

bezeichnet das Konzept der Strategiewahl als unrealistisch und fordert ein neues Verfahren.

Auf der Basis von Satisficing lässt sich ein binäres Bewertungssystem bilden, das Alternativen innerhalb aktivierter Image-Elemente daraufhin prüft, ob sie Anforderungen gerecht werden. Ein solcher *Brauchbarkeits-Test* registriert erfüllte Anforderungen und ermittelt aus den Erfüllungen geteilt durch die Anzahl Einzel-Tests eine Brauchbarkeitseinschätzung. Auf dieser Basis funktioniert der Stopp-Mechanismus *Satisficing*: Eine Entscheidung wird getroffen, wenn eine Lösung innerhalb aller Image-Elemente allen Ansprüchen gerecht wird und die Brauchbarkeitseinschätzung den Wert von 1 erreicht. Weil das Anspruchsniveau dabei immer den Wert 1 hat, wird von einem Festwertregler gesprochen. Eine Anpassung an die Entscheidungs-Wichtigkeit erfolgt über die Anzahl der aktivierten Image-Elemente und der Höhe der darin gestellten Anforderungen. Bei einer ebenfalls denkbaren Folgeregelung würde die Wichtigkeitsadaptation noch zusätzlich über eine Variation des Anspruchsniveaus erfolgen. Entscheide würden dann auch fallen können, wenn die Brauchbarkeitseinschätzung einer Alternative den Wert von kleiner als 1 erreicht.

Wie sind mögliche Zusammenhänge zwischen den Variablen beider Stopp-Mechanismen zu interpretieren? Urteilssicherheit verlangt ein Verständnis von Zukunft, da der Grad des Vertrauens abgeschätzt wird, mit dem ein Ereignis eintreffen wird. Satisficing bewertet Objekte im Hinblick auf Kriterien, eine Fähigkeit, die schon bei einfachen Organismen vorhanden ist. Satisficing ist somit stammesgeschichtlich älter als Urteilssicherheit, und mögliche Korrelationen zwischen den Größen beider Mechanismen sind deshalb am ehesten als ein Effekt des Anspruchsniveaus auf die gewünschte Urteilssicherheit, bzw. der Brauchbarkeitseinschätzung auf die aktuelle Urteilssicherheit, zu interpretieren.

Mit dem Stopp-Mechanismus Satisficing steht ein konkurrierendes Modell zur Verfügung, das dem Stopp-Mechanismus Urteilssicherheit entgegen gestellt werden kann, um damit den Erklärungserfolg beider Mechanismen miteinander vergleichen zu können.

Teil II: Empirische Analysen

3 Ein Experiment mit unerfahrenen Entscheidern

3.1 Zielsetzung und Hypothesen

Das Hauptziel dieser Untersuchung bestand darin, herauszufinden, ob eher der Stopp-Mechanismus Urteilssicherheit oder eher der Stopp-Mechanismus Satisficing für das Ende von Entscheidungsprozessen verantwortlich ist. Zusätzlich interessierte aber auch, ob der Brauchbarkeits-Test die Bildung der Brauchbarkeitseinschätzung beschreibt bzw. das Contingency Model von Beach und Mitchell (1978), die Support Theory von Tversky und Koehler (1994) und das Model of Belief Processing von Curley und Benson (1994) die Entstehung der Urteilssicherheit erklären.

Im ersten Abschnitt 3.1.1 werden drei Funktionstests für Stopp-Mechanismen mit den dazugehörigen Hypothesen vorgestellt. Im Abschnitt 3.1.1 werden Hypothesen formuliert, mit denen die Theorien getestet werden, welche die Entstehung der Brauchbarkeitseinschätzung bzw. der Urteilssicherheit beschreiben.

3.1.1 Funktionstests für Stopp-Mechanismen

Die Funktionsweise eines Regelkreises lässt sich wie folgt beschreiben: Der Regelkreis misst die Abweichung eines Ist-Werts (Regelgrösse) von einem Soll-Wert (Führunggrösse) und steuert das Verhalten angesichts der Mess-Ergebnisse (Regelabweichung) so, dass diese Abweichung möglichst klein bleibt (Bischof, 1995). Damit ein Stopp-Mechanismus als Regelkreis funktionieren kann, muss er drei Anforderungen erfüllen:

- (1) Ein Stopp-Mechanismus muss über eine feste oder eine adaptive Führunggrösse verfügen. Eine adaptive Führunggrösse muss die Wichtigkeit eines Problems für den Entscheider zum Ausdruck bringen können. Wenn die Wichtigkeit einer Entscheidung manipuliert wird, muss sich die Führunggrösse gleichsinnig ändern.
- (2) Ein Stopp-Mechanismus muss eine Regelgrösse bestimmen können, die den Stand der Problemlösung widerspiegelt. Die Verarbeitung nutzbarer Information, welche die Lösungsfindung erleichtert, muss diese Grösse erhöhen.

- (3) Ein Stopp-Mechanismus muss die Suche nach Information beenden können, wenn der Ist-Wert mit dem Soll-Wert übereinstimmt.

Aus den drei Anforderungen an Stopp-Mechanismen lassen sich drei Funktionstests mit den entsprechenden Hypothesen ableiten:

Funktionstest 1: Reaktion der Führungsgrösse

Dieser Test untersucht, ob bei wichtigeren Entscheidungen ein höherer Soll-Wert angegeben wird als bei weniger wichtigen Entscheidungen. Trifft dies zu, so ist dies ein Indiz für die Existenz eines $Assessor_{Soll}$ und eine erste Voraussetzung für eine Folgeregelung ist gegeben. Wird keine Wirkung der Aufgaben-Wichtigkeit festgestellt, so ist dies ein Hinweis auf eine Festwertregelung. Daraus ergaben sich folgende Hypothesen:

- (1a) Die Aufgaben-Wichtigkeit hat einen Einfluss auf die Höhe der Führungsgrösse (Folgeregelung).
- (1b) Die Aufgaben-Wichtigkeit hat keinen Einfluss auf die Höhe der Führungsgrösse (Festwertregelung).

Funktionstest 2: Reaktion der Regelgrösse

Dieser Test prüft, ob nach der Verarbeitung von entscheidungsrelevanter Information der Ist-Wert steigt. Ist dies der Fall, so ist dies ein Zeichen für das Vorhandensein eines $Assessor_{Ist}$. Daraus liess sich folgende Hypothese formulieren:

- (2) Mit jedem Verarbeitungsschritt steigt die Regelgrösse.

Funktionstest 3: Wirkung der Regelabweichung

Wenn Entscheidungen in einzelne Prozessschritte unterteilt werden, erforscht dieser Test bei einer Regelabweichung die Fähigkeit festzustellen, ob es sich bei einem Prozessschritt um ein definitives oder um ein provisorisches Urteil handelt. Wenn eine Regelabweichung diese Prognosefähigkeit besitzt, darf aber noch nicht auf die Wirkung eines Stopp-Mechanismus geschlossen werden. Wird das Verarbeitungsende bei einer endlichen Zahl von möglichen Prozessschritten durch den Zufall bestimmt, dann wird es um so wahrscheinlicher, dass ein Schritt ein endgültiger ist, je mehr Schritte ihm vorangegangen sind. Besteht ein Zusammenhang zwischen Informationsnutzung

und einem Ist-Wert (Funktionstest 2), dann werden neben den absolvierten Verarbeitungsschritten auch automatisch die Regelgrösse und die dazugehörige Regelabweichungen zu signifikanten Prädiktoren für das Verarbeitungsende. Damit von einer Wirkung der Regelabweichung auf die Informationsnutzung gesprochen werden kann, muss deshalb gefordert werden, dass die Regelabweichung eine bessere Vorhersage macht als die absolvierten Verarbeitungsschritte. Daraus ergab sich folgende Hypothese:

- (3) Eine Regelabweichung erkennt „letzte“ Schritte und ist in der Vorhersage besser als die Summe der absolvierten Verarbeitungsschritte.

3.1.2 Erwartete Zusammenhänge zwischen den Variablen

3.1.2.1 Brauchbarkeitseinschätzung

Der in Abschnitt 2.4 vorgeschlagene *Brauchbarkeits-Test* prüft Alternativen darauf, ob sie innerhalb von aktivierten Image-Elementen den Anforderungen gerecht werden oder nicht. Dabei werden um so mehr Elemente geprüft und höhere Anforderungen gestellt, je wichtiger eine Entscheidung ist. Unter der Voraussetzung, dass Versuchspersonen brauchbare Alternativen vorgelegt werden, lassen sich daraus zwei Erwartungen ableiten: Zum Einen sollte die Brauchbarkeitseinschätzung im Verlaufe eines Entscheidungsprozesses kontinuierlich ansteigen und zum Anderen sollte dieser Anstieg um so langsamer geschehen, je wichtiger eine Entscheidung bewertet wird. Wie die Beispiele in Tabelle 3-1 zeigen, wird ein entgültiges Urteil gefällt, wenn die Brauchbarkeit den Wert von 1 erreicht. Bei den wichtigeren Entscheidungen wird dieses Ziel in drei Schritten erreicht und in den weniger wichtigen Entscheidungen in einem oder in zwei Schritten.

Tabelle 3-1. Beispiele von möglicher Brauchbarkeitseinschätzung in Abhängigkeit von der Aufgaben-Wichtigkeit und des Entscheidungsforgangs (Verarbeitungsschritte)

	Entscheider	Schritt 1	Schritt 2	Schritt 3
Wichtigere Entscheidung	A	0.33	0.66	1
	B	0.33	0.66	1
Unwichtigere Entscheidung	C	0.5	1	
	D	1		
	F	1		

Daraus lässt sich die Erwartung ableiten, dass Entscheider eine favorisierte Alternative als umso brauchbarer einschätzen, je mehr Verarbeitungsschritte sie absolviert haben und dass sie in ihrer Einschätzung um so vorsichtiger sind, je wichtiger die Entscheidung für sie ist. Wird für das vorliegende Beispiel der Einfluss der Aufgaben-Wichtigkeit und der Verarbeitungsschritte auf die Brauchbarkeitseinschätzung mittels einer Regressionsanalyse ermittelt, so ergibt sich ein sehr enger positiver Zusammenhang zwischen Verarbeitungsschritten und Brauchbarkeitseinschätzung ($\beta = .88$) sowie ein enger negativer Zusammenhang zwischen Aufgaben-Wichtigkeit und Brauchbarkeitseinschätzung ($\beta = -.79$). Daraus ergab sich folgende Hypothese:

- (B) Es besteht ein positiver Zusammenhang zwischen den Verarbeitungsschritten und der Brauchbarkeitseinschätzung sowie ein negativer Zusammenhang zwischen Aufgaben-Wichtigkeit und Brauchbarkeitseinschätzung.

3.1.2.2 Aktuelle Urteilssicherheit

Im *Contingency Model for the Selection of Decision Strategies* von Beach und Mitchell (1978) wird die Urteilssicherheit vom Ausmass an analytischem Aufwand bestimmt, der betrieben wurde, um zu einer Entscheidung zu gelangen (Abschnitt 1.3.1). Wenn Entscheider wählen können, wie viel Verarbeitungsschritte sie absolvieren wollen, dann lässt sich der analytische Aufwand anhand der absolvierten Verarbeitungsschritte erheben. Wird die Urteilssicherheit von der analytischen Qualität der verwendeten Entscheidungsstrategien bzw. -heuristiken bestimmt, dann sollte auch unter Berücksichtigung anderer möglicher Einflussfaktoren ein positiver Zusammenhang zwischen den Verarbeitungsschritten und der aktuellen Urteilssicherheit festzustellen sein. Daraus liess sich folgende Hypothese formulieren:

- (P) Es besteht ein positiver Zusammenhang zwischen den Verarbeitungsschritten und der aktuellen Urteilssicherheit.

Laut der *Support Theory* von Tversky und Koehler (1994) errechnet sich die aktuelle Urteilssicherheit aus den akkumulierten Argumenten für eine bevorzugte Alternative, geteilt durch die Summe aller Argumente für alle Alternativen (Abschnitt 1.3.2.1). Zwei Erwartungen liessen sich daraus formulieren:

(1) Wenn zwei Entscheidungen zu fällen sind und sich die Alternativen in einer Entscheidungsaufgabe stärker voneinander unterscheiden als in der anderen, dann sollten sich auch die Urteilssicherheitsangaben der beiden Entscheidungen unterscheiden. Daraus ergab sich die Hypothese:

- (R₁) Der Differenzierungsgrad von Alternativen hat einen Einfluss auf die Urteilssicherheit.

(2) Wenn Entscheider zwei Alternativen zur Auswahl haben und jeweils die Brauchbarkeit der besseren Alternative (B) sowie die Differenz in der Brauchbarkeitseinschätzung zwischen der besseren und der schlechteren Alternative angeben (D), dann lässt sich die Urteilssicherheit (US) aus der Formel

$$US(\text{bessere Alternative}) = \frac{B(\text{bessere Alternative})}{2 \times B(\text{bessere Alternative}) - D}$$

berechnen. Tabelle 3-2 zeigt neun einfache Bewertungsbeispiele. Wird eine Regressionsanalyse mit diesen Angaben gemacht, so ergibt sich ein sehr enger positiver Zusammenhang zwischen der Differenzwahrnehmung und der Urteilssicherheit ($\beta = .84$) und ein etwas schwächerer negativer Zusammenhang zwischen der Brauchbarkeitseinschätzung und der Urteilssicherheit ($\beta = -.47$).

Tabelle 3-2. Beispiele von möglichen Angaben der aktuellen Urteilssicherheit in Abhängigkeit von der Braubarkeits- und der Differenzeinschätzung

Bewertung	Brauchbarkeit	Differenz	Urteilssicherheit
1	10	2	56%
2	10	4	63%
3	10	6	71%
4	8	2	57%
5	8	4	67%
6	8	6	80%
7	6	2	60%
8	6	4	75%
9	6	6	100%

Folgende Hypothese wurde aufgestellt:

- (R₂) Es besteht ein sehr enger positiver Zusammenhang zwischen der Differenzwahrnehmung und der Urteilssicherheit und ein schwächerer negativer Zusammenhang zwischen der Brauchbarkeitseinschätzung und der Urteilssicherheit.

Im *Model of Belief Processing* von Curley und Benson (1994) basiert die Urteilssicherheit im Gegensatz zur Support Theory nicht auf einer relativen sondern auf einer *absoluten* argumentativen Unterstützung (Abschnitt 1.3.2.2). Bezüglich des Differenzierungsgrades von Alternativen ergab dies die Erwartung, dass Entscheidungen mit unterschiedlich stark differenzierenden Alternativen sich in der Urteilssicherheiten nicht unterscheiden sollten. Bezüglich den Zusammenhängen zwischen den Variablen ergab sich die Forderung nach einem engen positiven Zusammenhang zwischen der Brauchbarkeitseinschätzung – die die kumulative Stärke der Argumente für die bevorzugte Alternative wiedergibt – und der Urteilssicherheit. Folgende Hypothesen wurden formuliert:

- (A₁) Der Differenzierungsgrad von Alternativen hat keinen Einfluss auf die Urteilssicherheit.
- (A₂) Es besteht ein enger positiver Zusammenhang zwischen der Brauchbarkeitseinschätzung und der Urteilssicherheit.

3.2 Methode

Einen Grund dafür, dass das Wissen über Stopp-Mechanismen noch recht bescheiden ist, sehen Gigerenzer und Todd (1999) sowie Huber (1997) darin, dass typische Entscheidungs-Experimente alle Informationen bereitstellen, obschon gerade die Suche nach Informationen einen wesentlichen Bestandteil realer Entscheidungen darstellt. Im vorliegenden Experiment konnten die Versuchspersonen deshalb selber bestimmen, wie viele Informationen sie nutzen wollten. 35 Studentinnen und Studenten hatten jeweils eine Informatik- (IT-) und eine Personalentscheidung (P-Entscheidung) zu fällen (Abbildung 3-1). Sowohl bei der IT- als auch bei der P-Entscheidung wurde nach einer kostenlosen Vorinformation drei kostenpflichtige Informations-Berichte angeboten. Nach der Bearbeitung jedes Berichts konnte entweder weitere Infor-

mation gekauft oder die definitive Entscheidung getroffen werden. In zwei Versuchsbedingungen war entweder die IT-Entscheidung sehr wichtig und die P-Entscheidung weniger wichtig (erstes Treatment „IT/p“) oder die IT-Entscheidung weniger wichtig und die P-Entscheidung sehr wichtig (zweites Treatment „it/P“).

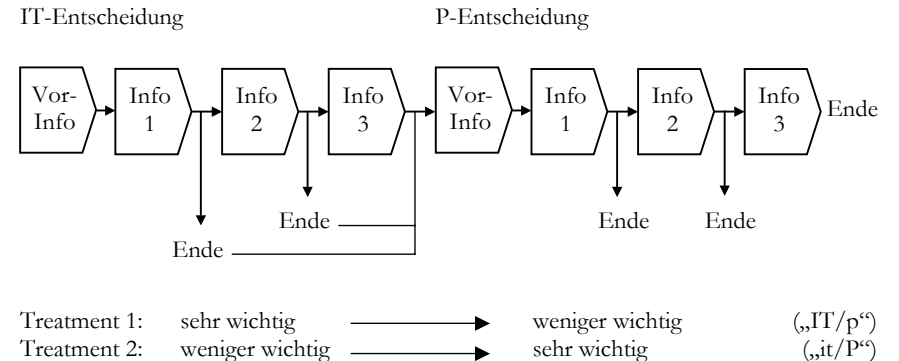


Abbildung 3-1. Ablauf des Experiments mit den dazugehörigen Treatments

3.2.1 Versuchspersonen

An dem Experiment nahmen 35 Studierende (13 Frauen und 22 Männer) von zwei Betriebsökonomie-Klassen des ersten Jahrgangs der Fachhochschule Aargau teil. Das Durchschnittsalter betrug 23.0 Jahre bei einer Standardabweichung von $SD = 1.4$, wobei die jüngste Versuchsperson 21, die älteste 26 Jahre alt war. Alle Studierenden verfügten über eine abgeschlossene Berufsausbildung oder eine Mittelschulbildung.

3.2.2 Materialien

Die Informatik- und Personalaufgabe simulierten gemeinsam eine Management-Übung, die in ähnlicher Form in Assessment-Centers verwendet wird. Mit Assessment-Centers ermitteln Firmen üblicherweise das Karrierepotential von Nachwuchsleuten und Führungskräften. Das Verfahren ist in Europa und den USA weit verbreitet (Schmidt & Hunter, 1998). In beiden Entscheidungsaufgaben wurden die relevanten Argumente in ähnlich strukturierten Berichten geliefert. Im Differenzierungsgrad der Alternativen waren die beiden Aufgaben aber verschieden: In der IT-Entscheidung war nach jedem Bericht mindestens eine Firma besser als eine andere und in der P-Entscheidung waren alle Kandidaten bezüglich der vorgegebenen Kriterien jeweils gleich gut qualifiziert.

3.2.2.1 Informatikentscheidung

Die Versuchspersonen mussten aus drei Software-Firmen (AA, BB und CC) die geeignetste auswählen. Die Urteile sollten auf den folgenden fünf Entscheidungskriterien basieren: Marktstellung, Leistungsnachweis, Dienstleistungsbreite, Systemerfahrung und Systemumfang (Tabelle 3-3). Die Kriterien entsprachen weitgehend den Kompetenzdimensionen, die bei der Bewertung von Angeboten im Systemgeschäft verwendet werden (Backhaus, 1990).¹

Drei Consultant-Berichte lieferten Information über die zu bewertenden Software-Firmen. Jeder stammte von einem anderen externen Berater. Die drei Berater beschrieben, was sie über die drei Firmen hinsichtlich der Bewertungskriterien wussten. So stand beispielsweise im dritten Consultant-Bericht über eine Software-Firma: „Ich weiss, dass erfolgreiche Projekte bereits bei der Hofer AG sowie der Kuster AG realisiert wurden.“ „Erfahrung mit Computern von Sun und Compaq sind mir bekannt.“ „Bei der Realisierung arbeiten sie in den mir bekannten Projekten mit Peter Software AG und Riter Software AG als Unterlieferanten zusammen.“

¹ Die Bezeichnungen „Dienstleistungsbreite“ und „Systemerfahrung“ ersetzen die vorgegebenen Begriffe „Dienstleistungspaket“ und „Kompatibilität“.

Tabelle 3-3. Entscheidungskriterien zur Bewertung von Softwarefirmen

1. Marktstellung = Anzahl Beschäftigte
Je mehr Beschäftigte ein Realisierungspartner hat, desto geringer ist das Risiko, dass die Firma durch Mitarbeiterabgänge ihre Leistungsfähigkeit verliert.
2. Leistungsnachweis = Anzahl Referenzen
Je mehr Projekte eine Softwarefirma bereits erfolgreich realisiert hat, desto wahrscheinlicher ist es, dass sie ein neues Projekt zu einem erfolgreichen Abschluss bringt.
3. Dienstleistungsbreite = Anzahl angebotene Lösungen
Je breiter die Angebotspalette einer Softwarefirma ist, desto grösser ist die Wahrscheinlichkeit, dass sie als Partnerin auch in zukünftigen Projekten mitwirken kann.
4. Systemerfahrung = Erfahrung mit Computerplattformen
Das Ausmass an Erfahrung mit System-Plattformen (z. B. IBM, HP usw.) erhöht die Wahrscheinlichkeit für eine erfolgreiche Implementation.
5. Systemumfang = Anzahl Unterlieferanten im Projekt (<i>je weniger, desto besser</i>)
Je weniger Unterlieferanten eine Firma überwachen muss, desto besser wird die Qualität der Realisierung.

Jeder neue Bericht lieferte neue Information über die zur Auswahl stehenden Alternativen. Damit stieg mit jedem Verarbeitungsschritt das verfügbare Wissen der Versuchspersonen. Tabelle 3-4 zeigt, welche Argumente in den drei Verarbeitungsschritten abgegeben wurden und welche Firma sich jeweils als Favoritin herauskristallisierte. Der erste Bericht beinhaltete in der Summe aller Argumente einen Vorteil der Firma BB beim „Leistungsnachweis“ und einen Vorsprung der Firma CC bei der „Systemerfahrung“ und stellte die Firma AA als unattraktivste Wahl dar. Die Berichte 2 und 3 machten die Firma CC zur eindeutig besten Wahl, da neue positive Argumente zugunsten von CC in den Kriterien „Leistungsnachweis“ und „Systemerfahrung“ erschienen. Bericht 2 nannte zudem bei den Firmen AA und BB ein zusätzlich negatives Argument beim „Systemumfang“. Diese beiden Firmen hatten je einen dritten Unterlieferanten zu kontrollieren. Jeder Bericht lieferte einerseits neue Erkenntnisse und verursachte andererseits hypothetische Kosten von je Fr. 15'000.-. Die Versuchspersonen konnten selber bestimmen, ob sie ihre definitive Entscheidung nach der Lektüre von einem, zwei oder drei Consultant-Berichten treffen wollten.

Tabelle 3-4. Verteilung der Argumente auf die drei Firmen in den Consultant-Berichten

Entscheidungskriterien	Bericht 1			Bericht 2			Bericht 3		
	Firma A	BB	CC	A	BB	CC	A	BB	CC
1. Marktstellung	x	x	x	x	x	x	x	x	x
2. Leistungsnachweis									
▪ Referenzfirma 1	x	x	x	x	x	x		x	
▪ Referenzfirma 2	x	x	x		x		x	x	x
▪ Referenzfirma 3		x		x	x	x	x	x	x
▪ Referenzfirma 4						x			
▪ Referenzfirma 5									x
▪ Referenzfirma 6						x			x
3. Dienstleistungsbreite									
▪ Angebot 1	x	x		x	x		x	x	x
▪ Angebot 2	x			x	x	x	x	x	x
▪ Angebot 3			x	x		x	x	x	x
▪ Angebot 4		x	x		x	x	x	x	x
4. Systemerfahrung									
▪ IBM	x	x	x	x	x				x
▪ Bull			x		x			x	
▪ HP		x	x			x		x	
▪ Unisys						x			x
▪ Sun	x					x		x	
▪ Compaq				x			x		x
5. Systemumfang									
▪ Unterdienstleister 1	x	x	x	x	x	x			x
▪ Unterdienstleister 2	x	x	x			x	x	x	x
▪ Unterdienstleister 3				x	x		x	x	

3.2.2.2 Personalentscheidung

Die Versuchspersonen hatten im Weiteren aus drei Verkäufern den geeignetsten auswählen. Das Ideal-Profil eines Bewerbers wurde wie folgt vorgegeben: Verkaufsgeschick/Überzeugungskraft, Leistungsbereitschaft, Umsetzen von Fachinformation, Planung und Organisation, Teamfähigkeit, Flexibilität, Kontaktfähigkeit und Frustrationstoleranz (Tabelle 3-5). Die acht Eigenschaften entsprechen einem Anforderungsprofil für Vertriebsbeauftragte (Jetter, 1996). Es sollte jene Person gefunden werden, die von diesen acht Verkäufereigenschaften im Vergleich zu den anderen Personen am meisten besitzt. Was die Information zu den drei Kandidaten betraf, so

wurde diese mit Hilfe von drei Interview-Berichte geliefert, die jeweils von einer anderen Interviewerin erstellt worden waren. Diese Zusammenfassungen enthielten Informationen, wie sie nach der Durchführung von unstrukturierten Interviews erwartet werden konnten. So war beispielsweise zu lesen: „Kandidat A versteht die Probleme seiner Kunden und geht darauf ein“. „Seine Besuche bei Kunden bereitet er gut vor“. „Er ist von seiner Sache überzeugt“.

Tabelle 3-5. Ideal-Profil eines Verkäufers

1. Verkaufsgeschick/Überzeugungskraft	<ul style="list-style-type: none"> ist von seiner Sache überzeugt kann andere für eigene Ideen begeistern
2. Leistungsbereitschaft	<ul style="list-style-type: none"> setzt sich für die Erreichung seiner Ziele voll ein setzt sich und anderen hohe Ziele
3. Umsetzen von Fachinformation	<ul style="list-style-type: none"> kann Fachinformation verständlich präsentieren berücksichtigt Vorwissen der Zuhörer und geht darauf ein
4. Planung und Organisation	<ul style="list-style-type: none"> arbeitet systematisch und zielorientiert setzt sich klare Ziele und verfolgt sie konsequent
5. Teamfähigkeit	<ul style="list-style-type: none"> wird im Kollegenkreis voll akzeptiert stellt sich in den Dienst der Gruppe
6. Flexibilität	<ul style="list-style-type: none"> kommt mit unvorhergesehenen Situationen gut zurecht ist in seiner Arbeitsweise beweglich
7. Kontaktfähigkeit	<ul style="list-style-type: none"> kann leicht den Kontakt zu anderen herstellen geht von sich aus auf andere zu und wartet nicht, bis andere auf ihn zugehen
8. Frustrationstoleranz	<ul style="list-style-type: none"> lässt sich durch Misserfolge nicht entmutigen geht mit Enttäuschungen konstruktiv um

Tabelle 3-6 gibt die Eigenschaften wieder, welche den drei Kandidaten nach den einzelnen Interviews zugesprochen wurden. Nach der ersten Interview-Serie erfüllten alle drei Bewerber je drei Kriterien: Bei den Kandidaten A und C waren es „Verkaufsgeschick“, „Leistungsbereitschaft“ und „Kontaktfähigkeit“, bei Kandidat B waren es „Verkaufsgeschick“, „Umsetzen von Fachinformation“ und „Kontaktfähigkeit“. Keiner der Kandidaten hatte da-

mit gegenüber den beiden anderen einen objektiven Vorsprung. Die Information aus den Interview-Berichten 2 und 3 zeigte, dass jeder Kandidat je ein neues Kriterium erfüllte. Der zweite Bericht enthielt bei Kandidat A „Umsetzen von Fachinformation,“ bei B und C „Teamfähigkeit“. Der letzte Bericht brachte als neues Kriterium bei Kandidat A „Planung und Organisation,“ bei B „Flexibilität“ und bei C „Umsetzen von Fachinformation“. Da alle Kriterien gleich viel zählten, war es auch nach dem zweiten und dritten Verarbeitungsschritt *nicht* möglich, einen der Kandidaten als objektiv Besten zu erkennen.

Tabelle 3-6. Zugesprochne Eigenschaften der Kandidaten gemäss Interview-Berichten

Entscheidungskriterien	Interviewserie Kandidaten	1			2			3		
		A	B	C	A	B	C	A	B	C
▪ Verkaufsgeschick/Überzeugungskraft		x	x	x	x	x	x	x	x	x
▪ Leistungsbereitschaft		x		x	x		x			x
▪ Umsetzen von Fachinformation			x		x	x			x	x
▪ Planung und Organisation								x		
▪ Teamfähigkeit						x	x			
▪ Flexibilität									x	
▪ Kontaktfähigkeit		x	x	x	x	x	x	x	x	x
▪ Information je Interviewserie		3	3	3	4	4	4	4	4	4
▪ Akkumuliert Information		3	3	3	4	4	4	5	5	5

Wie bei der IT-Aufgabe lieferte jeder Bericht neue Erkenntnisse und verursachte hypothetische Kosten von Fr. 15'000.-. Die Studierenden bestimmten wiederum selber, ob sie ihr endgültiges Urteil nach der Lektüre von einem, zwei oder drei Berichten fällen wollten.

3.2.3 Versuchsplan

Der Studie lag ein einfaktorieller Versuchsplan¹ (unabhängige Variable *Aufgaben-Wichtigkeit*) in einem kombinierten „Between-Subjects-“ und „Within-Subjects-Design“ zugrunde: Die eine Hälfte der Versuchspersonen begann mit der „sehr wichtigen“ Aufgabe und löste dann die „weniger wichtige“ Aufgabe, die andere Hälfte der Studierenden bearbeitete zunächst die „weniger wichtige“ Aufgabe und anschliessend die „sehr wichtige“ Aufgabe (Abbildung 3-1). Als nicht manipulierbare Grössen wurden die Variablen Need for Cognition und Selbstwirksamkeitserwartung erhoben.

3.2.3.1 Unabhängige Variable Aufgaben-Wichtigkeit

Die Wichtigkeit der beiden Aufgaben wurde anhand des Gewinnes dargestellt, welcher der Firma entgeht, falls nicht die beste Option gewählt wird. Bei der sehr wichtigen Entscheidung entgeht der Firma eine halbe Million Franken, wenn man sich anstelle der besten für die zweitbeste Alternative entschied und eine Million Franken, wenn man anstelle der besten die drittbeste Alternative wählte. Bei der weniger wichtigen Entscheidung waren die Beträge Fr. 30'000.- bzw. Fr. 50'000.- (Tabelle 3-7). Die Wichtigkeit der zu lösenden Aufgaben war an drei Stellen symbolisch mit Sternen markiert und an zwei Stellen wörtlich ausformuliert. Bei der sehr wichtigen Aufgabe stand auf dem Deckblatt die Bezeichnung der Entscheidung (Informatik- oder Personalentscheidung) in roten Lettern und war mit fünf Sternen gekennzeichnet. Bei der weniger wichtigen Aufgabe war die Bezeichnung schwarz beschriftet und hatte nur einen Stern. In jeder Aufgabenbeschreibung standen im Titel entweder ein oder fünf Sterne und die Unterschiede zwischen den Alternativen wurden im Text genannt. So stand bei der sehr wichtigen P-Entscheidung:

„Unterschied zwischen dem besten und dem zweitbesten Kandidaten: Fr. 500'000.-, Unterschied zwischen dem besten und dem drittbesten Kandidaten: Fr. 1'000'000.-“.

¹ Obschon sich die zwei Entscheidungsaufgaben in einem unterschiedlichen Differenzierungsgrad der Alternativen unterschieden, konnte nicht von einem zweifaktoriellen Versuchsplan gesprochen werden, da mit einem Positionseffekt und einer Wirkung der Thematik gerechnet werden musste. Die Aufgabenstellung erscheint deshalb nicht als unabhängige Variable.

Bei der weniger wichtigen Entscheidung war notiert:

„Unterschied zwischen dem besten und dem zweitbesten Kandidaten: Fr. 30'000.-, Unterschied zwischen dem besten und dem drittbesten Kandidaten: Fr. 50'000.-“.

Tabelle 3-7. Unterschiede zwischen den Optionen in den zwei Versuchsbedingungen

	IT-Entscheidung	P-Entscheidung
IT/p		
beste / zweitbeste Option	Fr. 500'000.-	Fr. 30'000.-
beste / drittbeste Option	Fr. 1'000'000.-	Fr. 50'000.-
it/P		
beste / zweitbeste Option	Fr. 30'000.-	Fr. 500'000.-
beste / drittbeste Option	Fr. 50'000.-	Fr. 1'000'000.-

3.2.3.2 Nicht manipulierbare Variablen

3.2.3.2.1 Need for Cognition (NFC)

Eagly und Chaiken (1993) sind der Meinung, dass neben der Wichtigkeit und der Relevanz einer Aufgabenstellung auch Need for Cognition – das Engagement und die Freude beim Denken – einen Einfluss auf die gewünschte Urteilssicherheit und damit auf die Informationsnutzung hat. Need for Cognition wurde mittels eines Tests von Bless, Wänke, Bohner, Fellhauer und Schwarz (1994) erfasst. Der Test basiert auf einer von Cacioppo und Petty (1982) entwickelten Skala. Er weist gute Skalenkennwerte auf und erlaubt eine ökonomische Durchführung (Bless et al., 1994). Verwendet wurde eine 7-Punkte-Skala, die von 0 für „völlig unzutreffend“ bis 6 für „trifft ganz genau zu“ reichte.¹

¹ In Vorversuchen konnte jeweils kein Zusammenhang zwischen Need for Cognition und der gewünschten Urteilssicherheit bzw. Informationsnutzung festgestellt werden. Für die Hauptuntersuchung wurden deshalb von den 16 Fragen die Fragen Nummer 4, 10 und 11 umgepolt, weil bei oberflächlicher Lektüre das Risiko bestand, die Verneinung zu übersehen. So wurde z. B. „Die Vorstellung, mich auf mein Denkvermögen zu verlassen, um es zu etwas zu bringen, spricht mich *nicht* an“ umformuliert zur Aussage „Die Vorstellung, mich auf mein Denkvermögen zu verlassen, um es zu etwas zu bringen, spricht mich an.“

3.2.3.2.2 Selbstwirksamkeitserwartung

Bei jeder Entscheidungsaufgabe (IT- bzw. P-Entscheidung) mussten die Versuchspersonen im Fragebogen 0 auf einer 10-Punkte-Skala angeben, ob sie glaubten, die vorliegende Informatikentscheidungsaufgabe bzw. Personalentscheidungsaufgabe lösen zu können (0 = „gar nicht“ bis 9 = „vortrefflich“).

3.2.3.3 Abhängige Variablen

3.2.3.3.1 Frage nach der wichtigeren Entscheidung

Um festzustellen, ob das Treatment funktionierte, mussten die Studierenden im IT-Fragebogen 0 angeben, welche der zu lösenden Aufgabenstellungen mehr als zehnmal wichtiger sei als die andere. Die richtige Antwort lautete je nach Treatment: Informatik- oder Personalentscheidung.

3.2.3.3.2 Relevanz für die Firma

Bei jeder Entscheidungsaufgabe mussten die Versuchspersonen im Fragebogen 0 auf einer 10-Punkte-Skala angeben, wie hoch sie die Wichtigkeit der im Moment bearbeiteten Aufgabe für die Firma einschätzten (0 = „unwichtig“ bis 9 = „sehr wichtig“).

3.2.3.3.3 Persönliche Relevanz

Bei jeder Entscheidungsaufgabe mussten die Versuchspersonen im Fragebogen 0 auf einer 10-Punkte-Skala angeben, wie wichtig sie selber die vorliegende Entscheidungsaufgabe nehmen würden (0 = „unwichtig“ bis 9 = „sehr wichtig“).

3.2.3.3.4 Verarbeitungsende bzw. Verarbeitungsschritte

Nachdem die Versuchspersonen auf jeden Verarbeitungsschritt hin ihre Präferenz angegeben hatten, mussten sie entscheiden, ob dies ihre endgültige Wahl sei oder ob sie einen weiteren Bericht zu Fr. 15'000.- kaufen wollten.

Die Anzahl Verarbeitungsschritte ergab sich aus dem Total der bezogenen Berichte. Bei den Versuchspersonen, die im Fragebogen 3 noch einmal Information wünschten, wurden 4 Arbeitsschritte erfasst.

3.2.3.3.5 Verarbeitungszeit

Auf jedem Fragebogen notierten die Versuchspersonen die genaue Uhrzeit. Berechnet wurde die Verarbeitungszeit jeder Entscheidung aus der Zeitangabe auf dem zuletzt ausgefüllten Fragebogen minus der Zeitangabe auf dem Fragebogen 0.

3.2.3.3.6 Gewünschte Urteilssicherheit

Zu Beginn jeder Entscheidungsaufgabe mussten die Versuchspersonen im Fragebogen 0 auf einer 15-Punkte-Skala angeben, wie sicher sie bei ihrem endgültigen Entscheid sein möchten (33% = „sehr unsicher“ bis 100% = „ganz sicher“).

3.2.3.3.7 Aktuelle Urteilssicherheit

In den Fragebögen 0 bis 3 wurden die Versuchspersonen jedes Mal gefragt, wie sicher sie sind, dass sich die im Moment bevorzugte Firma bzw. der momentan bevorzugte Kandidat in der endgültigen Entscheidung als die beste Wahl herausstellt. Die 15-Punkte-Skala ging von 33% für „sehr unsicher“ bis 100% für „ganz sicher“.

3.2.3.3.8 Sicherheitsdefizit

Das Sicherheitsdefizit wurde für jeden Verarbeitungsschritt aus der gewünschten Urteilssicherheit minus der aktuellen Urteilssicherheit berechnet.

3.2.3.3.9 Anspruchsniveau

Bei jeder Entscheidungsaufgabe mussten die Versuchspersonen im Fragebogen 0 auf einer 10-Punkte-Skala angeben, wie geeignet die Firma bzw. der Kandidat sein sollte, der die Zusage erhält (0 = „gar nicht“ bis 9 = „vortrefflich“).

3.2.3.3.10 Brauchbarkeitseinschätzung

Auf jedem Fragebogen (0 bis 3) wurden die Versuchspersonen gefragt, wie geeignet die bevorzugte Firma für den zu vergebenden Auftrag bzw. wie geeignet der bevorzugte Kandidat für die zu besetzende Position sei (0 = „gar nicht“ bis 9 = „vortrefflich“).

3.2.3.3.11 Brauchbarkeits-Lücke

Die Brauchbarkeits-Lücke wurde für jeden Verarbeitungsschritt aus dem Anspruchsniveau minus der jeweiligen aktuellen Brauchbarkeit berechnet.

3.2.3.3.12 Differenzwahrnehmung

Auf jedem Fragebogen (0 bis 3) wurden die Versuchspersonen gefragt, wie stark sich die bevorzugte Alternative von der zweitbesten Alternative unterscheidet (0 = „gar nicht“ bis 9 = „sehr stark“).

3.2.4 Durchführung

3.2.4.1 Instruktion

Die Instruktion beinhaltete folgende acht Punkte: (1) Assessment-Center, (2) Dilemma bei Entscheidungen, (3) optimales Entscheidungsverhalten, (4) Teilaufgaben, (5) unterschiedliche Wichtigkeiten, (6) Informationskosten, (7) Vorgehen und (8) Feedback sowie Belohnung.

1. Das *Assessment-Center* wurde als Verfahren vorgestellt, das mit Hilfe möglichst realitätsnaher Aufgaben das Managementpotential von Personen bewertet.
2. Das *Dilemma bei Entscheidungen* wurde als Konflikt von Führungspersonen dargestellt, einerseits möglichst gute Entscheidungen fällen zu wollen und andererseits jedoch nur über begrenzte Zeit und Ressourcen zu verfügen.
3. *Optimales Entscheidungsverhalten* wurde als die Suche nach einem Kompromiss zwischen Entscheidungsqualität und Entscheidungskosten beschrieben. Den Studierenden wurde gesagt, die Untersuchung hätte zum Ziel, festzustellen, wie sie sich in den Rollen von Geschäftsleiterinnen und Geschäftsleitern verhalten, wenn sie Entscheidungen von unterschiedlicher Bedeutung zu fällen und gleichzeitig für die entscheidungsrelevante Information ein Preis zu bezahlen haben.
4. Den Versuchspersonen wurden die drei *Teilaufgaben* der Übung vorgestellt: (1) IT-Entscheidung, (2) P-Entscheidung und (3) Lektüre eines Textes als Sitzungsvorbereitung. Dabei wurde betont, dass nur die zwei ersten Aufgaben bewertet würden.
5. Die *unterschiedlichen Wichtigkeiten* von Entscheidungen wurden anhand des Gewinnes erklärt, der einem entgeht, wenn man anstelle einer besten, eine zweitbeste bzw. drittbeste Alternative wählt.
6. Auf die *Informationskosten* von Fr. 15'000.- pro Bericht wurde hingewiesen.
7. Das schrittweise *Vorgehen* bei den zwei Entscheidungen wurde erklärt. Dabei wurde auf die Möglichkeit hingewiesen, nach jedem Verarbeitungsschritt den Entscheidungsprozess beenden zu können. Die Reihenfolge der zu lösenden Aufgaben wurde vorgegeben: Zuerst sollte die IT-Entscheidung und dann die P-Entscheidung gefällt werden.
8. Ein *Feedback* und eine leistungsbezogene *Belohnung* wurden angekündigt (vgl. Abschnitt 3.2.4.8).

Abschliessend wurde eine Zusammenfassung der Instruktion verteilt. Nach der Lektüre erhielten die Versuchspersonen die Gelegenheit, Fragen zu stellen.

3.2.4.2 Zufällige Zuordnung der Versuchspersonen zu den Versuchsbedingungen

Die Versuchspersonen wurden einer von zwei Versuchsbedingungen zugelost: „IT/p“ und „it/P“. Abschliessend wurden weitere Unterlagen verteilt: ein Deckblatt, das als Notizblatt verwendet werden konnte, je eine Anleitung zur IT- und P-Entscheidung sowie der IT-Fragebogen 0.

3.2.4.3 Vorinformation IT-Entscheidung

In der schriftlichen Anleitung zur IT-Entscheidung wurden zuerst die Entscheidungskriterien und die Nutzenunterschiede zwischen den Optionen beschrieben, danach wurde die Zielsetzung noch einmal genannt: „Finden sie die beste Softwarefirma, und versuchen sie, die Informationskosten möglichst tief zu halten.“ Der abschliessend auszufüllende IT-Fragebogen 0 enthielt zwei Spalten. In der ersten standen pro Firma einige Vorinformationen und in der zweiten waren die Fragen aufgelistet.

Die Vorinformationen enthielten zu jeder Firma Angaben über das Gründungsdatum, den Gründungsort und die Verteilung der beschäftigten Fachleute auf die Firmenstandorte. Entscheidungsrelevant war einzig die Summe der Beschäftigten: Bei allen drei Firmen waren es 60 Fachleute, die an unterschiedlichen Standorten arbeiteten (Tabelle 3-8).

Tabelle 3-8. Vorinformation über die drei zur Auswahl stehenden Firmen

▪	AA-Software AG wurde 1978 in Zürich gegründet und beschäftigt heute 10 Fachleute in Bern, 30 in Zürich und 20 in Luzern.
▪	BB-Software AG wurde 1979 in Bern gegründet und beschäftigt heute 20 Fachleute in Bern, 30 in Zürich und 10 in Luzern.
▪	CC-Software AG wurde 1980 in Luzern gegründet und beschäftigt heute 20 Fachleute in Bern, 20 in Zürich und 20 in Luzern.

Jede Versuchsperson bestimmte ihr Arbeitstempo selber und meldete sich beim Versuchsleiter, sobald sie den Fragebogen ausgefüllt hatte. Der Versuchsleiter sammelte dann den Fragebogen ein und übergab den ersten Consultant-Bericht und den dazugehörigen Fragebogen.

3.2.4.4 Verarbeitungsschritte 1-3 der IT-Entscheidung

Alle Versuchspersonen erhielten zunächst den ersten Consultant-Bericht, verarbeiteten diesen und beantworteten abschliessend die Fragen des Fragebogens 1 (Tabelle 3-9). Bei einem endgültigen Urteil oder wenn die Versuchspersonen beim dritten Verarbeitungsschritt angelangt waren, wurde ihnen der Fragebogen 0 der P-Aufgabe ausgeteilt. In allen anderen Fällen erhielten sie einen weiteren Bericht mit neuem Fragebogen. Im letzten Fragebogen konnten die Versuchspersonen angeben, ob sie noch einen vierten Consultant-Bericht kaufen möchten. Dabei wurde der Hinweis gemacht, dass bei einer positiven Antwort auf diese Frage, keine zusätzlichen Kosten angerechnet würden.

Tabelle 3-9. Fragen der Fragebögen 1-3 bei der IT-Entscheidung

-
1. Wenn Sie jetzt eine Firma wählen müssten, welche würden Sie favorisieren?
 2. Ist dies Ihre endgültige Entscheidung?
 3. Wie geeignet ist die von Ihnen bevorzugte Firma für den zu vergebenden Auftrag?
 4. Wie sicher sind Sie, dass sich die von Ihnen im Moment bevorzugte Firma nach Ihrem endgültigen Entscheid als die beste Wahl herausstellt [%]?
 5. Welche Firma ist im Moment für Sie die Nr. 2 bzw. die Nr. 3?
 6. Wie stark unterscheidet sich die bevorzugte Firma von der zweitbesten Firma?
-

3.2.4.5 Vorinformation P-Entscheidung

Die Anleitung enthielt eine Beschreibung der Entscheidungskriterien, eine Bezifferung des entgangenen Gewinns bei der Wahl eines zweitbesten bzw. drittbesten Verkäufers sowie den folgendermassen ausformulierten Auftrag: „Finden sie den besten Kandidaten und versuchen sie, die Interviewkosten möglichst tief zu halten“. Der P-Fragebogen 0 war zweiseitig und enthielt in der ersten Spalte einige Vorinformationen über die Bewerber und in der zweiten Spalte die aufgelisteten Fragen. Die Vorinformation umfasste eine kurze Beschreibung der drei Kandidaten. Die darin genannten Charaktermerkmale betrafen bei allen das Kriterium „Verkaufsgeschick/Überzeugungskraft“ und lieferten damit keinen objektiven Grund, einen der drei Kandidaten zu bevorzugen (Tabelle 3-10). Nach dem Beantworten aller Fragen, erhielten die Versuchspersonen den ersten Interview-Bericht und den dazugehörigen Fragebogen.

Tabelle 3-10. Vorinformation über die drei zur Auswahl stehenden Kandidaten

-
- Kandidat A kann andere für seine Ideen begeistern. Er ist ein anregender und sympathischer Gesprächspartner.
 - Kandidat B ist über Konkurrenzaktivitäten gut informiert. Er präsentiert Fachinformation verständlich.
 - Kandidat C versteht Kundenprobleme und geht darauf ein. Er ist bereit, mehr zu leisten als von ihm gefordert wird.
-

3.2.4.6 Verarbeitungsschritte 1-3 der P-Entscheidung

Alle Studierenden bearbeiteten den ersten Interview-Bericht, beantworteten dann den dazugehörigen Fragebogen (Tabelle 3-11). Wenn noch kein endgültiger Entscheid vorlag, wurde ein weiterer Bericht mit Fragebogen verteilt. Im letzten Fragebogen konnten die Versuchspersonen angeben, ob sie noch eine vierte Interviewserie in Auftrag geben möchten, ohne dass ihnen zusätzliche Kosten angerechnet würden.

Tabelle 3-11. Fragen der Fragebögen 1-3 bei der P-Entscheidung

-
1. Wenn Sie jetzt einen Kandidaten wählen müssten, welchen würden Sie favorisieren?
 2. Ist dies Ihre endgültige Entscheidung?
 3. Wie geeignet ist der von Ihnen bevorzugte Kandidat für die zu besetzende Position?
 4. Wie sicher sind Sie, dass sich der von Ihnen im Moment bevorzugte Kandidat nach Ihrem endgültigen Entscheid als die beste Wahl herausstellt [%]?
 5. Welcher Kandidat ist im Moment für Sie Nr. 2 bzw. Nr. 3?
 6. Wie stark unterscheidet sich der bevorzugte Kandidat vom zweitbesten Kandidaten?
-

3.2.4.7 Abschlussfragebogen und Lektüre

Sowohl nach einer definitiven Entscheidung als auch nach dem dritten Verarbeitungsschritt wurde der Abschlussfragebogen sowie ein mehrseitiger Bericht zum Thema Personalauswahl ausgehändigt. Der letzte Fragebogen enthielt den *Need for Cognition*-Test. Die ausgeteilte Lektüre hatte für die Übung keine Bedeutung. Sie sollte einzig verhindern, dass schnelle Versuchspersonen den Raum früher verlassen und dadurch einen Nachahmungseffekt bei den vorsichtigeren Entscheidern erzeugten. Diesbezüglich wurde nämlich in Vorexperimenten ohne Beschäftigungslektüre beobachtet, dass es Perso-

nen gab, die mit Nervosität und einer Beschleunigung ihrer Entscheidung reagierten, wenn eine Mehrheit die Übung abgeschlossen und den Unterrichtsraum verlassen hatte.

3.2.4.8 Feedback und leistungscontingente Belohnung

Nach Ablauf von zwei Wochen erhielten die Versuchspersonen ein Feedback und die angekündigte leistungsbezogene Belohnung wurde ausbezahlt. Als Kriterien zur Leistungsbewertung und Bezahlung dienten in erster Priorität Entscheidungsqualität und Kosten, in zweiter Priorität Zeitbedarf. Pro Klasse wurden zwei erste Preise à Fr. 20.-, zwei zweite Preise à Fr. 10.- und vier dritte Preise à Fr. 5.- verteilt.

3.2.5 Datenanalyse

Die Auswertung der Daten erfolgte auf drei Ebenen: (1) Ebene der Versuchspersonen, (2) Ebene der Entscheidungen und (3) Ebene der einzelnen Meinungsäußerungen (Abbildung 3-2). Die Reaktion der Soll-Werte auf die Aufgaben-Wichtigkeit (Funktionstest 1) wurde auf den Ebenen der Versuchspersonen (Within-Subjects) und der Entscheidungen (Between-Subjects) geprüft. Der Anstieg der Ist-Werte nach der Verarbeitung der Informations-Berichte (Funktionstest 2) wurde auf der Ebene der einzelnen Meinungsäußerungen untersucht. Ebenfalls auf dieser Ebene wurde geprüft, ob die Regelabweichungen „letzte“ Schritte erkennen (Funktionstest 3). Die Theorien zur Entstehung der Brauchbarkeitseinschätzung und der Urteilssicherheit wurden auf der Ebene der Meinungsäußerungen getestet.

Mittelwertvergleiche wurden mit Hilfe des Wilcoxon-Tests (Within-Subjects) und des U-Tests von Mann-Whitney (Between-Subjects) gemacht, da sich die Daten häufig als *nicht* normalverteilt erwiesen. Die Analyse der multivariaten Struktur der Variablenbeziehungen erfolgte mit schrittweisen multiplen Regressionsanalysen. Es sollte geprüft werden, ob und welche Varianzanteile insbesondere die Führungsgrößen, Regelgrößen, Regelabweichungen und die Informationsnutzung erklären. Dazu wurden die Variablen in theoretisch begründeten Blöcken, und innerhalb der Blöcke schrittweise nach ihrem Vorhersagebeitrag, in die Regressionsgleichungen aufgenommen.

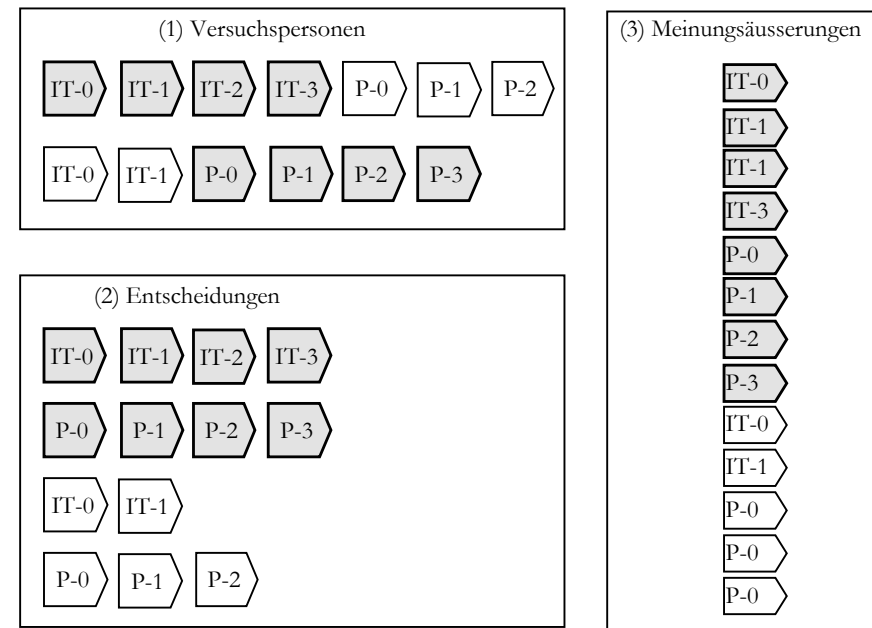


Abbildung 3-2. Auswertung der Daten innerhalb (1) der Ebene der Versuchspersonen, (2) der Ebene der Entscheidungen und (3) der Ebene der einzelnen Meinungsäußerungen

Trotz der Tatsache, dass *nicht* alle beteiligten Variablen jeweils multivariat normalverteilt waren, konnte davon ausgegangen werden, dass die Voraussetzungen für eine interferenzstatistische Absicherung der multiplen Korrelationen gegeben waren, weil der Stichprobenumfang im Verhältnis zur Anzahl der Variablen ($n > 40$ bei $k < 10$) genügte (Bortz, 1999). Dort, wo andere Vorgehensweisen zur Auswertung der Daten verwendet wurden, werden diese im Zusammenhang mit der Ergebnisdarstellung beschrieben.

3.3 Ergebnisse

3.3.1 Manipulations-Überprüfung

3.3.1.1 Frage nach der wichtigeren Entscheidung

In jeder Versuchsbedingung wurde jeweils eine Aufgabe als sehr wichtig bezeichnet. Im Treatment „IT/p“ war die IT-Aufgabe die sehr wichtige und im Treatment „it/P“ die P-Aufgabe (Abbildung 3-1). Zur Prüfung der Manipulation wurden die Versuchspersonen in einem ersten Fragebogen gefragt, welche der beiden zu lösenden Aufgaben mehr als zehnmal wichtiger sei als die andere. Die richtige Antwort musste lauten: „Informatikentscheidung“ in der Bedingung „IT/p“ und „Personalentscheidung“ in der Bedingung „it/P“. 34 der 35 Versuchspersonen beantworteten diese Frage und in 29 Fällen war sie zutreffend. Die Wichtigkeits-Manipulation wurde somit verstanden; $\chi^2 = 17.6$, $df = 1$, $p < .001$.

3.3.1.2 Relevanz für Firma

Die Versuchspersonen wurden gefragt, wie wichtig sie aus Sicht der Firma die zu lösenden Entscheidungsaufgaben bewerten. Es wurde geprüft, ob die jeweils wichtigere Aufgabe als relevanter eingeschätzt wurde.

Der *Within-Subjects-Vergleich* in der Versuchsbedingung „IT/p“ ergab für die IT-Entscheidung eine höhere Relevanz ($M = 8.4$) als für die P-Entscheidung ($M = 6.1$); $T = 3$, $p < .01$ (Tabelle 3-12). 10 von 17 Personen fanden die IT-Aufgabe sei wichtiger für die Firma als die P-Aufgabe, eine Person hatte eine gegenteilige Präferenz und 7 Personen machten keinen Unterschied in ihrer Relevanzeinschätzung. In der Versuchsbedingung „it/P“ zeigte das Treatment ebenfalls die erwartete Wirkung: Die IT-Entscheidung ($M = 5.5$) wurde als weniger wichtiger betrachtet als die P-Entscheidung ($M = 8.5$); $T = 7$, $p < .01$. 15 von 17 Personen bewerteten die P-Aufgabe als wichtiger und zwei die IT-Aufgabe.

Der *Between-Subjects-Vergleich* zwischen wichtigen und weniger wichtigen Entscheidungen ergab auch ein signifikantes Ergebnis ($M = 8.5$ versus $M = 5.8$); $\xi = -5.01$, $p < .001$.

Tabelle 3-12. Mittelwerte und Standardabweichungen der Wichtigkeit der Entscheidung für die Firma (IT- und P-Entscheidung) in den zwei Versuchsbedingungen

	IT-Entscheidung	P-Entscheidung
IT/p	8.4 (.9)	6.1 (2.5)
it/P	5.5 (2.2)	8.5 (.9)

Standardabweichungen sind in Klammern dargestellt

3.3.1.3 Persönliche Relevanz

Die Versuchspersonen gaben an, wie wichtig sie selber die beiden Entscheidungsaufgaben nahmen. Damit sollte geprüft werden, ob das Treatment eine Wirkung auf die persönlich wahrgenommene Relevanz hatte.

Der *Within-Subjects-Vergleich* in Tabelle 3-13 zeigt in der Bedingung „IT/p“ signifikant höhere Werte bei der IT-Entscheidung ($M = 8.4$) verglichen mit der P-Entscheidung ($M = 6.8$); $T = 9$, $p < .05$. 12 von 18 Personen bewerteten IT wichtiger als P und 6 Personen machten keinen Unterschied. In der Bedingung „it/P“ war der Unterschied ebenfalls signifikant (IT-Entscheidung: $M = 5.5$ versus P-Entscheidung: $M = 8.7$); $T = 0$, $p < .01$. 15 von 17 Personen machten die erwartete Bewertung (P wichtiger als IT) und nur zwei gaben gleich hohe Wichtigkeitswerte an.

Tabelle 3-13. Mittelwerte und Standardabweichungen der persönlichen Relevanz (IT- und P-Entscheidung) in den zwei Versuchsbedingungen

	IT-Entscheidung	P-Entscheidung
IT/p	8.4 (1.2)	6.8 (2.5)
it/P	6.2 (1.9)	8.7 (.8)

Standardabweichungen sind in Klammern dargestellt

Der *Between-Subjects-Vergleich* zwischen wichtigen und weniger wichtigen Entscheidungen ergab ebenfalls ein signifikantes Ergebnis ($M = 8.5$ versus $M = 6.5$); $\xi = -4.60$, $p < .001$.

3.3.2 Stopp-Mechanismen im Test

3.3.2.1 Reaktion der Führungsgrösse (Funktionstest 1)

3.3.2.1.1 Gewünschte Urteilssicherheit

Ob die Aufgaben-Wichtigkeit einen Einfluss auf die gewünschte Urteilssicherheit hatte, wurde mittels zweier Within-Subjects-Vergleiche und eines Between-Subjects-Vergleichs geprüft.

In der Versuchsbedingung „IT/p“ gaben die Versuchspersonen bei der IT-Entscheidung zwar einen durchschnittlich höheren Wert für die gewünschte Urteilssicherheit an ($M = 95.1$) als bei der P-Entscheidung ($M = 91.4$), der Unterschied war aber nicht signifikant; $T = 10$ (Tabelle 3-14). Sechs von 18 Personen wünschten bei der endgültigen IT-Entscheidung sicherer zu sein als bei der P-Entscheidung, Neun Personen machten keinen Unterschied und drei wollten bei der P-Entscheidung sicherer sein. Auch in der Bedingung „it/P“ war der Unterschied (IT-Entscheidung: $M = 90.6$ versus P-Entscheidung: $M = 94.7$) nicht signifikant; $T = 10.5$. Sieben von 17 Personen strebten eine höhere Urteilssicherheit bei der P-Entscheidung an, sieben waren indifferent und drei wollten bei der IT-Entscheidung sicherer werden.

Tabelle 3-14. Mittelwerte und Standardabweichungen der gewünschten Urteilssicherheit bei IT-Entscheidung und P-Entscheidung in den zwei Versuchsbedingungen

	IT-Entscheidung	P-Entscheidung
IT/p	95.1	91.4
	(5.2)	(8.2)
it/P	90.6	94.7
	(10.4)	(6.2)

Standardabweichungen sind in Klammern dargestellt

Abbildung 3-3 zeigt die Mittelwerte der gewünschten Urteilssicherheit bei den weniger wichtigen und den wichtigeren Entscheidungen aus der *Between-Subjects-Perspektive*. In den weniger wichtigen Entscheidungen wurde eine signifikant tiefere Urteilssicherheit angestrebt ($M = 91.0\%$) als in den wichtigeren Entscheidungen ($M = 94.9\%$); $z = -1.72$, $p < .05$ (einseitig).

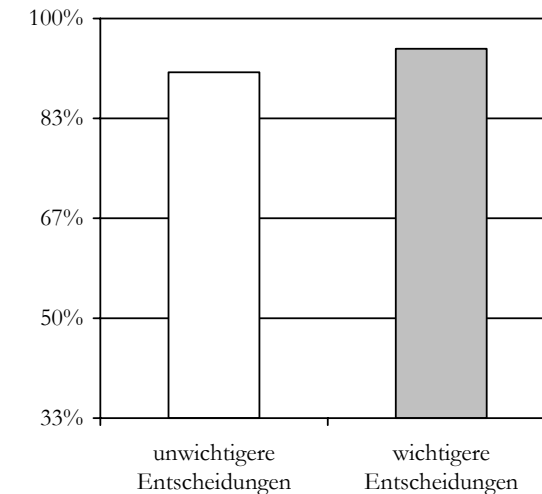


Abbildung 3-3. Mittelwerte der gewünschten Urteilssicherheit bei den weniger wichtigen und den wichtigeren Entscheidungen

3.3.2.1.2 Anspruchsniveau

Der *Within-Subjects-Vergleich* in der Versuchsbedingung „IT/p“ ergab zwar für die IT-Entscheidung ein höheres Anspruchsniveau ($M = 8.5$) als bei der P-Entscheidung ($M = 8.2$), aber der Unterschied war nicht signifikant; $T = 15.5$ (Tabelle 3-15). Sieben von 17 Personen gaben bei der IT-Entscheidung ein höheres Anspruchsniveau an als bei der P-Entscheidung, sieben Personen machten keinen Unterschied und drei hatten höhere Ansprüche bei der P-Aufgabe. In der Bedingung „it/P“ gaben die Versuchspersonen bei der IT-Entscheidung ein signifikant tieferes Anspruchsniveau an ($M = 8.4$) als bei der P-Entscheidung ($M = 8.8$), $T = 0$, $p < .05$. Sechs von 17 Personen verzeichneten ein höheres Anspruchsniveau bei der P-Entscheidung und elf waren indifferent.

Tabelle 3-15. Mittelwerte und Standardabweichungen des Anspruchsniveaus bei IT-Entscheidung und P-Entscheidung in den zwei Versuchsbedingungen

	IT-Entscheidung	P-Entscheidung
IT/p	8.5 (.86)	8.2 (.95)
it/P	8.4 (.44)	8.8 (.70)

Standardabweichungen sind in Klammern dargestellt

Der *Between-Subjects-Vergleich* der Mittelwerte des Anspruchsniveaus ergab bei den weniger wichtigen Entscheidungen ein signifikant tieferes Anspruchsniveau ($M = 8.3$) als bei den wichtigeren Entscheidungen ($M = 8.6$); $z = -2.44$, $p < .05$ (Abbildung 3-4).

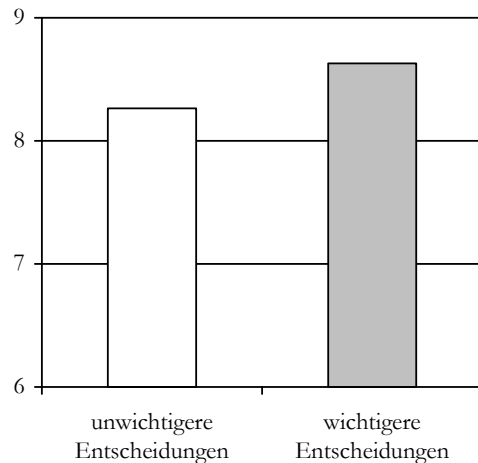


Abbildung 3-4. Mittelwerte des Anspruchsniveaus bei den weniger wichtigen und den wichtigeren Entscheidungen

3.3.2.2 Reaktion der Regelgröße (Funktionstest 2)

3.3.2.2.1 Aktuelle Urteilssicherheit

Ob die Urteilssicherheit mit zunehmender Informationsnutzung ansteigt (Within-Subjects-Annahme), wurde mit Hilfe eines Friedmann-Tests geprüft. Alle Entscheidungen wurden in Abhängigkeit vom Verarbeitungsende in drei Gruppen eingeteilt: Die Gruppe der *Ein-Schritt-Entscheidungen* enthielt alle IT- und P-Entscheidungen, welche nach dem ersten Verarbeitungsschritt beendet wurden, die *Zwei-Schritt-Entscheidungen* jene, die nach dem 2. Schritt ein Ende fanden und die *Drei-Schritt-Entscheidungen*, diejenigen, welche nach dem 3. Schritt endeten.

Abbildung 3-5 gibt die Mittelwerte der *Urteilssicherheit* bei allen Gruppen von Entscheidungen wieder. Der Anstieg der Urteilssicherheit war bei den 17 Ein-Schritt- ($M_0 = 58\%$, $M_1 = 79\%$; $\chi^2(1) = 12.25$, $p < .01$), den 26 Zwei-Schritt- ($M_0 = 56\%$, $M_1 = 67\%$, $M_2 = 91\%$; $\chi^2(2) = 44.42$, $p < .001$) und den 21 Drei-Schritt-Entscheidungen ($M_0 = 44\%$, $M_1 = 58\%$, $M_2 = 71\%$, $M_3 = 89\%$; $\chi^2(3) = 54.23$, $p < .001$) signifikant.

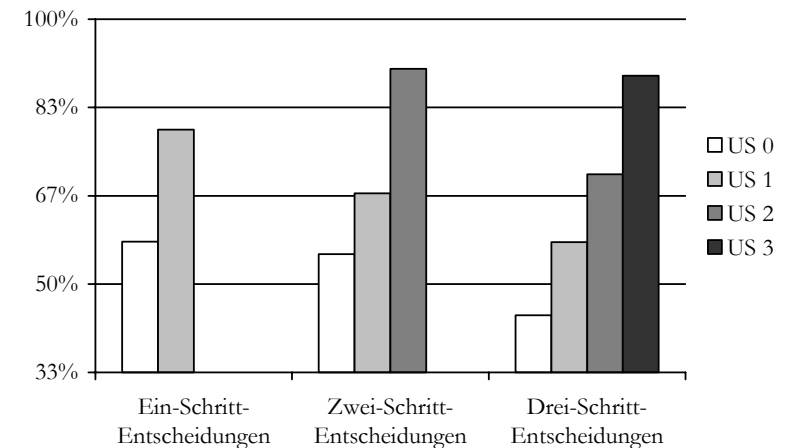


Abbildung 3-5. Mittelwerte der Urteilssicherheit nach der Vorinformation (US0) und den Verarbeitungsschritten 1 bis 3 (US1 bis US3) bei Ein- bis Drei-Schritt-Entscheidungen

Tabelle 3-16 stellt einen Vergleich zwischen den Gruppen dar. Sowohl nach dem ersten als auch nach dem zweiten Verarbeitungsschritt wurden die definitiven Urteile mit einer höheren Sicherheit abgegeben als die provisorischen. Nach dem ersten Verarbeitungsschritt wiesen die Ein-Schritt-Entscheidungen eine signifikant höhere Urteilssicherheit auf ($M = 79.1$) als die Zwei- ($M = 67.1$) und Drei-Schritt-Entscheidungen ($M = 57.9$); Tukey-HSD auf 5%-Niveau. Nach dem zweiten Verarbeitungsschritt war die Urteilssicherheit bei den Zwei-Schritt-Entscheidungen signifikant höher ($M = 90.6$) als bei den Drei-Schritt-Entscheidungen ($M = 71.1$); $z = -4.73, p < .001$.

Tabelle 3-16. Mittelwerte der Urteilssicherheit nach den Verarbeitungsschritten 1 und 2

	Ein-Schritt-Entsch.	Zwei-Schritt-Entsch.	Drei-Schritt-Entsch.
nach Schritt 1	79.1 _a	67.1 _b	57.9 _b
nach Schritt 2		90.6 _a	71.1 _b

Suffixe stehen für sich unterscheidende Mittelwerte (Tukey-HSD auf 5%-Niveau)

Die Wirkung der genutzten Informationsmenge auf die Höhe der *definitiven Urteilssicherheit* wurde ebenfalls geprüft. Dieser Einfluss war signifikant, wie ein Kruskal-Wallis-Test ergab; $\chi^2(2) = 7.0, p < .05$. Bei den IT-Entscheidungen war die definitive Urteilssicherheit um so höher, je mehr Information bis zur Entscheidung genutzt wurde (Tabelle 3-17). Dabei waren die Unterschiede zwischen den 1- ($M = 75.5$) und Drei-Schritt-Entscheidungen ($M = 96.0$) signifikant. Bei den P-Entscheidungen bewegten sich die Unterschiede in einem kleineren Rahmen (zwischen $M = 84.3$ für die Ein-Schritt- und $M = 90.8$ für die Zwei-Schritt-Entscheidungen). Die Unterschiede erwiesen sich als nicht signifikant. Insgesamt wurden die Ein-Schritt-Entscheidungen mit einer signifikant tieferen Urteilssicherheit gefällt ($M = 79.1$) als die Zwei-Schritt- ($M = 90.6$) und Drei-Schritt-Entscheidungen ($M = 89.8$).

Tabelle 3-17. Mittelwerte der Urteilssicherheit nach Abgabe des definitiven Urteils bei den Ein- bis Drei-Schritt-Entscheidungen

	Ein-Schritt-Entsch.	Zwei-Schritt-Entsch.	Drei-Schritt-Entsch.
IT-Entscheidungen	75.5 _a	90.4 _{ab}	96.0 _b
P-Entscheidungen	84.3 _a	90.8 _a	84.6 _a
Insgesamt	79.1 _a	90.6 _b	89.8 _b

Suffixe stehen für sich unterscheidende Mittelwerte (Games-Howell auf 5%-Niveau)

3.3.2.2.2 Brauchbarkeitseinschätzung

Abbildung 3-6 zeigt, dass die Mittelwerte der *Brauchbarkeitseinschätzung* bei allen Gruppen von Entscheidungen mit fortlaufender Informationsnutzung zunahm. Dieser Anstieg war bei den 17 Ein-Schritt- ($M_0 = 5.5, M_1 = 7.4$; $\chi^2(1) = 9.0, p < .01$), den 26 Zwei-Schritt- ($M_0 = 5.2, M_1 = 6.5, M_2 = 8.3$; $\chi^2(2) = 44.33, p < .001$) und den 21 Drei-Schritt-Entscheidungen ($M_0 = 5.1, M_1 = 6.4, M_2 = 7.0, M_3 = 8.3$; $\chi^2(3) = 48.35, p < .001$) signifikant.

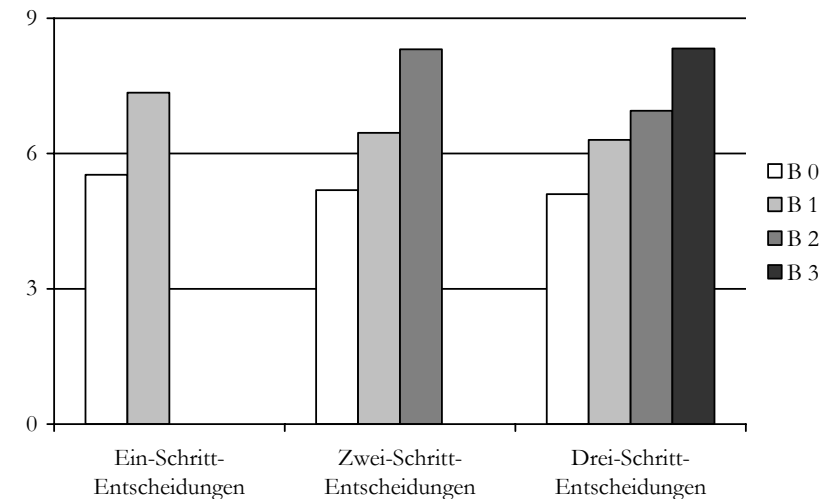


Abbildung 3-6. Mittelwerte der aktuellen Brauchbarkeitseinschätzung nach der Vorinformation (B 0) und den Verarbeitungsschritten 1 bis 3 (B 1 bis B 3) Ein- bis Drei-Schritt-Entscheidungen

Der Vergleich zwischen den Gruppen ergibt nach dem ersten Verarbeitungsschritt bei den Ein-Schritt-Entscheidungen eine signifikant höhere Brauchbarkeitseinschätzung ($M = 7.4$) als bei den Zwei- ($M = 6.5$) und Drei-Schritt-Entscheidungen ($M = 6.4$); Tukey-HSD auf 5%-Niveau (Tabelle 3-18). Nach dem zweiten Verarbeitungsschritt war die wahrgenommene Qualität bei den Zwei-Schritt-Entscheidungen signifikant höher ($M = 8.3$) als bei den Drei-Schritt-Entscheidungen ($M = 7.0$); $z = -4.31, p < .001$.

Tabelle 3-18. Mittelwerte der Brauchbarkeitseinschätzung nach den Verarbeitungsschritten 1 und 2 bei den Ein- bis Drei-Schritt-Entscheidungen

	Ein-Schritt-Entsch.	Zwei-Schritt-Entsch.	Drei-Schritt-Entsch.
Urteil nach Schritt 1	7.4 _a	6.5 _b	6.4 _b
Urteil nach Schritt 2		8.3 _a	7.0 _b

Suffixe stehen für sich unterscheidende Mittelwerte (Tukey-HSD auf 5%-Niveau)

Die Verarbeitungsschritte hatten ebenfalls eine Wirkung auf die Höhe der definitiven Brauchbarkeitseinschätzung, $\chi^2(2) = 14.8, p < .001$. Tabelle 3-19 gibt die Ergebnisse des Post-Hoc-Mehrfachvergleichs wieder. Bei den IT-Entscheidungen wurde die gewählte Alternative um so besser eingeschätzt, je mehr Information bis zur Entscheidung verarbeitet wurde. Dabei waren die Unterschiede zwischen den Ein-Schritt- ($M = 7.3$) bzw. den Zwei- ($M = 8.3$) und Drei-Schritt-Entscheidungen ($M = 8.9$) signifikant. Bei den P-Entscheidungen waren die Unterschiede kleiner (minimal $M = 7.4$ bei den Ein-Schritt- und maximal $M = 8.3$ bei den Zwei-Schritt-Entscheidungen). Zudem erwiesen sie sich als nicht signifikant. Über beide Aufgaben gesehen, wurde die definitive Brauchbarkeit bei den Ein-Schritt-Entscheidungen signifikant tiefer eingeschätzt ($M = 7.4$) als bei den Zwei- ($M = 8.3$) und Drei-Schritt-Entscheidungen ($M = 8.4$).

Tabelle 3-19. Mittelwerte der Brauchbarkeitseinschätzung nach Abgabe des definitiven Urteils bei den Ein- bis Drei-Schritt-Entscheidungen

	Ein-Schritt-Entsch.	Zwei-Schritt-Entsch.	Drei-Schritt-Entsch.
IT-Entscheidungen	7.3 _a	8.3 _a	8.9 _b
P-Entscheidungen	7.4 _a	8.3 _a	7.9 _a
Insgesamt	7.4 _a	8.3 _b	8.4 _b

Suffixe stehen für sich unterscheidende Mittelwerte (Games-Howell bzw. Tukey-HSD auf 5%-Niveau)

3.3.2.3 Wirkung der Regelabweichung (Funktionstest 3)

3.3.2.3.1 Erkennen „letzter“ Schritte

Die Erwartung, dass das Sicherheitsdefizit und die Brauchbarkeits-Lücke „letzte“ Schritte erkennen, wurde anhand der 149 Prozessschritte untersucht,

bei denen die Versuchspersonen eine definitive Entscheidung fällen konnten (Verarbeitungsschritte 1 bis 3).

Wie Abbildung 3-7 (a) zeigt war das Sicherheitsdefizit bei den „letzten“ Verarbeitungsschritten deutlich kleiner, ($M = 5.3\%$), als bei den Prozessschritten, die fortgesetzt wurden ($M = 30.5\%$); $\bar{\kappa} = -8.79, p < .001$).

Bei der Brauchbarkeits-Lücke ergab sich ein sehr ähnliches Bild. Hier war die Brauchbarkeits-Lücke bei den „letzten“ Schritten ($M = .33$) deutlich kleiner als bei den Schritten mit Fortsetzung ($M = 2.05$); $\bar{\kappa} = -7.96, p < .001$ (Abbildung 3-7 b).

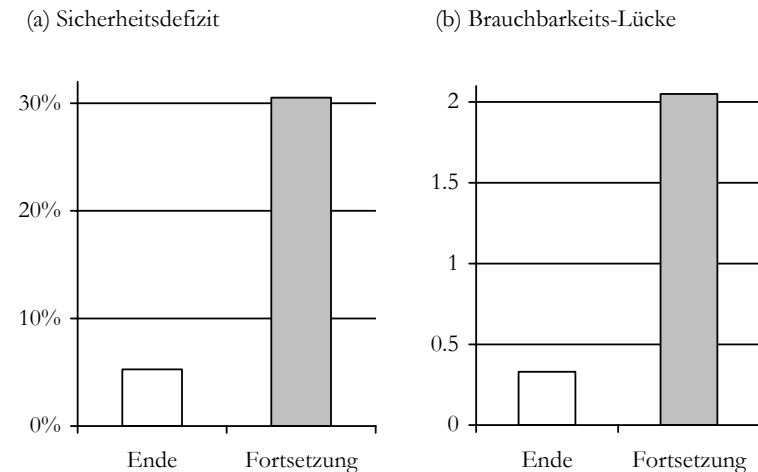


Abbildung 3-7. Mittelwerte der (a) Sicherheitsdefizit und (b) Brauchbarkeits-Lücke bei Verarbeitungsschritten ohne bzw. mit Fortsetzung

Vergleich der Regelabweichungen mit anderen Prädiktoren

Neben dem Sicherheitsdefizit und der Brauchbarkeits-Lücke erwiesen sich die Verarbeitungsschritte ($\bar{\kappa} = -5.11, p < .001$), die aktuelle Urteilssicherheit ($\bar{\kappa} = -7.87, p < .001$) und die Brauchbarkeitseinschätzung ($\bar{\kappa} = -7.17, p < .001$) als signifikante Prädiktoren (Tabelle 3-20). Gemessen am $\bar{\kappa}$ -Wert schlug sowohl das Sicherheitsdefizit als auch die Brauchbarkeits-Lücke die konkurrierenden Variablen.

Tabelle 3-20. Mittelwerte, Standardabweichungen und z -Wert der absolvierten Verarbeitungsschritte, Sicherheitsdefizit, aktuelle Urteilssicherheit, Brauchbarkeits-Lücke und Brauchbarkeitseinschätzung bei abschliessenden bzw. provisorischen Urteilen

	<i>M (SD) Ende</i>	<i>M (SD) Fortsetzung</i>	z -Wert
Verarbeitungsschritte	2.1 (.8)	1.4 (.6)	-5.11, $p < .001$
Sicherheitsdefizit	5% (11%)	31% (15%)	-8.79, $p < .001$
aktuelle Urteilssicherheit	87% (13%)	65% (15%)	-7.87, $p < .001$
Brauchbarkeits-Lücke	.3 (.9)	2.0 (1.0)	-7.95, $p < .001$
Brauchbarkeitseinschätzung	8.1 (1.0)	6.6 (1.1)	-7.01, $p < .001$

Vergleich der Regelabweichungen untereinander

Die z -Werte aus Tabelle 3-20 zeigen auch, dass das Sicherheitsdefizit ($z = -8.79$) besser diskriminierte als die Brauchbarkeits-Lücke ($z = -7.95$).

Mittels Diskriminanzanalyse wurde geprüft, wie gut die Vorhersage beider Regelabweichungen war. Die Klassifizierung war beim Sicherheitsdefizit in 83.1% der Fälle richtig, bei der Brauchbarkeits-Lücke in 80.3% der Fälle.

Abschliessend wurde geklärt, ob die Brauchbarkeits-Lücke neben dem Sicherheitsdefizit noch einen Vorhersagebeitrag leistet. Dies war der Fall und die Zuordnungsqualität stieg auf 84.4%.

3.3.2.3.2 Vorhersage späterer Informationsnutzung

Sowohl das *Sicherheitsdefizit nach der Vorinformation* ($\beta = .47, p < .001$) als auch die *Brauchbarkeits-Lücke nach der Vorinformation* ($\beta = .34, p < .01$) waren signifikante Prädiktoren der Kriteriumsvariablen Verarbeitungsschritte. Da aber bereits das Treatment *Aufgaben-Wichtigkeit* ($\beta = .59, p < .001$) einen signifikanten Einfluss auf die Informationsnutzung hatte, wurde geprüft, ob die Regelabweichungen über das Treatment hinaus einen Effekt auf die Informationsnutzung ausüben.

Sicherheitsdefizit

Neben der Aufgaben-Wichtigkeit (Modell 1) vermochte allein das Sicherheitsdefizit nach der Vorinformation einen Beitrag zur Vorhersage der Verarbeitungsschritte zu leisten. Die Ergebnisse in Tabelle 3-21 zeigen, dass das Sicherheitsdefizit die Varianzaufklärung um 11% steigerte; $F_{change}(1, 66) = 13.4, p < .001$ (Modell 2).

Tabelle 3-21. Sicherheitsdefizit und andere Prädiktoren zur Vorhersage der Verarbeitungsschritte

	R^2	ΔR^2	β	r
Modell 1				
Aufgaben-Wichtigkeit (Treatment)	.35		.59***	.59
Modell 2				
Aufgaben-Wichtigkeit (Treatment)	.45	.11***	.51***	.55
Sicherheitsdefizit nach Vorinformation			.34***	.41
Need for Cognition				.06
Selbstwirksamkeitserwartung				-.02
Relevanz für Firma				.25
Persönliche Relevanz				.18
Differenzwahrnehmung nach Vorinformation				-.06
Gewünschte Urteilssicherheit				.22
Aktuelle Urteilssicherheit nach Vorinformation				.22

*** $p < .001$

Abbildung 3-8 fasst die Ergebnisse zusammen: Die Aufgaben-Wichtigkeit hatte eine stärkere Wirkung auf die absolvierten Verarbeitungsschritte ($\beta = .51, p < .001$) als das Sicherheitsdefizit nach der Vorinformation ($\beta = .34, p < .001$). Neben dem direkten Effekt der Aufgaben-Wichtigkeit war aber auch ein indirekter über das Sicherheitsdefizit nach der Vorinformation ($\beta = .24, p < .001$) zu beobachten.

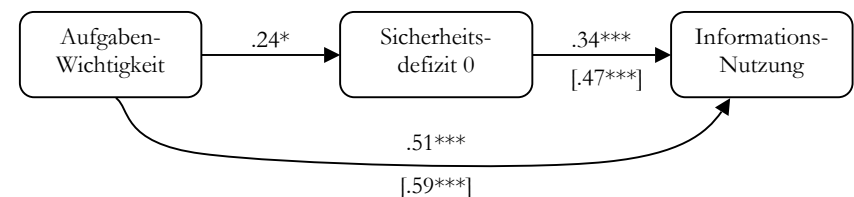


Abbildung 3-8. Direkter und indirekter Zusammenhang zwischen Aufgaben-Wichtigkeit, anfänglicher Sicherheitsdefizit und Informationsnutzung (β -Gewichte; *** $p < .001$; * $p < .05$; [] = β -Gewicht ohne Berücksichtigung des alternativen Pfades)

Brauchbarkeits-Lücke

Die Ergebnisse in Tabelle 3-22 zeigen, dass die Brauchbarkeits-Lücke nach der Vorinformation auch neben der Aufgaben-Wichtigkeit eine Wirkung auf die Informationsnutzung hatte und die Varianzaufklärung um 7% verbesserte; $F_{change}(1, 66) = 7.4, p < .01$ (Modell 2).

Bei Berücksichtigung der zusätzlichen Variablen Need for Cognition, Selbstwirksamkeitserwartung, Relevanz für Firma, Persönliche Relevanz, Differenzwahrnehmung nach der Vorinformation, Anspruchsniveau sowie Brauchbarkeitseinschätzung nach der Vorinformation erschien anstelle der Brauchbarkeits-Lücke der Prädiktor *Anspruchsniveau* in der Gleichung; $F_{change}(1, 66) = 8.1, p < .01$ (Modell 3). Verglichen mit Modell 2 wurde damit aber keine zusätzliche Varianzaufklärung erzielt.

Anders verhielt es sich innerhalb der Versuchsbedingung sehr wichtige Entscheidungen. Hier erwies sich die Brauchbarkeits-Lücke ($\beta = .50, p < .01$) als zuverlässigerer Prädiktor als die des Anspruchsniveaus ($\beta = .38, p < .05$).

Tabelle 3-22. Brauchbarkeits-Lücke und andere Prädiktoren zur Vorhersage der Verarbeitungsschritte

	R^2	ΔR^2	β	r
Modell 1				
Aufgaben-Wichtigkeit (Treatment)	.34		.58***	.58
Modell 2				
Aufgaben-Wichtigkeit (Treatment)	.41	.07***	.55***	.57
Brauchbarkeits-Lücke nach Vorinformation			.26**	.32
Modell 3				
Aufgaben-Wichtigkeit (Treatment)	.41	.07***	.52***	.55
Anspruchsniveau			.28**	.33
Need for Cognition				.06
Selbstwirksamkeitserwartung				-.06
Relevanz für Firma				.08
Persönliche Relevanz				.10
Differenzwahrnehmung nach Vorinformation				-.19
Brauchbarkeits-Lücke nach Vorinformation				.29
Brauchbarkeitseinschätzung nach Vorinformation				.29

*** $p < .001$; ** $p < .01$

Aus Abbildung 3-9 zeigt, dass von der Aufgaben-Wichtigkeit die grösste Wirkung ($\beta = .55, p < .001$) auf Informationsnutzung ausgeht. Der Einfluss der anfänglichen Brauchbarkeits-Lücke ($\beta = .26, p < .01$) ist deutlich schwächer, aber immer noch signifikant.

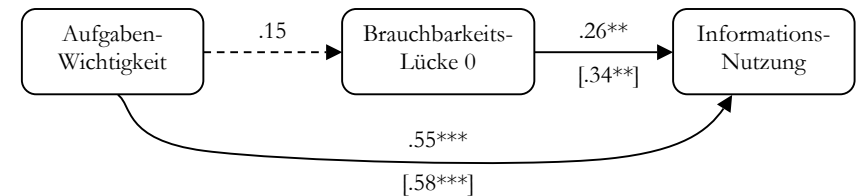


Abbildung 3-9. Direkter und indirekter Zusammenhang zwischen Aufgaben-Wichtigkeit, anfänglicher Brauchbarkeits-Lücke und Informationsnutzung (β -Gewichte; *** $p < .001$; ** $p < .01$; [] = β -Gewicht ohne Berücksichtigung des alternativen Pfades)

Vergleich von Sicherheitsdefizit und Brauchbarkeits-Lücke

Werden alle in den beiden letzten Abschnitten betrachteten Variablen in der Regressionsgleichung berücksichtigt, dann erweisen sich nur die Aufgaben-Wichtigkeit und das Sicherheitsdefizit als signifikante Vorhersagevariablen; $F(2, 66) = 27.3, p < .001$ (Tabelle 3-23).

Tabelle 3-23. Prädiktoren zur Vorhersage der Verarbeitungsschritte

	R^2	β	r
Aufgaben-Wichtigkeit (Treatment)	.45	.51***	.55
Sicherheitsdefizit nach Vorinformation		.34***	.41
Relevanz für Firma			.24
Anspruchsniveau			.23
Gewünschte Urteilssicherheit			.23
Aktuelle Urteilssicherheit nach Vorinform.			.23
Persönliche Relevanz			.17
Brauchbarkeits-Lücke nach Vorinformation			.09
Need for Cognition			.07
Differenzwahrnehmung nach Vorinformation			-.06
Brauchbarkeitseinschätzung nach Vorinform.			.03
Selbstwirksamkeitserwartung			.00

*** $p < .001$

3.3.3 Zusammenhänge zwischen den Variablen

3.3.3.1 Anspruchsniveau

Bei der Analyse aller 70 Entscheidungen fanden, wie Tabelle 3-24 zeigt, die Variablen *Selbstwirksamkeitserwartung* ($\beta = .36, p < .01$) und *persönliche Relevanz* ($\beta = .27, p < .05$) Eingang in der Regressionsgleichung; $F(1, 66) = 13.8, p < .001$. Die Varianzaufklärung machte 29% aus.

Tabelle 3-24. Prädiktoren zur Vorhersage des Anspruchsniveaus

	R ²	β	r
Selbstwirksamkeitserwartung	.29	.36**	.36
Persönliche Relevanz		.27*	.28
Aufgaben-Wichtigkeit (Treatment)			.06
Relevanz für Firma			.13
Need for Cognition			.17

** $p < .01$, * $p < .05$

Die Regressionsanalyse mit der Kriteriumsvariablen *persönliche Relevanz* ergab die signifikanten Vorhersagevariablen *Relevanz für Firma* ($\beta = .64, p < .001$) und *Selbstwirksamkeitserwartung* ($\beta = .28, p < .01$); $F(1, 68) = 65.3, p < .001$. Die Varianzaufklärung betrug 57% (Tabelle 3-25).

Tabelle 3-25. Prädiktoren zur Vorhersage der persönlichen Relevanz

	R ²	β	r
Relevanz für Firma	.57	.64***	.69
Selbstwirksamkeitserwartung		.28**	.38
Aufgaben-Wichtigkeit (Treatment)			-.17
Need for Cognition			.12

*** $p < .001$, ** $p < .01$

In der Regressionsanalyse mit der Kriteriumsvariablen *Relevanz für Firma* fand nur der Prädiktor *Aufgaben-Wichtigkeit* ($\beta = .60, p < .001$) Eingang in die Gleichung; $F(1, 68) = 38.4, p < .001$. Die Varianzaufklärung machte 36% aus (Tabelle 3-26).

Tabelle 3-26. Prädiktoren zur Vorhersage der Relevanz für Firma

	R ²	β	r
Aufgaben-Wichtigkeit (Treatment)	.36	.60***	.60
Need for Cognition			.00
Selbstwirksamkeitserwartung			.18

*** $p < .001$

Aus der Regressionsanalyse mit der Kriteriumsvariablen *Selbstwirksamkeitserwartung* und den Prädiktorvariablen *Aufgaben-Wichtigkeit*, *Need for Cognition* resultierte *Need for Cognition* ($\beta = .29, p < .05$) als einzige Vorhersagevariable; $F(1, 68) = 6.0, p < .05$.

3.3.3.2 Gewünschte Urteilssicherheit

Tabelle 3-27 zeigt, dass die Variablen *Anspruchsniveau* ($\beta = .45, p < .001$) und *persönliche Relevanz* ($\beta = .34, p < .01$) die angestrebte Urteilssicherheit vorher sagten; $F(1, 66) = 27.5, p < .001$. Sie erklärte 45% der Varianz der gewünschten Urteilssicherheit.

Tabelle 3-27. Prädiktoren zur Vorhersage der gewünschten Urteilssicherheit

	R ²	β	r
Anspruchsniveau	.45	.45***	.48
Persönliche Relevanz		.34**	.38
Aufgaben-Wichtigkeit (Treatment)			-.07
Relevanz für Firma			-.05
Need for Cognition			-.12
Selbstwirksamkeitserwartung			-.06

*** $p < .001$, ** $p < .01$

3.3.3.3 Brauchbarkeitseinschätzung

Bei der Brauchbarkeitseinschätzung und der aktuellen Urteilssicherheit umfassten die Analysen 218 Verarbeitungsschritte und es wurden nur Variablen in die Regressionsgleichungen aufgenommen, welche die Varianzaufklärung um mehr als 1% verbesserten. Es wurde zunächst geprüft, welche Prädiktorvariablen die Brauchbarkeitseinschätzung sowie die aktuelle Urteilssicherheit

vorhersagen und anschliessend wurden bei den signifikanten Prädiktorvariablen deren Vorhersagevariablen bestimmt.

Die Variable *Verarbeitungsschritte* stellte sich als beste Prädiktorvariable der Brauchbarkeitseinschätzung heraus (Tabelle 3-28, Modell 1). Sie erklärte 40% der Varianz; $F(1, 216) = 145.9, p < .001$. Als zweitstärkster Prädiktor erschien die *Selbstwirksamkeitserwartung* (Modell 2), welche die Varianzaufklärung um 3% verbesserte; $F_{change}(1, 215) = 10.4, p < .01$. In einem dritten Schritt tauchte die *Aufgaben-Wichtigkeit* in der Regressionsgleichung auf (Modell 3) und erhöhte die Vorhersage um weitere 3%; $F_{change}(1, 214) = 11.0, p < .01$. Als letzte Variable erschien die *Differenzwahrnehmung* in der Regressionsgleichung (Modell 4) und verbesserte die Vorhersage um zusätzliche 1.5%; $F_{change}(1, 213) = 6.0, p < .05$.

Tabelle 3-28. Prädiktoren zur Vorhersage der Brauchbarkeitseinschätzung

	R ²	ΔR ²	β	r
Modell 1				
Verarbeitungsschritte	.40		.64***	.64
Modell 2				
Verarbeitungsschritte	.43	.03**	.63***	.64
Selbstwirksamkeits-Erwartung			.17***	.22
Modell 3				
Verarbeitungsschritte	.46	.03**	.66***	.66
Selbstwirksamkeits-Erwartung			.19***	.25
Aufgaben-Wichtigkeit			-.17**	-.22
Modell 4				
Verarbeitungsschritte	.47	.15	.62***	.62
Selbstwirk-Erwartung			.19***	.25
Aufgaben-Wichtigkeit			-.16**	-.20
Differenzwahrnehmung			.13*	.17
Need for Cognition				-.09
Pers. Relevanz				.02
Relevanz Firma				.09
Aufgabenstellung (IT oder P)				.06

*** $p < .001$, ** $p < .01$, * $p < .05$

3.3.3.4 Aktuelle Urteilssicherheit

Die *Brauchbarkeitseinschätzung* war der beste Prädiktor der Urteilssicherheit und erklärte 61% der Varianz; $F(1, 216) = 339.6, p < .001$ (Tabelle 3-29, Modell 1).

Die *Differenzwahrnehmung* belegte den zweiten Platz und erhöhte die Varianzaufklärung um 6%; $F_{change}(1, 215) = 40.4, p < .001$ (Tabelle 3-29, Modell 2). Als dritte Variable tauchte die *Selbstwirksamkeitserwartung* in der Regressionsgleichung auf und erhöhte die bereits aufgeklärte Varianz um weitere 2%; $F_{change}(1, 214) = 15.7, p < .001$ (Tabelle 3-29, Modell 3).

Tabelle 3-29. Prädiktoren zur Vorhersage der aktuellen Urteilssicherheit

	R ²	ΔR ²	β	r
Modell 1				
Brauchbarkeitseinschätzung	.61		.78***	.78
Modell 2				
Brauchbarkeitseinschätzung	.67	.06***	.70***	.75
Differenzwahrnehmung			.26***	.40
Modell 3				
Brauchbarkeitseinschätzung	.70	.02***	.73***	.77
Differenzwahrnehmung			.25***	.40
Selbstwirksamkeits-Erwartung			-.15***	-.26
Need for Cognition				.01
Relevanz Firma				-.07
Pers. Relevanz				.07
Aufgaben-Wichtigkeit				-.08
Aufgabenstellung (IT oder P)				.05
Verarbeitungsschritte				.15

*** $p < .001$

3.3.3.5 Differenzwahrnehmung

Da keine Variablen erhoben wurden, um die Selbstwirksamkeitserwartung vorherzusagen, wurde nur noch die Differenzwahrnehmung einer weiteren Analyse unterzogen.

Tabelle 3-30 zeigt die Ergebnisse: Die Variable *Verarbeitungsschritte* erwies sich als bester Prädiktor (Modell 1). Sie erklärte 9% der Varianz; $F(1, 216) =$

22.7, $p < .001$ Als zweitstärkste Vorhersagevariable erschien die *Aufgabenstellung* in der Gleichung (Modell 2), welche die Aufklärung um 1% erhöhte; $F_{change}(1, 215) = 4.1, p < .05$.

Tabelle 3-30. Prädiktoren zur Vorhersage der Differenzwahrnehmung

	R^2	ΔR^2	β	r
Modell 1				
Verarbeitungsschritte	.11		.31***	.31
Modell 2				
Verarbeitungsschritte	.12	.01*	.31***	.31
Aufgabenstellung (IT oder P)			-.13*	-.14
Aufgaben-Wichtigkeit				-.12
Need for Cognition				-.01
Selbstwirksamkeits-Erwartung				0
Pers. Relevanz				.05
Relevanz Firma				.09

*** $p < .001, * p < .05$

3.3.3.6 Aufgabenstellung

Bei der IT-Entscheidung war die Firma BB nach dem ersten Bericht und die Firma CC nach dem zweiten und dritten Bericht als beste Wahl zu erkennen (vgl. Abschnitt 3.2.2.1). Wie Tabelle 3-31 zeigt, bevorzugten die Versuchspersonen sowohl nach dem ersten ($\chi^2(2) = 10.1, p < .01$) als auch nach dem zweiten ($\chi^2(1) = 14.4, p < .001$) und dritten Verarbeitungsschritt ($\chi^2(1) = 8.3, p < .01$) jeweils die Firma mit der stärksten argumentativen Unterstützung.

Tabelle 3-31. Präferenzen nach den einzelnen Verarbeitungsschritten bei der IT-Aufgabe

	Vorinformation	Schritt 1	Schritt 2	Schritt 3
Firma AA	21	8	0	1
Firma BB	8	20*	3	0
Firma CC	6	6	22*	11*

* richtige Lösung

In der P-Entscheidung wurde in jedem Interview-Bericht bei jedem Kandidaten ein neues Kriterium genannt. Objektiv gesehen, blieben die Alternativen in der Summe der zugesprochenen Charaktermerkmale nach jedem Bericht gleich gut (vgl. Abschnitt 3.2.2.2). Die Präferenzen der Versuchspersonen reflektierten die Tatsache, dass kein Bericht einen der Kandidaten bevorzugte (Tabelle 3-32). Weder bei der Entscheidung ($\chi^2(2) = .2$) noch nach dem ersten ($\chi^2(2) = 3.5$), zweiten ($\chi^2(2) = 1.6$) und dritten Verarbeitungsschritt ($\chi^2(1) = 3.6$) wurde ein Bewerber überzufällig häufig favorisiert.

Tabelle 3-32. Präferenzen nach den einzelnen Verarbeitungsschritten und beim endgültigen Urteil bei der P-Aufgabe

	Vorinformation	Schritt 1	Schritt 2	Schritt 3	Entscheid
Kandidat A	13	7	12	8	13
Kandidat B	5	16	8	5	11
Kandidat C	17	12	7	2	11

3.3.3.6.1 Brauchbarkeitseinschätzung

Bevor die Wirkung des unterschiedlichen Differenzierungsgrades der Optionen auf die Brauchbarkeitseinschätzung analysiert wurde, erfolgte eine Prüfung auf den Einfluss der Aufgaben-Abfolge bzw. der Thematik (Informatik bzw. Personal) auf die Brauchbarkeitseinschätzung. Da die Vorinformation im Gegensatz zu den später verteilten Informations-Berichten bei beiden Aufgaben sehr ähnlich war, boten sich die Daten aus dem Fragebogen 0 für diese Analyse an. Die Brauchbarkeitseinschätzung war nach der Vorinformation bei der IT- ($M = 5.2$) und bei der P-Entscheidung ($M = 5.2$) praktisch gleich; $\bar{\kappa} = -.34$. Somit konnte ein Positions- und Thematieffekt weitgehend ausgeschlossen werden.

Die Analyse der Verarbeitungsschritte 1 bis 3 in Abhängigkeit von der Informationsnutzung ergab ein unterschiedliches Bild: Bei den Ein-Schritt-Entscheidungen unterschieden sich die Werte nach dem ersten Schritt bei den IT-Entscheidungen ($M = 7.3$) nicht signifikant von derjenigen der P-Entscheidungen ($M = 7.4$); $\bar{\kappa} = -.21$ (Tabelle 3-33). Tabelle 3-33 zeigt ein ähnliches Resultat bei den Zwei-Schritt-Entscheidungen nach dem ersten (IT: $M = 6.3$ versus P: $M = 6.6$; $\bar{\kappa} = -1.27$) und nach dem zweiten Verarbeitungsschritt (IT: $M = 8.3$ versus P: $M = 8.3$; $\bar{\kappa} = -.87$). Anders war es bei den Drei-Schritt-Entscheidungen. Hier war der Unterschied nur nach dem ersten Ver-

arbeitungsschritt (IT: $M = 6.2$ versus P: $M = 6.5$; $\xi = -.37$) nicht signifikant. Nach dem zweiten Schritt war die wahrgenommene Brauchbarkeit bei der IT-Aufgabe ($M = 7.5$) signifikant höher als bei der P-Aufgabe ($M = 6.7$); $\xi = -2.02$, $p < .05$. Nach dem dritten Schritt war dieser Unterschied noch deutlicher ($M = 8.9$ versus $M = 7.9$); $\xi = -3.3$, $p < .01$ (Tabelle 3-33).

Tabelle 3-33. Mittelwerte der Brauchbarkeitseinschätzung in Abhängigkeit von der Aufgabenstellung bei den Ein- bis Drei-Schritt-Entscheidungen

	Vorinformation	Schritt 1	Schritt 2	Schritt 3
Ein-Schritt-Ent. IT	5.8	7.3		
Ein-Schritt-Ent. P	5.1	7.4		
Zwei-Schritt-Ent. IT	5.1	6.3	8.3	
Zwei-Schritt-Ent. P	5.3	6.6	8.3	
Drei-Schritt-Ent. IT	5.2	6.2	7.5	8.9
Drei-Schritt-Ent. P	5.3	6.5	6.7	7.9

3.3.3.6.2 Aktuelle Urteilssicherheit

Ein Positionseffekt oder ein Effekt der Thematik auf die aktuelle Urteilssicherheit konnte praktisch ausgeschlossen werden, da nach der Vorinformation bei beiden Aufgaben beinahe gleich Werte angegeben wurden (IT: $M = 51.6\%$ versus P: $M = 52.1\%$; $\xi = -.08$).

Bei den Ein-Schritt-Entscheidungen unterschieden sich nach dem ersten Verarbeitungsschritt die Werte der Urteilssicherheit (IT: $M = 65.4\%$ versus P: $M = 68.8\%$) kaum voneinander; $\xi = -.64$ (Tabelle 3-34).

Tabelle 3-34. Mittelwerte der aktuellen Urteilssicherheit in Abhängigkeit von der Aufgabenstellung bei den Ein- bis Drei-Schritt-Entscheidungen

	Vorinformation	Schritt 1	Schritt 2	Schritt 3
Ein-Schritt-Ent. IT	52.9	65.4		
Ein-Schritt-Ent. P	58.4	68.8		
Zwei-Schritt-Ent. IT	52.9	65.4	90.4	
Zwei-Schritt-Ent. P	58.4	68.8	90.8	
Drei-Schritt-Ent. IT	45.7	54.8	76.0	96.0
Drei-Schritt-Ent. P	47.4	60.2	67.1	84.6

Bei den Zwei-Schritt-Entscheidungen lagen die Werte nach dem ersten Verarbeitungsschritt sehr nahe beieinander (IT: $M = 65.4\%$ versus P: $M = 68.8\%$; $\xi = -.83$). Nach dem zweiten und letzten Verarbeitungsschritt war die aktuelle Urteilssicherheit bei der IT-Aufgabe ($M = 90.4\%$) praktisch gleich hoch wie bei der P-Aufgabe ($M = 90.8\%$); $\xi = -.75$ (Tabelle 3-34 und Abbildung 3-10).

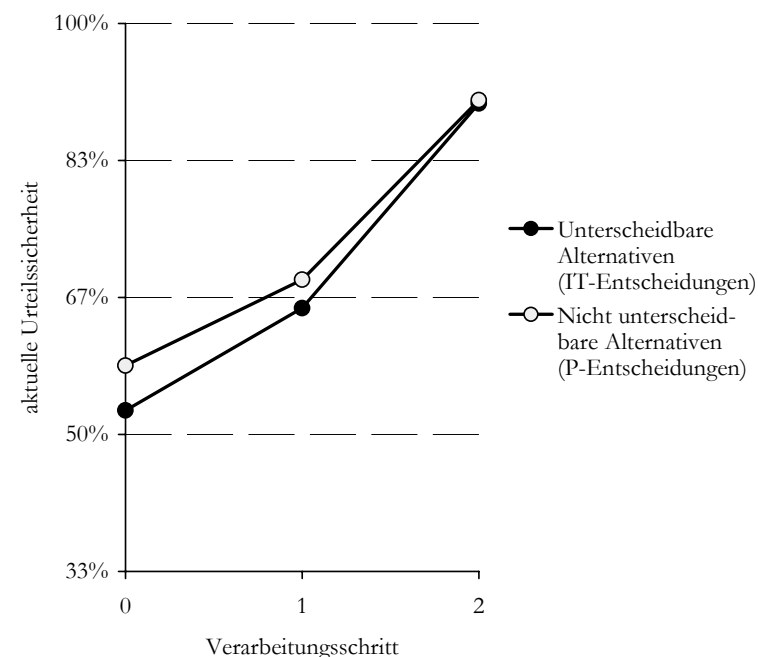


Abbildung 3-10. Wirkung des unterschiedlichen Differenzierungsgrades auf Mittelwerte der aktuellen Urteilssicherheit bei den Zwei-Schritt-Entscheidungen

Bei den Drei-Schritt-Entscheidungen waren die Unterschiede nach dem ersten Verarbeitungsschritt nicht signifikant (IT: $M = 54.8\%$ versus P: $M = 60.2\%$; $\xi = -.72$). Ein anderes Bild zeigte sich nach den Verarbeitungsschritten 2 und 3: Nach dem zweiten Schritt war die aktuelle Urteilssicherheit bei der IT-Aufgabe ($M = 76.0\%$) signifikant höher als bei der P-Aufgabe ($M =$

67.1%); $\bar{\kappa} = -1.74$, $p < .05$ (einseitig). Nach dem dritten Schritt war dieser Unterschied noch ausgeprägter ($M = 96.0\%$ versus $M = 84.6\%$); $\bar{\kappa} = -2.86$, $p < .01$ (Tabelle 3-34 und Abbildung 3-11).

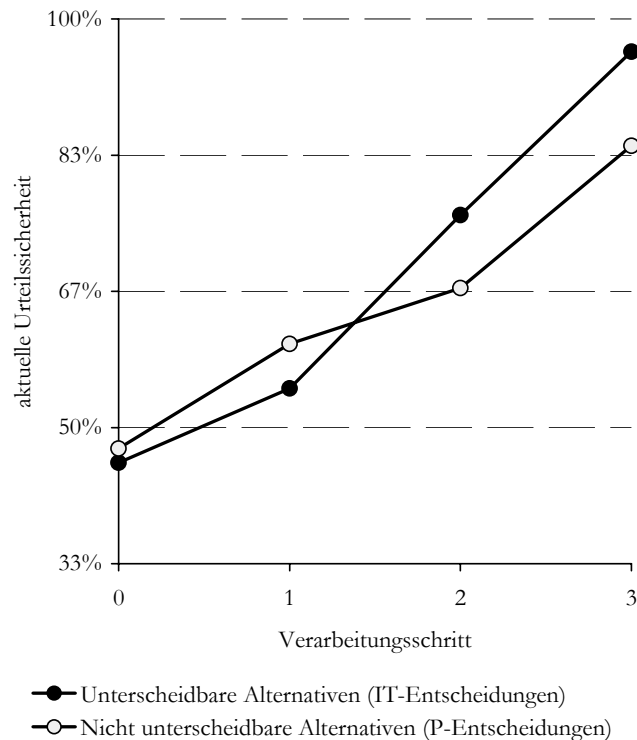


Abbildung 3-11. Wirkung des unterschiedlichen Differenzierungsgrades auf Mittelwerte der aktuellen Urteilssicherheit bei den Drei-Schritt-Entscheidungen

3.4 Diskussion

3.4.1 Stopp-Mechanismen im Test

3.4.1.1 Reaktion der Führungsgrösse (Funktionstest 1)

Die gewünschte Urteilssicherheit war bei den wichtigeren Entscheidungen signifikant höher als bei den weniger wichtigen Entscheidungen. In den Within-Subjects-Vergleichen waren die Unterschiede zwei mal knapp nicht signifikant. Die Aufgaben-Wichtigkeit leistete mit einer Varianzaufklärung von 6% nur einen geringen Beitrag zu Vorhersage der gewünschten Urteilssicherheit. Dieser Befund spricht eher für die Hypothese 1a (Folgeregelung) als 1b (Festwertregelung). Damit wird die Vorstellung unterstützt, es gäbe einen Assessor_{soll}, der die Höhe der gewünschten Urteilssicherheit bestimmt und der Stopp-Mechanismus Urteilssicherheit sei – sofern er die weiteren Tests besteht – ein Folgeregler.

Beim Anspruchsniveau ergab sich ein ähnliches Bild: Bei den wichtigeren Entscheidungen wurde eine signifikant höhere Brauchbarkeit der Lösung gewünscht als bei den weniger wichtigen Entscheidungen. In den Within-Subjects-Vergleichen war die Differenz einmal auf dem 5%-Niveau (Versuchsbedingung „it/P“) und einmal deutlich nicht signifikant (Versuchsbedingung „IT/p“). Die Aufgabe-Wichtigkeit erklärte 6% der Varianz des Anspruchsniveaus. Dieses Ergebnis widerlegte Hypothese 1b, das Treatment hat keinen Einfluss auf das Anspruchsniveau. Dies ist ein Hinweis darauf, dass ein Assessor_{soll} existiert, der ein generelles Anspruchsniveau bestimmt. Somit handelt es sich auch beim Stopp-Mechanismus Satisficing – sofern er die weiteren Tests besteht – um einen Folgeregler.

3.4.1.2 Reaktion der Regelgrösse (Funktionstest 2)

Im Verlauf aller Entscheidungsprozesse stiegen die Urteilssicherheit und die Brauchbarkeitseinschätzung kontinuierlich an. Nach dem ersten Arbeitsschritt waren die Ein-Schritt-Entscheider sicherer und von der Brauchbarkeit der bevorzugten Alternative überzeugter als die Mehr-Schritt-Entscheider. Nach dem zweiten Schritt waren die Werte der Zwei-Schritt-Entscheider signifikant höher als diejenigen der Drei-Schritt-Entscheider. Insgesamt fühlten sich die Versuchspersonen, die ein endgültiges Urteil gefällt hatten, sicherer und

schätzten die Brauchbarkeit der bevorzugten Alternative höher ein als die noch unentschlossenen Personen.

Die definitive Urteilssicherheit und die definitive Brauchbarkeitseinschätzung waren bei den IT-Entscheidungen um so höher, je mehr Information bis zur Entscheidung genutzt wurde. Signifikante Unterschiede zwischen den Ein-Schritt- und Drei-Schritt-Entscheidungen wurden festgestellt. Bei den P-Entscheidungen war dies nicht der Fall. Die Ursache dafür dürfte in der Tatsache begründet sein, dass sich die Kandidaten in der Summe der erfüllten Kriterien in keinem der Berichte unterschieden. Im Gegensatz dazu lieferte jeder neue Informations-Bericht in der IT-Entscheidung neue Argumente, mit der sich die Firmen besser differenzieren liessen.

Diesen zweiten Test bestanden beide Stopp-Mechanismen und die Hypothese 2, dass mit zunehmender Informationsnutzung die Ist-Grössen steigen, wurde unterstützt. Damit ist bei beiden Mechanismen der Nachweis für einen Assessor_{ist} und eine Regelgrösse erbracht.

3.4.1.3 Wirkung der Regelabweichung (Funktionstest 3)

Die Erwartung, die Regelabweichungen unterscheiden definitive Urteile von provisorischen, wurde von beiden Stopp-Mechanismen erfüllt: Das Sicherheitsdefizit und die Brauchbarkeits-Lücke erkannten „letzte“ Schritte. Da es um so wahrscheinlicher wird, dass es sich bei einem Verarbeitungsschritt um einen endgültigen handelt, je mehr Schritte ihm vorangegangen sind und die absolvierten Verarbeitungsschritte beide Ist-Werte beeinflussten (Test 2), war das Vorhersagevermögen der Regelabweichungen keine Überraschung. Sowohl das Sicherheitsdefizit als auch die Brauchbarkeits-Lücke erfüllten aber auch die zusätzliche Forderung, dass sie bessere Vorhersagen machen als die Variable Verarbeitungsschritte. Die Regelabweichungen schlugen aber nicht nur die Verarbeitungsschritte im Vorhersagevermögen des Verarbeitungsendes sondern auch die ebenfalls signifikanten Regelgrössen Urteilssicherheit und Brauchbarkeitseinschätzung. Somit unterstützten diese Ergebnisse die Hypothesen 3 und sie lassen den Schluss zu, dass sich die Versuchspersonen entschieden haben, weil sie „sicher genug“ geworden waren bzw. weil sie eine Lösung als „gut genug“ bewertet hatten.

Zusätzlich wurde geprüft, ob die anfänglichen Regelabweichungen die später absolvierten Verarbeitungsschritte vorhersagen konnten wie sie Bohner et al. (1998) für das Sicherheitsdefizit vorhersagen konnte (Abschnitt 1.4). Dies war der Fall: Die Versuchspersonen absolvierten um so mehr Verarbei-

tungsschritte, je grösser die anfängliche Diskrepanz zwischen der gewünschten und der aktuellen Urteilssicherheit bzw. je ausgeprägter die anfängliche Differenz zwischen dem Anspruchsniveau und der Brauchbarkeitseinschätzung war. Dieser Effekt zeigte sich auch, wenn das Treatment „Aufgaben-Wichtigkeit“ in die Analyse mit einbezogen wurde. Sowohl das Sicherheitsdefizit nach der Vorinformation als auch die Brauchbarkeits-Lücke nach der Vorinformation hatten einen Einfluss auf die Informationsnutzung, der über die Wirkung des Treatments hinaus ging.

3.4.1.4 Vergleich beider Stopp-Mechanismen

Da beide Mechanismen die Anforderungen an einen funktionierenden Regelkreis erfüllten, wurde im Sinne von Popper (1973) geprüft, welcher der beiden Mechanismen den höheren Erklärungserfolg aufweist und damit für die Beendigung der Entscheidungsprozesse eher verantwortlich war. Keinen Einfluss auf diesen Vergleich hatte Funktionstest 1, da er allein die Frage prüfte, ob es sich bei den Stopp-Mechanismen um einen Festwert- oder einen Folgeregler handelte.

Funktionstest 2 wurde von beiden Mechanismen uneingeschränkt bestanden und es waren keine Unterschiede feststellbar. Die genutzte Information wirkte sowohl auf die aktuelle Urteilssicherheit als auch auf die Brauchbarkeitseinschätzung sehr deutlich. Bei der IT-Entscheidung – nur hier lieferte jeder neue Bericht zusätzliche alternativendifferenzierende Argumente – erklärten die Verarbeitungsschritte bei beiden Grössen 46% der Varianz (Tabelle 3-35).

Tabelle 3-35. Vergleich der Stopp-Mechanismen Urteilssicherheit und Satisficing

	Urteilssicherheit	Brauchbarkeitseinschätzung
Reaktion der Regelgrösse (Test 2)	.46 ¹⁾	.46 ¹⁾
Erkennen „letzter“ Schritte (Test 3)	.46	.39

Zahlenangaben entsprechen der Varianzaufklärung; ¹⁾ IT-Entscheidungen

Im Funktionstest 3 erkannte das Sicherheitsdefizit „letzte“ Schritte in 83.1% der Fälle richtig, bei der Brauchbarkeits-Lücke waren es 80.3% der Fälle. Gemessen an der Varianzaufklärung lag der Wert für das Sicherheits-

defizit bei 46% und derjenige der Brauchbarkeits-Lücke bei 39% (Tabelle 3-35). Somit ergab sich hier ein Vorteil für den Stopp-Mechanismus Urteils-sicherheit.

Ein weiterer Vorteil für den Stopp-Mechanismus Urteils-sicherheit machte sich bei der Betrachtung des Vorhersagevermögens der Informationsnutzung bemerkbar. Das Sicherheitsdefizit nach der Vorinformation erklärte 22% der Varianz der Verarbeitungsschritte die Brauchbarkeits-Lücke nach der Vorinformation nur 12%. Bei Berücksichtigung aller möglicher Vorhersagevariablen erreichte neben der Aufgaben-Wichtigkeit nur das Sicherheitsdefizit die Signifikanzschwelle.

Bei allen Gegenüberstellungen schnitt der Stopp-Mechanismus Urteils-sicherheit gleich gut oder besser ab als der Stopp-Mechanismus Satisficing. Somit liefert der Stopp-Mechanismus Urteils-sicherheit einen höheren Erklärungserfolg. Die eingangs gestellte Frage, wann wissen wir genug, lässt sich deshalb für diese Untersuchung wie folgt beantworten: Entscheidungsprozesse werden eher dann beendet, wenn Entscheider „sicher genug“ sind, als wenn eine Lösung „gut genug“ ist.

3.4.2 Zusammenhänge zwischen den Variablen

3.4.2.1 Anspruchsniveau

Da eher von einer Wirkung des Anspruchsniveaus auf die gewünschte Urteils-sicherheit als von einem umgekehrten Effekt auszugehen ist (vgl. Kapitel 2.5), werden zunächst die Zusammenhänge, die zur Bildung des Anspruchsniveaus führten, betrachtet: Die Versuchspersonen wünschten eine umso bessere Lösung, je relevanter sie die Entscheidungsaufgabe für sich selber wahrnahmen und je höher sie ihre Fähigkeit einstuften, solche Entscheidungen zu fällen. Die persönliche Relevanz war um so grösser, je bedeutsamer die Studierenden die Entscheidungsaufgabe im Hinblick auf die Firma begriffen. Die Höhe dieser firmenbezogenen Relevanz war bei der sehr wichtigen Aufgabe? grösser als bei der weniger wichtigen Aufgabe. Die Selbstwirksamkeitserwartung war um so grösser, je höher die Versuchspersonen ihr Engagement und ihre Freude beim Denken einstuften (Abbildung 3-12).

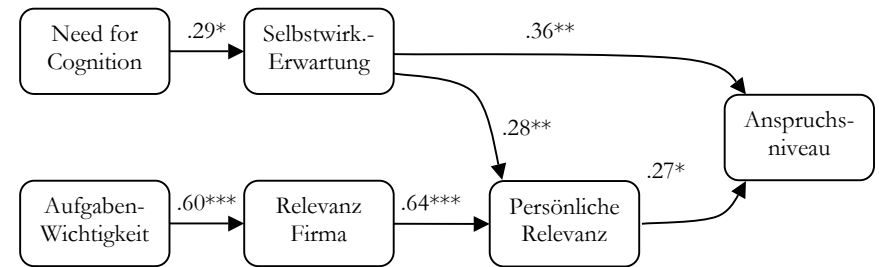


Abbildung 3-12. Zusammenhänge zwischen Aufgaben-Wichtigkeit, Relevanz für die Firma, persönliche Relevanz bzw. Need for Cognition, Selbstwirksamkeitserwartung und Anspruchsniveau (β -Gewichte; *** $p < .001$, ** $p < .01$, * $p < .05$)

3.4.2.2 Gewünschte Urteils-sicherheit

Das Wirkungsgefüge der gewünschten Urteils-sicherheit in Abbildung 3-13 zeigt, dass die Versuchspersonen bei ihrem endgültigen Urteil um so sicherer sein wollten, je relevanter sie die Entscheidung für sich selber bewerteten und je besser die Lösung sein sollte. Die persönliche Relevanz war um so höher, je bedeutungsvoller die Entscheidung für die Firma bewertet wurde und diese firmenbezogene Relevanz war bei der sehr wichtigen Aufgabe höher als bei der weniger wichtigen Aufgabe. Die Lösung sollte um so besser sein, je bedeutsamer die Entscheidung und je höher das eigene Lösungsvermögen bewertet wurden. Die Selbstwirksamkeitserwartung war um so grösser, je mehr Engagement und Freude beim Denken bekundet wurde.

Die Tatsache, dass die individuellen Faktoren Need for Cognition und Selbstwirksamkeitserwartung eine Wirkung auf die gewünschte Urteils-sicherheit zu entfalten vermochten, bestätigt die Erwartung, persönliche Faktoren würden die Führungsgrösse im Stopp-Mechanismus Urteils-sicherheit beeinflussen (Eagly & Chaiken, 1993; Maheswaran & Chaiken, 1991).

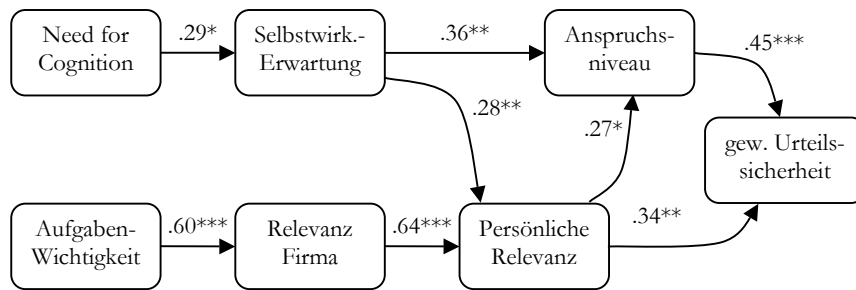


Abbildung 3-13. Zusammenhänge zwischen Aufgaben-Wichtigkeit, Relevanz für die Firma, persönliche Relevanz bzw. Need for Cognition, Selbstwirksamkeitserwartung, Anspruchsniveau und gewünschter Urteils-sicherheit (β -Gewichte; *** $p < .001$, ** $p < .01$, * $p < .05$)

3.4.2.3 Brauchbarkeitseinschätzung

Die Versuchspersonen beurteilten die Brauchbarkeit der bevorzugten Alternative um so höher, je mehr Information sie verarbeitet hatten, je grösser sie die Differenz zur zweitbesten Alternative wahrnahmen und je höher sie das eigene Lösungsvermögen bewerteten. Und sie waren bei der wichtigeren Aufgabe in ihren Einschätzungen vorsichtiger als bei der weniger wichtigen Aufgabe (Abbildung 3-14). Mit dem positiven Zusammenhang zwischen den Verarbeitungsschritten und der Brauchbarkeitseinschätzung und dem negativen Zusammenhang zwischen Aufgaben-Wichtigkeit und Brauchbarkeitseinschätzung wurden die Erwartungen an den *Brauchbarkeits-Test* erfüllt (Hypothese B). Verglichen mit allen anderen Prädiktoren ging von den *Verarbeitungsschritten* die grösste Wirkung auf die Brauchbarkeitseinschätzung aus. Diese Tatsache ist verständlich, lieferte doch jeder zusätzliche Informations-Bericht in beiden Entscheidungsaufgaben Argumente, die eine Firmen bzw. alle Kandidaten bezüglich der relevanten Kriterien besser ausschauen liessen (vgl. Abschnitt 3.2.2.1 und 3.2.2.2). Ob einfach Argumente registriert und gezählt wurden oder ob im Sinne des vorgeschlagenen Brauchbarkeits-Tests die Alternativen im Hinblick auf aktivierte Image-Elemente geprüft wurden, lässt sich aufgrund dieses Zusammenhangs aber nicht sagen.

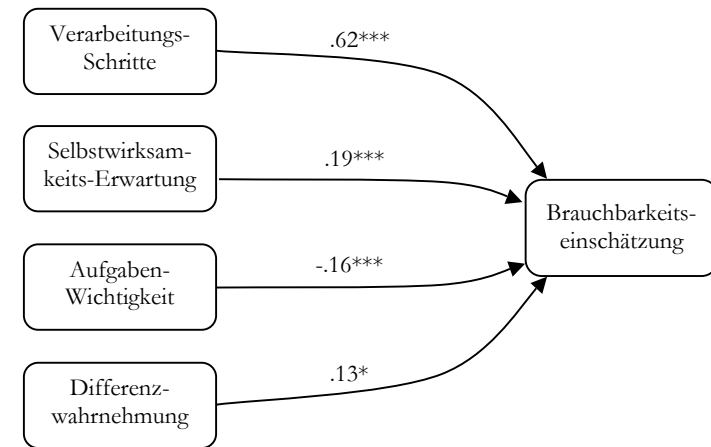


Abbildung 3-14. Zusammenhänge zwischen Verarbeitungsschritten, Selbstwirksamkeits-Erwartung, Aufgaben-Wichtigkeit, Differenzwahrnehmung und Brauchbarkeitseinschätzung (β -Gewichte; *** $p < .001$)

Der direkte negative Zusammenhang zwischen *Aufgaben-Wichtigkeit* und Brauchbarkeitseinschätzung kann dahin interpretiert werden, dass die Versuchspersonen bei der wichtigeren Entscheidung mehr Image-Elemente aktiviert und/oder höhere Anforderungen innerhalb der Image-Elemente gestellt hatten und deshalb tiefere Brauchbarkeitseinschätzungen abgaben. Mit einem β -Gewichte von $-.16$ ist der Zusammenhang aber so schwach, dass die daraus abgeleitete Unterstützung für den postulierten Brauchbarkeits-Test nur gering ist. Zudem lässt sich der Zusammenhang auch anders erklären: Die Versuchspersonen könnten die einzelnen Argumente bei der wichtigeren Aufgabe vorsichtiger aufgenommen haben als bei der weniger wichtigen Aufgabe.

Wie lässt sich die Korrelation zwischen *Differenzeinschätzung* und Brauchbarkeit interpretieren? Dieser Zusammenhang ist ein Anzeichen dafür, dass bei der Brauchbarkeitseinschätzung auch noch ein vergleichender Mechanismus mitspielt. Dabei wirken möglicherweise die nicht gewählten Alternativen als Referenz oder Anker (Tversky & Kahneman, 1974). Weil die Alternativen sich bei der IT-Entscheidung stärker voneinander unterscheiden als bei der P-Entscheidung, konnte die Wirkung des vergleichende Mechanismus auch

daran geprüft werden, ob die Aufgabenstellung einen Einfluss auf die Brauchbarkeitseinschätzung hatte. Falls der vergleichende Mechanismus beteiligt war, sollten die Versuchspersonen die bevorzugte Firma qualitativ besser einschätzen als den bevorzugten Kandidaten. Bei den Ein- und Zweischritt-Entscheidungen wurden keine signifikanten Unterschiede in den Brauchbarkeitsangaben festgestellt. Bei den Drei-Schritt-Entscheidungen unterschieden sich dagegen die Brauchbarkeitseinschätzungen nach dem zweiten und dritten Verarbeitungsschritt. Dieses Resultat zeigt, dass der vergleichende Mechanismus nur in den letzten zwei Dritteln der längeren Entscheidungen zum Zug kam. Der vergleichende Mechanismus könnte somit ein Hinweis für eine intensive Elaboration sein.

Der Einfluss der Variablen *Selbstwirksamkeitserwartung* auf die Brauchbarkeitseinschätzung zeigt, dass die Brauchbarkeit der favorisierten Alternative um so höher eingeschätzt wurde, je besser die Versuchspersonen ihre Fähigkeit bewerteten, Entscheidungen zu treffen. Ob die geschehen ist, weil eine spezifische Selbstwirksamkeitserwartung die Leistung erhöhte (Bandura, 1997) oder ob einfach die selbstsichereren Personen in der Einschätzung etwas mutiger waren als die weniger selbstsicheren, lässt sich aufgrund der vorliegenden Daten nicht beurteilen.

3.4.2.4 Aktuelle Urteilssicherheit

Die Versuchspersonen waren in ihren Urteilen um so sicherer, je brauchbarer sie die bevorzugte Lösung einschätzten, je grösser sie die Differenz zur zweitbesten Alternative wahrnahmen und je bescheidener sie das eigene Lösungsvermögen bewerteten (Abbildung 3-15). Der sich aus dem *Contingency Model for the Selection of Decision Strategies* (Beach & Mitchell, 1978) ergebende positive Zusammenhang zwischen den Verarbeitungsschritten und der aktuellen Urteilssicherheit wurde nicht festgestellt (Hypothese P). Damit widersprach der vorliegende Befund der Vorstellung, die Höhe des analytischen Aufwandes beeinflusse die Urteilssicherheit.

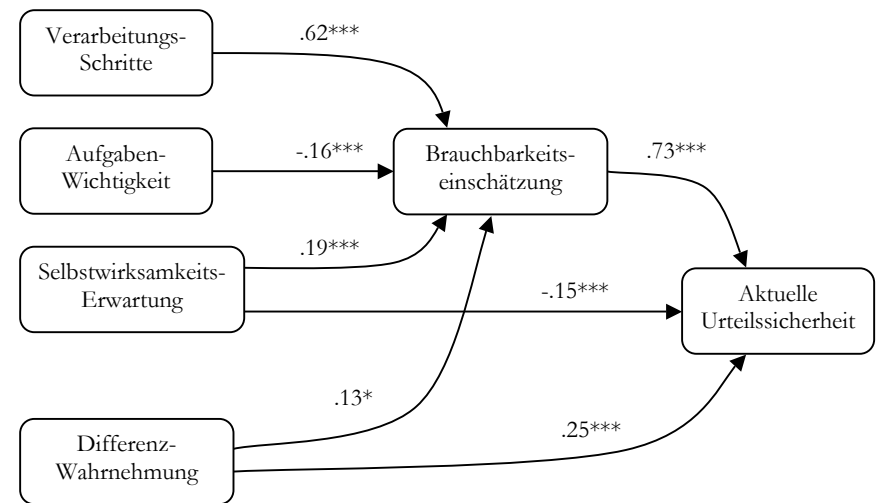


Abbildung 3-15. Zusammenhänge zwischen Verarbeitungsschritten, Aufgaben-Wichtigkeit, Selbstwirksamkeitserwartung, Brauchbarkeitseinschätzung, Differenzwahrnehmung und aktueller Urteilssicherheit (β -Gewichte; *** $p < .001$, * $p < .05$)

Bei der *Support Theory* (Tversky & Koehler, 1994) wurde zunächst geprüft, ob der unterschiedliche Differenzierungsgrad der Alternativen einen Einfluss auf die Urteilssicherheit hatte (Hypothese R₁). Die Ergebnisse waren uneinheitlich: Obschon in der IT-Entscheidung nach jedem Bericht mindestens eine Firma besser war als eine andere und in der P-Entscheidung alle Kandidaten bezüglich der vorgegebenen Kriterien immer gleich gut qualifiziert waren, unterschieden sich die Werte der Urteilssicherheit weder bei den Ein-Schritt-Entscheidungen nach Verarbeitungsschritt 1 noch bei den Zweischritt-Entscheidungen nach Verarbeitungsschritten 1 und 2. Einzig bei den Drei-Schritt-Entscheidungen wurden die Unterschiede nach dem sie nach Verarbeitungsschritt 1 noch nicht signifikant waren nach Verarbeitungsschritten 2 und 3 signifikant. Dieser Befund war ein Indiz dafür, dass die Versuchspersonen zumindest in der zweiten Hälfte der Drei-Schritt-Entscheidungen bei der Urteilssicherheits-Bildung die relative argumentative Unterstützung berücksichtigt haben. Ein starke Unterstützung für die Support Theory war dieses Indiz aber nicht, da genau der gleiche Effekt schon bei der

Brauchbarkeitseinschätzung zu beobachten war und von dieser Variablen der grösste Einfluss auf die Urteilssicherheit ausging. Klärung ergab sich aus der Prüfung der Hypothese R₂. Diese forderte, es bestehe ein sehr enger positiver Zusammenhang zwischen der Differenzwahrnehmung und der Urteilssicherheit und ein schwächerer negativer Zusammenhang zwischen der Brauchbarkeitseinschätzung und der Urteilssicherheit. Die Ergebnisse widersprachen dieser Forderung recht deutlich: Zum Einen war der Einfluss der Brauchbarkeit deutlich grösser als derjenige der Differenz und nicht umgekehrt, zum Anderen war der Zusammenhang zwischen Brauchbarkeitseinschätzung und Konfidenz positiv und nicht negativ (Abbildung 3-15). Die Vorstellung von Tversky und Koehler (1994), die aktuelle Urteilssicherheit bilde sich aus den akkumulierten Argumenten für eine bevorzugte Alternative, geteilt durch die Summe aller Argumente für alle Alternativen, wurde somit widerlegt.

Nach dem *Model of Belief Processing* (Curley & Benson, 1994) sollte der unterschiedliche Differenzierungsgrad der Alternativen keinen Einfluss auf die Urteilssicherheit haben (Hypothese A₁). Wie schon bei der Support Theory waren die Ergebnisse uneinheitlich: Einerseits entsprach die ausbleibende Wirkung bei den Ein- und Zwei-Schritt-Entscheidungen den Erwartungen, und andererseits widersprach die Wirkung bei den Drei-Schritt-Entscheidungen nach Verarbeitungsschritten 2 und 3 der Vorhersage. Weiteren Aufschluss über den Erklärungswert des Belief Processing Modells ergab die Prüfung der Zusammenhangshypothese (A₂). Die Erwartung, es bestehe ein enger positiver Zusammenhang zwischen der Brauchbarkeitseinschätzung und der Urteilssicherheit, erfüllte sich voll und ganz: Von der Brauchbarkeitseinschätzung ging die weitaus stärkste Wirkung auf die Urteilssicherheit aus (Abbildung 3-15). Der dominierende Einfluss der Brauchbarkeitseinschätzung ist aus Dicht des Belief Processing Modells sehr verständlich: Die Herausforderung bei beiden Entscheidungsaufgaben bestand darin, sich anhand einer Vielzahl von Argumenten ein Bild über die Qualität der Alternativen zu machen. Weil die kumulative Stärke der Argumente über die Wahl der Alternative bestimmen musste, lag es nahe, dass diese Erkenntnis auch für die Bildung der Urteilssicherheit verwendet wurde. Abbildung 3-15 zeigt, dass von der kumulativen Stärke der Argumente für die bevorzugte Alternative, wie sie in der Brauchbarkeitseinschätzung zum Ausdruck kam, zwar die stärkste, aber nicht die alleinige Wirkung auf die aktuelle Urteilssicherheit ausging. Die Sicherheit, eine richtige Wahl getroffen zu haben, hing auch noch vom wahrgenommenen Abstand zur zweitbesten Alternative ab und sie war um so kleiner, je ausgeprägter die Selbstwirksamkeitserwartung der Versuchspersonen war. Die Brauchbarkeitseinschätzung bzw. das Belief Processing Modell

kann zwar einen grossen Teil, aber nicht alles bei der Urteilssicherheits-Bildung erklären.

Insgesamt sprechen die Ergebnisse dafür, dass die Urteilssicherheit nicht auf einer Bewertung des Entscheidungsprozesses beruht, sondern sich auf gewonnene Erkenntnisse abstützt und dabei zwei Cues verwendet: Brauchbarkeitseinschätzung und Differenzwahrnehmung. Beide Cues werden nicht im Sinne der Wahrscheinlichkeitsrechnung zueinander in Relation gesetzt, sondern sie wirken im direkten Verhältnis zu ihrer Stärke.

3.4.3 Zusammenwirken beider Mechanismen

Sowohl zwischen Anspruchsniveau und gewünschter Urteilssicherheit als auch zwischen Brauchbarkeitseinschätzung und aktueller Urteilssicherheit wurden signifikante Korrelationen festgestellt. Es erscheint daher plausibel, dass bei der Genese eines Stopp-Mechanismus alle nutzbaren Funktionen verwendet wurden, die ein Entscheidungsprozess zur Verfügung stellte. Ein integrierter Mechanismus kommt der Wahrheit damit wahrscheinlich näher als ein Stopp-Mechanismus, der nur auf Urteilssicherheit oder Satisficing beruht. Abbildung 3-16 zeigt ein Entscheidungs-System mit dem integrierten Stopp-Mechanismus. Darin erkennt das Verarbeitungssystem anhand innerer und äusserer Signale, ob eine Entscheidungssituation vorliegt. Ist dies der Fall, so wirkt das äussere Signal Aufgaben-Wichtigkeit auf die Variable Relevanz für Firma. Die Relevanz für Firma wirkt gemeinsam mit der Selbstwirksamkeitserwartung wiederum auf die Variable persönliche Relevanz. Die Selbstwirksamkeitserwartung wird durch die Variable Need for Cognition beeinflusst. Die persönliche Relevanz wirkt gemeinsam mit der Selbstwirksamkeitserwartung über den Brauchbarkeits-Assessor_{soll} auf das Anspruchsniveau. Das Anspruchsniveau und die persönliche Relevanz liefern den Input für den Urteilssicherheits-Assessor_{soll}, der daraus die Führungsgrösse Gewünschte Urteilssicherheit bildet. Das Verarbeitungssystem verarbeitet entscheidungsrelevante Information. Bei dieser Elaboration wirkt die Information – gemessen an den Verarbeitungsschritten – gemeinsam mit den Variablen Selbstwirksamkeitserwartung, Aufgaben-Wichtigkeit und Differenzwahrnehmung auf den Brauchbarkeits-Assessor_{ist}. Die Differenzwahrnehmung wird durch die Variable Aufgabenstellung, hinter der sich ein unterschiedlicher Differenzierungsgrad der Alternativen verbirgt, und Verarbeitungsschritte beeinflusst.

Die *Brauchbarkeitseinschätzung* war um so höher, je mehr Information die Versuchspersonen genutzt hatten, je grösser sie die Differenz zwischen bester und zweitbesten Alternative wahrnahmen und je höher sie das eigene Problemlösungsvermögen bewerteten und sie waren bei der wichtigeren Aufgabe in ihren Einschätzungen vorsichtiger als bei der weniger wichtigen Aufgabe. Der erste und der letzte Zusammenhang entsprach der Vorhersage des Brauchbarkeits-Tests. Der Zusammenhang zwischen Aufgaben-Wichtigkeit und Brauchbarkeitseinschätzung war aber schwach, so dass nur von einer geringen Unterstützung für den postulierten Brauchbarkeits-Test gesprochen werden kann.

Die *aktuelle Urteilssicherheit* war um so höher, je brauchbarer die Versuchspersonen die bevorzugte Lösung einschätzten, je grösser sie die Differenz zur zweitbesten Alternative wahrnahmen und je bescheidener sie das eigene Lösungsvermögen bewerteten. Die Vorstellung, der betriebene analytische Aufwand bestimme die Urteilssicherheit, wurde widerlegt, da kein direkter Zusammenhang zwischen den Verarbeitungsschritten und der Urteilssicherheit festgestellt wurde. Somit fand das Contingency Model for the Selection of Decision Strategies von Beach und Mitchell (1978) keine Unterstützung. Ebenfalls widerlegt wurde die Annahme der Support Theory von Tversky und Koehler (1994), derzufolge sich die aktuelle Urteilssicherheit aus den akkumulierten Argumenten für die bevorzugte Alternative geteilt durch die Summe aller Argumente für alle Alternativen ergebe. Einerseits war der Einfluss der Brauchbarkeit deutlich grösser als derjenige der Differenz statt umgekehrt und andererseits war der Zusammenhang zwischen Brauchbarkeitseinschätzung und Konfidenz positiv statt negativ. Unterstützung fand dagegen die Erwartung des Model of Belief Processing von Curley und Benson (1994), es bestehe ein enger positiver Zusammenhang zwischen der Brauchbarkeitseinschätzung und der Urteilssicherheit. Dies entspricht der Vorstellung, dass bei Entscheidungen der Argumentationsbildung die zentrale Bedeutung zukommt und dass deshalb nicht nur die Wahl, sondern auch die Urteilssicherheit auf der absoluten argumentativen Unterstützung basiert. Insgesamt spricht dieser Befund für eine erkenntnisbezogene und gegen eine prozessbezogene Bildung der Urteilssicherheit. Die dabei verwendeten Cues Brauchbarkeitseinschätzung und Differenzwahrnehmung wirken im direkten Verhältnis zu ihrer Stärke, ohne dass dabei die Regeln der Wahrscheinlichkeitsrechnung berücksichtigt werden.

Diese vorliegenden Ergebnisse basieren auf einer relativ kleinen homogenen Stichprobe. Kritisieren lässt sich insbesondere die Stichprobenauswahl, handelte es sich doch bei den Versuchspersonen, um Studentinnen und Stu-

denten, die in ihrer Berufskarriere noch keine Entscheidungen in der vorgelegten Form fällen mussten. Das vorliegende Experiment warf deshalb die Frage auf, ob ähnliche Ergebnisse gefunden werden, wenn Entscheidungs-Experten diese Aufgaben lösen. In einer zweiten Studie wurden erfahrene Managerinnen und Manager mit denselben Entscheidungsaufgaben konfrontiert.

4 Ein Experiment mit erfahrenen Entscheidern

4.1 Zielsetzung und Hypothesen

Das vorangegangene Experiment wurde unter ähnlichen Bedingungen, aber mit erfahrenen Entscheidern wiederholt. Die Zielsetzung blieb dieselbe: (1) Herausfinden, ob eher der Stopp-Mechanismus Urteilssicherheit oder eher der Stopp-Mechanismus Satisficing für das Ende von Entscheidungsprozessen verantwortlich gemacht werden kann, und (2) herausfinden, ob der Brauchbarkeits-Test die Bildung der Brauchbarkeitseinschätzung beschreibt bzw. ob das Contingency Model von Beach und Mitchell (1978), die Support Theory von Tversky und Koehler (1994) und das Model of Belief Processing von Curley und Benson (1994) die Entstehung der Urteilssicherheit erklären. Die Hypothesen blieben die gleichen wie in Experiment 2:

- (1a) Die Aufgaben-Wichtigkeit hat einen Einfluss auf die Höhe der Führungsgrösse (Folgeregelung).
- (1b) Die Aufgaben-Wichtigkeit hat keinen Einfluss auf die Höhe der Führungsgrösse (Festwertregelung).
- (2) Mit jedem Verarbeitungsschritt steigt die Regelgrösse.
- (3) Regelabweichung erkennt „letzte“ Schritte und ist in der Vorhersage besser als die Summe der durchgeführten Verarbeitungsschritte.
- (B) Es besteht ein positiver Zusammenhang zwischen den Verarbeitungsschritten und der Brauchbarkeitseinschätzung sowie ein negativer Zusammenhang zwischen Aufgaben-Wichtigkeit und Brauchbarkeitseinschätzung.
- (P) Es besteht ein positiver Zusammenhang zwischen den Verarbeitungsschritten und der aktuellen Urteilssicherheit.
- (R) Es besteht ein enger positiver Zusammenhang zwischen der Differenzwahrnehmung und der Urteilssicherheit und ein schwächerer negativer Zusammenhang zwischen der Brauchbarkeitseinschätzung und der Urteilssicherheit.

4.2 Methode

Wie in Experiment 1 mussten nacheinander zwei Entscheidungsaufgaben gelöst werden. Die Argumente lieferten je drei kostenpflichtige Informations-

Berichte und die Versuchspersonen konnten nach der Bearbeitung jedes Berichts wählen, ob sie ein definitives Urteil fällen oder weitere Information beziehen wollten. In zwei Treatment-Bedingungen war entweder die erste Entscheidung sehr wichtig und die zweite Entscheidung weniger wichtig („IT/p“) oder umgekehrt („it/P“). Die wesentlichen Unterschiede zum vorangegangenen Experiment lassen sich wie folgt zusammenfassen:

- 19 statt 35 Versuchspersonen
- Durchschnittsalter von 36 statt 23 Jahren
- Fachhochschul- oder Hochschulausbildung statt Berufslehre oder Mittelschulausbildung
- Andere Verteilung der Argumente in der P-Entscheidung
- Veränderte Reihenfolge der Fragen innerhalb der Fragebögen
- Need for Cognition Test wurde aus Zeitmangel nicht durchgeführt
- Doppelt so hohe Beträge bei der leistungskontingenten Belohnung

4.2.1 Versuchspersonen

19 Studierende (drei Frauen und 16 Männer) eines Nachdiplomkurses an der Fachhochschule Aargau nahmen teil. Das Durchschnittsalter betrug 36.2 Jahre ($SD = 6.0$) bei einer Altersspannweite von 28 bis 47. Alle Studierenden verfügten über eine Fachhochschul- bzw. Hochschulausbildung und besaßen im Durchschnitt 12.1 Jahre Berufserfahrung ($SD = 6.8$).

4.2.2 Materialien

Die IT- und P-Entscheidung wurden wiederum in eine Assessment-Center-Aufgabe eingebettet. Bei beiden Entscheidungsaufgaben musste aus einer Auswahl von drei Alternativen, die beste gefunden werden. Aufgrund der geringeren Zahl von Versuchspersonen wurde darauf verzichtet, eine unterschiedliche Differenzierung der Alternativen in den zwei Entscheidungen vorzunehmen. Da die Aufgabenstellung im ersten Experiment nur einen geringen Effekt auf die Brauchbarkeitseinschätzung und die aktuelle Urteilsicherheit hatte (vgl. Abschnitt 3.3.3.6), war kein Nachteil bezüglich der Vergleichbarkeit von Experiment 2 mit Experiment 1 zu erwarten.

4.2.2.1 IT-Entscheidung

Die IT-Entscheidung wurde unverändert aus Experiment 1 übernommen.

4.2.2.2 P-Entscheidung

Im Gegensatz zu Experiment 1 war es bei dieser Untersuchung möglich, nach jedem Verarbeitungsschritt einen der drei Kandidaten (A, B und C) als geeignetsten zu erkennen (Tabelle 4-1): Die Kandidaten A und C erfüllten im ersten Interview-Bericht je drei Kriterien (Verkaufsgeschick, Leistungsbereitschaft und Kontaktfähigkeit), bei Kandidat B waren es vier (Verkaufsgeschick, Umsetzen von Fachinformation, Teamfähigkeit und Kontaktfähigkeit). Interview-Bericht 2 nannte bei Kandidat B dieselben vier bekannten Kriterien, bei Kandidat C ein neues Kriterium (Teamfähigkeit) und bei Kandidat A zwei neue Kriterien (Umsetzen von Fachinformation und Planung & Organisation) und machte Kandidat A zur besten Wahl. Bericht 3 lieferte einzig beim Kandidaten A ein weiteres neues Kriterium (Flexibilität) und machte ihn damit endgültig zur besten Option.

Tabelle 4-1. Den Kandidaten zugesprochne Eigenschaften gemäss Interview-Berichten

Entscheidungskriterien	Interviewserie 1			2			3		
	A	B	C	A	B	C	A	B	C
• Verkaufsgeschick/Überzeugungskraft	x	x	x	x	x	x	x	x	x
• Leistungsbereitschaft	x		x	x		x	x		x
• Umsetzen von Fachinformation		x		x	x		x	x	
• Planung und Organisation				x			x		
• Teamfähigkeit		x			x	x		x	x
• Flexibilität							x		
• Kontaktfähigkeit	x	x	x	x	x	x		x	x
Information je Interviewserie	3	4	3	5	4	4	5	4	4
Akkumulierte Information	3	4	3	5	4	4	6	5	5

4.2.3 Versuchsplan

Der Studie lag ein einfaktorieller Versuchsplan (unabhängige Variable *Aufgaben-Wichtigkeit*) in einem kombinierten „Between-Subjects-“ und „Within-Subjects-Design“ zugrunde: Die eine Hälfte der Versuchspersonen begann mit der „sehr wichtigen“ Aufgabe und löste dann die „weniger wichtige“ Aufgabe, die andere Hälfte der Studierenden bearbeitete zunächst die „weniger wichtige“ Aufgabe und anschliessend die „sehr wichtige“ Aufgabe.

4.2.3.1 Unabhängige Variable Aufgaben-Wichtigkeit

Die Wichtigkeit wurde wiederum anhand des potentiell entgehenden Gewinns bei der Wahl einer zweitbesten bzw. drittbesten Alternative dargestellt: Bei der „sehr wichtigen Entscheidung“ entging der Firma eine halbe Million Franken, wenn man sich anstelle der besten für die zweitbeste Alternative entschied und eine Million Franken, wenn man anstelle der besten die drittbeste Alternative wählte. Bei der „weniger wichtigen Entscheidung“ waren die Beträge Fr. 30'000.- bzw. Fr. 50'000.-.

4.2.3.2 Abhängige Variablen

Die Reihenfolge der Fragen innerhalb der Fragebögen 0-4 wurde geändert; es sollte festgestellt werden, ob sich die Resultate verändern, wenn die Versuchspersonen zuerst die Urteilssicherheits- und Brauchbarkeitsangaben machen und dann über die Endgültigkeit Ihres Urteils entscheiden, statt wie im ersten Experiment vorab über die Beendigung zu entscheiden und dann die Fragen zur Urteilssicherheit und Brauchbarkeit zu beantworten (Tabelle 4-2).

Tabelle 4-2. Fragen der Fragebögen 1-3 bei der IT-Entscheidung

1. Wenn Sie jetzt eine Firma wählen müssten, welche würden Sie favorisieren? (1)
2. Wie geeignet ist die von Ihnen bevorzugte Firma für den zu vergebenden Auftrag? (3)
3. Welche Firma ist im Moment für Sie die Nr. 2 bzw. die Nr. 3? (5)
4. Wie stark unterscheidet sich die bevorzugte Firma von der zweitbesten Firma? (6)
5. Wie sicher sind Sie, dass sich die von Ihnen im Moment bevorzugte Firma nach Ihrem endgültigen Entscheid als die beste Wahl herausstellt [%]? (4)
6. Ist dies Ihre endgültige Entscheidung? (2)

Die Angaben in Klammern entsprechen der Reihenfolge der Fragen in Experiment 1

4.2.4 Durchführung

Die Durchführung blieb grundsätzlich dieselbe wie in Experiment 2: Nach der Bearbeitung jedes Berichts konnten die Versuchspersonen wiederum wählen, ob sie entscheiden oder weitere Information beziehen wollten. In zwei Treatment-Bedingungen war entweder die erste Entscheidung sehr wichtig und die zweite weniger wichtig oder vice versa.

Als Kriterien zur Leistungsbewertung und Bezahlung wurden die selben Kriterien wie in Experiment 1 verwendet: mit erster Priorität Entscheidungsqualität und Kosten, mit zweiter Priorität Zeitbedarf. Es wurden zwei erste Preise (Bücher-Bons) à Fr. 40.-, zwei zweite Preise à Fr. 25.- und vier dritte Preise à Fr. 15.- ausbezahlt.

4.3 Ergebnisse

4.3.1 Manipulations-Überprüfung

4.3.1.1 Frage nach der wichtigeren Entscheidung

Im ersten Fragebogen wurde gefragt, welche der zu lösenden Aufgaben mehr als zehn Mal wichtiger als die andere sei. 18 der 19 Versuchspersonen gaben eine Antwort auf diese Frage und in allen Fällen war sie richtig.

4.3.1.2 Relevanz für Firma

Das Treatment zeigte sowohl im Between- als auch im Within-Subjects-Vergleich die erwartete Wirkung (Tabelle 4-3). Die Versuchspersonen bewerteten in der Versuchsbedingung „IT/p“ die IT-Entscheidung ($M = 8.5$) als wichtiger für die Firma als die P-Entscheidung ($M = 5.5$); $T = 4$, $p < .05$ (einseitig). Nur eine von zehn Personen machte eine unerwartete Bewertung (P wichtiger als IT). In der Bedingung „it/P“ war das Verhältnis gerade umgekehrt: die IT-Entscheidung ($M = 3.7$) wurde als weniger wichtiger als die P-Entscheidung ($M = 8.8$) betrachtet; $T = 0$, $p < .01$ (einseitig). Eine von neun Personen machte keinen Unterschied in ihrer Wichtigkeitseinschätzung. Unabhängig von der Aufgabenstellung unterschieden sich die wichtigen ($M = 8.7$) von den unwichtigen Entscheidungen ($M = 4.7$); $\chi = -4.87$, $p < .001$.

Tabelle 4-3. Mittelwerte und Standardabweichungen der Wichtigkeit der Entscheidung für die Firma (IT- und P-Entscheidung) in den zwei Versuchsbedingungen

	IT-Entscheidung	P-Entscheidung
IT/p	8.5 (.85)	5.5 (2.4)
it/P	3.7 (2.4)	8.9 (.8)

Standardabweichungen sind in Klammern dargestellt

4.3.1.3 Persönliche Relevanz

Im Within-Subjects-Vergleich zeigten die Mittelwerte der persönlichen Relevanz in der Bedingung „IT/p“ in die erwartete Richtung (IT-Entscheidung: $M = 8.1$ versus P-Entscheidung: $M = 5.8$). Der Unterschied war aber nicht signifikant; $T = 5$ (Tabelle 4-4). Drei von zehn Personen gaben keinen Unterschied an und eine Person machte eine unerwartete Bewertung (P wichtiger als IT). In der Bedingung „it/P“ waren die Unterschiede (IT-Entscheidung: $M = 4.8$ versus P-Entscheidung: $M = 8.8$) signifikant; $T = 0$, $p < .01$ (einseitig), obschon zwei von neun Personen keinen Unterschied feststellten.

Der Between-Subjects-Vergleich zwischen wichtigen und unwichtigen Entscheidungen ergab ein signifikantes Ergebnis ($M = 8.4$ versus $M = 5.3$); $\chi = -4.10$, $p < .001$.

Tabelle 4-4. Mittelwerte und Standardabweichungen der persönlichen Relevanz (IT- und P-Entscheidung) in den zwei Versuchsbedingungen

	IT-Entscheidung	P-Entscheidung
IT/p	8.1 (1.3)	5.8 (2.4)
it/P	4.8 (2.2)	8.8 (.8)

Standardabweichungen sind in Klammern dargestellt

4.3.2 Stopp-Mechanismen im Test

4.3.2.1 Reaktion der Führungsgrösse (Funktionstest 1)

4.3.2.1.1 Gewünschte Urteilssicherheit

Tabelle 4-5 zeigt den Within-Subjects-Vergleich. Hier strebten die Versuchspersonen in der Bedingung „IT/p“ für die IT-Entscheidung eine höhere Urteilssicherheit ($M = 89.8$) an als für die P-Entscheidung ($M = 79.0$). Der Unterschied war aber nicht signifikant; $T = 3.5$. Bei drei von neun Personen waren die Angaben für beide Entscheidungen gleich und eine Person wünschte bei der P-Entscheidung eine höhere Sicherheit. In der Bedingung „it/P“ wurde bei der IT-Entscheidung eine tiefere Urteilssicherheit ($M = 78.1$) angestrebt als bei der P-Entscheidung ($M = 92.5$). Der Unterschied war signifikant, $T = 1$, $p < .05$ (einseitig), obschon zwei von neun Personen gleiche Angaben machten und eine Person eine höhere Sicherheit bei der IT-Entscheidung wünschte.

Tabelle 4-5. Mittelwerte und Standardabweichungen der gewünschten Urteilssicherheit bei IT-Entscheidung und P-Entscheidung in den zwei Versuchsbedingungen

	IT-Entscheidung	P-Entscheidung
IT/p	89.8 (7.9)	79.0 (18.8)
it/P	78.3 (16.2)	90.6 (8.8)

Die Mittelwerte der gewünschten Urteilssicherheit in der *Between-Subjects-Perspektive* in Abbildung 4-1 zeigt, dass der Wert in den weniger wichtigen Entscheidungen signifikant tiefer war ($M = 78.7\%$) als in den wichtigeren Entscheidungen ($M = 90.2\%$); $z = -28$, $p < .01$.

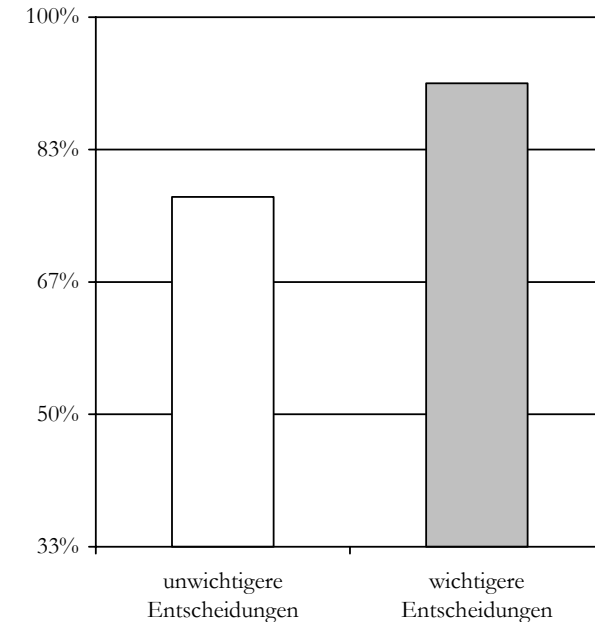


Abbildung 4-1. Mittelwerte der gewünschten Urteilssicherheit bei den weniger wichtigen und den wichtigeren Entscheidungen

4.3.2.1.2 Anspruchsniveau

Im Within-Subjects-Vergleich zeigten die Mittelwerte des Anspruchsniveaus in der Bedingung „IT/p“ bei der IT-Entscheidung zwar höhere Werte ($M = 8.6$) als bei der P-Entscheidung ($M = 7.8$), aber der Unterschied war aufgrund der hohen Zahl von indifferenten Personen (fünf von zehn) nicht signifikant; $T = 0$ (Tabelle 3-15). Ein ähnliches Bild ergab sich in der Bedingung „it/P“. Der Unterschied zwischen der IT-Entscheidung ($M = 7.2$) und der P-Entscheidung ($M = 8.3$) war ebenfalls nicht signifikant; $T = 2$. fünf von neun Personen verzeichneten ein höheres Anspruchsniveau bei der P-Entscheidung und 3 waren indifferent.

Tabelle 4-6. Mittelwerte und Standardabweichungen des Anspruchsniveaus bei IT-Entscheidung und P-Entscheidung in den zwei Versuchsbedingungen

	IT-Entscheidung	P-Entscheidung
IT/p	8.6 (.7)	7.8 (1.1)
it/P	7.2 (1.7)	8.3 (1.2)

Standardabweichungen sind in Klammern dargestellt

Abbildung 4-2 gibt die Ergebnisse des *Between-Subjects-Vergleichs* der Mittelwerte des Anspruchsniveaus wieder. Er ergab bei den weniger wichtigen Entscheidungen ein signifikant tieferes Anspruchsniveau ($M = 7.6$) als bei den wichtigeren Entscheidungen ($M = 8.4$); $U = 113, p < .05$ (einseitig).

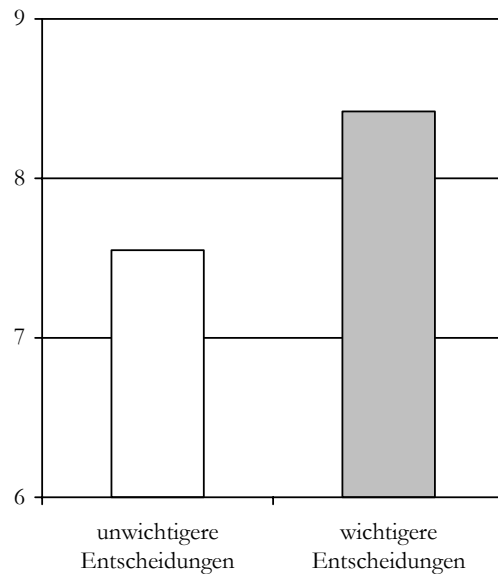


Abbildung 4-2. Mittelwerte des Anspruchsniveaus bei den weniger wichtigen und den wichtigeren Entscheidungen

4.3.2.2 Reaktion der Regelgröße (Funktionstest 2)

4.3.2.2.1 Aktuelle Urteilssicherheit

Alle 38 Entscheidungen wurden in Abhängigkeit vom Verarbeitungsende in drei Gruppen eingeteilt (Ein- bis Drei-Schritt-Entscheidungen). Abbildung 4-3 zeigt, dass die Urteilssicherheit in allen Gruppen kontinuierlich zunahm. Der Anstieg der Urteilssicherheit war bei den Ein-Schritt- ($M_0 = 49\%$, $M_1 = 67\%$; $\chi^2(1) = 5.44, p < .05$), den Zwei-Schritt- ($M_0 = 57\%$, $M_1 = 70\%$, $M_2 = 86\%$; $\chi^2(2) = 12.58, p < .01$) und den Drei-Schritt-Entscheidungen ($M_0 = 53\%$, $M_1 = 63\%$, $M_2 = 72\%$, $M_3 = 92\%$; $\chi^2(3) = 28.19, p < .001$) signifikant.

Die Wirkung der Verarbeitungsschritte auf die Höhe der *definitiven Urteilssicherheit* war signifikant, wie ein Kruskal-Wallis-Test ergab; $\chi^2(2) = 17.1, p < .001$. Die definitive Urteilssicherheit war um so höher, je mehr Information bis zur Entscheidung genutzt wurde. Die Unterschiede zwischen den Ein-Schritt- ($M = 67.2$), Zwei-Schritt- ($M = 85.7$) und Drei-Schritt-Entscheidungen ($M = 92.1$) waren signifikant (Games-Howell auf 5%-Niveau).

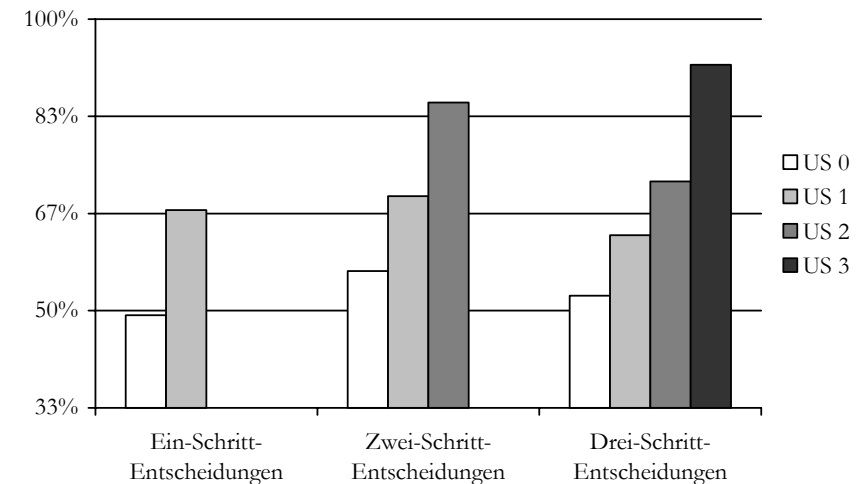


Abbildung 4-3. Mittelwerte der Urteilssicherheit nach der Vorinformation (US 0) und den Verarbeitungsschritten 1 bis 3 (US 1 bis US 3) bei Ein- bis Drei-Schritt-Entscheidungen

4.3.2.2.2 Brauchbarkeitseinschätzung

Abbildung 4-4 zeigt, dass die Mittelwerte der *Brauchbarkeitseinschätzung* bei allen Gruppen von Entscheidungen mit fortlaufender Informationsnutzung zunahm. Dieser Anstieg war bei den neun Ein-Schritt- ($M_0 = 5.0, M_1 = 6.5; \chi^2(1) = 4.5, p < .05$), den 15 Zwei-Schritt- ($M_0 = 5.3, M_1 = 7.0, M_2 = 7.9; \chi^2(2) = 17.8, p < .001$) und den elf Drei-Schritt-Entscheidungen ($M_0 = 5.9, M_1 = 6.8, M_2 = 7.2, M_3 = 8.4; \chi^2(3) = 19.4, p < .001$) signifikant.

Die Verarbeitungsschritte hatten ebenfalls eine Wirkung auf die Höhe der *definitiven Brauchbarkeitseinschätzung*, $\chi^2(2) = 16.8, p < .001$. Die gewählte Alternative wurde um so besser eingeschätzt, je mehr Information bis zur Entscheidung verarbeitet wurde. Dabei unterschieden sich die Ein-Schritt- ($M = 6.6$) sowohl von den Zwei-Schritt- ($M = 7.9$) als auch den Drei-Schritt-Entscheidungen ($M = 8.4$) auf dem 5%-Niveau (Games-Howell).

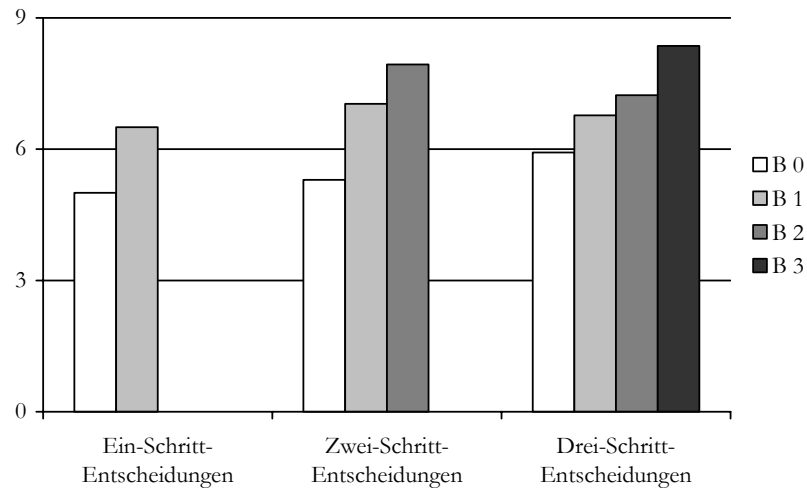


Abbildung 4-4. Mittelwerte der aktuellen Brauchbarkeitseinschätzung nach der Vorinformation (B 0) und den Verarbeitungsschritten 1 bis 3 (B 1 bis B 3) bei Ein- bis Drei-Schritt-Entscheidungen

4.3.2.3 Wirkung der Regelabweichung (Funktionstest 3)

4.3.2.3.1 Erkennen „letzter“ Schritte

Die Erwartung, dass das Sicherheitsdefizit und die Brauchbarkeits-Lücke „letzte“ Schritte erkennen würden, wurde anhand der 75 bzw. 74 Prozessschritte untersucht, bei denen die Versuchspersonen definitive Entscheidungen fällen konnten (Verarbeitungsschritte 1 bis 3).

Bei den Prozessschritten, bei denen es sich um den „Letzten“ handelte, war das *Sicherheitsdefizit* im Durchschnitt signifikant kleiner, ($M = .6\%$), als bei den Prozessschritten, wo weitere Information verlangt wurde ($M = 23.7\%$); $\chi = -6.05, p < .001$ (Abbildung 4-5a).

Bei der *Brauchbarkeits-Lücke* ergab sich ein sehr ähnliches Bild (Abbildung 4-5b): Die Brauchbarkeits-Lücke war bei den „letzten“ Prozessschritten ($M = .1$) deutlich kleiner als bei den „fortgesetzten“ Prozessschritten ($M = 1.4$); $\chi = -3.90, p < .001$.

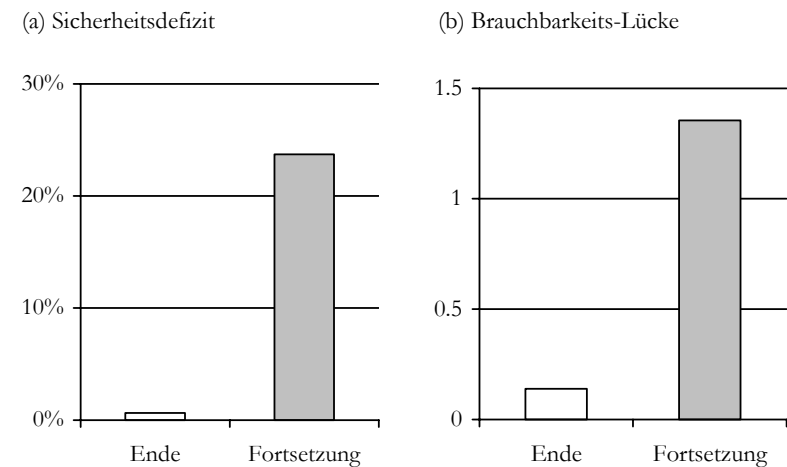


Abbildung 4-5. Mittelwerte des (a) Sicherheitsdefizits und der (b) Brauchbarkeits-Lücke bei Verarbeitungsschritten ohne bzw. mit Fortsetzung

Ein Vergleich mit dem ξ -Wert der Verarbeitungsschritte ($\xi = -4.42, p < .001$) ergab, dass einzig das Sicherheitsdefizit ($\xi = -6.05$) diese als Prädiktorvariable zu schlagen vermochte (Tabelle 4-7). Alle anderen Variablen waren deutlich schlechter: aktuelle Urteilssicherheit ($\xi = -3.94, p < .001$), Brauchbarkeits-Lücke ($\xi = -3.9, p < .001$) und Brauchbarkeitseinschätzung ($\xi = -3.00, p < .01$).

Zusätzlich wurde mittels Diskriminanzanalyse geprüft, wie gut die Vorhersage beider Regelabweichungen war. Die Klassifizierung war beim Sicherheitsdefizit in 82.7% und bei der Brauchbarkeits-Lücke in 73.0% der Fälle richtig.

Tabelle 4-7. Mittelwerte, Standardabweichungen und ξ -Wert der absolvierten Verarbeitungsschritte, Sicherheitsdefizit, aktuelle Urteilssicherheit, Brauchbarkeits-Lücke und Brauchbarkeitseinschätzung bei abschliessenden bzw. provisorischen Urteilen

	<i>M (SD) Ende</i>	<i>M (SD) Fortsetzung</i>	ξ -Wert
Verarbeitungsschritte	2.1 (.8)	1.3 (.5)	-4.42, $p < .001$
Sicherheitsdefizit	.6% (8.5%)	23.7% (16.6%)	-6.05, $p < .001$
aktuelle Urteilssicherheit	84% (14%)	68% (18%)	-3.94, $p < .001$
Brauchbarkeits-Lücke	.1 (1.3)	1.4 (1.3)	-3.90, $p < .001$
Brauchbarkeitseinschätzung	7.8 (1.0)	7.0 (1.1)	-3.00, $p < .01$

4.3.2.3.2 Vorhersage späterer Informationsnutzung

Nur das Sicherheitsdefizit nach der Vorinformation ($\beta = .34$) konnte die später absolvierten Verarbeitungsschritte vorhersagen; $F(1, 36) = 4.7, p < .05$. Die Brauchbarkeits-Lücke nach der Vorinformation erwies sich als nicht signifikanter Prädiktor; $F(1, 35) = .9$. Vom Treatment *Aufgaben-Wichtigkeit* ($\beta = .79$) ging erwartungsgemäss der stärkste Einfluss auf die Informationsnutzung aus; $F(1, 36) = 59.5, p < .001$.

Das *Sicherheitsdefizit nach der Vorinformation* vermochte über das Treatment hinaus keinen Effekt auf die Informationsnutzung ausüben, wie Tabelle 4-8 zeigt.

Tabelle 4-8. Aufgaben-Wichtigkeit und Sicherheitsdefizit nach der Vorinformation als Prädiktoren zur Vorhersage der Verarbeitungsschritte

	R^2	β	r
Aufgaben-Wichtigkeit (Treatment)	.62	.79***	.79
Sicherheitsdefizit nach der Vorinformation			.10

*** $p < .001$

Werden alle möglichen Prädiktorvariablen berücksichtigt, so erscheint neben der Aufgaben-Wichtigkeit ($\beta = .66$) nur noch die gewünschte Urteilssicherheit ($\beta = .33$) in der Regressionsgleichung (Tabelle 4-9, Modell 2). Sie erhöht die Varianzaufklärung um 9%; $F_{change}(1, 34) = 10.5, p < .01$.

Tabelle 4-9. Alle Prädiktoren zur Vorhersage der Verarbeitungsschritte

	R^2	ΔR^2	β	r
Modell 1				
Aufgaben-Wichtigkeit (Treatment)	.62		.79***	.79
Modell 2				
Aufgaben-Wichtigkeit (Treatment)	.71	.09**	.66***	.75
Gewünschte Urteilssicherheit			.33**	.49
Selbstwirksamkeitserwartung				-.08
Relevanz für Firma				.09
Persönliche Relevanz				.16
Anspruchsniveau				.14
anfängliche aktuelle Urteilssicherheit				.14
anfängliche Brauchbarkeitseinschätzung				.24
anfängliche Differenzwahrnehmung				-.14
Sicherheitsdefizit nach der Vorinformation				-.14
Brauchbarkeits-Lücke nach der Vorinformation				-.17

*** $p < .001$; ** $p < .01$

4.3.3 Zusammenhänge zwischen den Variablen

4.3.3.1 Anspruchsniveau

Tabelle 4-10 zeigt, dass nur die Variable *persönliche Relevanz* ($\beta = .54$) auf das Anspruchsniveau wirkte; $F(1, 36) = 14.5, p < .001$.

Tabelle 4-10. Prädiktoren zur Vorhersage des Anspruchsniveaus

	R^2	β	r
Persönliche Relevanz	.29	.54***	.54
Selbstwirksamkeitserwartung		.	-.14
Aufgaben-Wichtigkeit (Treatment)			-.07
Relevanz für Firma			.08

*** $p < .001$

In der Regressionsanalyse mit der Kriteriumsvariablen *persönliche Relevanz* fand nur der Prädiktor *Relevanz für die Firma* ($\beta = .92$) Eingang in die Gleichung; $F(1, 36) = 188.1, p < .001$ (Tabelle 4-11).

Tabelle 4-11. Prädiktoren zur Vorhersage der persönliche Relevanz

	R^2	β	r
Relevanz für Firma	.84	.91***	.91
Selbstwirksamkeitserwartung			.08
Aufgaben-Wichtigkeit (Treatment)			-.15

*** $p < .001$

Die Relevanz für die Firma wurde einzig durch die *Aufgaben-Wichtigkeit* ($\beta = .76$) beeinflusst; $F(1, 36) = 49.8, p < .001$ (Tabelle 4-12).

Tabelle 4-12. Prädiktoren zur Vorhersage der Relevanz für Firma

	R^2	β	r
Aufgaben-Wichtigkeit (Treatment)	.58	.76***	.60
Selbstwirksamkeitserwartung			.07

*** $p < .001$

4.3.3.2 Gewünschte Urteilssicherheit

Die Analyse der 38 Entscheidungen zeigte, dass nur die Variablen *persönliche Relevanz* ($\beta = .49$) und *Anspruchsniveau* ($\beta = .43$) die gewünschte Urteilssicherheit beeinflussten; $F(2, 35) = 30.1, p < .001$ (Tabelle 4-13).

Tabelle 4-13. Prädiktoren zur Vorhersage der gewünschten Urteilssicherheit

	R^2	β	r
Persönliche Relevanz	.61	.48***	.55
Anspruchsniveau		.43**	.51
Aufgaben-Wichtigkeit (Treatment)			-.10
Relevanz für Firma			-.07
Selbstwirksamkeitserwartung			.10

*** $p < .001, ** p < .01$

4.3.3.3 Brauchbarkeitseinschätzung

Die Analysen wurden mit 112 Verarbeitungsschritten durchgeführt. Verglichen mit dem zweiten Experiment stand damit eine geringere Zahl von Auswertungseinheiten zur Verfügung. Die Schwelle, an der Variablen in die Regressionsgleichungen aufgenommen wurden, lag deshalb bei 2% zusätzlicher Varianzaufklärung (zweites Experiment: 1%).

Nur die Variable *Verarbeitungsschritte* ($\beta = .63$) erschien in der Regressionsgleichung mit der Kriteriumsvariablen Brauchbarkeitseinschätzung; $F(1, 107) = 71.7, p < .001$ (Tabelle 4-14).

Tabelle 4-14. Prädiktoren zur Vorhersage der Brauchbarkeitseinschätzung

	R^2	β	r
Verarbeitungsschritte	.40	.63***	.63
Differenzwahrnehmung			.09
Aufgaben-Wichtigkeit			.04
Relevanz Firma			.08
Persönliche Relevanz			-.05
Selbstwirksamkeits-Erwartung			.25
Aufgabe (IT oder P)			-.10

*** $p < .001$

4.3.3.4 Aktuelle Urteilssicherheit

Die *Verarbeitungsschritte* erschienen als erste in der Gleichung und erklärten 45% der Varianz der aktuellen Urteilssicherheit; $F(1, 106) = 86.1, p < .001$ (Tabelle 4-15, Modell 1). Als zweite Variable tauchte die *Brauchbarkeitseinschätzung* auf und verbesserte die Varianzaufklärung um 9%; $F_{change}(1, 105) = 20.6, p < .001$ (Tabelle 4-15, Modell 2). Als letzte Variable erschien die *Differenzwahrnehmung* in der Gleichung und erhöhte die aufgeklärte Varianz um weitere 5%; $F_{change}(1, 104) = 12.6, p < .001$ (Tabelle 4-15, Modell 3).

Tabelle 4-15. Prädiktoren zur Vorhersage der aktuellen Urteilssicherheit

	R^2	ΔR^2	β	r
Modell 1				
Verarbeitungsschritte	.45		.67***	.67
Modell 2				
Verarbeitungsschritte	.54	.09***	.42***	.43
Brauchbarkeitseinschätzung			.39***	.40
Modell 3				
Verarbeitungsschritte	.56	.05***	.34***	.36
Brauchbarkeitseinschätzung			.36***	.40
Differenzwahrnehmung			.25***	.33
Aufgaben-Wichtigkeit				.20
Relevanz Firma				-.06
Persönliche Relevanz				.16
Selbstwirksamkeits-Erwartung				.16
Aufgabe (IT oder P)				.12

*** $p < .001$

4.3.3.5 Differenzwahrnehmung

Für eine weitere Analyse blieb als signifikanter Prädiktor der aktuellen Urteilssicherheit nur die Differenzwahrnehmung übrig. Tabelle 4-16 zeigt, dass sich einzig die Variable *Verarbeitungsschritte* als signifikante Vorhersagevariable erwies. Sie erklärten 17% der Varianz; $F(1, 110) = 22.3, p < .001$.

Tabelle 4-16. Prädiktoren zur Vorhersage der Differenzwahrnehmung

	R^2	β	r
Verarbeitungsschritte	.17	.41***	.41
Aufgaben-Wichtigkeit			.08
Relevanz Firma			.01
Persönliche Relevanz			.01
Selbstwirksamkeits-Erwartung			-.02
Aufgabe (IT oder P)			.03

*** $p < .001$

4.4 Diskussion

4.4.1 Stopp-Mechanismen im Test

4.4.1.1 Reaktion der Führungsgrösse (Funktionstest 1)

Die *gewünschte Urteilssicherheit* war bei den wichtigeren Entscheidungen signifikant höher als bei den weniger wichtigen Entscheidungen. In den Within-Subjects-Vergleichen waren die Unterschiede einmal signifikant und einmal nicht signifikant. Die Aufgaben-Wichtigkeit erklärte 16% der Varianz der gewünschten Urteilssicherheit. Hypothese 1a, der Stopp-Mechanismus Urteilssicherheit ist ein Folgerregler, fand damit Unterstützung.

Beim *Anspruchsniveau* ergab sich bei den weniger wichtigen Entscheidungen ein signifikant tieferer Durchschnittswert verglichen mit den wichtigeren Entscheidungen. In den Within-Subjects-Vergleichen waren die Differenzen zwar nicht signifikant aber mit einer Varianzaufklärung von 10% wurde die Annahme widerlegt (Hypothese 1b), die Aufgaben-Wichtigkeit habe keine Wirkung auf das Anspruchsniveau. Somit scheint der Stopp-Mechanismus Satisficing auf dem Prinzip der Folgereglung zu beruhen.

4.4.1.2 Reaktion der Regelgrösse (Funktionstest 2)

Sowohl die *aktuelle Urteilssicherheit* als auch die *Brauchbarkeitseinschätzung* stiegen im Verlauf aller Entscheidungsprozesse kontinuierlich an. Die definitive Urteilssicherheit und die definitive Brauchbarkeitseinschätzung waren um so höher, je mehr Information bis zur Entscheidung genutzt wurde. Bei der de-

finitiven Urteilssicherheit unterschieden sich zwischen sowohl die Ein-Schritt- als auch die Zwei-Schritt- und Drei-Schritt-Entscheidungen voneinander. Bei der definitiven Brauchbarkeitseinschätzung war nur der Unterschied zwischen den Ein-Schritt- und den Mehr-Schritt-Entscheidungen signifikant. Beide Stopp-Mechanismen bestanden diesen Test (Hypothesen 2) und der Nachweis für einen Assessor_{ist} wurde somit erbracht.

4.4.1.3 Wirkung der Regelabweichung (Funktionstest 3)

Das *Sicherheitsdefizit* und die *Brauchbarkeits-Lücke* erkannten „letzte“ Schritte und erfüllten damit die Erwartung, die Regelabweichung unterscheide definitive von provisorischen Urteilen. Da die Wahrscheinlichkeit, dass ein endgültiges Urteil gefällt wird, mit jedem Verarbeitungsschritt steigt und die Verarbeitungsschritte beide Ist-Werte beeinflussten (Funktionstest 2), waren diese Ergebnisse zu erwarten. Hypothesen 3 forderte deshalb nicht nur ein Erkennen „letzter“ Schritte, sondern auch eine verglichen mit den Verarbeitungsschritten bessere Vorhersage. Diese Anforderung erfüllte nur das Sicherheitsdefizit. Die Brauchbarkeits-Lücke machte eine deutlich schlechtere Vorhersage als die Variable Verarbeitungsschritte. Somit unterstützten diese Ergebnisse Hypothese 3 nur im Falle des Stopp-Mechanismus Urteilssicherheit und es lässt sich annehmen, dass (sich) die Versuchspersonen entschieden haben, weil sie „sicher genug“ geworden waren.

Die Erwartung, die anfänglichen Regelabweichungen beeinflusse die Informationsnutzung, bestätigte sich nur für das *Sicherheitsdefizit nach der Vorinformation*. Dieser Effekt verschwand aber, wenn das *Treatment Aufgaben-Wichtigkeit* in die Analyse miteinbezogen wurde.

4.4.1.4 Vergleich beider Stopp-Mechanismen

Aufschluss darüber, welcher der beiden Mechanismen den Anforderungen eines funktionierenden Regelkreises besser entsprach, ergaben die Resultate von Funktionstest 2 und Funktionstest 3.

Funktionstest 2 zeigte, dass beide Mechanismen mit ihren Regelgrößen deutlich auf die genutzte Information reagierten. Die Wirkung der Verarbeitungsschritte auf die aktuelle Urteilssicherheit war aber stärker als diejenige der Verarbeitungsschritte auf die Brauchbarkeitseinschätzung. Im ersten Fall betrug die Varianzaufklärung 45% und im zweiten Fall 40% (Tabelle 4-17).

Tabelle 4-17. Vergleich der Stopp-Mechanismen Urteilssicherheit und Satisficing

	Urteilssicherheit	Brauchbarkeitseinschätzung
Reaktion der Regelgröße (Test 2)	.45	.40
Erkennen „letzter“ Schritte (Test 3)	.43	.19

Zahlenangaben entsprechen der Varianzaufklärung;

Funktionstest 3 bestand nur der Stopp-Mechanismus Urteilssicherheit in vollem Umfang. Einzig das Sicherheitsdefizit erkannte „letzte“ Schritte und war in der Vorhersage besser als die Summe der durchgeführten Verarbeitungsschritte. Die Brauchbarkeits-Lücke wurde dem zweiten Teil dieser Forderung nicht gerecht. Die unterschiedlich starke Wirkung auf das Verarbeitungsende zeigt sich einerseits an der Varianzaufklärung, hier lag der Wert für das Sicherheitsdefizit bei 43% und derjenige der Brauchbarkeits-Lücke bei 19% (Tabelle 4-17) und andererseits an der Trefferquote, das Sicherheitsdefizit klassifizierte in 82.7% und die Brauchbarkeits-Lücke in 73.0% der Fälle richtig.

Ein weiterer Hinweis auf die grössere Bedeutung des Stopp-Mechanismus Urteilssicherheit gab die Prüfung, ob die anfängliche Regelabweichung die spätere Informationsnutzung vorhersagen könne. Dies war nur bei dem anfänglichen Sicherheitsdefizit der Fall, sie erklärte 12% der Varianz der Verarbeitungsschritte.

In dieser Untersuchung schnitt der Stopp-Mechanismus Urteilssicherheit bei allen Gegenüberstellungen besser ab als der Stopp-Mechanismus Satisficing. Der Stopp-Mechanismus Urteilssicherheit liefert somit einen höheren Erklärungserfolg. Die Frage, wann Entscheidungsprozesse beendet werden, lässt sich deshalb für diese Untersuchung wie folgt beantworten: Erfahrenen Entscheider wissen eher genug, wenn sie „sicher genug“ sind, als wenn eine Lösung „gut genug“ ist.

4.4.2 Zusammenhänge zwischen den Variablen

4.4.2.1 Anspruchsniveau

Für das Anspruchsniveau ergeben sich folgende Zusammenhänge: Die Versuchspersonen wünschten eine umso bessere Lösung, je relevanter sie die Entscheidungsaufgabe für sich selber wahrnahmen. Die persönliche Relevanz

war um so grösser, je bedeutsamer die Studierenden die Entscheidungsaufgabe im Hinblick auf die Firma begriffen. Die Höhe dieser firmenbezogenen Relevanz war bei der sehr wichtigen Aufgabe grösser als bei der weniger wichtigen Aufgabe. (Abbildung 4-6).

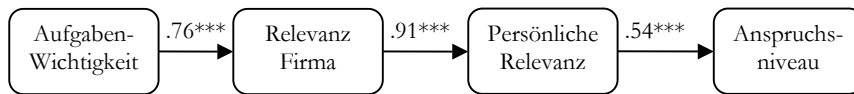


Abbildung 4-6. Zusammenhänge zwischen Aufgaben-Wichtigkeit, Relevanz für die Firma, persönliche Relevanz und Anspruchsniveau (β -Gewichte; *** $p < .001$)

4.4.2.2 Gewünschte Urteilssicherheit

Das Wirkungsgefüge der gewünschten Urteilssicherheit in Abbildung 4-7 zeigt, dass die Versuchspersonen bei ihrem endgültigen Urteil um so sicherer sein wollten, je relevanter sie die Entscheidung für sich selber bewerteten und je besser die Lösung sein sollte.

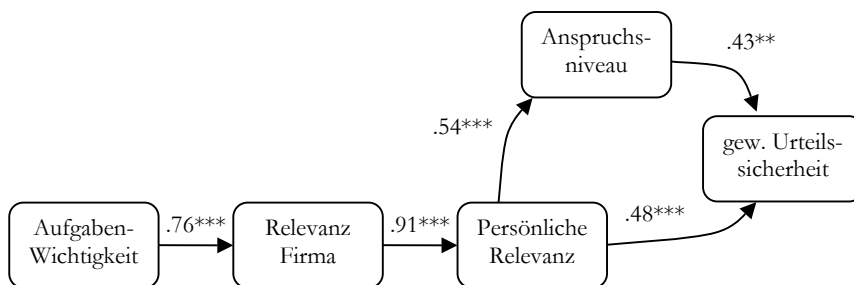


Abbildung 4-7. Zusammenhänge zwischen Aufgaben-Wichtigkeit, Relevanz für die Firma, persönliche Relevanz, Anspruchsniveau und gewünschter Urteilssicherheit (β -Gewichte; *** $p < .001$, ** $p < .01$, * $p < .05$)

Die persönliche Relevanz war um so höher, je bedeutungsvoller die Entscheidung für die Firma bewertet wurde und diese firmenbezogene Relevanz war bei der sehr wichtigen Entscheidungsaufgabe höher als bei der weniger wichtigen Aufgabe.

Der individuelle Faktor Selbstwirksamkeitserwartung erscheint nirgends in diesem Wirkungsgefüge. Somit wurde die Erwartung, persönliche Faktoren beeinflussen die Führungsgrösse im Stopp-Mechanismus Urteilsicherheit (Eagly & Chaiken, 1993; Maheswaran & Chaiken, 1991), nicht bestätigt.

4.4.2.3 Brauchbarkeitseinschätzung

Die sich aus dem *Brauchbarkeits-Test* ergebende Erwartung, es bestehe ein positiver Zusammenhang zwischen den Verarbeitungsschritten und der Brauchbarkeitseinschätzung sowie ein negativer Zusammenhang zwischen Aufgaben-Wichtigkeit und Brauchbarkeitseinschätzung, erfüllte sich nicht: Zwar bewerteten die Versuchspersonen die Brauchbarkeit der bevorzugten Alternative um so höher, je mehr Information sie verarbeitet hatten, aber sie liessen sich von der Aufgaben-Wichtigkeit nicht beeinflussen (Abbildung 4-8). Der fehlende negative Zusammenhang zwischen Aufgaben-Wichtigkeit und Brauchbarkeitseinschätzung spricht gegen die Vorstellung, dass Entscheider bei wichtigeren Problemen mehr Image-Elemente aktivieren und/oder darin höhere Anforderungen stellen als bei weniger wichtigen Problemen. Der vorliegende Befund widerlegt somit den vorgeschlagenen Brauchbarkeits-Test.

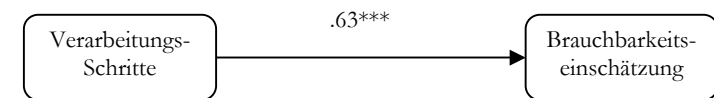


Abbildung 4-8. Zusammenhänge zwischen Verarbeitungsschritten und Brauchbarkeitseinschätzung (β -Gewichte; *** $p < .001$)

Wie lässt sich die Bildung der Brauchbarkeitseinschätzung erklären? Gemäss Burnstein & Vinokur (1977) sucht eine Person, die einen Entscheid

fällen muss, nach Eigenschaften, welche die einzelnen Wahlmöglichkeiten charakterisieren. Sie wägt Vorzüge und Nachteile der einzelnen Wahlmöglichkeiten gegeneinander ab und entscheidet sich für jene Wahlmöglichkeit, für die mehr und bessere Argumente sprechen. Da mit jedem Verarbeitungsschritt zusätzliche Argumente geliefert wurden, welche die beste Alternative innerhalb beider Entscheidungen präferabler ausschauen liess (vgl. Abschnitte 4.2.2.1 und 4.2.2.2), reflektierte die Brauchbarkeitseinschätzung die akkumulierte Summe an Evidenz für die beste Alternative.

4.4.2.4 Aktuelle Urteilssicherheit

Die Versuchspersonen waren um so sicherer, je mehr Verarbeitungsschritte sie absolviert hatten, je brauchbarer sie die bevorzugte Lösung einschätzten und je unterschiedlicher sie die beste und die zweitbesten Alternative wahrnahmen (Abbildung 4-9). Die aus dem *Contingency Model for the Selection of Decision Strategies* (Beach & Mitchell, 1978) abgeleitete Forderung nach einem positiven Zusammenhang zwischen den Verarbeitungsschritten und der aktuellen Urteilssicherheit wurde erfüllt. Dieser Befund spricht für eine prozessbezogene Entstehung der Urteilssicherheit, die den geleisteten analytischen Aufwand zur Bildung der aktuellen Urteilssicherheit heranzieht. Anspruch auf alleinige Gültigkeit konnte das Contingency Modell jedoch nicht erheben, da die Variable Verarbeitungsschritte nicht die einzige Einflussgrösse bildete. Neben dem prozessbezogenen Einfluss spielten auch noch erkenntnisbezogene Grössen in Form der Variablen Brauchbarkeitseinschätzung und Differenzwahrnehmung eine Rolle.

Die Wirkung der beiden letzt genannten Variablen wurde mit den Vorhersagen der beiden erkenntnisbezogenen Theorien verglichen. Widerlegt wurde die *Support Theory* von Tversky und Koehler (1994). Weder wurde ein enger positiver Zusammenhang zwischen der Differenzwahrnehmung und der Urteilssicherheit noch wurde ein schwächerer negativer Zusammenhang zwischen der Brauchbarkeitseinschätzung und der Urteilssicherheit festgestellt. Der Einfluss der Brauchbarkeit war grösser als derjenige der Differenz und der Zusammenhang zwischen Brauchbarkeitseinschätzung und Konfidenz war positiv statt negativ. Der Glaube an die Richtigkeit einer Hypothese war somit nicht von der relativen Unterstützung für diese Hypothese abhängig.

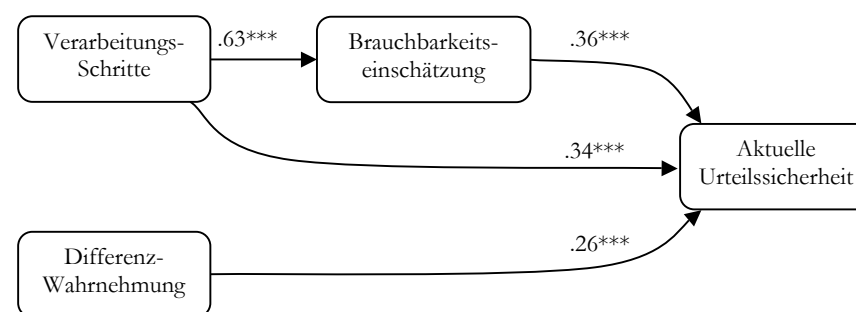


Abbildung 4-9. Zusammenhänge zwischen Verarbeitungsschritten, Brauchbarkeitseinschätzung, Differenzwahrnehmung und aktueller Urteilssicherheit (β -Gewichte; $*** p < .001$)

Unterstützung fand dagegen das *Model of Belief Processing* von Curley und Benson (1994). Die Erwartung, es bestehe ein enger positiver Zusammenhang zwischen der Brauchbarkeitseinschätzung und der Urteilssicherheit, wurde erfüllt. In Form der Brauchbarkeitseinschätzung zeigte damit die kumulative Stärke der Argumente ihre – wenn auch nicht alleinige – Wirkung auf die Urteilssicherheit.

Insgesamt sprechen die Ergebnisse dafür, dass die Versuchspersonen sich bei der Bildung ihrer Urteilssicherheit auf den Entscheidungsprozess und die dabei gewonnenen Erkenntnisse abstützten. Die dabei verwendeten Cues waren: Brauchbarkeitseinschätzung, Differenzwahrnehmung und Verarbeitungsschritte. Alle Cues wirkten proportional zu ihrer Stärke, ohne dass beim erkenntnisbezogenen Einfluss die relative Unterstützung berücksichtigt wurde.

4.4.3 Zusammenwirken beider Mechanismen

Die Zusammenhänge zwischen den Variablen beider Regelkreise deuten darauf hin, dass der Stopp-Mechanismus Urteilssicherheit den Input aus dem Regelkreis Satisficing verwendet. Daraus ergibt sich ein Entscheidungs-System mit einem integrierten Stopp-Mechanismus (Abbildung 4-10).

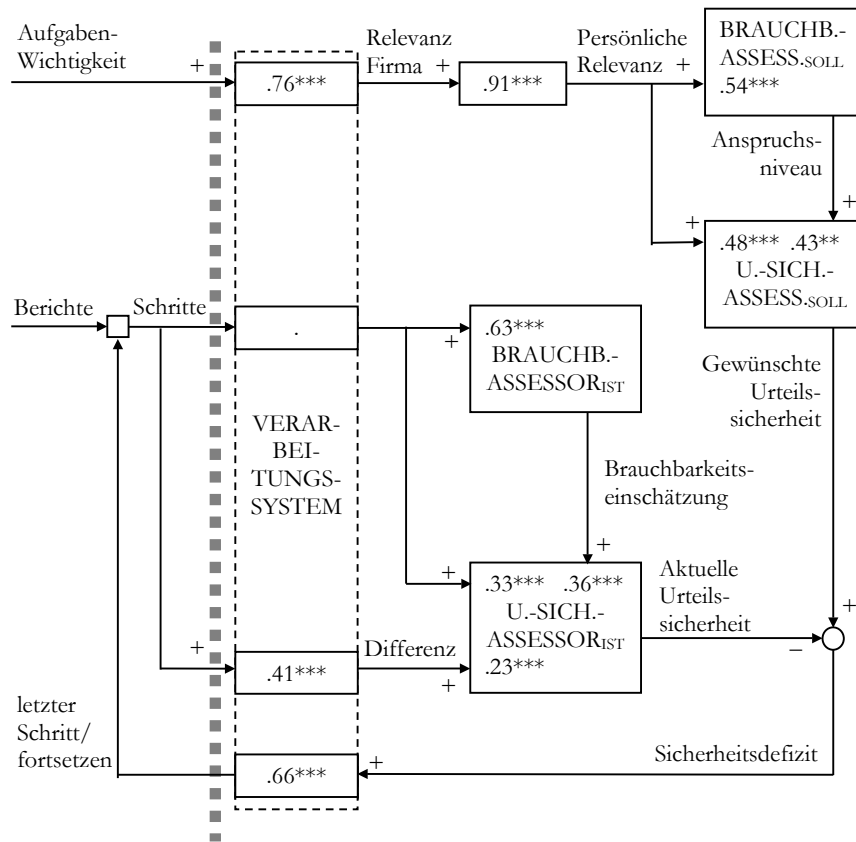


Abbildung 4-10. Entscheidungs-System mit einem integrierten Stopp-Mechanismus (Zahlenangaben sind β -Gewicht; $*** p < .001$, $** p < .01$)

Das Verarbeitungssystem verarbeitet äussere Signale. Dabei wirkt die Aufgaben-Wichtigkeit auf die Variable Relevanz für Firma. Relevanz für Firma wirkt wiederum auf die Variable persönliche Relevanz. Die persönliche Relevanz wirkt über den Brauchbarkeits-Assessor_{soll} auf das Anspruchsniveau. Das Anspruchsniveau und die persönliche Relevanz liefern den Input für den Urteilssicherheits-Assessor_{soll}, der daraus die Führungsgrösse gewünschte Urteilssicherheit bildet. Im Verarbeitungssystem wird die angeforderte Infor-

mation verarbeitet. Bei dieser Elaboration wirkt die Variable Verarbeitungsschritte auf den Brauchbarkeits-Assessor_{ist}. Der Brauchbarkeits-Assessor_{ist} bildet die Brauchbarkeitseinschätzung, welche gemeinsam mit den Variablen Verarbeitungsschritte und Differenzwahrnehmung auf den Urteilssicherheits-Assessor_{ist} wirkt. Die Differenzwahrnehmung wird durch die Variable Verarbeitungsschritte beeinflusst. Im Urteilssicherheits-Assessor_{ist} entsteht die Regelgrösse aktuelle Urteilssicherheit. Aus der Subtraktion Führungsgrösse minus Regelgrösse entsteht die Regelabweichung Sicherheitsdefizit. Das Sicherheitsdefizit beeinflusst das Informationsnutzungsverhalten des Produktionssystem. Ist das Sicherheitsdefizit klein oder gleich null, dann wird eine Entscheidung gefällt (letzter Schritt), ist es grösser, dann wird noch weitere Information gesucht und der Prozess setzt sich fort.

Weil organische Strukturen in ihrer Genese stammesgeschichtlich ältere Adaptationen weiterverwenden, umbauen oder einbeziehen (Bischof, 1993), kommt ein Stopp-Mechanismus, der alle innerhalb eines Entscheidungsprozesses verfügbaren Funktionen verwendet, der Wahrheit wahrscheinlich näher als ein Mechanismus, der auf einem einzigen Prinzip beruht.

4.5 Zusammenfassung

Wie in der ersten Untersuchung wurde in dieser Studie erneut geprüft, ob eher der Stopp-Mechanismus Urteilssicherheit oder eher der Stopp-Mechanismus Satisficing für das Ende der Entscheidungsprozesse verantwortlich ist und welche Theorien die Bildung der Brauchbarkeitseinschätzung und die Entstehung der Urteilssicherheit am besten erklären können.

Im ersten Funktionstest reagierten beide *Stopp-Mechanismen* auf die Aufgaben-Wichtigkeit, was darauf schliessen liess, dass beide Regelkreise nach dem Prinzip der Folgeregelung funktionierten. Der zweite Funktionstest ergab mit jedem Verarbeitungsschritt einen kontinuierlichen Anstieg der Urteilssicherheit und der Brauchbarkeitseinschätzung. Dabei reagierte die Urteilssicherheit etwas stärker auf Information als die Brauchbarkeitseinschätzung. Den Anforderungen des dritten Funktionstests wurde nur der Stopp-Mechanismus Urteilssicherheit vollständig gerecht. Einzig das Sicherheitsdefizit erkannte „letzte“ Schritte und war in der Vorhersage besser als die Variable Verarbeitungsschritte. Die Brauchbarkeits-Lücke wurde in der Vorhersage des Verarbeitungsendes von den Verarbeitungsschritten geschlagen. Somit liess sich die Frage, wann Entscheidungen beendet werden, für diese Untersuchung recht

eindeutig beantworten: Erfahrene Entscheider wissen genug, wenn sie „sicher genug“ sind.

Die *Brauchbarkeitseinschätzung* war um so höher, je mehr Information die Versuchspersonen verarbeitet hatten. Die Erwartung des Brauchbarkeits-Tests, dass Alternativen bei der sehr wichtigen Aufgabe zurückhaltender bewertet werden als bei der weniger wichtigen Aufgabe, bestätigte sich nicht. Das Fehlen des negativen Zusammenhangs zwischen Aufgaben-Wichtigkeit und Brauchbarkeitseinschätzung sprach gegen die Vorstellung, dass um so mehr Image-Elemente geprüft und um so höhere Anforderungen gestellt werden, je wichtiger eine Entscheidung ist.

Die Werte der *aktuellen Urteilssicherheit* lagen um so höher, je mehr Information die Versuchspersonen verarbeitet hatten, je brauchbarer sie die bevorzugte Lösung einschätzten und je unterschiedlicher sie die beste und die zweitbesten Alternative wahrnahmen. Der Zusammenhang zwischen Verarbeitungsschritten und Urteilssicherheit entsprach der Vorhersage des Contingency Model for the Selection of Decision Strategies (Beach & Mitchell, 1978), das den geleisteten analytischen Aufwand als bestimmend für die Urteilssicherheit betrachtet. Die Ergebnisse widerlegten die Support Theory (Tversky & Koehler, 1994), die die relative argumentative Unterstützung für die bevorzugte Alternative als urteilssicherheitsbestimmend ansieht und deshalb einen negativen Zusammenhang zwischen der Brauchbarkeitseinschätzung und der Urteilssicherheit vorhersagte. Erfüllt wurde dagegen die Erwartung des Model of Belief Processing (Curley & Benson, 1994), es bestehe ein enger positiver Zusammenhang zwischen der Brauchbarkeitseinschätzung und der Urteilssicherheit. Die kumulative Stärke der Argumente beeinflusste damit die Urteilssicherheit dem Belief Processing Modell entsprechend. Insgesamt zeigte sich, dass die erfahreneren Versuchspersonen sowohl prozessbezogene – Verarbeitungsschritte – als auch erkenntnisbezogene Aspekte – Brauchbarkeitseinschätzung und Differenzwahrnehmung – zur Bildung der Urteilssicherheit heranzogen.

Teil III: Fazit