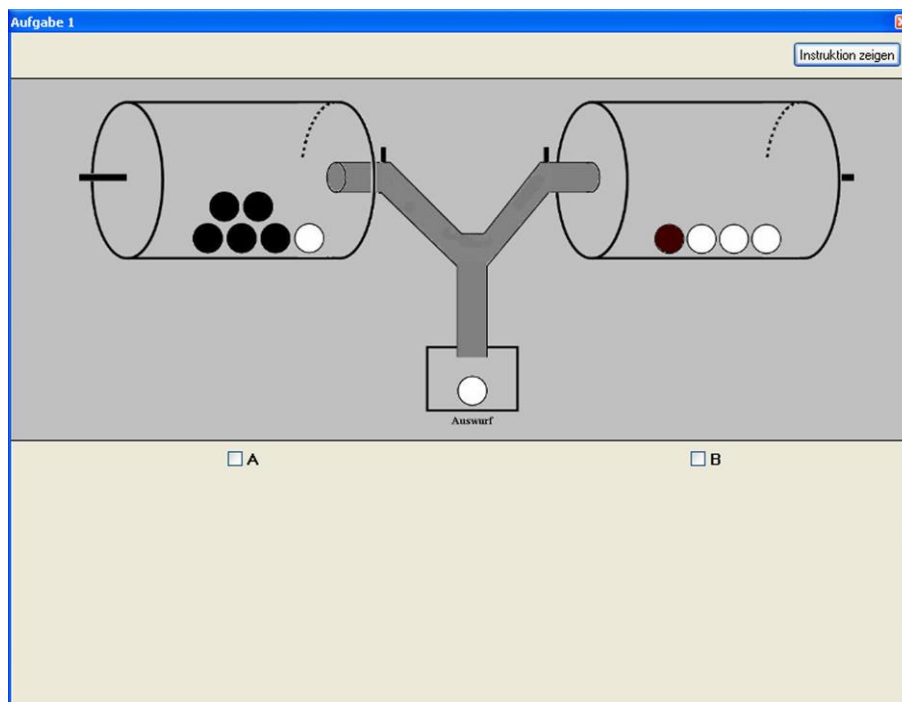


Abbildung 13: Wahrscheinlichkeitsschätzung (Stufe 3)

Aufgabe 1

Instruktion zeigen

Geben Sie nun eine Schätzung ab, in wie vielen Fällen von 100 Durchgängen, in denen nur weiße Kugeln im Auswurf zählen, die weiße Kugel aus Ihrem gewählten Behälter stammt.

Geben Sie Ihre Schätzung am Zahlenband an:

0% Schätzung: 71 % 100%

Neu entscheiden Nächste Aufgabe

Abbildung 14: Itembeispiel¹ unter Zeitdruck

Aufgabe 21

Verbleibende Zeit: 21 Sekunden

Instruktion zeigen

Eine weiße Kugel ist in den Auswurf gefallen.
Aus welchem Behälter stammt die weiße Kugel höchstwahrscheinlich?

Die Abbildung der beiden Behälter zeigt den Zeitpunkt vor der Ziehung, d.h. bevor eine Kugel in den Auswurf fiel!

Hilfen

Entscheidung der letzten drei Personen Computersimulation Farbe nächste Kugel selber Behälter

Computer soll entscheiden Farbe vorherige Kugel selber Behälter

Jetzt entscheiden

Nach Betätigung dieses Buttons stehen keine Hilfen mehr zur Verfügung

¹ Itembeispiel frei erfunden

Abbildung 15: sinnlose Hilfe

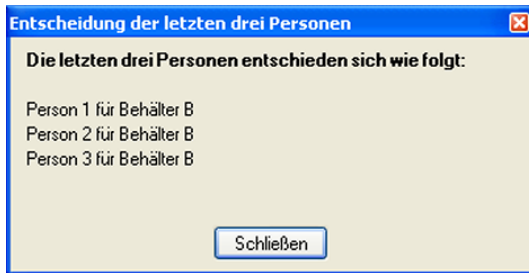


Abbildung 16: sinnvolle Hilfe

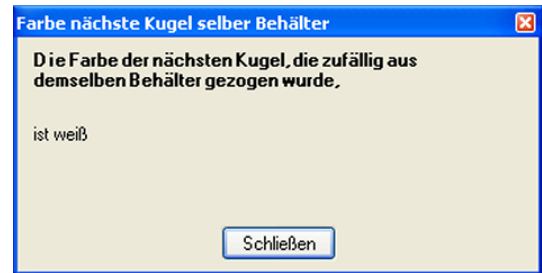


Abbildung 17: Hinweis auf Zeitlimitierung

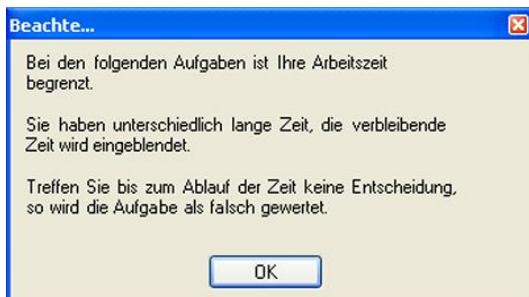
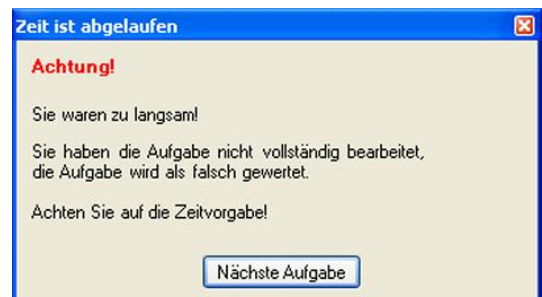


Abbildung 18: Beachtung der Zeitvorgabe



Beim EUcO 2.0 werden zwei Reaktionszeiten registriert, die Zeit für die Behälterwahl einschließlich der Inanspruchnahme der Hilfen (Stufe 1) und die Zeit, welche für die Schätzung der Wahrscheinlichkeit (Stufe 3) benötigt wird. Die Zeit für einen etwaigen Behälterwechsel wird mit der zweiten Reaktionszeit erfasst.

Während der gesamten Testphase besteht für die Testperson die Möglichkeit des wiederholten Instruktionsabrufes.

Beim EUcO 2.0 wird, wie bereits in der Vorgängerversion, die Konsistenz des Entscheidungsverhaltens beurteilt, in dem einige Items mit vertauschten Behältern oder vertauschten Farben bzw. mit oder ohne Zeitdruck vorgegeben werden.

Die Testperson kann für eine Entscheidung zwischen den Alternativen Behälter A oder B wählen, wobei sie keinen Einfluss auf die gezogene Kugel hat. Alleine der Zufall bestimmt für die Testperson die augenscheinlich gezogene Kugel. Die Unsicherheit ist durch die verschiedenen Wahrscheinlichkeiten bedingt.

Folgende Testkennwerte können aus der EUcO 2.0-Vorgabe abgeleitet werden und repräsentieren somit das Entscheidungsverhalten einer Person:

- » Anzahl richtiger Entscheidungen
- » Entscheidungsgüte anhand der Wahrscheinlichkeitsschätzung
- » Konsistenz des Entscheidungsverhaltens
- » Verwendung von Hilfen
- » Attributionsstil der Person
- » Einfluss von Zeitdruck
- » Entscheidungsgeschwindigkeit

3. Empirischer Teil

3.1 Fragestellungen und Hypothesen

Das Ziel dieser Diplomarbeit ist die Verbesserung eines Computertests zum Entscheidungsverhalten. Radinger (2002) hat dazu einen Test entwickelt, der im Zuge der praktischen Anwendung und wissenschaftlichen Beschäftigung Kritik auslöste. Die unter Kapitel 2.4.1 behandelten Kritikpunkte sollen aufgenommen und bereinigt werden und durch eine Neukonstruktion zum EUcO 2.0 führen.

Der EUcO bzw. der EUcO 2.0 wurde als Objektiver Persönlichkeitstest konstruiert bzw. weiterentwickelt. Durch die Verwendung von Objektiver Persönlichkeitstests wird das Messprinzip von Seiten des Getesteten nicht durchschaut: damit wird dem Gütekriterium der „Un-Verfälschbarkeit“ Rechnung getragen.

Fragestellungen:

1. Führte die Neukonstruktion des EUcO 2.0 zu einer besseren Verständlichkeit?
2. Kann durch die Vorgabe von Beispielaufgaben die Einschätzung der Wahrscheinlichkeit auf Testpersonenseite genauer gelingen?
3. Kann durch die Verwendung von Hilfen ein neuer theoriegeleiteter Testkennwert extrahiert werden?
4. Wie veränderten sich die Testkennwerte im Vergleich zur alten Version?

Die daraus folgenden Hypothesen:

1. Unterscheiden sich die Testkennwerte in den verschiedenen Gruppen?
2. Gibt es Unterschiede in den Testkennwerten zwischen EUcO (Version Radinger, 2002) und EUcO 2.0?
3. Unterscheiden sich die Personen in der Wahl der verwendeten Hilfskategorie?
4. Gibt es einen negativen Zusammenhang zwischen den Hilfskategorien?

3.2 Datenerhebung, Stichprobenbeschreibung und Testdurchführung

Im Rahmen der Lehrveranstaltung „ Psychologie II“ an der Fachhochschule (FH) Linz des Bachelorstudium „Soziale Arbeit (Vollzeitstudium)“ fand im Juni 2007 die Datenerhebung statt. Der Studiengang „Soziale Arbeit“ ist eine Ausbildung für Personen zur Beratung, Begleitung und Unterstützung von Menschen sowie zur Konzeption und Durchführung von Hilfsprozessen. Das Ausbildungsziel ist zu lernen, wie Helfen gelingt.

Zum Zeitpunkt der Testung waren die FH-Studenten im Alter von 21 bis 36 Jahren, durchschnittlich waren sie 24,22 Jahre alt. Der EUcO 2.0 wurde 32 FH-studenten vorgegeben; 65,63 % (21 Personen) der getesteten FH-Studenten waren weiblich. Die Aufnahme an Fachhochschulen setzt nicht zwingend die Hochschulreife voraus, durch eine facheinschlägige Berufsausbildung mit Zusatzprüfung oder die Ablegung der Studiumsberechtigungsprüfung können die Voraussetzungen und Anforderungen ebenfalls erfüllt werden. Die Schulbildung der FH-Studenten wurde nicht gesondert ermittelt: auf Grund der vorher genannten Zugangsvoraussetzungen für Hochschulen wird Hochschulreife niveau argumentiert. Ihnen wurde bei der statistischen Auswertung die Bildungsgruppe 3 zugewiesen.

Die Testung wurde im Computerraum des Fachhochschulcampus durchgeführt. Auf Grund der begrenzten Anzahl an Computerplätzen wurden die FH-Studenten in zwei Gruppen unterteilt. Indem die Studenten in der Nacht zuvor von einem einwöchigen Studienaufenthalt aus Schweden zurückkamen, zeigten viele von ihnen noch Müdigkeitssymptome. Manche Studenten blieben der Lehrveranstaltung fern.

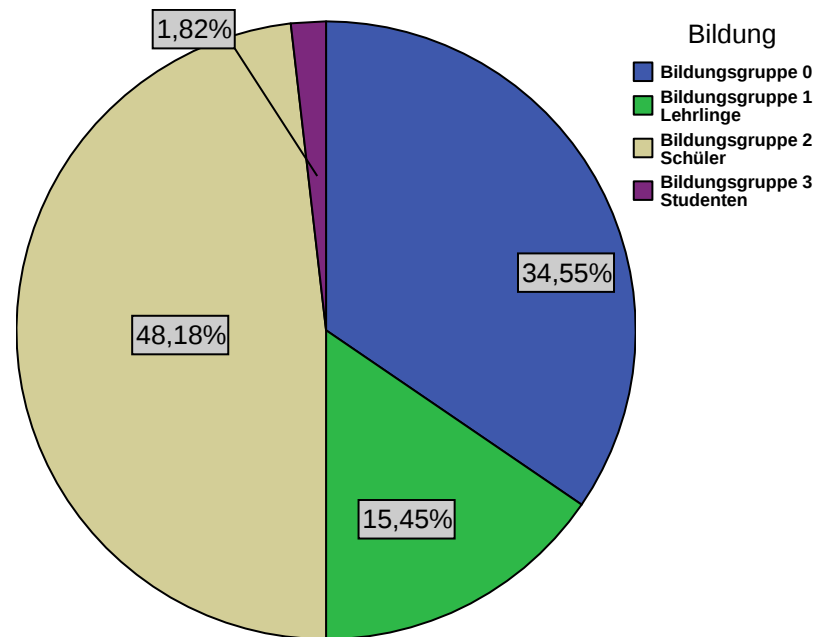
Von September bis Oktober 2009 wurden bei der Stellungskommission des österreichischen Bundesheeres in Wien 110 Stellungspflichtige mit Genehmigung des Heerespsychologischen Dienstes getestet. Im Rahmen der Untersuchung zur Eignung der gesetzlichen Wehrpflicht konnten Stellungspflichtige zwischen den

Pausen von psychologischer und medizinischer Befundung freiwillig an der EUcO 2.0-Vorgabe teilnehmen. Es wurden mit den Psychologen vor Ort Kriterien besprochen, welche für eine Einladung an der Teilnahme der Testung ausschlaggebend waren. Im Hinblick auf das verbale Verständnis der Instruktion des EUcO 2.0 und der unterschiedlich interindividuellen Sprachbegabung der Stellungspflichtigen wurden Personen mit auffälligen Sprachdefiziten nicht angesprochen. Eine Minderbegabung konnte sowohl anhand der Testbatterie, die jeder Stellungspflichtige durchlaufen muss oder durch den Psychologen diagnostiziert werden. Als weitere Kriterien wurden die psychische und soziale Unauffälligkeit sowie Alkohol- und Drogenabstinenz (Alkoholranke bzw. Drogenabhängige) festgelegt. Zuletzt war das Ergebnis der Testbatterie insgesamt entscheidend, ob Stellungspflichtige auffällige Minderbegabung für eine Mitwirkung in Frage kamen und somit eingeladen wurden oder nicht. Die Testbatterie des österreichischen Bundesheeres enthält neben Fragen zur Persönlichkeit (z.B. prosoziales Verhalten, emotionale Stabilität, positive Lebenseinstellung) auch leistungsbezogene Dimensionen (z.B. Wortgewandtheit, Raumvorstellung, Wahrnehmungsfähigkeit, Belastbarkeit). Körperliche Beeinträchtigung war kein Ausschlusskriterium.

78,2 % (das sind 86) der Personen waren zum Zeitpunkt der Testung 18 Jahre alt, 18,2 % mit 17 Jahren (das sind 20), eine Person war 16 Jahre alt. In Einzelfällen, wegen Verletzungen oder anderer besonderer Gründe kann die Eignung zum festgesetzten Stellungstermin nicht erfolgen. Darum werden Stellungspflichtige zu einem anderen Zeitpunkt zur so genannten Nachstellung vorgeladen. Während der EUcO 2.0-Vorgabe waren dies drei Personen im Alter von 20 Jahren. Das Durchschnittsalter der Personen war 17.9 Jahre. Im Zeitraum der Testvorgabe wurden ausschließlich Männer getestet. Die Stellungspflichtigen wurden in vier verschiedene Schulbildungsgruppen eingeteilt. Der Bildungsgruppe 0 wurden 38 Stellungspflichtige zugeschrieben. Sie bestand aus den Personen, deren Bildungsniveau dem Testergebnis nicht zugewiesen werden konnte. 17 Stellungspflichtige wurden der Bildungsgruppe 1 zugeschrieben: sie umfasste Personen, die eine Lehre machten. Die Bildungsgruppe 2 bestand ausschließlich

aus Schülern (HTL, HAK, AHS, Gymnasium, usw.); sie war die größte Gruppe mit 53 Personen. Der Gruppe der Studierenden (Bildungsgruppe 3) konnten 2 Personen zugeordnet werden. In der Abbildung 19 werden Prozentangaben dargestellt.

Abbildung 19: Bildungsgruppen in der Teilstichprobe der Stellungspflichtigen



Auf den Computern des österreichischen Bundesheeres durfte keine externe Computersoftware ohne rigorose Prüfung und Zustimmung der Verantwortlichen installiert werden. Für die Testdurchführung wurden deshalb in einem Raum mobile Computer aufgebaut, an denen die Stellungspflichtigen nach deren Einverständnis den EUcO 2.0 absolvieren konnten. Aufgrund der vorgeschriebenen Rahmenbedingung zur Eignungsfeststellung konnte keine Gruppentestung durchgeführt werden, die Freiwilligen wurden nacheinander getestet. Nur gelegentlich konnten einige Personen den EUcO 2.0 gleichzeitig bearbeiten.

Die Psychologen des Heerespsychologischen Dienstes schätzten, dass der Prozentsatz an Stellungspflichtigen die eine Teilnahme am EUcO 2.0 ablehnten, geringer als 10 % gewesen sei.

Die Altersverteilung ist in Abbildung 20 dargestellt. Tabelle 1 gibt eine Übersicht der Zusammensetzung der Gesamtstichprobe.

Abbildung 20: Häufigkeitsverteilung des Alters von Bundesheer und FH-Studenten

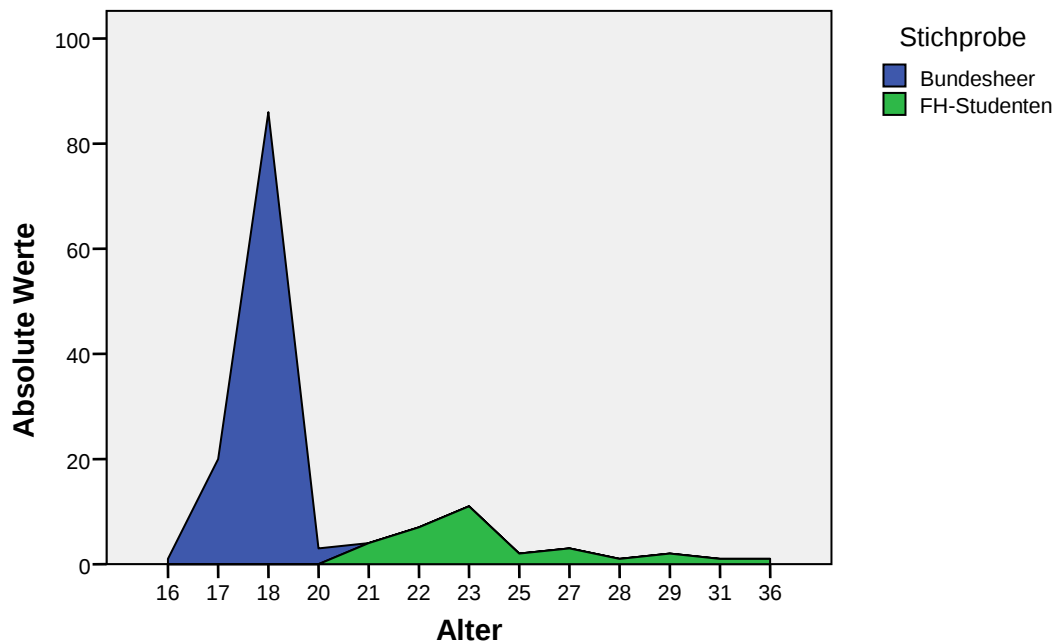


Tabelle 1: Stichprobenzusammensetzung

	Stichprobengröße	Alter	Geschlecht		Bildungsgruppe			
	N	in Jahren	männlich	weiblich	0	1	2	3
FH-Studenten	32	21 - 36	11	21	0	0	0	32
Stellungspflichtige	110	16 - 20	110	0	38	17	53	2
Gesamt	142	16 - 36	121	21	38	17	53	34

In der Absicht, den EUcO 2.0 als objektiven Persönlichkeitstest zu verwenden, wurden die Testpersonen bei der Testvorgabe nur dahingehend informiert, dass der Anlass der EUcO 2.0-Vorgabe eine Studie zur Einschätzung von Wahrscheinlichkeiten sei.

Die Täuschung der tatsächlichen Testintention sollte Annahmen über das wahre Bestreben des EUcO 2.0 und diesbezügliche Verfälschungen vorbeugen.

Den teilnehmenden Testpersonen wurden später die Ergebnisse und die wahre Testintention präsentiert.

3.3 Ergebnisdarstellung

Der EUcO 2.0 liefert eine Menge an Informationen, die zu Beginn dargestellt werden. Diese Informationen werden zu den verschiedenen Testkennwerten zusammengefasst: Rohscore, Entscheidungsgenauigkeit (geschätzte Wahrscheinlichkeit), Entscheidungskonsistenz und –geschwindigkeit, Verwendung von Hilfen, der Einfluss von Zeitdruck und der Attributionsstil einer Person.

Nach der Ergebnisdarstellung erfolgt die Überprüfung der Hypothesen bzgl. Unterschiede zwischen den Gruppen. Weiters wird untersucht, ob tatsächlich ein neuer Testkennwert gewonnen werden konnte.

Zum Schluss werden die Ergebnisse des EUcO 2.0 der alten Version (Radinger, 2002) gegenübergestellt.

3.3.1 Auswertung der Testkennwerte des EUcO 2.0

Die Auswertung der Informationen die der EUcO 2.0 liefert, erfolgt mit der Software SPSS für Windows – Version 15.0.

3.3.1.1 Rohscore

Auf den folgenden Seiten erfolgt die Darstellung des Rohscores in Form eines Histogramms für die gesamte Stichprobe (Abbildung 21) mit Normalverteilungskurve. Auf Grund des Aufgabenmaterials liegt die theoretische Ratewahrscheinlichkeit je Item bei 0,5. Der Erwartungswert bei 30 Items beträgt

infolgedessen 15, wenn nur geraten wird. Nur 10 Personen weisen einen Rohscore von kleiner gleich 15 auf.

In Tabelle 2 und 3 werden die einzelnen Mittelwerte der Teilstichproben dargestellt. In der Teilstichprobe Geschlecht werden nur Personen der Bildungsgruppe 3 berücksichtigt, um Verzerrungen zu vermeiden. Die Mittelwerte des Rohscores unterscheiden sich nicht signifikant voneinander.

Abbildung 21: Gesamtstichprobe mit Normalverteilungskurve

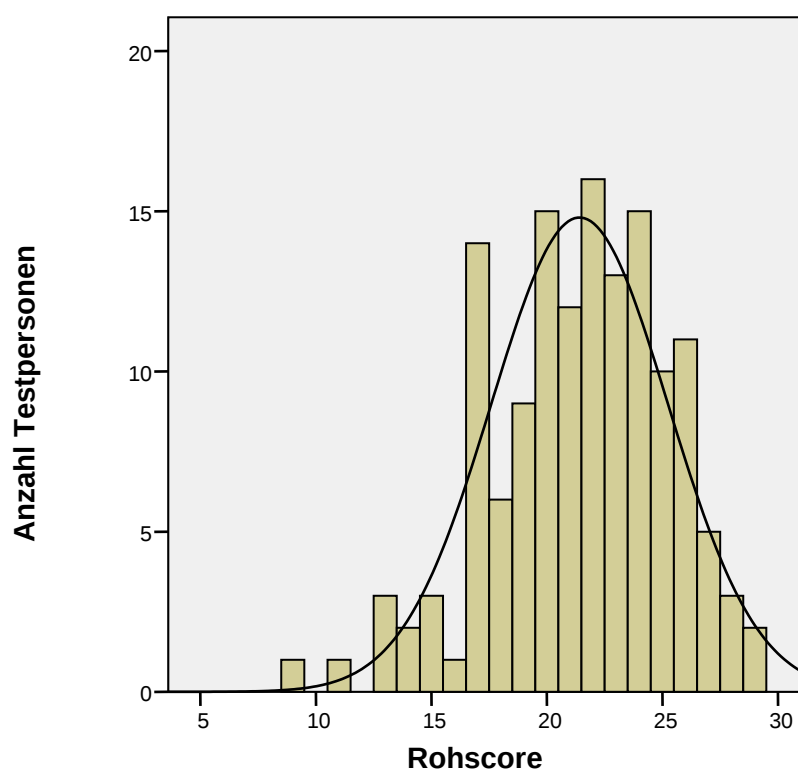


Tabelle 2: Mittelwert und Standardabweichung des Rohscores der Gesamtstichprobe und der Teilstichproben

	Mittelwert	N	Standardabweichung
Gesamtstichprobe	21,42	142	3,83
Bundesheer	21,37	110	3,96
FH-Studenten	21,56	32	3,39

Tabelle 3: Mittelwert und Standardabweichung des Rohscores nach Bildungsgruppe und Geschlecht

	Mittelwert	N	Standardabweichung
Bildungsgruppe 0	21,1	38	3,37
Bildungsgruppe 1	20,24	17	4,75
Bildungsgruppe 2	21,75	53	4,04
Bildungsgruppe 3	21,82	34	3,46
Männlich	23,15	13	3,87
Weiblich	21	21	2,98

Radinger (2002) hat in ihrer Arbeit einen so genannten Cut-Off-Score mittels Clusteranalyse berechnet. Anhand des Cut-Off-Scores sollen jene Testpersonen identifiziert werden, die bei der Beantwortung der Aufgaben eine „rationale Strategie“ anwenden. Der Cut-Off-Score wurde bei 22 gelösten Items festgelegt. Jene Personen, die einen Rohscore von 21 und weniger erreichen, werden zur Gruppe der Personen gezählt, die eine „intuitive Strategie“ wählen. Tabelle 4 gibt Auskunft darüber, wie hoch der Prozentsatz der Personen mit rationaler bzw. intuitiver Strategie in der Gesamtstichprobe und den Teilstichproben ist.

Tabelle 4: Prozentsatz der Personen mit intuitiver und rationaler Strategie

	intuitive Strategie		rationale Strategie	
Gesamtstichprobe	67	47,18%	75	52,82%
Bundesheer	52	47,27%	58	52,73%
FH-Studenten	15	46,86%	17	53,14%

3.3.1.2 Entscheidungsgüte

Ein weiterer Testkennwert ist die Entscheidungsgüte. Sie wird anhand der geschätzten Wahrscheinlichkeit aus den Testinformationen entnommen. Die mathematischen Wahrscheinlichkeiten der 30 Items liegen zwischen 51 % und 86%. Der Mittelwert beträgt 58 %. Im Durchschnitt beträgt die geschätzte Wahrscheinlichkeit über alle Personen und Items 63,03 %. Der Mittelwert der Schätzungen aller Testpersonen liegt im Bereich von 0 bis 80,33 %.

Die Testpersonen werden in der Instruktion darauf hingewiesen, dass die Wahrscheinlichkeitsschätzung nicht unter 50 % liegen kann, weil dies einen Widerspruch des gewählten Behälter darstellen würden. Zwei Personen geben während der gesamten Testphase keine einzige Wahrscheinlichkeitsschätzung ab. Und zwei weitere Personen geben nur Schätzungen ab, welche unter 50 % liegen. Sechs Testpersonen markieren den Wahrscheinlichkeitsschätzer bei mindestens 10 Aufgaben unter 50 %. 86 Personen machen mindestens einmal eine Schätzung unter 50 %, nur 44 Testpersonen handeln instruktionskonform und markieren alle Itemschätzungen über 50 %.

Aus der Tabelle 5 können die Mittelwerte entnommen werden, wenn nur jene Datensätze für die Auswertung verwendet werden, bei denen die Testpersonen instruktionskonform gearbeitet haben. Das heißt, dass die Schätzung der Wahrscheinlichkeit über 50 % liegt und das Item richtig gelöst wurde.

Tabelle 5: Mittelwerte bei instruktionskonformer Testbearbeitung

	Mittelwert	Minimum	Maximum	N
Item 1 - 30	68,49	58,85	87,18	138
Item 1 - 20	67,93	59,5	81,3	138
Item 21 - 30	69,67	51,00	100	136

Betrachtet man nur die gelösten Aufgaben, wobei die Schätzungen über 50 % liegen müssen, so werden 80,25 % der Items über- bzw. 19,25 % der Items unterschätzt. Nur eine Person hat mehr Unterschätzungen als Überschätzungen. Bei allen anderen Personen überwiegen die Überschätzungen.

Die mittleren absoluten Abweichungen betragen im Durchschnitt 4 % bei der Unterschätzung und 16,45 % bei der Überschätzung. Die höchste Überschätzung liegt bei 27 Prozentpunkten und die höchste Unterschätzung bei 11 Prozentpunkten.

Auf Grund der Tatsache, dass die Wahrscheinlichkeiten nicht genau geschätzt werden können, werden die Schätzungen auf drei Intervalle aufgeteilt.

Im Intervall 1 liegen alle „Treffer“ (richtige Schätzungen) einer Testperson, welche die statistisch berechnete Wahrscheinlichkeit genau schätzte oder deren absolute Abweichungen ± 1 betragen. Im Intervall 2 werden absolute Abweichungen bis ± 3 abgebildet und im Intervall 3 absolute Abweichungen bis ± 5 , d.h. im Intervall 3 sind auch die „Treffer“ der Intervalle 1 und 2. In Tabelle 6 werden die einzelnen Mittelwerte der Absolutabweichungen über die Gesamtstichprobe und der Teilstichproben abgebildet. Im Durchschnitt hat eine Person in der Gesamtstichprobe 1,78 richtige Schätzungen im Intervall 1 und 3,7 richtige Schätzungen im Intervall 2.

Tabelle 6: Mittelwerte der absoluten Abweichungen

	Mittelwert			N
	Intervall 1	Intervall 2	Intervall 3	
Gesamtstichprobe	1,78	3,7	6,12	142
Bundesheer	1,68	3,5	5,97	110
FH-Studenten	2,13	4,41	6,63	32

Tabelle 7 gibt die Mittelwerte über die Bildungsgruppen und das Geschlecht wieder. Um das Ergebnis nicht zu verzerren, wird beim Geschlecht nur die Bildungsgruppe 3 zur statistischen Berechnung herangezogen. Bei der statistischen Überprüfung der Mittelwerte, kann in den Teilstichproben kein signifikanter Mittelwertsunterschied festgestellt werden.

Tabelle 7: Mittelwerte der absoluten Abweichungen, Intervall 1 - 3

	Mittelwert			N
	Intervall 1	Intervall 2	Intervall 3	
Bildungsgruppe 0	1,87	3,37	5,95	38
Bildungsgruppe 1	1,24	3,47	5,65	17
Bildungsgruppe 2	1,74	3,64	5,65	53
Bildungsgruppe 3	2,03	4,29	6,41	34
Männlich	2,15	4,54	6,62	13
Weiblich	1,95	4,14	6,29	21

3.3.1.3 Konsistenz

Um die Konsistenz des Entscheidungsverhaltens zu überprüfen, werden zehn Items in derselben oder in ähnlicher Form zweimal vorgegeben. Mit dem χ^2 -Test nach McNemar wird überprüft, ob die Itempaare für die Testpersonen gleich schwierig sind. Testpersonen, die eine Computerentscheidung wählen oder deren

Zeit abgelaufen ist, werden bei der statistischen Auswertung nicht berücksichtigt. In Tabelle 8 sind die Ergebnisse nachzulesen.

Bei vier Itempaaren ergibt der χ^2 -Test ein signifikantes Ergebnis. Diese sind für die Testpersonen nicht gleich schwierig, obwohl die mathematischen Wahrscheinlichkeiten gleich sind. Zwei Itempaare, die unter Zeitdruck vorgegeben werden, erscheinen für die Testpersonen als gleich schwierig bzw. wenn man sich die Häufigkeiten ansieht als gleich leicht. 129 Personen haben Itempaar 5-23 bearbeitet; davon haben 122 Personen beide richtig. Das Itempaar 13-27 markierten 111 Testpersonen; 105 haben beide richtig.

Bis auf die Itempaare 9-15 und 4-19 werden alle anderen Paare von mindestens 2/3 der Testpersonen gelöst. Mit Ausnahme des Itempaares 4-19 wird das zweite Item öfter gelöst.

Tabelle 8: χ^2 - Test Ergebnisse

Itempaar			χ^2	Sig.
2	20	gleiches Item	0,15	0,985
3	18	gleiches Item	10,22	0,041
4	19	gleiches Item	8,15	0,322
7	17	Farbe der Kugeln vertauscht	1,05	0,581
8	16	Behälter vertauscht	2,98	0,049
9	15	doppelte bzw. halbe Anzahl der Kugeln	12,69	<0,001
10	14	gleiches Item	17,49	0,029
Itempaar				
5	23	zweites Item unter Zeitdruck	23,57	1.000
12	25	zweites Item unter Zeitdruck	13,46	0,581
13	27	zweites Item unter Zeitdruck	0,342	1.000

3.3.1.4 Verwendung von Hilfen

Durchschnittlich werden circa 10 Hilfen in Anspruch genommen; der Median beträgt 7. Für den gesamten Test beträgt das 1. Quartil 1,75 und 13,25 Hilfen im 3. Quartil. 24 Testpersonen bzw. 16,9 % wählen keine Hilfe aus; 14 % beanspruchen mehr als 25 Hilfen. In Tabelle 9 sind dazu die Mittelwerte aufgelistet.

Am häufigsten wird die Hilfe „Expertenrat“ in Anspruch genommen, am wenigsten die Hilfe „Farbe vorheriger Kugel selber Behälter“. In Tabelle 10 werden die Mittelwerte jener Personen, die mindestens eine Hilfe verlangten, nach erster und zweiter Hilfe zusammengefasst.

Tabelle 9: Mittelwert und Standardabweichung der Hilfen, n = 142

Hilfen	Mittelwert	Standardabw.
Hilfen gesamt	9,98	10,78
Expertenrat	3,44	4,65
Computersimulation	2,18	3,21
Entscheidung der letzten drei Personen	1,65	2,39
Computer soll entscheiden	1,05	1,88
Farbe nächste Kugel selber Behälter	1,03	2,25
Farbe vorherige Kugel selber Behälter	0,63	1,57

Tabelle 10: Mittelwerte der ersten und zweiten Hilfe, n = 118

Hilfen	Mittelwert	
	1. Hilfe	2. Hilfe
Hilfen gesamt	8,73	3,28
Expertenrat	3,21	0,92
Computersimulation	1,88	0,74
Entscheidung der letzten drei Personen	1,55	0,44
Computer soll entscheiden	1,05	0,21
Farbe nächste Kugel selber Behälter	0,68	0,56
Farbe vorherige Kugel selber Behälter	0,36	0,41

Berücksichtigt man nur jene Personen, die mindestens eine Hilfe verwenden, so werden im Durchschnitt bei Items ohne Zeitlimitierung jeweils 0,73 Hilfen verwendet. Bei Items, die unter Zeitdruck bearbeitet werden, sinkt dieser Wert auf 0,19.

Nach inhaltlichen Überlegungen können die Hilfen in sinnvolle und sinnlose unterteilt werden. Sinnvolle Hilfen sind:

- Expertenrat einholen,
- Computersimulation und
- Farbe nächste Kugel selber Behälter.

Diese Hilfen geben der Testperson keine ehrliche Auskunft über den richtigen Behälter. Ausgehend von der Prämisse, dass sie verlässliche Informationen im Sinne von „fachgerechten“ Hinweisen geben, werden sie zu den sinnvollen Hilfen gezählt. Die Korrelation zwischen der Anzahl der beanspruchten sinnvollen und sinnlosen Hilfen ist signifikant und beträgt $r = 0,598$. Conclusio dieses positiven Zusammenhangs: Testpersonen, die viele sinnvolle Hilfen beanspruchen, wählen auch eher sinnlose Hilfen aus.

Eine weitere Hilfe, welche aber nicht explizit einen Hinweis auf einen Behälter gibt, ist das Abrufen der Instruktion. Sie kann während der gesamten Testphase wiederholt beansprucht werden. 93,7 % der Personen haben die Instruktion nach der Einführung nicht, 4,9 % einmal und 1,4 % wiederholt abgerufen.

3.3.1.5 Attributionsstil einer Person

Um den Attributionsstil einer Person abzuleiten, werden die Korrelationen zwischen den beanspruchten Hilfen untersucht. Die positive Korrelation zwischen sinnvollen und sinnlosen Hilfen legt die Vermutung nahe, dass Testpersonen bestimmte Hilfearten präferieren. Die Hilfen werden dazu in die Kategorien „Person“, „Situation“ (Computer) und „Stimulus“ (Kugel) eingeteilt.

Am häufigsten wird von den Testpersonen (n=107), welche über die gesamte Testphase mindestens zwei Hilfen verlangen, die Hilfekategorie „Person“ (mit einem Mittelwert von 6,71) in Anspruch genommen. Etwas weniger Verwendung findet die Kategorie „Situation“ (Mittelwert 4,22) und am geringsten wird die Hilfekategorie „Stimulus“ benutzt (Mittelwert 2,21).

Erwartungsgemäß müssten die Korrelationen zwischen den Hilfekategorien signifikant **und** negativ sein. Dass dem nicht so ist, zeigt Tabelle 11.

Tabelle 11: Korrelationen zwischen den Kategorien

			Korrelation	Sig.
Person	zw.	Situation	0,336	<0,001
Person	zw.	Stimulus	0,208	0,032
Situation	zw.	Stimulus	0,248	0,01

3.3.1.6 Einfluss von Zeitdruck

Rohscore

Ein Item zählt als vollständig gelöst, wenn beide Teilaufgaben innerhalb der vorgegebenen Zeit bearbeitet werden. Hat eine Person einen Behälter ausgewählt, jedoch zu wenig Zeit für die Wahrscheinlichkeitsschätzung, so wird das gesamte Item als „Zeit abgelaufen“ verrechnet. Für 40,8 % der Testpersonen reicht der vorgegebene Zeitrahmen, um alle Items komplett zu bearbeiten. 35,9 % überschreiten einmal das vorgegebene Zeitlimit und für 23,3 % ist die Zeit bei bis zu acht Items zu kurz.

Auf 10,56 % der Testpersonen hat die Zeitlimitierung keinen Einfluss; sie lösen gleich viele Items ohne bzw. unter Zeitdruck. 30,99 % lösen unter Zeitdruck mehr Aufgaben und 58,45 % der Personen haben ohne Zeitdruck einen höheren Anteil von richtig gelösten Items. Mittels Wilcoxon-Test ($p < 0,001$) konnten signifikante Mittelwertsunterschiede zwischen dem Anteil gelöster Items ohne und unter Zeitdruck festgestellt werden. Die Korrelation beträgt 0,393 und ist signifikant.

Wahrscheinlichkeitsschätzung

67,93 % beträgt die durchschnittlich geschätzte Wahrscheinlichkeit ohne Zeitdruck, unter Zeitdruck 69,66 %. Mittels t-Test konnte ein signifikanter Mittelwertsunterschied ($p = 0,001$) festgestellt werden. Allerdings sind auch die mathematischen Wahrscheinlichkeiten der Items ohne Zeitdruck um einen Prozentpunkt niedriger. Die Korrelation der Schätzungen ohne und unter Zeitdruck beträgt 0,650 ($p < 0,001$).

Mittels Wilcoxon-Test konnten in allen drei Intervallen (siehe Kapitel 3.3.1.2) für die absoluten Abweichungen signifikante Mittelwertsunterschiede ($p \leq 0,001$) ohne und unter Zeitdruck festgestellt werden. Die Korrelation beträgt bei Intervall 1 ohne und unter Zeitdruck $r = 0,179$; bei Intervall 2 $r = 0,435$ und bei Intervall 3 $r = 0,468$.

Einsatz von Hilfen

Unter Zeitdruck werden durchschnittliche 1,87 Hilfen verwendet, ohne Zeitdruck 8,11 Hilfen. Mittels Wilcoxon-Test konnten signifikante Mittelwertsunterschiede festgestellt werden. Unter Zeitdruck werden weniger Hilfen beansprucht. Die Korrelation zwischen der Anzahl der verwendeten Hilfen ohne und unter Zeitdruck beträgt 0,648.

3.3.1.7 Entscheidungsgeschwindigkeit

Die benötigte Zeit wird bei jedem Item erfasst. Dieser Zeitraum unterteilt sich in den 1. Zeitabschnitt für die Behälterwahl, einschließlich der optionalen Inanspruchnahme von Hilfen und in den 2. Zeitabschnitt für die Wahrscheinlichkeitsschätzung. Abbildung 20 zeigt die Verteilung der benötigten Zeit über die gesamte Testdauer.

Die benötigte Gesamtzeit liegt zwischen 2 Min. und 9 Sek. und 17 Min. und 32 Sek. Der Median für die durchschnittliche Gesamtzeit ergibt 8 Min. 9 Sek. In Tabelle 12 sind die Zeiten für die einzelnen Abschnitte nachzulesen. Tabelle 13 gibt die Zeiten für korrekte bzw. falsche Itembearbeitung wieder. Mittels Wilcoxon-Test werden die einzelnen Mittelwertsunterschiede untersucht. Sie unterscheiden sich alle signifikant voneinander. Es besteht ein signifikanter negativer Zusammenhang ($r = -0,411$) zwischen der benötigten Gesamtzeit bei gelösten und nicht gelösten Items.

Die Ergebnisse sind in Tabelle 14 nachzulesen.

Abbildung 20: Verteilung der benötigten Gesamtzeit

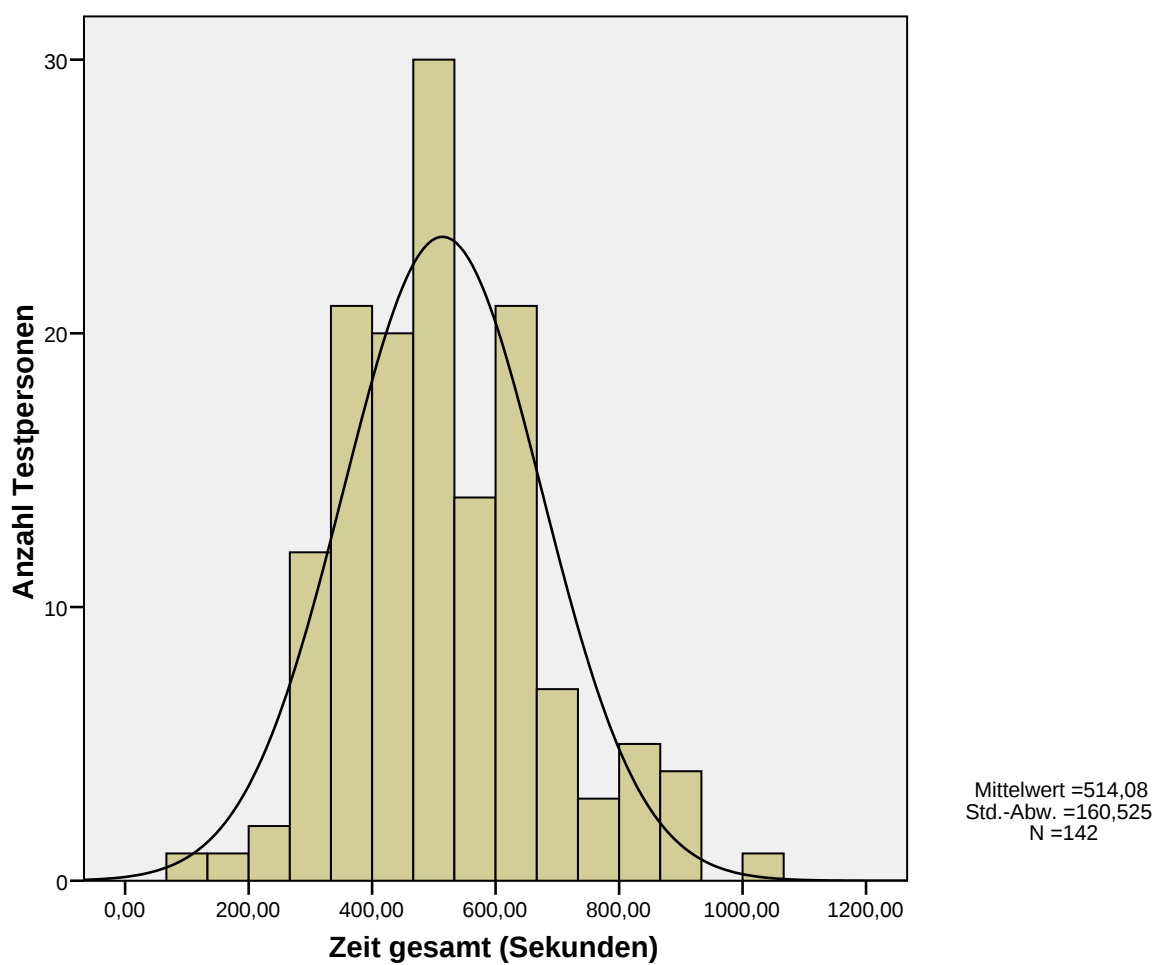


Tabelle 12: Zeitdauer - 1. Abschnitt, 2. Abschnitt und gesamt über alle Items

	Median	Standardabw.	1. Quartil	3. Quartil
Zeit gesamt	8 Min 9 Sek	2 Min 40 Sek	6 Min 39 Sek	8 Min 9 Sek
1. Zeitabschnitt	4 Min 0 Sek	1 Min 37 Sek	2 Min 58 Sek	5 Min 10 Sek
2. Zeitabschnitt	3 Min 52 Sek	1 Min 42 Sek	2 Min 58 Sek	5 Min 2 Sek

Tabelle 13: Benötigte Zeit für korrekte und falsche Itembeantwortung, n=142

	Item	Item richtig vs. falsch	Median	Standardabw.
1. und 2. Zeitabschnitt	1 bis 30	richtig	357 Sek	175 Sek
1. und 2. Zeitabschnitt	1 bis 30	falsch	129 Sek	79 Sek
1. Zeitabschnitt	1 bis 30	richtig	171 Sek	82 Sek
1. Zeitabschnitt	1 bis 30	falsch	55 Sek	38 Sek
2. Zeitabschnitt	1 bis 30	richtig	185 Sek	109 Sek
2. Zeitabschnitt	1 bis 30	falsch	52 Sek	29 Sek
1. Zeitabschnitt	1 bis 20	richtig	141 Sek	76 Sek
1. Zeitabschnitt	1 bis 20	falsch	49 Sek	35 Sek
1. Zeitabschnitt	21 bis 30	richtig	22 Sek	12 Sek
1. Zeitabschnitt	21 bis 30	falsch	7 Sek	7 Sek
2. Zeitabschnitt	1 bis 20	richtig	148 Sek	99 Sek
2. Zeitabschnitt	1 bis 20	falsch	38 Sek	26 Sek
2. Zeitabschnitt	21 bis 30	richtig	35 Sek	17 Sek
2. Zeitabschnitt	21 bis 30	falsch	11 Sek	8 Sek

Tabelle 14: Mittelwertsunterschiede und Korrelationen zwischen den einzelnen Zeitabschnitten, n= 142

	Item	a ¹		Mittelwertsunterschiede	Sig.	Korrelation	Sig.
Zeit gesamt	1 bis 30	f	r	Sig.	<0,001	-0,411	<0,001
1. Zeitabschnitt	1 bis 30	f	r	Sig.	<0,001	-0,192	0,022
2. Zeitabschnitt	1 bis 30	f	r	Sig.	<0,001	n. s.	0,883
1. Zeitabschnitt	1 bis 20	f	r	Sig.	<0,001	0,179	0,033
2. Zeitabschnitt	1 bis 20	f	r	Sig.	<0,001	n. s.	0,44
1. Zeitabschnitt	21 bis 30	f	r	Sig.	<0,001	n. s.	0,14
2. Zeitabschnitt	21 bis 30	f	r	Sig.	<0,001	n. s.	0,063

a¹...statistische Auswertung über falsche und richtige Items

3.3.2 Untersuchung der Mittelwerte in den beiden Teilstichproben

Zur Analyse der Mittelwertsunterschiede wurde jeweils jenes Verfahren angewendet, welches die statistischen Voraussetzungen erfüllt. In den Tabellen wird das angewendete Verfahren angegeben. Als Signifikanzniveau wurde $\alpha=0,05$ verwendet.

Die verschiedenen Mittelwerte der Teilstichproben Bundesheer und FH-Studenten sind in der Tabelle 15 nachzusehen. In den Abbildungen 21 bis 24 werden die verschiedenen Mittelwerte (gelöste Items, geschätzte Wahrscheinlichkeit, Item gelöst und geschätzte Wahrscheinlichkeit $p \geq 50$ und durchschnittlich benötigte Zeit) in Boxplots dargestellt.

Unterschiede in den Teilstichproben

In den Teilstichproben können keine signifikanten Mittelwertsunterschiede in Bezug auf die Anzahl der gelösten Items festgestellt werden. Weiters gib es keine signifikanten Ergebnisse bei der geschätzten Wahrscheinlichkeit, der Anzahl der absoluten Abweichungen und bei der Verwendung von sinnvollen und sinnlosen Hilfen.

Die Hilfe „Expertenrat einholen“ nehmen die FH-Studenten signifikant öfters in Anspruch als die Stellungspflichtigen. Sowohl bei der Zeit für die Entscheidung, als auch bei der Zeit für die Wahrscheinlichkeitsschätzung und somit bei der Gesamtzeit, kann ein weiterer signifikanter Unterschied zwischen den Teilstichproben festgestellt werden. Die Studenten benötigen mehr Zeit für die Bearbeitung.

Tabelle 15: Mittelwertsvergleiche zwischen Bundesheer (BH) und FH-Studenten (FH), signifikante Ergebnisse sind farbig unterlegt (U-Test = Mann-Whitney-Test)

Kennwert	Mittelwert		Signifikanz	stat. Verfahren
	BH	FH		
	n=110	n=32		
Anzahl richtig gelöste Items	21,37	21,56	0,806	t-Test
Anzahl falsch gelöster Items	6,51	6,78	0,67	t-Test
Mittelwert der geschätzten Wahrscheinlichkeit ¹	72,19	69,14	0,712	U-Test
Mittelwert der geschätzten W. der richtigen Items ²	68,67	67,89	0,391	t-Test
Anzahl mittlere „Treffer“ Intervall 1	1,86	2,13	0,063	U-Test
Anzahl mittlere „Treffer“ Intervall 2	3,5	4,41	0,156	U-Test
Anzahl mittlere „Treffer“ Intervall 3	5,97	6,63	0,535	U-Test
Konsistenz ohne Zeitdruck	6,46	6,75	0,18	U-Test
Konsistenz unter Zeitdruck	2,62	2,59	0,632	U-Test
Konsistenz gesamt	9,08	9,34	0,771	U-Test
Hilfe "Expertenrat einholen"	2,98	5	0,015	U-Test
Hilfe "Computersimulation	2,05	2,59	0,274	U-Test
Hilfe "Entscheidung der letzten drei Personen"	1,73	1,41	0,91	U-Test
Hilfe "Farbe nächste Kugel selber Behälter"	1,09	0,81	0,297	U-Test
Hilfe "Computer soll entscheiden"	1,23	0,44	0,128	U-Test
Hilfe "Farbe vorherige Kugel selber Behälter"	0,66	0,53	0,934	U-Test
Sinnvolle Hilfen	6,13	8,41	0,085	U-Test
Sinnlose Hilfen	3,62	2,38	0,337	U-Test
Hilfen gesamt	9,75	10,78	0,482	U-Test
Hilfekategorie "Person"	4,71	6,41	0,119	U-Test
Hilfekategorie "Stimulus"	1,75	1,34	0,671	U-Test
Hilfekategorie "Situation"	3,28	3,03	0,813	U-Test
Anzahl richtige Items ohne Zeitdruck	13,3	15,06	0,089	U-Test
Anzahl richtige Items unter Zeitdruck	6,28	6,28	0,49	U-Test
Zeit für Entscheidung inkl. Hilfebeanspruchung ³	165,42	226,81	0,001	U-Test
Zeit für Schätzung der Wahrscheinlichkeit ³	189,3	243,06	0,011	U-Test
Zeit gesamt (richtige und falsche Items)	486,56	608,66	0,001	U-Test

¹ Mittelwert der geschätzten Wahrscheinlichkeit ohne Berücksichtigung der Behälterwahl

² Stichprobengröße: BH=106, FH=32

³ Nur richtig gelöste Items werden analysiert

Abbildung 21: Mittelwerte der gelösten Items

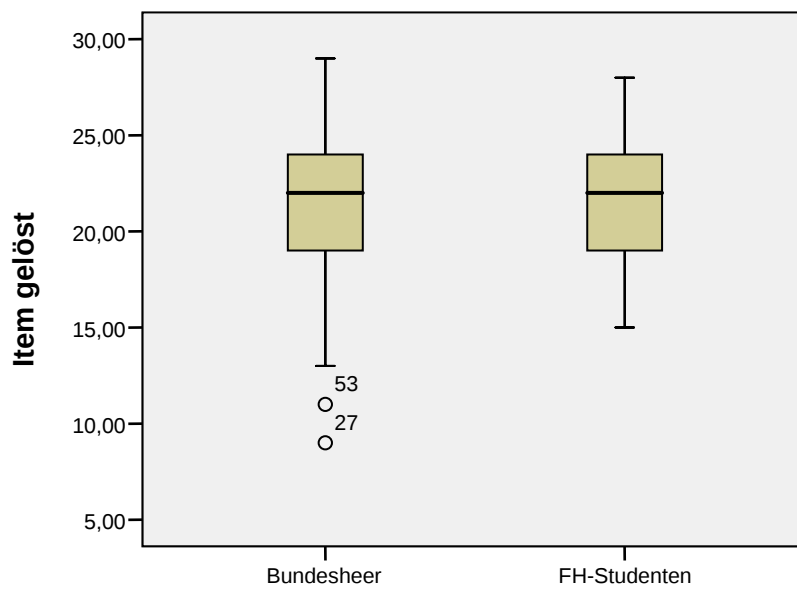


Abbildung 22: Mittelwerte der geschätzten Wahrscheinlichkeiten

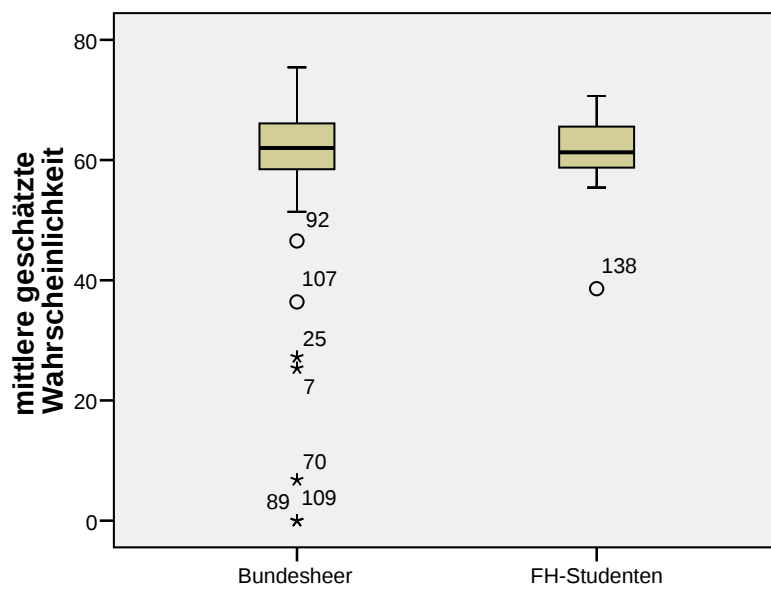


Abbildung 23: Mittelwerte der geschätzten
Wahrscheinlichkeiten bei instruktionskonformer Bearbeitung

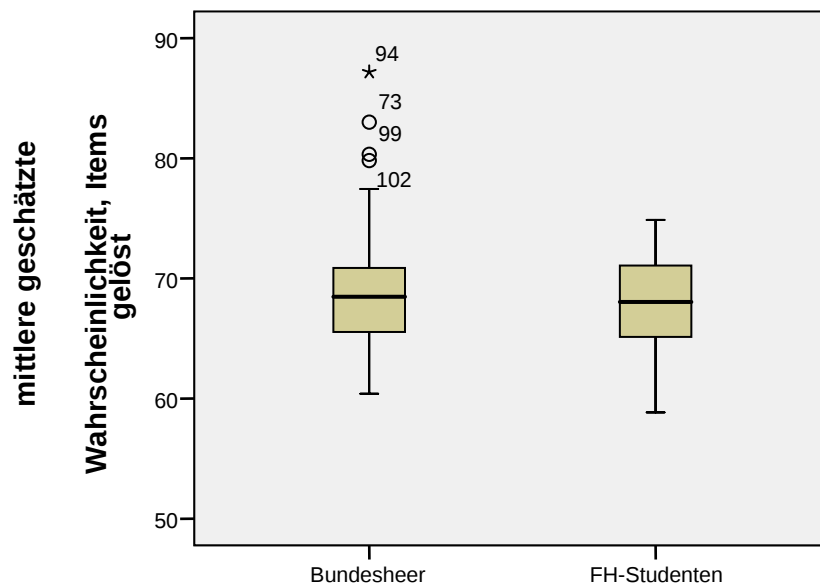
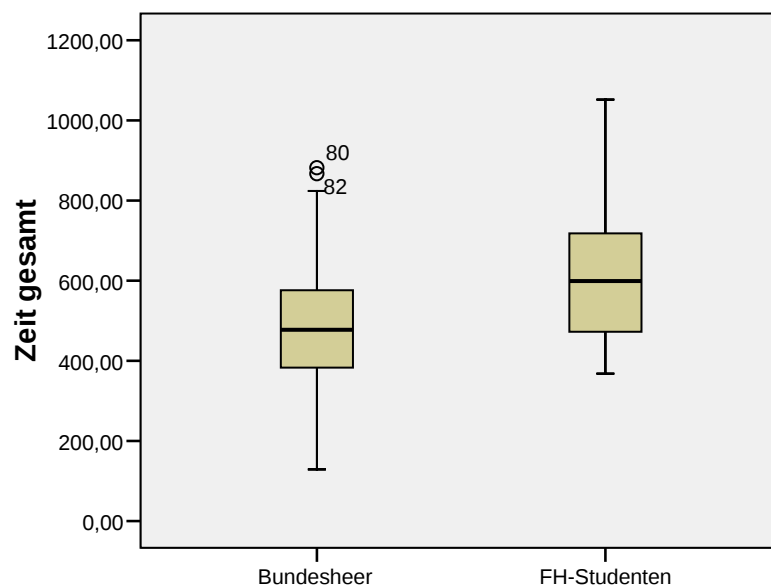


Abbildung 24: Benötigte Zeit in Sekunden



Unterschiede in den Bildungsgruppen

Die Untersuchung der Mittelwertsunterschiede in Bezug auf die Verwendung von sinnlosen Hilfen ergibt ein signifikantes Ergebnis bei den Bildungsgruppen. Personen der Bildungsgruppe 1 beanspruchen mehr sinnlose Hilfen. Eine genaue Aussage über die Bildungsgruppe 0 kann nicht erfolgen, da in dieser alle Personen zusammengefasst werden, die in keine der drei anderen Gruppen zugeordnet werden können. Weitere signifikante Unterschiede können bei der benötigten Zeit festgestellt werden. Dies ist insofern nicht überraschend, als 94 % der Bildungsgruppe 3 aus FH-Studenten bestehen. Für die Itembearbeitung beanspruchen FH-Studenten signifikant mehr Zeit.

In der Tabelle 16 erfolgt eine Auflistung der statistischen Auswertung der Mittelwerte. In den Abbildungen 25 und 26 werden die Mittelwerte zwischen den Bildungsgruppen bezüglich Rohscore und geschätzter Wahrscheinlichkeit ($p \geq 50$) in grafischen Form dargestellt.

Tabelle 16: Mittelwertsvergleiche zwischen den Bildungsgruppen, signifikante Ergebnisse sind farblich unterlegt (K-W-Test = Kruskal-Wallis-Test)

Kennwert	Mittelwert				Sig. $\alpha=0.05$	stat. Verfahren
	BG 0	BG 1	BG 2	BG 3		
	n=38	n=17	n=53	n=34		
Anzahl richtig gelöste Items	21,11	20,24	21,75	21,82	0,672	K-W-Test
Anzahl falsch gelöster Items	6	8,06	6,47	6,62	0,477	K-W-Test
Mittelwert der geschätzten Wahrscheinlichkeit ¹	63,17	61,01	58,73	59,8	0,648	K-W-Test
Mittelwert der geschätzten W. der richtigen Items ²	69,11	70,12	67,9	67,91	0,422	ANOVA
Anzahl mittlere "Treffer" Intervall 1	1,87	1,24	1,74	2,03	0,146	K-W-Test
Anzahl mittlere "Treffer" Intervall 2	3,37	3,47	3,64	4,29	0,673	K-W-Test
Anzahl mittlere "Treffer" Intervall 3	5,95	5,65	6,21	6,41	0,948	K-W-Test
Konsistenz ohne Zeitdruck	6,26	6,59	6,55	6,76	0,123	K-W-Test
Konsistenz unter Zeitdruck	2,39	2,76	2,72	2,62	0,095	K-W-Test
Konsistenz gesamt	8,66	9,35	9,26	9,38	0,063	K-W-Test
Hilfe "Expertenrat einholen"	4,13	2,18	2,49	4,76	0,078	K-W-Test
Hilfe "Computersimulation"	2,42	2,12	1,85	2,44	0,474	K-W-Test
Hilfe "Entscheidung der letzten drei Personen"	1,74	2,29	1,6	1,32	0,527	K-W-Test
Hilfe "Farbe nächste Kugel selber Behälter"	1,89	0,88	0,6	0,79	0,086	K-W-Test
Hilfe "Computer soll entscheiden"	1,71	0,94	1,02	0,41	0,088	K-W-Test
Hilfe "Farbe vorherige Kugel selber Behälter"	1,08	1	0,28	0,5	0,068	K-W-Test
Sinnvolle Hilfen	8,45	5,18	4,94	8	0,117	K-W-Test
Sinnlose Hilfen	4,53	4,24	2,91	2,24	0,041	K-W-Test
Hilfen gesamt	12,97	9,41	7,85	10,24	0,173	K-W-Test
Hilfekategorie "Person"	5,87	4,47	4,09	6,09	0,328	K-W-Test
Hilfekategorie "Stimulus"	2,97	1,88	0,89	1,29	0,048	K-W-Test
Hilfekategorie "Situation"	4,13	3,06	2,87	2,85	0,355	K-W-Test
Anzahl richtige Items ohne Zeitdruck	14,26	12,59	13,02	14,68	0,457	K-W-Test
Anzahl richtige Items unter Zeitdruck	6,42	6	6,32	6,18	0,821	K-W-Test
Zeit für Entscheidung inkl. Hilfebeanspruchung ³	182,53	164	152,03	225,65	<0,001	K-W-Test
Zeit für Schätzung der Wahrscheinlichkeit ³	203,29	148,76	195	235,65	0,036	K-W-Test
Zeit gesamt (richtige und falsche Items)	524,89	455,12	472,08	596,94	0,02	K-W-Test

¹ Mittelwert der geschätzten Wahrscheinlichkeit ohne Berücksichtigung der Behälterwahl

² Stichprobengröße: BH=106, FH=32

³ Nur richtig gelöste Items werden analysiert

Abbildung 25: Mittelwerte der gelösten Items zwischen den Bildungsgruppen

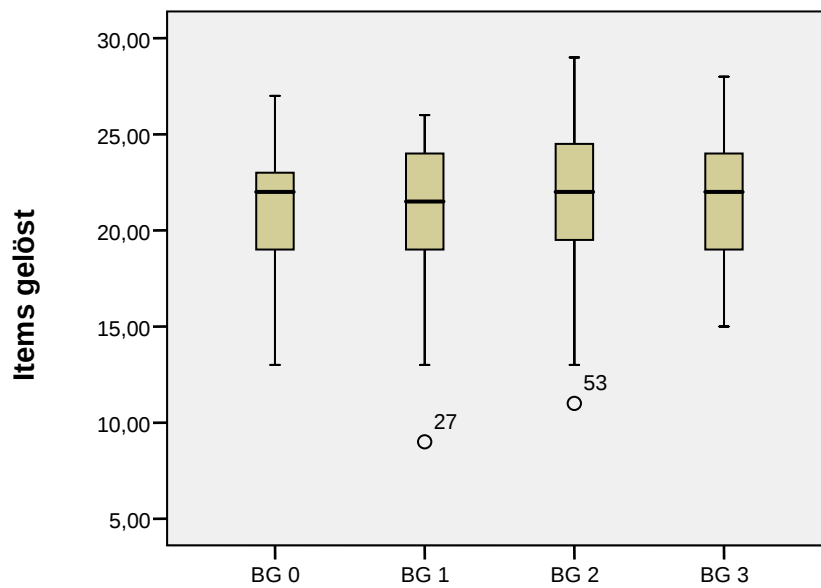
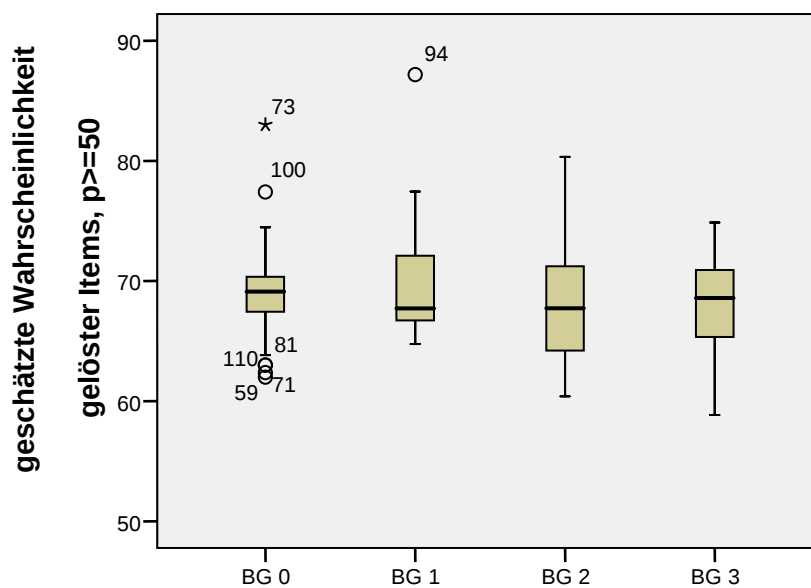


Abbildung 26: Mittelwerte der richtig gelösten Items bei instruktionskonformer Bearbeitung zwischen den Bildungsgruppen



Geschlechtsunterschiede

Bei der Auswertung über das Geschlecht werden nur die FH-Studenten berücksichtigt, zumal die Stellungspflichtigen ausschließlich aus männlichen Probanden bestehen.

Es sind keine signifikanten Mittelwertsunterschiede (siehe Tabelle 17) zu beobachten. Abbildung 27 zeigt den Rohscore von Männern und Frauen. Auf Grund der geringen Stichprobe ($n=32$) muss die Interpretation der Ergebnisse mit Bedacht erfolgen.

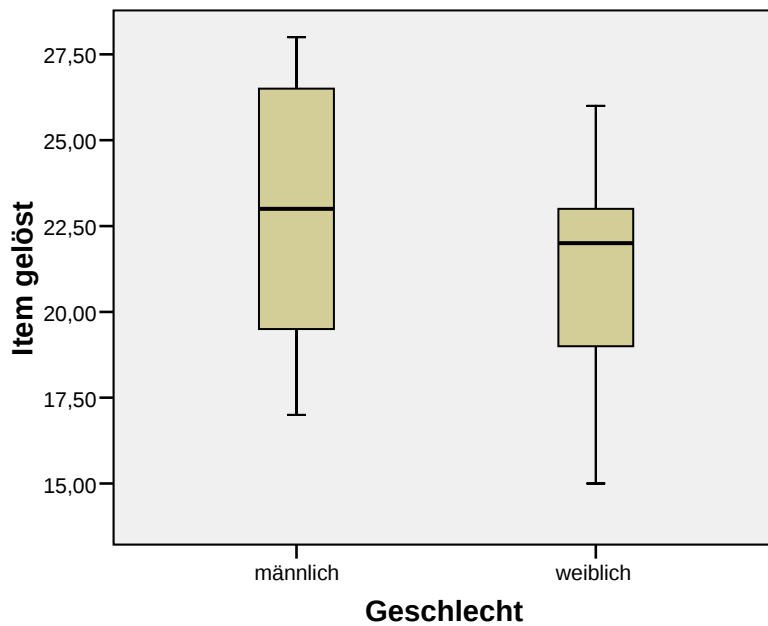
Tabelle 17: Mittelwertsvergleich zwischen männlich und weiblich

Kennwert	Mittelwert		Sig.	stat. Verfahren
	männlich	weiblich		
	n=11	n=21	$\alpha=0.05$	
Anzahl richtig gelöste Items	22,64	21	0,201	t-Test
Anzahl falsch gelöster Items	6,1	7,14	0,351	t-Test
Mittelwert der geschätzten Wahrscheinlichkeit ¹	63,12	60,67	0,272	t-Test
Mittelwert der geschätzten W. der richtigen Items	68,82	67,4	0,371	t-Test
Anzahl mittlere „Treffer“ Intervall 1	2,45	1,95	0,354	t-Test
Anzahl mittlere „Treffer“ Intervall 2	4,91	4,14	0,511	t-Test
Anzahl mittlere „Treffer“ Intervall 3	7,27	6,29	0,584	U-Test
Konsistenz ohne Zeitdruck	6,55	6,86	0,289	U-Test
Konsistenz unter Zeitdruck	2,64	2,57	0,696	U-Test
Konsistenz gesamt	9,18	9,43	0,639	U-Test
Hilfe "Expertenrat einholen"	3,18	5,95	0,177	t-Test
Hilfe "Computersimulation"	1,82	3	0,349	t-Test
Hilfe "Entscheidung der letzten drei Personen"	0,64	1,81	0,104	U-Test
Hilfe "Farbe nächste Kugel selber Behälter"	0,18	1,14	0,506	U-Test
Hilfe "Computer soll entscheiden"	0,45	0,43	0,907	U-Test
Hilfe "Farbe vorherige Kugel selber Behälter"	0,18	0,71	0,434	U-Test
Sinnvolle Hilfen	5,18	10,1	0,13	t-Test
Sinnlose Hilfen	1,27	2,95	0,88	U-Test
Hilfen gesamt	6,45	13,05	0,103	t-Test
Hilfekategorie "Person"	3,82	7,76	0,123	t-Test
Hilfekategorie "Stimulus"	0,36	1,86	0,531	U-Test
Hilfekategorie "Situation"	2,27	3,43	0,379	t-Test
Anzahl richtige Items ohne Zeitdruck	15,28	14,95	0,716	t-Test
Anzahl richtige Items unter Zeitdruck	7,09	5,86	0,076	t-Test
Zeit für Entscheidung inkl. Hilfebeanspruchung ²	253,73	212,71	0,531	U-Test
Zeit für Schätzung der Wahrscheinlichkeit ²	305,27	210,48	0,155	U-Test
Zeit gesamt (richtige und falsche Items)	689,55	566,29	0,104	U-Test

¹ Mittelwert der geschätzten Wahrscheinlichkeit ohne Berücksichtigung der Behälterwahl

² Nur richtig gelöste Items werden analysiert

Abbildung 27: Mittelwerte von Männern und Frauen



3.3.3 EUcO 2.0 Ergebnisse im Vergleich zur alten Version

Das Ziel der EUcO 2.0-Neukonstruktion war die Bereinigung der in Kapitel 2.4.1 angeführten Kritikpunkte. Die Frage, ob und inwiefern dies gelungen ist, soll nun durch den Vergleich der Ergebnisse der beiden Versionen beantwortet werden. In Tabelle 18 werden die Resultate zwischen EUcO 2.0 und EUcO „alt“ gegenübergestellt.

Die durchschnittliche Anzahl der gelösten Items ist auf 21,42 gestiegen. Abweichend von der „alten“ Version können keine signifikanten Mittelwertsunterschiede zwischen den Teilstichproben und den Bildungsgruppen festgestellt werden.

Vergleicht man die geschätzten Wahrscheinlichkeiten, so fallen beim EUcO 2.0 die Schätzungen um circa 7 % geringer aus. Hingegen liegen die

durchschnittlichen Wahrscheinlichkeitsschätzer der richtig beantworteten Items nur etwas darunter, ungefähr 2,5 Prozentpunkte.

Die von den Testpersonen in Anspruch genommenen Hilfen sind von 22 (alte Version) auf etwa 10 Hilfen gesunken. An dieser Stelle ist jedoch anzuführen, dass nach der ersten Aufgaben nur fünf Hilfen angeboten werden und maximal zwei Hilfen pro Item beansprucht werden können. Stellt man jene Hilfen, die in beiden Verfahren zur Auswahl stehen, einander gegenüber, ergeben sich bei den Testpersonen annähernd gleiche Präferenzen.

Die durchschnittliche Bearbeitungszeit ist beim EUcO 2.0 um ungefähr 2 ½ Minuten kürzer, die Testpersonen nehmen indessen mehr Zeit für die Schätzung in Anspruch.

Stellt man die Ergebnisse der Entscheidungskonsistenz gegenüber, werden nur mehr vier von zehn Itempaare, auf Basis der Lösungshäufigkeit, ungleich schwierig erlebt und zwei als sehr leicht empfunden. Insgesamt ist eine Erhöhung der Konsistenz zu beobachten.

Die untersuchten Korrelationen bezüglich sinnvoller vs. sinnloser Hilfen, gelöste Items bzw. Wahrscheinlichkeitsschätzung mit vs. ohne Zeitdruck sind geringer ausgefallen.

Tabelle 18: Ergebnisvergleich der beiden Testversionen

Kennwert	Mittelwert	
	EUcO 2.0	EUcO "alt"
Anzahl richtig gelöster Items	21,42	17,99
Mittelwert der geschätzten W.	63,03	70,3
Mittelwert der geschätzten W. der richtigen Items	68,49	70,7
Hilfen gesamt	9,98	22
	Median (Sekunden)	
Zeit gesamt	489,5	639
	Korrelation	
Zwischen sinnvollen und sinnlosen Hilfen	0,598	0,757
Zwischen richtig gelöster Items mit und ohne Zeitdruck	0,393	0,631
Zwischen Wahrscheinlichkeitsschätzung mit und ohne Zeitdruck	0,65	0,882

3.3.4 Beantwortung der Fragestellungen und Hypothesenüberprüfung

Die Ergebnisse des EUcO 2.0 lassen darauf schließen, dass es zu einer besseren Verständlichkeit auf Seite der Testpersonen gekommen ist. Einerseits aufgrund der einfachen und klaren Instruktionsgestaltung, die zu keinen auffälligen Problemen während der Beschäftigung in der Instruktionsphase führte. Andererseits weil die Instruktion in der Testphase kaum abgerufen wird.

Durch die Vorgabe von Beispielaufgaben wird ein minimales Basiswissen über das Wahrscheinlichkeitsprinzip aufgebaut.

Ein weiteres Indiz dafür, dass die Verständlichkeit verbessert wurde, ist im Vergleich der Testkennwerte zu finden. Der Mittelwert der gelösten Items fällt höher aus und jener über die durchschnittlich benötigte Zeit niedriger.

Ein Überschätzen der Wahrscheinlichkeit wird auch in der neu konstruierten Version beobachtet; allerdings fällt sie geringer aus. Analysiert man nur gelöste

Items, erfolgen die Schätzungen genauer. Es kommt aber ebenfalls zu Überschätzungen.

Durch die statistische Auswertung über die Inanspruchnahme von Hilfen (Hilfekategorien) kann kein neuer Testkennwert abgeleitet werden.

Bezüglich der Hypothesenprüfung kann folgendes festgestellt werden:

- a.) Die Testkennwerte in den verschiedenen Gruppen unterscheiden sich nur im Hinblick auf die Zeit. Vereinzelt in der Inanspruchnahme von Hilfen (siehe Kapitel 3.3.2).
- b.) Es gibt Unterschiede in den Testkennwerten zwischen EUcO 2.0 und EUcO (siehe Kapitel 3.3.3).
- c.) Die Personen unterscheiden sich nicht in der Wahl der verwendeten Hilfen.
- d.) Der Zusammenhang zwischen den Hilfekategorien ist positiv. Auf Grund des positiven Zusammenhanges zwischen den Hilfskategorien kann der Attributionsstil einer Person nicht festgestellt werden.

3.4 Ausblick

In der Neukonstruktion des EUcO 2.0 wird den Kritikpunkten angemessen begegnet:

1. Die Instruktion wurde auf das nötigste verkürzt und der Text durch grafische Elemente aufgefrischt.
2. Die Testperson lernt den Umgang mit Wahrscheinlichkeiten durch die Vorgabe von Beispielaufgaben kennen.
3. Durch das „Lottoprinzip“ wird eine analoge Verbindung zu den eigentlich unlogischen Items hergestellt.

Natürlich sind nach der Überarbeitung Überlegungen getätigt worden, welche Aufgaben gemacht werden müssen, damit der Test einer sinnvollen Anwendung in der Praxis gerecht wird. Die Animation der Beispielaufgaben müsste konsequenterweise ebenso auf die Items erfolgen. Weiters sollte noch intensiver auf die Wahrscheinlichkeitsschätzung eingegangen werden. Entweder durch die Vorgabe von zusätzlichen Beispielaufgaben oder durch eine stärkere Beschäftigung auf Testpersonenseite während der Bearbeitung der Beispielaufgaben.

In der aktuellen Version besteht für die Testperson die Möglichkeit, ihre Entscheidung rückgängig zu machen. Tatsächlich sind in der Realität Entscheidungsänderungen nicht so leicht zu bewerkstelligen, zumal sie zu unterschiedliche Konsequenzen führen. Die vom Testautor geforderte Umsetzung einer irreversiblen Entscheidung wird in der Neukonstruktion nicht realisiert. Eine Änderung der Entscheidung fließt auch deshalb nicht in die Testauswertung ein; sie wird erst gar nicht erhoben.

Die Analyse des EUcO 2.0 erfolgte an Hand von einer eigentlich kleinen Stichprobe. Um valide Aussagen zu generieren sollte er an einer größeren Stichprobe mit entsprechenden Bildungsgruppen und einer repräsentativen Geschlechtsverteilung untersucht werden.

Weiters müssen noch Untersuchungen zu den Gütekriterien getätigt werden. Der EUcO 2.0 muss ebenfalls an einer Extremgruppe validiert werden.

Einige Testpersonen geben Schätzungen ab, welche unter 50 % liegen, obwohl bei der Instruktion auf diesen Widerspruch hingewiesen wird. Im wissenschaftlichen Diskurs sollte erörtert werden, wie diesem Phänomen begegnet werden kann. Die Testpersonen ständig darauf hinzuweisen, wäre nur eine Möglichkeit.

Bei einer Neuprogrammierung sollen Instruktion und Test zusammengeführt und auf der gesamten Bildschirmoberfläche dargestellt werden. Dadurch wird eine Ablenkung des auf dem Desktop befindlichen Materials verhindert.

4. Zusammenfassung

Das Thema dieser Arbeit war die Verbesserung und Neukonstruktion des objektiven Persönlichkeitstests zum Entscheidungsverhalten EUcO (Entscheiden unter Unsicherheit computerisierter objektiver Persönlichkeitstest). Es sollten einerseits den Kritikpunkten begegnet, andererseits die in der Praxis aufgetretenen Schwächen kompensiert werden. Die neue Version erhält die Bezeichnung EUcO 2.0.

Kritik wurde unter anderem in Bezug auf die lange und unverständliche Instruktion geäußert sowie auf das Fehlen von Beispielaufgaben. Ebenso verzerrt die hohe Anzahl an Hilfen bzw. deren optischen Dominanz die Wahrnehmung der Testperson.

Der EUcO 2.0 besteht aus insgesamt 30 Aufgaben, wovon die letzten 10 unter Zeitdruck vorgegeben werden. Die Aufgaben basieren auf dem Bayes-Theorem, einem aus der Wahrscheinlichkeitstheorie stammenden Lehrsatz, der eingetretene Ereignisse (a- posteriori-Wahrscheinlichkeit) berücksichtigt. Entscheidungen bzw. deren Wahrscheinlichkeiten können nach diesem Theorem entsprechend berechnet werden. Eine Person muss eine zufällig gezogene Kugel (schwarz oder weiß) einem von zwei Behältern zuweisen. Die Behälter enthalten eine unterschiedliche Anzahl von schwarzen und weißen Kugeln. In der Realität existieren bei der Entscheidungsfindung eine Vielzahl an Handlungsmöglichkeiten, meistens erfolgt eine Reduktion, so dass nur mehr zwei Möglichkeiten übrig bleiben. Diese Handlungsalternativen werden durch die beiden Behälter repräsentiert. Die Umsetzung der Unsicherheit erfolgt durch die zufällige Ziehung der Kugel. Der Eigenschaft eines Objektiven Persönlichkeitstest wird dadurch Rechnung getragen, dass der Testperson mitgeteilt wird, sie müsse eine gezogene Kugel zuordnen und anschließend die Wahrscheinlichkeiten dafür einschätzen. Die wahre Testintention, Messung des individuellen Entscheidungsverhaltens, kann auf Grund des Aufgabenmaterials nicht erahnt werden. Sie wird verschwiegen. Für die Zuordnung der Kugel stehen

verschiedene Hilfen z.B. „Expertenrat einholen“ bereit. Dadurch erfolgt eine reale Abbildung von Entscheidungssituationen, bei der gewisse (sinnvolle und sinnlose) Hinweise vorhanden sein können.

Bei der Neukonstruktion wurden unter Berücksichtigung wissenschaftlicher Erkenntnisse die grafischen Möglichkeiten des Computers besser ausgenützt. Die Instruktion wurde verständlicher gestaltet, und Beispielaufgaben wurden für ein besseres „Wahrscheinlichkeitsgefühl“ konstruiert. Die Anzahl und Dominanz der Hilfen wurde reduziert und auf zwei pro Aufgabe begrenzt.

An dieser Studie nahmen 110 Stellungspflichtige Personen des österreichischen Bundesheers und 32 FH-Studenten des Zweiges „Soziale Arbeit“ teil.

Der EUcO 2.0 weist folgende Testkennwerte aus: Rohscore, Entscheidungsgüte, Entscheidungskonsistenz, Verwendung von Hilfen, Einfluss von Zeitdruck und die Gesamtzeit.

Ein neuer Testkennwert „Attributionsstil einer Person“ konnte durch die statistische Analyse nicht gefunden werden. Der Zusammenhang zwischen den Hilfskategorien ist positiv; nicht wie in der Alternativhypothese gefordert, negativ.

Eine Verbesserung des Testverständnisses kann auf Grund der höheren Mittelwerte bei den gelösten Aufgaben angenommen werden.

In den einzelnen Teilstichproben konnten keine Mittelwertsunterschiede bei den gelösten Aufgaben festgestellt werden.

Bei der neu konstruierten Version wurde die Wahrscheinlichkeit wiederum überschätzt, die Überschätzungen fielen aber geringer aus. Weiters gab es keine Unterschiede in Bezug auf die Anzahl und Präferenz der Hilfen zwischen den Stichproben. Signifikante Mittelwertsunterschiede sind bei der benötigten Zeit zu beobachten. Je gebildeter eine Person ist, desto mehr Zeit beansprucht sie für die Entscheidung und für die Schätzung der Wahrscheinlichkeit.

Die Neukonstruktion des EUcO 2.0 darf auf Grund der vorliegenden Ergebnisse als erfolgreich angesehen werden, jedoch müssen noch weitere Untersuchungen im Hinblick auf die Gütekriterien vorgenommen werden.

Literaturverzeichnis

A

Amelang, M. & Bartussek, D. (2001). *Differentielle Psychologie und Persönlichkeitsforschung*. Stuttgart: Kohlhammer.

Amthauer, R., Brocke, B., Liepmann, D. & Beauducel, A. (2001). *Intelligenz-Struktur-Test 2000 R (I-S-T 2000 R)*. Göttingen: Hogrefe.

Anastasi, A. (1954). *Psychological Testing*. New York: The Macmillan Company.

Anderson, J. R. (2001). *Kognitive Psychologie*. Heidelberg: Spektrum Akademischer Verlag.

Angleitner A. & Wiggins J. S. (1986). *Personality Assessment via Questionnaires*. Heidelberg: Springer Verlag.

B

Bamberg, G. & Coenenberg, A.G. (2006). *Betriebswirtschaftliche Entscheidungslehre*. München: Verlag Vahlen.

Becker, P. (1989). *Trierer Persönlichkeitsfragebogen (TRF)*. Göttingen: Hogrefe.

Bellmann, M. & Sarges, W. (1990). Verständlichkeitsverbesserung der Testinstruktion und ihre Auswirkung auf die Leistung in einem Schulintelligenztest. *Zeitschrift für differenzielle und diagnostische Psychologie*, 1, 137-148.

Berg, C. C. (1972). *Adaptive Entscheidungen unter Ungewissheit*. Unveröffentlichte Dissertation, Universität Mannheim.

Bitz, M. (1981). *Entscheidungstheorie*. München: Verlag Vahlen.

Booth, J. (1994). Computerdiagnostik. In R. S. Jäger & F. Petermann (Hrsg.), *Psychologische Diagnostik* (S. 186-197). Weinheim: Psychologie Verlags Union.

Brengelmann, J.C. & Brengelmann, L. (1960). Deutsche Validierung von Fragebogen der Extraversion, neurotischen Tendenz und Rigidität. *Zeitschrift für experimentelle und angewandte Psychologie*, 7, 291-331.

Bronner, R. (1989). *Planung und Entscheidung*. München: R. Oldenbourg Verlag

Bukasa, B., Kisser, R. & Wenninger, U. (1990). Computergestützte Leistungsdiagnostik bei verkehrspsychologischen Eignungsuntersuchungen. *Diagnostica*, 36, 148-165.

Bühlman, H. (1969). *Einführung in die Theorie und Praxis der Entscheidung bei Unsicherheit*. Berlin: Springer.

Bühner, M. (2004). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson Studium.

C

Camerer, C. (1992). *Recent tests of generalizations of expected utility theory*. In W. Edwards (Hrsg.), *Utility theories: Measurements and applications* (S. 207-251). Boston: Kluwer.

Campbell, D. T. & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.

Cattell, R. B. (1973). *Die empirische Erforschung der Persönlichkeit*. Weinheim: Beltz Verlag.

D

Debus, G. (1977). Gefühle. In T. Herrmann, P. R. Hofstätter, H. P. Huber & F. E. Weinert (Hrsg.), *Handbuch psychologischer Grundbegriffe* (S.156-168). München: Verlag Kösel.

E

Edwards, W. (1954). The theory of decision making. *Psychological Bulletin*, 51, 380-417.

Eisenführ, F. & Weber, M. (2003). *Rationales Entscheiden*. Berlin: Springer Verlag.

F

Feger, H. (1968). *Untersuchung zur Bedeutsamkeit von Entscheidungssituationen- Ein experimenteller Beitrag*. Unveröffentlichte Dissertation, Universität Bonn.

Fischer, K. & Jungermann, H. (1996). Rarely occurring, headaches and rarely occurring blindness: Is rarely = rarely? The meaning of verbal frequentistic labels in specific medical context. *Journal of Behavioral Decision Making*, 9, 153-172.

Fisseni, H. (1997). *Lehrbuch der psychologischen Diagnostik*. Göttingen: Hogrefe.

Fisseni, H. (2004). *Lehrbuch der psychologischen Diagnostik*. Göttingen: Hogrefe.

Friedrichs, J. (1997). *Methoden empirischer Sozialforschung*. Opladen: Westdeutscher Verlag.

G

Gertzen, H. (1990). *Entscheidungen bei sequenzierter Informationsdarbietung am Bildschirm*. Münster: Waxmann.

Gibbons, R. (1992). *A Primer in Game Theorie*. Harlow: Verlag Finacial Times.

Gigerenzer, Gerd. (1993). Die Repräsentation von Information und ihre Auswirkung auf statistisches Denken. In W. Hell, K. Fiedler & G. Gigerenzer (Hrsg.), *Kognitive Täuschungen* (S. 99-127). Heidelberg: Spektrum Verlag.

Gigerenzer, G. & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, 102, 684-704.

Gigerenzer, G., Szijtink, Z., Porter, T., Daston L., Beatty, J. & Krüger, L. (1999). *Das Reich des Zufalls*. Heidelberg: Spektrum Verlag.

Gigerenzer, G. & Todd, P. M. (2001). *Simple heuristics that make us smart*. Oxford: Oxford University Press.

Gigerenzer, G. (2007). *Bauch-Entscheidungen*. München: Bertelsmann Verlag.

Gittler, G. (1989). *Drei-dimensionaler Würfeltest (3DW)*. Weinheim: Beltz.

Gluck, M. A. & Bower, G. H. (1988). Form conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General*, 8, 37-50.

Goldsmith, R. W. (1967). *Persönlichkeitsspezifische Komponenten des Entscheidungsverhaltens*. Unveröffentlichte Dissertation, Universität Berlin.

Guthke, J. & Sprung, L. (1990). In J. Guthke, H. R. Böttcher & L. Sprung (Hrsg.), *Psychodiagnostik, Bd. 1*. Berlin: Deutscher Verlag der Wissenschaft.

Grubitzsch, S. (1991). *Testtheorie und Testpraxis*. Reinbek bei Hamburg: Rowohlt Verlag.

H

Hartley, J., Lindsey, D. & Burnhill, P. (1977). Alternatives in the typographic design of questionnaires. *Journal of Occopational Psychology*, 50, 209-304.

Häcker, H. O. & Stapf, K. (2004). *Dorsch Psychologisches Wörterbuch*. Bern: Verlag Hans Huber.

Häcker, H., Schmidt, L. R., Schwenkmezger, P. & Utz, H. E. (1975). *Objektive Testbatterie (OA-TB 75)*. Weinheim: Beltz.

Hergovich, A. & Hörndler, H. (1994). *Gestaltwahrnehmungstest - Ein computerbasiertes Verfahren zur Messung der Feldartikulation*. Frankfurt/M.: Swets Test Services.

Hergovich, A. (2006). In T. M. Ortner, R. T. Proyer & K. D. Kubinger (Hrsg.), *Theorie und Praxis Objektiver Persönlichkeitstests* (S. 124-132). Bern: Huber.

Herkner, W. (1980). *Attribution-Psychologie der Kausalität*. Bern: Verlag Huber.

Herkner, W. (2001). *Lehrbuch Sozialpsychologie*. Bern: Verlag Huber.

Hoffrage, U. (1993). Die Illusion der Sicherheit bei Entscheidungen unter Unsicherheit. In W. Hell, K. Fiedler & G. Gigerenzer (Hrsg.) *Kognitive Täuschungen* (S. 73-97). Heidelberg: Spektrum Verlag.

Holler, M. J. & Illing, G. (1996). *Einführung in die Spieltheorie*. Berlin: Springer.

Horn, R. (2003). *Probleme des interkulturellen Einsatzes von Fragebogen. Vergleichende Analyse von Benchmarks anhand von Daten aus den USA, Großbritannien, Frankreich und Deutschland – eine empirische Analyse zum Bias-Problem*. Landau: Verlag Empirische Pädagogik.

Hörmann, H. (1960). *Konflikt und Entscheidung*. Göttingen: Verlag Hogrefe.

Huber, O. (1982). *Entscheiden als Problemlösen*. Bern: Verlag Huber.

J

Jäger, R. S. & Petermann, F. (1992). *Psychologische Diagnostik*. Weinheim: Psychologie Verlags Union.

Jäger, R. S. & Krieger, W. (1994). Zukunftsperspektiven der computergestützten Diagnostik, dargestellt am Beispiel der treatmentorientierten Diagnostik. *Diagnostica*, 40, 217-243.

Jungermann, H. (1997). Reasons for uncertainty: From frequencies to stories. *Psychologische Beiträge*, 39, 126-139.

Jungermann, H., & Fischer, K. (2005). Using expertise and experience for giving and taking advice. In T. Betsch & S. Haberstroh (Hrsg.), *The routines of decision making* (S. 157-173). Mahwah, NJ.: Lawrence Erlbaum Associates.

Jungermann, H. (2005). Entscheiden: In D. Frey, L. Rosenstiel, C. G. Hoyos (Hrsg.), *Wirtschaftspsychologie* (S. 72-78). Basel: Beltz Verlag.

Jungermann, H., Pfister, H. & Fischer, K. (2010). *Die Psychologie der Entscheidung*. Heidelberg: Spektrum Verlag.

K

Kahneman, D. & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237-251.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263-291.

Kahneman, D., Slovic, P. & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, MA: University Press.

Kahneman, D. & Snell, J. (1992). Predicting a changing taste: Do people know what they will like? *Journal of Behavioral Decision Making*, 5, 187-200.

Karmasin, F. & Karmasin, H. (1977). *Einführung in Methoden und Probleme der Umfrageforschung*. Wien: Verlag Böhlau.

Kelley, H. H. (1967). Attribution theory in social psychology. In Levine, D. (Hrsg.), *Nebraska symposium on motivation*, Bd. 15. Lincoln: University of Nebraska Press.

Kersting, M. (2003). Augenscheinvalidität. In K. D. Kubinger & R. S. Jäger (Hrsg.) *Schlüsselbegriffe der Psychologischen Diagnostik* (S. 54-55). Weinheim: Beltz.

Kirchler, E. (2005). *Arbeits- und Organisationspsychologie*. Wien: Facultas.

Klieme, E. & Stumpf, H. (1990). Computereinsatz in der pädagogisch-psychologischen Diagnostik. In K. Ingenkamp & R. S. Jäger (Hrsg.), *Tests und Trends* 8 (S. 13-63). Weinheim: Beltz.

Klinck, D. (2002). Computerdiagnostik. In F. Peterman & M. Eid (Hrsg.), *Handbuch der psychologischen Diagnostik* (S. 209-232). Göttingen: Verlag Hogrefe.

Koriat, A., Lichtenstein, S. & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 107-118.

Kreutz, H. & Titscher, S. (1974): Die Konstruktion von Fragebögen. In J. Koolwijk, & M. Wieken-Mayser (Hrsg.), *Techniken der empirischen Sozialforschung, Bd 4: Erhebungsmethoden: Die Befragung* (S. 255-314). München/Wien: Verlag Oldenbourg.

Kroeber-Riel, W. (1993). *Bildkommunikation*. München: Verlag Vahlen.

Kroeber-Riel, W. & Esch, F. (2000). *Strategie und Technik der Werbung, Verhaltenswissenschaftliche Ansätze*. Stuttgart: Verlag W. Kohlhammer.

Krumm, S. & Schmidt-Atzert, L. (2009). *Leistungstests im Personalmanagement*. Göttingen: Hogrefe.

Kubinger, K. D. & Ebenhöf, J. (1996). *Arbeitshaltungen*. Frankfurt/M.: Swets Test Services.

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

Kubinger, K. D. (2003b). (Un-) Verfälschbarkeit. In K. D. Kubinger & R. S. Jäger (Hrsg.), *Schlüsselbegriffe der Psychologischen Diagnostik* (S. 429-443). Weinheim: Beltz Verlag.

Kubinger, K. D. Schrott, A. & Ortner, T. (2000). *Entwicklung Objektiver Persönlichkeitstest zu den Eignungsmerkmalen „Belastbarkeit“ und „Entscheidungsverhalten“*. Arbeitsberichte Psychologischer Dienst der Bundeswehr. Bonn: Bundesministerium der Verteidigung.

Kubinger, K. D. (2006a). *Psychologische Diagnostik*. Göttingen: Hogrefe.

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

L

Laux, H. (2003). *Entscheidungstheorie*. Berlin: Springer Verlag.

Lee, W. (1977). *Psychologische Entscheidungstheorie*. Weinheim: Beltz.

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

M

Manz, K., Dahmen, A. & Hoffmann, L. (2000). *Kompaktstudium Wirtschaftswissenschaften, Bd. 10. Entscheidungstheorie*. München: Verlag Vahlen.

Menges, G. (1970). On Subjective Probability and Related Problems - Theory and Decision. *International Journal for Philosophy and Methodology of the Social Sciences*, 1, (S. 40-60).

Moosbrugger, H. & Kelava, A. (2007). *Testtheorie und Fragebogenkonstruktion*. Heidelberg: Springer.

Morgan, M.G. und Henrion, M. (1990). *Uncertainty. A guide to dealing with uncertainty in qualitative risk and policy analysis*. Cambridge: Cambridge University Press.

Mummendey, H. D. & Grau I. (2008). *Die Fragebogenmethode*. Göttingen: Hogrefe.

N

Niegemann, H. M. (2008). *Kompendium multimediales Lernen*. Berlin: Verlag Springer.

Nisbett, R. E. & Ross, L. (1980). *Human inference: Strategies and shortcomings*. Englewood Cliffs: Prentice-Hall.

O

Onken, J., Hastie, R. & Revelle, W. (1985). Individual differences in the use of simplification strategies in a complex, decision-making task. *Journal of Experimental Psychology: Human Perception and Performance*, 11, 14-27.

Oppenheim, A. N. (1966). *Questionnaire design and attitude measurement*. New York: Basic Books.

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

Ortner, T. M., Proyer, R. T. & Kubinger, K. D. (2006). *Theorie und Praxis Objektiver Persönlichkeitstests*. Bern: Huber.

Ortner, T., Horn, R., Kersting, M., Krumm, S., Kubinger, K. D., Proyer, R. T., Schmidt-Atzert, L., Schuhfried, G., Schütz, A., Wagner-Menghin, M. M. & Westhoff, K. (2007). Zitieren von Internetquelle. (2010, Juni 27) [„Standortbestimmung und Zukunft Objektiver Persönlichkeitstests“]. URL http://www.tu-chemnitz.de/hsw/psychologie/professuren/diffpsy/oefiles/artikel/s64_s75_fach_R2_07-ortner.pdf

P

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

Petermann, F. & Eid M. (2006). *Handbuch der psychologischen Diagnostik*. Göttingen: Verlag Hogrefe.

Piaggio, L. (1973). Die empirische Erforschung der Persönlichkeit. In R. B. Cattell (Hrsg.), *The scientific analysis of personality* (S. 12). Harmondsworth: Penguin Books.

R

Radinger, R. (2002). *Evaluierung eines neuen Computertests zum Entscheidungsverhalten*. Unveröffentlichte Diplomarbeit, Universität Wien.

Richter, H. J. (1983). Grundlagen schriftlicher Massenbefragungen: ein verhaltenstheoretischer Beitrag zur Methodenkritik. In H. Feger & J. Bredenkamp (Hrsg.), *Enzyklopädie der Psychologie. Bd. 2 Datenerhebung* (S. 378-403). Göttingen: Hogrefe.

Robinson, G.H. (1964). Continuous estimation of a time-varying probability. *Ergonomics*, 7, 7-21.

Roger, T. B. (1974). An analysis of two central stages underlying responding to personality items: The self-referent decision and response selection. *Journal of Research in Personality*, 8, 128-138.

Rost, J. (1996). *Lehrbuch Testtheorie, Testkonstruktion*. Bern: Hans Huber.

Russo, J. E. & Doshier, B. A. (1983). Strategies for multiattribute binary choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 676-696.

Russo, J. E. und Schoemaker, P. J. H. (1989). *Decision traps. The ten barriers to brilliant decision-making & how to overcome them*. Doubleday: Currency.

Rütter, T. (1978). Formen der Testaufgaben. In K. J. Klauer (Hrsg.), *Handbuch der Pädagogischen Diagnostik. Bd. 1* (S. 143-161). Düsseldorf: Schwann.

S

Sader, M. (1960). Untersuchungen über Instruktionsgebung. *Bericht über den 22. Kongreß der Deutschen Gesellschaft für Psychologie*. Göttingen.

Savage, L. J. (1954). *The foundations of statistics*. New York: John Wiley & Sons.

Schäfer, R. E. (1976). *Probabilistische Informationsverarbeitung: Theorie, Methoden und ein Experiment*. Bern: Huber.

Schmidt, L. R. (1975). *Objektive Persönlichkeitsmessung in diagnostischer und klinischer Psychologie*. Weinheim: Beltz Verlag.

Sidiropoulou, E., (1997). Computerdiagnostik. In H. J. Fisseni (Hrsg.), *Lehrbuch der psychologischen Diagnostik* (S. 383-410). Göttingen: Hogrefe.

Stift, H. (1995). *Die Beeinflussung eines kognitiven Urteils durch einen affektiven Kontext bei feldabhängigen versus feldunabhängigen Personen*. Unveröffentlichte Diplomarbeit, Universität Wien.

Svenson, O. (1996). Decision making and the search for fundamental psychological regularities: What can be learned from a process perspective? *Organizational Behavior and Human Decision Processes*, 65, 252 – 267.

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

T

Teigen, K. H. (1994). Variants of subjective probabilities: Concept, norms, and biases. In K. Gilhooly, M. Keane, R. Logan & G. Erdos (Hrsg.), *Lines of thinking: Reflections on the psychology of thought* (S. 71-97). Chichester: John Wiley & Sons.

Teuschner, U. (2003). Evaluation of a decision training program for vocational guidance. *International Journal for Educational and Vocational Guidance*, 3, 177-192.

Tränkle, U. (1983). Fragebogenkonstruktion. In H. Feger & J. Bredenkamp (Hrsg.), *Datenerhebung* (S. 222-301). Göttingen: Hogrefe.

V

Volz-Sidiropoulou, E. (2004). Computerbasierte Psychodiagnostik. In H. Fisseni (Hrsg.), *Lehrbuch der psychologischen Diagnostik* (S. 279-297). Göttingen: Hogrefe.

Vukovich, A. F. (1964). Psychologische Diagnostik. In R. Heiss, K. J. Groffmann & L. Michel (Hrsg.), *Handbuch der psychologischen Diagnostik*, Bd. 6, (S. 113-147). Göttingen: Hogrefe.

W

Wendt, D. (1983). Statistische Entscheidungstheorie und Bayes-Statistik. In J. Bredenkamp & H. Feger (Hrsg.), *Enzyklopädie der Psychologie*. Bd. 5, *Hypothesenprüfung* (S. 471-529). Göttingen: Hogrefe.

Westhoff, K., Hellfritsch, L. J., Hornke, L. F., Kubinger, K. D., Lang, F., Moosbrugger, H., Püschel, A. & Reimann G. (2004). *Grundwissen für die berufsbezogene Eignungsbeurteilung nach DIN 33430*. Lengerich: Pabst Science Publishers.

Woike, J., K. (2003). Testökonomie. In K. D. Kubinger & R. S. Jäger (Hrsg.), *Schlüsselbegriffe der Psychologischen Diagnostik* (S. 405-407). Wertheim: Beltz Verlag.

Y

Yates, J. F., Veinott, X. & Patalano, A. L. (2003). Hard decision, bad decision: On decision quality and decision aiding. In S. L. Schneider & J. Shanteau (Hrsg.), *Emerging perspectives in judgement and decision research* (S. 13-63). Cambridge: Cambridge University Press.

Z

Zimbardo P. G. (1995). *Psychologie*. Berlin: Springer Verlag.

Lebenslauf

Persönliche Daten:

Name	Josef Past
Geburtsdatum	03.12.1972
Staatsbürgerschaft	Österreich
Familienstand	ledig
Adresse	4642 Sattledt Goldstraße 9a

Bildungsweg:

1997 – 2010	Psychologiestudium Universität Wien
1996 – 1997	Studiumsberechtigungsprüfung Universität Wien
1987 – 1991	Fachschule für Elektrotechnik, Linz
1983 – 1987	Hauptschule, 4142 Hofkirchen
1979 – 1983	Volksschule, 4144 Oberkappel