

Ruprecht-Karls-Universität Heidelberg
Psychologisches Institut

Instanzenbasiertes Lernen: Der Einfluss von Zeitdruck und Persönlichkeitsmerkmalen beim Komplexen Problemlösen mit dem Water-Purification-Plant-Szenario

Diplomarbeit

Heidelberg, August 04

vorgelegt von:

Sebastian Töpfer

Hans-Junginger-Strasse 3

69214 Eppelheim

bastiantoepfer@web.de

Erstbetreuer: Prof. Dr. Joachim Funke (Allgemeine und Theoretische Psychologie)

Zweitgutachter/in: Dr. Claudia Schmidt-Rathjens (Differenzielle Psychologie und Persönlichkeitsforschung) delegiert von Prof. Dr. H. J. Ahrens (Differenzielle Psychologie und Persönlichkeitsforschung, Biopsychologie und Methodenlehre)

DANKSAGUNG

Besonders danken möchte ich Cleotilde Gonzalez vom Department of Social and Decision Sciences der Carnegie Mellon University in Pittsburgh, die das computersimulierte komplexe Problemlöseszenario „Water-Purification-Plant“ zur Verfügung gestellt hat. Für eine gute Betreuung Dr. Claudia Schmidt-Rathjens und Prof. Dr. Joachim Funke. Für Ideen und Anregungen, die mir interessante Perspektiven ermöglichten, Dr. Hans Rudi Fischer sowie Dr. Markus Gumbel. Für die wertvolle Unterstützung bei Schwierigkeiten mit dem Computer möchte ich Bärbel Maier-Schicht und Marion Lamarsch danken. Ebenso danken möchte ich dem dm-Drogerie-Markt in Heidelberg und den SISI-Werken in Eppelheim, welche Incentives für die Versuchspersonen bereitgestellt haben. Schließlich natürlich auch Dank an alle Teilnehmern/innen am Experiment.

*„...und auf der Strasse steigen immer wieder Frauen in die Taxen,
alle mit der gleichen Bewegung,
mit der sie den Kopf einziehen und dann hinter
sich den Mantel festklemmen:
allmählich malt man sich diese verschiedenen Frauen schon
als etwas Mythisches aus
- altes Schluckauf seinstrunkener Poeten -: als eine Frau mit Wasser
in den Beinen einsteigt,
mühseliger als die anderen,
und heilsam das leichtfertige BILD zerstört...
Und womit kehrst du am Abend nach Hause zurück? -
mit solchen Anblicken zum Beispiel, antwortet der Anblicksammler stolz.*

*Und womit ordnest du sie? -
Weil die Angst vor dem Unsinn vorbei ist,
brauchen sie keine Ordnung mehr.
Und der eigene Eindruck? -
Weil der Unsinn vorbei ist, ist der Anblick
zugleich schon ein Eindruck geworden.“
Peter Handke (S. 116f)*

INHALT

<i>Danksagung</i>	2
<i>Inhalt</i>	4
<i>Zusammenfassung</i>	7
<i>Abstract</i>	8
<i>Tabellenverzeichnis</i>	9
<i>Abbildungsverzeichnis</i>	11
<i>Anhangsverzeichnis</i>	14
1 Einleitung	15
2 Theoretischer Hintergrund	17
2.1 Zur Entscheidungsforschung	17
2.2 Komplexes Problemlösen	20
2.2.1 Geschichte und Hintergrund des komplexen Problemlösens	20
2.2.2 Merkmale und Definition eines komplexen Problems	24
2.2.3 Zusammenfassung	25
2.3 Die Instanzenbasierte Lerntheorie	25
2.3.1 Grundlagen	25
2.3.2 Die Lernmechanismen	29
2.3.3 Ein Modell für Entscheidungsprozesse	29
2.3.4 Zusammenfassung	31
2.4 Die Water-Purification-Plant (WPP)	32
2.5 Auswirkungen von Zeitdruck in Entscheidungssituationen	35
2.6 Zur Auswahl der Kovariaten und Logik des Designs	38
2.7 Forschungsfragen und Ziele	45
2.8 Hypothesen	47
2.9 Zusammenfassung	50

3	Methoden	51
3.1	Design und Materialien	51
3.1.1	Das WPP-Szenario	52
3.1.2	Der NEO-FFI	53
3.1.3	Computervorerfahrung	55
3.2	Stichprobe	55
3.2.1	Festlegung der Stichprobengröße	55
3.2.2	Versuchsteilnehmer	56
3.3	Ablauf der Untersuchung	57
3.4	Statistische Hypothesen	59
3.5	Datenanalyseverfahren	61
4	Ergebnisse	64
4.1	Konventionen	65
4.2	Der NEO-FFI	66
4.3	Hypothesenüberprüfung	67
4.4	Explorative Datenanalyse	75
4.4.1	Neurotizismus	75
4.4.2	Extraversion/Introversion	75
4.4.3	Offenheit für Erfahrungen	77
4.4.4	Geschlechtseffekte	77
4.4.5	Computervorerfahrung	78
4.4.6	Zusammenfassung	80
4.5	Qualitative Datenanalyse	80
4.5.1	Verschlechterung und Verbesserung – wie wirken sie?	80
4.5.2	Die Analyse der Verlaufskurven	81
4.5.3	Zusammenfassung	87
4.6	Zusammenfassung der Ergebnisse	88
5	Diskussion	91
5.1	Allgemeines	91

5.2	Grenzen der Studie und Kritik am IBLT-Modell	92
6	<i>Ausblick</i>	95
7	<i>Literatur</i>	96
8	<i>Anhang</i>	104

ZUSAMMENFASSUNG

Das komplexe Problemlöse-Szenario Water-Purification-Plant (WPP) wird in der vorliegenden Diplomarbeit verwendet, um das Modell der Instanzenbasierten Lerntheorie von Gonzalez et al. zu überprüfen. Hierfür spielen zwei Gruppen das Szenario mehrmals an zwei aufeinander folgenden Tagen ohne und mit Zeitdruck durch. Als Kovariaten fließen in die repeated measurement Analyse drei ausgewählte Persönlichkeitsmerkmale aus dem NEO-FFI (Neurotizismus, Extraversion, Offenheit für Erfahrungen) und die Computervorerfahrung mit ein. Im Gegensatz zu den Befunden von Gonzalez et al. finden sich keine signifikanten Effekte bei den Kovariaten. Erst in der explorativen Analyse mittels einzelner t-Tests der Daten zeigt sich ein starker Einfluss der Variable Computervorerfahrung an beiden Test-Tagen. Des Weiteren erreichen die männlichen Teilnehmer signifikant bessere Werte bei der WPP am zweiten Tag, wobei aber der Lerngewinn für männliche und weibliche Teilnehmer/innen von Tag1 zu Tag2 fast gleich ist. Eine qualitative Analyse der Daten zeigt keinen Effekt für zwei gebildete Gruppen, die sich durch mehr Gewinne versus mehr Verluste bei den Durchgängen unterscheiden. Die Analyse der Verlaufskurven zeigt, dass erfolgreiches Agieren in der WPP mit niedrigen Neurotizismus-Werten zusammenhängt. Des Weiteren scheinen vorausschauendes Handeln und nur leichte Veränderungen der Strategie wichtig für gute Leistung bei der Bearbeitung der WPP zu sein. Das IBLT-Modell kann aufgrund der empirischen Daten als nicht bestätigt angesehen werden. Die WPP aber ist ein gut programmiertes Szenario, das viele Möglichkeiten zur weiteren Erforschung bietet.

ABSTRACT

The instance-based learning theory from Gonzalez et al. will be tested in this study using the complex problem solving scenario water-Purification-Plant (WPP). Two groups, with and without time constraints, perform in several trials using the scenario on two following days. Covariates for the repeated measurement analyses are three selected personality characteristics out of the NEO-FFI (neuroticism, extraversion, openness for experience) and the know-how people had in dealing with a computer. In contrast to the results of Gonzalez et al. there are no significant effects with the covariates. It is only in the explorative data analysis by using some t-tests that there is a significant effect for the variable computer knowledge on both days. Furthermore the male participants perform significantly better within the WPP scenario on the second day, whereby the learning progress for males and females between the first and second day is nearly the same. A qualitative analysis of the data shows no significant effects for two groups differing in that the one group had more gains and the other one more losses during the trials they made. The analysis of the curves of missed gallons suggests that successful performance in the WPP goes together with low scores on neuroticism. Above that to look forward and take future effects into account as well as only little changes in the strategy seem to correlate with succeeding in the WPP. The IBLT-model could not be verified in this study. But the WPP is a good-programmed scenario with many possibilities for future research.

TABELLENVERZEICHNIS

<i>Tabelle</i>	<i>Seite</i>
1. <i>Experimentelles Design</i>	52
2. <i>Konventionen für Effektgrößen (vgl. Bortz, 1999, S. 140)</i>	65
3. <i>Interkorrelationen und Alpha-Koeffizienten der ausgewählten Skalen für die Gesamtstichprobe (n = 34)</i>	67
4. <i>Deskriptive Maße bei der Bearbeitung der Water-Purification-Plant</i>	68
5. <i>Repeated measurement Analyse mit vier Kovariaten: globale Analyse. p-Werte sind anhand des Huynh-Feldt Sphärizitätsschätzers berichtet.</i>	71
6. <i>Mittelwerte, Standardabweichungen und die Ergebnisse des t-Tests für unabhängige Stichproben zwischen den beiden Treatment-Gruppen an Tag 1 und Tag 2</i>	72
7. <i>Mittelwerte, Standardabweichungen und die Ergebnisse des t-Tests für abhängige Stichproben zwischen den beiden Tagen für die beiden Bedingungen</i>	73
8. <i>Mittelwerte, Standardabweichungen und die Ergebnisse des t-Tests für abhängige Stichproben zwischen den beiden Tagen für die Gesamtwerte.</i>	74
9. <i>Mittelwerte und Standardabweichungen des Persönlichkeitsmerkmals Neurotizismus mit den Ergebnissen des t-Tests für unabhängige Stichproben an beiden Tagen zwischen den mediangesplitteten Gruppen.</i>	75
10. <i>„Slow“-Bedingung: Mittelwerte und Standardabweichungen des Persönlichkeitsmerkmals Extraversion mit den Ergebnissen des t-Tests für unabhängige Stichproben an beiden Tagen zwischen den mediangesplitteten Gruppen.</i>	76

11. „Fast“-Bedingung: Mittelwerte und Standardabweichungen des Persönlichkeitsmerkmals Extraversion mit den Ergebnissen des t-Tests für unabhängige Stichproben an beiden Tagen zwischen den mediangesplitteten Gruppen.	76
12. Mittelwerte und Standardabweichungen des Persönlichkeitsmerkmals Offenheit für Erfahrungen mit den Ergebnissen des t-Tests für unabhängige Stichproben an beiden Tagen zwischen den mediangesplitteten Gruppen.	77
13. Mittelwerte und Standardabweichungen der Variable „Geschlecht“ mit den Ergebnissen des t-Tests für unabhängige Stichproben an beiden Tagen zwischen den Gruppen „Frauen“ und „Männer“.	77
14. Mittelwerte und Standardabweichungen der Extremwerte der Variable „Computervorerfahrung“ mit den Ergebnissen des t-Tests für unabhängige Stichproben an beiden Tagen.	79
15. Mittelwerte und Standardabweichungen der Variable „loss-gain“ mit den Ergebnissen des t-Tests für unabhängige Stichproben an beiden Tagen.	81
16. Beispiel-Items aus dem NEO-FFI (Borkenau und Ostendorf, 1993)	104
17. Daten der ersten Sortierung	107
18. Daten der zweiten Sortierung	108
19. Daten der dritten Sortierung	110

ABBILDUNGSVERZEICHNIS

<i>Abbildung</i>	<i>Seite</i>
1. <i>Das IBLT Prozessmodell</i>	30
2. <i>Screenshot der Water Purification Plant-Aufgabe</i>	35
3. <i>Die Auswirkungen von Zeitdruck auf das IBLT Prozessmodell</i>	36
4. <i>Das konzeptuelle Modell zu Leistung und Erfahrung beim Entscheiden in komplexen Problemlösesituationen von Omodei und Wearing (1995)</i>	43
5. <i>Schematisch dargestellter Ablauf des Experiments</i>	59
6. <i>Graphische Abbildung der Leistungskurven zwischen der „slow“-Bedingung und der „fast“-Bedingung über 12 Durchgänge. Wenige missed gallons bedeuten eine gute Leistung.</i>	69
7. <i>Graphische Abbildung der Leistungskurven zwischen der „slow“-Bedingung und der „fast“-Bedingung an beiden Tagen. Wenige missed gallons bedeuten eine gute Leistung.</i>	70
8. <i>Zwei Beispiele von Verlaufskurven. Oben in der „slow“-Bedingung mit insgesamt acht Durchläufen und unten in der „fast“-Bedingung mit insgesamt zwölf Durchläufen. Die missed gallons werden kumulativ pro Deadline (halbstündlich ab fünf Uhr simulierter Zeit) angegeben.</i>	82
9. <i>Bewertung der Kategorien der ersten Sortierung anhand der gesamten missed gallons</i>	84
10. <i>Verhältnis von Teilnehmer/innen mit hohen vs. niedrigen Werten beim Persönlichkeitsmerkmal Neurotizismus in der ersten Sortierung nach Kategorien.</i>	84

11. *Verhältnis von Teilnehmer/innen mit hohen vs. niedrigen Werten beim Persönlichkeitsmerkmal Extraversion in der ersten Sortierung nach Kategorien.* 85
12. *Verhältnis von Teilnehmer/innen mit hohen vs. niedrigen Werten beim Persönlichkeitsmerkmal Offenheit für Erfahrungen in der ersten Sortierung nach Kategorien.* 86
13. *Wirkbeziehungen zwischen Variablengruppen im Problemlöseprozess nach Schoppek und Putz-Osterloh (2003).* 92
14. *Bewertung der Kategorien der zweiten Sortierung anhand der gesamten missed gallons.* 108
15. *Verhältnis von Teilnehmer/innen mit hohen vs. niedrigen Werten beim Persönlichkeitsmerkmal Neurotizismus in der zweiten Sortierung nach Kategorien.* 109
16. *Verhältnis von Teilnehmer/innen mit hohen vs. niedrigen Werten beim Persönlichkeitsmerkmal Extraversion in der zweiten Sortierung nach Kategorien.* 109
17. *Verhältnis von Teilnehmer/innen mit hohen vs. niedrigen Werten beim Persönlichkeitsmerkmal Offenheit in der zweiten Sortierung nach Kategorien.* 110
18. *Bewertung der Kategorien der dritten Sortierung anhand der gesamten missed gallons.* 111
19. *Verhältnis von Teilnehmer/innen mit hohen vs. niedrigen Werten beim Persönlichkeitsmerkmal Neurotizismus in der dritten Sortierung nach Kategorien.* 111
20. *Verhältnis von Teilnehmer/innen mit hohen vs. niedrigen Werten beim Persönlichkeitsmerkmal Extraversion in der dritten Sortierung nach Kategorien.* 112

21. *Verhältnis von Teilnehmer/innen mit hohen vs. niedrigen Werten beim Persönlichkeitsmerkmal Extraversion in der dritten Sortierung nach Kategorien.*

112

ANHANGSVERZEICHNIS

<i>Anhang</i>	<i>Seite</i>
1. <i>Beispiel-Items aus dem NEO-FFI (Borkenau und Ostendorf, 1993)</i>	104
2. <i>Instruktion für die Teilnehmer/innen.</i>	105
3. <i>Datentabellen und Abbildungen der qualitativen Auswertung</i>	107

1 EINLEITUNG

Bei der Lektüre der Zeitung am Sonntag morgen (Frankfurter Allgemeine Sonntagszeitung vom 22 Februar 2004, Nr. 8) prangt im Wirtschaftsteil ein Artikel über Frankfurt am Main. Der Flughafenausbau wird gestoppt, es gibt immer weniger Erwerbstätige und teure Büroflächen stehen leer. Vorbei die Zeiten von einem boomenden Standort, in dem die enorm hohe Gewerbesteuer von der Bürgermeisterin mit den Worten: „Das Gute hat eben seinen Preis“, kommentiert wird. Hätte man solch eine Dynamik vorhersehen können? Gleich auf der nächsten Seite werden wir darüber aufgeklärt, was passieren kann, wenn keine eindeutige Aufgabentrennung zwischen Staat und Wirtschaft vorherrschen – am Beispiel der misslungenen Einführung der LKW-Maut durch die Firma „Toll Collect“. Und drei Seiten weiter werden mögliche Ursachenfaktoren diskutiert – geknüpft an Personen. Irgend jemand muss ja Schuld sein. Wie lässt sich die Dynamik in solch einem System beschreiben? Können Politiker und Manager da noch einen guten Überblick behalten und effiziente Entscheidungen treffen? Welche Informationsquellen werden dafür genutzt und wie lassen sich die (Persönlichkeits-) Merkmale der Leute, die besonders effiziente Entscheidungen treffen, beschreiben? Komplexes Problemlösen ist ein wichtiger Zweig der Denkpsychologie und die Chancen liegen hoch, dass ihm eine rosige Zukunft vergönnt ist. Die Relevanz des Themas und die Bedeutung zeigt sich an der immer größeren Anzahl veröffentlichter Artikel (vgl. Funke, 1999a) in Psychologischen Zeitschriften. Des weiteren lässt ein Blick in Zeitungen die Fülle an Anknüpfungspunkten zu aktuellen politischen, wirtschaftlichen und gesellschaftlichen Fragen erahnen. Einige Entscheidungen mit ungutem Ausgang wären vielleicht nicht getroffen worden, wenn Wissen über die Prozesse, die in Interaktion mit komplexen Problemstellungen auftreten, vorhanden gewesen wäre.

Nach Heiko Ernst (2000) haben wir in der heutigen Gesellschaft bei vielen Dingen eine schier unbegrenzte Entscheidungsfreiheit, und um damit umzugehen auch – sozusagen als „Gratiszugabe“ – jede Menge an Informationen. Schon Adam und Eva mussten sich entscheiden, mussten wählen zwischen Unschuld oder Erkenntnis, Gehorsam oder Eigenwille. Wie gehen wir aber in der heutigen Zeit damit um? Wird Entscheidungsfreudigkeit, wie Ernst vorschlägt, zu einer neuen Schlüsselqualifikation, weil im Zeitalter der Flexibilität der Trend besteht, Entscheidungen eher auszusitzen oder abzuwarten?

Wie lernen wir, in einem komplexen System Entscheidungen zu treffen? Und was ist das: ein komplexes System? Ein gutes Beispiel mit Bezug zu der vorliegenden Arbeit ist der Erwerb eines neuen Handys von einem neuen Hersteller. Das Menü funktioniert anders, ist anders aufgeteilt und es gibt viel mehr Funktionen die größtenteils in ihrer Funktion unbekannt sind. Vielleicht liest man nun zuerst die Gebrauchsanleitung, oder lässt sich die relevanten Funktionen im Laden persönlich erklären. Vielleicht probiert man aber auch einfach Verschiedenes aus, nach dem „trial-and-error“ Prinzip? Vielleicht werden sogar manche Funktionen nie entdeckt, obwohl sie viel genützt hätten? Wie ein komplexes Problem beschrieben werden kann und welche Lernmechanismen im Umgang damit wirken, wird im Laufe der nächsten Seiten diskutiert.

Je nachdem, in welchem Kontext man Entscheidungssituationen sieht, ergeben sich unterschiedliche Implikationen und Möglichkeiten die beteiligten Prozesse zu untersuchen. In der vorliegenden Diplomarbeit soll der Fokus auf der kognitiven Verarbeitung liegen und auf dem Einfluss von Zeitdruck beim Lernen von Entscheidungen in komplexen Systemen. Außerdem sollen die Wechselwirkungen von ausgewählten Persönlichkeitsmerkmalen mit der Bearbeitung einer komplexen Problemlöseaufgabe untersucht werden. Den Rahmen gibt die Instanzenbasierte Lerntheorie, die von Gonzalez, Lerch und Lebiere (2003) als eine allgemeine Theorie zum Entscheiden in komplexen Situationen vorgeschlagen wird.

Der Aufbau der Arbeit erfolgt in Anlehnung an die Arbeiten der Arbeitsgruppe um Cleotilde Gonzalez an der Carnegie Mellon University in Pittsburgh/USA (department of social and decision sciences). Das zur Datenerhebung verwendete computersimulierte Szenario wurde von ihr für diese Untersuchung freundlicherweise bereitgestellt.

Die geplante Arbeit beschäftigt sich mit einer Thematik, die sowohl Bereiche der Kognitiven Psychologie als auch der Differenziellen Psychologie (vgl. hierzu Süß, 1999) beinhaltet. Der Schwerpunkt liegt in der Kognitiven Psychologie auf den Teilgebieten Komplexes Problemlösen, Entscheidungs- und Lernforschung und in der Differenziellen Psychologie auf den Fünf-Faktoren-Persönlichkeitsmodellen.

2 THEORETISCHER HINTERGRUND

Im folgenden Abschnitt wird ein Überblick über die *Entscheidungsforschung* gegeben, die einen Eckpunkt der Studie markiert. Als weiterer Eckpunkt wird das Forschungsgebiet *Komplexes Problemlösen* besprochen, dessen Geschichte, Hintergrund und Bezug zur vorliegenden Arbeit, sowie die Merkmale und die Definition eines komplexen Problems. Anschließend folgt eine Beschreibung der *Instanzenbasierten Lerntheorie*, wie sie bei Gonzalez et al. (2003) verwendet wird. Die Auswahl der *Kovariaten* – der NEO-FFI und ausgewählte Persönlichkeitsdimensionen – wird darauffolgend erläutert und die *Forschungsfragen und Ziele* werden formuliert. Schließlich werden die *Hypothesen* aus vorliegenden Forschungsarbeiten abgeleitet.

2.1 Zur Entscheidungsforschung

Ein Mensch erhofft sich fromm und still,
dass er einst das kriegt, was er will.
Bis er dann doch dem Wahn erliegt,
und schließlich das will, was er kriegt.

Eugen Roth

(zitiert nach Jungermann, Pfister & Fischer, 1998, S. 34)

Die Entscheidungsforschung hat ihre Wurzeln in der Mathematik (Wahrscheinlichkeitstheorie), der Philosophie (Utilitarismus) und der Ökonomie (homo oeconomicus) (vgl. Jungermann et al., 1998). Seit etwa 200 Jahren werden entscheidungstheoretische Probleme in verschiedenen Wissenschaftsbereichen diskutiert. Das Zentrum dieses Forschungsbereiches lässt sich momentan am ehesten der Ökonomie zuschreiben (z.B. Schäfer, 1998; Beck, 2001), was auch an den verliehenen Nobelpreisen deutlich wird (vgl. Leahey, 2003; Kahneman, 2003). Wenn man Entscheidungsforschung betreibt, gibt es zwei mögliche Richtungen, die eingeschlagen werden können: die präskriptive Entscheidungsforschung und die deskriptive. Erstere beschäftigt sich mit der Frage, welche Entscheidungen man vernünftigerweise wählen sollte, wenn man bestimmte Grundpostulate rationalen Denkens für richtig hält (vgl. hierzu auch Fischer, in press), und man kann mit formalisierten Regeln und Verfahren zur Strukturierung

von Entscheidungssituationen Menschen unterstützen (z.B. Arztentscheidungen). Die deskriptive Richtung möchte das menschliche Entscheidungsverhalten beschreiben und produziert hierzu Modelle und Theorien. Mit ihnen kann Verhalten erklärt und vorhergesagt werden. Wichtig hierbei ist, dass der Mensch sich oftmals nur begrenzt rational verhält. Das dazu passende Konzept der „*bounded rationality*“, wurde von Herbert Simon 1978 eingeführt (vgl. Leahey, 2003).

Ein anderes bedeutendes Konzept wurde nach der „kognitiven Wende“ die Sicht auf den *Menschen als ein informationsverarbeitendes System*. Natürlich gibt es daran auch einige Kritik (vgl. Speiser, 2004), aber dieses Konzept spannt einen Rahmen, in dem „Entscheiden“ als spezifischer kognitiver Prozess angesehen werden kann, der nach bestimmten Regeln operiert – also als eine vergleichende Beurteilung und Wahl. Dabei spielen sowohl andere kognitive Prozesse eine Rolle (Wahrnehmung, Denken, Gedächtnis) als auch die Verwendung von Sprache.

Nach Jungermann et al. (1998) setzen Entscheidungsprozesse Wissen und Motivation voraus und außerdem sind Emotionen bedeutsam (vgl. auch Vollmeyer & Funke, 1999; Buchner, 1999, S. 210). In Entscheidungssituationen wird auf Vorwissen aus Erfahrungen mit ähnlichen Situationen zurückgegriffen (vgl. Gonzalez et al., 2003; Logan, 1988). Dies kann sowohl förderlich als auch hinderlich sein, wie man zum Beispiel beim Umgang mit einem neuen Betriebssystem für den Computer selbst erfahren kann (vgl. auch Brehmer, 1992; Van der Linden, Sonnentag, Frese & van Dyck, 2001). Motivation kann als Voraussetzung angesehen werden, eine Entscheidung überhaupt anzugehen und hängt auch mit den erwarteten Konsequenzen und dem Nutzen von Entscheidungen zusammen (vgl. Heckhausen, 1980, S. 170ff). Der Zusammenhang zwischen Emotionen und Entscheidungen ist aber noch nicht ausreichend erforscht (Jungermann, et al., 1998; Sperring, 2001).

Verschiedene Sichtweisen auf das Thema Entscheidungen zeigen zum Beispiel die Denkpsychologie, die Motivationspsychologie sowie die Lernpsychologie auf. Aus denkpsychologischer Sicht interessieren die Operatoren, die zur Lösung einer Entscheidung entwickelt wurden. Die Motivationspsychologie beschäftigt sich mit den Anreizen in Konfliktsituationen, also den Zielen und den Gründen von bestimmten Entscheidungen. Die letztgenannte Perspektive untersucht den Einfluss von Erfahrungen in Entscheidungssituationen, der auch in der vorliegenden Diplomarbeit eine zentrale Rolle spielt. Es geht also nicht um „static snapshots“ bei Entscheidungen, sondern um Ketten von

Entscheidungen, die sowohl *Zeit* als auch *Kontext* stärker berücksichtigen als viele Aufgaben der klassischen psychologischen Entscheidungsforschung (vgl. Brehmer, 1996).

Die vorliegende Diplomarbeit verknüpft – in Ahnlehnung an die Arbeiten der Forschungsgruppe um Gonzalez – mehrere Trends in der Entscheidungsforschung (vgl. Jungermann et al., 1998):

Aus kognitionspsychologischer Sicht greift sie das *Strukturkonzept der mentalen Modelle* auf, in dem Überlegungen zur Repräsentation des Lernprozesses während des Entscheidens angestellt werden. Weiterhin nimmt sie Bezug zum *Prozesskonzept des „matching“*, wegen der Annahme von SDUs – Wissensseinheiten, die in „situation“, „decision“ und „utility“ unterteilt sind (vgl. Kapitel 2.3). Aufgrund dieser Annahme folgt, dass im Entscheidungsprozess auf passende, früher gemachte Erfahrung, zurückgegriffen wird und dabei diejenige ausgewählt wird, die unter den gegebenen Umständen am besten passt. Zentral ist in diesem Gebiet das *Entscheidungsverhalten in dynamischen Situationen*, und auch der Nutzen *inferenzieller Urteilsstrategien*, also Heurismen.

Die emotionspsychologische Sicht wird in gewisser Weise durch die Rolle der Persönlichkeitsdimension „Neurotizismus“ berücksichtigt. Dies wird später genauer erläutert. Des weiteren kann Zeitdruck beim Entscheiden als emotionsinduzierendes Moment gesehen werden (vgl. dazu auch Brehmer, 1992, 1996; Kreuzig, 1983; Spering, 2001; Schwarz, 2000). Dies ermöglicht eine neue Sichtweise auf die vorliegende Studie.

Die differenzialpsychologische Perspektive kommt durch den Einbezug von Persönlichkeitsmerkmalen und die Überlegungen, wie Expertenwissen entsteht und wie es genutzt wird, ins Spiel. In der vorliegenden Diplomarbeit wird davon ausgegangen, dass Expertenwissen in der WPP relativ schnell aufgebaut werden kann (vgl. Logan, 1988). Zuerst beruhen die Entscheidungen auf Heuristiken; später wird davon abgesehen und die Entscheidungen beruhen dann auf Erfahrungen, dem angehäuften und modifizierten Wissen (beschrieben als *SDUs* bei Gonzalez et al. 2003). Geschlechtsspezifische Aspekte werden explorativ untersucht. Weiterhin kann auch ein interkultureller Vergleich mit den in den USA gewonnenen Daten von Gonzalez stattfinden. Dies liegt allerdings nicht im Fokus der vorliegenden Diplomarbeit. Dass aber kulturübergreifende Untersuchungen interessante Gedanken anregen können, zeigen Strohschneider und Güss (1999). Jedoch ist die vorliegende Diplomarbeit eine „modifizierte Replikation“ der Studien von Gonzalez und Vergleiche können insofern nur in begrenztem Rahmen vorgenommen werden.

Schließlich mündet die Idee der Verbindung verschiedener Ansätze in der praktischen Relevanz der vorliegenden Diplomarbeit. Ziel ist es, mit den Ergebnissen der Untersuchung das Lernen von Entscheidungen in verschiedenen Anwendungsbereichen unterstützen zu können. Denkbar sind hier Gebiete, wie die Mensch-Computer-Interaktion sowie die Konstruktion von Lernumgebungen in komplexen Systemen. So können Implikationen für das Vermeiden von Monotonie bei Arbeitsprozessen abgeleitet werden, genauso wie Ideen für eine effiziente Vorbereitung von Menschen auf komplexe Problemlösesituationen. Schließlich könnte die Fehlerforschung – im Zusammenhang mit Zeitdruck und bestimmten Persönlichkeitsvariablen – daraus Anregungen für Trainings gewinnen (vgl. auch Hüther, 2004).

2.2 Komplexes Problemlösen

"Am meisten Energie vergeudet der Mensch mit der Lösung von
Problemen, die niemals auftreten werden."

William Somerset Maugham

2.2.1 Geschichte und Hintergrund des komplexen Problemlösens

Die Forschungsrichtung „komplexes Problemlösen“ beschäftigt sich allgemein mit dem Denken und Handeln von Menschen in unbestimmten, dynamischen und komplexen Situationen (Dörner, Schaub & Strohschneider, 1999). Ursprünglich wurde sie mit der Absicht gebildet, die traditionelle Denk- und Problemlöseforschung neu zu orientieren, da diese sich zunehmend von der Realität zu entfernen schien. Dies bedeutet, dass die Problemlöseforschung den Kontext, in den die Handlungen und das Denken von Menschen eingebettet sind, nicht ausblendete, sondern wieder mit einfügte. Wichtig hierbei war die Verbreitung von Computern, die es erlaubte komplexe Simulationen zu konstruieren, welche sich durch Nähe zu alltäglichen Situationen auszeichnen. Die frühen Studien zum komplexen Problemlösen stammen aus den 70er und 80er Jahren (Funke, 1999b). Donald Broadbents Forschung (80er und 90er Jahre) über verschiedene Gedächtnissysteme und die Arbeiten der Gruppe um Dietrich Dörner (ab Mitte der 70er Jahre) markieren den Beginn der komplexen Problemlöseforschung in Europa (vgl. Funke, 2001). Schon in den 60er Jahren gab es zwar in den USA Strömungen, die über „business simulations“ und „management games“ forschten. Sie wurden aber mangels Internationalisierung in der Psychologie nicht beachtet und somit nicht aufgegriffen. Heute (im Mai 2004) verzeichnet die Datenbank „PsychInfo“ bei der Eingabe des Begriffes „complex problem solving“ 252 Treffer (vgl. Funke, 1999b).

Dörners System „Lohhausen“ zählt zu den bekanntesten (vgl. Dörner, 1989). Darin muss eine Versuchsperson die Rolle eines Bürgermeisters einer Kleinstadt übernehmen und diesbezügliche Entscheidungen fällen. Dieses System hat mehr als 2000 Variablen. Interessant ist die aus diesem Experiment resultierende Kritik an der traditionellen Intelligenzforschung und das Einführen einer neuen Art und Weise psychologische Experimente durchzuführen: mit dem Computer, mit simulierten Szenarien, und dadurch mit der Option einen Wirklichkeitsbereich effizient (wirklichkeitsnäher) abbilden zu können. Kritisch ist häufig die Vielzahl der Variablen und somit eine Unklarheit bezüglich der relevanten Parameter. Konkret also: wie bestimmt man, wer ein eher guter und wer eine eher schlechter Entscheider innerhalb eines Szenarios ist?

Brehmer (1992) beschreibt dazu zwei Ansätze, den (1) Ansatz individueller Unterschiede und den (2) experimentellen Ansatz. Das Vorgehen beim Ansatz individueller Unterschiede ist, dass einige Teilnehmer/innen eine Computersimulation durchspielen und anschließend anhand ihrer Ergebnisse in Extremgruppen (z.B. erfolgreich vs. nicht erfolgreich bezüglich einer bestimmten Variable) eingeteilt werden. Diese Gruppen werden dann mit anderen Verhaltensmaßen oder verschiedenen Daten aus psychometrischen Testverfahren korreliert und verglichen. So möchte man herausfinden, welche Anforderungen eine bestimmte Simulation an die Teilnehmer/innen stellt, und wie man Verhalten beim komplexen Problemlösen vorhersagen kann. Im zweiten Ansatz wird versucht Systemmerkmale zu variieren. Das Verhalten der Teilnehmer/innen lässt dann Rückschlüsse auf den Aufbau mentaler Modelle zu. Dieser Ansatz liegt der vorliegenden Diplomarbeit zugrunde. Der Zeitdruck im Umgang mit der Water Purification Plant (WPP) wird manipuliert und es wird angenommen, dass in den beiden Bedingungen verschieden effektive mentale Modelle gebildet werden können, was am Verlauf der Leistungen in den verschiedenen aufeinanderfolgenden Durchgängen der WPP-Aufgabe abgelesen werden kann (siehe auch 2.3).

Momentan befassen sich mehrere Forschungsgruppen in Deutschland mit unterschiedlichen Foki mit dem Gebiet des komplexen Problemlösens (vgl. Funke, 2003). Von Bemühungen eine Gesamtheorie über psychische Strukturen und Prozesse zu entwerfen, zu Bezügen von Problemlösefähigkeit mit klassischen Intelligenztest bis zur Einflechtung eines grundlegenden formalen Rahmens in die Forschung im Sinne eines Qualitätsmanagement zur besseren Vergleichbarkeit der Ergebnisse und zum Vermeiden von Fehlern in vergangenen Studien (Funke, 1999b, 2001; Buchner, 1999; Funke & Buchner, 1992).

So formulieren Dörner et al. (1999) als Ziel ihrer Forschung die Integration verschiedener Teilgebiete der Allgemeinen Psychologie. Dazu legen sie ihre PSI-Theorie vor, deren Grundlage das Nachbilden von Absichten und Handlungen via Computersimulation ist. Diese Idee ähnelt der Forschung von Gonzalez et al. (2003), die Lernprozesse (Wissensbildung) – ausgehend von der Instanzenbasierten Lerntheorie – simuliert. Ein wichtiges Konzept innerhalb der PSI-Theorie ist das der *Selbstreflexion*, welche Tisdale (1998, zitiert nach Dörner et al., 1999) eingehend untersucht hat. Nach ihm hilft Selbstreflexion bei der Modifikation von Denkstrukturen und –abläufen und dient zum Planen und zur Organisation des Denkens im Voraus. Versuchspersonen zeigen in seinen Experimenten, dass nicht nur das Explorationsverhalten verbessert wird, sondern die Denkprozesse plastischer und flexibler erscheinen. Allerdings ist Selbstreflexion mühsam, benötigt Zeit und sie kann das Selbstwertgefühl des Individuums beeinträchtigen; insbesondere, wenn die Aufmerksamkeit auf missglückte Lösungsversuche gelenkt wird. Diese Aspekte sind für die für die experimentelle Manipulation der vorliegenden Diplomarbeit wichtig. Durch Zeitdruck soll in der „fast“-Bedingung Selbstreflexion verhindert werden (vgl. Abbildung 3, S.36). Dies deckt sich mit der Hypothese (1), wonach Versuchspersonen in der „fast“-Bedingung eine andere Lernkurve zeigen als Versuchspersonen in der „slow“-Bedingung (vgl. 2.8). Letztere sollten (laut Gonzalez et al., submitted) am zweiten Tag, trotz weniger Durchläufen, schneller sein, dafür mit dem Bonus von genügend Zeit (und somit der Möglichkeit) zur Selbstreflexion.

Bezüglich der Alltagsnähe der Forschung zum komplexen Problemlösen lässt sich festhalten, dass Denken und Entscheiden oft unter Zeitdruck stattfindet, teils unter hohen emotionalen Belastungen und Entscheidungen eben auch Nebenwirkungen und Langzeiteffekte hervorbringen. Weiterhin stoßen Menschen an verschiedene Limits (kognitive Fähigkeiten, verfügbare Zeit, zur Verfügung stehende Informationen, Erfahrungswerte, etc.) im Umgang mit komplexen Systemen, was mit dem Konzept „*bounded rationality*“ von Herbert Simon (Kahneman, 2003; Leahey, 2003) beschrieben wird. Diese Limitierungen führen zu interessanten Auswirkungen, welche dem Konzept der *Heuristik* besondere Bedeutung geben. Solche Daumenregeln sind beim Entscheiden oftmals erstaunlich effektiv (vgl. Gigerenzer & Todd, 1999).

Als allgemeine Ziele der Komplexen Problemlöseforschung nennt Funke (2003) (1) das Finden von Regularitäten bei der Interaktion von verschiedenen Prozessen (vgl. auch Brehmer, 1996, S. 228) und (2) das Untersuchen der Interaktion von kognitiven, emotionalen und motivationalen Prozessen bei der Bearbeitung von computersimulierten Szenarien. Auch gehört das Erforschen

von Fehlerquellen, Katastrophen, Dogmenbildung und anderer Phänomene im Zusammenhang mit computersimulierten Szenarien zum Gegenstandsbereich der komplexen Problemlöseforschung.

Als zentrale Befunde der angesprochenen Lohhausen-Studie wurden Selbstsicherheit und Extraversion als Erfolgsprädiktoren herausgestellt. Dies passt zu den Hypothesen (2) und (4) (siehe 2.8), da Extraversion als ein Persönlichkeitsmerkmal untersucht wird und der Autor annimmt, dass man zwischen Selbstsicherheit und dem Persönlichkeitsmerkmal „Neurotizismus“ – das in der Diplomarbeit erhoben wird – enge Beziehungen herstellen kann (vgl. hierzu auch Omodei & Wearing, 1995). Die Logik des Designs wird ebenso unterstützt durch die Vorschläge Buchners (1995) bezüglich des Zugangs zum komplexen Problemlösen durch die Suche nach individuellen Unterschieden (zum Beispiel bei der Lernfähigkeit), der heuristischen Kompetenz, und dem Vergleich von Experten und Novizen. Aufbauend auf die Resultate von Gonzalez et al. (2003) wird ein Zugang, der individuelle Unterschiede untersucht, in der vorliegenden Diplomarbeit angewandt.

Als weiterer interessanter Bezug zur Logik des gewählten Designs sei Dörners Spezifizierung von intellektuellen Notfallreaktionen beim Umgang mit komplexen Problemen erwähnt (vgl. auch Van der Linden et al., 2001). Diese sind (1) die Senkung des intellektuellen Niveaus, (2) die Tendenz zu schnellem Handeln und (3) die Degeneration der Hypothesenbildung. Die Effekte des Absinkens von Selbstreflexion, werden in der vorliegenden Arbeit durch die zwei experimentellen Bedingungen („slow“ versus „fast“) untersucht. Eine Degeneration der Hypothesenbildung mag – wenn man die Idee hinter der vorliegenden Arbeit mit den Beschreibungen von Dörner verknüpfen will – auf einen ineffizienten Aufbau der SDUs rückführbar sein.

Mögliche Anwendungsfelder findet die komplexe Problemlöseforschung zum Beispiel bei der Eignungsdiagnostik von Führungskräften, beim immer bedeutender werdenden Training von Systemkompetenz (Kriz, 2000; Capra, 1996, Vester, 2002), in internationalen Studien wie z.B. PISA (Programme for Internationale Student Assessment) oder INES (International Life Skills Survey), bei Fragen zu Lernen und Transfer sowie bei der kognitiven Modellierung komplexer Prozesse.

2.2.2 Merkmale und Definition eines komplexen Problems

Zur Ordnung des Gegenstandsbereichs „Komplexes Problemlösen“ schlägt Funke (2003) eine Taxonomie vor, in der er zwischen Personen-, Situations- und Aufgabenmerkmalen unterscheidet. Kognitive Personenmerkmale sind zum Beispiel Verarbeitungsmerkmale, wie in der vorliegenden Arbeit die *Selbstreflexion* oder bei nicht-kognitiven Merkmalen die *Selbstsicherheit*, welche nach Ansicht des Autors eng mit dem Persönlichkeitsmerkmal „Neurotizismus“ verbunden ist. Ein situationsspezifisches Merkmal wäre *Zeitdruck* oder das *Ausmaß des Trainings*. Systemmerkmale beschreiben die Barrieren zwischen dem Ist- und Soll-Zustand, also hier das *effiziente Anwenden* von Heuristiken und – im späteren Verlauf – des erworbenen Wissens zum Öffnen und Schließen verschiedener Pumpen zu den richtigen Zeitpunkten. Merkmale von komplexen Problemen sind nach Funke (2003):

- (1) Komplexität,
- (2) *Vernetztheit*,
- (3) *Dynamik*,
- (4) Intransparenz, und
- (5) *Politelie*.

Funke (2001) erläutert, wie man durch einen formalen Hintergrund bei computersimulierten Problemlösenszenarien, sauberer experimentell arbeiten kann und vor diesem Hintergrund – aus Strukturgleichungsmodellen oder der Theorie der Finiten Automaten – dann die Oberfläche der Szenarien an die Erfordernisse anpassen kann. In dieser Diskussion stellt er heraus, dass die zwei Merkmale „*Dynamik*“ und „*Vernetztheit*“ die entscheidenden sind. *Vernetztheit* als Abhängigkeit zwischen zwei oder mehr Variablen beinhaltet das Merkmal Komplexität. Intransparenz und *Politelie* sind keine systemimmanenten Merkmale, sondern sind abhängig von der Entscheidung des Versuchsleiters, wie viel Informationen er preisgeben möchte und welche Ziele in der Instruktion vorgegeben werden. Diese zwei Merkmale sind oft Teil der Präsentation eines komplexen Problemlösenszenarios, aber sie sind nicht Teil der Computersimulation.

2.2.3 Zusammenfassung

Zusammengefasst lässt sich sagen, dass die komplexe Problemlöseforschung als relativ neues und fruchtbares Gebiet einige reizvolle Ansätze und Ideen zur Verfügung stellt. Da laut Funke (1999b) die amerikanische Problemlöseforschung wenig interessante Forschungsansätze bietet, besteht die Möglichkeit diesem Forschungsgebiet einen europäischen Schwerpunkt zu geben. In der vorliegenden Diplomarbeit liegt der Fokus auf dem Bezug von Persönlichkeitsvariablen zu komplexen Problemlösen, also einer Verknüpfung von Kognitionspsychologischer und Differenzieller Psychologie.

2.3 Die Instanzenbasierte Lerntheorie

2.3.1 Grundlagen

Die Idee hinter der Instanzenbasierten Lerntheorie ist, dass durch Übung separate Instanzen im Gehirn abgespeichert werden, die eine Situation (S = situation) mit den möglichen Entscheidungsalternativen (D = decision) und dem letztendlichen Nutzen (U = utility) der vollzogenen Entscheidung enthält. Diese Einheiten werden von Gonzalez et al. demzufolge SDU's genannt. Etymologisch hängt das Wort „Instanz“ zusammen mit „der zuständigen Stelle (insbesondere bei Behörden oder Gerichten)“ und auch mit der Bedeutung „auf etwas stehen“ (vgl. Duden Bd.7). Beim Ausführen von Fertigkeiten wie zum Beispiel Auto fahren, benutzen wir also zuerst Heuristiken und dann immer mehr die durch Erfahrung gewonnenen und abgespeicherten Instanzen – also die spezifischen Lösungen zu verschiedenen erlebten Situationen. Die Wissensplattform, das sichere Terrain, wächst.

Das Abrufen solcher Instanzen läuft, wenn sie gebildet sind – laut Logan (1988) – ohne Anstrengung ab und wird nicht durch Aufmerksamkeitsressourcen limitiert. Somit ist Automatisierung nicht dem Forschungsgebiet der Aufmerksamkeit zugeordnet, sondern vielmehr der Forschung der Gedächtnis- und Lernpsychologie.

Logan beschreibt das Phänomen wie Menschen Aktivitäten auf „Autopilot“ ausführen näher. Wie verhält sich überhaupt unsere Aufmerksamkeit? Wie entstehen Routinen, Gewohnheiten, Fertigkeiten etwas zu tun, ohne dass wir uns über die Einzelschritte bei der Ausführung bewusst sein müssen? Was sind die Gefahren solcher Routinen? Ist der automatische Prozessein anstrengender oder geschieht er einfach, ist er autonom? Wie wird aus einem Novizen in einer bestimmten Sache ein Experte? Dies sind Fragen, die Logan in seinem Artikel behandelt und

die einen Eckpfeiler bei der Instanzenbasierten Lerntheorie darstellen, die in der vorliegenden Diplomarbeit als Grundlage dient und den Rahmen für die unabhängige Variable „Zeitdruck“ bildet (vgl. Abbildung 3).

Ein gutes Beispiel ist neben dem Lernen zu Lesen oder dem Rechnen im Zahlenraum bis 100 das Lernen, Auto zu fahren: In den ersten Fahrstunden achtet man genau darauf, die Füße auf den Pedalen mit dem Schalten der Hand zu koordinieren und gleichzeitig den Schulterblick nicht auszulassen. Dabei wird auch noch auf Relevantes geachtet und darauf, angemessen zu reagieren. Dies erscheint als sehr viel auf einmal und fühlt sich auch anstrengend an. Schon nach einiger Zeit aber scheinen die Abläufe wie von selbst zu funktionieren und man kann neben dem Diskutieren mit dem Freund oder der Freundin auf dem Beifahrersitz gleichzeitig Musik hören und noch eine Zigarette rauchen. Wie solch eine enorme Leistung des Gehirns vonstatten geht, kann mit der Instanzenbasierten Lerntheorie (IBLT) erklärt werden.

Um Lernprozesse vorhersagen zu können, muss laut Regina Vollmeyer eine Theorie vorhanden sein (Funke, 1999a). Die Instanzenbasierte Lerntheorie stellt ein Rahmenmodell für das Lernen von Entscheiden in komplexen Problemlösesituationen zur Verfügung. Gonzalez et al. (2003) beschreiben das Lernen und Aneignen von Fertigkeiten in der IBLT vor dem Hintergrund von vier anderen Theorien.

Sie beziehen sich unter anderem auf die chunking theory von Chase und Simon (1973, zitiert nach Gonzalez et al., 2003). Experten suchen selektiver nach relevanter Information im Vergleich zu Novizen. Diese Theorie wurde durch die Annahme erweitert, dass Leute mit einem hohen Fertigniveau auch andere Langzeitgedächtnisstrukturen, sogenannte „templates“, benutzen. Im IBLT-Modell bezieht sich diese Theorie auf die Lernmechanismen „recognition-based retrieval“ und „adaptive strategies“.

Das instance-based recognition model von Hintzman (1984, 1986, zitiert nach Gonzalez et al., 2003) ist ein Gedächtnisspurenmodell. Eine Gedächtnisspur beinhaltet darin eine Liste der relevanten Aufgabenmerkmale als „record of an experience“. Im Modell bezieht sich diese Theorie auf den Prozess des „recognition-based retrieval“.

Medin und Schaffner (1978, zitiert nach Gonzalez et al. 2003) schlagen eine Theorie zur Klassifikation von eingespeicherten Beispielen unter Berücksichtigung des Kontexts vor. Menschen lernen also Objekte zu klassifizieren anhand der abgelegten Beispiele im Gedächtnis. Wichtig hierbei ist das Konzept „Ähnlichkeit“, das von Nosofsky (1984, zitiert nach Gonzalez

et al., 2003) als mathematische Funktion beschrieben werden kann. Im Modell bezieht sich diese Theorie auf den „recognition-based retrieval“, die „adaptive strategies“ sowie auf evaluative Funktionen.

Die instance theory of automatization von Logan (1988) postuliert ebenso, dass Beispiele aus dem Gedächtnis abgerufen werden. Entscheidungen basieren zuerst auf einem (generellen) Algorithmus und später auf einem (spezifisch) instanzenbasierten Mechanismus. Mehrere Studien belegen, dass automatisches Handeln durch das Nutzen abgelegter Strukturen im Gedächtnis geschieht. Wenn wir also auf „Autopilot sind“, Fertigkeiten wie Auto fahren ausführen, werden Instanzen in einen „single-step direct-access retrieval Mechanismus“ abgerufen. Anders gesagt: direkter Gedächtnisabruf liegt automatisiertem Verhalten zugrunde. Die „enge“ Forderung von Logan, die abgelegten Instanzen müssen mit der angetroffenen Situation genau übereinstimmen, also „identisch sein“, wird von Nosofsky und Palmeri (1997, zitiert nach Gonzalez et al. 2003) geändert zu „ähnlich sein“. Das *Prinzip* dieser Theorie ist, dass sich im Laufe der Zeit eine Wissensbasis ansammelt, aber die grundlegenden Prozesse gleich bleiben. So geht Logan auch nicht von einer Begrenzung der Aufmerksamkeit aus, sondern bezieht die Veränderung in der Leistung auf die Veränderung der Wissensbasis in Größe und Qualität. Er nimmt an, dass (1) Instanzen obligatorisch und unvermeidbar abgelegt werden, der (2) Vergleichsprozess, ob ähnliche Instanzen vorliegen, obligatorisch abläuft und (3) die Repräsentation der Eindrücke durch Instanzen beschrieben werden kann. Gonzalez et al. (2003) führen die oben erwähnten SDU-Einheiten ein. Sie nehmen an, dass Wissensseinheiten die folgenden drei Informationen enthalten: *Situation*, also die Situation in der man sich befindet wird abgespeichert. *Decision* heißt, die verfügbaren Entscheidungsmöglichkeiten sind vermerkt und *Utility*, dass der Nutzen der resultierenden Entscheidung festgehalten und durch Übung modifiziert wird. Weiterhin nimmt Logan an, dass (4) jedes Zusammentreffen mit einem Stimulus separat eingespeichert, abgelegt und abgerufen wird. Der (5) Lernmechanismus zeichnet sich durch ein Akkumulieren von episodischen Gedächtnisspuren durch Erfahrung (von algorithmischer Verarbeitung zu gedächtnisbasierter) aus. Es findet (6) quasi ein Wettrennen zwischen dem Algorithmus und dem Abrufprozess statt. Dies äußert sich in einer immer schnelleren Verarbeitung und einer Reduktion der Variabilität – die Entscheidungen werden immer spezifischer, die instanzenbasierte Verarbeitung gewinnt die Oberhand. Logan und auch Gonzalez et al. zeigen, dass die Variabilität mit zunehmender Übung abnimmt und auch die Entscheidungen immer schneller ablaufen. Hypothese (3) bezieht sich eben auf diese Abnahme.

Logan macht Vorschläge, wie diese erste Version seiner Theorie noch verbessert werden könnte, zum Beispiel indem vorhandene Instanzen miteinander vermischt werden. Dies wird in Gonzalez et al.'s Modell aufgegriffen und umgesetzt. So werden die SDU's angepasst, wenn der Nutzen einer Entscheidung höher oder niedriger bewertet und dann abgespeichert wird. Die Aussage seiner Theorie, umschrieben mit: „Übung macht den Meister“, wird von Gonzalez et al. (submitted) geändert zu der Aussage „nur die *richtige* Übung macht den Meister“ – zumindest beim Entscheiden beim komplexen Problemlösen (vgl. auch Hypothese (1)).

Die Automatisierung wird nach Logan nie komplett sein, da jede zusätzliche Instanz das Gedächtnis ändert, wie auch im IBLT-Modell dargestellt. Hilfreiche Faktoren für Automatisierung sind nach Logan, wenn die betreffenden Stimuli in (1) geringen Mengen vorhanden sind, wenn sie (2) einfach zu merken sind und eine (3) große Anzahl von Durchgängen, in denen man ihnen ausgesetzt ist, vorliegen. Im IBLT-Modell bezieht sich die instance theory of automatization auf die Mechanismen „adaptive strategies“ und evaluative Funktionen.

Inwieweit das IBLT-Modell im Rahmen komplexen Problemlösens Sinn macht, soll in dieser Diplomarbeit überprüft werden.

Funke (2001) schlägt im Rahmen der Wissensaneignung und -anwendung das Zwei-Räume-Modell von Simon und Lea (1974, zitiert nach Funke, 2001) vor, welches eine Differenzierung des Problemraums möglich macht. Der Raum wird in einen Regelraum und einen Instanzenraum aufgeteilt. Ersterer enthält alle möglichen Regeln zu einem Problem, der zweite verschiedene Instanzen, also SDU-Einheiten. Die Raum-Metapher passt sehr gut und lässt sich ohne weiteres auch auf die Arbeiten von Gonzalez et al. (2003) übertragen. Vielleicht würden sie die Räume „Heurismenraum“ und „SDU-Raum“ nennen? Und der/die mit dem Problem konfrontierte durchschreitet im Laufe der Zeit einen Raum und kommt irgendwann in den anderen. Wenn er/sie wenig Zeit beim Durchschreiten hatte, konnte er/sie viel mitnehmen und startet das nächste Mal – in einer ähnlichen Situation – sofort im anderen Raum. Gonzalez et al. (2003) haben versucht herauszufinden, ob sie feststellen können, wann jemand durch die Tür von einem in den anderen Raum wechselt und quasi diesen Weg mit dem Computer simuliert.

2.3.2 Die Lernmechanismen

Simon and Langley (1981, zitiert nach Gonzalez et al., 2003, S. 592) definieren „Lernen“ als: „...a process that modifies a system as to improve, more or less irreversibly, its subsequent performance of the task or of tasks drawn from the same population.“ Sie schlagen aufgrund der vielen Wege, wie ein kognitives System beeinflusst werden kann, eine Theorie von Lernmechanismen vor, statt einer Lerntheorie. Diese fünf Lernmechanismen, die das Kernstück der Instanzenbasierten Lerntheorie bilden, werden bei Entscheidungsprozessen wirksam. Gleichzeitig postuliert die Theorie vier *Phasen*, die bei Entscheidungsprozessen durchlaufen werden. Die IBLT kann also als ein *Lernmechanismus-Paket* verstanden werden. Charakteristisch ist dabei der Bezug auf eine sich verändernde Wissensbasis (vgl. 2.3.1).

Die Lernmechanismen werden folgendermaßen genannt: (1) *instance-based knowledge*, d.h. das Anhäufen einer Wissensbasis in Entscheidungsprozessen und, (2) *recognition-based retrieval*, was bedeutet, dass Experten zielsicherer Informationen finden, unterstützt und gelenkt durch ihre Erfahrungen (die große Ähnlichkeit der abgelegten SDUs zu den vorliegenden Situationen), (3) *adaptive strategies*, was meint, dass zu Beginn Heuristiken verwendet werden, um zu einer Entscheidung zu gelangen, aber später (mit Übung) eine Anpassung erfolgt, bei der mehr auf Erfahrung – auf *die abgelegten Instanzen* – zurückgegriffen wird. (4) *necessity-based choice*, eine Methode, die Suche nach Alternativen (nach Nutzen oder Notwendigkeit) zu kontrollieren, und (5) *feedback updates*, um Kausalbeziehungen zwischen Entscheidungen und Ergebnissen zu konstruieren und um den Nutzen (utility) einer SDU zu modifizieren.

2.3.3 Ein Modell für Entscheidungsprozesse

Die vier Phasen im Entscheidungsprozess, die von Gonzalez et al. ausgeführt werden (die Phase „Execution“ wird bei ihr nicht weiter erläutert), (Abbildung 1) sind:

1. Recognition
2. Judgement
3. Choice

4. Feedback

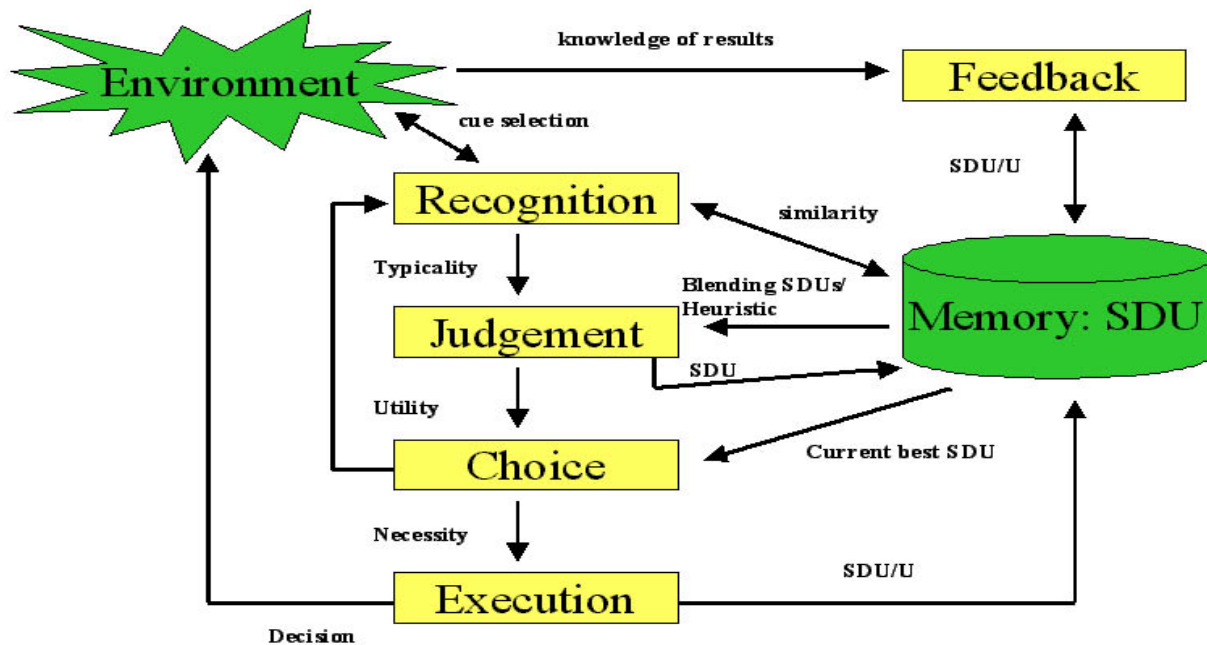


Abbildung 1: Das IBLT Prozessmodell

Ad (1) *recognition*: Der Prozess beginnt mit dem Suchen nach möglichen Alternativen in der Umwelt. Diese sind nicht offensichtlich gegeben. Die recognition-Fähigkeit entwickelt sich mit der Zeit. Experten suchen relevante Information gezielter und erfolgreicher. Dies geschieht durch Chunking, dem Zusammenfassen von Arbeitsschritten, oder „templates“, eine Art Langzeitgedächtnisstrukturen. Die information-reduction Hypothese von (Haider und Frensch, 1996, zitiert nach Gonzalez et al., 2003) besagt, dass im Laufe der Zeit mit zunehmender Übung Information selektiver genutzt wird und relevante Faktoren einer Aufgabe besser erkannt werden. Es wird in diesem Schritt, kurz gesagt, die Typikalität der Situation beleuchtet und ein Abgleich mit schon vorhandenen SDUs vorgenommen.

Ad (2) *judgement*: Nachdem eine Situation als typisch oder untypisch eingestuft worden ist, wird die Angemessenheit einer möglichen Handlung evaluiert. Heuristiken kommen bei atypischen Situationen zum Zug und auf typische Situationen kann mit vorhandenem Wissen, sprich Instanzen, reagiert werden. Der Nutzen einer möglichen Handlung wird durch Heuristiken auf der einen Seite oder auf der anderen durch den Nutzenwert aller relevanten ähnlichen Instanzen geschätzt. Die verschiedenen Nutzenwerte werden dann kombiniert in einer neuen Einschätzung der vorliegenden Situation. Eine mögliche Heuristik ist hierbei auf die vorhandene Zeit als Handlungskriterium zu achten. Schon in dieser Stufe kann also die

unabhängige Variable „Zeitdruck“ Auswirkungen zeigen (vgl. Hypothese (1)), indem sie zu einem Aufbau von qualitativ schlechteren SDU-Einheiten führt.

Ad (3) *choice*: In diesem Schritt wird der beste weitere Verlauf der Handlungen bestimmt. Nach jeder Alternative wird entschieden, ob gehandelt nun oder weiter evaluiert wird. Dies hängt vom Notwendigkeitslevel (necessity) ab. Wenn wenig Zeit zur Verfügung steht, wird schneller evaluiert (vgl. auch Kerstholt, 1994; Benson & Beach, 1996) und auch unsauberer. Auch hier hat die ausgewählte unabhängige Variable „Zeitdruck“ Einfluss. In rationalen Modellen sind alle möglichen Alternativen bekannt. Dies passt nicht zum Bearbeiten eines komplexen Problemlöseszenarios, in welchem dem/der Teilnehmer/in sozusagen per Definition viele Dinge unbekannt sind. Die IBLT geht davon aus, dass eine mittlere Strategie zwischen optimalem Vorgehen und befriedigendem Vorgehen bei Entscheidungen gewählt wird. Nur bei genügend Zeit werden alle möglichen Alternativen in Betracht gezogen. Das Gefühl für den richtigen Entscheidungszeitpunkt, erwächst mit der Zeit.

Ad (4) *feedback*: Um die Leistung zu verbessern, müssen die Ergebnisse der Entscheidungen beachtet werden. Feedback hilft, die SDU-Einheiten zu verbessern, sie neu zu kombinieren und zu verfeinern. Dadurch wird Expertise erreicht. So wird der Erfolgswahrscheinlichkeitswert der Instanzen mit Übung verbessert. Oft ist dies allerdings schwierig, da Feedback in komplexen Problemlösesituationen verzögert ist. Auch in diese Stufe wirkt die unabhängige Variable ein.

2.3.4 Zusammenfassung

„Wer entdeckt, wie Erinnerungen wirklich fixiert werden, verdient immer noch den Nobelpreis“, sagt Hans Markowitsch (Speiser, 2004). Die Instanzenbasierte Lerntheorie liefert ein mögliches Modell für das Anlegen von Wissen und dessen Anwendung, welches aber natürlich auch seine Schwächen hat. In jedem Fall dient es aber einer Annäherung an Forschungsfragen rund um das Forschungsgebiet „Gedächtnis“ und scheint insbesondere beim komplexen Problemlösen gut anwendbar zu sein. Zusammenfassend lässt sich sagen, dass die IBLT einen Ansatz für den Umgang mit Erfahrung beim Interagieren mit komplexen Systemen darstellt. Entscheidungsfindungen werden durch Wissen – dargestellt als akkumulierte Instanzen – ermöglicht. Dies geschieht nach einer ersten Phase, in der Heuristiken verwendet werden. Später, mit der Expertise und Erfahrung fußen Entscheidungen auf Instanzen, welche in SDU's (situation (Merkmale der Situation), decision (Entscheidungsmöglichkeiten), utility (Nutzen von Ergebnissen nach Entscheidungen)) beschrieben werden. Die Ähnlichkeit

zwischen Situationen in komplexen Systemen steigt mit Übung an und somit auch die Erfahrung. Auf diese Art und Weise werden Entscheidungen präziser und ähnlicher zu früheren Entscheidungen (Varianzverringern) in vergleichbaren Situationen und kognitive Prozesse laufen schneller ab (speed-up). Anders gesagt: „Denkgleise“ werden initiiert. Vertrautes gewinnt an Bedeutung, Altbewährtes wird verwendet, Alternativen werden je nach Dringlichkeit, Nutzen oder Verfügbarkeit von Zeit weniger gesucht und probiert. Es geschieht also eine Mischung zwischen optimierender und befriedigender Entscheidung. Durch Feedback kann Wissen verfeinert und mit größerem Nutzen eingesetzt werden. Die vorliegende Arbeit versucht diesen Prozess zu stören, um das IBLT-Modell zu falsifizieren.

Sieht man Lernen als Erweiterung von unserem Wissen über die Welt (vgl. Fischer, 2001; Fischer, (in press)), so erscheinen die Annahmen der IBLT und das dargestellte Modell auch von einem philosophischen Standpunkt plausibel. In der Mathematik werden ebenso instanzbasierte Modelle diskutiert (Aha, Kibler & Albert, 1991; Luger, Lewis & Stern, 2002), die mit der IBLT eng verknüpft sind. So stellen zum Beispiel auch Jungermann et al. (1998) eine sehr gute Übersicht über in der Psychologie untersuchte Entscheidungsregeln zusammen (S. 126), die letztendlich auf mathematischen Algorithmen beruhen. Diese enge geschichtliche Verknüpfung zwischen Entscheidungsforschung und Mathematik wird auch von Retzer (2002) diskutiert. Die Verwendung des Instanzbasierten Modells beim Treffen von Entscheidungen und Ausbilden von Fertigkeiten hat also im Bezug zu komplexen Szenarien einige Vorzüge. Interessante Ansätze zur Einbettung der Instanzbasierten Theorie in die Wahrscheinlichkeitstheorie (vgl. Hüllermeier, Dubois & Prade, 2001) lassen einen gespannt sein, welche weiteren Entwicklungen dieser Ansatz in der Zukunft bestreiten wird.

2.4 Die Water-Purification-Plant (WPP)

Das in dieser Diplomarbeit verwendete komplexe Problemlöse-Szenario zeichnet sich dadurch aus, dass es nur eine abhängige Variable hat: die Anzahl der verpassten Wassermenge (mised gallons). Andere Daten, wie die Zeiten, während der die Pumpen aktiv sind und während der sie stillstehen oder geputzt werden, werden in ein Log-File geschrieben. Es werden auch die missed gallons pro Deadline berechnet und es besteht die Möglichkeit, einen Durchgang im Replay-Modus – auch bei erhöhter oder verlangsamter Geschwindigkeit – nochmals anzuschauen.

Durch diese Programmierung der WPP-Aufgabe wird die Schwierigkeit der Interpretation von effizientem Problemlöseverhalten bei verschiedenen möglichen Ziele (Politelie) umgangen. Andererseits kann die vorhandene hohe Zielspezifität kritisch gesehen werden (vgl. Vollmeyer & Funke, 1999; Funke, 2001). Leider wird so die Auswertung ergebnisorientiert und nicht prozessorientiert (vgl. Strohschneider & Güss, 1999; und Dietrich Dörners Kritik an den verwendeten Methoden in der komplexen Problemlöseforschung, 1999). Eine genauere Analyse ist aber durch die vorbildliche Programmierung der WPP möglich, da man sich die einzelnen Abläufe nochmals in Zeitlupe oder Zeitraffer anschauen kann. Dies wird im explorativen Teil der Datenanalyse (4.4) anhand der missed gallons pro Deadline untersucht. Der Weg zum Ziel, die Entscheidungen, die getroffen werden, die Art wie gelernt wird, die Bearbeitung der WPP und der Bezug der Leistungen bei der WPP zu ausgewählten Persönlichkeitsmerkmalen, stehen im Mittelpunkt.

Abbildung 2 (S. 35) zeigt einen Screenshot der WPP. Gonzalez et al. beschreiben die WPP als eine „resource allocation task“. Sie wurde als Isomorph einer „real world“-Aufgabe im Bereich der Logistik, wie beispielsweise bei der Post, entworfen. Die Aufgabe kann schnell beschrieben werden (die Anleitung für die Teilnehmer/innen findet sich in Anhang B) und die Dauer des Durchgangs kann eingestellt werden. Es wird ein System zum Verteilen von Wasser dargestellt. Es gibt Tanks und Rohrleitungen, die diese Tanks verbinden. Der/die Teilnehmer/in verteilt das Wasser durch das Öffnen und Schließen der Pumpen per Mausclick. Das System besitzt, dadurch, dass zusätzliches Wasser von außen in das System fließen kann und auch einfach in Tanks gefüllt wird, ohne dass der/die Teilnehmer/in davon vorher bescheid weiß, dynamische Eigenschaften. Die dynamischen Komponenten können alle vorher festgelegt werden. So können auch verschiedene Schwierigkeitsgrade – neben der Erhöhung der Geschwindigkeit – modelliert werden. Genauer: es können Zeit, Wassermenge und der Zieltank für das zusätzliche Wasser definiert werden. Die maximale Anzahl Pumpen, die gleichzeitig geöffnet sein kann, beträgt fünf. Somit sind Entscheidungen innerhalb der WPP miteinander verknüpft. Wenn man an einer bestimmten Stelle eine Pumpe öffnet, geschieht dies zu Lasten einer anderen Stelle. Die Pumpen stellen also eine knappe, zu verwaltende Ressource dar. Der Hauptleistungsindikator ist – wie oben beschrieben – die verpasste Wassermenge (missed gallons). Das ist das Wasser, das nicht durch das Leitungssystem gepumpt werden konnte, bevor die jeweiligen Deadlines erreicht wurden. Nach jeder Deadline werden die momentanen missed gallons angezeigt und upgedated, bis zur letzten Deadline um 10:00 Uhr.

Die beste mögliche Leistung, die erreicht werden kann, wäre 0 missed gallons und die schlechteste 1080 missed gallons. Eine optimale Lösung ist durch eine bestimmte Folge von Entscheidungen des Öffnens und Schließens der Pumpen anhand der gegebenen Möglichkeiten gekennzeichnet. Eine vernünftige Leistung haben Gonzalez et al. (submitted) durch das Programmieren des „zero intelligence schedulers“ festgelegt. Hierbei werden Pumpen nach dem Zufallsprinzip geöffnet und geschlossen aber jeweils so, dass immer fünf Pumpen aktiv sind. Der genaue erhaltene Mittelwert bei 30 Wiederholungen ist 182,9 missed gallons (SD = 28,4); somit liegt eine vernünftige Leistung zwischen 0 und 200 missed gallons.

Die Simulation läuft von 2:00 Uhr bis 10:00 Uhr in simulierter Zeit. Die Geschwindigkeit wird vorher eingestellt. Die 23 miteinander verbundenen Tanks sollten möglichst kein Wasser mehr beinhalten, nachdem alle Deadlines abgelaufen sind. Für die Berechnung der Deadlines wird sowohl die Position des Tanks (chain value) herangezogen, als auch die darin verbliebene Wassermenge. Rot markierte, still stehende Pumpen können durch Anklicken aktiviert und durch erneutes Anklicken deaktiviert werden (maximal fünf gleichzeitig). Wenn eine Pumpe abgeschaltet ist – was auch passiert, wenn das Wasser im Tank davor leer wird – wechselt ihre Farbe zu gelb. Die gelbe Phase gibt an, dass die Pumpe gereinigt wird, was 10 Minuten simulierter Zeit in Anspruch nimmt. Wird während dieser Zeit eine andere Pumpe aktiviert, wechselt deren Farbe nach lila und wenn die, die gerade gereinigt wird, tatsächlich abgeschaltet ist, nach rot.

Dem Szenario können im Sinne von Funke (2001) die Eigenschaften Dynamik und Vernetztheit zugesprochen werden, die beiden entscheidenden Charakteristika eines komplexen Problems (vgl. auch 2.2.2).

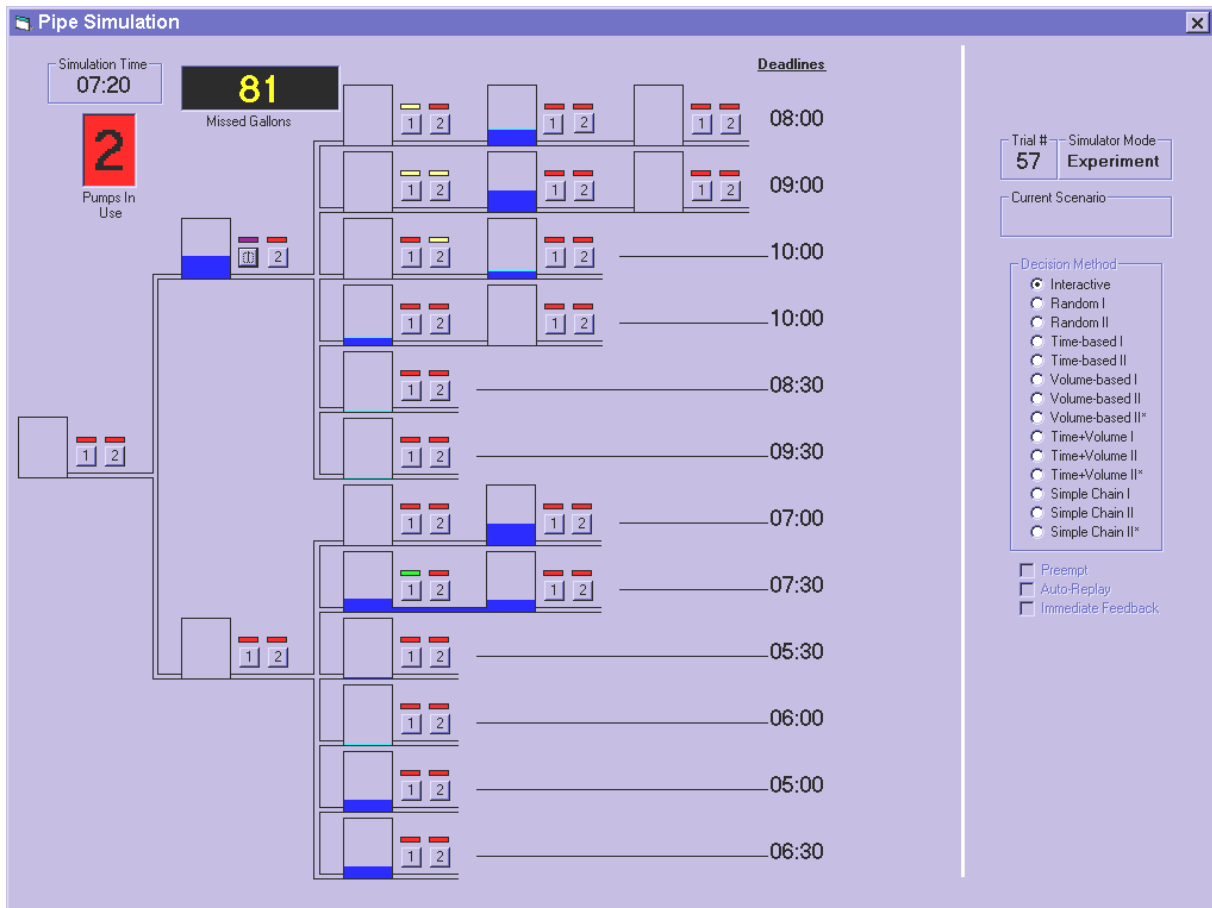


Abbildung 2: Screenshot der Water Purification Plant-Aufgabe

2.5 Auswirkungen von Zeitdruck in Entscheidungssituationen

Gonzalez et al. (2003) definieren Zeitdruck als “difference between the amount of available time and the amount of time required to solve a decision task.” Zeitdruck ist relativ in Abhängigkeit von der Stabilität – der Eigendynamik – eines Problemlöseszenarios zu sehen.

Ben Zur und Breznitz (1981) definieren Zeitdruck als Informationsmenge, die in einem bestimmten Zeitintervall beachtet und verarbeitet werden muss, oder als einen bestimmte benötigte Zeitdauer, um einen bestimmten Informationsbetrag zu verarbeiten.

Die Auswirkungen im Instanzenbasierten Modell sind in Abbildung 3 dargestellt. Der Aufbau von SDUs wird durch das Einwirken in die Prozesse bei der Stufe „Choice“ und „Feedback“ gestört (vgl. auch 2.3). Entscheidungen werden schnell, ohne Alternativen evaluieren zu können und ohne ausreichende Informationen zu haben, gefällt. Außerdem ist es unter Zeitdruck schwieriger, im Laufe des Prozesses über mehrere Durchgänge qualitativ gute SDUs aufzubauen, da kaum Zeit für Feedback, also das Verfeinern und Ändern vorhandener SDUs

bleibt. Folgen sind in diesem Modell schlechtere SDUs wenn Zeitdruck vorherrscht und damit (als Operationalisierung der SDUs) ineffizientere Entscheidungen und schlechtere Testleistung.

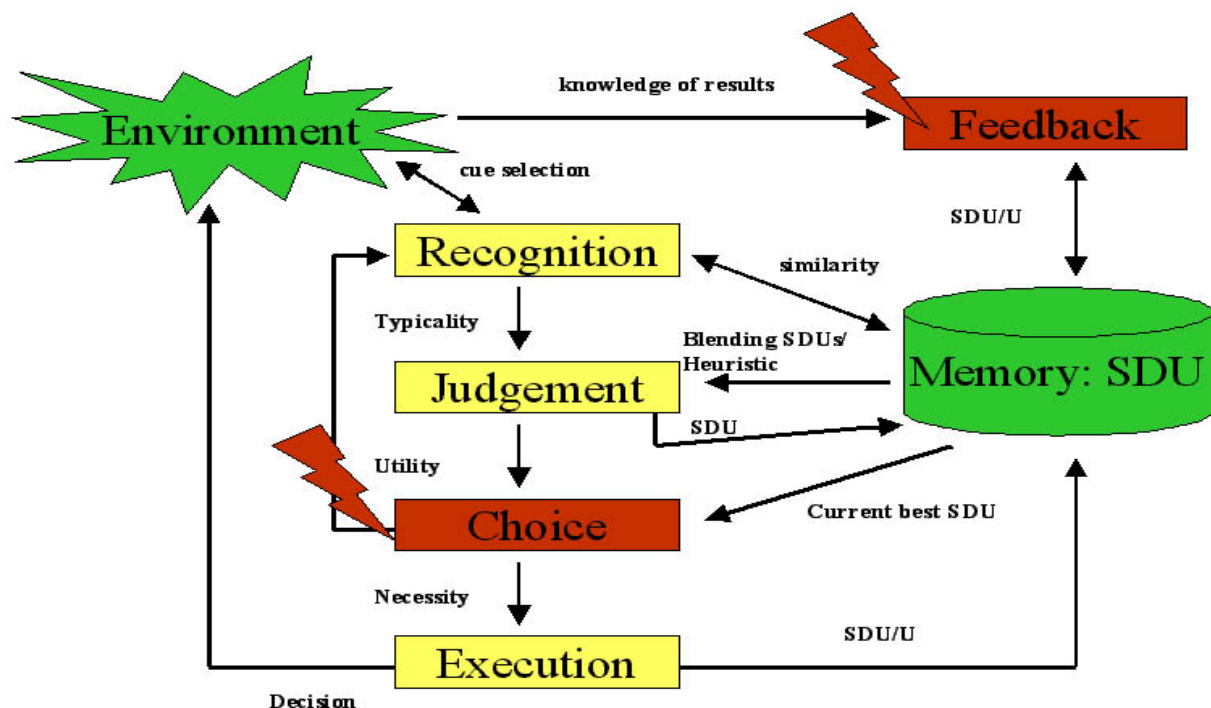


Abbildung 3: Die Auswirkungen von Zeitdruck auf das IBLT Prozessmodell

Benson und Beach (1996) betonen die Wichtigkeit, die Variable „Zeitdruck“ zu standardisieren. Sie wählen aus Pilottests einen Zeitabstand zwischen der normalen Bedingung und der Zeitdruckbedingung von einer Standardabweichung unterhalb des Mittelwertes. Das Design dieser Diplomarbeit lehnt sich an das von Gonzalez et al. (submitted) an und somit wird der dort verwendete Zeitabstand, der sich auf Basis von Pilottests entwickelt hat, übernommen. Im Umgang mit Zeitdruck bei Entscheidungen schlagen Benson und Beach vier Möglichkeiten zu reagieren vor. Diese sind (1) das Beschleunigen der Entscheidungsstrategie, (2) das Wechseln zu einer anderen Strategie bzw. Abkürzung der benutzten, (3) das Beibehalten der Art und Weise, wie entschieden wird und den Zeitdruck ignorieren und (4) die Entscheidungsstrategie nur ungenau ausführen. Im IBLT-Modell bedeutet dies den Aufbau ineffizienter SDUs.

Ben Zur und Breznitz (1981) schlagen drei verschiedenen Coping-Methoden in Situationen mit Zeitdruck vor: Avoidance, Acceleration und Filtration. Dies meint, dass Teilnehmer/innen unter hohem Zeitdruck eher Entscheidungen vermeiden, ihre Informationssuche beschleunigen und dass sie bei multiplen Entscheidungen subjektiv bedeutsame Information verwenden („conservatism under stress“, S. 102), während andere Information nicht in die Entscheidung miteinbezogen wird. Dies spricht dafür, dass Feedback im IBLT-Modell unter Zeitdruck kaum

verwendet werden wird. Ben Zur und Bresnitz postulieren einen U-förmigen Zusammenhang zwischen Zeitdruck und Leistung, d.h. es gibt einen Punkt des optimalen Zeitdrucks für eine bestimmte Leistung.

Rawlings und Carnie (1989) legen Eysenck's „arousal theory of extraversion“ (Eysenck, 1967, zitiert nach Rawlings & Carnie) zu Grunde und zeigen, dass eine Interaktion zwischen Zeitdruck und Persönlichkeit beim Lösen von vier Subtests aus dem WAIS-R besteht. Extravertierte zeigen eine höhere Leistung, wenn sie Zeitdruck empfinden. Genau umgekehrt zeigen Introvertierte eine höhere Leistung, sobald sie geringen Zeitdruck wahrnehmen (vgl. Hypothese (2)). Dies führen die Autoren auf das Yerkes-Dodson-Gesetz – der umgekehrt U-förmigen Kurve – zurück, das besagt, dass optimale Leistung bei einem mittleren Maß von Erregung oder Stress möglich ist. Bei zuviel oder auch zu wenig Stress sind Menschen weniger leistungsstark. Eysenck relativiert diese Aussage mit seiner Theorie, indem er Unterschiede zwischen extravertierten und introvertierten Menschen in Bezug auf die Stressverarbeitung mit einbezieht. Zu einem ähnlichen Ergebnis kommen Revelle, Amaral und Turriff (1976), die die Interaktion zwischen Zeitdruck, Extraversion und Introversion sowie Koffein im Bezug auf die verbale Leistung eines Tests, des GRE (Graduate Record Examination), untersuchen. Ihr Ergebnis steht in Einklang mit der obigen Feststellung, dass extravertierte Menschen durch Zeitdruck und Koffein (sprich mehr „Arousal“) zu besseren Testleistungen gelangen können, während introvertierte Menschen ohne Koffein und ohne Zeitdruck bessere verbale Leistungen beim GRE zeigen.

Van der Linden et al. (2001) stellen die Bedeutsamkeit von Reflexion über das Geschehene heraus. Wenn die Möglichkeit zur Reflektion verhindert wird – wie durch Zeitdruck in der vorliegenden Studie – ist das Lernen weniger effizient. Im weiteren wird keine systematische Exploration eingesetzt, sondern es werden weniger nützliche Strategien im Umgang mit komplexen Problemen verwendet. Zum Beispiel die Versuch-und-Irrtums-Strategie, die rigide Exploration, das Verbleiben bei nutzloser Informationssuche – um ein Kompetenzgefühl zu bewahren – oder das Auftreten negativer Selbstbeurteilungen. Das Verwenden dieser Strategien wird in Zusammenhang mit ungenügender Planung und schlechter Feedback-Nutzung gebracht (vgl. dazu Abschnitt 4.5). Es ist sehr wahrscheinlich, dass die Teilnehmer/innen in der vorliegenden Arbeit weniger nützliche Strategien verwenden und es so zu schlechteren Ergebnissen in der „fast“-Bedingung – mit hohem Zeitdruck und wenig Möglichkeit zur Selbstaufmerksamkeit und hier zur Feedbackanalyse und zur erschöpfenden Überprüfung aller Alternativen – kommt (vgl. Hypothese (1)). Weiterhin wird unter Zeitdruck das wichtige

Lernen aus Fehlern und deren Konsequenzen erschwert („*To learn from an error it is necessary to process the information that the error can provide.*“, Van der Linden et al., S. 191). Auch das Herausfinden von Kausalbeziehungen – und somit der Aufbau adäquater mentaler Modelle – ist schwer möglich (vgl. auch Brehmer, 1992, 1996). Eine wichtige Rolle spielen hierbei die negativen Selbstbewertungen und demzufolge auch die Ausprägung des Persönlichkeitsmerkmals „Neurotizismus“, also der emotionalen Labilität. Je emotional stabiler, umso besser wird man in Situationen mit Zeitdruck zurecht kommen (vgl. Hypothese (4)). Des Weiteren besprechen Van der Linden et al. die Wichtigkeit des Ausprobierens von verschiedenen Möglichkeiten beim Umgang mit komplexen Problemen. Dies spricht für einen positiven Zusammenhang zwischen der Leistung in der WPP-Aufgabe und einem hohen Wert auf dem Persönlichkeitsmerkmal „*Offenheit für Erfahrungen*“.

2.6 Zur Auswahl der Kovariaten und Logik des Designs

Die Lohhausen-Studie (vgl. Dörner, 1989) war eine der ersten Studien, die Zusammenhänge zwischen Persönlichkeitsmerkmalen und komplexen Entscheidungssituationen untersucht hat. Die vorliegende Diplomarbeit schließt in gewisser Weise an deren Ergebnisse an, unterscheidet sich aber durch einige Merkmale. Hier liegen Entscheidungsprozesse und die Auswirkungen von Zeitdruck im Lichte der Instanzenbasierten Lerntheorie sensu Gonzalez et al. (2003) im Fokus. Es wird versucht, Persönlichkeitsmerkmale mit kognitiven Variablen im Sinne von Omodei und Wearing (1995) zu verquicken.

Einen Überblick über Fünf-Faktoren-Modelle der Persönlichkeit – die auch die Grundlage des NEO-FFI bilden – geben zum Beispiel Amelang und Bartussek (1997) und es lassen sich dort folgende Facetten für die drei ausgewählten Dimensionen finden:

- *Neurotizismus* meint das Gegenteil von emotionaler Stabilität. Menschen mit hohen Werten auf dieser Dimension können als ängstlich, gespannt, reizbar, emotional und labil beschrieben werden. Menschen mit niedrigen Werten hingegen, als zufrieden, ruhig und stabil.
- Menschen mit hohen Werten auf der Dimension *Extraversion* sind eher gesprächig, bestimmt, aktiv und sozial. *Introvertierte*, also Menschen mit niedrigem Score auf dieser Dimension, eher reserviert, scheu und still.

- Die Dimension *Offenheit für Erfahrungen* wird charakterisiert durch Attribute, wie interessiert, phantasievoll, intelligent und geschickt. Menschen mit niedrigen Werten auf dieser Dimension werden als gewöhnlich, einfach und ohne Tiefgang beschrieben.

Fünf-Faktoren-Modelle sind das Ergebnis mehrerer Dekaden faktorenanalytischer Forschung im Zusammenhang mit der Trait-Persönlichkeit. Ursprünglich wurde das Modell von Galton (1884, zitiert nach Zhang, 2003) vorangetrieben. Es hat seine Wurzeln im psycho-lexikalischen Ansatz. Fünf Dimensionen, die heute als „Big Five“ bekannt sind, dienen als guter Rahmen die Persönlichkeit zu beschreiben und sind Bestandteil vieler Theorien. Die Dimensionen haben ihrer Namen im Laufe der Zeit etwas geändert. Am gebräuchlichsten scheinen die Arbeiten von Costa und McCrae (1992) zu sein. Der NEO-FFI ist mit insgesamt 60 Items eine Kurzform des NEO Personality Inventory. Er wurde von Borkenau und Ostendorf (1993) für deutschsprachige Länder angepasst.

Nach Li-Fang Zhang (2003, S. 1433) – der untersucht, ob Lernfortschritte durch die „Big Five“ vorhergesagt werden können – wurde die Rolle der „Big Five“ für das Lernen bisher nicht genügend gewürdigt. Er sieht durch deren Erforschung Möglichkeiten, Lernenden zu neuen Einsichten und zu angenehmerem Lernen zu verhelfen. Zhang (2002) untersucht in seinem Artikel Zusammenhänge zwischen Denkstilen und den „Big Five“. Er findet Korrelationen zwischen einem Fragebogen zu Denkstilen (Thinking Styles Inventory) nach einem neuen Modell von Sternberg (mental self-government model) und dem NEO-FFI. Er schlussfolgert aber, dass jedes der beiden Verfahren seinen eigenen Bereich abdeckt. Extravertierte scheinen in seiner Studie – wie bei Ergebnissen anderer Studien auch – eher Aktivisten zu sein, während Introvertierte eher reflektieren. Dies wird bei den Hypothesen berücksichtigt.

Zhang unterscheidet zwischen *Stilen* (wobei er zwischen Lern-, Kognitiven und Denkstilen trennt) und *Fähigkeiten* (wie z.B. Größe des Arbeitsgedächtnisses). Stile beziehen sich auf unsere bevorzugte Art, unsere Fähigkeiten zu nutzen. Dies würde sich interessanterweise mit Kreuzig's Definition von prozesshaften Persönlichkeitsmerkmalen weitgehend decken (vgl. S. 40). Stile sind am besten an der Schnittstelle zwischen Intelligenz und Persönlichkeit eingeordnet (Sternberg, 1994, zitiert nach Zhang, 2002). Dies unterstreicht den Nutzen einer Verwendung von Persönlichkeitsdimensionen mit komplexem Problemlösen in dieser Diplomarbeit. Sie ist ein Versuch interessante Zusammenhänge zu finden, welche die Bedeutung von Persönlichkeitsdimensionen in der kognitiven Forschung stützen (vgl. Süß, 1999; Kreuzig, 1981) und so zu einem umfassenderen Gesamtbild des Menschen im Bereich

des komplexen Problemlösens führen. Die Arbeitsgedächtniskapazität wurde in den Studien von Gonzalez et al. (submitted) und Gonzalez (submitted) – im Unterschied zur vorliegenden Arbeit – miteinbezogen.

Als mögliche Anwendung seiner Ergebnisse bringt Zhang an, dass Menschen mit hohen Werten auf der Dimension „Neurotizismus“ offenbar strukturiertere Lernmethoden brauchen, als Menschen mit niedrigen Werten. Solche Ergebnisse könnten in Zukunft genutzt werden, um zum Beispiel Computeroberflächen (Fahrkartenautomaten, Regelsysteme in Kraftwerken oder Kläranlagen) benutzerfreundlicher anzupassen. Es wäre interessant, inwieweit die in dieser Studie überprüften Hypothesen mit dem Forschungsgebiet der kognitiven Stile zusammenhängen. In den Artikeln von Peterson, Deary und Austin (2003a), Riding (2003) und Peterson, Deary and Austin (2003b) wird aber deutlich, dass es noch keine Einigung über ein Instrument zur Erhebung von kognitiven Stilen gibt. Ansätze sind jedoch vorhanden. In den drei genannten Artikeln wird der Riding's cognitive style analyses test – insbesondere seine Reliabilität – sehr kontrovers diskutiert.

Kreuzig (1981) nimmt in seinem Artikel Bezug auf die Lohhausen-Studie von 1980 und unterstreicht die Bedeutung des Einbeziehens von Persönlichkeitsmerkmalen in die kognitive Forschung. Diese sollten aber prozessorientiert sein, aufgrund des ebenso prozesshaften Ablaufs kognitiver Aktivitäten. In der vorliegenden Studie werden die Persönlichkeitsmerkmale nur am Anfang erhoben. Somit werden Vergleiche zu den Arbeiten von Gonzalez et al. möglich. Bei der Bestimmung der Leistungsgüte in der Lohhausen-Studie gab es einige Probleme (z.B.: Eine Vielzahl möglicher Zielzustände), was hochkomplexe Probleme auch zu einem schwierigen Forschungsgegenstand macht. Noch liegt auch keine umfassende Taxonomie von Problemtypen vor (Funke & Hussy, 1979, zitiert nach Kreuzig; Funke, 2003), was zur Folge hat, dass Studien zur Validierung der Befunde via Korrelation (also der Orientierung an anderen bewährten Konzepten, wie dem Intelligenzkonstrukt zum Beispiel) konsistent keine bedeutenden Befunde erbracht haben. Doch die Anzahl der Forschungsarbeiten auf diesem Gebiet gedeiht und erfolgsversprechende Tendenzen (vgl. Süß, 1999) erwachsen daraus.

Kreuzig (1981) definiert Persönlichkeitsmerkmale nicht als Traits oder States, sondern als *häufig wiederkehrender relativ überdauernder Einsatz gleicher kognitiver Prozesse*. Diese können dann die Stabilisierung bestimmter Prozesse bewirken. Man kann davon ausgehen, dass die Erkenntnisse zu kognitiven Stilen auch in der komplexen Problemlöseforschung spannende Anwendung finden würden (vgl. Zhang 2002, 2003). Vorwissen im Sinne von abstrakten

Schemata, die durch Neugier und Training entstanden sind (man könnte sagen: auch durch *Offenheit für Erfahrungen* bedingt), halten Kreuzig und auch einige andere Autoren (z.B. Van der Linden et al, 2001) für einen der wichtigsten Prädiktoren der Leistungsgüte beim komplexen Problemlösen. Deshalb wird Computervorerfahrung in die vorliegende Studie miteinbezogen. Ebenso ist für Kreuzig die *willkürliche* Steuerung von divergenten versus konvergenten Denkprozessen entscheidend. In seiner Arbeit stellt er fest, dass zur Problemlösegüte tendenziell negative Korrelationen zur Persönlichkeitsdimension *Neurotizismus* bestehen (vgl. Hypothese (4)), zur *Extraversion* signifikant positive (vgl. Hypothese (2)) und zur *Selbstsicherheit* – die allerdings in dieser Diplomarbeit nicht erhoben wird – hoch signifikante, gemessen mit dem ENRL von Janke (1996, zitiert nach Kreuzig, 1981). Extraversion enthält für ihn Aspekte von *Offenheit für Erfahrungen*. Im Bezug auf Selbstsicherheit beschreiben sich in seiner Studie Frauen signifikant als weniger selbstsicher. Anhand dieser Ergebnisse kann in der folgenden Diplomarbeit geschlussfolgert werden, dass ein Unterschied zwischen Männern und Frauen im Bezug auf den Umgang mit der Problemlösegüte besteht (vgl. Hypothese (6)). Dies wird allerdings nur explorativ untersucht. Des weiteren erscheint die Auswahl der Kovariaten *Neurotizismus*, *Extraversion* und *Offenheit für Erfahrungen* im Lichte dieser Studie als eine adäquate Wahl. Eine Verwendung von prozessorientierten Persönlichkeitsmerkmalen kann also zugunsten der Vergleichbarkeit mit den Arbeiten von Gonzalez in dieser Diplomarbeit nicht geleistet werden, aber Implikationen für deren Nutzung und Einbau in ein ähnliches Design können für zukünftige Forschung abgeleitet werden (vgl. auch Abbildung 21, S. 112).

Omodei und Wearing (1995), die ein ähnliches Design zur vorliegenden Arbeit verwenden, diskutieren in ihrem Artikel ein theoriebasiertes Modell, welches Motivation, Affekte, Intentionen und kognitive Leistung integriert. Sie monieren:

„Research in the psychology of decision making has usually been more concerned with the cognitive processes persons employ than with the role of motivation and/or affect. This is partly because the latter variables are not seen as influencing the cognitive processes involved in the decision task studied. Researchers generally take motivation as given, and consider affect only as a factor which degrades decision making performance, but is otherwise irrelevant.“

Es war in der Vergangenheit nicht möglich, die Leistung in komplexen Aufgaben nur auf kognitive Merkmale zurückzuführen. Deshalb beinhaltet ihr Modell auch verschiedene nicht-kognitive Einflussfaktoren, wie: Motivation, Intention, Aufmerksamkeit, Langeweile, Involvement, Angst, Arousal sowie Feedback. Feedback (*„a person's direct cognitive evaluations of their progress in that environment“*) im Sinne vom Gewahrsein der Leistung in der komplexen Problemlöseaufgabe und dessen Auswirkungen auf Emotionen, Motivation und Kognitionen (vgl. Abbildung 4). Omodei und Wearing's Annahme ist, dass emotionsregulierte Prozesse zentrale Bedeutung bei der Handlungsregulierung in komplexen Problemlöseaufgaben haben. Das bedeutet, Emotionen bestimmen die Aufmerksamkeitskapazität und -richtung. Emotionen selbst und Informationsverarbeitungsressourcen werden wiederum durch psychophysiologische Mechanismen mediiert, die Arousal zugrunde legen. Hieraus leiten sich Konsequenzen für die in dieser Diplomarbeit vorgeschlagenen Hypothesen und für die unabhängige Variable „Zeitdruck“ ab. Intention steuert in ihrem Modell auch die Aufmerksamkeit. Diese Komponenten fehlen im IBLT-Modell, haben aber nach Omodei und Wearing große Relevanz zur Erklärung der kognitiven Leistung in einer komplexen Problemlöseaufgabe.

Die Bedeutung, die Omodei und Wearing Emotionen zumessen, wird in der vorliegenden Diplomarbeit nur durch das Einbeziehen des Persönlichkeitsmerkmals „Neurotizismus“ berücksichtigt. Emotionen haben einen informativen (kognitiven) Gehalt – wie z.B. bei Bower (1981, zitiert nach Omodei & Wearing, 1995) im Bezug auf Lerntheorien oder Schwarz (2000) ausgeführt ist – und sind somit bei der Steuerung einer komplexen Problemlöseaufgabe relevant. Emotionale Stabilität sollte also zu besseren Leistungen bei der Bearbeitung führen (vgl. Hypothese (4)). Dies gilt insbesondere, da Emotionen die Bemühungen, die man in eine Aufgabe weiterhin steckt, moderieren.

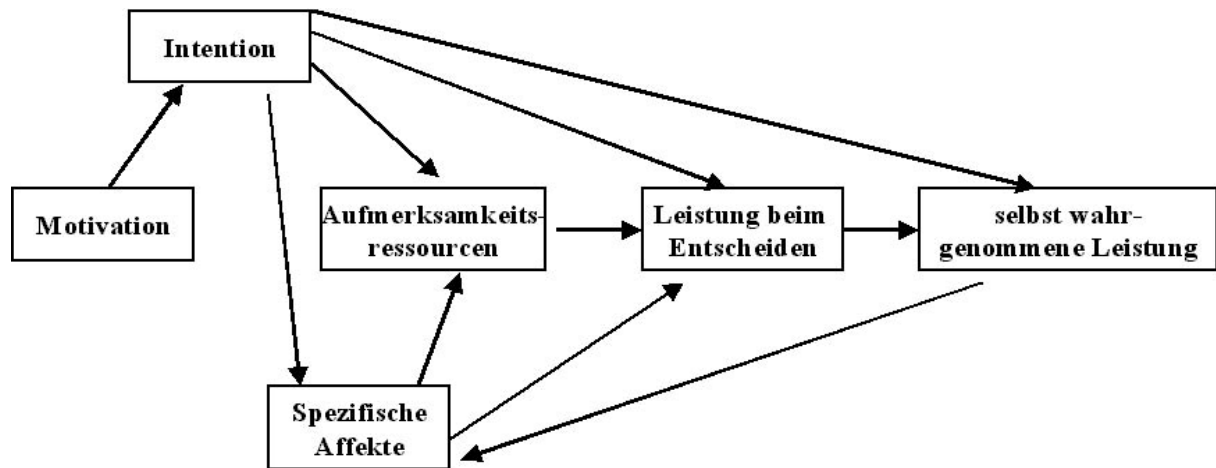


Abbildung 4: Das konzeptuelle Modell zu Leistung und Erfahrung beim Entscheiden in komplexen Problemlösesituationen von Omodei und Wearing (1995)

Die Schlussfolgerung, die Omodei und Wearing aus ihrer Studie ziehen, ist, dass hohes Involvement zu besseren Leistungen führt, wohingegen mehr Aufmerksamkeit nicht unbedingt bessere Leistungen nach sich zieht. Wichtig ist, ob die Aufmerksamkeit auf aufgabenrelevante Merkmale gelenkt wird. Die IBLT postuliert, dass dieses Finden der relevanten Merkmale im Laufe der Zeit mit zunehmender Übung in einer Aufgabe immer besser funktioniert. Omodei und Wearing meinen, dass es bisherigen Untersuchungen nicht gelungen ist, relevante Persönlichkeitsmerkmale herauszufinden, die mit guter Leistung in einer komplexen Problemlöseaufgabe einhergehen. Die vorliegende Diplomarbeit soll hierzu durch die drei ausgewählten Merkmale aus dem NEO-FFI einen Beitrag leisten.

In seiner Studie untersucht Blickle (1996) die Beziehungen zwischen Persönlichkeitsmerkmalen und Lernstrategien mit einem Pfadanalysemodell. Sein Argument ist, dass Lernstrategien und Anstrengung koordiniert werden müssen, um optimale Leistungsergebnisse zu erhalten, wie in ähnlicher Art und Weise auch Kreuzig (1981) – mit dem willkürlichen Wechseln zwischen divergenten und konvergenten Denken – betont.. Nach einem Modell von Mumford und Gustafson (1988, zitiert nach Blickle, 1996) können bestimmte Persönlichkeitsmerkmale das Nutzen von Strategien vereinfachen oder aber auch erschweren. Sie können die nötigen Impulse geben oder abblocken und schließlich auch mit darüber entscheiden, ob sich ein Mensch nach einem Fehlschlag von einer Aufgabe abwendet oder aber „am Ball bleibt“. Nach Ansicht des Autors der vorliegenden Diplomarbeit lassen sich hier Bezüge zur Persönlichkeitsdimension „Neurotizismus“ finden.

Für eine Kooperation zwischen differenziell-psychometrischer und kognitionspsychologischer Forschung spricht sich Süß (1999) aus. Er findet Zusammenhänge zwischen Intelligenz und der Problemlöseleistung der Probanden bei verschiedenen komplexen Problemlöseszenarien. Wichtig sei der große Einfluss von Vorwissen. Des Weiteren moniert Süß aber mehrere methodische Probleme in einigen Problemlösestudien, die aus verschiedener – auch kritischer – Perspektive anders gesehen werden können (vgl. Funke et al., 1999; Feyerabend, 1978, S. 246-247, 309-310, 364-367; Wagner 2001, S. 108ff). Viel mehr scheint sich gerade die komplexe Problemlöseforschung zur Erfindung neuer methodischer Ansätze anzuregen. Man kann sich zum Beispiel Verlinkungen mit der Chaostheorie (vgl. Zeemann, 1976; Kriz, 2000) oder neueren mathematischen Entwicklungen wie dem Indextheorem vorstellen. Auch Süß betont die Möglichkeiten einer gegenseitigen Bereicherung der oben genannten Gebiete in der Psychologie, was sich in interessanten integrativen Ideen zur Auswertung (vgl. Canas, Quesada, Antoli, & Fajado, 2003; Van der Linden et al., 2001) zeigen kann. Aus Ressourcengründen wird aber zumeist auf einfachere Modelle zurückgegriffen, die leicht Zielscheibe methodischer Kritik werden können. Die hohe ökologische Validität computersimulierter Szenarien beschreibt Süß gleichzeitig als deren „Achillesferse“, da sie das Vorwissen der Teilnehmer/innen am Experiment als kritische Varianzquelle für Leistungsdifferenzen, die nur statistisch kontrollierbar ist, mit einführt.

In seinem Buch über Systemkompetenz – also wie Menschen effizient mit (komplexen) Systemen umgehen – beschreibt Kriz (2000) verschiedene Voraussetzungen, die eben dafür nötig sind und verschiedene Planspiele, um Systemkompetenz zu trainieren. Darunter fallen auch Simulationen am Computer, die Menschen ein Gespür für komplexe Systeme und potentielle Fehlerquellen vermitteln können. Verschiedene Teilqualifikationen (z.B.: Umgang mit der emotionalen Dimension; Soziale Kontaktfähigkeit), die starken Bezug zu Persönlichkeitsmerkmalen haben, werden von Kriz erläutert und sprechen für das verwendete Design. Nach Kriz (vgl. Kriz, S. 149) ist der Prozess der Selbstreflexion einer der wichtigsten, um effizient mit Systemen umgehen zu können. Genau dies ist auch ein kritischer Punkt im Lernprozess bei der IBLT. Durch Zeitdruck soll die Möglichkeit zur Reflexion verhindert werden, was sich in schlechten Leistungen bei der Bearbeitung der WPP-Aufgabe niederschlagen soll.

Wenn man davon ausgeht, dass mit dem Erwerb von Wissen ein neurologisches Korrelat einhergeht, bestätigen die Befunde von Spitzer (1996) die Logik des Designs. Er merkt an, dass Lernen von komplexen Zusammenhängen eine bestimmte Zeit benötigt, die nicht verkürzt

werden kann und sollte. Evolutionspsychologisch gesehen ist eine wichtige Frage, warum Menschen mit einem noch nicht ausgereiften Gehirn auf die Welt kommen und sich dies dann entwickelt („extrauterine Embryonalentwicklung“). Spitzer, der über neuronale Netzwerke forscht, beschreibt dies, obwohl es auf den ersten Blick als ein Unglück wirkt, als notwendige Voraussetzung für das Erlernen komplexer Sachverhalte. Ein sich *entwickelndes* Gehirn, das mit komplexen Inputmustern konfrontiert wird, kann so erst die notwendigen Abstraktionsleistungen für eine gute innere Repräsentation vollziehen. Anders gesagt: komplexer Input, der nicht verarbeitet werden kann, stellt für ein System nur Rauschen dar. Das Lernen kann man sich vorstellen als das Herausfinden und -filtern von grundlegenden Mustern. Erst später ist es dann möglich, auch komplexen Input zu verarbeiten.

Ein weiterer Punkt, der für die Verwendung des Designs in der vorliegenden Arbeit spricht, findet sich in dem Artikel „Die Logik des Niedergangs“ von Probst und Raisch (2004). Sie diskutieren den Umgang von Managern mit komplexen Situationen, also dem Steuern von Unternehmen. Sie besprechen sowohl die kritische Rolle der Persönlichkeit des CEOs als auch die Bedeutung von gesundem Wachstum und Wandel, was Übertragungen auf das in der vorliegenden Diplomarbeit vorgestellte Design zulässt. Insbesondere bei der Einführung neuer Computerprogramme in Unternehmen müssen die Mitarbeiter/innen neu lernen und ihr Wissen ändern (vgl. Simon, 1993; Schein & Coutu, 2002, Schein, 2003). Geschieht dies angepasst an die Art, wie Menschen lernen, so kann die Organisation ihre Stabilität bewahren. In ihrem Beitrag erwähnen Probst und Raisch auch die große Bedeutung der Selbstkritik, die notwendig ist, damit Reflexion stattfindet und auch verwertet wird. Zusammengefasst spiegelt sich in verschiedenen Perspektiven die Logik des verwendeten Designs – also der Verwendung der Kovariaten aus dem NEO-FFI sowie der Variable „Computervorerfahrung“ und das Einbringen von Zeitdruck als unabhängige Variable zum Testen der IBLT – wider.

2.7 Forschungsfragen und Ziele

Das Ziel der vorgestellten Studie ist die Überprüfung des Zusammenhanges zwischen der Leistung in einer computersimulierten komplexen Problemlöseaufgabe (der Water-Purification-Plant) und ausgewählten Persönlichkeitsdimensionen unter zwei verschiedenen Zeitdruck-Bedingungen. Die Persönlichkeitsdimensionen sollen als Kovariaten mit in die Berechnungen einfließen und sind: „Neurotizismus“, „Extraversion/Introversion“ sowie „Offenheit für Erfahrungen“. Die Verwendung dieser drei Variablen lässt sich theoretisch und durch die

Unterstützung anderer empirische Studien begründen (vgl. Conelly et al., 2000, Humphreys & Revelle, 1984, Omodei & Wearing, 1995, Zhang, 2002, Zhang, 2003, Revelle et al., 1976, Blickle, 1996, Rawlings & Carnie, 1989, Ngidi & Sibaya, 2002, Kriz, 2000).

Die Auswahl der Kovariaten und des Designs wurde von den Studien von Gonzalez et al. (2003), Gonzalez et al. (submitted) und Gonzalez (submitted) abgeleitet. Die vorliegende Diplomarbeit prüft das von Gonzalez et al. (2003) vorgeschlagene Modell des Instanzenbasierten Lernens im Kontext des komplexen Problemlösens. Des weiteren werden ihre Ergebnisse kritisch hinterfragt und versucht durch das Einbringen von Persönlichkeitsmerkmalen als Kovariaten auszuweiten.

Es wird angenommen, dass sowohl hemmende als auch fördernde Zusammenhänge zwischen Persönlichkeitsmerkmalen und den Leistungen in einem komplexen computersimulierten Szenario bestehen. Diese können wertvolle Hinweise auf die Modifikation von Lehrgängen, bei denen es um die Ausbildung von Fertigkeiten (Routinen) geht, geben. Auch können Hinweise zum Einführen von neuen Computerprogrammen in Unternehmen erschlossen werden. Es soll versucht werden zu zeigen, dass zum Aufbau von Routinen eine gewisse Zeit vonnöten ist – also eine gewisse Übung, welche nicht beschleunigt werden kann, da sich ansonsten ein Lernniveau einpendelt, welches schwierig zu erweitern ist. Ein Beispiel dafür wäre das Lernen des Zehn-Finger-Systems. Lernt man es unter Zeitdruck, leidet die Qualität der späteren Leistung. D.h. das anfängliche typische Verwenden von Heuristiken muss sauber abgelöst werden können durch ein sorgfältiges Entstehen von SDUs (situation (Ist die Situation neu oder ähnlich zu einer schon gehabt?), decisions (Welche Möglichkeiten habe ich zu entscheiden?), utility (Hat sich die Entscheidung das letzte Mal bewährt/ ...tut sie dies auch dieses Mal?)). Dies sind die grundlegenden Ideen, die Gonzalez et al. in ihren Studien beschrieben haben. Man kann sich fragen, ob die Ergebnisse auch in Deutschland replizierbar sind und ob es alternative Erklärungsmöglichkeiten dafür gäbe.

2.8 Hypothesen

Somit lassen sich folgende Hypothesen formulieren:

Hypothesen über Zeitdruck

Hypothese 1: Ein ähnliches Ergebnismuster bezüglich der Leistung in der Water-Purification-Plant wie bei Gonzalez et al. 2003, Gonzalez, Qudrat-Ullah und Lerch (submitted), und Gonzalez (submitted) wird sich für die beiden Treatment-Gruppen ergeben.

D.h. die Teilnehmer/innen in der „slow“-Bedingung werden am zweiten Tag – wenn beide Gruppen unter Zeitdruck spielen – besser sein, als die Teilnehmer/innen der „fast“-Bedingung, da sie laut Gonzalez et al. am ersten Tag die effizientere Übung hatten. Außerdem wird die Interaktion der zwei Bedingungen mit den Kovariaten signifikant sein.

Hypothese 2: Es wird angenommen, dass sich das Persönlichkeitsmerkmal Extraversion als bedeutsame Kovariate herausstellen. Im Sinne von Van der Linden et al. (2001) kann angenommen werden, dass sich diese Persönlichkeitsdimension positiv auf ein effektives Explorationsverhalten auswirkt (vgl. auch Conelly et al., 2000).

Explorativ wird untersucht: Für das Persönlichkeitsmerkmal „Extraversion“ ergibt sich ein komplexer Zusammenhang im Bezug auf die Leistung in der WPP (vgl. Revelle et al., 1976; Humphreys & Revelle, 1984). Dieser komplexe Zusammenhang spiegelt sich vermutlich auch in den Ergebnissen von Zhang (2002, 2003) wider, der keine eindeutigen Korrelationen dieser Persönlichkeitsdimension mit Lernfortschritten und kognitiven Stilen finden konnte. In der „fast“-Bedingung sind Extravertierte besser, da ihnen ihr niedriges Aktivationsniveau (Arousal) hilft, genau die passende Aktivierung für gute Leistungen zu erhalten. Im Gegensatz dazu haben Introvertierte von vorneherein ein höheres Aktivationsniveau und profitieren deshalb von der „slow“-Bedingung am meisten, haben dort also die bessere Leistung (vgl. auch Rawlings & Carnie, 1989). Dies ist in Abbildung 5 schematisch dargestellt. Es kommt hierbei auch die Beziehung zum in 2.6 diskutierten Modell von Omodei und Wearing (1995) zum Ausdruck, die in einem pfadanalytischen Modell einen starken positiven Zusammenhang zwischen Arousal und Performance finden können. Die Persönlichkeitsdimension Extraversion/Introversion würde also im Sinne von Baron und Kenny (1986) als ein Moderator wirken, der je nach Ausprägung, die Leistung in der WPP erhöht oder verringert.

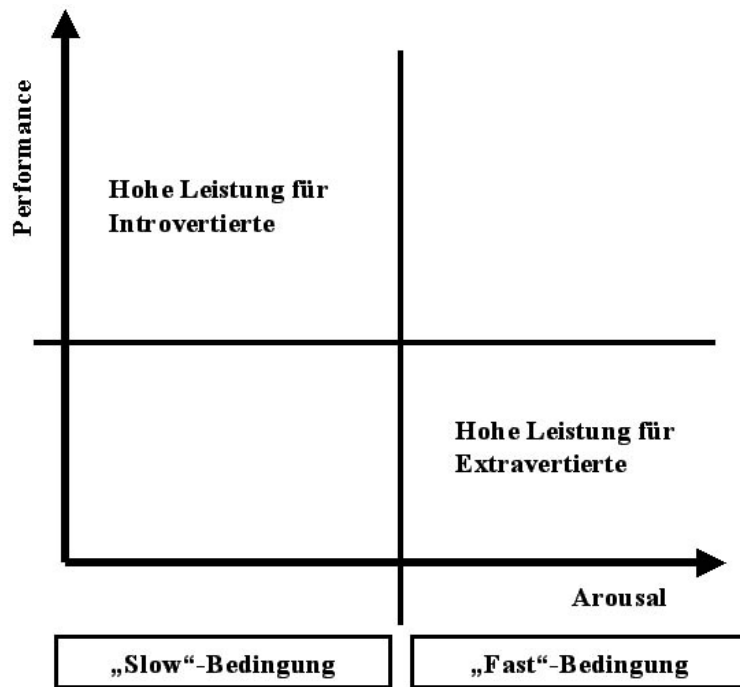


Abbildung 5: Visualisierung der Hypothese (2)

Hypothesen über Lerneffekte

Hypothese 3: In beiden Bedingungen wird der Verlauf der Kurve der missed gallons über die Durchgänge 1 bis 8 asymptotisch sein. Dabei werden im Sinne von Logan (1988) die Mittelwerte und Standardabweichungen signifikant geringer pro Durchgang (d.h. die Leistung der Teilnehmer/innen in dieser Computersimulation wird besser, es findet „Lernen“ statt). Das bedeutet auch, dass die Mittelwerte und Standardabweichungen am zweiten Tag im Durchschnitt niedriger sind als am ersten Tag.

Hypothesen über Persönlichkeitsdimensionen

Hypothese 4: Zhang (2002) bringt an, dass Menschen mit hohen „Neurotizismus-Werten“ strukturierte Lernmethoden brauchen, somit ein stabileres Umfeld. Neurotizismus wird sich demzufolge als bedeutsame Kovariate, die einige Varianz aufklärt, in einer komplexen (nicht-strukturierten) Computersimulation erweisen.

Explorativ wird untersucht: Es wird angenommen, dass Versuchspersonen mit hohen „Neurotizismus-Werten“ in der "slow condition" besser sind, da sie am ersten Tag nicht dem Druck schnelle Entscheidungen treffen zu müssen – also großer Dynamik und Unsicherheit –

ausgesetzt sind. Zeitdruck erzeugt negative Emotionen (vgl. Brehmer, 1992, 1996; Schwarz, 2000) und Veränderungen im Entscheidungsverhalten (Kerstholt, 1994; Benson & Beach 1996; Ben Zur & Breznitz, 1981). Teilnehmer/innen mit hohen Werten auf der Dimension „Neurotizismus“ werden mit dieser komplexen Computersimulation schlechter zurecht kommen als Teilnehmer/innen mit niedrigen Werten (vgl. auch Ngidi & Sibaya, 2002).

Hypothese 5: Auch das Persönlichkeitsmerkmal „Offenheit für Erfahrungen“ wird sich als bedeutsame Kovariate herausstellen. Die Ergebnisse von Zhang (2002, 2003) stützen diese Hypothese. „Offenheit für Erfahrungen“ ist Lernen zuträglich, was ebenso Blickle (1996) darstellt. Eine weitere Untersuchung von Conelly et al. (2000) betont die Bedeutung des Faktors „Offenheit für Erfahrungen“ für den Erfolg beim komplexen Problemlösen.

Explorativ wird untersucht: Teilnehmer/innen mit hohen Werten auf der Dimension "Offenheit für Erfahrungen" zeigen eine höhere Leistung in der WPP-Aufgabe als Probanden/innen mit niedrigen Werten. Diese Hypothese geht zurück auf die Untersuchungen von Zhang (2003) und Zhang (2002), der Beziehungen zwischen den „Big Five“ und Lerneffekten untersuchte.

Hypothesen über Geschlechtseffekte

Hypothese 6: Männliche Versuchsteilnehmer werden bessere Ergebnisse erzielen als weibliche. Allerdings wird erwartet, dass weit mehr Frauen am Experiment teilnehmen, was eine saubere statistische Auswertung problematisch macht. Somit kann diese Hypothese nur explorativ untersucht werden.

Hypothesen zur Computervorerfahrung

Hypothese 7: Es wird davon ausgegangen, dass Teilnehmer/innen, die im Umgang mit Computern vertraut sind, bessere Leistungen mit dem computersimulierten Szenario erreichen (vgl. Van der Linden et al. 2001; Schoppek & Putz-Osterloh, 2003). Diese Hypothese soll ebenso explorativ untersucht werden.

2.9 Zusammenfassung

Diese Studie beschäftigt sich mit Lernprozessen beim komplexen Problemlösen. Wie findet Wissenserwerb beim Bearbeiten computersimulierter Szenarien statt und passen die hier gewonnenen Ergebnisse zu den Untersuchungen von Gonzalez et al. (2003), die versucht haben, Wissenserwerb bei Menschen mit mathematischen Modellen zu vergleichen? Das theoretische Rahmenmodell des Lernprozesses – was nach Vollmeyer und Funke (1999) Voraussetzung für eine Modellierung ist – liefert die Instanzenbasierte Lerntheorie und ACT-R, was von Gonzalez et al. verwendet wird, aber nicht Fokus der vorliegenden Arbeit ist. Es kann nach Vollmeyer und Funke (1999) zwischen Personen- und Aufgabenmerkmalen beim komplexen Problemlösen unterschieden werden. Als *Personenmerkmale* werden in diese Studie „Neurotizismus“, „Extraversion“ und „Offenheit für Erfahrungen“ aufgenommen. *Aufgabenmerkmale* betreffen zum Beispiel die semantische Einkleidung und die Zielspezifität. Es gibt aber keine besondere Story wie z.B. bei Moro (vgl. Strohschneider & Güss, 1999) zur Einleitung für die Water-Purification-Plant. Das Ziel ist, möglichst wenige „missed gallons“ zu erreichen. Dieses Ziel kann als spezifisch beschrieben werden, da es den Teilnehmern/innen bekannt ist. Die Studie lässt sich als „modifizierte Replikation“ zu den Arbeiten von Gonzalez et al. beschreiben, da einige Charakteristika sehr ähnlich sind, aber sich der Versuchsablauf in manchen Bereichen unterscheidet. Auf die Gemeinsamkeiten und Unterschiede wird im Abschnitt 3 und 4 genauer eingegangen.

3 METHODEN

Methodische Vergleiche können mit den Studien von Cleotilde Gonzalez et al. vorgenommen werden, an die sich die vorliegende Diplomarbeit anlehnt. Auf Unterschiede und Gemeinsamkeiten zu diesen Studien wird hingewiesen.

Ziel der Studie ist zu zeigen, inwieweit Zeitdruck eine Auswirkung auf die Leistungen in einer komplexen computersimulierten Problemlöseaufgabe hat und wie viel Varianz ausgesuchte Persönlichkeitsdimensionen im Bearbeitungsprozess aufklären. Hierfür wird ein experimentelles Setting gewählt.

Im Folgenden werden das Design und die verwendeten Materialien vorgestellt. Außerdem wird die Stichprobe beschrieben, der Ablauf des Experiments verdeutlicht und die Datenanalyseverfahren genannt.

3.1 Design und Materialien

Die Studie ist ein „single-variable experimental Design mit Messwiederholung und Kovariate“ (vgl. Graziano & Raulin, 2000). Die unabhängige Variable ist Zeitdruck und sie ist zweifach gestuft: in eine „slow-Bedingung“ ohne Zeitdruck und eine „fast-Bedingung“ mit Zeitdruck. Es wird an zwei Messzeitpunkten erhoben, an zwei verschiedenen aufeinanderfolgenden Tagen. Während dieser Messungen werden mehrere Durchläufe an der computersimulierten Water-Purification-Plant durchgespielt, wobei Leistungs- und Verlaufswerte in ein Log-File abgespeichert werden. Als Kovariaten werden drei Persönlichkeitsdimensionen aus dem NEO-FFI (Ostendorf & Borkenau, 1993) verwendet, die zu Beginn der Erhebung erhoben werden. Des weiteren wird die Variable „Computervorerfahrung“ als Kovariate eingeführt. Die Studie beinhaltet sowohl explorative Fragestellungen als auch eine weitergehende qualitative Untersuchung der Verlaufsdaten der WPP. Tabelle 1 bietet einen Überblick über das Design.

Tabelle 1: Experimentelles Design

	Tag 1	Tag 2	Anzahl der Durchgänge	Verwendete Zeit pro Aufgabe (Min.)
„slow“-Bedingung	2 Durchgänge á 24 Minuten	6 Durchgänge á 8 Minuten	8	96
„fast“-Bedingung	6 Durchgänge á 8 Minuten	6 Durchgänge á 8 Minuten	12	96
Verwendete Zeit pro Aufgabe (Min.)	48	48		

3.1.1 Das WPP-Szenario

Die Grundlagen der Computersimulation wurden in Kapitel 2.4 besprochen. Da es verschiedene computersimulierte Szenarien gibt, die auch zur Erforschung unterschiedlicher Schwerpunkte dienen, ist es wichtig, zwischen dem Konstrukt „komplexes Problemlösen“ und dessen Operationalisierung zu unterscheiden (Spering, 2001). Möglichkeiten zur Steigerung der Validität durch Strukturgleichungsmodelle oder durch Finite Automaten wurden in 2.2 besprochen. Die Water-Purification-Plant zeichnet sich durch eine geschickte semantische Einkleidung (emotional neutral) und einfache Bedienbarkeit via Maus-Klicks aus.

Für die Auswertung ist entscheidend, dass das Programm die Werte der „missed gallons“ pro Durchlauf abspeichert, sowie die Werte der „missed gallons“ pro Deadline. Die Simulation läuft von 2:00 Uhr bis 10:00 Uhr in einer konstanten Geschwindigkeit, die vorher vom Versuchsleiter eingestellt wurde. Ab 5:00 Uhr einschließlich gibt es jede halbe Stunde eine Deadline. Nach einer Deadline wird die zusätzliche verpasste Wassermenge im Tank direkt davor dazugerechnet und erscheint auf dem Bildschirm. Der/die Teilnehmer/in erhält also Feedback über seine/ihre momentane Leistung, sobald die Deadlines (ab 5:00 Uhr simulierter Zeit) beginnen.

Nach der letzten Deadline um 10:00 Uhr werden alle restlichen missed gallons noch hinzugerechnet. Das sind diejenigen, die während der Bearbeitung in weiter hinten liegenden Tanks zurückgeblieben sind. Interessant für die qualitative Datenanalyse Abschnitt 4.5 ist, dass dieser Wert als die *Fähigkeit vorausschauend zu handeln, (und Zeitabläufe mit in Entscheidungen einzubeziehen)* bezeichnet werden könnte. Man sieht dies anhand eines Peaks in den Verlaufsdiagrammen der missed gallons pro Deadline (vgl. 4.5). Dies führt zu einem charakteristischen Muster, das bei fast allen Teilnehmer/innen beobachtbar ist. Die Güte des Vorausplanen eines/r Teilnehmers/in lässt sich also anhand der Größe dieses Peaks bewerten. Je größer, je schlechter und je kleiner, um so besser wurde sich frühzeitig schon um das Bearbeiten der kritischen Stellen (dies sind die insbesondere die Deadlines bei 8:00 und 9:00 Uhr im hier verwendeten Szenario) im Rohrleitungssystem gekümmert.

Pilotstudie:

Zwei Studenten des Faches Psychologie und zwei aus anderen Fachrichtungen wurden gebeten die WPP-Aufgabe durchzuführen, um Anregungen für einen optimalen Versuchsablauf zu bekommen. Dies trug erheblich zur Gestaltung der Instruktion (siehe Anhang B) bei.

Variable Zeitdruck:

Die Teilnehmer/innen in der „fast“-Bedingung spielen das Szenario drei mal so schnell durch, wie diejenigen in der „slow“-Bedingung. Gonzalez et al. (submitted) haben in Pilotstudien herausgefunden, dass dieses Zeitdruckmaß ein angemessenes ist (vgl. 2.5).

3.1.2 Der NEO-FFI

Der NEO-FFI (Borkenau & Ostendorf, 1993) ist ein paper and pencil Selbstreport-Fragebogen, der die deutschsprachige Übersetzung und Adaptation des NEO-Five-Factor-Inventory von Costa und McCrae (1992) darstellt. Es handelt sich um einen faktorenanalytisch konstruierten Test zur Beschreibung der wichtigsten Dimensionen individueller Persönlichkeitsunterschiede. Der NEO-FFI zeichnet sich herbei durch eine gelungene Mischung zwischen Ökonomie und Differenzierungsgrad bei den fünf Faktoren aus und wird gerne auch in angewandten Bereichen der Psychologie verwendet, zum Beispiel in der markt- und werbepsychologischen Forschung, einem klassischen Gebiet der Entscheidungsforschung.

Der NEO-FFI ist eine Kurzform des NEO-PI. Er umfasst fünf Skalen zu je 12 Items, also 60 Items insgesamt. Die fünf Skalen des NEO-FFI sind Neurotizismus, Extraversion, Offenheit für Erfahrungen, Verträglichkeit und Gewissenhaftigkeit. Für jedes Item steht eine 5-fach abgestufte Likert-Skala zur Verfügung, mit den Stufen „starke Ablehnung“, „Ablehnung“, „Neutral“, „Zustimmung“ und „starke Zustimmung“. Die Skalen werden mit einer Schablone ausgewertet. Der Gesamtwert für jedes Persönlichkeitsmerkmal ist der summierte Wert der jeweils 12 Items einer Skala, geteilt durch die Anzahl der beantworteten Items (zumeist geteilt durch 12). In Anhang A sind zwei Beispiel-Items pro Persönlichkeitsmerkmal angegeben. Die Durchführungszeit des NEO-FFI beträgt ungefähr 10 Minuten. Die Reliabilität und die Validität wird von Amelang und Zielinski (1994) als „sehr ordentlich“ bezeichnet. Die alpha-Koeffizienten liegen bei über .80 für die Skalen Neurotizismus, Extraversion und Gewissenhaftigkeit und bei .70 für Offenheit für Erfahrungen und Verträglichkeit.

Die Skalen lassen sich wie folgt kurz darstellen:

Neurotizismus: Tendenz nervös, erregbar, unsicher und verlegen zu sein und sich Sorgen um die eigene Gesundheit zu machen. Unfähigkeit, Bedürfnisse zu kontrollieren sowie auf Stressreaktionen angemessen zu reagieren.

Extraversion: Leute mit hohen Werten zeigen sich gesellig, aktiv, gesprächig, optimistisch und unternehmungslustig. Sie suchen Anregung und Aufregung.

Offenheit für Erfahrungen: Abwechslung, Kreativität und Unabhängigkeit im Urteil werden bevorzugt. Neue Erfahrungen werden gewertschätzt.

Verträglichkeit: Gutmütige, altruistische und mitfühlende Leute haben hohe Werte auf dieser Skala. Kooperation und Nachgiebigkeit sind weitere Charakteristika.

Gewissenhaftigkeit: Diese Skala unterscheidet zuverlässige, ordentliche, hart arbeitende und ehrgeizige Leute von nachlässigen und gleichgültigen.

Für die vorliegende Diplomarbeit wurden aufgrund der in Abschnitt 2.5 diskutierten Studien Neurotizismus, Extraversion und Offenheit für Erfahrungen ausgewählt. Kritisch anzumerken ist, dass – wie bei nahezu jedem Persönlichkeitstest – die Verfälschbarkeit des Tests z.B. durch soziale Erwünschtheit als relativ hoch einzustufen ist.

Aufgrund der Ökonomie und der Kürze des Verfahrens erscheint der NEO-FFI als ideales Instrument im Zusammenhang mit dem gegebenen Kontext.

3.1.3 Computervorerfahrung

Vorerfahrungen mit Computern sind eine wichtige Variable bei computersimulierten Problemlöseszenarien (vgl. Van der Linden et al., 2001, S. 193). Es wurden folgende drei selbstkonstruierte Fragen gestellt (vgl. Anhang B):

Ich habe gute Computerkenntnisse.
 Starke Ablehnung *Ablehnung* *Neutral* *Zustimmung* *Starke Zustimmung*

Ich benutze einen Computer...
 Fast nie *ein paar mal im Monat* *mehrmals in der Woche* *täglich*

Seit wieviel Jahren ungefähr arbeitest du schon mit einem Computer?

Die ersten vier Teilnehmer/innen bekamen nur die erste Frage vorgelegt, die anderen zwei wurden später erst hinzugefügt. Letztendlich wurden die Werte der ersten Frage verwendet, da diese alle Teilnehmer/innen beantwortet hatten und sie auf der gleichen Art der Skala beruht, wie der NEO-FFI, der zuvor ausgefüllt wurde.

3.2 Stichprobe

3.2.1 Festlegung der Stichprobengröße

Die Anwendung des Hypothesentestens wird seit einiger Zeit kontrovers diskutiert (vgl. Cohen, 1994, Wilkinson et al., 2000, Wilcox, 1998). Wichtig scheint in jedem Fall, auf bestimmte Rahmenbedingungen zu achten. Dazu gehört eine angemessene Stichprobengröße.

Die Power in einem Experiment beschreibt – als quantitativer Index – die Sensitivität eines Experiments (Czienskowski, 1996). Sie ist als Gegenwahrscheinlichkeit zum beta-Fehler, also dem Fehler zweiter Ordnung definiert. Wird der Fehler erster Ordnung durch das Festlegen eines Signifikanzniveaus festgelegt, muss der Fehler zweiter Ordnung durch experimentelle Methoden, die eine Reduzierung der Fehlervarianz nach sich ziehen, möglichst klein gehalten

werden. Der alpha-Fehler, der beta-Fehler, die Stichprobengröße und auch die Effektstärke hängen voneinander ab (Czienskowski, 2000). Deshalb wird empfohlen, sich an vorausgegangenen Experimenten oder Pilotstudien zu orientieren.

Gonzalez et al. (2003) verwenden in ihrer ersten Studie 14 Teilnehmer/innen, um Vergleiche zwischen computergenerierten Daten mit experimentell Gewonnenen anzustellen. In der Studie von Gonzalez et al. (submitted) nehmen 33 Teilnehmer/innen teil. In der Studie, die im Laufe der Diplomarbeit erst entstand (Gonzalez, submitted), sind schließlich 52 Teilnehmer/innen beteiligt. Eine ähnliche Teilnehmer/innenzahl von $n = 34$ wird demnach in der vorliegenden Diplomarbeit als ausreichend angenommen. Diese Entscheidung erscheint auch vor dem Hintergrund eines begrenzten Rahmens einer Diplomarbeit gerechtfertigt.

3.2.2 Versuchsteilnehmer

Die Rekrutierung der Teilnehmer/innen fand am Psychologischen Institut der Ruprecht-Karls Universität Heidelberg im Frühjahr 2004 statt. Die Teilnehmer/innen wurden in Lehrveranstaltungen auf das Experiment aufmerksam gemacht sowie über ein Plakat oder durch direktes Ansprechen von Personen in der Fakultät und per Mundpropaganda. Die Stichprobe von 34 Teilnehmer/innen besteht aus diesem Grund zum größten Teil ($n = 31$; 91,2%) aus Studenten/innen des Faches Diplom-Psychologie. Des weitern nahmen mehr Frauen ($n = 27$; 79,4%) als Männer ($n = 7$; 20,6%) an der Studie teil. An der Pilotstudie nahmen zwei Psychologiestudenten und zwei Studenten aus andern Fächern teil. Dies wird vom Autor als wichtig angesehen, da Psychologiestudenten/innen häufig schon im Umgang mit Experimenten vertraut sind und somit „Reibungen“ im Versuchsablauf besser von Studenten/innen anderer Fachrichtungen erkannt werden.

Als Incentives wurden Kleinigkeiten vom dm-Drogerie-Markt, den SISI-Werken sowie die Verlosung von drei Ravensburger Spielen angeboten. Dies ist ein wichtiger Unterschied zu den Studien von Gonzalez et al. die pro Teilnehmer/in 50\$ vergeben konnten.

Die Erhebung wurde im Februar 2004 innerhalb von drei Wochen, kurz vor den Semesterferien, durchgeführt. Die Nacherhebung einer Teilnehmerin fand ungefähr eine Woche nach der Haupterhebung statt. Ein Teilnehmer wurde ausgeschlossen, da er den NEO-FFI-Fragebogen nicht vollständig ausgefüllt hatte. Eine weitere Teilnehmerin wurde im nachhinein

ausgeschlossen, da sie die Simulation nicht an zwei hintereinanderfolgenden Tagen, sondern an einem Freitag und an dem darauf folgenden Montag durchgespielt hatte.

Die Teilnehmer/innen wurden randomisiert auf die beiden Bedingungen verteilt. Dies gilt nicht für die Nacherhebung – diese hatte zum Ziel, wieder gleiche Zellhäufigkeiten herzustellen (vgl. dazu Keselman & Keselman, 1990; Field, 1998).

Jeweils drei Versuchspersonen konnten das Szenario gleichzeitig an Computern durchspielen, die in einem Computerraum der Universität aufgestellt waren. Es war genügend Licht vorhanden und der Versuchsleiter war anwesend, um eventuelle Fragen zu klären.

Zusammengefasst: Der Altersmittelwert aller Teilnehmer/innen liegt bei 22,24 (SD = 3,53). 79,4% (n = 27) sind weibliche, 20,6% (n = 7) männliche Versuchspersonen. 91,2% (n = 31) studieren Psychologie im Hauptfach und 8,8% (n = 3) eine andere Fachrichtung. Die Altersmittelwerte der zwei verschiedenen Treatment-Bedingungen sind 21,94 (SD = 3,67) für die „slow“-Bedingung und 22,53 (SD = 3,49) für die „fast“-Bedingung und somit nicht signifikant voneinander verschieden ($t(32) = -0,58$; $p = 0,57$). Die Daten über die genaue Zusammensetzung der Stichproben werden in den Artikeln von Gonzalez et al. nicht erwähnt.

3.3 Ablauf der Untersuchung

Die Teilnehmer/innen nahmen vor dem Computer Platz und ihnen wurde standardisiert eine Übersicht über die Untersuchung vom Versuchsleiter gegeben. Sie bekamen eine schriftliche Anleitung und den NEO-FFI-Fragebogen mit Kugelschreiber und wurden instruiert, sich bei Fragen oder Schwierigkeiten an den Versuchsleiter zu wenden, der im selben Raum etwas abseits saß. Nach dem Ausfüllen des NEO-FFI und der Fragen zum Computervorwissen lasen die Teilnehmer/innen eine Beschreibung der WPP (siehe Anhang B). Dann wurde ihr Code eingegeben, sowie einige demographische Merkmale, die das Programm verlangt (class status, study subject, age, sex) und im Anschluss daran ein Probedurchlauf von ungefähr 5 Minuten Länge gestartet.

Hierbei konnten die Teilnehmer/innen die Simulation (nur) bis zur ersten Deadline erproben. Die Simulation wurde hierbei auf eine mittlere Geschwindigkeit eingestellt (Stufe: 15 Minuten). Dies geschah im Gegensatz zu Gonzalez (2003, in press), bei der die Probephase ca. 10 Minuten dauerte, in der langsamst möglichen Geschwindigkeit (Stufe: 30 Minuten). Der

Versuchsleiter entschied sich für eine Verkürzung der Probephase aufgrund der langen Dauer des Experiments mit mehreren wiederholten Durchgängen und weil die Geschwindigkeit bei Stufe: „30 Minuten“ sehr langsam ist. Der Versuchsablauf schien reibungslos zu sein und einige Teilnehmer/innen wollten sogar schon nach kürzerer Zeit mit dem „richtigen“ Versuch beginnen, hielten also einen noch kürzeren Zeitraum für ausreichend zum Eingewöhnen in die Bedingung der WPP. Dies spricht für die Benutzerfreundlichkeit der WPP. Nach dieser Probephase hatten die Teilnehmer/innen noch die Möglichkeit Fragen zu stellen, die – wie bei Gonzalez et al. (2003) – begrenzt beantwortet wurden. Sie wurden während der Probephase auf Deadlines, Wasserleitungen, Höchstzahl von aktivierbaren Pumpen und die Anzeigen (Uhrzeit, missed gallons, aktivierte Pumpenanzahl) hingewiesen. Keine Angaben wurden gemacht über das Verwenden von Heuristiken oder etwaige geschickte Lösungsstrategien. Bei den Teilnehmer/innen der „slow“-Bedingung wurde die Simulation am ersten Tag nach dem ersten Durchlauf vom Versuchsleiter gestartet. Die Teilnehmer/innen in der „fast“-Bedingung wurden nach dem zweiten Durchlauf aufgefordert, die Simulation selbst neu zu starten nach einem Durchlauf, so dass der Versuchsleiter möglichst wenig interagieren musste und somit nicht nach Feedback zum erhaltenen Ergebnis befragt wurde. Am zweiten Tag wurden die Teilnehmer/innen der „slow“-Bedingung ebenfalls aufgefordert, die Simulation alle acht Minuten selbst neu zu starten. An diesem Tag fingen die Teilnehmer/innen sofort mit der Simulation an. Nach der Erhebung am zweiten Tag wurde den Teilnehmern/innen mündlich nähere Information zur Untersuchung gegeben. Nach dem ersten und dem zweiten Tag bekamen die Teilnehmer/innen jeweils kleine Incentives (z.B. Seifen, Duschgel, Kugelschreiber, Vitaminkapseln) die vom dm-Markt in Heidelberg und den SISI-Werken in Eppelheim gesponsert wurden. Die Ravensburger Spiele wurden erst nach der gesamten Datenerhebung unter den Teilnehmer/innen verlost.

Wichtig ist, dass die Untersuchung an zwei aufeinanderfolgenden Tagen stattgefunden hat, und nicht wie bei Gonzalez et al. an drei hintereinanderfolgenden Tagen. Diese Einschränkung geschah aufgrund der schon erwähnten begrenzten Ressourcen im Kontext einer Diplomarbeit.

Der Ablauf des Experiments ist nochmals in Abbildung 5 schematisch dargestellt.

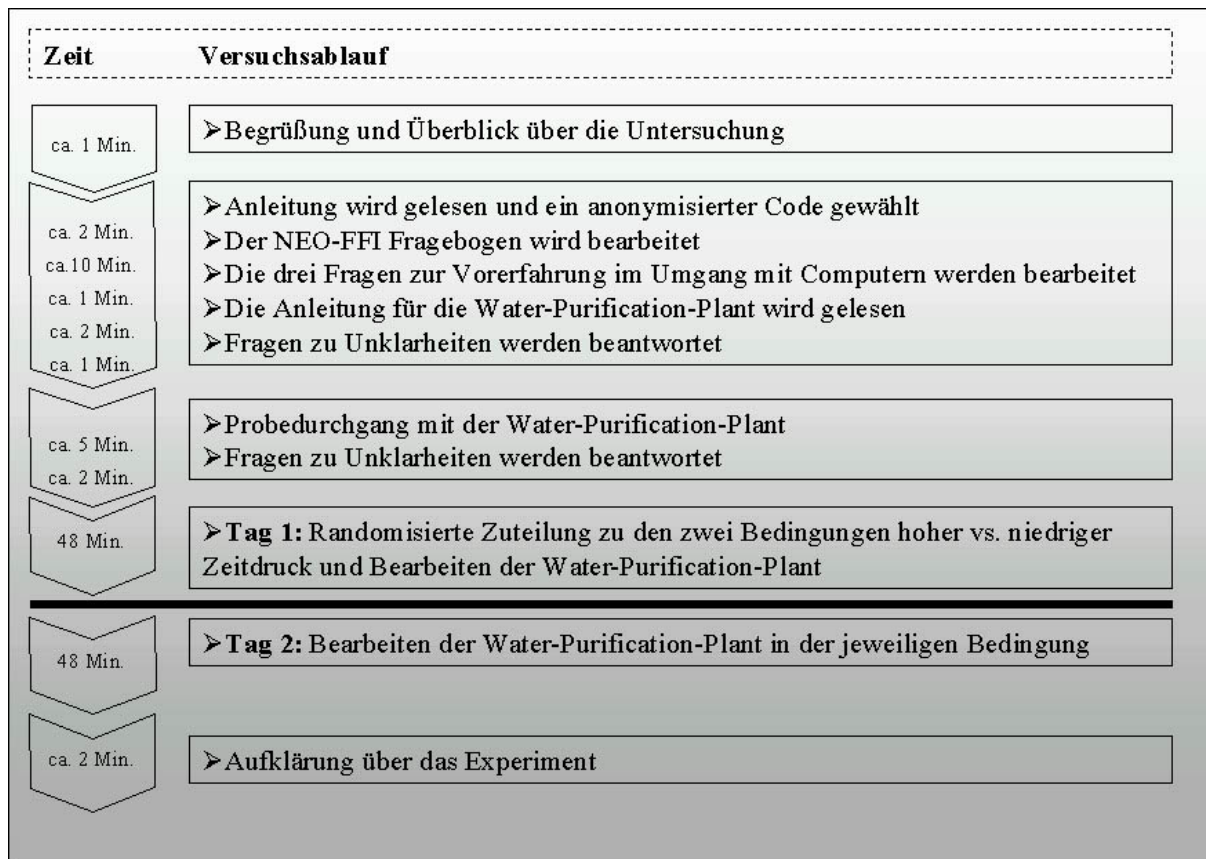


Abbildung 5: Schematisch dargestellter Ablauf des Experiments

3.4 Statistische Hypothesen

Die Studie beinhaltet sowohl Aspekte des „individual approach“ als auch des „experimental approach“ sensu Brehmer (1996). Die hier beschriebenen statistischen Hypothesen leiten sich aus 2.8 ab. Die Auswertungsverfahren, die jeweils verwendet werden, sind an den entsprechenden Stellen erwähnt. Wichtig ist, dass in die repeated measurement Analyse nur die ersten acht Durchgänge eingehen. Die weiteren Durchgänge der „fast“-Bedingung werden wie bei Gonzalez verworfen, um eine Konfundierung mit der erhaltenen Übung zu vermeiden und die Teilnehmer/innen bei einem gleichen Lernlevel zu testen (vgl. Gonzalez et al., submitted, S.15). In Gonzalez' zweiter experimentellen Studie (Gonzalez, submitted) unterscheidet sie zwischen einer Übungs- und einer Testphase. Die Übungsphase dauert bei ihr zwei Tage (in der vorliegenden Arbeit ein Tag), die Testphase einen Tag. Analog kann bei der vorliegenden Diplomarbeit der erste Tag als Übungsphase und der zweite Tag als Testphase gesehen werden. Darauf wird insbesondere für die explorativen Analysen aus praktischen Gründen zurückgegriffen.

Hypothesen über Zeitdruck

Hypothese 1: Die Mittelwerte des Leistungsmaßes „missed gallons“ werden sich signifikant voneinander unterscheiden, und zwar so, dass die Teilnehmer/innen der „slow“-Bedingung am zweiten Tag geringere Mittelwerte erreichen, also bessere Leistung zeigen. Das Ergebnismuster wird anhand einer repeated measurement ANOVA sowie graphisch mit Liniendiagrammen untersucht. Für die Berechnung des Unterschieds der Mittelwerte zwischen den beiden Gruppen wird ein t-Test verwendet. Die Werte der „slow“-Bedingung sind dabei am zweiten Tag besser als am ersten.

Hypothese 2: Das Persönlichkeitsmerkmal „Extraversion“ wird sich als signifikante Kovariate herausstellen. Dazu wird eine repeated measurement ANOVA mit Kovariate gerechnet (vgl. Gonzalez et al.).

Explorativ: Die Teilnehmer/innen werden per Mediansplit (vgl. Gonzalez) anhand dieser Persönlichkeitsdimension in zwei Gruppen „Ehigh“ und „Elow“ geteilt. Die Mittelwerte der eher Introvertierten werden an beiden Tagen in der „slow“-Bedingung geringer sein im Vergleich zu denen in der „fast“-Bedingung. Ebenso gilt: Die Mittelwerte der eher Extravertierten werden an beiden Tagen in der „fast“-Bedingung besser sein im Vergleich zu denen in der „slow“-Bedingung. Dies wird anhand zweier t-Tests geprüft (vgl. Abbildung 4, S. 43).

Hypothesen über Lerneffekte

Hypothese 3: Die Mittelwerte der missed gallons in beiden Treatment-Gruppen werden am zweiten Tag signifikant geringer sein als am ersten. Ein t-Test wird für jede Gruppe gerechnet, um dies zu überprüfen. Der Verlauf der Mittelwerte und der Standardabweichungen wird graphisch anhand von Liniendiagrammen untersucht. Die Mittelwerte werden im Mittel signifikant geringer, von Durchgang zu Durchgang. Dies wird mittels einer repeated measurement ANOVA getestet.

Hypothesen über Persönlichkeitsdimensionen

Hypothese 4: Das Persönlichkeitsmerkmal „Neurotizismus“ wird sich als signifikante Kovariate der repeated measurement ANOVA mit Kovariate erweisen.

Explorativ: Die Teilnehmer/innen werden auch hier per Mediansplit (vgl. Gonzalez) anhand dieser Persönlichkeitsdimension in zwei Gruppen „Nhigh“ und „Nlow“ geteilt. Teilnehmer/innen der „Nlow“-Gruppe werden niedrigere Mittelwerte (also bessere Ergebnisse) am zweiten Tag erreichen. Ein t-Tests wird dafür gerechnet.

Hypothese 5: Auch das Persönlichkeitsmerkmal „Offenheit für Erfahrungen“ wird sich als bedeutsame Kovariate in der repeated measurement ANOVA mit Kovariate herausstellen.

Explorativ: Die Teilnehmer/innen werden ebenso per Mediansplit (vgl. Gonzalez) anhand dieser Persönlichkeitsdimension in zwei Gruppen „Ohigh“ und „Olow“ geteilt. Teilnehmer/innen der „Ohigh“-Gruppe werden niedrigere Mittelwerte (also bessere Ergebnisse) am zweiten Tag erreichen. Wiederum wird hierfür ein t-Tests gerechnet.

Hypothesen über Geschlechtseffekte

Hypothese 6: Männliche Versuchsteilnehmer werden weniger missed gallons erreichen als weibliche. Ein t-Test über die Mittelwerte der missed gallons am zweiten Tag wird signifikant geringere Werte (also bessere Leistung) für die männlichen Teilnehmer/innen ergeben.

Hypothesen zur Computervorerfahrung

Hypothese 7: Computervorerfahrung wird sich als signifikante Kovariate herausstellen. Die Werte der Frage nach der persönlichen Einschätzung der Computerkenntnisse werden als Kovariate in der repeated measurement ANOVA mit Kovariate verwendet. Es wird auch – im Sinne der Übungs- und Testphase nach Gonzalez – ein t-Test zwischen zwei Gruppen mit den Extremwerten auf dieser Variablen gerechnet. Die Teilnehmer/innen, die ihre Vorerfahrung größer einschätzen, werden signifikant bessere Ergebnisse an beiden Tagen haben.

3.5 Datenanalyseverfahren

Zur Auswertung der Daten wurden die Programme SPSS 11.5 für Windows sowie SPSS für Windows Student Version Release 9.0.1 (SPSS Inc. 1989 - 2003) und Excel 2000 (Microsoft Corporation) verwendet. Für intervallskalierte Werte werden Mittelwerte und Standardabweichungen angegeben, für ordinalskalierte Werte Rangfolgen, absolute Werte sowie Prozentwerte.

Für die verwendeten Verfahren werden die Voraussetzungen (Varianzhomogenität mit Levene's Test; Normalverteiltheit anhand des Kolmogorov-Smirnov-Tests) geprüft und Verletzungen gegebenenfalls angegeben. Auf potentielle Ausreißer wird via graphischer Analyse sowie verschiedener Maße (Studentized Residuals, Cook's distance und Mahalanobis distance, siehe Field, S. 124ff) geachtet.

Neben verschiedenen t-Tests wurde eine repeated measurement ANOVA ohne Kovariate und mit den Persönlichkeitsmerkmalen "Extraversion-Introversion", „Neurotizismus“, „Offenheit für Erfahrungen“ und „Computervorerfahrung“ als Kovariaten gerechnet. Nach Werner (1997) zählen diejenigen Regressionsmodelle zur Kovarianzanalyse, in denen auf der Prädiktorseite quantitative und qualitative Variablen auftreten. Es geht in diesem Verfahren um das Ermitteln weiterer Variablen, die die abhängige Variable potentiell beeinflussen. Deren Einfluss wird dann kontrolliert und die Messergebnisse davon bereinigt (herauspartialisiert). Die Auswahl der Tests orientierte sich an den Arbeiten von Gonzalez et al. und den Ausführungen von Vorberg und Blankenberger (1999).

Werner (1997, S. 415ff) kritisiert vier nicht nötige Voraussetzungen und eine einschränkende Beschreibung der Verwendung einer Kovarianzanalyse. Die drei nötigen Voraussetzungen sind nach Werner (1997) auf Grundlage des Allgemeinen Linearen Modells:

1. Der Erwartungswert des Fehlers ist null;
2. Homoskedastizität liegt vor, d.h. die Kovarianzmatrix der Fehler ist eine Diagonalmatrix mit konstanter Fehlervarianz;
3. Die Fehler müssen normalverteilt ein.

Während die letzten beiden Voraussetzungen gegeben sind, ist die erste Voraussetzung durch die stattfindenden Lerneffekte kritisch zu sehen. Laut Werner ist es in solchen Fällen nötig, den Zeitverlauf mit ins das Modell einzubauen. Dies wird durch die Verwendung einer repeated measurement Analyse getan. Des weitern ist Box' Test nicht signifikant, womit diese Annahme als bestätigt angesehen werden kann.

Bei repeated measurement Verfahren ist die Annahme der Sphärizität (z.B. Field, 1998) wichtig. Kurz kann dies als ein Analogon zur Varianzhomogenität bei anderen statistischen Methoden beschrieben werden. Die Sphärizitätsannahme ist weniger streng als die der „compound symmetry“ (bestimmte Charakteristika der Datenmatrix), aber wird darauf

zurückgeführt. Die Sphäritätsannahme bezieht sich auf die Gleichheit der Varianzen der Unterschiede zwischen den Treatment Levels (Field, 2000). Oder anders gesagt, sie bedeutet „homogene Treatment-Differenz-Varianzen“ (Vorberg & Blankenberger). Da das vorliegende Experiment allerdings explizit eine Lernkurve abbildet, muss diese Annahme zwangsläufig verletzt werden. Dies wird mit Mauchly's Test (Field, 2000) überprüft, der tatsächlich bei den Analysen signifikant ist, was eine Abweichung von der Sphärität in den Daten bedeutet. SPSS bietet hierfür Schätzer der Sphärität an. Field (1998) gibt praktische Tipps zum Umgang mit solchen Situationen. Keselman, Algina, Kowalchuk und Wolfinger (1999) sowie Wilcox, Keselman, Muska und Cribbie (2000) diskutieren den Umgang mit verschiedenen Datensätzen und Methoden der Analyse bei repeated measurement Designs. Letztendlich erscheint das Verwenden des univariaten Ansatzes in der vorliegenden Arbeit als passend (da $\epsilon > 0.7$):

“As a rule it seems that when you have a large violation of sphericity ($\epsilon < 0.7$) and your sample is greater than $(n+10)$ the multivariate procedures are more powerful, whilst with small sample sizes or when sphericity holds ($\epsilon > 0.7$) the univariate approach is preferred.” (Field, 1998, S. 338).

Des Weiteren wird auf den Huynh-Feldt Sphäritäts-Schätzer zurückgegriffen, den Field bei Werten größer 0.75 nahe legt (Field, S. 333f). Field diskutiert weiterhin bestimmte Verfahren für post-hoc Analysen. Er empfiehlt Kontraste gegenüber post-hoc Verfahren, da bei der Verwendung von Kontrasten die Power größer ist. Bei den univariaten Techniken scheint, wenn die Sphäritätsannahme nicht zutrifft, die Bonferroni-Methode die beste in Bezug auf Power und Kontrolle des Fehlers erster Ordnung zu sein (Field, 2000, S. 330).

Zu den erwähnten statistischen Analysen findet eine qualitative Auswertung der Verlaufskurven der einzelnen Teilnehmer/innen als Erweiterung des Ansatzes von Gonzalez statt. Diese weitergehende Analyse der einzelnen Daten lässt sich im Sinne von Creswell (1998) als „Case Study“ im weitesten Sinne beschreiben. Beziehungen zu den ausgewählten Persönlichkeitsmerkmalen sollen dabei diskutiert werden.

4 ERGEBNISSE

Die vorliegende Diplomarbeit lehnt sich an das Design der Arbeiten von Gonzalez an. Jedoch gibt es einige wichtige Unterschiede: Zuerst macht es einen Unterschied, in welcher Kultur ein Problemlöseszenario getestet wird, wie Strohschneider & Güss (1999) gezeigt haben. Es ist unklar, auf welchen Dimensionen sich die deutsche und die amerikanische Kultur mit Relevanz für den hier behandelten Kontext unterscheiden. Es kann aber davon ausgegangen werden, dass das kulturelle Umfeld, in dem man aufgewachsen ist, einen Einfluss auf die Art und Weise sich mit einem komplexen Problem zu befassen, hat. Ein weiterer sehr wichtiger Unterschied ist, dass Gonzalez ihre Teilnehmer/innen die WPP an drei verschiedenen Tagen durchspielen lässt, wohingegen die Untersuchung in der vorliegenden Diplomarbeit die Teilnehmer/innen nur zwei Tage veranschlagte. Dies hängt vor allem mit den zur Verfügung stehenden Ressourcen (Zeit, sowie Geld für angemessene Incentives) zusammen. Obwohl laut Logan (1988) wenige Durchgänge ausreichen, um Routinen zu etablieren, mag es sehr wohl sein, dass hierbei Unterschiede in den Ergebnissen Unterschiede im Design widerspiegeln. Als weiterer Unterschied sei noch die Art der Kovariate genannt. Gonzalez et al. verwenden als Kovariate den Wert im Ravenstest als Indiz für die kognitiven Fähigkeiten der Teilnehmer/innen. In der vorliegenden Diplomarbeit wurden drei Persönlichkeitsmerkmale des NEO-FFI ausgewählt. Außerdem fließt zusätzlich die Computervorerfahrung in die Analyse mit ein. Diese Auswahl geschah aufgrund von Erfahrungen aus anderen Studien und der Idee, Persönlichkeitsfaktoren und kognitive Maße in Verbindung zu bringen (vgl. Süß, 1999).

Das klassische H_0 -Hypothesentesten muss sich gegen immer stärkere Kritik behaupten (vgl. Cohen, 1994). Oft wird das Testen der H_0 fehlerhaft angewandt, bzw. fehlerhafte Schlüsse aus den angewandten Tests gezogen. Dies geschieht meist aus Gründen Unwissenheit, aber auch mangels überzeugender Alternativen. Cohen meint, dass wir uns momentan auf Replikationen verlassen müssen, um für bestimmte Effekte einen signifikanten Unterschied postulieren zu dürfen. Die vorliegende Diplomarbeit soll hierzu einen Beitrag leisten, indem sie (auch interkulturell) einen Vergleich – allerdings mit einem modifizierten Design – anstrebt, welcher die Ergebnisse von Gonzalez et al. (2003) ergänzen soll.

Zum klassischen H_0 -Hypothesentesten nach dem alpha-Signifikanzkriterium werden zusätzliche Maße aus einer qualitativen Analyse in die Beurteilung der Ergebnisse miteinbezogen.

Der Umgang mit den Ergebnissen einer repeated measurement Analyse ist nicht trivial und teils werden kleine Verletzungen der Voraussetzungen für bestimmte statistische Methoden schon mit unreliaiblen Resultaten bestraft (vgl. Keselman et al., 1999, Field, 2000, Keselman & Keselman, 1990, Wilcox et al., 2000). Deshalb wurde zum Beispiel auf gleiche Zellhäufigkeiten geachtet. Nach der Festlegung von Darstellungskonventionen werden die Daten des NEO-FFI dargestellt. Danach werden deskriptive Maße der Water-Purification-Plant vorgestellt sowie die Ergebnisse der repeated measurement Analyse. Hiernach folgen t-Tests und die qualitative Auswertung. In der Darstellung wird immer wieder der Bezug zu den Arbeiten von Cleotilde Gonzalez hergestellt werden.

4.1 Konventionen

Aufgrund der Daten der Studien von Gonzalez et al. und der verfügbaren Ressourcen wurde die Stichprobengröße auf $n = 17$ pro Bedingung festgelegt. Das Signifikanzniveau auf dem getestet wird, beträgt $\alpha = .05$.

Es werden neben Freiheitsgraden und Signifikanzniveaus bei den t-Tests Effektgrößen (d) berichtet. Sie können nach Tabelle 2 klassifiziert werden. Nach Hays (1994, S. 410) werden sie anhand der Formel $d = y_{\max} - y_{\min} / s$ für t-Tests mit abhängigen Stichproben berechnet. Hierbei steht „s“ für die gepoolte within-group Standardabweichung und „ y_{\max} “ und „ y_{\min} “ stellen die Stichprobenmittelwerte dar. Für t-Tests mit abhängigen Stichproben lautet die Formel $d^2 = t^2 / (2(1-r)/n)$ wobei t der Wert des t-Tests ist, r die Korrelation zwischen beiden Stichproben und $n = df+1$ (nach Dunlap, Cortina, Vaslaw & Buke, 1996, S. 171). Die Schätzer für Effektgrößen, die SPSS berechnet, sind nach Field (2000, S.299) zumeist verzerrt, weshalb auf ihre Darstellung verzichtet wird.

Tabelle 2: Konventionen für Effektgrößen (vgl. Bortz, 1999, S. 140)

Effekt	D
schwach	.20
mittel	.50
stark	.80

Für die repeated measurement Analyse werden nur die Daten der Durchgänge 1 bis 8 verwendet. Somit werden die Teilnehmer/innen bei einem gleichen Lernlevel getestet und Konfundierung mit der erhaltenen Übung wird vermieden (vgl. Gonzalez et al. (submitted), S. 15). Die Daten der restlichen Durchläufe, die aufgrund des Designs nur für die Teilnehmer/innen der „fast“-Bedingung vorliegen, werden bei den gerechneten t-Tests miteinbezogen. Dies geschieht aufgrund einer Einteilung des Experiments von Gonzalez in eine Übungs- und Testphase.

Bei den statistischen Analysen wurden die Voraussetzungen überprüft. Verletzungen werden im Text berichtet. Welche Datensätze nicht in die Analyse einbezogen wurden, ist in Abschnitt 3.2 erläutert. Von den Empfehlungen Konfidenzintervalle anzugeben (Cohen, 1994) wird abgesehen, um die Komplexität nicht weiter zu erhöhen. Bei der Reihenfolge der Darstellung werden die Kovariaten in folgender Reihenfolge behandelt: Neurotizismus, Extraversion, Offenheit für Erfahrungen und schließlich Computervorerfahrung.

4.2 Der NEO-FFI

Wie oben beschrieben, zeichnet sich der NEO-FFI durch eine ordentliche Reliabilität und gute Validität aus. Die Alpha-Koeffizienten für die verschiedenen Skalen liegen für die Gesamtstichprobe im Manual in der Gesamtstichprobe bei Neurotizismus 0.85, Extraversion 0.80 und Offenheit für Erfahrungen 0.71. Nach Geschlecht aufgeteilt liegen sie im Manual für Frauen für Neurotizismus bei 0.85 (Männer 0.85), Extraversion bei 0.80 (Männer 0.79) und für Offenheit für Erfahrungen bei 0.70 (Männer 0.72). Die Koeffizienten im Experiment sind also in einem akzeptablen Bereich. Daher kann von einer gelungenen Anwendung des NEO-FFI ausgegangen werden. Die Mittelwerte und Standardabweichungen und Alpha-Koeffizienten sind in Tabelle 3 abgebildet.

Tabelle 3: *Interkorrelationen und Alpha-Koeffizienten der ausgewählten Skalen für die Gesamtstichprobe (n = 34)*

Skala	1.	2.	3.	M	SD
1. Neurotizismus	0.75			2.28	0.48
2. Extraversion	-0.32	0.85		2.37	0.57
3. Offenheit für Erfahrungen	0.17	0.43*	0.75	2.81	0.46

Anmerkung: Die Alpha-Koeffizienten sind fettgedruckt in der Diagonale dargestellt.

** signifikanter Korrelationskoeffizient mit $p < 0.05$*

Insgesamt gab es kaum Schwierigkeiten im Umgang mit dem NEO-FFI. In einem Fall wurde eine Frage übersehen, was aber laut Manual keine Probleme bei der Auswertung nach sich zieht (Borkenau & Ostendorf, 1993). Ein anderer Teilnehmer füllte nur die erste und letzte Seite aus. Diese Daten wurden deshalb in der Analyse nicht berücksichtigt.

4.3 Hypothesenüberprüfung

Verteilungseigenschaften und deskriptive Statistiken der Werte des Water-Purification-Plant-Szenarios werden zuerst dargestellt. Darauf folgen die Ergebnisse der repeated measurement ANOVAs. Die deskriptiven Maße der WPP sind getrennt für die beiden Gruppen in Tabelle 4 angegeben. Sowohl die Mittelwerte als auch die Standardabweichungen nehmen im Laufe der Zeit ab. Dies entspricht dem Ergebnismuster von Gonzalez et al. (vgl. Hypothese (3)).

Tabelle 4: Deskriptive Maße bei der Bearbeitung der Water-Purification-Plant

Durchgang	"slow"-Bedingung					„fast“-Bedingung				
	N	Minimum	Maximum	M	SD	N	Minimum	Maximum	M	SD
1	17	113	434	185.41	78.67	17	89	384	211.65	76.71
2	17	79	360	172.00	82.37	17	77	233	160.24	49.91
3	17	92	269	167.00	62.60	17	129	383	189.82	64.19
4	17	73	286	147.94	55.26	17	73	438	164.71	81.53
5	17	65	266	140.47	52.52	17	32	358	150.35	68.44
6	17	69	217	138.94	55.72	17	47	289	126.76	56.05
7	17	70	310	146.53	60.89	17	65	288	131.71	57.54
8	17	85	218	132.06	33.43	17	46	231	113.12	48.33
9						17	48	347	114.06	67.12
10						17	59	334	113.53	65.04
11						17	44	329	116.82	69.43
12						17	38	341	103.82	69.53

In der Abbildung 6 sieht man den Verlauf der missed gallons, also der Leistung in der WPP, über alle Durchgänge. Interessant ist der Vergleich der beiden Kurven beim zweiten Durchgang. Dieser „Zacke“ in der Kurve der Teilnehmer/innen der „fast“-Bedingung ist schwer erklärbar. Er führt aber in einem full-factorial model in der repeated measurement ANOVA zu einem signifikanten Effekt bei wiederholten Kontrasten, an dieser Stelle zwischen dem ersten und zweiten Durchgang (vgl. Field, 2000, S.330) ($F(1,28) = 4.72, p < 0.05$). Dies gilt auch für

die Stelle zwischen dem fünften und sechsten Durchgang mit $F(1,28) = 4.03$, $p < 0.05$, wo die Teilnehmer/innen der „fast“-Bedingung diejenigen der „slow“-Bedingung überholen.

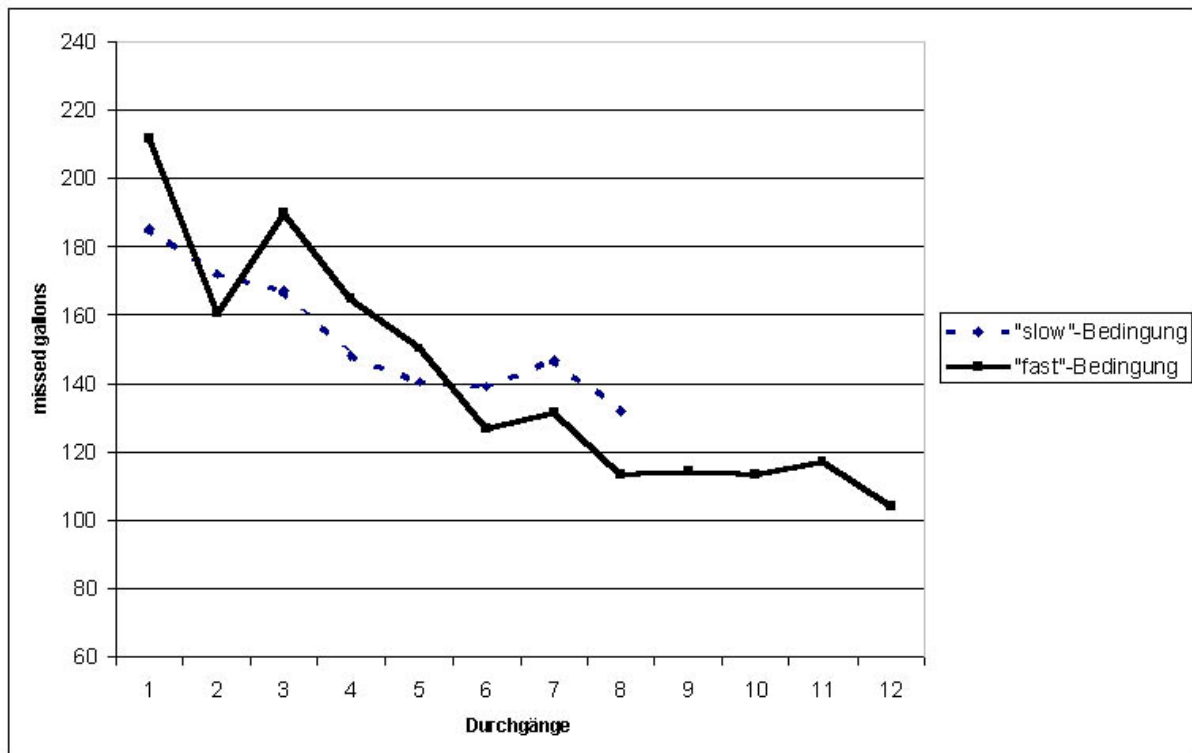


Abbildung 6: Graphische Abbildung der Leistungskurven zwischen der „slow“-Bedingung und der „fast“-Bedingung über 12 Durchgänge. Wenige missed gallons bedeuten eine gute Leistung.

Interessant ist auch, dass die Teilnehmer/innen in der „fast“-Bedingung diejenigen der „slow“-Bedingung im Laufe der Durchgänge überholen. Dies zeigt Abbildung 6 sehr anschaulich. Laut Gonzalez et al. sollte dies genau anders herum verlaufen. Sie findet letztendlich zwar auch keinen signifikanten Effekt zwischen den beiden Bedingungen, dennoch sind bei ihr die Teilnehmer/innen am zweiten Tag in der „slow“-Bedingung besser. D.h. die Teilnehmer/innen der „slow“-Bedingung sollten durch ihre bessere, effizientere Übung immer bessere Leistungen erbringen können. Die Hypothese (1), die das Ergebnismuster von Gonzalez et al. replizieren wollte, kann also als nicht bestätigt angesehen werden. Abbildung 7 zeigt die Veränderungen bei den missed gallons für die beiden Gruppen an den zwei Tagen. Die Teilnehmer/innen der „slow“-Bedingung sind im Durchschnitt am ersten Tag schlechter und können diesen Rückstand am zweiten Tag nicht aufholen.

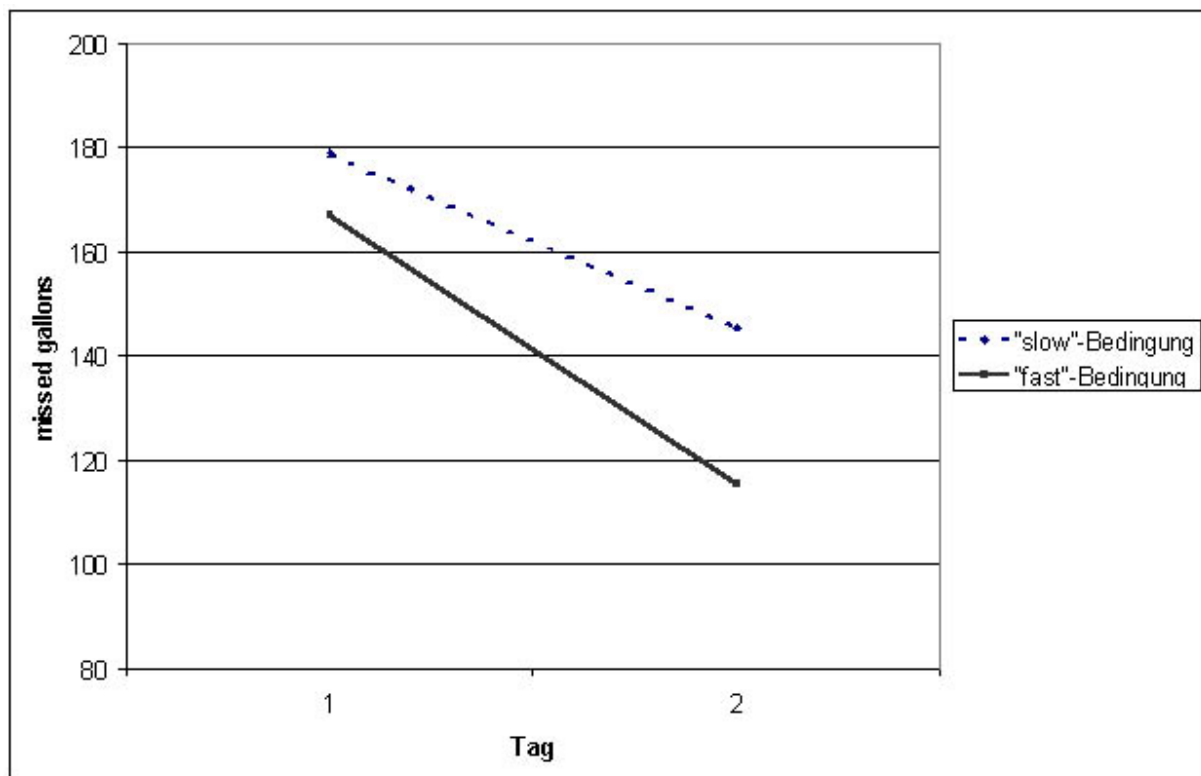


Abbildung 7: Graphische Abbildung der Leistungskurven zwischen der „slow“-Bedingung und der „fast“-Bedingung an beiden Tagen. Wenige missed gallons bedeuten eine gute Leistung.

Die Ergebnisse der globalen Analyse der repeated measurement ANOVA mit Interaktionen und mit vier Kovariaten (Neurotizismus, Extraversion, Offenheit für Erfahrungen, Computervorerfahrung) gerechnet, finden sich in Tabelle 5. Die Within-Subjects Analyse der Leistung bei der WPP zeigt wie bei Gonzalez et al. einen signifikanten Unterschied bei der Interaktion von Durchgang und Treatmentgruppe (Huynh-Feldt Sphärizitätsschätzer: $F(6,171) = 2.39, p < 0.05$; anhand des konservativen Greenhouse-Geisser Sphärizitätsschätzers: $F(4,121) = 2.39, p = 0.05$). Genaueres Aufschlüsseln mittels Kontrasten zeigt wie oben beschrieben, dass dieser Effekt auf Unterschiede der Gruppen zwischen dem ersten und zweiten sowie dem fünften und sechsten Durchgang (vgl. auch Abbildung 6) zurückzuführen ist. Keine signifikanten Ergebnisse finden sich im Gegensatz zu Gonzalez für die Kovariaten. Bemerkenswert ist, dass die Kovariate Computervorerfahrung die meiste Varianz unter den Kovariaten aufklärt.

Tabelle 5: *Repeated measurement Analyse mit vier Kovariaten: globale Analyse. p-Werte sind anhand des Huynh-Feldt Sphärizitätsschätzers berichtet.*

	Leistung		
	df	F	p
Within-Subjects			
Durchgang	6,171	0.76	0.61
Durchgang x Bedingung	6, 171	2.39	0.03*
Durchgang x Neurotizismus	6, 171	0.83	0.59
Durchgang x Extraversion	6, 171	0.94	0.47
Durchgang x Offenheit für Erfahrung	6, 171	0.65	0.70
Durchgang x Computervorerfahrung	6, 171	1.14	0.34
Durchgang x Bedingung x Neurotizismus	14,175	0.63	0.84
Durchgang x Bedingung x Extraversion	14,175	0.88	0.58
Durchgang x Bedingung x Offenheit für	14,175	0.60	0.86
Durchgang x Bedingung x Computervorerfahrung	14,175	1.43	0.14
Between-Subjects			
Bedingung	1,28	0.19	0.67
Neurotizismus	1,28	0.32	0.57
Extraversion	1,28	0.43	0.52
Offenheit für Erfahrungen	1,28	0.22	0.64
Computervorerfahrung	1,28	0.83	0.37
Bedingung x Neurotizismus	2,25	0.18	0.84
Bedingung x Extraversion	2,25	0.39	0.68
Bedingung x Offenheit für Erfahrungen	2,25	0.78	0.47
Bedingung x Computervorerfahrung	2,25	1.25	0.30

Anmerkung: * $p < 0.05$

Ein t-Test für unabhängige Stichproben (bei dem alle 12 Durchgänge in die Analyse mit einfließen) um den Unterschied der beiden Treatmentgruppen am zweiten Tag zu untersuchen, ergibt ein insignifikantes Ergebnis mit $t(32) = 1.64$, $p = 0.6$. Allerdings ist der Unterschied in der den Hypothesen entgegengesetzten Richtung zu finden (vgl. Tabelle 6). Abbildung 7 illustriert dieses Ergebnis graphisch. Es ist darin auch deutlich zu sehen, dass sich die Teilnehmer/innen der „fast“-Bedingung mehr verbessern, als diejenigen der „slow“-Bedingung.

Tabelle 6: Mittelwerte, Standardabweichungen und die Ergebnisse des t-Tests für unabhängige Stichproben zwischen den beiden Treatment-Gruppen an Tag 1 und Tag 2

Messzeitpunkte	Slow (n = 17)		Fast (n = 17)		Freiheitsgrade	t-Test	Effektgröße
	M	SD	M	SD			
Tag 1	178.71	73.56	167.15	59.78	32	0.50	0.17
Tag 2	145.49	46.63	115.51	59.26	32	1.64	0.57

Diese Ergebnisse bestätigen, dass die Annahmen der Hypothese (1) empirisch nicht bestätigt wurden. Die Teilnehmer/innen der „fast“-Bedingung sind am zweiten Tag besser trotz Zeitdruck.

Eine repeated measurement ANOVA, die ohne die Kovariaten gerechnet wird, zeigt einen signifikanten Within-Subject Unterschied pro Durchgang ($F(5,171) = 16.07, p < 0.001$). Mit den Kovariaten tritt dieser Effekt nicht auf. Dies lässt auf einen Moderator-Effekt der Kovariaten schließen. Kontraste zeigen, dass dies auf die Unterschiede zwischen den Durchgängen eins und zwei ($F(1,32) = 9.21, p < 0.001$), drei und vier ($F(1,32) = 8.43, p < 0.01$) sowie sieben und acht ($F(1,32) = 8.27, p < 0.01$) zurückgeht. Dies sind also diejenigen Durchgänge, bei denen von den Teilnehmer/innen im Durchschnitt am meisten Verbesserung in der WPP erreicht wird.

Gonzalez testet die Güte ihrer Kovariate mittels einer linearen Regression über die Mittelwerte aller Durchgänge. Sie erhält ein signifikantes Ergebnis für ihre Kovariate „Raventestwert“. In der vorliegenden Arbeit sind die Ergebnisse mit allen vier Kovariaten allerdings nicht signifikant ($F(4,29) = 0.21, p > 0.05$) und das Regressionsmodell klärt nur ungefähr 3 Prozent der Varianz in den Daten auf. Es zeigt sich aber wieder das Muster, dass die Variable Kenntnis offensichtlich am meisten Varianz aufklärt. Field betont die Wichtigkeit die Residuen zu untersuchen. Bezüglich der Ausreißer lässt sich festhalten, dass eine Teilnehmerin als Ausreißer (kritisch dazu Wilcox (1998), der zum Beispiel getrimmte Mittelwerte vorschlägt, was aber aufgrund der eindeutig der nicht-signifikanten Ergebnisse hier nicht angewandt wird) kategorisiert werden kann, da der Wert der Studentized Residuals knapp über dem kritischen Wert 3 liegt. Allerdings sind Cook's Distance (Maß für den Gesamteinfluss eines einzelnen

Falles auf das Modell) und Mahalanobis Distance (Maß für die Abstände der einzelnen Fälle zu den Mittelwerten der Prädiktorvariablen) in einem guten Bereich, so dass dies vernachlässigt wird. Das Ergebnismuster, wie es oben dargestellt ist, kann somit im Folgenden diskutiert und die Daten auf einem anderen Auflösungs niveau und mit einem anderen Blickwinkel analysiert werden.

Hypothese (3), die signifikante Lerneffekte und damit eine signifikante Verbesserung pro Durchgang annimmt, kann somit als teilweise bestätigt angesehen werden (vgl. auch Tabelle 7 und 8). Bezüglich der Hypothese (3) ist weiterhin festzuhalten, dass vier der Teilnehmer/innen am zweiten Tag schlechtere Mittelwerte bei der Bearbeitung der WPP haben. Womit dies zusammenhängt, ist unklar. Eventuell waren die Teilnehmer/innen nicht motiviert gewesen, gute Ergebnisse zu erzielen? Die Werte der missed gallons an den beiden Tagen unterscheiden sich also signifikant, wie Tabelle 8 zeigt ($t(33) = 5.54, p < 0.001$).

Tabelle 7: Mittelwerte, Standardabweichungen und die Ergebnisse des t-Tests für abhängige Stichproben zwischen den beiden Tagen für die beiden Bedingungen

Treatment- Gruppen	Tag 1 (n = 17)		Tag 2 (n = 17)		Freiheitsgrade df	t-Test t	Effektgröße d
	M	SD	M	SD			
Slow	178.71	73.56	145.49	46.63	16	2.62**	0.34
Fast	167.15	59.78	115.51	59.26	16	6.14***	0.61

Anmerkung: ** $p < 0.01$ (einseitig); *** $p < 0.001$ (einseitig)

Aufgeschlüsselt nach den beiden Gruppen ist ein t-Test für abhängige Stichproben (Tabelle 8) signifikant mit $t(16) = 2.62, p < 0.01$ für die „slow“-Bedingung und für die „fast“-Bedingung mit $t(16) = 6.14, p < 0.001$.

Tabelle 8: Mittelwerte, Standardabweichungen und die Ergebnisse des t-Tests für abhängige Stichproben zwischen den beiden Tagen für die Gesamtwerte.

missed gallons	Tag 1 (n = 34)		Tag 2 (n = 34)		Freiheitsgrade	t-Test	Effektgröße
	M	SD	M	SD			
Gesamt	172.98	66.26	130.50	54.67	33	5.54***	0.68

Anmerkung: *** $p < 0.001$ (einseitig)

Die Hypothesen über signifikante Kovariaten (Hypothese (2), (4), (5) und (7)) müssen allesamt verworfen werden. Im Gegensatz zu Gonzalez' et al. Experimenten konnten keine signifikanten Interaktionen von Kovariaten mit den Treatment-Bedingungen gefunden werden.

Zusammenfassend kann gesagt werden, dass die Befunde von Gonzalez et al. nicht bestätigt werden konnten. Es fanden sich signifikante Ergebnisse bezüglich der Unterschiede an beiden Tagen, was aber trivial ist, und wofür verschiedenste Lern- und Entscheidungsmodelle (außer der IBLT) angenommen werden können. Das Modell in Abbildung 21 (S. 112) und die Ausführungen von Schoppek und Putz-Osterloh können eine mögliche Erklärung für die nicht-signifikanten Ergebnisse mit den Kovariaten liefern. Persönlichkeitsmerkmale wirken nicht direkt auf die Problemlöseleistung, sondern über lange (Um-)Wege (Motivation, verwendete Strategien) und müssten möglichst nicht nur am Anfang, sondern zu mehreren Zeitpunkten während eines Experiments gemessen werden. Es ist nicht klar, wie der „Zacke“ in der Verlaufskurve der Leistungen (Abbildung 6, S. 69) zustande kommt, der schließlich zu einem signifikanten Kontrast führt. Es ist möglich, dass er durch die Struktur der WPP zustande kommt. Ein paar befragte Teilnehmer/innen gaben an, dass sie, nachdem sie sich mit der Aufgabe vertraut gemacht hatten, begannen Strategien zu testen. Es ist möglich, dass dies zu einem abrupten Leistungsabfall führte, was mit stabilisierendem Verhalten beantwortet wurde. Um solche Effekte näher zu untersuchen, werden in Abschnitt 4.5 einige qualitative Analysen durchgeführt. Es stellt sich also die Frage, inwieweit die WPP hilfreich ist, das Lernen von Entscheidungen in komplexen Situationen zu untersuchen.

4.4 Explorative Datenanalyse

Explorativ wurden im Sinne des „individual approach“ beim komplexen Problemlösen (vgl. Brehmer, 1996) die Unterschiede in der Leistung zwischen Gruppen, die durch Mediansplits je Kovariate entstanden sind, via t-Tests untersucht (Hypothese (2), (4), und (5)). Außerdem wurde auf Geschlechtseffekte (Hypothese (6)) hinsichtlich der Leistung in der WPP getestet. Die Zweiteilung der Datenerhebungszeitpunkte von Gonzalez (submitted) in eine Übungs- und eine Testphase ist an dieser Stelle insbesondere bedeutend. Deshalb werden in die Analysen die Werte aller Durchgänge miteinbezogen.

4.4.1 Neurotizismus

Zur Überprüfung der Hypothese (4) wurden die Teilnehmer/innen in gleich große Gruppen „Nhigh“ und „Nlow“ eingeteilt. Ein t-Test zur Überprüfung des Unterschieds ist nicht signifikant mit $t(32) = -0.98$, $p = 0.17$. In Tabelle 9 ist zu sehen, dass die Standardabweichungen für die „Nhigh“-Gruppe deutlich größer sind. Das bedeutet, dass Teilnehmer/innen in dieser Gruppe größere Variabilität in ihren Werten aufweisen, was zu einem hohen Wert auf der Persönlichkeitsdimension „Neurotizismus“ passt. Zumindest tendenziell zeichnen die Mittelwerte und Standardabweichungen die Idee der Hypothese (4) nach. Da die Varianzen am zweiten Tag nicht homogen sind (Levene’s Test ist signifikant), werden die Freiheitsgrade zur Korrektur verändert und somit der t-Test mit 26.39, anstatt 32 Freiheitsgraden gerechnet.

Tabelle 9: Mittelwerte und Standardabweichungen des Persönlichkeitsmerkmals Neurotizismus mit den Ergebnissen des t-Tests für unabhängige Stichproben an beiden Tagen zwischen den mediangesplitteten Gruppen.

Messzeitpunkt	Nlow (n = 17)		Nhigh (n = 17)		Freiheitsgrade	t-Test	Effektgröße
	M	SD	M	SD			
Tag 1	165.80	57.45	180.16	75.14	32	-0.63	0.22
Tag 2	121.32	40.16	139.68	66.12	26.39	-0.98	0.35

4.4.2 Extraversion/Introversion

Um die Hypothese (2) explorativ zu überprüfen, wird ebenso ein t-Test gerechnet. Die Ergebnisse sind in Tabelle 10 dargestellt. Levene’s Teststatistik ist am ersten Tag in der „slow“-

Bedingung signifikant, weshalb die Annahme der Varianzhomogenität nicht gehalten werden kann. Deshalb ist der dargestellte t-Wert nicht mit 15 Freiheitsgraden, sondern mittels 9,83 Freiheitsgraden berechnet.

Da vier t-Tests durchgeführt werden, wird ein Bonferronikorrigiertes Signifikanzniveau mit $\alpha = 0.05/4 = 0.01$ je Test verwendet (vgl. Beck-Bornholdt & Dubben, S.58). Die Mittelwerte der eher Extravertierten in der „fast“-Bedingung sind an beiden Tagen niedriger (Tabelle 11) als die der eher Introvertierten. Allerdings finden sich keine signifikanten Ergebnisse bei den Vergleichen sowohl für die „slow“-Bedingung als auch für die „fast“-Bedingung. Somit wird der explorative Teil der Hypothese (2) als nicht bestätigt angesehen.

Tabelle 10: „Slow“-Bedingung: Mittelwerte und Standardabweichungen des Persönlichkeitsmerkmals Extraversion mit den Ergebnissen des t-Tests für unabhängige Stichproben an beiden Tagen zwischen den mediangesplitteten Gruppen.

Messzeitpunkt	Elow (n = 9)		Ehigh (n = 8)		Freiheitsgrade	t-Test	Effektgröße
	M	SD	M	SD			
Tag 1	169.22	47.20	189.38	97.86	9.83	0.53	0.28
Tag 2	155.48	46.88	134.25	46.76	15	0.91	0.45

Tabelle 11: „Fast“-Bedingung: Mittelwerte und Standardabweichungen des Persönlichkeitsmerkmals Extraversion mit den Ergebnissen des t-Tests für unabhängige Stichproben an beiden Tagen zwischen den mediangesplitteten Gruppen.

Messzeitpunkt	Elow (n = 8)		Ehigh (n = 9)		Freiheitsgrade	t-Test	Effektgröße
	M	SD	M	SD			
Tag 1	177.35	78.94	158.28	38.77	15	0.65	0.32
Tag 2	129.27	82.01	103.28	27.94	15	0.90	0.47

4.4.3 Offenheit für Erfahrungen

Um die Hypothese (5) zu überprüfen, wird wiederum ein t-Test für die beiden mediangesplitteten Gruppen des Persönlichkeitsmerkmals „Offenheit für Erfahrungen“ gerechnet. Die Ergebnisse der Analyse, die in Tabelle 12 dargestellt sind, zeigen aber keinerlei signifikante Ergebnisse.

Tabelle 12: Mittelwerte und Standardabweichungen des Persönlichkeitsmerkmals Offenheit für Erfahrungen mit den Ergebnissen des t-Tests für unabhängige Stichproben an beiden Tagen zwischen den mediangesplitteten Gruppen.

Messzeitpunkt	Low (n = 17)		High (n = 17)		Freiheitsgrade	t-Test	Effektgröße
	M	SD	M	SD			
Tag 1	164.42	62.90	181.54	70.30	32	-0.75	0.26
Tag 2	131.68	65.38	129.32	43.42	32	0.12	0.04

4.4.4 Geschlechtseffekte

Nach Bortz (1999, S.138) verursachen selbst deutlich verschiedene Stichprobenumfänge beim t-Test keine Verzerrungen, wenn die Varianzhomogenität gegeben ist. Levene´s Test ist für beide Tage nicht signifikant, womit Varianzhomogenität angenommen werden kann, also die Ergebnisse des t-Tests in Tabelle 13 reliabel sind.

Tabelle 13: Mittelwerte und Standardabweichungen der Variable „Geschlecht“ mit den Ergebnissen des t-Tests für unabhängige Stichproben an beiden Tagen zwischen den Gruppen „Frauen“ und „Männer“.

Messzeitpunkt	Frauen (n = 27)		Männer (n = 7)		Freiheitsgrade	t-Test	Effektgröße
	M	SD	M	SD			
Tag 1	181.51	69.07	140.10	43.45	32	1.50	0.74
Tag 2	139.29	54.70	96.60	42.20	32	1.91*	0.88

Anmerkung: * $p < .05$ (einseitig)

Die Ergebnisse zeigen einen signifikanten Unterschied am zweiten Tag zwischen Frauen und Männern $t(32) = 1.91$, $p < 0.05$ (einseitig). Männer zeigen also signifikant bessere Resultate am zweiten Tag. Betrachtet man die Mittelwerte der beiden Gruppen, stellt man aber fest, dass der Lerngewinn von Tag 1 zu Tag 2 gleich ist. Männer erreichen also einen besseren Mittelwert, aber lernen nicht effizienter im Umgang mit der WPP. Erklärt werden könnte der signifikante Unterschied zwischen beiden Gruppen damit, dass mehr Männer das Internet nutzen, sie mehr Affinität zu Computern haben und somit auch die größere Vorerfahrung mitbringen (vgl. auch 4.4.5). Dass Männer die besseren komplexen Problemlöser sind, kann man aus den Ergebnissen nicht schlussfolgern.

4.4.5 Computervorerfahrung

Verschiedene Leistungen bei der Bearbeitung eines komplexen Problems aufgrund von Unterschieden bei der Computervorerfahrung sind in einigen Studien berichtet worden (vgl. Schoppek & Putz-Osterloh, 2003). Die Teilnehmer/innen konnten in der vorliegenden Diplomarbeit auf drei Skalen angeben, wie hoch sie ihre Vorerfahrung im Umgang mit Computern einschätzten (vgl. 3.1.3). In die Analyse gehen die Werte nur einer Skala ein:

Ich habe gute Computerkenntnisse.

Starke Ablehnung Ablehnung Neutral Zustimmung Starke Zustimmung

Die Extremgruppen, also die Leistungen der Teilnehmer/innen, die „Starke Ablehnung“ und „Ablehnung“ angekreuzt haben, werden in der Analyse mit den Leistungen derjenigen, die „Zustimmung“ und „starke Zustimmung“ angekreuzt haben, verglichen (vgl. Brehmer, 1996). Die Teilnehmer/innen, die „Neutral“ angekreuzt haben ($n = 12$), werden in dieser Extremgruppen-Analyse ausgeschlossen.

Tabelle 14 zeigt einen signifikanten Unterschied mit $t(20) = 1.85$, $p < 0.05$ am ersten Tag und einen signifikanten ($t(12.40) = 2.50$, $p < 0.01$) Unterschied für die beiden Extremgruppen am zweiten Tag. Der t-Wert am zweiten Tag wurde mit 12.40, anstatt 20 Freiheitsgraden berechnet, da Levene's Test signifikant ist.

Hypothese (7) kann somit als teilweise bestätigt angesehen werden. Computervorerfahrung hat offensichtlich einen entscheidenden Einfluss auf die Leistung bei einer computersimulierten

Problemlöseaufgabe. Dies gilt selbst, wenn die Bedienung der Aufgabe, wie in der WPP, sehr einfach gehalten ist.

Interessant ist, dass die Teilnehmer/innen im Durchschnitt seit 5,3 Jahren ($n = 30$) einen Computer benutzen. Diese relativ lange Vorerfahrung könnte zumindest teilweise den Einfluss der beiden Treatment-Bedingungen verdecken. Die Vorerfahrung bezüglich der Nutzung eines Computers reichte von 2 Wochen bis zu 14 Jahren ($n = 30$).

Bei der Häufigkeit der Benutzung ($n = 30$) gaben drei (8,8 %) Teilnehmer/innen an, den Computer ein paar mal im Monat zu benutzen. Dreizehn (38,2 %) benutzen ihn mehrmals pro Woche und vierzehn (41,2 %) Teilnehmer/innen täglich. Die Werte liegen bei der Frage nach der Häufigkeit der Benutzung genauso wie bei der Anzahl der Jahre, die man schon mit Computern arbeitet, für nur 30 Teilnehmer/innen vor.

Tabelle 14: *Mittelwerte und Standardabweichungen der Extremwerte der Variable „Computervorerfahrung“ mit den Ergebnissen des t-Tests für unabhängige Stichproben an beiden Tagen.*

Messzeitpunkt	Low (n = 9)		High (n = 13)		Freiheitsgrade	t-Test	Effektgröße
	M	SD	M	SD			
Tag 1	188.89	69.89	147.38	34.86	20	1.85*	0.79
Tag 2	152.31	48.55	106.69	30.65	12.40	2.50**	1.15

Anmerkung: * $p < .05$ (einseitig); ** $p < .01$ (einseitig)

Insgesamt kann die Variable „Computervorerfahrung“ als sehr bedeutend bei der Analyse von komplexen Problemszenarien bewertet werden. Es wäre wünschenswert, eine standardisierte Erhebungsmethode hierfür bereitzustellen, zum besseren Vergleich von Ergebnissen (vgl. auch Schoppek & Putz-Osterloh, 2003). Kritisch anzumerken ist, dass Student/innen einen leichteren Zugang zu einem Computerarbeitsplatz haben als viele andere Menschen. Deshalb lassen sich die gewonnenen Ergebnisse nicht generalisieren.

4.4.6 Zusammenfassung

Die Auswertung der t-Tests bestätigt die Ergebnisse der globalen Analyse anhand der repeated measurement ANOVA (Tabelle 5) zum großen Teil. Es finden sich keine signifikanten Ergebnisse zwischen den Extremgruppen der Persönlichkeitsmerkmale. Jedoch gibt es signifikante Unterschiede bei der Variable Geschlecht und Computervorerfahrung. Beide wurden hier aber nur explorativ untersucht und könnten Gegenstand genauerer Analysen zukünftiger Arbeiten sein. Wichtig ist festzuhalten, dass Computervorerfahrung, wie in anderen Studien auch gezeigt wird (vgl. Schoppek & Putz-Osterloh, 2003), am meisten Varianz unter den Kovariaten aufklärt.

4.5 Qualitative Datenanalyse

Nachdem die Daten quantitativ untersucht wurden, schließt sich nun eine qualitative Analyse an. Dies geschieht im Sinne von Cohen (1994) als ein Versuch, die Daten näher zu verstehen und Muster erkennen zu können. Nachdem die Werte der Analyse von Verlusten versus Gewinnen im Verlauf der WPP besprochen wurden, werden in diesem Abschnitt grundlegende Aspekte für die qualitative Analyse der „missed gallons pro Deadline“ besprochen. Die verschiedenen Kategorien und die erste der drei verschiedenen Sortierungen wird dargestellt und Ideen zur Interpretation gegeben. Die Daten und Abbildungen der zweiten und dritten Sortierung befinden sich im Anhang C (S.107), da sie keine zusätzliche Information geben.

4.5.1 Verschlechterung und Verbesserung – wie wirken sie?

In Tabelle 15 sind die Ergebnisse eines t-Tests dargestellt, bei dem die beiden Gruppen danach eingeteilt wurden, wie oft sie im Verlauf der Durchgänge weniger (gain) oder mehr (loss) missed gallons erhielten. Die Idee dahinter ist – analog zu Schoppek und Putz-Osterloh's Modell (Abbildung 20, S.112) – , dass ein Feedback über die Leistung während der Bearbeitung bedeutende Effekte auf die Motivation und somit die Problemlöseleistung hat. Es wurde ausgezählt, ob sich die Teilnehmer/innen zum nächsten Durchgang verbessert oder verschlechtert haben. Danach wurden sie, je nachdem, welcher Wert größer war, in die Gruppe „loss“ (mehr Verschlechterungen) oder „gain“ (mehr Verbesserungen) eingeteilt. Die Ergebnisse des t-Tests sind allerdings weder am ersten noch am zweiten Tag signifikant (Tabelle 15). Bemerkenswert ist, dass die Teilnehmer/innen, die mehr Verschlechterungen im Laufe der Bearbeitung zu verzeichnen hatten, die besseren Werte sowohl am ersten als auch am

zweiten Tag haben. D.h. trotz negativem Feedback über ihre Problemlöseleistung erreichen sie am Ende die besseren Werte. Es ist schwer dieses Ergebnis zu erklären. Eventuell hat ein negatives (Teil-)Feedback in der WPP einen anspornenden Charakter? Es mag aber auch sein, dass eine Verschlechterung bei einem Durchgang keine starke Wirkung auf die Motivation hat, da sie auf die schwierige Aufgabe attribuiert werden kann.

Tabelle 15: *Mittelwerte und Standardabweichungen der Variable „loss-gain“ mit den Ergebnissen des t-Tests für unabhängige Stichproben an beiden Tagen.*

Messzeitpunkt	loss (n = 11)		gain (n = 23)		Freiheitsgrade	t-Test	Effektgröße
	M	SD	M	SD			
Tag 1	161.21	44.61	178.61	74.69	32	-0.71	0.29
Tag 2	115.94	39.35	137.46	60.19	32	-1.077	0.43

4.5.2 Die Analyse der Verlaufskurven

Die WPP speichert die „missed gallons pro Deadline“ in ein Log-File. Diese Daten sind interessant, da jede/r Teilnehmer/in ein charakteristisches Ergebnismuster (Verlaufskurve) zu haben scheint. Zwei Verlaufskurven als Beispiel findet sich in Abbildung 8. Zur Interpretation der Verlaufskurve ist anzumerken, dass die meisten einen Peak am Ende eines Durchlaufs aufweisen. Die missed gallons werden in der WPP fortlaufend nach jeder Deadline berechnet. Dies bezieht sich aber auf die missed gallons in den Tanks unmittelbar vor der Ausleitung. Wasser, das in weiter hinten im Rohrsystem liegenden Tanks verblieben ist, wird ganz am Ende eines Durchlaufs gesammelt aufgerechnet (vgl. hierzu Abbildung 2, S. 35).

Ist ein Peak also klein, so wurde sich frühzeitig um die Leitungen gekümmert, in denen das Wasser einen langen Weg durch mehrere Tanks hat. Man könnte einen niedrigen Peak somit als die Fähigkeit beschreiben vorausschauend in der WPP zu handeln, also Zeitabläufe effizient zu berücksichtigen. Die beispielhaften Kurven in Abbildung 8 illustrieren dies. Der Peak am Ende eines Durchgangs lässt sich deutlich erkennen und auch, dass er sich im Laufe der Durchgänge in seiner Größe verändert.

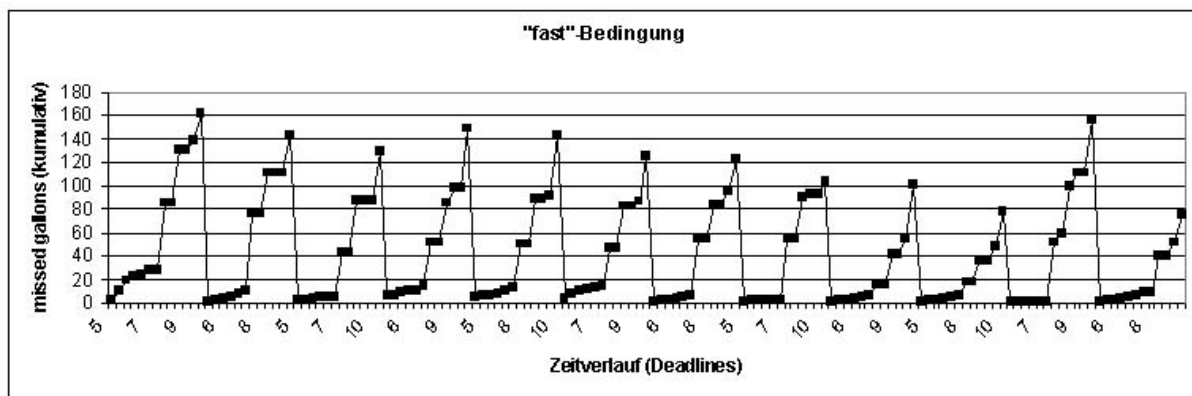
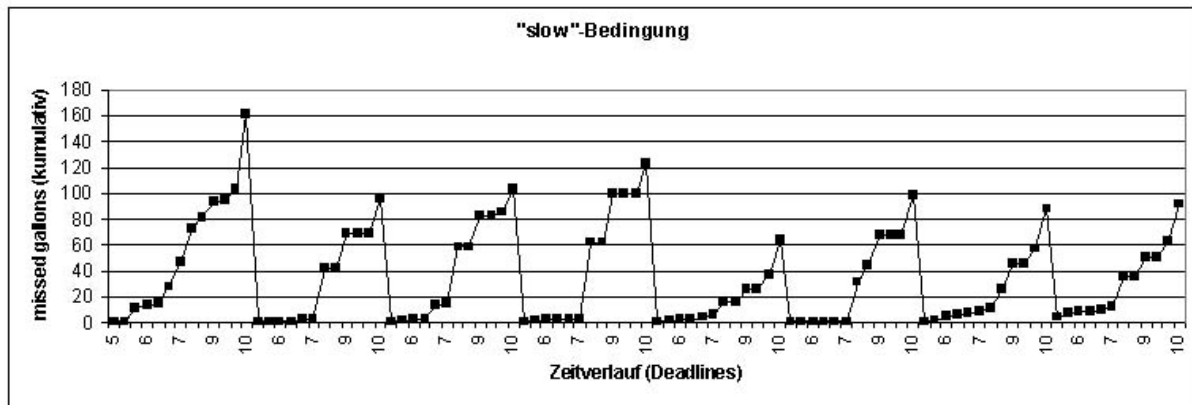


Abbildung 8: Zwei Beispiele von Verlaufskurven. Oben in der „slow“-Bedingung mit insgesamt acht Durchläufen und unten in der „fast“-Bedingung mit insgesamt zwölf Durchläufen. Die missed gallons werden kumulativ pro Deadline (halbstündlich ab fünf Uhr simulierter Zeit) angegeben.

Für die qualitative Analyse wurden die Verlaufskurven in verschiedene Kategorien aufgeteilt. Dies geschah für drei unterschiedlichen Sortierungen, die unten beschrieben sind. Nach den Aufteilungen wurde versucht, Muster – insbesondere im Zusammenhang mit den drei ausgewählten Persönlichkeitsdimensionen – in den Daten zu finden. Dieser Ansatz stellt einen Versuch dar, dem „Betriebsgeheimnis“ von guten Problemlösern in der WPP-Aufgabe auf die Spur zu kommen. Als besonderes charakteristisch gilt das auch in Abbildung 8 zu sehende „2-3-1“-Muster. Ungefähr bei der 8:00 Uhr Deadline steigen die missed gallons, bleiben aber bis zur nächsten Deadline gleich. Danach erfolgt wieder Zuwachs der missed gallons, der sich über drei Deadlines stabil hält. Schließlich endet die Aufgabe mit dem schon besprochenen Peak und somit einer weiteren Erhöhung der missed gallons.

Die qualitativen Daten wurden also in Bezug auf ihre Beziehungen zu den drei ausgewählten und mittels NEO-FFI erhobenen Persönlichkeitsmerkmalen untersucht (vgl. auch Schoppek &

Putz-Osterloh, 2003). Weiterhin wird die Gesamtleistung der Teilnehmer/innen als Indikator für die Güte der Bearbeitung in die Analyse einfließen. Die Ergebnisse werden zusammengefasst präsentiert. Die genauen Daten finden sich in Anhang C (S. 107).

Die erste Sortierung bezieht sich auf die Form der Kurve, die in fünf Kategorien eingeteilt wurde:

1. Das klassische „2-3-1“-Muster (n = 10; 30 %)
2. Annähernd „2-3-1“-Muster mit hohem Peak (n = 10; 30 %)
3. Annähernd „2-3-1“-Muster mit niedrigem Peak (n = 6; 18 %)
4. Eher gewölbte Kurve (n = 5; 14,71 %)
5. Kurve ohne Peak (n = 3; 8,82 %)

Die Leistung der Teilnehmer/innen in den beschriebenen Kategorien, operationalisiert durch die Anzahl der gesamten missed gallons, ergibt folgende Reihenfolge (siehe Abbildung 9): Beste Leistung erreichen die Teilnehmer/innen in der Gruppe 3 (annähernd „2-3-1“-Muster mit niedrigem Peak (n = 6)), gefolgt von der Gruppe 1 (das klassische „2-3-1“-Muster (n = 10)) und dann 2 (annähernd „2-3-1“-Muster mit hohem Peak (n = 10)). Anschließend folgt Gruppe 5 (Kurve ohne Peak (n = 3)) und schließlich Gruppe 4 (eher gewölbte Kurve (n = 5)). Abbildung 9 illustriert dieses Ergebnis. Es muss kritisch angemerkt werden, dass für die Operationalisierung auch andere Maße genommen werden können. So mag die Verbesserung vom Durchschnittswert des ersten Tages auf den Durchschnittswert des zweiten Tages ein interessantes Maß sein. Allerdings sollen hier die vorgeschlagenen „gesamten missed gallons“ im Mittelpunkt stehen. Zukünftige Arbeiten mögen andere Leistungs- bzw. Bewertungsmaße mit einbeziehen, die sich aus dieser Studie ableiten lassen.

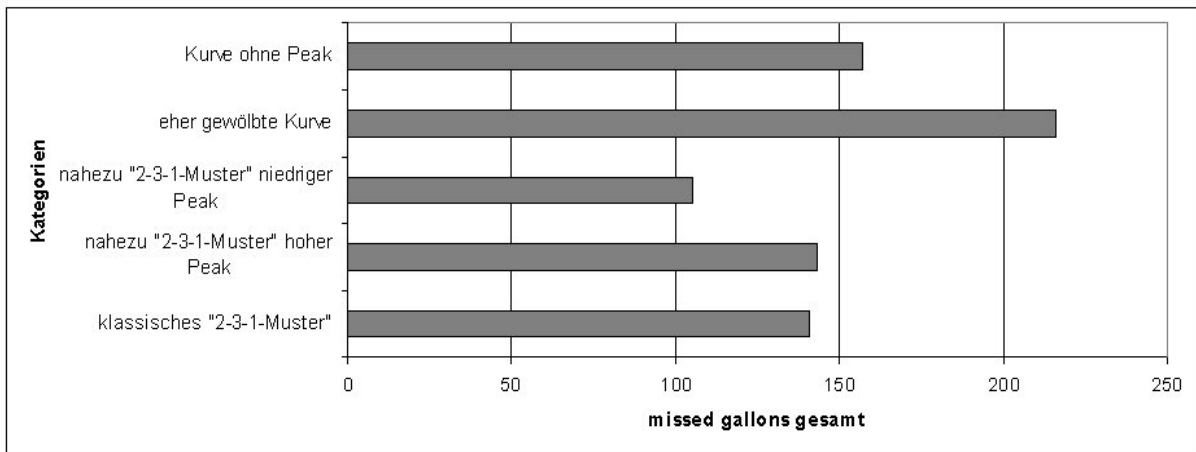


Abbildung 9: Bewertung der Kategorien der ersten Sortierung anhand der gesamten missed gallons

Für das Persönlichkeitsmerkmal Neurotizismus kann festgehalten werden, dass Teilnehmer/innen in der Gruppe mit den besten Leistungen (nahezu „2-3-1“ Muster mit niedrigem Peak) überwiegend einen niedrigen Neurotizismus-Wert aufweisen (vgl. Abbildung 10). Entsprechend zeichnen sich Teilnehmer/innen in der letzten Gruppe (eher gewölbte Kurve) durch hohe Neurotizismus-Werte aus. Die höchsten Neurotizismus-Werte finden sich in der Gruppe, bei denen das klassische „2-3-1“-Muster beobachtbar ist.

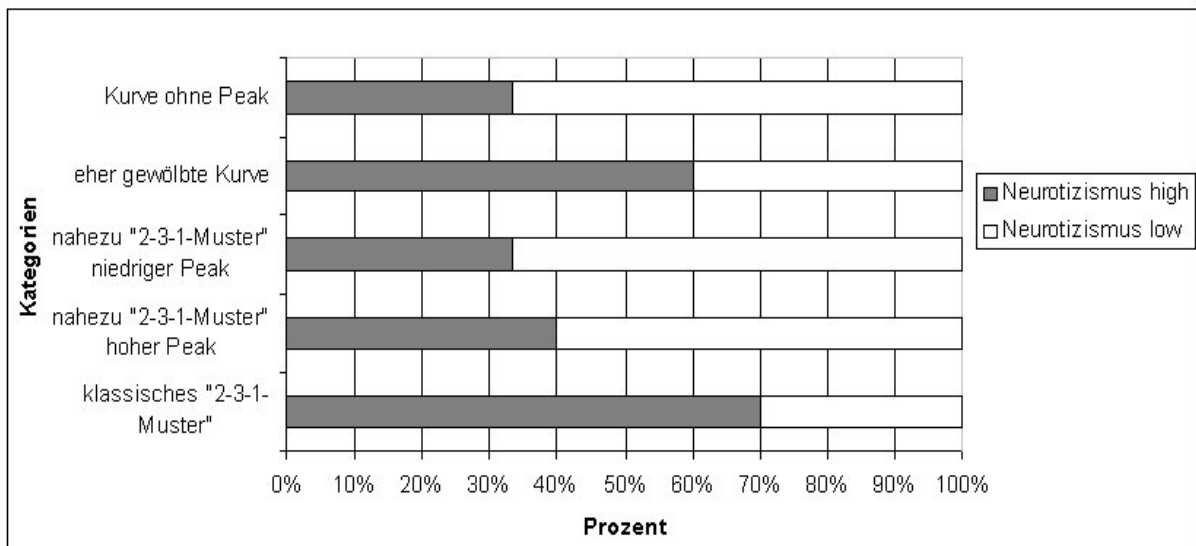


Abbildung 10: Verhältnis von Teilnehmer/innen mit hohen vs. niedrigen Werten beim Persönlichkeitsmerkmal Neurotizismus in der ersten Sortierung nach Kategorien.

Für das Persönlichkeitsmerkmal Extraversion kann gesagt werden, dass Teilnehmer/innen in der Gruppe mit den besten Leistungen überwiegend einen hohen Extraversions-Wert aufweisen

(vgl. Abbildung 11). Entsprechend zeichnen sich Teilnehmer/innen in der Gruppe mit den schlechtesten Leistungen durch niedrige Extraversions-Werte aus. Wiederum finden sich die Teilnehmer/innen mit dem höchsten Extraversions-Wert in der Gruppe, in der das klassische „2-3-1“-Muster vorherrscht. Ein hoher Extraversions-Wert hängt also in der qualitativen Analyse mit besseren Leistungen in der WPP und dem klassischen „2-3-1“-Muster zusammen.

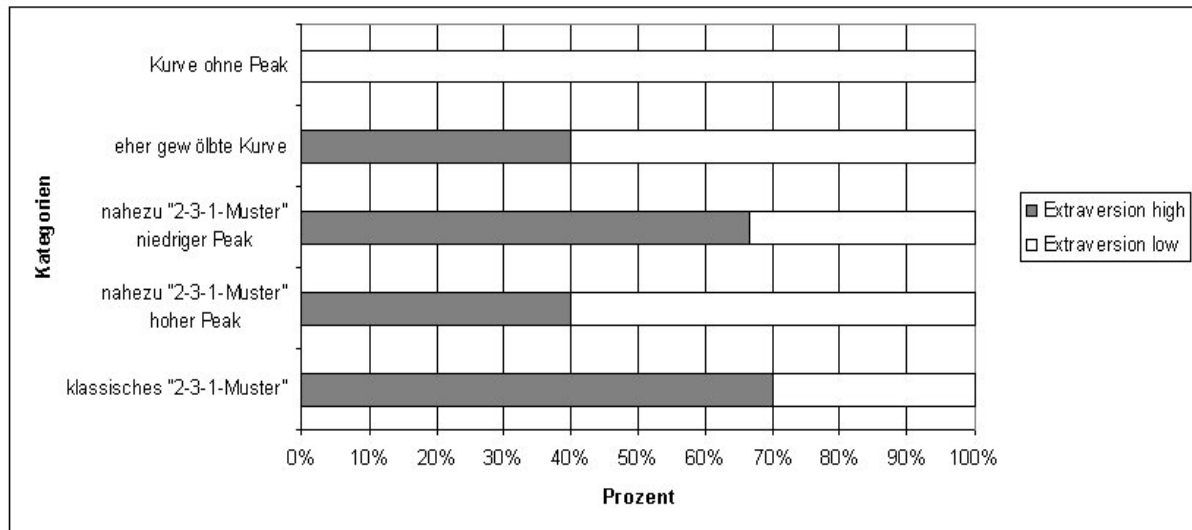


Abbildung 11: Verhältnis von Teilnehmer/innen mit hohen vs. niedrigen Werten beim Persönlichkeitsmerkmal Extraversion in der ersten Sortierung nach Kategorien.

Beim Persönlichkeitsmerkmal Offenheit für Erfahrungen sind Teilnehmer/innen in der Gruppe mit den besten Leistungen überwiegend einen niedrigen Offenheit für Erfahrungen-Wert aufweisen (vgl. Abbildung 12). Entsprechend zeichnen sich Teilnehmer/innen in der letzten Gruppe durch hohe Offenheit für Erfahrungs-Werte aus. Auch auf dieser Variable korrelieren hohe Werte zum klassischen „2-3-1“-Muster.

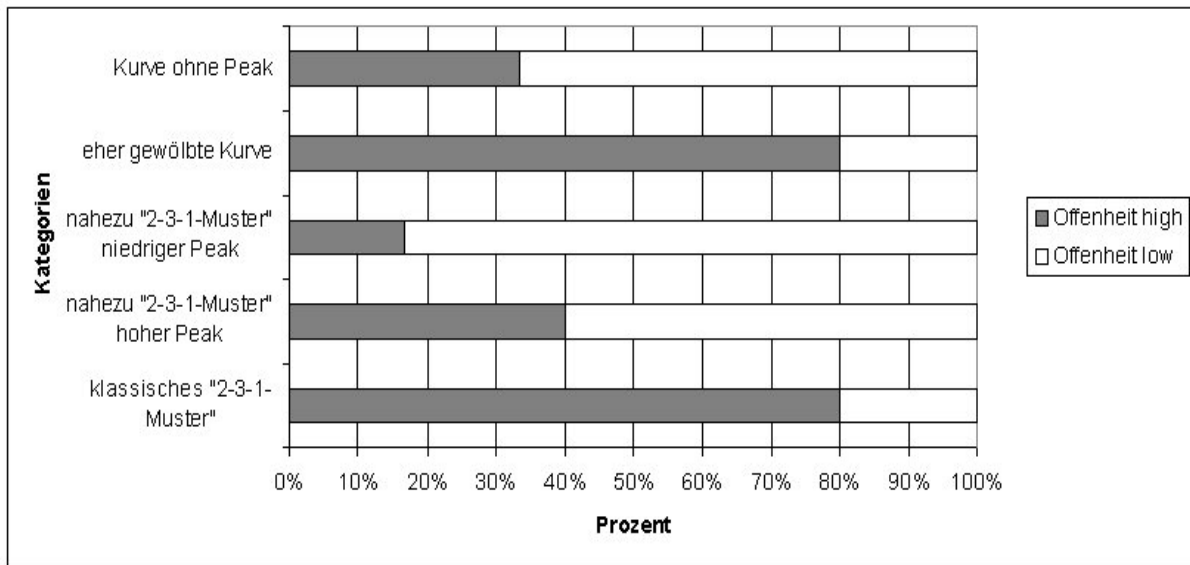


Abbildung 12: Verhältnis von Teilnehmer/innen mit hohen vs. niedrigen Werten beim Persönlichkeitsmerkmal Offenheit für Erfahrungen in der ersten Sortierung nach Kategorien.

Fazit: Das „nahezu „2-3-1-Muster“ mit niedrigem Peak“ scheint eine/n erfolgreiche/n Teilnehmer/in auszuzeichnen. Erfolgreiches Agieren in der WPP zeigten bei dieser Sortierung am häufigsten die Teilnehmer/innen mit niedrigen Neurotizismus-Werten, hohen Extraversions-Werten und niedrigen Offenheit für Erfahrungs-Werten. Interessant ist, dass die Teilnehmer/innen, die das klassische „2-3-1“-Muster zeigen, sich außerdem durch hohe Werte auf allen drei Persönlichkeitsvariablen auszeichnen. Dies ist bemerkenswert, da sich 30 % der Teilnehmer/innen in dieser Gruppe befinden, die außerdem das zweitbeste Gesamtergebnis anhand der missed gallons erreicht (siehe Abbildung 9).

Für die zweite Sortierung, die sich auf die Veränderungen der Kurven im Laufe der Bearbeitung bezieht, kann folgendes festgehalten werden. Erfolgreiches Agieren in der WPP ist bei den Teilnehmer/innen der Gruppe, die nur leichte Veränderungen in ihren Verlaufskurven zeigen, beobachtbar. Offenbar begünstigen die Einstellungen der WPP Strategiewechsel und starke Änderungen eher nicht. Möglichst wenige Entscheidungen scheinen also den Königsweg zu einem erfolgreichen Ergebnis zu bilden. Bei der Verallgemeinerung von Ergebnissen ist dies auf jeden Fall zu beachten. Bezeichnend ist, dass Teilnehmer/innen mit hohen Neurotizismus-Werten nur geringe Veränderungen vornehmen, die sich in den missed gallons pro Deadline widerspiegeln. Diese Leute scheinen lieber auf den schon belauften Wegen zu wandern und eher weniger Neues auszuprobieren. Allerdings sind selbst Teilnehmer/innen mit hohen Offenheits-Werten häufig in dieser Gruppe vertreten und nicht in der Gruppe die viele

Veränderungen zeigt. Insgesamt also differenziert diese Sortierung nicht sehr effektiv zwischen den drei gebildeten Gruppen.

Für die dritte Sortierung, die eine Modifikation (Reduzierung auf nur drei Gruppen) der ersten darstellt, zeigt sich, dass die Unterschiede zwischen den einzelnen Gruppen geringer werden. Des weiteren unterscheidet sich nun das Ergebnis von dem der ersten Sortierung. Erfolgreiche Teilnehmer/innen zeichnen sich hier durch einen niedrigen Extraversions-Wert und eine hohen Offenheits-Wert aus. Für die Persönlichkeitsvariable Neurotizismus ergeben sich keine Unterschiede bei allen drei Kategorien

4.5.3 Zusammenfassung

Alle drei Sortierungen hatten zum Ziel, die WPP von Gonzalez et al. unter einem anderen Blickwinkel zu beleuchten und insbesondere zu testen, inwieweit Effekte in der quantitativen Untersuchung auf Merkmale in der Bedienbarkeit oder des Ablaufs des Problemlöseszenarios zurückgeführt werden können. Artefakte und Verzerrungen der Ergebnisse können mittels solch einer Art der Herangehensweise herausgefunden werden.

Natürlicherweise spiegeln qualitative Untersuchungen subjektive Einstellungen wider, was sich hauptsächlich bei der Art der Kategorienbildung niederschlägt. Es wurden drei verschiedene Sortierungen verwendet, wobei die dritte eine Modifikation der ersten Sortierung darstellt.

Die Ergebnisse der drei Sortierungen unterscheiden sich bezüglich der Einflüsse der drei ausgewählten Persönlichkeitsmerkmale. Erfolgreiche Teilnehmer/innen zeichnen sich in der ersten Sortierung durch niedrige Neurotizismus-, hohe Extraversions- und niedrige Offenheits-Werte aus. In der zweiten Sortierung bezüglich der Veränderung der Kurve, durch niedrige Neurotizismus-, niedrige Extraversion- und hohe Offenheits-Werte. Erfolgreiche Teilnehmer/innen in der dritten Sortierung haben niedrige Extraversion- und hohe Offenheits-Werte wohingegen sich die Teilnehmer/innen im Bezug auf Neurotizismus nicht unterscheiden.

Zusammengefasst scheinen niedrige Neurotizismus-Werte am ehesten hilfreich für ein erfolgreiches Bearbeiten dieser komplexen Computersimulation zu sein. Für die beiden anderen Persönlichkeitsmerkmale ergibt sich je nach Sortierung ein anderes Bild. Bezüglich der Muster der Verlaufskurven lässt sich festhalten, dass das „2-3-1-Muster“ am häufigsten zu sehen ist. Man kann vermuten, dass die Einstellung der WPP dieses Muster begünstigt. Um aber

festzustellen, inwieweit die Teilnehmer/innen schon durch die Aufgabe in bestimmte Bearbeitungsmodi gezwungen werden, müsste man die Ergebnisse der vorliegenden Arbeit mit Daten vergleichen, die mit einer anderen Einstellung der WPP gewonnen wurden. Generell scheint eine erfolgreiche Strategie zu sein, den Peak möglichst gering zu halten, also sich möglichst frühzeitig um kritische Deadlines zu kümmern, die viel Zeit brauchen, bis in ihnen das Wasser abgelaufen ist. Dies gelingt den Teilnehmer/innen in der „fast“-Bedingung am besten (vgl. Anhang C). Weiterhin scheint es ratsam, bei der Bearbeitung möglichst wenig Änderungen vorzunehmen. Offenheit für Erfahrung und hohe Werte auf der Variable Extraversion sind in einem solchen Fall eher störend, was sich in der ersten Sortierung widerspiegelt. Bei der ersten Sortierung unterscheiden sich die Kategorien am meisten bezüglich der gesamten missed gallons. Dies ist für die beiden anderen Sortierungen nicht der Fall, was deren Interpretation uneindeutig macht.

4.6 Zusammenfassung der Ergebnisse

Im folgenden werden zunächst die Ergebnisse der quantitativen und dann der qualitativen Analysen zusammengefasst.

Das Ergebnismuster von Gonzalez et al. konnte zum Großteil nicht repliziert werden. Die Teilnehmer/innen der „fast“-Bedingung zeigen die besseren Ergebnisse bei der Bearbeitung der WPP. Somit muss man das IBLT-Modell von Gonzalez et al. in Frage stellen. Außerdem scheint die WPP je nach Einstellung bestimmte Bearbeitungsmodi vorzugeben. Das müsste aber weitergehend erforscht werden. Die Ergebnisse der repeated measurement Analyse zeigen eine signifikante Veränderung der Leistung über die Durchgänge innerhalb der Teilnehmer/innen. Diese Leistung variiert mit der Bedingung, weshalb ein signifikanter Effekt für die Interaktion zwischen Durchgang und Bedingung – wie bei Gonzalez – beobachtbar ist. Allerdings zeigt sich bei der vorliegenden Analyse kein signifikanter Effekt für die ausgewählten Kovariaten.

Als Erklärung für die Ergebnisse kann das Modell von Omodei und Wearing herangezogen werden. Sie schlagen vor, dass mehr Erregung auch zu höherer Aufmerksamkeit führt. Dies wäre eine mögliche Erklärung dafür, dass die Teilnehmer/innen in der „fast-Bedingung“ bessere Leistungen vorweisen können.

Da Gonzalez keine genauen Angaben zu ihrer Stichprobe macht, könnte ein Vergleich der Zusammensetzung der beiden Stichproben aufschlussreich sein. Des Weiteren ist die Vergütung

des Experiments bei Gonzalez mit 50\$ pro Teilnehmer/in ungleich höher als in der vorliegenden Diplomarbeit. Es mag weiterhin interessant sein, die Ergebnisse bezüglich der verschiedenen Kulturen (im Sinne von Strohschneider & Güss, 1999) zu vergleichen.

Man könnte auch eine andere Sicht der Auswertung testen: So wäre es interessant angesichts der Ergebnisse von Canast, Quesada, Antoli und Fajardot (2003) zu überprüfen, welchen Unterschied es macht, dass nur für die Teilnehmer/innen in der „slow“ Bedingung am zweiten Tag die Art und Weise der Simulation geändert ist. Es ist denkbar, dass dies insbesondere bei Teilnehmern/innen mit einem hohen Neurotizismus-Wert einen Effekt verursacht, der zu Leistungseinbußen führt. Für die Teilnehmer/innen der „fast“-Bedingung war die Aufgabe an beiden Tagen die gleiche, sie mussten sich nicht an etwas anderes (einen anderen Kontext oder eine andere Geschwindigkeit) anpassen. Dies würde eine völlig andere Sicht auf die Resultate werfen und es wäre plausibel, dass die Teilnehmer/innen in der „slow“-Bedingung mit hohen Neurotizismus-Werten im Durchschnitt schlechtere Ergebnisse erzielten, als die in der „fast“-Bedingung. Tatsächlich haben für die „slow“-Bedingung die Teilnehmer/innen mit hohen Neurotizismus-Werten einen höheren Mittelwert (132.24 (43.99)) als diejenigen mit niedrigen Neurotizismus-Werten (109.04 (33.90)). Dieser Unterschied ist allerdings mit einem t-Test für unabhängige Stichproben gerechnet nicht signifikant ($t(15) = 1.22, p = 0.12$).

Ähnlich wie bei Gonzalez et al. (2003) gab es signifikante Verbesserungen zwischen den beiden Tagen, es wurde also gelernt. Einige Fragestellungen wurden explorativ untersucht. Für die ausgewählten Persönlichkeitsmerkmale ergibt sich keinerlei signifikanter Effekt. So muss auch die Hypothese für das vorhergesagte Ergebnismuster beim Persönlichkeitsmerkmal Extraversion als nicht bestätigt angesehen werden.

Männliche Teilnehmer zeigen bei der Bearbeitung der WPP am zweiten Tag eine bessere Leistung als Frauen. Dieses Ergebnis müsste man aber weitergehend erforschen, da es hier nur explorativ untersucht wurde. Außerdem ist der Lerngewinn zwischen Tag1 und Tag2 für Frauen und Männer gleich. Inwieweit also andere Faktoren, wie Vorwissen oder größeres Interesse für Technik im Allgemeinen dort mit hineinspielen, ist unklar und muss bei der Interpretation beachtet werden.

Weiterhin gab es bei der explorativen Analyse der Computervorerfahrung signifikante Ergebnisse für beide Tage. Das Computervorerfahrung eine wichtige Rolle spielt, zeigen viele andere Studien zum komplexen Problemlösen (vgl. Schoppek & Putz-Osterloh, 2003).

Die qualitative Analyse der Rolle von Gewinnen und Verlusten während der Bearbeitung ergab keinen signifikanten Effekt. Bei der Analyse der Verlaufskurven zeigte sich, dass Teilnehmer/innen mit niedrigen Neurotizismus-Werten am ehesten gute Ergebnisse bei der Bearbeitung der WPP erreichten. Diese Leute verändern kaum ihre Strategie, verbessern sich aber im Rahmen der Strategie mit jedem Durchlauf. Für die beiden anderen Persönlichkeitsmerkmale sind die Ergebnisse uneindeutig. Besonders häufig lässt sich ein bestimmtes Muster bei den missed gallons pro Deadline beobachten, das weiter oben als klassisches „2-3-1-Muster“ beschrieben wurde. Wichtig ist, dass es ratsam scheint, möglichst wenige Veränderungen bei der Bearbeitung vorzunehmen. Dies wirkt sich günstig auf die Gesamtleistung aus. Für eine Verwendung der WPP im diagnostischen Bereich müsste abgeklärt werden, inwieweit die Einstellungen der WPP solch ein Muster erzwingen, oder ob die Teilnehmer/innen ihre Möglichkeit Strategien zu wechseln eher nicht nutzen.

5 DISKUSSION

"...man's truth is never absolute
because the basis of fact is hypothesis."
(Ch. S. Pierce, zitiert nach Fischer, 2001)

5.1 Allgemeines

Die Diskussion, welche Komponenten zu guten Entscheidungen führen, kann ausgeweitet werden auf eine soziologische und auch epochale Dimension und berührt auch andere Forschungsgebiete bei einem ganzheitlichen Blick. Neben Kybernetik spielen auch Philosophie, Mathematik (vgl. Luger, Lewis und Stern, 2002), Politik und Pädagogik eine wichtige Rolle. Neben der Erforschung von Kognitionen scheinen die Emotionen eine bestimmende Funktion im Entscheidungsprozess einzunehmen (vgl. Damasio, 1994, zitiert nach Nuber, 2000; ebenso Ciompi, 1998, S. 68ff). Debatten um das Thema „Weisheit“ (Staudinger, 2001), aber auch um das Thema „Fehler“ (Schaub, 2001) und den Umgang von Menschen mit Zeit fließen ebenso mit ein, wie die Themen Technik/Mensch-Interaktion (vgl. Geißler, Töpfer und Zwickel, 1999), Wahrnehmungsexperimente und der Einfluss, den Entscheidungen auf Politik und Lebensgestaltung haben. Einzelne Experimente können sich aber nur auf Ausschnitte und bestimmte Foki konzentrieren.

Mathematik und Kybernetik spielen eine wichtige Rolle bei der Simulation von Lernprozessen, wie bei Luger, Lewis und Stern (2002), die ein spannendes mathematisches Modell zur Abbildung von Wissen und Lernprozessen entwickelt haben, das sich einer konstruktivistischen Perspektive bedient. Epistemologische Fragen – wie die Welt in unseren Kopf kommt, wie Wissen beschaffen ist und benutzt wird – werden von ihnen untersucht. „Model refinement“, also der Aufbau und das Ändern einer Repräsentation der Welt, wird von ihnen wie auch in den Arbeiten von Gonzalez et al. als eine wichtige Möglichkeit gesehen, Lernprozesse abzubilden und untersuchen zu können.

Schoppek und Putz-Osterloh diskutieren das in Abbildung 13 dargestellte Modell und die Gründe, warum Untersuchungen beim komplexen Problemlösen, die Korrelationen zu individuellen Unterschieden finden wollten, oft scheitern müssen. Im Modell sieht man, dass diese nicht direkt auf die Problemlöseleistung wirken. Das in der vorliegenden Arbeit untersuchte Design zeigt im Einklang mit diesem (post-hoc hinzugezogenem) Modell keine signifikanten Effekte für die ausgewählten Kovariaten, bis auf die Effekte der

Computervorerfahrung, die explorativ untersucht wurden. Anhand der Ergebnisse der repeated measurement Analyse lässt sich vermuten, dass die Kovariaten als Moderator für die Interaktion von Durchgang mit Bedingung wirken.

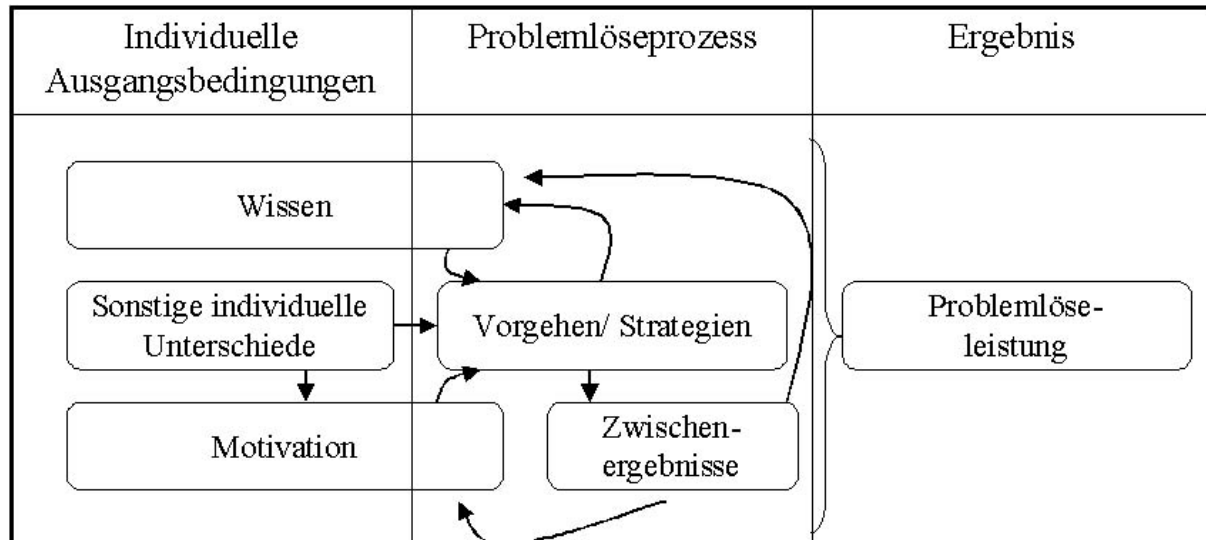


Abbildung 13: Wirkbeziehungen zwischen Variablengruppen im Problemlöseprozess nach Schoppek und Putz-Osterloh (2003).

5.2 Grenzen der Studie und Kritik am IBLT-Modell

Man kann aus der Studie von Ben Zur und Bresnitz (1981) folgern, dass sich Persönlichkeitsmerkmale insbesondere in der „slow“-Bedingung auswirken, da in der „fast“-Bedingung mehr auf die äußeren Erfordernisse reagiert wird. Da die Ergebnisse der „fast“-Bedingung ähnlich der von Gonzalez et al. (2003) simulierten Ergebnisse für die Zeit-Heuristik sind, liegt der Schluss nahe, dass die experimentelle Manipulation in der „slow“-Bedingung nicht gewirkt hat. D.h. ein Tag für die Übungsphase mag nicht genug sein, bei zwei Tagen, an denen jeweils zwei mal für 24 Minuten gespielt wird, könnte der Effekt sichtbar sein.

Ein weiterer Punkt ist, dass die Stichprobe ein „convenient sample“ war, die hauptsächlich aus Studenten/innen bestand. Interessant wäre die gleiche Studie mit als geistig behindert eingestuften Jugendlichen durchzuführen, mit Top-Managern oder auch mit älteren Menschen oder mit Menschen aus verschiedenen sozialen Schichten. Es kann erwartet werden, dass fehlende Vorerfahrungen im Umgang mit Computern und mit Testsituationen ganz andere Ergebnisse liefern könnten.

Aufgrund der Empirie kann das IBLT-Modell in dieser Arbeit als nicht bestätigt angesehen werden. Der Aufbau und die Ideen des Modells, die sich aus verschiedenen Disziplinen zusammensetzen, sind beeindruckend, dennoch konnten die Ergebnisse von Gonzalez nicht repliziert werden. Ein wichtiger Kritikpunkt am IBLT-Modell ist, dass das Auflösungs-niveau der einzelnen Wissens-einheiten nicht spezifiziert ist. D.h., wieso sollte das Gehirn ausgerechnet jede einzelne Entscheidung mit den angegebenen Phasen abspeichern, warum nicht auch Zwischenschritte (vgl. Retzer, 2002)? Eventuell sind im Sinne von Schwarz (2000), der in seinem Artikel erläutert, wie – durch Experimente und Theorien bestätigt – Emotionen und Stimmungen Einfluss auf Entscheidungssituationen nehmen können, vorhergehende oder nachfolgende Prozesse handlungsrelevanter? March (1978, zitiert nach Schwarz, 2000) geht so weit, dass er sagt, dass alle Entscheidungen Vorhersagen über Emotionen, die sie mit sich bringen, beinhalten. Somit sind Entscheidungen und Emotionen untrennbar miteinander verknüpft. Gonzalez et al. (2003) haben diese Komponente und die Komponente der Persönlichkeitsdimensionen am Entscheidungsprozess nicht berücksichtigt.

Kritisch ist weiterhin, dass Gonzalez ihre Stichprobe nicht genau beschreibt. Es mag sein, dass sich daraus Unterschiede in und zwischen den Ergebnissen erklären lassen, insbesondere, weil die Computervorerfahrung eine entscheidende Rolle spielt (vgl. auch Süß, 1999). Bei der Beurteilung und dem Vergleich der Ergebnisse ist es wichtig, die Unterschiede im Design zu beachten. Auch die 50\$ als Incentive bei Gonzalez mögen einen Unterschied bei der Bearbeitung der WPP ausmachen. Des weiteren rechnet Gonzalez in ihrer zweiten Studie mit Prozentwerten, welche die Lösegüte bei der WPP angeben. Durch diese Änderung ist ein Vergleich zwischen beiden Studien, denen ein ähnliches Design zu Grunde liegt, sehr schwer möglich. Bezüglich der SDU-Einheiten ist anzumerken, dass sie keine Differenzierung zwischen kurzzeitigem und langfristigem Nutzen vornehmen. Dieser Unterschied wäre aber wichtig, da der Umgang mit Zeit bei komplexen Problemen eine entscheidende Rolle spielt (Brehmer, 1996). In der qualitativen Analyse wurde dies berücksichtigt und ein Maß für den Umgang mit Zeit vorgeschlagen.

Kritikpunkte an der komplexen Problemlöseforschung stellt Funke (2003) dar. Diese gelten natürlich auch für die vorliegende Diplomarbeit gilt, werden aber nach und nach durch wachsende Forschungstätigkeiten reduziert. Sie werden auf die vorliegende Diplomarbeit bezogen.

- Theoriearmut: Durch das Einbringen der IBLT von Gonzalez et al. (2003) wird ein Versuch zur Abhilfe bei diesem Kritikpunkt unternommen.
- Mangelnde Integration anderer psychologischer Teildisziplinen: Hier sollen Kognitionspsychologie und Differentielle Psychologie vermischt werden. Außerdem fließen Ansätze aus der Mathematik bei der Konstruktion der IBLT mit ein.
- Fehlende fachübergreifende Kooperationen: Dies wird in der vorliegenden Arbeit kaum berücksichtigt. Gonzalez arbeitet mit einem Team von Programmierern, die die WPP entwickelten. Allerdings fehlen Bezüge zu Philosophie, Soziologie, Operational Research, Logik, Politik, und anderen Fachgebieten, die durchaus Nutzen aus der komplexen Problemlöseforschung ziehen können. So mag zum Beispiel der Einfluss von Werten – als soziologische Perspektive – eine große Rolle spielen.
- Unzureichende Nutzung systemtheoretischer Ansätze: Dies wird hier kaum beachtet.
- Fehlen einer Taxonomie von Problem- bzw. Systemtypen: Mit der WPP wird ein zusätzlicher Aufgabentyp bereitgestellt, der leider nicht die formalen Strukturen, die Funke (2001) vorschlägt, zu Grunde legt. Dennoch bietet die Einkleidung und Arbeitsweise des Szenarios die Möglichkeiten, durch das Ändern der Einstellungen verschiedene Problemtypen zu kreieren.
- Mangelnde Beachtung des Messfehlers und der Maßbestimmung: Die WPP gibt nur eine abhängige Variable aus, was Vorteile und Nachteile mit sich bringt. Deshalb wurde die WPP in der vorliegenden Arbeit auch zusätzlich qualitativ anhand der Verlaufskurven untersucht.
- Auch nicht-signifikante Befunde veröffentlichen: Die nicht-signifikanten Resultate sind in der Diplomarbeit veröffentlicht. Sie werden mit Gonzalez ausgetauscht, so dass diese als Feedback in die Forschungen von Gonzalez mit eingehen.
- Bessere Versuchsplanung und –auswertung: Die Auswertung einer repeated measurement Analyse mit Kovariaten ist ein nicht ganz unproblematisches Verfahren. Jedoch scheint sie aufgrund der prozesshaften Analyse insbesondere für die komplexe Problemlöseforschung geeignet.

6 AUSBLICK

„Zukunft ist das Produkt menschlicher Entscheidungen.“

Matthias Horx

Dörner et al. (1999) weisen auch auf kulturelle Unterschiede bei komplexen Denk- und Problemlöseprozessen hin. Lernerfahrungen spielen eine große Rolle und die Verfügbarkeit von Problemen, um sich ausprobieren zu können. Dies zu trainieren sei die Aufgabe vom Spielen bei Kindern laut Spitzer (1996). Die vorliegende Studie kann zum Verständnis des kulturellen Einflusses insofern beitragen als dass die erhobenen Daten von Gonzalez weiter verwendet und in kulturvergleichenden Studien eingesetzt werden können.

Bei der Anwendung der WPP könnte man die Kompatibilität zu anderen Betriebssystemen verbessern und eine Anleitung liefern. Die Installation ist nicht unproblematisch.

Bezüglich des Versuchsplans wäre es interessant, noch mehr verschiedene Zeitdruck-Levels zu testen, um den Effekt von Zeitdruck genauer differenzieren zu können. Als Theorierahmen lässt sich sehr gut vorstellen die „control theory“, wie sie Brehmer (1992, 1996) vorschlägt, einzuflechten. Die Kombination mit neueren prozessbezogenen Persönlichkeitsmerkmalen, die überdies in eine Theorie eingebettet sind, wie z.B. „need for cognition“, wären gut geeignet und interessant, um sie im Bezug auf die WPP zu verwenden.

Die weitere Forschung – insbesondere bei anderen Stichproben – mit der WPP scheint lohnenswert, jedoch muss man anhand der empirischen Ergebnisse der vorliegenden Studie das IBLT-Modell revidieren.

7 LITERATUR

- Amelang, M. & Bartussek, D. (1997). *Differentielle Psychologie und Persönlichkeitsforschung*. (4., überarbeitet und erweiterte Auflage). Stuttgart: Kohlhammer.
- Amelang, M. & Zielinski, W. (1994). *Psychologische Diagnostik und Intervention*. Berlin: Springer.
- Aha, D.W., Kibler, D., & Albert, M.K. (1991). Instance-based learning algorithms. *Machine Learning*, 6, 1, 37-66.
- Baron, R. M. & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychology research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173-1182.
- Beck, H. (2001). Das Recht auf Unvernunft. *Frankfurter Allgemeine Sonntagszeitung vom 4. November*, 44, 51.
- Beck-Bornholdt, H.-P. & Dubben, H.-H. (1997). *Der Hund, der Eier legt. Erkennen von Fehlinformation durch Querdenken*. Reinbeck bei Hamburg: Rowohlt.
- Benson, L., & Beach, L. R. (1996). The effects of time constraints on the prechoice screening of decision options. *Organizational Behavior and Human Decision Processes*, 67, 2, 222 - 228.
- Ben Zur, H., & Breznitz, S. J. (1981). The effect of time pressure on risky choice behavior. *Acta Psychologica*, 477, 89 – 104.
- Blickle, G. (1996). Personality traits, learning strategies, and performance. *European Journal of Personality*, 10, 337-352.
- Borkenau, P. & Ostendorf, F. (1993). NEO-Five-Factor Inventory according to Costa and McCrae Manual [NEO-Fünf-Faktoren Inventar (NEO-FFI) nach Costa and McCrae Manual], Göttingen: Hogrefe.
- Bortz, J. (1999). *Statistik für Sozialwissenschaftler* (5. Aufl.). Heidelberg: Springer.

- Buchner, A. (1999). Komplexes Problemlösen vor dem Hintergrund der Theorie finiter Automaten. *Psychologische Rundschau*, 50, 4, 206-212.
- Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta Psychologica*, 81, 211-241.
- Brehmer, B. (1996). Man as stabiliser of systems: From static snapshots of judgement processes to dynamic decision making. *Thinking And Reasoning*, 2 (2/3), 225-238.
- Canas, J.J., Quesada, J.F., Antoli, A., & Fajado, I. (2003) Cognitive flexibility and adaptability to environmental changes in dynamic complex problem-solving tasks. *Ergonomics*, 46, 5, 482-501.
- Capra, F. (1996). *Wendezeit. Bausteine für eine neues Weltbild* (5. Aufl.). München: Deutscher Taschenbuch Verlag.
- CiOMPI, L. (1998). *Affektlogik* (5. Aufl.). Stuttgart: Klett-Cotta.
- Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, 49, 12, 997-1003.
- Connelly, M.S., Gilbert, J.A., Zaccaro, S.J., Threfall, K.V., Marks, M.A., & Mumford, M.D. (2000). Exploring the relationship of leadership skills and knowledge to leader performance. *Leadership Quarterly*, 11, 1, 65-86.
- Costa, P. T. Jr. & McCrae, R. R. (1992). *Revised NEO Personality Inventory and NEO Five factor Inventory Professional Manual*. Odessa, FL: Psychological Assessment Resources.
- Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Sage: Thousand Oaks.
- Czienskowski, U. (1996). *Wissenschaftliche Experimente: Planung, Auswertung, Interpretation*. Weinheim: Psychologische Verlags Union.
- D'Amico, E.J., Neilands, T.B., & Zambarano (2001). Power analyses for multivariate and repeated measures design: a flexible approach using the SPSS MANOVA procedure. *Behavior Research Methods, Instruments, & Computers*, 33, 4, 479-484.

- Dörner, D. (1989). Die Logik des Mißlingens. Strategisches Denken in komplexen Situationen (1. Aufl.). Reinbek bei Hamburg: Rowohlt.
- Dörner, D., Schaub, H., & Strohschneider, S. (1999). Komplexes Problemlösen – Königsweg der Theoretischen Psychologie? *Psychologische Rundschau*, 50, 4, 198-205.
- Dunlap, Cortina, Vaslaw & Buke (1996). Einige Möglichkeiten zur Berechnung der Effektstärke. Verfügbar unter: <http://www.phil.uni-sb.de/~jakobs/seminar/vpl/bedeutung/effektstaerketool.htm> [05.05.2004]
- Duden, Etymologie (Bd. 7) (1997). Das Herkunftswörterbuch (Nachdruck der 2. Aufl.). Bearbeitet von Günther Drosdowski. Mannheim: Dudenverlag.
- Ernst, H. (2000). Für das Leben gibt es keine Kostümprobe. *Psychologie Heute*, 27, 1, 3.
- Field, A. (1998). A bluffer's guide to sphericity. Newsletter of the Mathematical, Statistical and Computing section of the British Psychological Society, 6, 1, 13-22. Verfügbar unter: <http://www.sussex.ac.uk/Users/andyf/research/articles/sphericity.pdf> [05.05.2004]
- Field, A. (2000). Discovering statisticc using SPSS for Windows. Advanced techniques for the beginner. London: Sage Publications.
- Fischer, H.R. (2001). Abductive reasoning as a way of worldmaking. *Foundations of Science*, 6, 361-383.
- Fischer, H.R. (Hrsg.) (in press). Die Wirklichkeit der Metapher. Weilerswist: Velbrück Wissenschaft.
- Feyerabend, P. (1978). Ausgewählte Schriften. Der wissenschaftstheoretische Realismus und die Autorität der Wissenschaften (Bd. 1). Vieweg: Braunschweig, Wiesbaden.
- Funke, J. & Buchner, A. (1992). Finite Automaten als Instrumente für die Analyse von wissensgeleiteten Problemlöseprozessen: Vorstellung eines neuen Untersuchungsparadigmas. *Sprache & Kognition*, 11, 1, 27-37.
- Funke, J., Buchner, A., Dörner, D., Süß, H.-M., & Vollmeyer, R. (1999a). Diskussionsrunde zum Themenheft „Komplexes Problemlösen“. *Psychologische Rundschau*, 50, 4, 229-233.

- Funke, J. (1999b). Komplexes Problemlösen – ein Blick zurück und ein Blick nach vorne –. *Psychologische Rundschau*, 50, 4, 229-233.
- Funke, J. (2001). Dynamic systems as tools for analysing human judgement. *Thinking And Reasoning*, 7, 1, 69-89.
- Funke, J. (2003). *Problemlösendes Denken*. Stuttgart: Kohlhammer.
- Frankfurter Allgemeine Sonntagszeitung vom 22. Februar 2004, Nr. 8, Seiten 27-33.
- Geißler, H., Töpfer, S. & Zwickel, J. (1999). Evaluation und Problemanalyse bei SAS- im Vergleich zu SPSS-Kursen an der Universität Heidelberg. Eine empirische Studie. Proceedings zur 3. Konferenz für SAS-Anwender in Forschung und Entwicklung 25.-26. Februar 1999 Ruprecht-Karls-Universität Heidelberg/Carina Ortseifen (Editor). Heidelberg, Universitätsrechenzentrum.
- Gigerenzer, G. & Todd, P. M. (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
- Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science*, 27, 591-635.
- Gonzalez, C., Qudrat-Ullah, H. & Lerch, F. J. Decision making in a dynamic, real-time environment : Learning under time pressure. *Manuscript submitted to Human Factors for publication*.
- Gonzalez, C. Decision making in a dynamic, real-time environment: Learning under high cognitive workload. *Manuscript submitted to Human Factors for publication*.
- Graziano, A. M. & Raulin, M. L. (2000). *Research methods. A process of inquiry* (4th edition). Needham heights, MA: Allyn & Bacon.
- Handke, P. (1974). *Als das Wünschen noch geholfen hat*. Frankfurt: Suhrkamp.
- Hays, W. L. (1994). *Statistics*. Orlando, FL: Harcourt Brace & Company.
- Heckhausen, H. (1980). *Motivation und Handeln. Lehrbuch der Motivationspsychologie*. Heidelberg: Springer.

- Helmreich, R. (2001). „Fehler sind meist eine Sache der Unternehmenskultur“. *Psychologie Heute*, 28, 1, 68-69.
- Hüther, G. (2004). *Bedienungsanleitung für ein menschliches Gehirn* (4. Aufl.). Göttingen: Vandenhoeck & Ruprecht.
- Humphreys M. S. & Revelle, W. (1984). Personality, motivation, and performance: a theory of the relationship between individual differences and information processing. *Psychological Review*, 91, 2, 153-184.
- Jungermann, H., Pfister, H.-R., & Fischer, K. (1998). *Die Psychologie der Entscheidung: Eine Einführung*. Heidelberg: Spektrum Akademischer Verlag.
- Kahneman, D. (2003). A perspective on judgement and choice. Mapping bounded rationality. *American Psychologist*, 58, 9, 697–720.
- Kerstholt, J.H. (1994). The effect of time pressure on decision-making behaviour in a dynamic task environment. *Acta Psychologica*, 86, 89-104.
- Keselman, J.C. & Keselman, H.J. (1990). Analysing unbalanced repeated measures designs. *British Journal of Mathematical and Statistical Psychology*, 43, 265-282.
- Keselman, H.J., Algina, J., Kowalchuk, R.K., & Wolfinger, R.D. (1999). A comparison of recent approaches to the analysis of repeated measurements. *British Journal of Mathematical and Statistical Psychology*, 43, 265-282.
- Kreuzig, H. W. (1981). Über den Zugang zu komplexem Problemlösen mittels prozeßorientierter kognitiver Persönlichkeitsmerkmale. *Zeitschrift für experimentelle und angewandte Psychologie*, 28, 2, S. 294-308.
- Kriz, W. K. (2000). *Lernziel: Systemkompetenz. Planspiele als Trainingsmethode*. Göttingen: Vandenhoeck & Ruprecht.
- Leahy, T. H. (2003). Herbert A. Simon. Nobel Prize in Economic Science 1978. *American Psychologist*, 58, 9, 753-755.
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, 95, 4, 492-527.

- Luger G. F., Lewis, J., & Stern, C. (2002). Problem solving as model refinement: towards a constructivist epistemology. *Brain, Behavior and Evolution*, 59, 87-100.
- Messik, S. (1996). Bridging cognition and personality in education: the role of style in performance and development. *European Journal of Personality*, 10, 353-376.
- Ngidi, D. P. & Sibaya, P. T. (2002). Black teacher's personality dimensions and work-related stress factors. *South African Journal of Psychology*, 32, 3, 7-15.
- Nuber, U. (2000). Mit Herz und Verstand. Wie Sie Entscheidungen treffen, die Sie später nicht bereuen müssen. *Psychologie Heute*, 27, 1, 20-27.
- Omodei, M.M. & Wearing, A.J. (1995). Decision making in complex dynamic settings: a theoretical model incorporating motivation, intention, affect, and cognitive performance. *Sprache & Kognition*, 14, 2, 75-90.
- Peterson, E. R., Deary, I. J., & Austin, E. J. (2003a). The reliability of Riding's cognitive style analytical test. *Personality and Individual Differences*, 34, 881-891.
- Peterson, E. R., Deary, I. J., & Austin, E. J. (2003b). On the assessment of cognitive style: four red herrings. *Personality and Individual Differences*, 34, 899-904.
- Probst, G. & Raisch, S. (2004). Die Logik des Niedergangs. *Harvard Business Manager*, 3, 37-45.
- Rawlings, D. & Carnie, D. (1989). The interaction of EPQ Extraversion with WAIS subtest performance under timed and untimed conditions. *Personality and Individual Differences*, 10, 4, 453-458.
- Retzer, A. (2002). *Passagen – Systemische Erkundungen*. Stuttgart: Klett-Cotta.
- Revelle, W., Amaral, P., & Turriff, S. (1976). Introversion/extraversion, time stress and caffeine: Effect on verbal performance. *Science*, 192, 149-150.
- Riding, R. J. (2003). On the assessment of cognitive style: a commentary on Peterson, Deary, and Austin. *Personality and Individual Differences*, 34, 893-897.
- Schäfer, A. (1998). Die Psychodynamik der Aktienkurse. *Psychologie Heute*, 25, 3, 38-43

- Schaub, H. (2001). Fehler sind menschlich...und doch oft vermeidbar. *Psychologie Heute*, 28, **1**, 62-67.
- Schein, E. H. & Coutu, D. L. (2002). Blut, Schweiß und Tränen – von der Angst zu lernen. *Harvard Business manager*, 5, 72-79.
- Schein, E. H. (2003). Angst und Sicherheit. *Organisationsentwicklung*, 3, 4-13.
- Schoppek, W. & Putz-Osterloh, W. (2003). Individuelle Unterschiede und die Bearbeitung komplexer Probleme. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 24, **3**, 163-173-
- Schwarz, N. (2000). Emotion, cognition, and decision making. *Cognition and Emotion*, 14, **4**, 433-440.
- Speiser, G. (2004). Der vergebliche Blick in Aladdins Lampe. *Frankfurter Allgemeine Sonntagszeitung vom 29 Februar*, 9, 62-63.
- Speiser, G. (2004). Da war doch noch was? *Frankfurter Allgemeine Sonntagszeitung vom 14 März*, 11, 64.
- Spring, M. (2001). Emotionen und Kontrollüberzeugungen beim komplexen Problemlösen. Eine experimentelle Untersuchung anhand des computersimulierten Problemlöseszenarios FSYS 2.0 *Unveröffentlichte Diplomarbeit am Psychologischen Institut der Ruprecht-Karls Universität Heidelberg*.
- Spitzer, M. (1996). Geist im Netz. Modelle für Denken, Lernen und Handeln. Heidelberg: Spektrum Akademischer Verlag.
- Staudinger, U. (2001). „Weisheit ist leicht zu erkennen – aber schwer zu erreichen“. *Psychologie Heute*, 28, **10**, 22-27.
- Strohschneider, S. & Güss, D. (1999). The fate of the moros: a cross-cultural exploration of strategies in complex and dynamic decision making. *International Journal of Psychology*, 34, **4**, 235-252.
- Süß, H.M. (1999). Intelligenz und komplexes Problemlösen. *Psychologische Rundschau*, 50, **4**, 220-228.

- Van der Linden, D., Sonnentag, S., Frese, M., & Van Dyck, C. (2001). Exploration strategies, performance, and error consequences when learning a complex computer task. *Behaviour & Information Technology*, 20, 3, 189-198.
- Vester, F. (2002). Die Kunst vernetzt zu denken. München: Deutscher Taschenbuch Verlag.
- Vollmeyer, R. & Funke, J. (1999). Personen- und Aufgabenmerkmale beim komplexen Problemlösen. *Psychologische Rundschau*, 50, 4, 213-219.
- Vorberg, D. & Blankenberger, S. (1999). Die Auswahl statistischer Tests und Maße. *Psychologische Rundschau*, 50, 3, 157-164.
- Wagner, H.-J. (2001). Objektive Hermeneutik und die Bildung des Subjekts. Mit einem Text „Die Philosophie von Charles Sanders Peirce als Philosophie der Krise“ von Ulrich Oevermann (Ulrich Oevermann (Hg.))(1. Aufl.). Weilerswist: Velbrück Wissenschaft.
- Werner, J. (1997). Lineare Statistik. Das Allgemeine Lineare Modell. Weinheim: Psychologische Verlags Union.
- Wilcox, R.R. (1998). How many discoveries have been lost by ignoring modern statistical methods? *American Psychologist*, 53, 3, 300-314.
- Wilcox, R.R., Keselman, H.J., Muska, J., & Cribbie, R. (2000). Repeated measures ANOVA: Some new results on comparing trimmed means and means. *British Journal of Mathematical and Statistical Psychology*, 53, 69-82.
- Wilkinson, L., & Task Force on Statistical Inference (2000). Statistical methods in Psychology Journals: Guidelines and explanations. *American Psychologist*, 21 S. Verfügbar unter: www.apa.org/journals/amp/amp54594.html [09.11.2003]
- Zeeman, E.C. (1976). Catastrophe theory. *Science*, 65-83.
- Zhang, Li-fang (2002). Thinking styles and the big five personality traits. *Educational Psychology*, 22, 1, 17-31.
- Zhang, Li-fang (2003). Does the big five predict learning approaches? *Personality and Individual Differences*, 34, 1431-1446.

8 ANHANG

Anhang A:

Tabelle 16: *Beispiel-Items aus dem NEO-FFI (Borkenau und Ostendorf, 1993)*

Skala	Beispiel-Item	Wert*
Neurotizismus	Ich bin nicht leicht beunruhigt.	0
	Wenn ich unter starkem Stress stehe, fühle ich mich manchmal, als ob ich zusammenbräche.	4
Extraversion	Ich bin kein gut gelaunter Optimist.	0
	Ich habe gerne andere Leute um mich herum.	4
Offenheit für Erfahrungen	Poesie beeindruckt mich wenig bis gar nicht.	0
	Ich habe Spass daran, mit Theorien oder abstrakten Ideen zu spielen.	4
Verträglichkeit	Ich bekomme häufiger Streit mit meiner Familie und meinen Kollegen.	0
	Ich würde lieber mit anderen zusammenarbeiten, als mit ihnen zu wetteifern.	4
Gewissenhaftigkeit	Ich verträdele eine Menge Zeit, bevor ich mit der Arbeit beginne.	0
	Ich versuche, alle mir übertragenen Aufgaben sehr gewissenhaft zu erledigen.	4

* Der Wert basiert auf der Antwort: „starke Zustimmung“

Anhang B: Instruktion für die Teilnehmer/innen.

Liebe/r Teilnehmer/in

Überblick: Bei diesem Versuch geht es darum herauszufinden, wie Menschen mit einem komplexen System (einer Anlage zum Verteilen von Wasser) umgehen. Dafür spielst du an zwei aufeinander folgenden Tagen das Computerszenario „Wasserpumpe“ jeweils ein paar mal durch. Am ersten Tag hast du eine kurze Probephase. Der kurze Fragebogen wird ebenso nur am ersten Tag bearbeitet.

Du bekommst 2,5 Versuchspersonenstunden sowie Kleinigkeiten vom dm-Markt. Außerdem werden drei original Ravensburger Spiele unter den (nur ca. 30) Teilnehmern/innen verlost. Nach den Semesterferien hängt am Vpn-Brett ein Zettel mit dem Code der Gewinner/innen.

Alle Daten werden anonym behandelt. Deshalb bitte bei „Name“ als **Code** auf dem Fragebogen eintragen:

- letzten Buchstabe des Vornamens deiner Mutter (z.B.: N)
- *Zahl des* Geburtsmonats deiner Mutter (z.B.: 01 für Januar)
- zweiter Buchstabe des Vornamens deines Vaters (z.B.: E)

Bitte fülle **jetzt** den **Fragebogen** aus.....

----- ... dann geht es wieder auf diesem Blatt weiter...-----

Bitte kreuze an (nach dem Fragebogen):

Ich habe gute Computerkenntnisse.

Starke Ablehnung *Ablehnung* *Neutral* *Zustimmung* *Starke Zustimmung*

Ich benutze einen Computer...

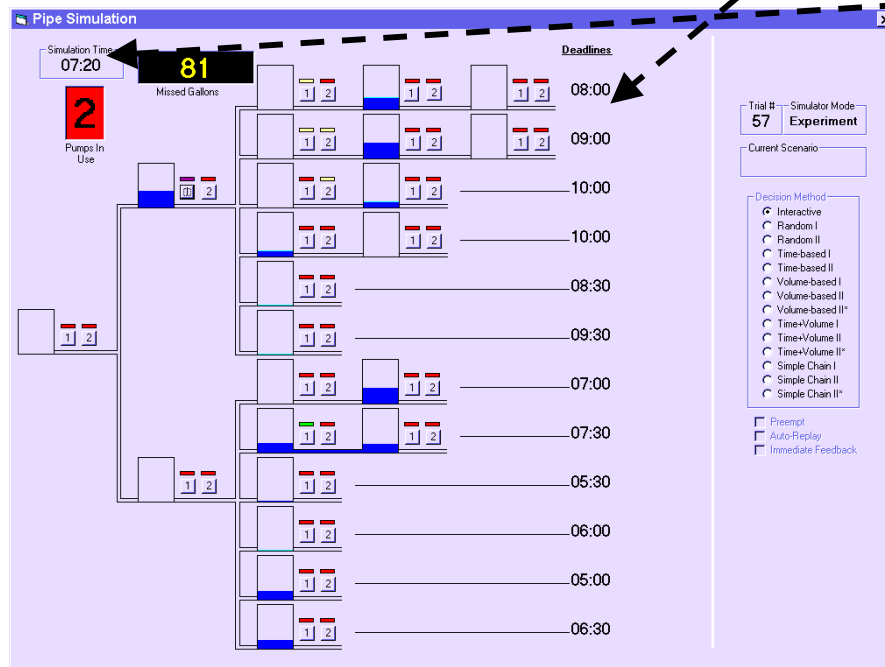
Fast nie *ein paar mal im Monat* *mehrmals in der Woche* *täglich*

Seit wieviel Jahren ungefähr arbeitest du schon mit einem Computer?

...dann drehe dieses Blatt um.

Genauer zum Computerszenarios „Wasserpumpe“:

Nach einer kurzen Probephase kann es los gehen. Die **Deadlines** sind an der rechten Bildschirmseite. Die Wasser-Tanks sind verbunden durch Leitungen, wobei das Wasser in blau dargestellt ist. Buttons für die zwei Pumpen je Tank befinden sich jeweils rechts daneben. Die **Zeit**



der Simulation ist in der linken oberen Ecke angezeigt, sowie die schon verpaßten Wasser-Gallonen. Außerdem befindet sich dort die Anzeige für die gerade in Gebrauch stehenden Pumpen (hier: zwei). Die Zeit läuft von 2:00 bis 10:00 (8 Simulationsstunden).

Du sollst in diesem Szenario Pumpen öffnen und schließen (durch **Anklicken der Kästchen** mit den Nummern drauf), um so das ganze Wasser durch das Leitungssystem zu befördern (**höchstens 5 Pumpen** zur gleichen Zeit). **Grüne** Farbe über Kästchen zeigt aktivierte Pumpen an.

Es kommt zu verschiedenen Zeiten Wasser von außen in das System in die Tanks, welches dann weitergeleitet werden muss. Nach dem Abstellen einer Pumpe dauert es kurze Zeit, bis sie „gewartet“ ist (angezeigt durch die Farbe **gelb**). Dies geschieht durch Anklicken, oder wenn der Tank davor leer geworden ist. Manchmal brauchen die Pumpen eine kurze Zeit zum Anlaufen (dann sind sie **lila**).

Ziel ist also, das ins Leitungssystem kommende Wasser **komplett** (von links nach rechts) zu befördern, noch bevor die Deadlines vorbei sind.

D.h. die „**missed gallons**“ (also das Wasser, das in Tanks zurückbleibt, wenn die jeweilige Deadline überschritten ist), sollen möglichst gering gehalten werden. Diese werden nochmals am Ende der Simulationszeit berechnet und angezeigt.

Nach jeder Runde muss das System vom Versuchsleiter neu gestartet werden. Gib dem Versuchsleiter bitte ein Zeichen, wenn du eine Runde gespielt hast.

Viel Spaß!

Bei Fragen kannst du dich gerne an den Versuchsleiter wenden.

Anhang C: Datentabellen und Abbildungen der qualitativen Auswertung

Tabelle 17: Daten der ersten Sortierung

Kategorie	Slow- Bedin- gung	Fast- Bedin- gung	Nlow	Nhigh	Elow	Ehigh	Olow	Ohigh	Missed gallons Tag 1	Missed gallons Tag 2	Missed gallons gesamt
Kurve ohne Peak	30%	70%	30%	70%	30%	70%	80%	20%	164.68	116.95	140.66
Eher gewölbte Kurve	60%	40%	60%	40%	60%	40%	60%	40%	161.35	132.47	142.87
nahezu "2- 3-1- Muster" niedriger Peak	33%	67%	67%	33%	33%	67%	83%	17%	123.17	88.61	104.94
nahezu "2- 3-1- Muster" hohem Peak	80%	20%	40%	60%	60%	40%	20%	80%	268.60	196.50	216.32
klassisches "2-3-1- Muster"	67%	33%	67%	33%	100%	0%	67%	33%	179.67	142.89	157.14

Die zweite Sortierung bezieht sich auf die Veränderungen, die die Teilnehmer/innen im Laufe der Bearbeitung vornahmen und anhand der Verlaufskurve der kumulativen missed gallons erschlossen werden können.

Tabelle 18: *Daten der zweiten Sortierung*

Kategorie	Slow- Bedin- gung	Fast- Bedin- gung	Nlow	Nhigh	Elow	Ehigh	Olow	Ohigh	Missed gallons Tag 1	Missed gallons Tag 2	Missed gallons gesamt
Viel verändert	50%	50%	30%	70%	40%	60%	30%	70%	174.70	132.50	151.48
Leicht verändert	14%	86%	57%	43%	57%	43%	43%	57%	164.43	107.12	133.53
Kaum verändert	65%	35%	59%	41%	53%	47%	65%	35%	175.49	139.22	154.14

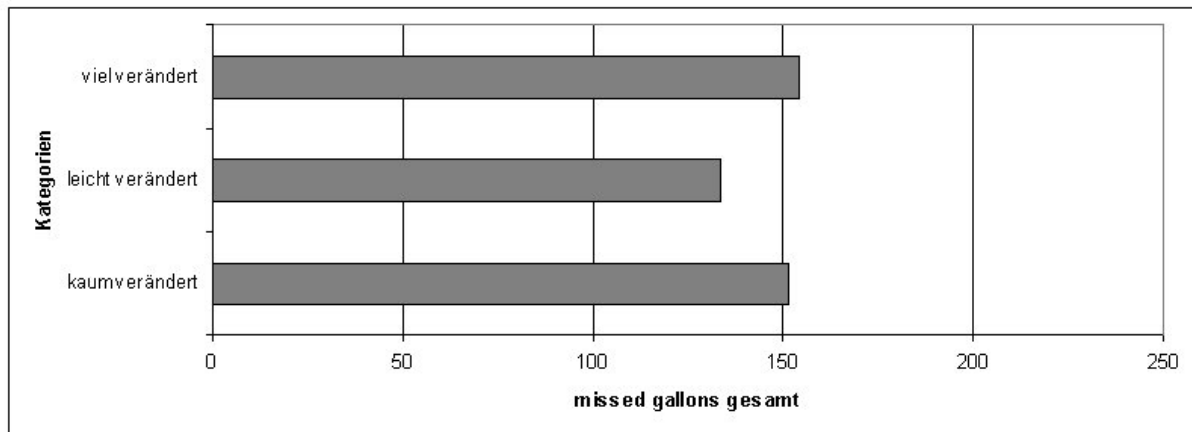


Abbildung 14: *Bewertung der Kategorien der zweiten Sortierung anhand der gesamten missed gallons.*

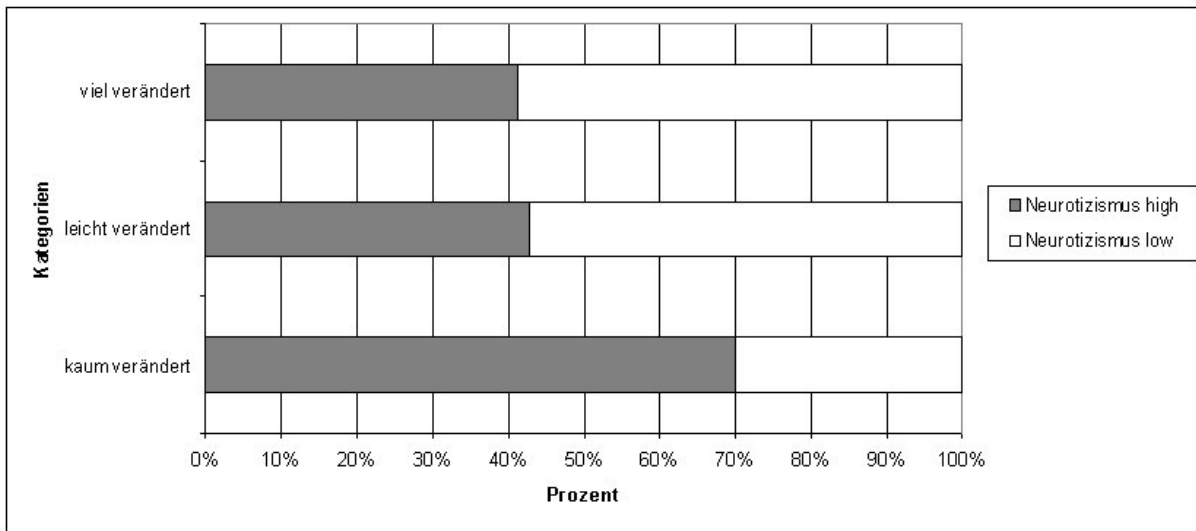


Abbildung 15: Verhältnis von Teilnehmer/innen mit hohen vs. niedrigen Werten beim Persönlichkeitsmerkmal Neurotizismus in der zweiten Sortierung nach Kategorien.

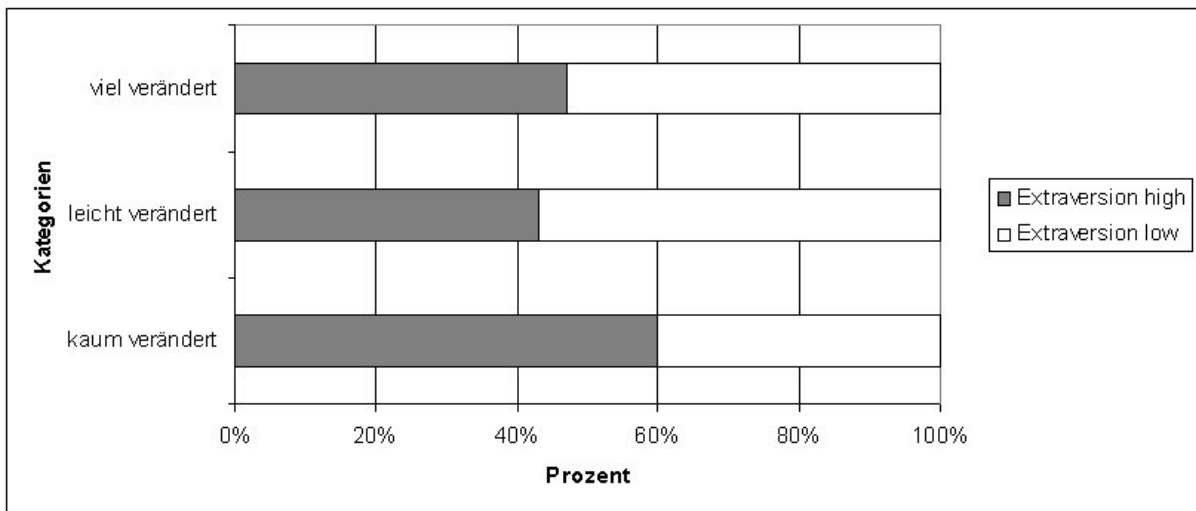


Abbildung 16: Verhältnis von Teilnehmer/innen mit hohen vs. niedrigen Werten beim Persönlichkeitsmerkmal Extraversion in der zweiten Sortierung nach Kategorien.

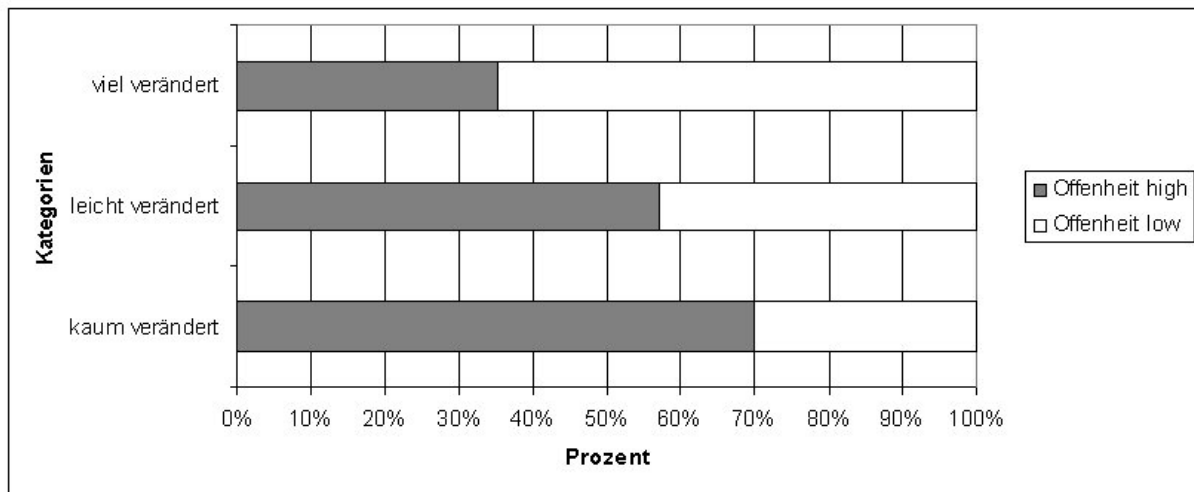


Abbildung 17: Verhältnis von Teilnehmer/innen mit hohen vs. niedrigen Werten beim Persönlichkeitsmerkmal Offenheit in der zweiten Sortierung nach Kategorien.

Die dritte Sortierung ist eine Überarbeitung und Zusammenfassung der ersten. Die fünf Kategorien wurden auf drei reduziert:

Tabelle 19: Daten der dritten Sortierung

Kategorie	Slow- Bedin- gung	Fast- Bedin- gung	Nlow	Nhigh	Elow	Ehigh	Olow	Ohigh	Missed gallons Tag 1	Missed gallons Tag 2	Missed gallons gesamt
Gewölbte Kurve	50%	50%	50%	50%	43%	57%	50%	50%	161.29	132.38	144.43
Niedriger Peak	33%	67%	50%	50%	75%	25%	42%	58%	145.11	97.71	120.24
Hoher Peak	75%	25%	50%	50%	25%	75%	62%	38%	235.25	176.40	194.13

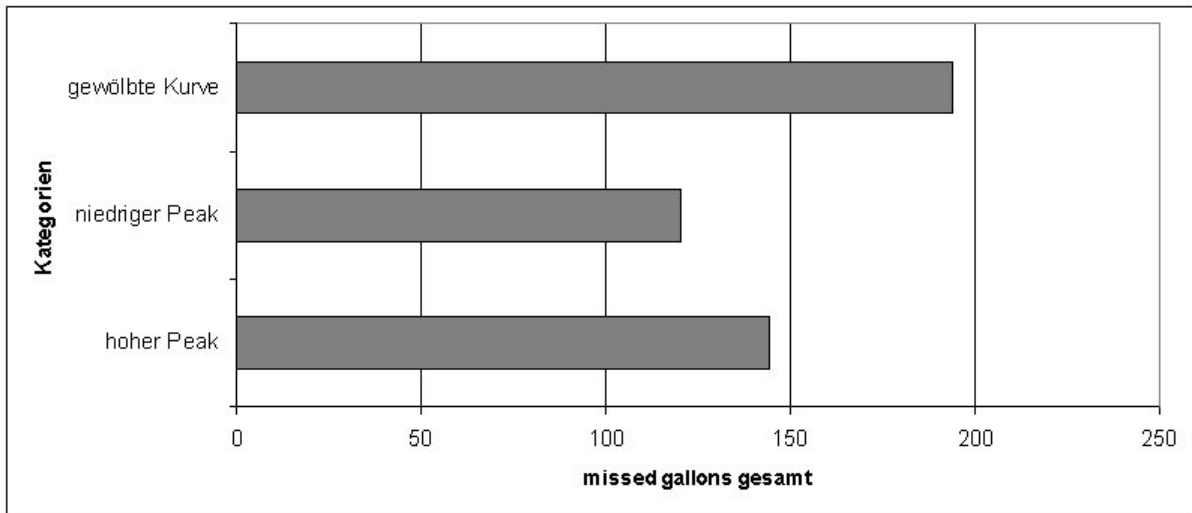


Abbildung 18: Bewertung der Kategorien der dritten Sortierung anhand der gesamten missed gallons.

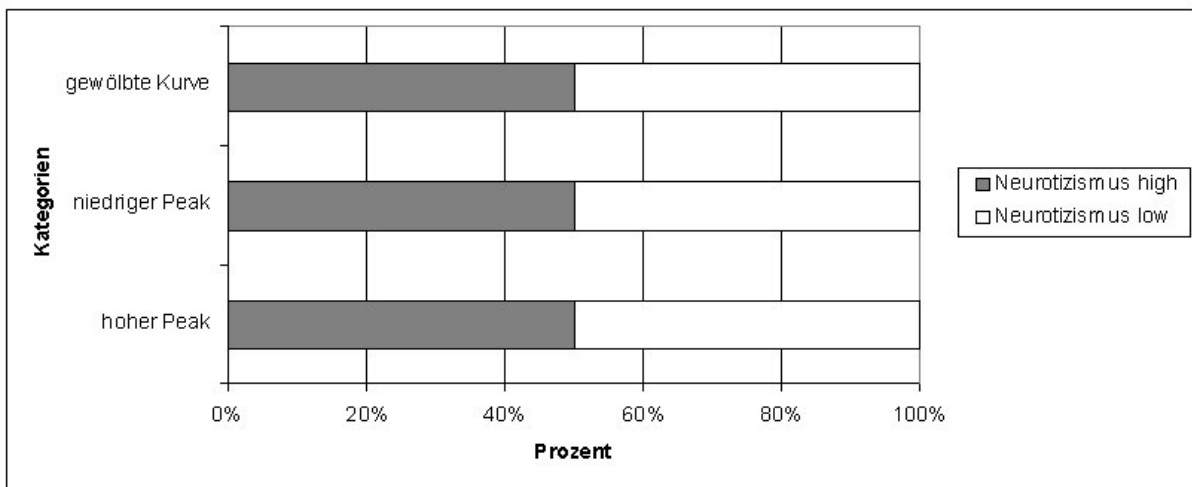


Abbildung 19: Verhältnis von Teilnehmer/innen mit hohen vs. niedrigen Werten beim Persönlichkeitsmerkmal Neurotizismus in der dritten Sortierung nach Kategorien.

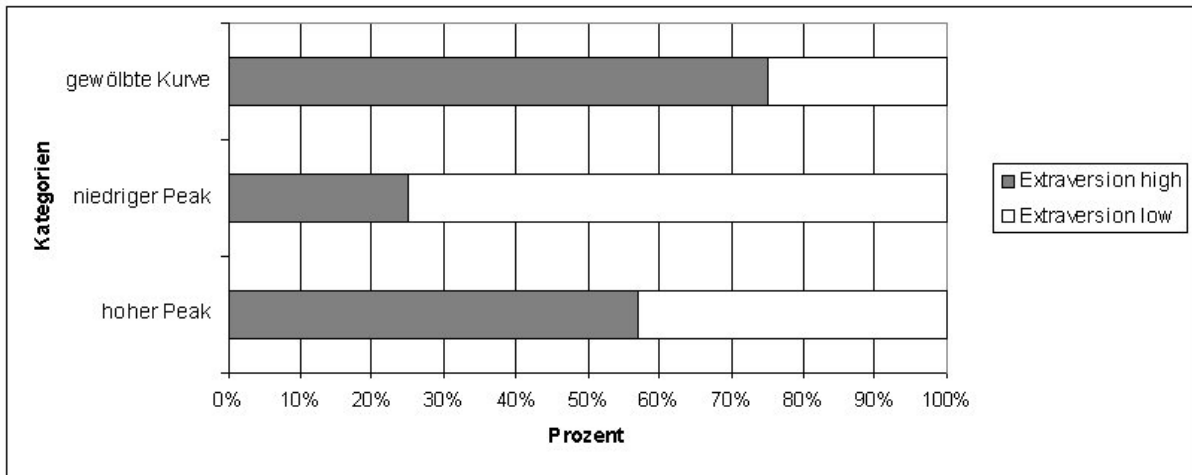


Abbildung 20: Verhältnis von Teilnehmer/innen mit hohen vs. niedrigen Werten beim Persönlichkeitsmerkmal Extraversion in der dritten Sortierung nach Kategorien.

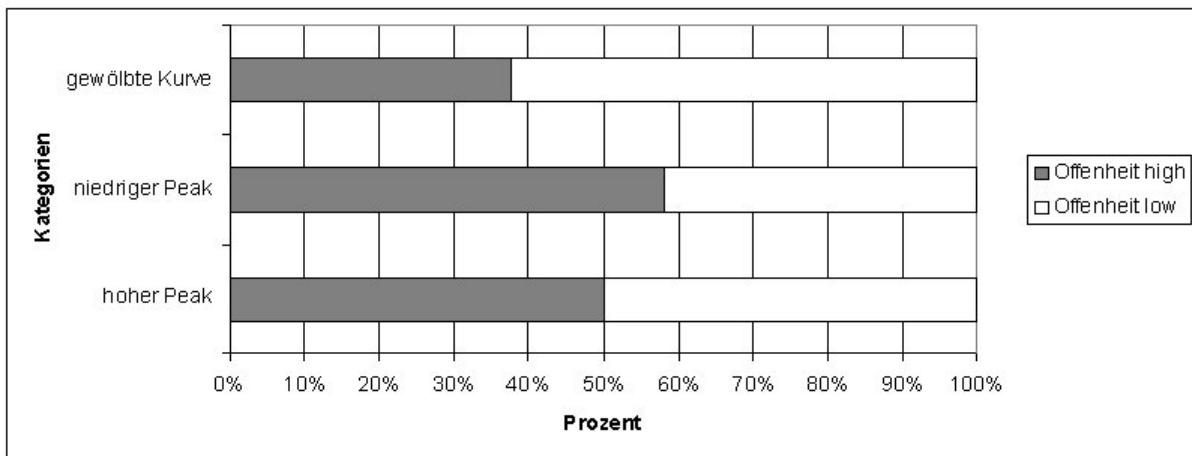


Abbildung 21: Verhältnis von Teilnehmer/innen mit hohen vs. niedrigen Werten beim Persönlichkeitsmerkmal Extraversion in der dritten Sortierung nach Kategorien.

Erklärung

Hiermit erkläre ich, dass ich die Diplomarbeit selbstständig verfasst habe und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Heidelberg, 18. August 2004

Unterschrift

Sebastian Töpfer

Hans-Junginger-Str. 3

69214 Eppelheim

Freiwillige Erklärung

Ich stimme ausdrücklich zu, dass meine durch Prof. Dr. J. Funke und Prof. Dr. H. J. Ahrens (Dr. C. Schmidt-Rathjens) betreute Diplomarbeit mit dem Titel „*Instanzenbasiertes Lernen: Der Einfluss von Zeitdruck und Persönlichkeitsmerkmalen beim Komplexen Problemlösen mit dem Water-Purification-Plant-Szenario*“ nach Beendigung der Diplomprüfung wissenschaftlichen Zwecken zugänglich gemacht und in der Institutsbibliothek ausgestellt wird (Veröffentlichungen nach § 6 Abs. 1 UrhG), sowie hieraus im Rahmen des § 51 UrhG zitiert werden kann. Sämtliche Verwertungsrechte nach § 51 UrhG verbleiben bei dem Verfasser der Diplomarbeit.

Heidelberg, 18. August 2004

Unterschrift

Sebastian Töpfer

Hans-Junginger-Str. 3

69214 Eppelheim