

Verhaltenswirkungen von Lesbarkeit und einer Fremdsprache im Management Reporting: Experimentelle Studien

INAUGURALDISSERTATION

zur

teilweisen Erlangung der Würde eines Doktors der Wirtschaftswissenschaft,
Doctor rerum oeconomicarum (Dr. rer. oec.)

der

Fakultät für Wirtschaftswissenschaft
– Schumpeter School of Business and Economics

der

Bergischen Universität Wuppertal

vorgelegt von

Ewgenij Besuglov, M. Sc.

Wuppertal, im August 2021

Vorgelegt bei dem Promotionsausschuss, vertreten durch Prof. Dr. Stefan Bock, Fakultät für Wirtschaftswissenschaft – Schumpeter School of Business and Economics der Bergischen Universität Wuppertal zur teilweisen Erlangung der Würde eines Doktors der Wirtschaftswissenschaft, Doctor rerum oeconomicarum (Dr. rer. oec.) gemäß der Prüfungsordnung vom 28.08.2019.

Dekan der Fakultät für Wirtschaftswissenschaft – Schumpeter School of Business and Economics: Prof. Dr. Nils Crasselt

1. Gutachter:

Prof. Dr. Nils Crasselt

Lehrstuhl für Controlling

Fakultät für Wirtschaftswissenschaft – Schumpeter School of Business and Economics
Bergische Universität Wuppertal

2. Gutachter:

Prof. Dr. André Betzer

Lehrstuhl für Finanzwirtschaft und Corporate Governance

Fakultät für Wirtschaftswissenschaft – Schumpeter School of Business and Economics
Bergische Universität Wuppertal

Die Dissertation kann wie folgt zitiert werden:

urn:nbn:de:hbz:468-20220405-094209-3

[<http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade%3A468-20220405-094209-3>]

DOI: 10.25926/gbs1-c614

[<https://doi.org/10.25926/gbs1-c614>]

Vorwort

Die vorliegende Arbeit wurde im September 2021 von der Fakultät für Wirtschaftswissenschaft – Schumpeter School of Business and Economics der Bergischen Universität Wuppertal als Dissertation angenommen. An dieser Stelle möchte ich die Gelegenheit nutzen, um all denen zu danken, die mich bei der Entstehung dieser Arbeit unterstützt haben.

Mein besonderer Dank gebührt meinem Doktorvater Prof. Dr. Nils Crasselt. Er hat mich in allen Phasen der Entstehung dieser Arbeit begleitet und gefördert. Stets hatte er ein offenes Ohr für meine Belange und ich konnte mich auf ihn als einen kompetenten sowie engagierten akademischen Lehrer verlassen. Seine herzliche und aufmunternde Art hat die Jahre am Lehrstuhl für Controlling nicht nur fachlich, sondern auch persönlich zu einer besonders schönen Zeit gemacht.

Prof. Dr. André Betzer hat im Rahmen des Promotionsverfahrens das Zweitgutachten übernommen. Für die freundliche Betreuung und seine konstruktiven Hinweise bei der Begutachtung meiner Dissertation möchte ich auch ihm sehr herzlich danken.

Dank der ausgesprochen kollegialen und freundschaftlichen Zusammenarbeit mit meinen Kolleginnen Helene Nickel-Poulkas und Rebekka Hilser fühlte ich mich von Anfang meiner Mitarbeiterzeit sehr wohl. Nicht zuletzt auch ihnen verdanke ich aufgrund fachlicher Diskussionen und weiterer Unterstützung das Gelingen meiner Forschungsprojekte. Mit großer Freude werde ich mich an die vielen schönen Momente mit euch innerhalb und außerhalb der Arbeit zurückerinnern.

Weil die angewandte empirische Forschung stets mit einem nicht zu vernachlässigenden organisatorischen Aufwand verbunden ist, war ich auf die tatkräftige und zuverlässige Hilfe bei der Durchführung meiner Experimente angewiesen. Dafür möchte ich mich insbesondere bei unseren ehemaligen studentischen Hilfskräften am Lehrstuhl Timo Hübel und Kajatheepan Navanathan sowie vielen weiteren Kolleginnen und Kollegen bedanken.

Im Juli 2019 hatte ich die Gelegenheit, einen Teil meiner Forschungsergebnisse auf der Summer School „Experimental Research in Management Accounting“ in Bayreuth zu präsentieren. Den dort anwesenden Professorinnen und Professoren sowie Doktorandinnen und Doktoranden danke ich für die hilfreichen Anmerkungen und Verbesserungsvorschläge.

Meinen Eltern Natalia und Oleg Besuglov, meinem Bruder Daniil Besuglov sowie meiner Frau Elena Kalinina verdanke ich mehr als allen anderen. Sie haben immer an mich geglaubt und mich in allen meinen Unterfangen bestärkt. Ich werde beispielsweise nie vergessen, wie meine Eltern buchstäblich dafür gekämpft haben, dass ich in der achten Klasse von der Hauptschule auf das Gymnasium wechseln durfte. Meiner liebevollen Frau – zu Beginn der Promotionszeit war sie noch meine Freundin – danke ich einerseits für ihre unvorstellbare Geduld und andererseits dafür, dass sie mir immer wieder vor

Augen geführt hat, dass es neben der Dissertation auch andere wichtige und schöne Dinge im Leben gibt. Die interdisziplinäre Ausrichtung meiner Doktorarbeit erlaubte es, dass Elena mit ihrer linguistischen Expertise mich fachlich sehr stark und gern unterstützt hat. Ganz lieben Dank dafür!

Meine beiden Großväter kannten sich seit ihrer Jugend, waren beide (in unterschiedlichen Gewichtsklassen) ukrainische Landesmeister im Boxen und sind leider beide im Jahr 2021 verstorben. Meine Großmütter durchleben zurzeit die Schrecken des Krieges in der Ukraine – und wir gedanklich mit ihnen. Meinen Großeltern möchte ich diese Arbeit widmen. Мир та свободу Україні!

Wuppertal, im März 2022

Ewgenij Besuglov

Inhaltsverzeichnis

| | |
|---|-----------|
| Abbildungsverzeichnis | IX |
| Tabellenverzeichnis | X |
| Abkürzungsverzeichnis | XI |
| 1. Einleitung | 1 |
| 1.1 Themenbegründung und Zielsetzung | 1 |
| 1.2 Gang der Untersuchung | 4 |
| 2. Theoretische Grundlagen | 5 |
| 2.1 Betriebswirtschaftliche Entscheidungstheorie..... | 5 |
| 2.1.1 Normative Entscheidungstheorie..... | 5 |
| 2.1.2 Deskriptive Entscheidungstheorie | 12 |
| 2.2 Dual-System-Theorien..... | 23 |
| 2.2.1 System 1 und System 2..... | 23 |
| 2.2.2 Heuristiken und Biases | 26 |
| 2.2.2.1 Zusammenhang mit System 1 und System 2..... | 26 |
| 2.2.2.2 Repräsentativitätsheuristik..... | 29 |
| 2.2.2.3 Verankerungsheuristik..... | 32 |
| 2.2.2.4 Verfügbarkeitsheuristik | 34 |
| 2.2.3 Schnelle und sparsame Heuristiken..... | 37 |
| 2.2.3.1 Charakteristika und Einordnung..... | 37 |
| 2.2.3.2 Take-the-Best-Heuristik | 39 |
| 2.2.3.3 Rekognitionsheuristik..... | 42 |
| 2.3 Ausgewählte kognitive Spracheffekte | 44 |
| 2.3.1 Lesbarkeit | 44 |
| 2.3.2 Fremdsprache..... | 49 |
| 2.3.2.1 Brain-Drain-Modell..... | 49 |
| 2.3.2.2 Foreign-Language-Effekt | 52 |

| | | |
|-----------|---|-----------|
| 2.3.2.2.1 | Wirkungsbereiche | 52 |
| 2.3.2.2.2 | Erklärungsansätze und Rolle der Sprachkompetenz..... | 56 |
| 2.3.2.2.3 | Offene Fragen | 59 |
| 2.4 | Management Reporting | 62 |
| 2.4.1 | Begriffsabgrenzung und Aufgaben..... | 62 |
| 2.4.2 | Gestaltungsdimensionen | 65 |
| 2.4.2.1 | Zweck | 66 |
| 2.4.2.2 | Inhalt..... | 67 |
| 2.4.2.3 | Träger und Empfänger..... | 68 |
| 2.4.2.4 | Zeit..... | 69 |
| 2.4.2.5 | Art und Form | 70 |
| 2.4.3 | Störungen und Verhaltenswirkungen | 72 |
| 2.4.4 | Vermeidung von unerwünschten Verhaltenswirkungen..... | 78 |
| 2.5 | Experimentelle Forschung | 83 |
| 2.5.1 | Einordnung | 83 |
| 2.5.2 | Merkmale und Designarten | 85 |
| 2.5.3 | Gütekriterien..... | 87 |
| 3. | Experiment 1: Risikoverhalten | 91 |
| 3.1 | Einleitung | 91 |
| 3.2 | Hypothesen | 93 |
| 3.2.1 | Hypothese 1 | 93 |
| 3.2.2 | Hypothese 2 | 94 |
| 3.2.3 | Hypothese 3 | 96 |
| 3.3 | Experimentdesign und Teilnehmer..... | 97 |
| 3.3.1 | Experimentdesign | 97 |
| 3.3.1.1 | Ablauf und Aufgabe | 97 |
| 3.3.1.2 | Unabhängige Variablen und Kontrollvariablen..... | 99 |
| 3.3.1.3 | Abhängige Variablen..... | 101 |
| 3.3.2 | Teilnehmer..... | 102 |
| 3.4 | Ergebnisse..... | 104 |
| 3.4.1 | Deskriptive Statistiken..... | 104 |

| | | |
|-----------|--|------------|
| 3.4.2 | Hypothesentests | 106 |
| 3.4.2.1 | H1 und H2 | 106 |
| 3.4.2.2 | H3 | 108 |
| 3.4.3 | Ergänzende Analysen | 110 |
| 3.4.3.1 | Manipulation-Checks und alternative Erklärungen..... | 110 |
| 3.4.3.2 | Überprüfung der Erklärungsansätze für den FLE..... | 111 |
| 3.5 | Diskussion der Ergebnisse..... | 114 |
| 4. | Experiment 2: Ethisches Verhalten | 117 |
| 4.1 | Einleitung | 117 |
| 4.2 | Hypothesen | 119 |
| 4.2.1 | Hypothese 1 | 119 |
| 4.2.2 | Hypothese 2 | 119 |
| 4.2.3 | Hypothese 3 | 121 |
| 4.3 | Experimentdesign und Teilnehmer..... | 121 |
| 4.3.1 | Experimentdesign | 121 |
| 4.3.1.1 | Ablauf und Aufgabe | 121 |
| 4.3.1.2 | Unabhängige Variablen und Kontrollvariablen..... | 123 |
| 4.3.1.3 | Abhängige Variablen..... | 124 |
| 4.3.2 | Teilnehmer..... | 124 |
| 4.4 | Ergebnisse..... | 126 |
| 4.4.1 | Deskriptive Statistiken..... | 126 |
| 4.4.2 | Hypothesentests | 128 |
| 4.4.2.1 | H1 und H2 | 128 |
| 4.4.2.2 | H3 | 129 |
| 4.4.3 | Ergänzende Analysen | 131 |
| 4.4.3.1 | Manipulation-Checks und alternative Erklärungen..... | 131 |
| 4.4.3.2 | Überprüfung der Erklärungsansätze für den FLE..... | 132 |
| 4.5 | Diskussion der Ergebnisse..... | 135 |
| 5. | Zusammenfassung und Ausblick | 137 |

| | |
|-----------------------------------|------------|
| Anhang I | 141 |
| Anhang II | 142 |
| Anhang III | 145 |
| Literaturverzeichnis | 148 |

Abbildungsverzeichnis

| | |
|---|-----|
| Abbildung 1: Beispiel für das Allais-Paradoxon..... | 11 |
| Abbildung 2: Wertfunktion der PT..... | 16 |
| Abbildung 3: Entscheidungsgewichtungsfunktion der PT | 18 |
| Abbildung 4: Vierfeldmuster der PT | 21 |
| Abbildung 5: Metakognitive Auslöser von System 2..... | 25 |
| Abbildung 6: Passende Entscheidungswerkzeuge bei Risiko und Ungewissheit..... | 39 |
| Abbildung 7: Ablauf der Take-the-Best-Heuristik..... | 40 |
| Abbildung 8: Grad der Lesbarkeit und kognitive (Lern-)Performance | 48 |
| Abbildung 9: Brain-Drain-Modell..... | 50 |
| Abbildung 10: Aufgaben im Management-Reporting-Prozess | 64 |
| Abbildung 11: Gestaltungsdimensionen des Berichtswesens | 65 |
| Abbildung 12: Verhaltenswirkungen von Berichten..... | 74 |
| Abbildung 13: Visualisierung von H1 und H2 (Experiment 1) | 96 |
| Abbildung 14: Interaktion zwischen Sprache und Lesbarkeit..... | 108 |
| Abbildung 15: Test des Kausalmodells (hohe Lesbarkeitsbedingung) | 113 |
| Abbildung 16: Visualisierung von H1 und H2 (Experiment 2) | 121 |

Tabellenverzeichnis

| | |
|--|-----|
| Tabelle 1: Intransitive Lotterien | 10 |
| Tabelle 2: Empirische Parameterschätzungen der PT | 20 |
| Tabelle 3: Ausgewählte Eigenschaften von System 1 und System 2 | 24 |
| Tabelle 4: Generelle vs. strategische Risikobereitschaft beim FLE | 60 |
| Tabelle 5: Ausgewählte Biases im Controlling | 76 |
| Tabelle 6: Deskriptive Statistiken zu den Teilnehmern (Experiment 1) | 103 |
| Tabelle 7: Deskriptive Statistiken zu den abhängigen Variablen (Experiment 1) | 105 |
| Tabelle 8: Statistische Analyseergebnisse – Teil 1 (Experiment 1) | 107 |
| Tabelle 9: Statistische Analyseergebnisse – Teil 2 (Experiment 1) | 109 |
| Tabelle 10: Deskriptive Statistiken zu den Teilnehmern (Experiment 2) | 125 |
| Tabelle 11: Deskriptive Statistiken zu den abhängigen Variablen (Experiment 2) | 127 |
| Tabelle 12: Statistische Analyseergebnisse – Teil 1 (Experiment 2) | 129 |
| Tabelle 13: Statistische Analyseergebnisse – Teil 2 (Experiment 2) | 130 |
| Tabelle 14: Regressionskoeffizienten zur Messung ethischer Orientierung | 134 |
| Tabelle 15: Beispiele für Lesbarkeitsmanipulationen (Experiment 1) | 142 |
| Tabelle 16: Ausschlusskriterien (Experiment 1) | 143 |
| Tabelle 17: ANCOVA – Volles Modell mit Rohstichprobe (Experiment 1) | 144 |
| Tabelle 18: Beispiele für Lesbarkeitsmanipulationen (Experiment 2) | 145 |
| Tabelle 19: Ausschlusskriterien (Experiment 2) | 146 |
| Tabelle 20: ANCOVA – Volles Modell mit Rohstichprobe (Experiment 2) | 147 |

Abkürzungsverzeichnis

| | |
|---------------|------------------------------------|
| ANCOVA | Kovarianzanalyse |
| ANOVA | Varianzanalyse |
| BDM | Brain-Drain-Modell |
| c. p. | ceteris paribus |
| CLE | Cognitive-Load-Effekt |
| CPT | Cumulative-Prospect-Theorie |
| CRT | Cognitive-Reflection-Test |
| DFE | Disfluency-Effekt |
| DST | Dual-System-Theorien |
| EU | Expected Utility |
| EW | Erwartungswert |
| FF-Heuristik | Fast-and-Frugal-Heuristik |
| Fg | Freiheitsgrade |
| FLE | Foreign-Language-Effekt |
| HB | Heuristiken und Biases |
| M | Mittelwert |
| MBA | Master of Business Administration |
| PT | Prospect-Theorie |
| SEC | Securities and Exchange Commission |
| TTB-Heuristik | Take the Best-Heuristik |

1. Einleitung¹

1.1 Themenbegründung und Zielsetzung

„*Wouldn't economics make a lot more sense if it were based on how people actually behave, instead of how they should behave?*“²

Das Management Accounting³ soll zur Steigerung des Unternehmenswerts beitragen, indem ein effektiver und effizienter Einsatz von knappen Ressourcen sichergestellt wird.⁴ Dieser Zielsetzung entspringen zwei wesentliche Funktionen von Controlling-Informationen: Entscheidungsfunktion (*decision-facilitating*) und Verhaltenssteuerungsfunktion (*decision-influencing*).⁵

Im Rahmen der Entscheidungsfunktion sollen Informationen bereitgestellt werden, die für die Planungs- und Entscheidungsaufgaben notwendig sind.⁶ Dabei steht vor allem die Frage im Fokus, welche Informationen einem Entscheider in einer bestimmten Situation zur Verfügung gestellt werden sollten, um den Entscheidungsfindungsprozess zu erleichtern und zu verbessern.⁷

Im Hinblick auf die Verhaltenssteuerungsfunktion sollen Controlling-Informationen helfen, das Verhalten der mit Führungsaufgaben betrauten Akteure an den Unternehmenszielen auszurichten.⁸ Dies ist notwendig, weil zum einen angenommen wird, dass Manager potenziell andere Ziele verfolgen als diejenigen, die ihnen für ihre Positionen von der Unternehmung vorgegeben werden (*Wollensprobleme*). Zum anderen wird – abweichend vom standard-ökonomischen Bild des rational handelnden Homo oeconomicus – davon ausgegangen, dass Manager kognitiven Beschränkungen unterliegen und folglich falsche Entscheidungen treffen können (*Könnensprobleme*). Dem Controlling obliegt es im Rahmen seiner Rationalitätssicherungsfunktion, die negativen Folgen dieser verhaltensbezogenen Wollens- und Könnensprobleme zu reduzieren bzw. im Idealfall ganz zu vermeiden.⁹

¹ Aus Gründen der besseren Lesbarkeit wird bei Personenbezeichnungen und personenbezogenen Hauptwörtern in der vorliegenden Arbeit ausschließlich die männliche Form verwendet. Im Sinne der Gleichbehandlung gelten entsprechende Begriffe grundsätzlich für alle Geschlechter. Die verkürzte Sprachform hat nur redaktionelle Gründe und beinhaltet keine Wertung.

² Ariely 2008, S. 239.

³ Management Accounting weist große Ähnlichkeiten mit dem deutschen Controlling auf, vgl. Hirsch 2008, S. 40. Daher werden die beiden Begriffe im Verlauf dieser Arbeit synonym verwendet.

⁴ Vgl. Sprinkle und Williamson 2007, S. 415.

⁵ Vgl. Demski und Feltham 1976, S. 8f., übersetzt nach Schwering und Uepping 2017, S. 158.

⁶ Vgl. Zimmerman 2017, S. 3.

⁷ Vgl. Sprinkle und Williamson 2007, S. 427f.

⁸ Vgl. Schwering und Uepping 2017, S. 158.

⁹ Vgl. Weber 2013, S. 220.

Forschungsarbeiten im Bereich der Verhaltenssteuerung lassen sich dem Behavioral Management Accounting – als Teildisziplin des Behavioral Accounting – zuordnen.¹⁰ Das Behavioral Accounting bedient sich neben ökonomischer Theorien vor allem psychologischer und soziologischer Ansätze und zeichnet sich somit durch seine hohe Interdisziplinarität aus.¹¹ In der verhaltenswissenschaftlichen Forschung wird Experimenten als Forschungsmethode ein besonderer Stellenwert eingeräumt.¹² Der Hauptvorteil der experimentellen Methode gegenüber anderen empirischen Forschungsansätzen besteht darin, dass eine isolierte Betrachtung der kausalen Beziehung zwischen verschiedenen Einflussfaktoren und zu untersuchenden abhängigen Variablen ermöglicht wird.¹³ Die vorliegende Arbeit knüpft an die Tradition des Behavioral Accounting an, indem der Einfluss von Lesbarkeit und einer Fremdsprache im Management Reporting experimentell untersucht wird.

Lesbarkeit ist ein allgegenwärtiges Anliegen nicht nur in der Rechnungslegungspraxis, sondern auch in vielen anderen Bereichen wie Militär, Gesundheitswesen und Recht.¹⁴ Die im Jahr 1998 von der US-amerikanischen Securities and Exchange Commission (SEC) verabschiedete Plain English Rule 421 (d) wurde initiiert, um die Emittenten von Finanzinformationen zur Verwendung klarer englischer Sprache zu verpflichten.¹⁵ Allein die Existenz einer solchen Regelung unterstreicht deutlich die Bedeutung der Lesbarkeit für die finanzielle Berichterstattung. Der Konsens in der Accounting-Forschung besteht darin, dass weniger lesbare Finanzdokumente negative Folgen für die Entscheidungsträger – wie z. B. erhöhte Verarbeitungsschwierigkeiten – nach sich ziehen.¹⁶

¹⁰ Zu den weiteren Teildisziplinen des Behavioral Accounting zählen: Behavioral Financial Accounting und Behavioral Tax Accounting, vgl. Gillenkirch und Arnold 2008, S. 129. Während das Behavioral Accounting in den angelsächsischen Ländern seit langem als etabliert gilt, fand die Einbeziehung der verhaltenswissenschaftlichen Perspektive im deutschsprachigen Raum bis in die späten 2000er-Jahre kaum statt, vgl. Hirsch 2008, S. 40.

¹¹ Einen Überblick über psychologische und soziologische Ansätze im Behavioral Accounting geben u. a. Birnberg und Shields 1989.

¹² Insbesondere die angelsächsische Forschung nimmt mit zahlreichen, in den letzten Jahrzehnten durchgeführten Experimenten eine Vorreiterrolle in diesem Bereich ein, während in der deutschsprachigen Controlling-Forschung die experimentelle Methode lange Zeit wenig beachtet wurde, vgl. Hirsch 2009, S. 167f.

¹³ Vgl. Gillenkirch und Arnold 2008, S. 130.

¹⁴ Vgl. Bonsall IV et al. 2017, S. 330.

¹⁵ Vgl. Rennekamp 2012, S. 1320.

¹⁶ Vgl. Tan et al. 2014, S. 280.

Dies kann sich in erschwertem Verständnis und reduzierter Bereitschaft, relevante Informationen zu extrahieren, äußern.¹⁷ Aber auch ohne Verständnisprobleme kann niedrige Lesbarkeit die kognitiven Prozesse beeinflussen, indem bspw. der Entscheider empfänglicher für heuristische Impulse wie Gefühle wird.¹⁸

Englisch gilt heute als die Lingua franca in fast allen Bereichen der globalisierten Welt wie Politik, Kultur und Wirtschaft.¹⁹ Einige multinationale Unternehmen aus nicht anglophonen Ländern wie bspw. Nissan oder Honda planen die Einführung der englischen Sprache als ihre offizielle Konzernsprache oder setzten dies bereits um.²⁰ Das bedeutet, dass die Mitarbeiter in solchen Organisationen in einer für sie fremden Sprache kommunizieren und folglich ihre Entscheidungen auf Grundlage dieser Sprache treffen müssen. Eine einheitliche Unternehmenssprache bringt viele Vorteile, daraus können aber auch einige potenzielle Probleme resultieren: Auf der einen Seite könnten die Empfänger von Informationen in einer Fremdsprache aufgrund mangelnder Sprachkompetenz mit Verarbeitungsschwierigkeiten konfrontiert sein, auf der anderen Seite stehen Mitarbeiter, die für die Aufbereitung von Informationen zuständig sind (wie z. B. Controller), häufig vor der Herausforderung, eine hohe Lesbarkeit in der fremdsprachlichen Berichterstattung zu gewährleisten.

Neuere psychologische Studien zeigen, dass sowohl die kognitive Belastung (z. B. infolge niedriger Lesbarkeit) als auch der Einsatz einer Fremdsprache das Entscheidungsverhalten beeinflussen können. Das Ziel der vorliegenden Arbeit ist es, den Einfluss dieser linguistischen Determinanten auf das Risikoverhalten und das ethische Verhalten speziell im Controlling-Kontext zu untersuchen.²¹

Die vorliegende Arbeit trägt zum wissenschaftlichen Erkenntnisfortschritt wie folgt bei: Für die Controlling-Forschung und -Praxis sind die Erkenntnisse aus dieser Arbeit insofern relevant, als einige nicht offensichtliche sprachbezogene Verhaltensphänomene im Management Reporting mit potenziell weitreichenden Konsequenzen aufgezeigt werden können. Aus Sicht der Forschung zu kognitiven Einflüssen auf das Entscheidungsverhalten verspricht die gemeinsame Untersuchung der kognitiven Belastung (infolge niedriger Lesbarkeit) und des Einflusses einer Fremdsprache und damit die Verknüpfung von zwei bislang separat untersuchten Effekten neue Erkenntnisse. Darüber hinaus handelt es sich bei den im Rahmen dieser Arbeit durchgeführten Experimenten – nach dem besten Kenntnisstand des Autors – um die ersten strukturierten Untersu-

¹⁷ Vgl. Bloomfield 2002, S. 233ff.

¹⁸ Vgl. Tan et al. 2014, S. 285ff.

¹⁹ Vgl. Tietze 2008, S. 72.

²⁰ Siehe u. a. Borzykowski 2017.

²¹ Die Begründung der Relevanz des Risikoverhaltens und des ethischen Verhaltens für das Controlling erfolgt jeweils in Kapiteln 3.1 und 4.1.

chungen, in denen führende Erklärungsansätze für den Spracheffekt gleichzeitig getestet werden und die Sprachkompetenz der Teilnehmer mittels eines objektiven Tests erhoben wird.

1.2 Gang der Untersuchung

Die vorliegende Arbeit ist wie folgt aufgebaut: In Kapitel 2 werden zunächst relevante theoretische Grundlagen erläutert, die für das Verständnis der experimentellen Studien im empirischen Teil dieser Arbeit wichtig sind. Dabei wird im ersten Schritt (Kapitel 2.1) der Unterschied zwischen dem normativen und dem deskriptiven Ansatz der betriebswirtschaftlichen Entscheidungstheorie beleuchtet. Im Zusammenhang mit den Dual-System-Theorien wird in Kapitel 2.2 auf das Konzept vom Denken in zwei sog. Systemen eingegangen, das eine zentrale Rolle bei der Ableitung von Hypothesen im empirischen Teil dieser Arbeit spielt. Dabei zeichnet sich das kognitiv anstrengende System 2 durch analytische Entscheidungsfindung aus. Die Eigenschaften des heuristischen Systems 1 führen (unter bestimmten Umständen) zu falschen bzw. systematisch verzerrten Urteilen. Im weiteren Verlauf werden in Kapitel 2.3 ausgewählte kognitive Effekte der Lesbarkeit und Fremdsprache präsentiert. Schließlich runden Kapitel 2.4 mit Gestaltungsdimensionen und Verhaltenswirkungen des Management Reporting sowie Kapitel 2.5 mit der Vorstellung der experimentellen Forschungsmethode den Grundlagenteil dieser Arbeit ab.

Kapitel 3 beinhaltet die erste experimentelle Untersuchung mit dem Schwerpunkt auf dem Risikoverhalten. Nach einer kurzen Einleitung in Kapitel 3.1 erfolgt im nachfolgenden Kapitel 3.2 die Ableitung der zu prüfenden Hypothesen. Das Experimentdesign und die Teilnehmer werden in Abschnitt 3.3 vorgestellt, bevor auf die Ergebnisse des Experiments in Kapitel 3.4 eingegangen wird. Das Oberkapitel schließt in Abschnitt 3.5 mit der Diskussion der gewonnenen Resultate.

Das zweite Experiment zum ethischen Verhalten wird in Kapitel 4 behandelt, das identisch strukturiert ist wie das vorangehende Oberkapitel.

Die Arbeit endet in Kapitel 5 mit einer thesenförmigen Zusammenfassung der zentralen Ergebnisse sowie einem Ausblick auf mögliche Anknüpfungspunkte für weitere Forschungsbemühungen.

2. Theoretische Grundlagen

2.1 Betriebswirtschaftliche Entscheidungstheorie

2.1.1 Normative Entscheidungstheorie

Entscheidungen von leitenden Angestellten haben oft weitreichende Konsequenzen und können über den Erfolg oder Misserfolg des ganzen Unternehmens entscheiden.²² Die Frage, wie sich die Wirtschaftsakteure formal-logisch verhalten sollten bzw. wie sie sich tatsächlich in der Realität verhalten, ist eines der zentralen Anliegen der wirtschaftswissenschaftlichen Forschung und wird im Rahmen der Entscheidungstheorie untersucht.²³ Entsprechend den Annahmen über die Entscheidungsprämissen der handelnden Individuen gliedert sich die Entscheidungstheorie in einen normativen und einen deskriptiven Ansatz.

In der *normativen Entscheidungstheorie*²⁴ wird ausgehend von gegebenen formal-logischen Entscheidungsprämissen abgeleitet, wie eine optimale Entscheidung getroffen werden „soll“. Dabei wird der eigentliche Entscheidungsprozess formalisiert, um eine systematische Behandlung von Entscheidungssituationen zu ermöglichen.²⁵ Im Zentrum der Betrachtung steht ein Entscheider,²⁶ der innerhalb der gegebenen Umwelt die ihm zur Verfügung stehenden Handlungsoptionen und die daraus resultierenden Konsequenzen bestimmt. Die Konsequenzen werden anschließend entsprechend den individuellen Präferenzen und in Übereinstimmung mit der formalen Entscheidungslogik in eine optimale Entscheidung integriert.²⁷ Die formale Entscheidungslogik fußt auf bestimmten Axiomen zur Präferenzordnung des Entscheidungsträgers. Die einzelnen Axiome werden im weiteren Verlauf dieses Abschnitts noch näher erläutert.

²² Beispielsweise gehören Investitionsentscheidungen aufgrund ihrer Häufigkeit und ihrer Auswirkungen weit in die Zukunft zu den wichtigsten Entscheidungen in den Unternehmen, vgl. Weber et al. 2003, S. 11; Weber 2014, S. 33.

²³ An dieser Stelle wird aus Vereinfachungsgründen unterstellt, dass die Entscheidungsträger in den Unternehmen stets im Sinne ihres Unternehmens handeln und es keine Interessenkonflikte gibt. Interessenkonflikte und Informationsasymmetrie werden in der von Jensen und Meckling 1976 begründeten Prinzipal-Agent-Theorie berücksichtigt. Einen kurzen und gut zugänglichen Überblick zu dieser Thematik geben u. a. Hochhold und Rudolph 2009.

²⁴ In der Literatur ist die Bezeichnung präskriptive Entscheidungstheorie ebenso geläufig, siehe z. B. Laux et al. 2014. Die normative Entscheidungstheorie gehört zur Familie der neoklassischen Theorie bzw. der ökonomischen Standardtheorie.

²⁵ Vgl. Langer und Rogowski 2009, S. 179.

²⁶ Die individuelle Sichtweise wird im Rahmen der Spieltheorie um strategische Interaktionen zwischen mehreren Akteuren erweitert. Einen kurzen und gut zugänglichen Überblick zu dieser Thematik gibt u. a. Meyer 2009.

²⁷ Vgl. Lang 2009, S. 163f.

Die Grundzüge der normativen Entscheidungstheorie werden in dieser Arbeit anhand des *Erwartungsnutzenmodells*²⁸ (engl. Expected Utility Model, kurz EU-Modell) skizziert. Das EU-Modell basiert auf dem wahrscheinlichkeitstheoretischen Konzept des Erwartungswerts.²⁹ Ferner wird angenommen, dass jede Entscheidungsalternative als eine abstrakte Lotterie mit gegebenen Eintrittswahrscheinlichkeiten für die einzelnen Konsequenzen angesehen werden kann.³⁰ Formal wird der erwartete Gesamtnutzen³¹ einer risikobehafteten Alternative (bzw. Lotterie) A – hier EU(A) – wie folgt ermittelt:

$$EU(A) = \sum_i^n p_i \times u(a_i)$$

Mit

p_i : Eintrittswahrscheinlichkeit der Konsequenz i

$u(a_i)$: Nutzen der Konsequenz i

Analog zur subjektiven Nutzenbewertung der Konsequenzen können die Eintrittswahrscheinlichkeiten ebenfalls als subjektive Werte angesehen werden. Wenn sowohl der Nutzen als auch die Wahrscheinlichkeit als subjektive Größen angenommen werden, wird vom Modell des subjektiv erwarteten Nutzens (engl. Subjective Expected Utility Model) gesprochen.³² In der ökonomischen Entscheidungstheorie wird davon ausgegangen, dass subjektive Wahrscheinlichkeiten als „echte“ Wahrscheinlichkeiten inter-

²⁸ Dieses Prinzip wurde zum ersten Mal von Bernoulli 1738 beschrieben (eine Übersetzung seiner Arbeit ins Englische erschien in der Zeitschrift *Econometrica* im Jahr 1954). Bernoulli wollte damit das Verhalten von Akteuren im Zusammenhang mit dem Sankt-Petersburg-Paradoxon erklären.

²⁹ Vgl. Lee 1977, S. 52ff.

³⁰ Im Allgemeinen wird in der entscheidungstheoretischen Literatur zwischen Sicherheit, Risiko, Unsicherheit und Ungewissheit (auch Unsicherheit i. e. S. oder Ambiguität genannt) unterschieden. Steht die Konsequenz einer Alternative im Vorfeld fest, handelt es sich um eine Entscheidung mit sicherem Ausgang. Bei mehreren möglichen und bekannten Konsequenzen einer Alternative mit offengelegten Eintrittswahrscheinlichkeiten wird vom Risiko und bei bekannten Konsequenzen aber unbekanntem Wahrscheinlichkeiten von Unsicherheit gesprochen. Schließlich bedeutet die Ungewissheit, dass weder die Konsequenzen noch die korrespondierenden Wahrscheinlichkeiten bekannt sind, vgl. Witt 2018, S. 169f. Im empirischen Teil dieser Arbeit (siehe Kapitel 3) bildet eine Risikosituation den Schwerpunkt der Untersuchung, daher wird nachfolgend ausschließlich auf Entscheidungen unter Risiko rekuriert.

³¹ Der Ausdruck Nutzen steht für eine quantitative Beziehung zwischen einer Konsequenz und dem subjektiven Wert, den ein Entscheidungsträger dieser Konsequenz beimisst (z. B. kann dieselbe Einkommenshöhe unterschiedliche Nutzenwerte bei verschiedenen Personen stiften). Eine Nutzenfunktion gibt den funktionalen Zusammenhang zwischen den Konsequenzen und den entsprechenden subjektiven Nutzenwerten wieder. Anhand des Verlaufs der Nutzenfunktion kann die Risikoeinstellung eines Entscheiders abgeleitet werden. Ist die Nutzenfunktion linear, so wird ein risikoneutraler Entscheidungsträger unterstellt. Verläuft die Nutzenfunktion streng konkav (konvex), so wird ein Entscheidungsträger als risikoscheu (risikofreudig) bezeichnet, vgl. Laux et al. 2014, S. 117f.

³² Siehe Edwards 1954.

pretiert werden können – d. h., die subjektiven Wahrscheinlichkeiten sind mit den Regeln der Wahrscheinlichkeitstheorie konform.³³ In anderen Aspekten unterscheiden sich die beiden Konzepte nicht wesentlich voneinander, weswegen im weiteren Verlauf der Arbeit nur noch das EU-Modell behandelt wird. Bei mehreren zur Auswahl stehenden Lotterien entscheidet sich ein rational handelnder Akteur gemäß dem EU-Modell stets für diejenige Alternative, die seinen individuellen erwarteten Gesamtnutzen maximiert. Im EU-Modell wird also eine konsequentialistische Handlungsweise unterstellt.

Damit eine optimale Entscheidung vom EU-Modell auf eine konsistente Weise repräsentiert werden kann, müssen bestimmte Anforderungen an die Präferenzen des Entscheiders erfüllt sein. Diese *Axiome* stellen das Rationalitätsverständnis der normativen Entscheidungstheorie dar. Von Neumann und Morgenstern 1947 legten als Erste eine axiomatische Fundierung des EU-Modells vor.³⁴ Vier Axiome gelten heute als zentral und werden nachfolgend näher erläutert. Es seien drei Lotterien A, B und C gegeben. Des Weiteren bezeichnet $>$ eine Präferenz- und \sim steht für eine Indifferenzrelation. Die vier Axiome lauten wie folgt:³⁵

- Vergleichbarkeit: Für jedes Paar aus Lotterien A und B gilt $A > B$, $B > A$ oder $A \sim B$.

Dieses Axiom ist auch unter der Bezeichnung Vollständigkeitsaxiom bekannt und fordert, dass beliebige Lotterien miteinander verglichen werden können. Von zwei Lotterien präferiert ein Entscheidungsträger entweder die Lotterie A gegenüber der Lotterie B, die Lotterie B gegenüber der Lotterie A oder er ist zwischen A und B indifferent.

- Transitivität: Aus $A > B$ und $B > C$ folgt $A > C$.

Wenn ein Entscheider die Lotterie A der Lotterie B und die Lotterie B der Lotterie C vorzieht, dann muss er gleichzeitig die Lotterie A gegenüber der Lotterie C präferieren.

- Unabhängigkeit: Aus $A > B$ folgt $p \times A + (1 - p) \times C > p \times B + (1 - p) \times C$, für alle p und C .

Eine Präferenz zwischen zwei Lotterien A und B bleibt auch dann erhalten, wenn die beiden Lotterien mit einer dritten Lotterie C in derselben Weise verknüpft werden. Mittels Streichung von identischen Komponenten in den zusammengesetzten

³³ Vgl. Pfister et al. 2017, S. 174.

³⁴ Erweiterungen und alternative Ansätze lassen sich u. a. bei Friedman und Savage 1952; Lee 1977; Luce und Raiffa 1957 finden. Übersichtliche und gut verständliche Darstellungen der Axiome geben z. B. Eisenführ et al. 2010; Laux et al. 2014; Plous 1993.

³⁵ Vgl. im Folgenden Pfister et al. 2017, S. 178f.

Lotterien kann wieder die Beziehung $A > B$ gewonnen werden. Aus diesem Grund trägt dieses Axiom im Englischen die Bezeichnung Cancellation Axiom.³⁶

- Kontinuität: Für alle $A > B > C$ gibt es ein p derart, dass $B \sim p \times A + (1 - p) \times C$ gilt.

Das Kontinuitätsaxiom besagt, dass für eine Lotterie B , die in der Präferenzordnung zwischen A und C liegt, immer eine Kombination aus A und C mit p existiert, bei der der Entscheidungsträger indifferent wird.

Nur wenn die obigen Präferenzpostulate erfüllt sind, existiert eine eindeutige Nutzenfunktion und das EU-Modell ist gültig. In der normativen Entscheidungstheorie etablierte sich unter dem Begriff *Homo oeconomicus*³⁷ ein Idealtypus eines Entscheidungsträgers, der sein Handeln streng nach den formal-logischen Modellprämissen des EU-Modells ausrichtet. Im Speziellen hat ein solches Individuum folgende Eigenschaften:³⁸

- Eigeninteresse: Das Individuum verfolgt eigene Interessen. Eigeninteresse soll jedoch nicht zwangsläufig mit einem negativ konnotierten Egoismus gleichgesetzt werden. Der *Homo oeconomicus* begegnet seinen Mitmenschen grundsätzlich neutral und ist bereit, mit ihnen zu kooperieren, solange ein solches Verhalten dem Eigeninteresse dient.³⁹
- Rationalität: Gemäß dem ökonomischen Prinzip handelt ein Akteur rational, wenn er bei gegebenen Ressourcen eine Zielgröße maximiert bzw. mit einem Minimum an Ressourcen eine vorgegebene Zielgröße erreicht. Im weiteren Sinne kann allerdings jede Handlung nach dem formalen Rationalitätsverständnis als rational angesehen werden, solange der Entscheidungsträger seine Handlung systematisch aus den ihm zur Verfügung stehenden und bekannten Handlungsalternativen auswählt und die eingeschlagene Handlung bei Kenntnis der Ziele, der Mittel und der Umwelt zu erwarten war.⁴⁰
- Nutzenmaximierung (unter Restriktionen): Die Prämisse der Nutzenmaximierung ist zentraler Ankerpunkt im EU-Modell. Demnach maximieren die Akteure ihren

³⁶ Vgl. z. B. Plous 1993, S. 81.

³⁷ Grundlegend zum Modell des *Homo oeconomicus* und der Anwendung dieses Konzepts in der wirtschafts- und sozialwissenschaftlichen Forschung siehe u. a. Kirchgässner 2008.

³⁸ Vgl. im Folgenden Franz 2004, S. 4ff.

³⁹ Vgl. Kirchgässner 2008, S. 41ff.

⁴⁰ Vgl. Tietzel 1981, S. 121f. Aus dieser Definition kann geschlussfolgert werden, dass das Rationalitätskonzept eher einen deskriptiven Charakter hat und sich problemlos auf zahlreiche Verhaltensmodelle übertragen lässt. In Abgrenzung zur Entscheidungslogik des *Homo oeconomicus* wird in Kapitel 2.2.3 eine alternative Auffassung des Rationalitätskonzepts vorgestellt.

erwarteten Nutzen unter Nebenbedingungen bzw. Restriktionen.⁴¹ Laut der makroökonomischen Theorie stellen das verfügbare Einkommen, die auf den Märkten geltenden Preise, sowie rechtliche Rahmenbedingungen typische Restriktionen dar, die im Rahmen eines Nutzenmaximierungsprozesses simultan beachtet werden müssen.⁴²

- Feststehende Präferenzen: Die Präferenzen müssen dem Individuum bekannt sein und dürfen sich im Zeitablauf nicht bzw. nur langsam verändern.⁴³ Ferner muss die beobachtbare Präferenzordnung immer im Sinne der oben aufgezeigten Axiome formal-logisch konsistent sein.
- Vollständige Informiertheit: Schließlich kann die optimale Handlungsalternative mithilfe der Zielfunktion nur dann bestimmt werden, wenn alle relevanten Konsequenzen und die korrespondierenden Eintrittswahrscheinlichkeiten bekannt sind bzw. hinreichend genau abgeschätzt werden können. Per Annahme verfügt der Homo oeconomicus über alle für die Entscheidung relevanten Informationen ohne dabei nennenswerte Transaktions- oder Informationskosten in Kauf nehmen zu müssen.

Das EU-Modell bedient sich des künstlichen Konstrukts des Homo oeconomicus, um damit die Bestimmung von aus normativer Sicht optimalen Entscheidungen zu ermöglichen. Im Gegensatz zum deskriptiven Theorieansatz steht dabei die Beschreibung von in der realen Welt tatsächlich beobachtbaren Verhaltens- und Entscheidungsmustern nicht im Vordergrund der Betrachtung. Somit wundert es nicht, dass das Verhalten von realen Akteuren mit dem Konzept des Homo oeconomicus im Allgemeinen nicht zu vereinbaren ist.⁴⁴ Stellvertretend für diverse Abweichungen vom Idealtypus des Homo oeconomicus soll im Folgenden auf Verstöße gegen die oben vorgestellten Präferenzaxiome eingegangen werden.⁴⁵

⁴¹ Vgl. Homann und Suchanek 2005, S. 364.

⁴² Vgl. Kirchgässner 2008, S. 59ff.

⁴³ Vgl. Eichenberger 1992, S. 7f.

⁴⁴ Nicht als tatsächliches Menschenbild, sondern als ein leistungsfähiges Analyseinstrument, findet das Arbeitsmodell des Homo oeconomicus eine breite Akzeptanz in der normativen Entscheidungsforschung. In diesem Kontext führen Suchanek und Kerscher 2007 jene Bedingungen an, unter denen sich der Homo oeconomicus als ein hilfreiches Konzept bei der Lösung von Koordinationsproblemen in modernen Organisationen erweisen kann.

⁴⁵ Aus der verhaltensorientierten Perspektive betrachtet, bedarf jede Annahme im Modell des Homo oeconomicus einer kritischen Auseinandersetzung. Im weiteren Verlauf wird an passenden Stellen der Arbeit auf die Tauglichkeit dieses Konzept zur Beschreibung des Verhaltens von realen Entscheidungsakteuren Bezug genommen.

Tabelle 1: Intransitive Lotterien

| Lotterie | Gewinnwahrscheinlichkeit (p) | Gewinnbetrag (\$) | Erwartungswert (p × \$) |
|----------|------------------------------|-------------------|-------------------------|
| A | 7/24 | 5,00 | 1,46 |
| B | 8/24 | 4,75 | 1,58 |
| C | 9/24 | 4,50 | 1,69 |
| D | 10/24 | 4,25 | 1,77 |
| E | 11/24 | 4,00 | 1,83 |

Quelle: In Anlehnung an Tversky 1969, S. 33.

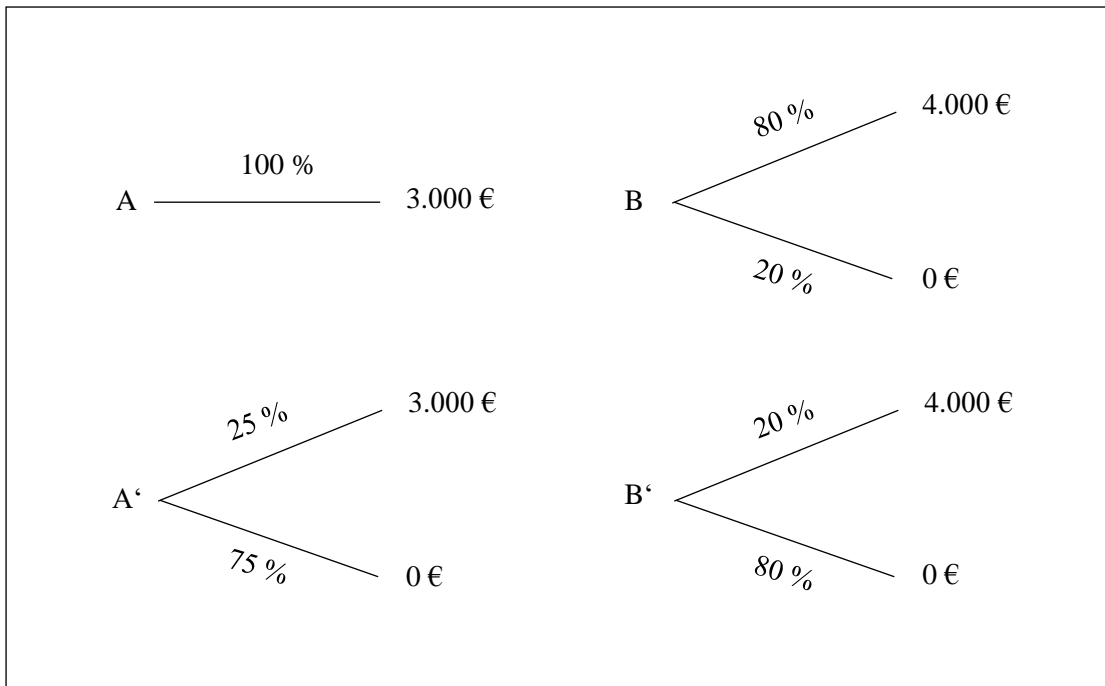
Zunächst sind zahlreiche Situationen vorstellbar, in denen es einem Entscheider schwerfallen oder gar unmöglich sein wird, zwei Alternativen miteinander zu vergleichen.⁴⁶ Solche Situationen würden zur Verletzung des Vergleichbarkeitsaxioms führen.

Eine systematische Verletzung des Transitivitätsaxioms wies Tversky 1969 nach. Der Autor verwendete in seiner über mehrere Wochen angelegten Studie fünf Lotterien, deren Gewinnwahrscheinlichkeiten, Gewinnbeträge und daraus resultierende Erwartungswerte in Tabelle 1 abgebildet sind. Die 18 Studienteilnehmer mussten insgesamt dreimal zwischen allen Lotterien wählen; dabei wurde jedem Probanden jeweils ein Lotterienpaar in zufälliger Reihenfolge vorgelegt. Acht Personen fielen durch intransitives Verhalten auf: Wenn zwei Lotterien sehr ähnliche Gewinnwahrscheinlichkeiten aufwiesen (z. B. Lotterie A und B), dann orientierten sich diese Probanden bei ihrer Wahl am höheren Gewinnbetrag (hier Lotterie A). Waren dagegen die Gewinnwahrscheinlichkeiten sehr unterschiedlich (z. B. Lotterie A und E), dann wurde die Alternative mit der höheren Gewinnwahrscheinlichkeit präferiert (hier Lotterie E). So konnte Tversky 1969 intransitive Präferenzen empirisch nachweisen (im vorliegenden Beispiel $A > B$, $B > C$, $C > D$, $D > E$ aber $E > A$).

Das Unabhängigkeitsaxiom – das besagt, dass die Präferenzrelation zwischen zwei Lotterien unabhängig von der Ergänzung dieser um gemeinsame (und somit irrelevante) Konsequenzen sein muss – wird ebenfalls unter bestimmten Voraussetzungen von realen Entscheidungsträgern systematisch verletzt. Allais 1953 zeigte als einer der Ersten, bei welcher Art von Ergänzungen zweier Lotterien um eine weitere dritte Lotterie es

⁴⁶ Vgl. Eisenführ et al. 2010, S. 250f.; Pfister et al. 2017, S. 179.

Abbildung 1: Beispiel für das Allais-Paradoxon



Anmerkung: $A' = 0,25 \times A + 0,75 \times C$ und $B' = 0,25 \times B + 0,75 \times C$, C entspricht 0 € mit Sicherheit.

Quelle: In Anlehnung an Eisenführ et al. 2010, S. 403.

zur Umkehrung der Präferenzen kommt. Das Phänomen einer auf solche Weise induzierten Präferenzänderung wird seitdem als das Allais-Paradoxon bezeichnet. Im Folgenden wird ein Zahlenbeispiel präsentiert, mit dessen Hilfe sich ein mit dem Unabhängigkeitspostulat unverträgliches Verhalten nachweisen lässt. In Abbildung 1 sind die Lotterien A und B bzw. A' und B' dargestellt, wobei die letzteren beiden Lotterien aus einer äquivalenten Verknüpfung von A und B mit einer weiteren Lotterie C resultieren. Bei der Wahl zwischen A und B würden sich die meisten Personen für den sicheren Gewinn der Option A entscheiden, um das Risiko, bei der Alternative B leer auszugehen, zu vermeiden. Da die Gewinnbeträge bei beiden Lotterien A' und B' mit einem vergleichbaren Risiko verbunden sind, ist eine Präferenzumkehrung zugunsten der Alternative B' mit dem höheren Gewinnbetrag bzw. Erwartungswert zu erwarten. Die Präferenzordnungen $A > B$ und $A' < B'$ sind zwar intuitiv nachvollziehbar, verletzen aber das Unabhängigkeitsaxiom.⁴⁷

Schließlich ist auch das Kontinuitätsaxiom nicht immer haltbar. Insbesondere wenn in der Rangordnung $A > B > C$ die rangniedrigste Option für den Entscheidungsträger

⁴⁷ Die verhaltenstheoretische Fundierung dieses Phänomens leistet die Prospect-Theorie, die in Kapitel 2.1.2 noch genauer beschrieben wird.

besonders aversiv ist, lässt sich praktisch für keine noch so geringe Eintrittswahrscheinlichkeit p eine sinnvolle Kombination aus A und C finden, bei der der Entscheidungsträger indifferent gegenüber der Alternative B wird.⁴⁸

Trotz der Versuche, das EU-Modell mit einer gelockerten Axiomatisierung einiger Anforderungen an die Präferenzpostulate zu modifizieren,⁴⁹ wird nicht zuletzt anhand der oben aufgezeigten Beispiele deutlich, dass das Konzept des Homo oeconomicus zur Modellierung des Verhaltens von realen Entscheidungsakteuren wenig geeignet ist. Vielmehr zielt der normative Ansatz auf die Ableitung von formal-logisch optimalen Entscheidungen unter der Annahme eines idealtypisch handelnden Individuums ab. Gegebene Entscheidungsprämissen ermöglichen dabei, Aussagen bezüglich des maximal möglichen Nutzenniveaus treffen zu können. Ungeachtet dessen spielen Methoden der normativen Entscheidungstheorie eine wichtige Rolle in der Unternehmenspraxis in Form von Optimierungsmodellen. So zählen bspw. Kapitalmarktmodelle, Modelle zur Versicherungsnachfrage und zu Auswirkungen von Hintergrundrisiken zu den konkreten Anwendungsfeldern der normativen Entscheidungstheorie.⁵⁰

2.1.2 Deskriptive Entscheidungstheorie

Der Forschungszweig der *deskriptiven Entscheidungstheorie* gliedert sich in zwei Bereiche auf: deskriptive oder offene Entscheidungsmodelle und deskriptive Präferenztheorien.⁵¹

Deskriptive oder offene Entscheidungsmodelle⁵² sind dadurch gekennzeichnet, dass der menschliche Entscheidungsprozess – d. h. die Informationsgewinnung, -verarbeitung sowie die Beeinflussung des Entscheidungsprozessablaufs von den Umweltstrukturen – im Fokus der Betrachtung stehen, während bei normativen bzw. geschlossenen Modellen die Entscheidungsprämissen feststehen und vielseitige Interdependenzen mit der

⁴⁸ Anschauliche Beispiele für solche Konstellationen geben u. a. Eisenführ et al. 2010, S. 251; Pfister et al. 2017, S. 180f.

⁴⁹ Insbesondere das Unabhängigkeitsaxiom gilt als eine besonders restriktive Anforderung, weil dadurch die Menge der zulässigen Präferenzen stark eingeschränkt wird, vgl. Eisenführ et al. 2010, S. 254. Dies veranlasste bspw. Allais 1953 dazu, das EU-Modell selbst im Rahmen von rationalen Entscheidungen gänzlich abzulehnen. In Anbetracht dieser Kritik versuchten spätere Ansätze – z. B. die Konzepte des Choquet-Erwartungsnutzens von Schmeidler 1989 oder der Repräsentation von Klibanoff et al. 2005 – eine weniger strenge Form des Unabhängigkeitsaxioms mit dem EU-Modell zu vereinbaren. Zur allgemeinen Kritik und Verteidigung des EU-Modells siehe u. a. Kühberger 1994 bzw. Pfister 1994.

⁵⁰ Vgl. Lang 2009, S. 172ff.

⁵¹ Vgl. Langer und Rogowski 2009, S. 182.

⁵² Vgl. Kirsch 1970, S. 61.

Umwelt weitgehend außer Acht gelassen werden.⁵³ Die Aufgabe der deskriptiven Entscheidungsmodelle, die teils als Konglomerat aus Erklärungsansätzen und teils als geschlossene Theorien betrachtet werden können, besteht in der Beschreibung und Erklärung von Zuständen und Prozessen menschlichen Entscheidungsverhaltens.⁵⁴ Hierbei wird explizit dem Umstand Rechnung getragen, dass der einer Entscheidung vorausgehende Suchprozess und die Informationsverarbeitungskapazitäten eines Individuums deutlichen Limitationen unterliegen.⁵⁵

Wie bereits erläutert, können in der realen Welt mit menschlichen Entscheidungsakteuren zahlreiche Abweichungen von idealtypischen Verhaltensmustern gemäß der normativen Entscheidungstheorie beobachtet werden,⁵⁶ die zu Konsistenzproblemen wie dem Allais-Paradoxon führen können. Deskriptive Präferenztheorien schlagen eine Brücke zwischen solchen Paradoxa und dem normativen Ansatz, indem psychologische Einflussfaktoren in das mathematische Kalkül integriert werden.⁵⁷ Die einzelnen Konzepte können jedoch nur bestimmte Phänomene erklären, ein allumfassendes deskriptives Entscheidungsmodell existiert bislang nicht.⁵⁸ Die Prospect-Theorie ist ein bekannter Vertreter der deskriptiven Präferenztheorien und wird im Folgenden näher erläutert.

Die *Prospect-Theorie* (PT)⁵⁹ von Tversky und Kahneman 1979 bzw. die Erweiterung dieser Theorie in Form der Cumulative-Prospect-Theorie (CPT) von Tversky und Kahneman 1992 gilt als die wichtigste Revision des EU-Modells. Die PT bzw. CPT

⁵³ Vgl. Kirsch 1994, S. 1.

⁵⁴ Vgl. Langer und Rogowski 2009, S. 183.

⁵⁵ Empirische Untersuchungen wiesen nach, dass Menschen auf beschränkte Informationsverfügbarkeit und Verarbeitungskapazitäten systematisch mit vereinfachten Entscheidungsstrategien (Heuristiken) reagieren. Dieses Thema wird schwerpunktmäßig in Kapitel 2.2 behandelt.

⁵⁶ Selbst die zentrale Maxime der Nutzenmaximierung kann nicht uneingeschränkt einen Anspruch auf Gültigkeit erheben. Experimentelle Studien zum Ultimatum-Spiel (siehe u. a. Güth et al. 1982) zeigten, dass Menschen lieber auf einen monetären Gewinn verzichten, wenn der Fairnessaspekt verletzt wird. Bei einem Ultimatum-Spiel kann Spieler 1 allein über die Aufteilung eines Geldbetrags (z. B. 10 €) zwischen ihm selbst und einem anderen Spieler 2 entscheiden. Spieler 2 kann aber bei der von Spieler 1 festgelegten Aufteilung sein Veto einlegen, sodass beide Spieler leer ausgehen. Nach dem Prinzip der Nutzenmaximierung sähe die erwartete Aufteilung wie folgt aus: Spieler 1 beansprucht für sich 9,99 € und Spieler 2 gibt sich mit 1 Cent zufrieden (1 Cent ist besser als nichts). Etwa 40 % der Teilnehmer in empirischen Tests entschieden sich in der Rolle von Spieler 1 für eine „gerechte“ 50:50-Aufteilung des Geldbetrags und die in der Rolle von Spieler 2 agierenden Probanden lehnten eine Aufteilung umso stärker ab, je „ungerechter“ diese von ihnen empfunden wurde, vgl. Homann und Suchanek 2005, S. 381.

⁵⁷ So sind in der Regret-Theorie (siehe u. a. Bell 1982) oder der Disappointment-Theorie (siehe u. a. Bell 1985) emotionale Determinanten wie Freude, Enttäuschung sowie Bedauern integrale Bestandteile des mathematischen Modells.

⁵⁸ Vgl. Langer und Rogowski 2009, S. 188.

⁵⁹ Im Deutschen ist die Bezeichnung Neue Erwartungstheorie ebenso geläufig, vgl. Kahneman 2014, S. 342.

ermöglicht es, viele – aus normativer Sicht unerklärliche – Verhaltensanomalien modelltechnisch abzubilden. In dieser Arbeit wird stellvertretend für beide Modelle die Bezeichnung PT verwendet. Die wesentlichen Unterschiede zwischen der originären PT und der CPT werden an passenden Stellen dieses Kapitels aufgegriffen.

Im weiteren Verlauf werden die Grundkomponenten der PT vorgestellt, und zwar die Editierungsmechanismen, die Wertfunktion und die Entscheidungsgewichtungsfunktion.⁶⁰ Des Weiteren soll anhand von Beispielen gezeigt werden, wie mithilfe der PT ausgewählte Verhaltensparadoxa erklärt werden können.

Bei den Editierungsmechanismen⁶¹ der PT wird zwischen zwei Phasen im Entscheidungsprozess mit risikobehafteten Optionen unterschieden: Editierung und Evaluation. In der Editierungsphase werden die vorhandenen Informationen mittels kognitiver Operationen transformiert, kodiert – d. h. vereinfacht – und es werden mentale Repräsentationen erzeugt, die als Grundlage der Präferenzbildung dienen. Anschließend werden in der zweiten Phase die editierten Handlungsoptionen evaluiert. Im Einzelnen umfassen die Editierungsmechanismen die folgenden Schritte:⁶²

- **Kodierung:** Der Entscheidungsträger legt einen Referenzpunkt fest, von dem aus die verschiedenen Konsequenzen beurteilt werden. Der Referenzpunkt hängt einerseits von individuellen Verhältnissen ab (z. B. dem derzeitigen Vermögen, der Erwartungshaltung in Bezug auf eine Zielgröße) und kann andererseits von externen Faktoren, wie sozialen Vergleichen, beeinflusst werden. Die Lage des Referenzpunktes bestimmt, ob der Entscheider die Veränderung seiner aktuellen Situation infolge des Eintretens einer Konsequenz als Gewinn oder Verlust wahrnimmt.
- **Kombination:** Die Wahrscheinlichkeiten von Lotterien mit identischen Konsequenzen werden addiert. So reduziert sich die Menge der entscheidungsrelevanten Konsequenzen.
- **Segregation:** Sichere Bestandteile von risikobehafteten Konsequenzen einer Lotterie werden abgetrennt. So kann eine Lotterie mit den beiden Konsequenzen 300 € mit $p = 80\%$ und 200 € mit $p = 20\%$ in die folgenden beiden Nutzenkomponenten segregiert werden: Zum einen 200 € mit Sicherheit und zum anderen die Aussicht auf einen zusätzlichen Gewinn von 100 € mit $p = 80\%$.
- **Streichung:** Entscheidungskomponenten, die bei allen Handlungsoptionen gleich sind, bleiben i. d. R. unberücksichtigt. Dies kann dazu führen, dass mehrstufige Entscheidungen ungeachtet möglicher Interdependenzen zwischen den einzelnen Stufen isoliert betrachtet werden.

⁶⁰ Vgl. hier und im Folgenden Pfister et al. 2017, S. 184.

⁶¹ Die Editierungsmechanismen sind bei der originären PT und der CPT deckungsgleich.

⁶² Vgl. im Folgenden Laux et al. 2014, S. 175f.; Wenig 2009, S. 198.

- Vereinfachung: Wahrscheinlichkeiten und Nutzenkomponenten werde aus Gründen der Vereinfachung ab- oder aufgerundet. So werden bspw. die Werte 101 € und $p = 49\%$ in den Augen eines Entscheiders als äquivalent zu 100 € bzw. $p = 50\%$ erachtet.
- Vorauswahl über Dominanzkriterien: Dominierte Alternativen bleiben vom Entscheidungsprozess gänzlich ausgeschlossen.
- Evaluierung: Nachdem die Konsequenzen der Handlungsoptionen, wie oben beschrieben, editiert wurden, folgt in der zweiten Phase die Evaluation, d. h., es wird analog zum EU-Modell die Alternative mit dem höchsten erwarteten subjektiven Nutzenwert bestimmt.

In Bezug auf die einzelnen Editierungsschritte besteht immer noch Unklarheit darüber, ob und inwieweit diese mathematisch abgebildet werden können. Die Tatsache, dass im Vorfeld einer Entscheidung die Editierung von zur Verfügung stehenden Alternativen tatsächlich stattfindet, wird indes nicht angezweifelt.⁶³

Tversky und Kahneman 1979 führten im Kontext der PT die Bezeichnung Wertfunktion (engl. Value Function)⁶⁴ ein und nahmen damit bewusst eine Abgrenzung zur klassischen Nutzenfunktion des EU-Modells vor. Sowohl die Wertfunktion als auch die Nutzenfunktion verfolgen die gleiche Grundidee, nämlich die funktionale Abbildung von subjektiven Nutzenwerten in Abhängigkeit von einer Zielgröße. Die Wertfunktion modifiziert jedoch die klassische Nutzenfunktion in mehreren Aspekten entscheidend. Die nachfolgenden Eigenschaften sind für eine deskriptiv gültige Wertfunktion (siehe Abbildung 2) gemäß der PT charakteristisch:⁶⁵

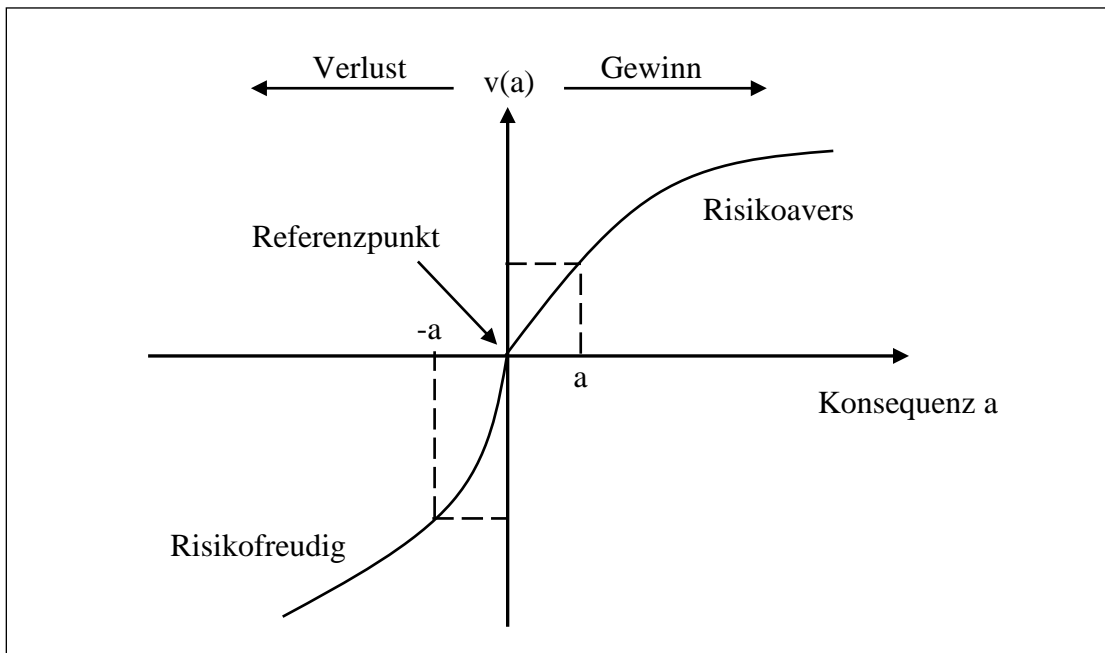
- Referenzpunktabhängigkeit: Dies ist ein fundamentaler Unterschied zum traditionellen EU-Modell. Wie bereits im Zusammenhang mit dem Kodierungsmechanismus in der Editierungsphase erwähnt, werden die Änderungen der Nutzenniveaus nicht absolut, sondern relativ zum aktuellen Status quo – der Referenzpunkt hat einen Wert von $v(0) = 0$ – bewertet. Positive Veränderungen relativ zum Referenzpunkt werden vom Entscheidungsträger als Gewinne und negative Veränderungen als Verluste wahrgenommen. So kann eine Gehaltserhöhung, je nachdem, wie hoch diese bei einem Kollegen ausfiel, einmal freudeinduzierend als Gewinn und einmal als Verlust mit einer daraus resultierenden Enttäuschung empfunden werden.

⁶³ Vgl. Wenig 2009, S. 198.

⁶⁴ Die Annahmen bezüglich der Wertfunktion in der CPT gehen auf die originäre PT zurück.

⁶⁵ Vgl. Eisenführ et al. 2010, S. 432ff.; Pfister et al. 2017, S. 186ff.

Abbildung 2: Wertfunktion der PT



Quelle: In Anlehnung an Tversky und Kahneman 1979, S. 279.

- Abnehmende Sensitivität: Rechts vom Referenzpunkt (Gewinnbereich) ist die Wertfunktion von einem abnehmenden Grenznutzen gekennzeichnet. Damit ist gemeint, dass z. B. die Verdoppelung einer Zielgröße einen weniger als doppelt so hohen Nutzenzuwachs bedingt: $v(2 \times a) < 2 \times v(a)$. In Erweiterung zum EU-Modell, in dem die Nutzenfunktion über den gesamten Bereich konkav verläuft, weist die Wertfunktion einen konvexen Funktionsverlauf links vom Referenzpunkt (Verlustbereich) auf. Das bedeutet eine abnehmende Sensitivität gegenüber Verlustzuwächsen: $v(2 \times (-a)) > 2 \times v(-a)$. Zusammengenommen kann anhand der Wertfunktion erklärt werden, warum Entscheidungsträger ihre Risikoeinstellung ändern, sobald Gewinne durch entsprechende Verluste ersetzt werden.⁶⁶
- Verlustaversion: Eine weitere Besonderheit der Wertfunktion äußert sich darin, dass ein Gewinn weniger stark als ein betragsmäßig gleicher Verlust bewertet wird: $v(a) < |v(-a)|$. Mit anderen Worten wird unterstellt, dass die Wertfunktion im Verlustbereich (zunächst) steiler verläuft als im Gewinnbereich.⁶⁷

⁶⁶ Dieses Phänomen trägt die Bezeichnung Reflexionseffekt (engl. Reflection Effect), vgl. Tversky und Kahneman 1979, S. 286f. So präferieren bspw. die meisten Menschen den sicheren Gewinn von 3.000 € gegenüber einer Wette bestehend aus 4.000 € mit $p = 80\%$ und 0 € mit $p = 20\%$ ($EW = 3.200$ €). Die Präferenz ändert sich zugunsten der risikobehafteten Alternative, sobald die Geldbeträge mit einem negativen Vorzeichen (Verlustkodierung) versehen werden. Im Gewinnbereich sind die Entscheider folglich risikoavers und im Verlustbereich risikofreudig.

⁶⁷ Vgl. Eisenführ et al. 2010, S. 432; Wenig 2009, S. 199.

Die mathematische Gleichung der Wertfunktion lautet wie folgt:

$$v(a) = \begin{cases} a^\alpha, & a \geq 0 \\ -\lambda \times (-a)^\beta, & a < 0 \end{cases}$$

Mit

$0 \leq \alpha, \beta \leq 1$: Grad der abnehmenden Sensitivität

$\lambda > 1$: Grad der Verlustaversion

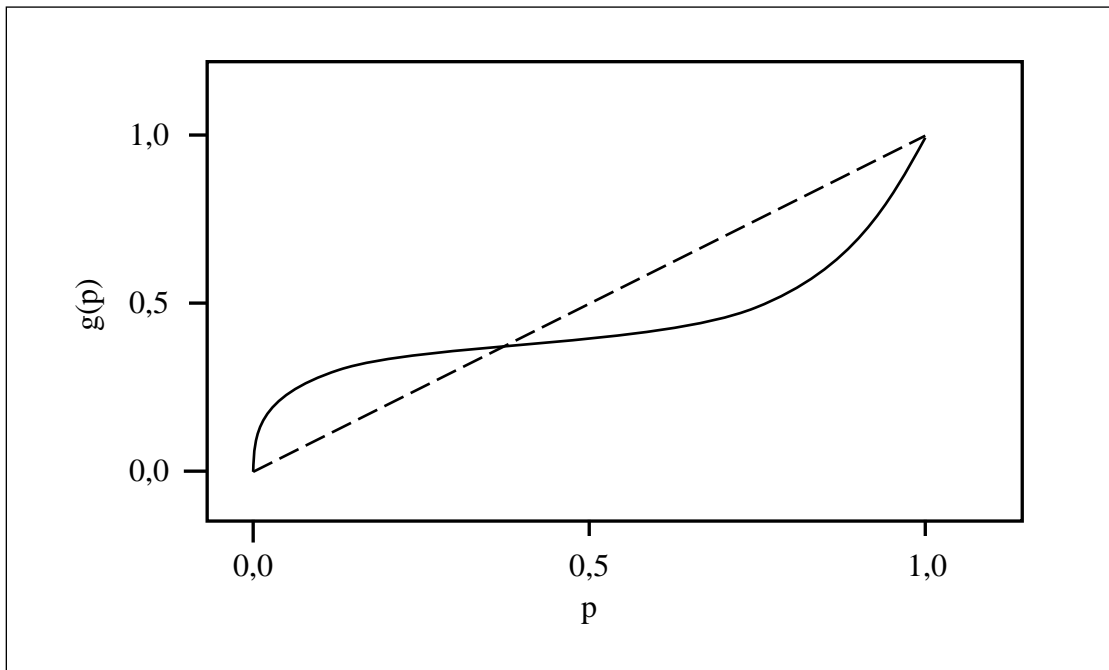
Eine weitere bedeutende Modifikation des EU-Modells vonseiten der PT betrifft die Entscheidungsgewichtungsfunktion.⁶⁸ Laut der PT durchlaufen die für das Eintreten einer Konsequenz relevanten Wahrscheinlichkeiten eine kognitive Anpassung. Die Eintrittswahrscheinlichkeit p wird zum Entscheidungsgewicht $g(p)$ transformiert, das die Bedeutung des Eintretens einer Konsequenz im Entscheidungsprozess repräsentiert. Gestützt auf empirische Untersuchungen konnten Tversky und Kahneman 1992 bzw. Gonzalez und Wu 1999 folgende Annahmen bezüglich der Entscheidungsgewichtungsfunktion (siehe Abbildung 3) ableiten:⁶⁹

- Monotonie: Die Funktion ist monoton steigend.
- Feste Endpunkte: Die Funktion ist an den Endpunkten fixiert mit folgenden Werten $g(p = 0) = 0$ und $g(p = 1) = 1$. Die beiden Werte stehen für die Randpunkte eines Kontinuums und symbolisieren die absolute Unmöglichkeit bzw. Sicherheit.

⁶⁸ Die wichtigste Erweiterung der originären PT durch die CPT spiegelt sich in der Entscheidungsgewichtungsfunktion wider. Die Schwäche der originären PT, dass die Transformation einzelner Wahrscheinlichkeiten die Wahl von dominierten Alternativen bedeuten kann, wird in der CPT mithilfe der rangplatzabhängigen Nutzentheorie in Kombination mit dem Referenzpunktdenken behoben, vgl. Wenig 2009, S. 201. Die Trennung in zwei rangplatzabhängige Funktionen für Gewinne und Verluste verleiht der Entscheidungsgewichtungsfunktion gemäß der CPT den typischen umgekehrt S-förmigen Verlauf. Weitere Vorteile der CPT gegenüber der originären PT umfassen insbesondere die Erweiterung des Modells auf Lotterien mit mehr als zwei Konsequenzen und die Anwendungsmöglichkeit auf Entscheidungssituationen unter Unsicherheit, vgl. Pfister et al. 2017, S. 193f.

⁶⁹ Vgl. hier und im Folgenden Pfister et al. 2017, S. 189f.

Abbildung 3: Entscheidungsgewichtungsfunktion der PT



Quelle: In Anlehnung an Tversky und Kahneman 1992, S. 310.

- Abnehmende Sensitivität: Die Entscheidungsgewichtungsfunktion hat zwei Referenzpunkte in Gestalt der beiden Endpunkte. Gemäß dem Möglichkeitseffekt bewirkt eine geringfügige Zunahme der Wahrscheinlichkeit von $p = 0$ auf z. B. $p = 0,05$ eine überproportionale Zunahme des Entscheidungsgewichts – d. h., kleine Wahrscheinlichkeiten werden übergewichtet. Gleichzeitig ist gemäß dem Sicherheitseffekt eine geringfügige Abnahme der Wahrscheinlichkeit von $p = 1$ auf z. B. $p = 0,95$ mit einer überproportionalen Abnahme des Entscheidungsgewichts verbunden – d. h., große Wahrscheinlichkeiten werden untergewichtet. Die abnehmende Sensitivität bedeutet, dass ausgehend von den Randpunkten sich die Stärke der Über- bzw. Untergewichtung zur Mitte der Kurve hin sukzessive abschwächt.
- Subadditivität: Es gilt $g(p) + g(1 - p) < 1$. Im Gegensatz zum EU-Modell gehorchen die Entscheidungsgewichte nicht den Regeln der Wahrscheinlichkeitstheorie und summieren sich nicht zur Eins auf.

Die einparametrische Form der mathematischen Gleichung der Entscheidungsgewichtungsfunktion lautet wie folgt:⁷⁰

$$g(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{\frac{1}{\gamma}}}$$

Mit

$0 < \gamma < 1$: Grad der Über- bzw. Untergewichtung

Einige empirische Untersuchungen ergaben, dass der Parameter γ für den Verlust- bzw. Gewinnbereich unterschiedlich ausfallen kann. Die typischen Verläufe der Wert- und Entscheidungsgewichtungsfunktion konnten in zahlreichen empirischen Untersuchungen bestätigt werden (Parameterschätzungen aus ausgewählten Studien befinden sich in Tabelle 2). Formal wird der gewichtete Gesamtnutzen einer risikobehafteten Alternative A – hier PT(A) – wie folgt ermittelt:

$$PT(A) = \sum_{i=1}^m v(a_i) \times w^-(p_i) + \sum_{i=m+1}^n v(a_i) \times w^+(p_i)$$

Mit

$$w^-(p_i) = g^- \times \left(\sum_{j=1}^i p_j \right) - g^- \times \left(\sum_{j=1}^{i-1} p_j \right)$$

$$w^+(p_i) = g^+ \times \left(\sum_{j=i}^n p_j \right) - g^+ \times \left(\sum_{j=i+1}^n p_j \right)$$

Die abweichende Notation für w^- bzw. w^+ ist dadurch bedingt, dass die Bildung des gewichteten Gesamtnutzens im Verlust- und Gewinnbereich unterschiedlich vonstattengeht. Im Verlustbereich erfolgt die Kumulierung der Wahrscheinlichkeiten „von links zum Referenzpunkt“, wohingegen diese im Gewinnbereich „von rechts zum Referenzpunkt“ abläuft.⁷¹

⁷⁰ Vgl. Tversky und Kahneman 1992, S. 309. Neben der einparametrischen Formel wird in der Literatur ebenso eine zweiparametrische Gleichung für die Entscheidungsgewichtungsfunktion vorgeschlagen, siehe z. B. Gonzalez und Wu 1999. Damit soll neben dem Grad der Über- bzw. Untergewichtung von Wahrscheinlichkeiten an verschiedenen Stellen der kumulierten Verteilung ein weiterer psychologischer Effekt modelliert werden. Der zweite Faktor steht für die generelle Über- bzw. Untergewichtung von Wahrscheinlichkeiten und beeinflusst die Höhe der Funktion, vgl. Eisenführ et al. 2010, S. 431.

⁷¹ Eine anschauliche Demonstration dieses Verfahrens anhand eines Zahlenbeispiels findet sich u. a. bei Eisenführ et al. 2010, S. 426ff.

Tabelle 2: Empirische Parameterschätzungen der PT

| Studie | Wertfunktion | | Entscheidungsge- wichtungsfunktion |
|---------------------------|-----------------------|-----------|---------------------------------------|
| | α bzw. β | λ | γ |
| Abdellaoui 2000 | 0,89 | – | 0,60 (0,60) |
| Abdellaoui et al. 2005 | 0,91 | – | (0,76) |
| Abdellaoui et al. 2008 | 0,86 | 2,61 | 0,60 |
| Abdellaoui et al. 2011 | 0,79 | 2,47 | 0,73 |
| Balaz et al. 2013 | 1,00-1,02 | 2,10-2,50 | 0,58-0,62 |
| Bleichrodt und Pinto 2000 | 0,77 | – | 0,67 (0,55) |
| Kilka und Weber 2001 | 0,76-1,00 | – | (0,30-0,51) |
| Stott 2006 | 0,19 | – | 0,96 |
| Tversky und Kahneman 1992 | 0,88 | 2,25 | 0,61 |
| Tversky und Fox 1995 | 0,88 | – | (0,69) |
| Wu und Gonzalez 1996 | 0,52 | – | 0,71 (0,68) |
| Wu und Gonzalez 1999 | 0,49 | – | 0,68 |

Parameter für den Verlustbereich sind in Klammer aufgeführt.

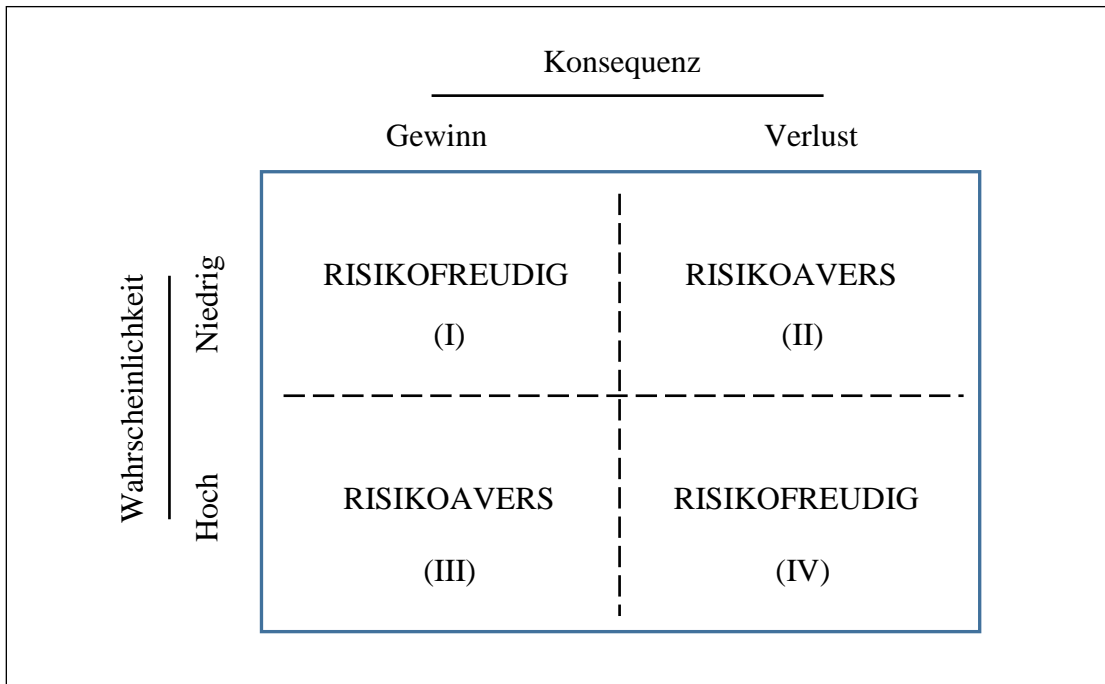
Quelle: In Anlehnung an Balaz et al. 2013, S. 661.

In Kapitel 2.1.1 im Zusammenhang mit dem Allais-Paradoxon wurde ein Zahlenbeispiel, das ein mit dem Unabhängigkeitsaxiom nonkonformes Verhalten induziert, vorgestellt.⁷² Mithilfe der PT kann ein solches Verhalten modelltechnisch abgebildet werden. Nach der Logik der Wahrscheinlichkeitstheorie weisen die Lotteriepaaire A und B bzw. A' und B' die gleiche Relation der Gewinnwahrscheinlichkeiten auf und somit wäre bei konstant gehaltenen Gewinnbeträgen von A zu A' bzw. B zu B' eine Präferenzumkehrung nicht zu erwarten.⁷³ Aus der psychologischen Perspektive sind allerdings diese Gewinnrelationen aufgrund der subjektiven Gewichtung der absoluten Wahrscheinlichkeiten unterschiedlich groß. Demnach fällt aufgrund des Sicherheitseffekts der Übergang von „sicher“ (100 %) zu „ziemlich wahrscheinlich“ (80 %) deutlich stärker ins Gewicht als der Übergang von „nicht besonders wahrscheinlich“ (25 %) zu „noch etwas weniger wahrscheinlich“ (20 %). Im Lotteriepaar A und B überwiegt der Sicherheitseffekt und der Entscheidungsträger präferiert die Option A. Auf der anderen

⁷² Siehe Abbildung 1.

⁷³ Das Verhältnis der Gewinnwahrscheinlichkeiten von A zu B bzw. A' zu B' beträgt $\frac{100\%}{80\%} = \frac{25\%}{20\%} = \frac{5}{4}$.

Abbildung 4: Vierfeldmuster der PT



Quelle: In Anlehnung an Kahneman 2014, S. 389; Pfister et al. 2017, S. 193.

Seite kann die etwas höhere Gewinnwahrscheinlichkeit von A' die Aussicht auf einen höheren Gewinn der Lotterie B' nicht kompensieren und der Entscheidungsträger ändert seine Wahl zugunsten der letztgenannten Option.

Anhand der Wert- und Entscheidungsgewichtungsfunktion kann die Risikoeinstellung in verschiedenen Situationen besser und viel flexibler als im Rahmen des EU-Modells prognostiziert werden. Die kombinierten Effekte aus einer Konsequenz (Gewinn vs. Verlust) und ihrer Eintrittswahrscheinlichkeit (hoch vs. niedrig) auf die Risikoeinstellung können gemäß der PT in einem Vierfeldmuster (siehe Abbildung 4) dargestellt werden.⁷⁴ Das Vierfeldmuster kann zur Erklärung einer Reihe von Verhaltensphänomenen herangezogen werden. Die Überschätzung von niedrigen Wahrscheinlichkeiten führt im Gewinnbereich (erster Quadrant) dazu, dass die Aussicht auf einen möglichen Gewinn überschätzt wird, was z. B. die Teilnahme an unvorteilhaften Glücksspielen erklären kann.⁷⁵ Im Verlustbereich kann die Überschätzung niedriger Wahrscheinlichkeiten (zweiter Quadrant) die Nachfrage nach überflüssigen Versicherungen als Ausdruck eines risikoaversen Verhaltens herbeiführen. Die Angst vor einem hohen Verlust

⁷⁴ Vgl. Tversky und Kahneman 1992, S. 306ff.; Tversky und Fox 1995, S. 269f.

⁷⁵ In der Konstellation des ersten Quadranten wirken die Wert- und Entscheidungsgewichtungsfunktionen eigentlich in gegensätzliche Richtungen, weil die Wertfunktion im Gewinnbereich konkav verläuft. Die Überschätzung kleiner Eintrittswahrscheinlichkeiten und die daraus resultierende übersteigerte Hoffnung verleitet aber zum risikofreudigen Verhalten umso mehr, je größer der potenzielle Gewinn ist.

ist so stark ausgeprägt, dass die bloße Möglichkeit dieses Verlusts – selbst bei einer zu vernachlässigenden Eintrittswahrscheinlichkeit – über eine Versicherung mit einer kalkulatorisch zu hoch angesetzten Prämie ausgeschlossen werden soll. Weitere Fälle eines risikoaversen Verhaltens können in Gerichtsprozessen beobachtet werden: Aus Furcht, nichts zu bekommen, sind viele Kläger bereit, einen ungünstigen Vergleich zu akzeptieren. Die Kombination aus der Unterschätzung hoher Wahrscheinlichkeiten und konkavem Funktionsverlauf der Wertfunktion im Gewinnbereich (dritter Quadrant) bedingt ebenfalls die Risikoaversion.⁷⁶ Schließlich kann die Unterschätzung großer Wahrscheinlichkeiten in Verbindung mit Verlusten (vierter Quadrant) ein risikofreudiges Verhalten induzieren. Mit der Risikowahl wird auf einen günstigeren, aber weniger wahrscheinlichen Ausgang spekuliert. Im Unternehmenskontext kann z. B. eine übersteigerte Risikobereitschaft in einem ruinösen Preiswettbewerb münden.⁷⁷

Die vielseitige Einsetzbarkeit der PT konnte auch dadurch unter Beweis gestellt werden, dass dieses Modell zu Beschreibungs- und Prognosezwecken vom individuellen Verhalten in zahlreichen Lebensbereichen verwendet wird. In der Ökonomik wird mithilfe der PT das Verhalten von Entscheidungsträgern auf dem Aktien- und Arbeitsmarkt, im Konsumgüterbereich und in der Versicherungsbranche erklärt, um nur einige Beispiele zu nennen.⁷⁸

Die PT prägte die verhaltensorientierte Wirtschaftsforschung wesentlich mit und trug zum hohen Bekanntheitsgrad dieser Forschungsrichtung maßgeblich bei.⁷⁹ Die vorliegende Arbeit bedient sich der experimentellen Forschungsmethode⁸⁰ und knüpft somit unmittelbar an den behavioristisch orientierten Ansatz der deskriptiven Entscheidungstheorie an.

⁷⁶ Gestützt auf das EU-Modell kann die gleiche Aussage bezüglich der Risikoeinstellung im dritten Quadranten hergeleitet werden. Im EU-Modell dient jedoch einzig der empirisch bestätigte konkave Funktionsverlauf der Nutzenfunktion zur Ableitung der Risikoeinstellung, was die Variabilität und das Erklärungspotenzial des EU-Modells im Vergleich zur PT wesentlich einschränkt.

⁷⁷ Vgl. Kahneman 2014, S. 391f.

⁷⁸ Eine ausführliche Behandlung der Anwendungsfelder der PT und weitere Beispiele finden sich u. a. bei Barberis 2013, S. 180ff.

⁷⁹ Vgl. Barberis 2013, S. 173f. In breiten Kreisen erlangte die PT spätestens seit der Verleihung des Wirtschafts-Nobelpreises an Daniel Kahneman im Jahr 2002 hohe Popularität.

⁸⁰ In Kapitel 2.5 wird die Methode der experimentellen Forschung vorgestellt.

2.2 Dual-System-Theorien

2.2.1 System 1 und System 2

*Dual-System-Theorien*⁸¹ (engl. Dual-Process Theories, kurz DST) gliedern kognitive Abläufe in zwei Gruppen von Prozessen auf. Obwohl die Unterschiede zwischen den kognitiven Prozessen im Allgemeinen auf mehr als nur zwei Merkmalen beruhen, erwies es sich für die psychologische Forschung als nützlich, eine kategoriale Einteilung der Denkprozesse in zwei qualitativ verschiedene Teilsysteme vorzunehmen: System 1 und System 2.⁸² In diesem Zusammenhang muss beachtet werden, dass die oben vorgenommene Dichotomie tatsächlich eine kontinuierliche Dimension mit vielen Zwischenformen ist. Außerdem kann ein System durchaus einige Züge des anderen Systems aufweisen.⁸³

Die für System 1 charakteristischen kognitiven Prozesse erfolgen schnell und unbewusst.⁸⁴ Da das Denken in diesem Modus leichtfällt, braucht System 1 keine willentliche Aktivierung und läuft bei jedem kognitiven Prozess automatisch im Hintergrund. Zu den typische Aufgaben von System 1 gehören: Sprachverstehen in eigener Muttersprache, Sinneswahrnehmungen sowie Verarbeitung von affektiven Bewertungen.⁸⁵

⁸¹ Synonym dazu werden Begriffe wie Dual-Prozess-Theorien oder duale Architekturen verwendet.

⁸² Die Bezeichnung System 1 bzw. System 2 geht auf Kahneman 2003, 2014; Morewedge und Kahneman 2010 sowie Stanovich und West 2000 zurück. Der System-1/System-2-Ansatz baut wiederum auf ähnlichen Konzepten wie der klassischen Theorie zum prozeduralen und deklarativen Wissen auf, siehe Anderson 1983; Chaiken und Trope 1999; Epstein et al. 1996; Sloman 1996. DST implizieren nicht, dass bestimmte Areale im menschlichen Gehirn streng nach einer dualen Systematik arbeiten. Vielmehr sollen der kognitive Denkprozess und das daraus resultierende Entscheidungsergebnis als charakteristisch für eines der beiden Systeme interpretiert werden. Neben einem Dual-Prozess-Modell wird in der Literatur auch ein Tri-Prozess-Modell vorgeschlagen, vgl. Stanovich 2011, S. 61ff. An dieser Stelle soll kritisch angemerkt werden, dass die theoretische Begründung und empirische Evidenz in Bezug auf die DST umstritten sind, siehe Grayot 2020.

⁸³ Vgl. Norman 2010, S. 196; Pfister et al. 2017, S. 346.

⁸⁴ Weitere Eigenschaften von System 1 sind in Tabelle 3 aufgeführt.

⁸⁵ Vgl. Hadjichristidis et al. 2017b, S. 646; Kahneman 2014, S. 31ff.

Tabelle 3: Ausgewählte Eigenschaften von System 1 und System 2

| System 1 | System 2 |
|----------------------------------|--|
| Affektiv | Rational |
| Assoziativ | Regelbasiert |
| Automatisch | Kontrolliert |
| Evolutionär alt | Evolutionär neu |
| Heuristisch | Analytisch |
| Kognitiv leicht | Kognitiv anstrengend |
| Unabhängig vom Arbeitsgedächtnis | Limitiert durch Arbeitsgedächtniskapazitäten |
| Pragmatisch | Logisch |
| Schnell | Langsam |
| Unbewusst | Bewusst |

Quelle: In Anlehnung an Evans 2008, S. 257; Witt 2018, S. 44.

Das Denken im System-2-Modus steht dem System-1-Denken diametral gegenüber. System 2 steht für kontrollierte und bewusst ablaufende kognitive Prozesse.⁸⁶ Es erfordert gezielte Aufmerksamkeitsallokation und beansprucht limitierte Ressourcen des Arbeitsgedächtnisses.⁸⁷ Kognitiv anstrengende Logik- und Rechenaufgaben (z. B. die Berechnung des Ausdrucks „744:3“), komplexe und wenig eingeübte motorische Abläufe (z. B. das Rückwärtseinparken) sowie ambivalente Präferenzen (Trade-offs) sind typische Situationen, in denen System 2 involviert ist.⁸⁸

Die folgenden Wechselbeziehungen zwischen System 1 und System 2 sind im Rahmen der DST von besonderer Bedeutung:⁸⁹

- System 1 als Default-System und System 2 als Interventionsinstanz: Die meisten Denkprozesse im Alltag laufen ganz automatisch im „Voreinstellungsmodus“ von System 1 ab. Die mit System 1 getroffenen Entscheidungen und das daraus resultierende Handlungsergebnis sind i. d. R. zufriedenstellend, solange die Umgebungssituation der gewohnten Normalität entspricht. Sobald aber der Entscheidungsträger

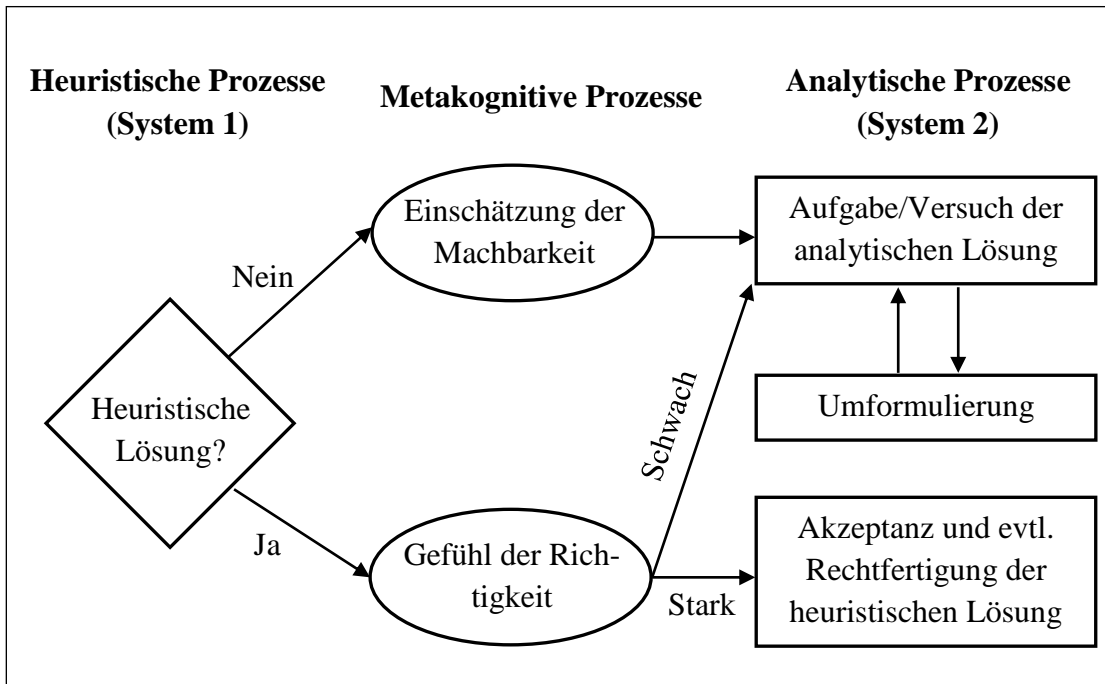
⁸⁶ Weitere Eigenschaften von System 2 sind in Tabelle 3 aufgeführt.

⁸⁷ Vgl. Whitney et al. 2008, S. 1179. Im Arbeitsgedächtnis werden Informationen, die für die anspruchsvollen kognitiven Denk- und Entscheidungsprozesse gleichzeitig benötigt werden, kurzfristig aufbewahrt. Das Arbeitsgedächtnis gilt als eine Zwischeninstanz zum Langzeitgedächtnis und wird mit bewussten bzw. kontrollierten Denkprozessen assoziiert, vgl. Baddeley 2003, S. 829.

⁸⁸ Vgl. Pfister et al. 2017, S. 345f.

⁸⁹ Vgl. im Folgenden Evans 2008, S. 261ff.; Pfister et al. 2017, S. 347f.

Abbildung 5: Metakognitive Auslöser von System 2



Quelle: In Anlehnung an Thompson 2009, S. 172.

mit Schwierigkeiten und Ausnahmesituationen konfrontiert wird, wechselt er in den aufmerksamen System-2-Denkmodus. So wird z. B. eine Routinefahrt ins Büro standardmäßig von System 1 ausgeführt. Taucht auf dem Weg zur Arbeit unerwartet eine Baustelle auf und muss folglich eine Alternativroute überlegt werden, kommt System 2 zum Einsatz. Weil kontrollierte und regelbasierte Denkprozesse kognitiv anstrengend sind, erfordert der Einsatz von System 2 mitunter erhebliche Willenskraft.⁹⁰ Gemäß dem Prinzip des geringsten Aufwands wird das mit Anstrengung verbundene System 2 sparsam eingesetzt und nur in solchen Fällen aktiviert, wenn subjektiv gute bzw. überzeugende Gründe dafür sprechen. In Abbildung 5 sind metakognitive Auslöser von System 2 schematisch dargestellt. Falls die heuristische Lösung eines Problems nicht zum Tragen kommen kann,⁹¹ wird zunächst eine Einschätzung im Hinblick auf die generelle Machbarkeit vorgenommen. Je nachdem, wie diese Einschätzung ausfällt, wird weiter im System-2-Modus das Problem gelöst oder es kommt zur Aufgabe von weiteren Lösungsversuchen.⁹² Falls sich bei der heuristischen Lösung eines Problems ein starkes inneres Gefühl der Richtigkeit

⁹⁰ Kahneman 2014, S. 55ff. spricht vom „faulen“ Kontrolleur im Zusammenhang mit System 2.

⁹¹ Beispielsweise aufgrund der hohen Komplexität oder mangelnden Expertise.

⁹² Gegebenenfalls wird noch vor der endgültigen Aufgabe versucht, die Problemstellung umzuformulieren, um so die Lösung des Problems doch noch zu erreichen.

einstellt,⁹³ wird diese Lösung ohne Weiteres von System 2 als Handlungsgrundlage akzeptiert oder System 2 wird lediglich dazu benutzt, die heuristische Lösung mit logischen Argumenten zu untermauern. Löst die heuristische Lösung ein schwaches inneres Gefühl der Richtigkeit aus, so wird mit dem langsameren System 2 ein weiterer Blick auf das Problem geworfen und es folgt die oben beschriebene Aufgabe/Lösung-Prozedur. In allen Fällen braucht System 2 im Gegensatz zu System 1 einen mentalen Auslösungsmechanismus.

- Konflikte zwischen System 1 und System 2: Die kontinuierliche Arbeit von System 1 läuft stets im Hintergrund ab und kann zwar von System 2 als Interventionsinstanz kontrolliert, aber nicht komplett abgestellt werden. Ein gutes Beispiel für solche unbewusste und zum Teil unkontrollierbare kognitive Prozesse sind optische Täuschungen, wie bspw. die Müller-Lyer-Illusion. Bei der Müller-Lyer-Illusion handelt es sich um zwei gleich lange Linien, die jedoch aufgrund der unterschiedlichen Formationen an den Enden (Pfeilspitze bzw. umgekehrte Pfeilspitze) visuell unterschiedlich lang erscheinen. Selbst wenn die optische Täuschung aufgedeckt und ins Bewusstsein geführt wird (System 2), lässt der visuelle Eindruck der unterschiedlichen Längen trotzdem nicht nach (System 1). Auf diese Weise können Konflikte zwischen den beiden Systemen entstehen. In solchen Konfliktsituationen wird System 2 manchmal dazu benutzt, die intuitiven Urteile von System 1 mit logischen Argumenten zu rationalisieren.

Der System-1/System-2-Ansatz bietet einen theoretischen Rahmen für weitere Konzepte, die zur Erklärung von Phänomenen im Zusammenhang mit heuristischem Denken dienen. Nachfolgend werden zwei Sichtweisen auf das Thema Heuristiken erläutert und die vorliegende Arbeit wird in diesen Kontext eingeordnet.

2.2.2 Heuristiken und Biases

2.2.2.1 Zusammenhang mit System 1 und System 2

Restriktionen in Bezug auf Informationsverfügbarkeit und kognitive Verarbeitungskapazitäten führen oft zum Einsatz von *Heuristiken* – d. h. vereinfachten Entscheidungsregeln – bei der Bildung von Urteilen und beim Fällen von intuitiven Entscheidungen.⁹⁴ Heuristiken stehen im unmittelbaren Zusammenhang mit kognitiven Prozessen von System 1. Sie werden intuitiv und ohne langes Nachdenken angewandt. In den meisten

⁹³ Das innere Gefühl der Richtigkeit stellt sich insbesondere in bereits bekannten Entscheidungssituationen und bei hoher mentaler Leichtigkeit, mit der eine Aufgabe bearbeitet werden kann, ein.

⁹⁴ Vgl. Langer und Rogowski 2009, S. 183.

Fällen sind heuristische Lösungsstrategien ausreichend, um zu richtigen⁹⁵ oder annähernd richtigen Urteilen und Entscheidungen zu gelangen.

Das von Tversky und Kahneman 1974 begründete Forschungsprogramm im Bereich der Heuristiken und Biases⁹⁶ (HB) zeigte, dass der Einsatz von Heuristiken in manchen Situationen zu systematischen *Urteilsverzerrungen*⁹⁷ (engl. Biases) führen kann. Gemäß dem HB-Ansatz ist das schnelle und affektive System 1 für diese Biases verantwortlich.⁹⁸ Insbesondere die folgenden Eigenschaften von System 1 begünstigen die Entstehung von Fehlurteilen:⁹⁹

- **Assoziative Aktivierung:** Das menschliche Gedächtnis besteht aus einem riesigen Netzwerk an assoziativ miteinander verbundenen Informationen. Ein Stimulus (z. B. Bilder, Gerüche oder Wörter) aktiviert automatisch und kaskadenartig alle mit diesem konkreten Stimulus assoziierten Informationselemente im Gedächtnis, die früher aufgrund von Erfahrungen oder infolge von Lernprozessen erworben wurden. Hierbei folgt der Prozess der Aktivierungsausbreitung keinen logischen oder syntaktischen Regeln.¹⁰⁰
- **Kohärenz:** Die mittels einer Abfolge von Stimuli aktivierten Assoziationen werden von System 1 spontan zu einem kohärenten Gesamtbild integriert, woraus insbesondere in einer beschränkten Informationslage voreilige Schlussfolgerungen gezogen werden können.¹⁰¹ Sollen bspw. basierend auf nur zwei Charaktereigenschaften wie Intelligenz und Stärke die Fähigkeiten einer Führungskraft beurteilt werden, fällt das Urteil intuitiv und augenblicklich sehr positiv aus. Die Beurteilung einer Führungskraft anhand von nur zwei Charakteristika ist nach logischen Gesichtspunkten

⁹⁵ „Richtig“ bedeutet im vorliegenden Kontext, dass der Entscheidungsprozess in Übereinstimmung mit den wahrscheinlichkeitstheoretischen Grundprinzipien erfolgt, vgl. Tversky und Kahneman 1974, S. 1124.

⁹⁶ Klassische Arbeiten auf diesem Gebiet lassen sich u. a. in den Sammelbänden von Gilovich et al. 2002 oder Kahneman et al. 1982 finden. Der HB-Ansatz ist innerhalb der deskriptiven Entscheidungstheorie dem Forschungszweig der offenen Entscheidungsmodelle (siehe Kapitel 2.1.2) zuzuordnen. Im Unterschied zu den deskriptiven Präferenztheorien, zu denen bspw. die bereits vorgestellte PT gehört, basieren individuelle Entscheidungen in der HB-Theorie nur auf Teilinformationen und sind nicht einer konsequentialistischen Optimierungslogik unterworfen.

⁹⁷ Synonym dazu werden in der Literatur auch die Begriffe Verhaltensanomalien oder Informationspathologien verwendet, siehe Hirsch 2007; Scholl 1992.

⁹⁸ Vgl. Stanovich und West 2000, S. 658f.

⁹⁹ Vgl. im Folgenden Kahneman 2014, S. 69ff.; Pfister et al. 2017, S. 348f.

¹⁰⁰ Vgl. Anderson 1983, S. 89ff.; Anderson 2013, S. 126ff.

¹⁰¹ Kahneman 2014 umschreibt dieses Phänomen mit dem Ausdruck: „What you see is all there is“.

selbstverständlich nicht ausreichend und sollte unterdrückt werden, bis alle relevanten Informationen vorliegen.¹⁰² Da aber System 1 autonom arbeitet, lässt sich ein gewisser Grad an Voreingenommenheit nicht gänzlich vermeiden.

- **Kausalität:** Anders als beim statistischen Denken, in dessen Rahmen die Schlüsse über Einzelfälle aus Kategorien und der Gesamtheit gezogen werden, vollzieht die intuitive Wahrnehmung von System 1 fortwährend Ursache-Wirkungs-Bewertungen unabhängig von der Fallzahl und der Sinnhaftigkeit solcher Bewertungen in einer gegebenen Situation. Wird z. B. eine Person beobachtet, die ihre Hand in eine bestimmte Richtung ausstreckt, wird dies sofort als eine Richtungsweisung interpretiert. Ein positiver medizinischer Test wird unkritisch als Bestätigung für die Krankheit akzeptiert, obwohl auch andere Ursachen für das Zustandekommen eines positiven Testergebnisses in Frage kämen.¹⁰³

Die oben aufgeführten kognitiven Mechanismen von System 1 erzeugen stets ein Bild von kohärent aufeinander folgenden Sequenzen von Ereignissen und Handlungen.¹⁰⁴ Umgekehrt sind Informationen mit der Struktur einer Geschichte für System 1 deutlich leichter zugänglich als abstrakte Sachverhalte, die sich nicht unmittelbar kausal interpretieren lassen. Zu der letzteren Kategorie gehören insbesondere Aufgaben auf dem Gebiet der Logik und statistisches Denken. Im Rahmen des System-1/System-2-Ansatzes fällt die Verarbeitung solcher Aufgaben in den Zuständigkeitsbereich von System 2. Weil System 2 nur sparsam eingesetzt wird,¹⁰⁵ kann es vor allem in solchen Situationen zu Urteilsverzerrungen kommen, in denen das logische und statistische Denken erforderlich ist, der Einsatz von System 2 aber trotzdem ausbleibt. Außerdem kann das von System 1 vermittelte innere Gefühl der Richtigkeit so stark sein, dass die Notwendigkeit einer Aktivierung von System 2 nicht erkannt bzw. System 2 ausschließlich zur Unterstützung der heuristischen Lösung verwendet wird.

Insgesamt kann der Entstehungsmechanismus von Biases gemäß dem HB-Ansatz wie folgt zusammengefasst werden: In bestimmten Situationen liefert das heuristische System 1 fehlerbehaftete Urteile und Lösungsvorschläge, die sich verfestigen, weil das rationale System 2 seine interventionistische Rolle nicht erfüllt.¹⁰⁶

¹⁰² Schließlich würden weitere (zum aktuellen Zeitpunkt nicht offengelegte) Eigenschaften wie Bestechlichkeit oder Engstirnigkeit die Führungskraft sehr schnell in einem schlechten Licht erscheinen lassen.

¹⁰³ Jeder medizinische Test birgt stets die Gefahr einer Fehldiagnose, die sich in Form der Falsch-positiv- bzw. Falsch-negativ-Rate ausdrückt, vgl. Gigerenzer 2013, S. 220.

¹⁰⁴ In diesem Zusammenhang siehe das Konzept der narrativen Strukturen von Schank und Ableson 1995.

¹⁰⁵ Siehe das „Prinzip des geringsten Aufwands“ im Zusammenhang mit System 2 weiter oben.

¹⁰⁶ Vgl. Kahneman 2003, S. 710. System 1 gilt zwar im Vergleich zu System 2 als deutlich schlechter im statistischen Denken, allerdings kann System 2 aufgrund der beschränkten Rechen- und Informationsverarbeitungskapazitäten des Menschen ebenfalls zu Fehlern führen, vgl. Witt 2018, S. 39.

Die Rollen der beiden Systeme beim Zustandekommen von Biases können am besten anhand der Logikaufgaben aus dem Cognitive-Reflection-Test (CRT) von Frederick 2005 verdeutlicht werden. Der CRT ist so konzipiert, dass die intuitive Antwort aus dem System-1-Denkmodus heraus falsch ist. Erst infolge von gewissen Überlegungen im System-2-Denkmodus kann die richtige Antwort erschlossen werden. Die berühmte Ball/Schläger-Aufgabe aus dem CRT lautet sinngemäß wie folgt: Ein Ball und ein Schläger kosten zusammen 1 € und 10 Cent. Der Schläger ist 1 € teurer als der Ball. Wie viel kostet der Ball?¹⁰⁷ Die intuitive und schnelle Antwort lautet i. d. R. 10 Cent. Zugleich ist diese Antwort falsch, weil dann der Schläger gemäß dem Wortlaut aus der Aufgabenstellung 1 € kosten müsste, damit die Gesamtsumme gewahrt bleibt, was aber einen Preisunterschied von 90 Cent und nicht 1 € bedeuten würde. Auf die richtige Lösung – 1,05 € für den Schläger und 5 Cent für den Ball – kommt ein Entscheider in den meisten Fällen nur dann, wenn er bewusst weitere plausible Preiskombinationen hinsichtlich der Voraussetzungen aus der Aufgabenstellung prüft.

Nachfolgend werden die „klassischen Heuristiken“¹⁰⁸ des HB-Ansatzes zusammen mit einigen daraus resultierenden Biases unter die Lupe genommen.¹⁰⁹

2.2.2.2 Repräsentativitätsheuristik

Die *Repräsentativitätsheuristik* (engl. Representativeness Heuristic) kommt bei der Beantwortung von zwei Fragetypen zum Einsatz.¹¹⁰ Erstens wird damit die Wahrscheinlichkeit für die Zugehörigkeit eines Einzelfalls zu einer bestimmten Kategorie geschätzt. Zweitens soll mithilfe der Repräsentativitätsheuristik die Frage beantwortet werden, mit welcher Wahrscheinlichkeit sich ein Ereignis innerhalb eines bestimmten Zufallsprozesses einstellen wird. Im ersten Fall wird eine unsichere Diagnose gemacht (z. B. eine Person wird eher als Vertreter einer bestimmten Berufsgruppe identifiziert) und im zweiten Fall eine Vorhersage getroffen (z. B. die Auslosung einer scheinbar zufällig angeordneten Zahlenkombination 6, 17, 21, 27, 34, 47 ist wahrscheinlicher als die Auslosung einer streng geordneten Zahlenreihe von 1 bis 6).¹¹¹ Die Verwendung der Repräsentativitätsheuristik erweist sich bei der Schätzung von Wahrscheinlichkeiten als besonders nützlich, weil

- die Schätzung sehr einfach ist und wenig an kognitiven Ressourcen erfordert,

¹⁰⁷ Vgl. Frederick 2005, S. 26f.

¹⁰⁸ Langer und Rogowski 2009, S. 183.

¹⁰⁹ Ergänzend dazu werden in Kapitel 2.4.3 Urteilsverzerrungen speziell im Kontext des Management Reporting behandelt.

¹¹⁰ Vgl. Tversky und Kahneman 1982, S. 4.

¹¹¹ Vgl. Pfister et al. 2017, S. 134.

- dieser Mechanismus sich auf ein breites Spektrum an Situationen übertragen lässt und
- das Ergebnis der Schätzung in vielen Fällen korrekt ist.¹¹²

Unter bestimmten Bedingungen ist die Repräsentativitätsheuristik aber auch für Fehlurteile verantwortlich. Folgende Biases können aus der Repräsentativitätsheuristik resultieren:¹¹³

- **Basisratenfehler (engl. Base Rate Fallacy):** Bei der Repräsentativitätsheuristik spielt es keine Rolle, wie häufig ein Ereignis, dessen Wahrscheinlichkeit es zu schätzen gilt, überhaupt a priori in der Grundgesamtheit vorkommt. Eine introvertierte und ordnungsliebende Person wird bspw. mit höherer Wahrscheinlichkeit als Bibliothekar und nicht als Landwirt eingeordnet, weil diese Merkmale mit den im Gedächtnis gespeicherten Informationen über einen prototypischen Bibliothekar übereinstimmen. Hierbei wird aber nicht dem Umstand Rechnung getragen, dass es in der Gesamtbevölkerung rein statistisch gesehen mehr Landwirte als Bibliothekare gibt. Bei der großen Zahl an Landwirten ist es durchaus wahrscheinlich, dass einige Vertreter dieser Berufsgruppe introvertiert und ordnungsliebend sind.¹¹⁴
- **Konjunktionsfehlschluss (engl. Conjunction Fallacy):** Nach den Regeln der Wahrscheinlichkeitstheorie kann die Verknüpfung von zwei disjunkten Ereignissen nicht wahrscheinlicher sein als jedes einzelne Ereignis für sich allein genommen. Aufgrund der Repräsentativitätsheuristik verstoßen aber viele Entscheidungsträger gegen dieses Prinzip und begehen den Konjunktionsfehlschluss. Tversky und Kahneman 1983 illustrieren diese Urteilsverzerrung mithilfe des berühmten Linda-Problems.¹¹⁵ Den Teilnehmern der Studie wurde die nachfolgende Beschreibung einer Person namens Linda vorgelegt, die sinngemäß wie folgt lautet:

Linda ist 31 Jahre alt, Single, freimütig und sehr intelligent. Sie hat Philosophie im Hauptfach studiert. Als Studentin interessierte sie sich sehr für Themen wie Diskriminierung und soziale Gerechtigkeit, und sie nahm auch an Anti-Atomkraft-Protesten teil.

Daraufhin wurden die Probanden gefragt, welche der folgenden Alternativen ihrer Meinung nach wahrscheinlicher ist:

(1) Linda ist eine Bankkassiererin.

¹¹² Vgl. Teigen 2016, S. 205.

¹¹³ Vgl. im Folgenden Pfister et al. 2017, S. 135f.

¹¹⁴ Mathematisch richtige Wahrscheinlichkeiten unter Berücksichtigung der Basisrate lassen sich mithilfe der Bayes-Regel bestimmen.

¹¹⁵ Das Linda-Problem wurde seitdem in zahlreichen weiteren empirischen Studien untersucht und dient oft als Anknüpfungspunkt in der Debatte um die Rationalität von Heuristiken, siehe Bar-Hillel und Neter 1993; Hertwig und Gigerenzer 1999.

(2) Linda ist eine Bankkassiererin und in der feministischen Bewegung aktiv.

Die überwiegende Mehrzahl der Befragten stufte die zweite Alternative als wahrscheinlicher ein. Dies ist ein klarer Beweis für den Konjunktionsfehlschluss, denn statistisch betrachtet kann die Konjunktion von Ereignissen (hier Bankkassiererin und Feministin) nie größer sein als jedes Einzelereignis für sich allein genommen (hier Bankkassiererin). Das gemeinsame Auftreten von beiden Merkmalen im oberen Beispiel erscheint deshalb als wahrscheinlicher, weil die beschriebene Person besser in das Bild einer feministischen Bankkassiererin als nur einer Bankkassiererin passt. Selbst bei der Wiederholung dieses Experiments mit statistikaffinen Doktoranden des Studiengangs Entscheidungswissenschaft der Stanford Graduate School of Business stuften ganze 85 % der Teilnehmer die zweite Alternative als wahrscheinlicher ein.¹¹⁶ Offenbar übt die Repräsentativitätsheuristik selbst bei Experten in manchen Situationen wie dem Linda-Problem einen viel stärkeren Einfluss aus als die Logik.

- Vernachlässigung der Regression zur Mitte: Die Repräsentativität von extremen Ereignissen eines Zufallsprozesses wird überschätzt. Mit anderen Worten, man nimmt beim Zustandekommen eines seltenen Ereignisses fälschlicherweise an, dass auch künftige Ereignisse ähnlich ausgeprägt sein werden. Die statistisch zu erwartende Tendenz zur Rückkehr zum Mittelwert nach einem seltenen Ergebnis wird übersehen, was die Bildung von falschen Schlussfolgerungen befördert. Kahneman 2014 illustrierte dieses Phänomen anhand eines Beispiels aus seiner persönlichen Laufbahn als Psychologe in der israelischen Armee.¹¹⁷ Demnach äußerten erfahrene Ausbilder von Piloten die Meinung, dass nach einem Lob für eine besonders weiche Landung i. d. R. eine harte Landung folge, während der umgekehrte Fall meist in Kombination mit einem Tadel auftritt. Die Ausbilder sprachen sich folglich gegen das verbale Lob und für die Kritik als Mittel zur Förderung des Lernprozesses aus. Dabei wurde aber außer Acht gelassen, dass sowohl besonders weiche als auch besonders harte Landungen eine Abweichung von der „durchschnittlichen“ Landung darstellen und gemäß dem Prinzip der Regression zur Mitte auf eine besonders harte Landung eine tendenziell weichere Landung et vice versa zu erwarten ist.
- Spielerfehlschluss (engl. Gambler's Fallacy bzw. Hot Hand Fallacy): Einerseits beschreibt dieses Phänomen, dass das „charakteristische“ Zufallsmuster auch in kurzen Abfolgen eines Zufallsprozesses erwartet wird. Wenn bspw. bei einer viermaligen Wiederholung eines Münzwurfs viermal nacheinander „Kopf“ geworfen wird, steigt die subjektive Erwartung dafür, dass beim nächsten Wurf eine „Zahl“ herauskommen wird. Hierbei wird die Tatsache ignoriert, dass bei einem fairen Münzwurf die statistische Wahrscheinlichkeit für einen der beiden Ausgänge immer exakt

¹¹⁶ Vgl. Tversky und Kahneman 1983, S. 298.

¹¹⁷ Vgl. Kahneman 2014, S. 219ff.

50 % beträgt und zwar unabhängig von den vorausgehenden Wiederholungen. Des Weiteren ist den Entscheidungsakteuren oft nicht bewusst, dass das für einen zugrundeliegenden Zufallsprozess repräsentative Ergebnismuster in Übereinstimmung mit dem Gesetz der großen Zahlen erst bei einer großen Zahl an Wiederholungen verlässlich zu erwarten ist. Außerdem wird der Spielerfehlschluss auch im sportlichen Kontext beschrieben. So werden im Basketball mehrere Treffer nacheinander im Sinne einer „Glücksträhne“ interpretiert, was aber in diesem lokalen Fall statistisch gesehen nicht mehr als eine illusorische Korrelation ist.¹¹⁸ Zusammenfassend werden beim Spielerfehlschluss scheinbare und nicht repräsentative Muster eines Zufallsprozesses erkannt und darauf aufbauend fehlerbehaftete Prognosen gemacht.

Die oben aufgeführten Beispiele stellen nur eine kleine Auswahl an Verhaltensphänomenen dar, welche auf die Repräsentativitätsheuristik zurückgeführt werden können. Trotz einiger Einwände theoretischer Natur erwies sich dieses Konzept als ein flexibles Instrument zur Beschreibung vieler intuitiver Urteile. Eine fundierte Darstellung der Repräsentativitätsheuristik zusammen mit weiteren anschaulichen Beispielen für mögliche Biases findet sich u. a. bei Teigen 2016.

2.2.2.3 Verankerungsheuristik

Ist bei der Schätzung einer quantitativen Größe die richtige Antwort nicht oder nur grob bekannt, wird bei der *Verankerungsheuristik* (engl. Anchoring) ein Ausgangswert – sprich Anker – für die zu approximierende Größe gesetzt und dieser im Laufe eines Adjustierungsprozesses sukzessive angepasst. Dabei kann der Ankerwert für die Ausgangsproblemstellung irrelevant sein oder die Anpassung erfolgt unzureichend. In solchen Fällen führt die Verankerungsheuristik systematisch zu verzerrten Einschätzungen. Für das Verankerungsphänomen existieren zwei gängige Erklärungsansätze. Zum einen wird davon ausgegangen, dass der Anker zunächst bewusst falsch gesetzt wird. Je nachdem, ob viele gute Gründe für weitere Anpassungen sprechen, fällt die Adjustierung entsprechend stärker oder schwächer aus. Dieser Ansatz unterstellt einen bewussten Adjustierungsprozess.¹¹⁹ Zum anderen kann der Ankereffekt mit dem sog. Priming erklärt werden. Hierbei löst der in einen bestimmten Kontext eingebettete Anker gemäß dem bereits bekannten Konzept der assoziativen Aktivierung alle mit ihm verknüpften Informationselemente im Gedächtnis aus, auch wenn manche davon für das

¹¹⁸ Vgl. Gilovich et al. 1985, S. 295. Gleichzeitig konnte in der Studie von Raab et al. 2012 für die Sportart Volleyball in 50 % der Fälle ein positiver Zusammenhang zwischen aufeinanderfolgenden Punktgewinnen statistisch nachgewiesen werden. Dieser Befund wird damit erklärt, dass im Gegensatz zum Basketball die gegnerischen Spieler beim Volleyball in einem weit geringeren Maße physischen Einfluss auf den Hot-Hand-Spieler ausüben können und dieser somit seinen Erfolgslauf viel besser zur Geltung bringen kann.

¹¹⁹ Vgl. Epley und Gilovich 2006, S. 312f.

eigentliche Problem irrelevant sind. Mittlerweile setzte sich der zweite Erklärungsansatz durch, welcher auf dem Prinzip der assoziativen Aktivierung basiert und mit System 1 in Verbindung steht.

Folgende Biases können aus der Verankerungsheuristik resultieren:

- Fehleinschätzung numerischer Größen: Tversky und Kahneman 1974 demonstrieren zum ersten Mal experimentell, wie ein willkürlich vorgegebener Ankerwert imstande ist, individuelle Entscheidungen zu beeinflussen. In ihrem Experiment setzten die Autoren im ersten Schritt zwei separate Gruppen von Studenten vor ein Glücksrad. Bei einer Gruppe blieb der Zeiger scheinbar zufällig auf der Zahl 65 und bei der anderen Gruppe auf der Zahl 10 stehen. Im zweiten Schritt musste jeder Teilnehmer den Anteil afrikanischer Staaten in der Organisation der Vereinten Nationen schätzen. Die Probanden aus der Gruppe mit der größeren Glücksradzahl schätzten den fragten Anteil im Durchschnitt auf etwa 45 %, wohingegen die Probanden aus der Gruppe mit der kleineren Glücksradzahl im Durchschnitt eine Schätzung von 25 % abgaben.¹²⁰ Weil in der Studie davon auszugehen war, dass die Mehrheit der Teilnehmer die richtige Antwort nicht wissen konnte, wurde offenbar eine für die Beantwortung der Frage irrelevante Zahl als Orientierungswert genommen und in den Entscheidungsprozess unbewusst integriert.
- Rückschaufehler (engl. Hindsight Bias): Pfister et al. 2017 verdeutlichen das Prinzip des Rückschaufehlers anhand eines fiktiven Beispiels. Gibt bspw. eine Person ohne Vorkenntnisse eine Schätzung für die Länge des Rheins in Höhe von 1 150 km ab, geht die Erinnerung an diese Zahl i. d. R. mit der Zeit nach und nach verloren. Erfährt also die Person nach einer gewissen Zeit die tatsächliche Länge des Rheins von 1 320 km, dient diese Angabe nunmehr als der Ankerwert für die ursprüngliche Schätzung der Rheinlänge. Der Rückschaufehler bedeutet, dass die Erinnerung in Richtung des Ankerwerts „korrigiert“ wird. Im vorliegenden Beispiel würde sich der Entscheidungsträger seine ursprüngliche Schätzung nach oben anpassen – z. B. 1 250 km. Läge die ursprüngliche Schätzung bei 1 730 km, würde die Person den Erinnerungswert nach unten anpassen – z. B. 1 630 km.¹²¹

Eine fundierte Darstellung der Verankerungsheuristik zusammen mit weiteren anschaulichen Beispielen für mögliche Biases findet sich u. a. bei Bahnik et al. 2016.

¹²⁰ Vgl. Tversky und Kahneman 1974, S. 1128.

¹²¹ Vgl. Pfister et al. 2017, S. 139f.

2.2.2.4 Verfügbarkeitsheuristik

Nach Tversky und Kahneman 1982 kommt die *Verfügbarkeitsheuristik* (engl. Availability Heuristic) bei zwei Fragetypen zum Einsatz.¹²² Zum einen wird damit die Häufigkeit der Exemplare aus einer bestimmten Klasse geschätzt, zum anderen kann damit die Frage nach der (bedingten) Wahrscheinlichkeit eines Ereignisses beantwortet werden. Die subjektive Einschätzung der Häufigkeit oder Wahrscheinlichkeit fällt intuitiv umso höher aus, je leichter bzw. schneller sich der Entscheidungsträger Beispiele für den fragten Sachverhalt in Erinnerung rufen kann.¹²³ Tversky und Kahneman 1973 illustrieren dieses Prinzip am folgenden Beispiel: In einem Experiment sollten die englischsprachigen Teilnehmer einschätzen, ob es wahrscheinlicher ist, dass ein zufällig ausgewähltes Wort aus einem englischen Buch eher ein „K“ am Anfang als an der dritten Stelle des Wortes hat. Die meisten Probanden hielten ein „K“ am Anfang des Wortes tatsächlich für wahrscheinlicher, obwohl es im Englischen doppelt so viele Wörter mit einem „K“ an dritter Stelle des Wortes als am Anfang gibt.¹²⁴ Dieses Antwortverhalten kann damit erklärt werden, dass Beispiele für Wörter, die mit einem „K“ anfangen, viel leichter in den Sinn kommen – und somit „verfügbarer“ sind – als Beispiele aus der anderen Wortkategorie.¹²⁵

Ähnlich wie die Repräsentativitätsheuristik ist die Verfügbarkeitsheuristik im Allgemeinen sehr effizient, weil häufig vorkommende Ereignisse i. d. R. auch besser im Gedächtnis abgespeichert werden als seltene Vorkommnisse. Unter bestimmten Bedingungen ist die Verfügbarkeitsheuristik zugleich für gewisse Fehlurteile verantwortlich. Folgende Biases können in diesem Zusammenhang genannt werden:¹²⁶

- Beeinflussung aufgrund von Lebhaftigkeit und Präsenz: Lebhaftere und einschneidende individuelle Erfahrungen prägen sich stark ins Gedächtnis ein und können so die objektive Wahrnehmung verzerren. Eine Person, die bspw. vor kurzem einen schweren Verkehrsunfall auf der Autobahn erlebte, wird die Wahrscheinlichkeit für schwere Autobahnunfälle tendenziell viel höher einschätzen als eine Person, die noch nie eine solche Erfahrung persönlich machen musste. Selbst wenn die persönliche Erfahrung fehlt, können Ereignisse, die dauerhaft – z. B. aufgrund medialer

¹²² Die Fragetypen überschneiden sich mit denen der Repräsentativitätsheuristik. Tatsächlich können, wie später noch gezeigt wird, manche Biases mithilfe von beiden Heuristikarten erklärt werden.

¹²³ Vgl. Tversky und Kahneman 1982, S. 11.

¹²⁴ Vgl. im Folgenden Tversky und Kahneman 1973, S. 211.

¹²⁵ Kognitivpsychologisch kann dies mit dem im Zusammenhang mit System 1 eingeführten Konzept der assoziativen Aktivierung begründet werden. Ein Wort, das mit dem Buchstaben „K“ anfängt, kann leichter aus dem Gedächtnis abgerufen werden. Das erinnerte Wort führt wiederum zur Aktivierung von weiteren Wörtern mit einem „K“ am Anfang usw., was schließlich in einem selbstverstärkenden Prozess mündet.

¹²⁶ Vgl. im Folgenden Pfister et al. 2017, S. 137f.

Berichterstattung – in der Wahrnehmung präsent sind, zur Überschätzung der Häufigkeit bzw. Wahrscheinlichkeit dieser Ereignisse führen. Dieser Effekt kam im Experiment von Russo und Schoemaker 1989 deutlich zum Vorschein. Die Teilnehmer der Studie mussten eine Einschätzung geben, ob in den USA jährlich mehr Menschen an Lungenkrebs oder an Autounfällen sterben. Die meisten Befragten schätzten die Autounfälle als das größere Risiko ein, obwohl tatsächlich in den USA zum damaligen Zeitpunkt jährlich dreimal so viele Menschen an Lungenkrebs wie infolge eines Autounfalls verstarben.¹²⁷ Die Autoren führen diese Überschätzung darauf zurück, dass in den Medien weitaus häufiger und intensiver über Verkehrsunfälle als über Fälle von Lungenkrebs berichtet wird. Insgesamt können also sehr lebhaft persönliche Erfahrungen und Ereignisse, die ständig in der Wahrnehmung präsent sind, zur Fehleinschätzung von relativen Häufigkeiten und dadurch zu Biases führen.

- Egozentrische Perspektive: Eigens erlebte Erfahrungen sind kognitiv verfügbarer und haben somit mehr Gewicht als abstrakte Vorstellungen aus der Perspektive einer anderen Person. Als bspw. Ross und Sicoly 1979 verheiratete Paare nach dem individuellen Beitrag jedes Einzelnen zur Hausarbeit (z. B. Müllbeseitigung, Staubsaugen usw.) befragten, summierten sich die Angaben regelmäßig auf über 100 %.¹²⁸ Aus einer egozentrischen Perspektive ist es offenbar viel einfacher und schneller, Beispiele für den eigenen Einsatz im Haushalt aus der Erinnerung zu rekonstruieren als für den Einsatz des Ehepartners. Dies führte systematisch zur Fehleinschätzungen der objektiven Häufigkeiten.
- Illusorische Korrelation: Ähnlich wie bei dem bereits im Zusammenhang mit der Repräsentativitätsheuristik vorgestellten Spielerfehlschluss kann die Verfügbarkeitsheuristik die Entstehung von illusorischen Korrelationen bei der Beurteilung von verknüpften Ereignissen begünstigen. Soll bspw. die Stärke des Zusammenhangs zwischen Haschischkonsum und Straffälligkeit beurteilt werden, wird vermutlich diesen beiden Merkmalen eine starke Korrelation bescheinigt.¹²⁹ Es fällt ziemlich leicht, sich Bilder mit Joints rauchenden Straftätern vorzustellen. Aus einer starken assoziativen Verknüpfung wird wiederum unreflektiert auf einen validen

¹²⁷ Vgl. Russo und Schoemaker 1989, S. 83. Eine Fehleinschätzung von Risiken kann zur ernststen Gefahr werden. So führte die Flugangst nach dem Terroranschlag vom 11. September 2001 dazu, dass viele Menschen Flugreisen gemieden haben und stattdessen auf das Auto umgestiegen sind. Weil aber Flugreisen statistisch betrachtet viel sicherer sind als Reisen mit dem Auto, kam es in den Folgemonaten nach dem Terroranschlag zu einem übermäßigen Anstieg an tödlichen Verkehrsunfällen auf den US-amerikanischen Straßen, vgl. Gigerenzer 2013, S. 20ff.

¹²⁸ Vgl. Ross und Sicoly 1979, S. 325f.

¹²⁹ Hierbei soll selbstverständlich der Haschischkonsum für sich allein genommen nicht als Straftat betrachtet werden.

statistischen Zusammenhang der beiden Merkmale geschlossen.¹³⁰ Bei der Beurteilung der obigen Frage müssten aber auch Fallkonstellationen, wie straffällig gewordene Personen, die kein Haschisch konsumieren, oder nicht straffällig gewordene Haschischkonsumenten, in die Betrachtung einbezogen werden.

Im Zusammenhang mit der Verfügbarkeitsheuristik stellt sich die Frage, ob die kognitive Leichtigkeit, mit der einem Entscheidungsträger Beispiele für einen Sachverhalt einfallen, oder die absolute Häufigkeit der erinnerten Beispiele letztlich für die Beurteilung ausschlaggebend ist. Diese Frage konnten Schwarz et al. 1991 in ihrer experimentellen Studie klären. In dieser Studie sollten sich die Teilnehmer in einer Gruppe an sechs und in einer anderen Gruppe an zwölf Fälle erinnern, in denen sie sehr selbstsicher aufgetreten sind und sich mit ihrer Meinung durchgesetzt konnten. Anschließend wurden alle Teilnehmer nach ihrer Einschätzung des eigenen Durchsetzungsvermögens gefragt. Weil es den meisten Menschen deutlich leichter fällt, sechs anstelle von zwölf Episoden aus der Erinnerung zu rufen, eignete sich dieses Experimentdesign besonders gut, um die Frage nach der Wirkungsweise der Verfügbarkeitsheuristik zu beantworten. Tatsächlich schätzten die Probanden aus der Gruppe mit sechs Erinnerungsepisoden ihr Durchsetzungsvermögen viel höher ein als die Probanden aus der anderen Gruppe.¹³¹ Dies ist ein Beweis dafür, dass nicht die Menge der erinnerten Beispiele, sondern die kognitive Leichtigkeit der Erinnerung hinter der Verfügbarkeitsheuristik operiert. Dabei können unterschiedliche Auslöser, wie Emotionen oder Zugreifbarkeit, für die kognitive Leichtigkeit des Erinnerungsprozesses verantwortlich sein. Eine fundierte Darstellung der Verfügbarkeitsheuristik zusammen mit weiteren anschaulichen Beispielen für mögliche Biases findet sich u. a. bei Reber 2016.

Mithilfe der *Attributsubstitutionstheorie* von Kahneman und Frederick 2002 werden die empirisch nachgewiesenen Verhaltensphänomene aus dem HB-Forschungsprogramm in einen übergeordneten theoretischen Zusammenhang gebracht. Die Attributsubstitutionstheorie postuliert, dass bei der Beantwortung eines komplexen Zielattributs dieses durch ein einfacher zu beantwortendes heuristisches Attribut substituiert wird.¹³² So wird eine schwierig einzuschätzende Wahrscheinlichkeit eines Ereignisses (Zielattribut) nach Maßstäben der Ähnlichkeit oder kognitiver Leichtigkeit (heuristisches Attribut) beurteilt. Anhand der Stärke des heuristischen Urteils werden Schlussfolgerungen im Hinblick auf die Ausprägung des Zielattributs gezogen, was bei fehlender Korrelation der beiden Größen zur Entstehung von Biases führt.¹³³

¹³⁰ Vgl. Kahneman 2014, S. 535.

¹³¹ Vgl. Schwarz et al. 1991, S. 196f.

¹³² Vgl. Kahneman 2003, S. 707.

¹³³ Vgl. Kahneman und Frederick 2002, S. 53f.

2.2.3 Schnelle und sparsame Heuristiken

2.2.3.1 Charakteristika und Einordnung

Während beim HB-Ansatz der Schwerpunkt der Untersuchung auf der Analyse von Urteilsverzerrungen liegt, betont das Forschungsprogramm zu den *schnellen und sparsamen* – engl. fast and frugal (FF) – Heuristiken das Potenzial von einfachen Heuristiken zur Lösung von komplexen Problemen.¹³⁴ Die beiden Ansätze unterscheiden sich außerdem darin, dass die Heuristiken gemäß dem HB-Ansatz ausschließlich auf die intuitiven kognitiven Prozesse zurückgeführt werden, während bei den FF-Heuristiken angenommen wird, dass diese sowohl intuitiv als auch überlegt ausgeführt werden können.¹³⁵ Ansonsten bestehen keine wesentlichen Unterschiede in der Konzeption des Heuristikbegriffs zwischen den beiden Forschungsprogrammen. Genau wie im HB-Ansatz werden unter FF-Heuristiken einfache Faustregeln zur Bildung von Urteilen verstanden. Insbesondere sind für eine FF-Heuristik die folgenden Eigenschaften charakteristisch:¹³⁶

- **Schnelligkeit:** Heuristiken sind schnell (engl. fast) und erfordern keine oder nur sehr wenig an kognitiven Ressourcen.
- **Sparsamkeit:** Heuristiken sind sparsam (engl. frugal), d. h. nur ausgewählte relevante Informationen werden bei der Entscheidungsfindung berücksichtigt und die restlichen Informationen werden ignoriert.¹³⁷
- **Genauigkeit:** Heuristiken resultieren in adäquaten Lösungsergebnissen, die manchmal sogar genauer sind als Ergebnisse von komplexen analytischen Verfahren (z. B. Regressionsanalysen).

Mithilfe der FF-Heuristiken gelangt also ein Entscheider in kurzer Zeit und unter Verwendung einer kleinen Informationsmenge zu hinreichend genauen Urteilen. Der FF-

¹³⁴ Die Begründung des FF-Ansatzes geht auf Gigerenzer und Todd 1999b; Gigerenzer 2004; Gigerenzer und Gaissmaier 2011; Gigerenzer et al. 2011 zurück.

¹³⁵ Die Heuristiken gemäß dem FF-Ansatz führen aber stets zur Entstehung von „Bauchgefühlen“. Bauchgefühle tauchen rasch im Bewusstsein auf und ein Entscheider ist bereit, danach zu handeln, obwohl er das Gefühl der inneren Richtigkeit nicht genau erklären kann, vgl. Gigerenzer 2007, S. 57.

¹³⁶ Vgl. im Folgenden Gigerenzer und Gaissmaier 2011, S. 454.

¹³⁷ Der FF-Ansatz fußt auf dem Konzept der begrenzten Rationalität (engl. bounded rationality) von Simon 1955; 1956. In Abgrenzung zu Modellen der unbegrenzten Rationalität – auf denen das Konstrukt des Homo oeconomicus basiert, siehe Kapitel 2.1.1 – bilden menschliche Entscheidungsträger mit eingeschränkten kognitiven Kapazitäten und das Vorhandensein von Informationskosten den Ausgangspunkt für Modelle der begrenzten Rationalität, vgl. Todd 2001, S. 51. In diesem Zusammenhang darf jedoch der Begriff der begrenzten Rationalität nicht mit der Optimierung unter Restriktionen verwechselt werden. Im Rahmen der begrenzten Rationalität dienen explizit nur wenige relevante Informationen als Entscheidungsgrundlage, wohingegen bei der Optimierung unter Restriktionen alle verfügbaren Informationen solange verarbeitet werden, bis die Kosten der weiteren Informationssuche die daraus resultierende Vorteile übersteigen, vgl. Gigerenzer und Todd 1999b, S. 5ff.

Ansatz beruht auf dem *Satisficing-Prinzip*, d. h., eine Alternative wird sofort akzeptiert, wenn das daraus resultierende Ergebnis in einer zufriedenstellenden Weise zur Befriedigung des Anspruchsniveaus führt.¹³⁸

Des Weiteren liegt dem FF-Ansatz das Konzept der *ökologischen Rationalität* zugrunde. Nach diesem Rationalitätsverständnis sind FF-Heuristiken erst dann rational und folgerichtig, wenn die für den Einsatz einer Heuristik benötigten Fähigkeiten eines Individuums im hohen Maße an die gegebenen Umweltstrukturen adaptiert sind.¹³⁹ So wird bspw. bei der Blickheuristik zum Fangen von hohen Bällen aus der Luft die evolvierte Fähigkeit zur Verfolgung von bewegten Objekten mit den Augen ausgenutzt.¹⁴⁰ Hierbei wandelt das Auge die komplexe Flugbahn, die der Ball in der Luft beschreibt, visuell in eine einfache gerade Linie um und die Blickheuristik nutzt dies aus, ohne dass dabei eine komplexe und zeitaufwendige Berechnung der parabelförmigen Flugbahn notwendig wäre. Alle Heuristiken sind bereichsspezifisch. So können mithilfe der Blickheuristik Probleme gelöst werden, die eine Kollision mit bewegten Objekten betreffen.¹⁴¹

Die Effizienz der FF-Heuristiken im Vergleich zu komplexen statistischen Lösungsmethoden kommt am meisten in Situationen der Ungewissheit zum Tragen (siehe Abbildung 6).¹⁴² Eine Situation der Ungewissheit zeichnet sich vor allem dadurch aus, dass weder alle Konsequenzen einer Alternative noch die entsprechenden Eintrittswahrscheinlichkeiten der Konsequenzen bekannt sind und sich somit ein Optimum nicht eindeutig bestimmen lässt.¹⁴³ In solchen Situationen sind einfache Heuristiken aufgrund ihrer Robustheit gegenüber komplexen statistischen Optimierungsmodellen mit vielen zu schätzenden Parametern im Vorteil, weil nur die tatsächlich relevanten Informationen in die Entscheidung einbezogen werden und keine bloße Anpassung an die vielen vom Zufall getriebenen Datenmuster erfolgt.¹⁴⁴ Auf der anderen Seite des Kontinuums sind klassische Risikosituationen angeordnet. Im Gegensatz zur Ungewissheit ist eine Risikosituation vor allem dadurch gekennzeichnet, dass alle möglichen Alternativen mit den daraus resultierenden Konsequenzen und entsprechenden Eintrittswahrscheinlich-

¹³⁸ Vgl. Gigerenzer und Goldstein 2011, S. 34.

¹³⁹ Vgl. Gigerenzer und Todd 1999a, S. 13.

¹⁴⁰ Eine ausführliche Beschreibung und Analyse der Blickheuristik finden sich u. a. bei McLeod und Dienes 1996.

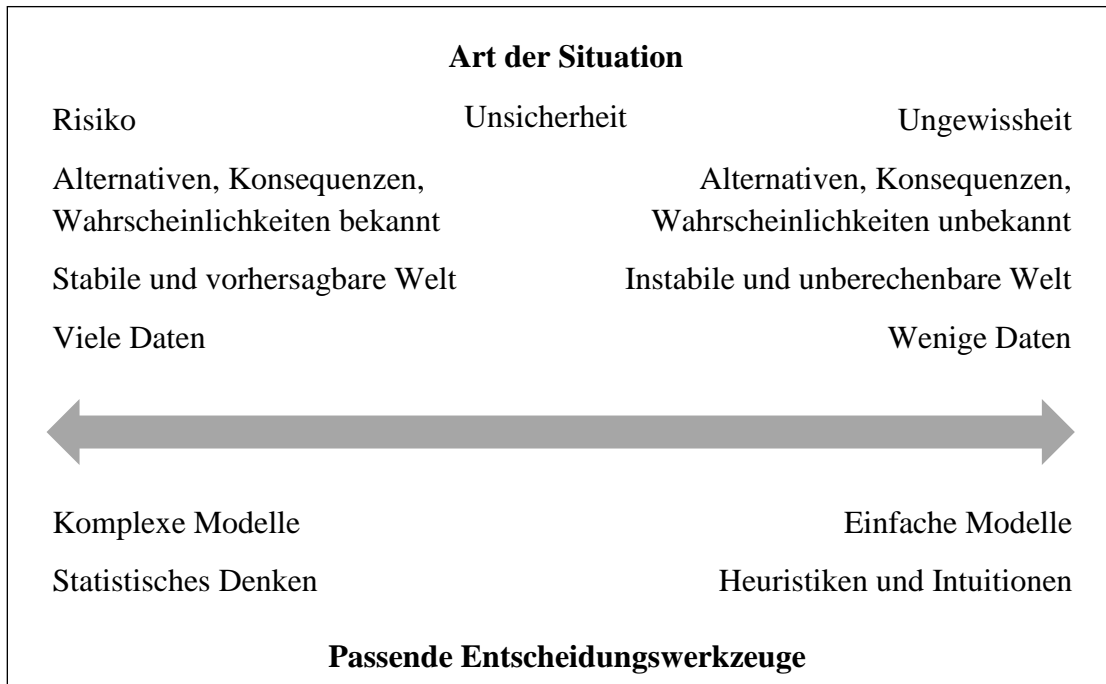
¹⁴¹ Vgl. Gigerenzer und Gaissmaier 2006, S. 332.

¹⁴² Die Begriffe Risiko, Unsicherheit und Ungewissheit wurden bereits in Kapitel 2.1.1 definiert und voneinander abgegrenzt.

¹⁴³ Viele Entscheidungen in der Realität müssen unter mehr oder weniger stark ausgeprägter Unsicherheit getroffen werden, vgl. Mousavi und Gigerenzer 2014, S. 1672; Volz und Gigerenzer 2012, S. 1.

¹⁴⁴ Vgl. Neth 2014, S. 51f. Dieses Phänomen trägt die Bezeichnung Weniger-ist-mehr-Effekt und wird ausführlicher in Kapitel 2.2.3.2 vorgestellt.

Abbildung 6: Passende Entscheidungswerkzeuge bei Risiko und Ungewissheit



Quelle: In Anlehnung an Gaissmaier und Neth 2016, S. 21; Gigerenzer 2013, S. 38.

keiten bekannt sind. In solchen Situationen sind komplexe statistische Verfahren überlegen und führen zu besseren Ergebnissen als vereinfachte Modelle und heuristische Entscheidungen.¹⁴⁵

Kognitionspsychologisch wird angenommen, dass Heuristiken aus den drei nachfolgenden Bausteinen bestehen: Such-, Stopp- und Entscheidungsregel.¹⁴⁶ Die Suchregel bestimmt die Kriterien zur Gewinnung von relevanten Informationen. Die Stoppregele legt fest, wann die Suche nach weiteren Informationen beendet werden soll. Schließlich wird anhand der Entscheidungsregel ein Urteil gebildet. Diese Bausteine gelten für alle Heuristiken und unterscheiden sich lediglich in der genauen Spezifikation. Beispielhaft für die FF-Heuristiken werden nachfolgend die Take-the-Best-Heuristik und die Rekognitionshuristik näher vorgestellt.

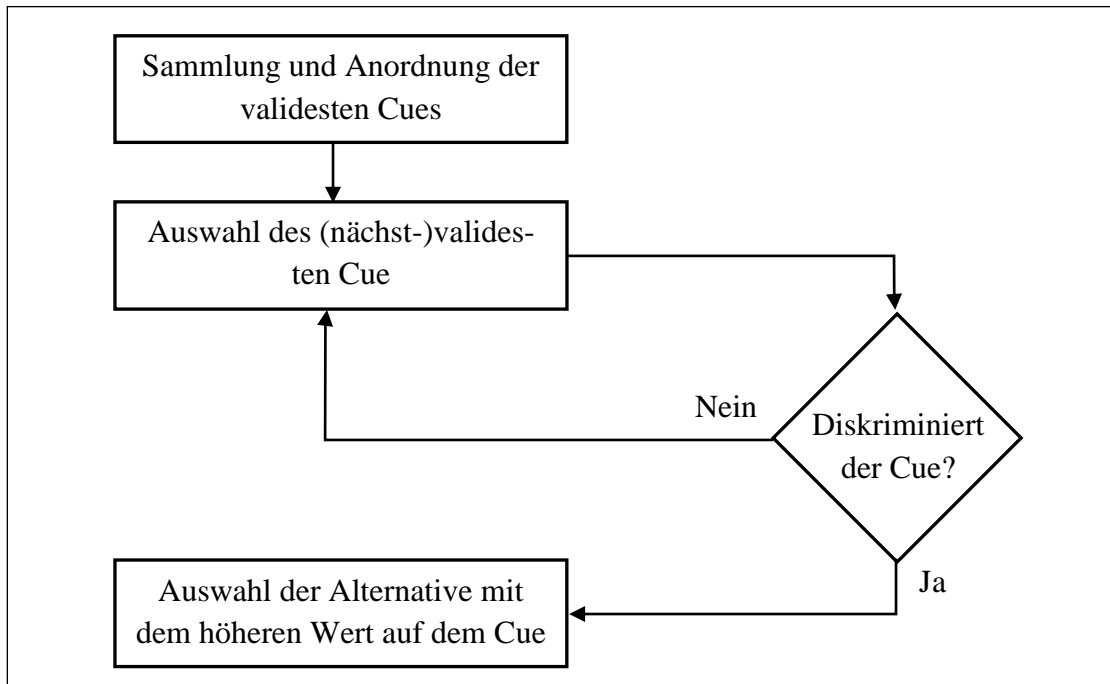
2.2.3.2 Take-the-Best-Heuristik

Bei der *Take-the-Best-Heuristik* (TTB-Heuristik) wird nur anhand des validesten Hinweises (engl. Cue) ein Urteil über das Zielobjekt gebildet. Diskriminiert der valideste

¹⁴⁵ Vgl. Gaissmaier und Neth 2016, S. 20f.

¹⁴⁶ Vgl. Pfister et al. 2017, S. 141.

Abbildung 7: Ablauf der Take-the-Best-Heuristik



Quelle: In Anlehnung an Gigerenzer und Goldstein 1996, S. 653; Pfister et al. 2017, S. 141.

Cue nicht zwischen den Alternativen, wird der zweitvalideste Cue verwendet usw.¹⁴⁷ Ein einziger guter Grund reicht also aus, um die finale Entscheidung treffen zu können und die restlichen Informationen werden bewusst vernachlässigt. Der Ablauf der TTB-Heuristik folgt im Allgemeinen dem folgenden Schema (siehe auch Abbildung 7):¹⁴⁸

- Suchregel: Es werden die relevantesten Unterscheidungsmerkmale im Hinblick auf das Zielobjekt festgelegt und eine Reihenfolge wird beginnend mit dem validesten Cue gebildet. Dabei stützt sich ein Entscheidungsträger bei der Bestimmung der Cues und ihrer Reihenfolge auf seine Erfahrungen.¹⁴⁹
- Stoppregel: Ist der Wert des validesten Cues bei einer Alternative höher als bei den anderen Alternativen, folgt die Entscheidung. Ansonsten werden die Alternativen unter Zuhilfenahme des nächstvalidesten Cues usw. geprüft, bis sich ein Unterschied in den Cue-Werten feststellen lässt oder es keine Cues mehr gibt.

¹⁴⁷ Ein solches Vorgehen wird auch als lexikografischer Prozess bezeichnet, weil die Suche von Begriffen in einem Lexikon nach demselben Prinzip wie die Evaluation von Cues bei der TTB-Heuristik erfolgt.

¹⁴⁸ Vgl. im Folgenden Gigerenzer und Goldstein 1996, S. 653; Pfister et al. 2017, S. 141f.

¹⁴⁹ Die (subjektive) Validität eines Cues i ist definiert als der Quotient aus der Anzahl der korrekten Vorhersagen aufgrund von Cue i und der Anzahl der Paare, bei denen die Werte von Cue i voneinander abwichen, vgl. Gigerenzer und Gaissmaier 2006, S. 351.

- Entscheidungsregel: Die Alternative mit dem höheren Wert auf dem Cue wird gewählt. Waren bei allen Alternativen die Cue-Werte gleich, wird eine Alternative zufällig ausgewählt.

Die TTB-Heuristik ist schnell und sparsam, weil hierfür relativ wenige Informationen erforderlich sind und eine Entscheidung i. d. R. nach wenigen Prüfschritten der Cues erreicht werden kann. Die Anwendung der TTB-Heuristik setzt jedoch immer voraus, dass der Entscheider imstande ist, Wissen über die Validitäten der Cues zu generieren und eine entsprechende Rangordnung mit den validesten Cues zu bilden.¹⁵⁰

Die Effizienz von FF-Heuristiken wird meist im Rahmen von Prognosen, die sich durch eine hohe Unsicherheit auszeichnen, untersucht. Gigerenzer 2007 gibt in diesem Kontext folgendes konkretes Beispiel für die TTB-Heuristik: Eine Gruppe von Studenten der New York University wurde gebeten, den Ausgang von zufällig ausgewählten Spielen aus der National Basketball Association der Saison 1996/97 vorherzusagen. Die Probanden sollten ihre Vorhersagen jeweils auf Grundlage aller gewonnenen Spiele einer Mannschaft (Basisrate) und des aktuellen Halbzeitstands treffen. In mehr als 80 % der Fälle wandten die Entscheider die nachfolgende spezielle Form der TTB-Heuristik an: War die Bilanz der gewonnenen Spiele bei einer Mannschaft um 15 oder mehr größer als bei der anderen Mannschaft, fiel die Entscheidung sofort zugunsten der siegreicheren Mannschaft aus. Ansonsten wurde diejenige Mannschaft ausgewählt, die zur Spielhalbzzeit vorne lag. Diese einfache Regel ist aus rationaler Sicht fehlerbehaftet, weil entweder lediglich die alte Information aller gewonnenen Spiele ohne Berücksichtigung des aktuellen Halbzeitstands herangezogen wird (Konservatismus) oder die Entscheidung nur anhand des Halbzeitstands unter Vernachlässigung der Grundgesamtheit in Form der insgesamt gewonnenen Spiele einer Mannschaft getroffen wird (Basisratenfehler). Eine Computersimulation ergab, dass die oben beschriebene TTB-Heuristik trotz ihrer Verstöße gegen die Rationalität in fast 80 % der Prognosen eine richtige Antwort lieferte und damit genauso gut wie die statistisch korrekte Bayes-Regel abschnitt.¹⁵¹

Wenn bestimmte Anwendungsvoraussetzungen erfüllt sind, kann der TTB-Heuristik insgesamt eine hohe Effizienz und Genauigkeit attestiert werden.¹⁵² Hierbei müssen neben der individuellen Fähigkeit einer Person zur Generierung der Cue-Validitäten auch

¹⁵⁰ Verfügt ein Entscheider über gewisse Cues, kann aber keine Aussage in Bezug auf die Validität treffen, bietet sich die Minimalist-Heuristik an, vgl. Gigerenzer und Goldstein 1996, S. 661. Hierbei wird irgendein Cue zufällig ausgesucht und geprüft, ob sich bei diesem Cue die Alternativen unterscheiden lassen.

¹⁵¹ Vgl. Gigerenzer 2007, S. 159ff. Ein ähnliches Ergebnis konnte für die Fußballspiele der Bundesligasaison 1998 bis 2000 reproduziert werden, vgl. Gröschner und Raab 2006, S. 27ff.

¹⁵² Die TTB-Heuristik funktioniert oft sogar besser als normative Verfahren, bei denen alle verfügbaren Informationen mit einer optimalen Gewichtung in das Modell einfließen, vgl. Gigerenzer und Goldstein 1996, S. 650.

passende Umweltstrukturen vorhanden sein. Das bedeutet, dass sich einerseits die validesten Cues tatsächlich zur Diskriminierung der Qualität des Zielobjekts eignen müssen und andererseits die restlichen Cues für die Entscheidung redundant sind.

2.2.3.3 Rekognitionsheuristik

Die *Rekognitionsheuristik*¹⁵³ lässt sich einfach wie folgt formulieren: Wenn eines von zwei Objekten wiedererkannt wird und das andere nicht, dann soll daraus die Schlussfolgerung gezogen werden, dass das wiedererkannte Objekt einen höheren Wert auf dem Zielkriterium hat.¹⁵⁴ Die Rekognitionsheuristik ist sogar noch schneller und sparsamer als die TTB-Heuristik, weil bei ihr keine Notwendigkeit zur Bildung einer Rangordnung besteht und nur ein Cue im Hinblick auf die Wiedererkennung geprüft werden muss.

Bei der Rekognitionsheuristik machen sich die Menschen ihre im Laufe der Evolution entwickelte Fähigkeit der Wiedererkennung von Bildern, Gesichtern und Stimmen zunutze. Personen können für gewöhnlich aus einem großen Strom an Informationen die neuen von den alten Inhalten sehr leicht unterscheiden. Die Mächtigkeit des menschlichen Wiedererkennungsvermögens wurde im psychologischen Experiment von Standing 1973 eindrucksvoll unter Beweis gestellt. Die Teilnehmer in dieser Studie bekamen bis zu 10 000 Bilder für je fünf Sekunden zu sehen und konnten nach zwei Tagen im Durchschnitt 7 280 Bilder richtig wiedererkennen.¹⁵⁵ Wie bei allen FF-Heuristiken hängt die Effizienz der Rekognitionsheuristik von ihrer ökologischen Rationalität ab. Die Rekognitionsheuristik ist nur dann erfolgreich, wenn das (fehlende) Wiedererkennen nicht zufällig ist und stark mit dem Zielkriterium korreliert, wobei das Wissen über die Richtung der Korrelation entweder bereits genetisch kodiert ist oder über Erfahrungen erlernt wird.

Unter der Annahme, dass die ökologische Rationalität erfüllt ist, beruht die Rekognitionsheuristik auf dem *Weniger-ist-mehr-Effekt*. Dieses Prinzip soll anhand eines Beispiels erläutert werden. Angenommen, ein Entscheidungsträger erzielt in 80 % der Fälle, in denen er die Rekognitionsheuristik anwenden kann, tatsächlich die richtige Antwort – die Rekognitionsvalidität beträgt hier 80 %. Des Weiteren soll angenommen

¹⁵³ Die Bezeichnung Wiedererkennungsheuristik ist ebenso geläufig.

¹⁵⁴ Die Rekognitionsheuristik sollte nicht mit der bereits im Rahmen des HB-Ansatzes vorgestellten Verfügbarkeitsheuristik verwechselt werden. Eine Wiedererkennung von Objekten und das Abrufen von verfügbaren Gedächtnisinhalten sind nämlich zwei verschiedene mentale Prozesse. Häufig bleibt das Wiedererkennungsvermögen weiterhin intakt, wenn das Erinnerungsgedächtnis beeinträchtigt wurde. So können ältere Menschen mit fortschreitender Demenz und Patienten mit bestimmten Gehirnschädigungen zwar gewisse Objekte wiedererkennen, haben aber oft gleichzeitig Schwierigkeiten damit, konkretes Wissen über die Objekte abzurufen, vgl. Gigerenzer 2007, S. 121; Warrington und McCarthy 1988, S. 184ff.

¹⁵⁵ Vgl. Standing 1973, S. 209.

werden, dass er in 60 % der Fälle die richtige Antwort weiß, wenn beide Alternativen vom Entscheidungsträger wiedererkannt werden – die Wissensvalidität beträgt hier 60 %. Schließlich, wenn keine von den beiden Alternativen wiedererkannt wird, beträgt die Treffergenauigkeit per Definition 50 %. Der höchste Wert für die richtigen Antworten stellt sich ein, wenn nicht alle, sondern genau die Hälfte aller Alternativenpaare wiedererkannt werden und bei der anderen Hälfte die Rekognitionsheuristik angewendet werden kann. In diesem Fall beträgt nämlich die Trefferquote gerundet 69 %.¹⁵⁶ Der Weniger-ist-mehr-Effekt drückt sich also dadurch aus, dass ein gewisser Grad an Unwissenheit in Kombination mit hoher Rekognitionsvalidität von Vorteil ist.¹⁵⁷

Der Weniger-ist-mehr-Effekt tritt in mindestens drei verschiedenen Formen auf.¹⁵⁸ Erstens kann er zwischen Gruppen von Personen beobachtet werden, wenn eine Gruppe mit weniger Wissen imstande ist, bessere Ergebnisse innerhalb eines Wissensbereichs zu generieren als eine andere Gruppe mit mehr Wissen.¹⁵⁹ In einer Studie entschieden sich bei der Frage, ob Detroit oder Milwaukee mehr Einwohner hat, rund 40 % der amerikanischen Studenten für die Stadt Milwaukee, wohingegen eine Gruppe von deutschen Studenten praktisch geschlossen die richtige Antwort Detroit gab.¹⁶⁰ Weil die meisten deutschen Studenten die Stadt Detroit, jedoch nicht Milwaukee wiedererkannten, konnten sie die Rekognitionsheuristik anwenden und die Frage richtig beantworten. Gleichzeitig kannten die amerikanischen Studenten beide Städte und konnten sich nicht eindeutig auf eine der beiden Antwortmöglichkeiten festlegen. Zweitens kann sich der Weniger-ist-mehr-Effekt auch zwischen Wissensbereichen einstellen, wenn dieselbe Personengruppe in einem Wissensbereich, in dem sie sich weniger auskennt, eine höhere Treffergenauigkeit erzielt als in einem anderen Bereich, in dem sie über mehr Wissen verfügt. Drittens schlägt sich der Weniger-ist-mehr-Effekt regelmäßig im Wissenserwerb nieder, wenn die individuelle Leistungskurve zunächst ansteigt und anschließend wieder fällt.

Beim FF-Ansatz werden je nach Umweltstrukturen die evolvierten Fähigkeiten wie Instrumente aus einem „adaptiven Werkzeugkasten“¹⁶¹ ausgewählt und zur Lösung von

¹⁵⁶ Die Berechnung dieses Werts erfolgte anhand der Formel aus Gigerenzer und Gaissmaier 2006, S. 344.

¹⁵⁷ Die Rekognitionsheuristik führt im Allgemeinen nur zum Weniger-ist-mehr-Effekt, wenn die Rekognitionsvalidität streng größer als die Wissensvalidität ist, vgl. Goldstein und Gigerenzer 2002, S. 79.

¹⁵⁸ Vgl. Gigerenzer und Gaissmaier 2006, S. 345f.

¹⁵⁹ Wissen bezieht sich im vorliegenden Kontext auf die Anzahl der wiedererkannten Objekte und nicht auf die Wissensvalidität.

¹⁶⁰ Vgl. Gigerenzer 2007, S. 15. Ein weiteres Beispiel in diesem Kontext stellt die Tatsache dar, dass die Prognosegenauigkeit von Laien bzw. Amateuren in Bezug auf Sportergebnisse oft auf dem gleichen Niveau wie bei den Experten oder sogar darüber liegt. Dies wurde u. a. anhand der Ergebnisse des Wimbledon-Turniers aus dem Jahr 2003 bestätigt, vgl. Serwe und Frings 2006, S. 324ff.

¹⁶¹ Vgl. Gigerenzer 2007, S. 71ff. Ausgewählte FF-Heuristiken mit einer kurzen Beschreibung ihrer Funktionsweise und der jeweils passenden Umweltstruktur finden sich u. a. bei Pfister et al. 2017, S. 356.

Problemen herangezogen. Dabei ist der Einsatz von FF-Heuristiken in komplexen Situationen mit hoher Unsicherheit von Vorteil und ökologisch rational, was in zahlreichen Experimenten und Simulationsuntersuchungen unter Beweis gestellt werden konnte.¹⁶² Demgegenüber erfordern Situationen mit kalkulierbaren Risiken statistisches Denken unter Verwendung aller verfügbaren Informationen, weil entweder eine weniger optimale oder sogar fehlerbehaftete Lösung – wie im Falle von HB-Heuristiken – zu erwarten ist.

2.3 Ausgewählte kognitive Spracheffekte

Sprachliche Merkmale üben gemäß den DST einen Einfluss auf den Denkmodus eines Entscheidungsträgers aus.¹⁶³ Folglich kann der durch die Sprachverarbeitung bedingte Effekt auf System 1 bzw. System 2 zur Entstehung von Biases und anderen verhaltensbestimmenden Phänomenen führen. Mit Lesbarkeit und Fremdsprache werden in diesem Kontext im Folgenden zwei wichtige linguistische Merkmale behandelt, die einerseits eine besondere Relevanz für die Controlling-Forschung und -Praxis haben und von denen andererseits bekannt ist, dass sie nachweislich mit einem starken Einfluss auf kognitive Denkprozesse einhergehen können.

2.3.1 Lesbarkeit

Die kognitiven Effekte der Lesbarkeit¹⁶⁴ wurden vor allem in der Psychologie und auf dem Gebiet der Lernforschung intensiv erforscht. Der Lesefluss in den Studien aus diesen Fachrichtungen wird insbesondere anhand der Schriftart, -farbe sowie -größe (z. B. *Ar Berkley* vs. *Arial*) manipuliert¹⁶⁵ und es werden die daraus resultierenden Effekte auf

¹⁶² Nicht hinreichend geklärt ist die Frage geblieben, ob und wann Personen die FF-Heuristiken tatsächlich anwenden, vgl. Hilbig 2010, S. 923ff. So zeigten die Untersuchungen von Bröder 2000 oder Newell et al. 2003, dass zum einen die TTB-Heuristik nicht immer konsistent verwendet wird und zum anderen das Verhalten von einigen Individuen nicht mit der TTB-Heuristik kompatibel ist.

¹⁶³ So wurde in Kapitel 2.2.1 bereits angemerkt, dass das Sprachverstehen in der eigenen Muttersprache eher im System-1-Denkmodus erfolgt.

¹⁶⁴ Eine einheitliche Definition des Begriffs Lesbarkeit existiert nicht. Im Allgemeinen wird unter der Lesbarkeit eines Textes die Leichtigkeit, mit der ein Leser den Text verstehen und verarbeiten kann, verstanden, vgl. Bonsall IV et al. 2017, S. 330. Eine ähnliche Definition gilt für den Lesefluss, vgl. Oppenheimer 2008, S. 238. Aus diesem Grund werden im Rahmen dieser Arbeit die Begriffe Lesbarkeit und Lesefluss synonym verwendet.

¹⁶⁵ In einigen Studien wurde der Lesefluss aber auch mittels sich reimender, komplexer oder sogar rotierter Wörter manipuliert, siehe McGlone und Tofighbakhsh 2000; Alter und Oppenheimer 2006; Oppenheimer 2006, Sungkhasettee et al. 2011.

die kognitiven Leistungen bzw. Lernperformance beobachtet.¹⁶⁶ Empirisch konnten dabei zwei verschiedene Phänomene festgestellt werden: Disfluency-Effekt und Cognitive-Load-Effekt. Beide Phänomene werden im Folgenden näher betrachtet.

Der *Disfluency-Effekt* (DFE) folgt dem Konzept der gewünschten Schwierigkeit¹⁶⁷ und bezeichnet die positive Wirkung der reduzierten Lesbarkeit. Alter et al. 2007 gelang zuerst der empirische Nachweis, dass ein erschwerter Lesefluss zum besseren Abschneiden im Cognitive-Reflection-Test und einem verbesserten syllogistischen Denken führen kann.¹⁶⁸ Diese Effekte begründen die Autoren mithilfe der DST. Demnach löst ein gehemmter Lesefluss ein metakognitives Gefühl von Verarbeitungsschwierigkeiten aus, was wiederum die Aktivierung des analytischen Systems 2 bedingt. Gleichzeitig sind flüssig zu verarbeitenden Informationen i. d. R. mit keinem subjektiven Gefühl der mentalen Anstrengung verbunden und der Denkprozess erfolgt überwiegend im automatischen System-1-Denkmodus. So reguliert die Lesbarkeit nicht nur die Zeit, die für die Verarbeitung von Informationen erforderlich ist, sondern darüber hinaus auch die generelle Herangehensweise an die Problemlösung.¹⁶⁹

Der positive Einfluss reduzierter Lesbarkeit auf das logische Denken wurde ferner von Song und Schwarz 2008 im Zusammenhang mit der Moses-Illusion bestätigt.¹⁷⁰ Die-
mand-Yauman et al. 2011 konnten den DFE nicht nur unter kontrollierten Laborbedingungen, sondern auch in einem echten Lernkontext einer Schulklasse reproduzieren und French et al. 2013 zeigten, dass auch Schüler mit Dyslexie vom DFE profitieren können. Im religiösen Kontext gelang Gervais und Norenzayan 2012 (Experiment 5) der Nachweis, dass Menschen allein infolge einer reduzierten Lesbarkeitsstufe eine skeptischere Haltung in Glaubensfragen einnehmen können. Die Autoren führen diesen Befund auf die Aktivierung des kritischen Denkens zurück. Ein weiteres Indiz dafür, dass die Lesbarkeit einen Einfluss auf System 2 ausübt, liefern Lehmann et al. 2016 mit ihrer Untersuchung der Rolle des Arbeitsgedächtnisses. Der DFE war nämlich insbesondere bei denjenigen Probanden am stärksten ausgeprägt, die über höhere Arbeitsgedächtniskapazitäten verfügten.¹⁷¹ Weil das Denken im System-2-Modus die eingeschränkten Res-

¹⁶⁶ Einen kurzen Überblick über relevante Studien findet sich u. a. bei Lehmann et al. 2016; Seufert et al. 2017.

¹⁶⁷ Vgl. Bjork 1994, S. 189ff. Neben der Lesbarkeit können zur Erzeugung der gewünschten Schwierigkeit auch andere Mittel, wie Bilder oder gesprochene Sprache, eingesetzt werden, vgl. Eitel et al. 2014, S. 490; Kühl et al. 2014, S. 192.

¹⁶⁸ Vgl. Alter et al. 2007, S. 569ff.

¹⁶⁹ Vgl. Seufert et al. 2017, S. 222.

¹⁷⁰ Die Frage bei der Moses-Illusion lautet sinngemäß wie folgt: „Wie viele Tiere von jeder Art nahm Moses mit auf die Arche?“. Die intuitive und schnelle Antwortet „zwei“ ist falsch, weil gemäß der Bibel Noah und nicht Moses die Arche baute.

¹⁷¹ Vgl. Lehmann et al. 2016, S. 89ff.

sources des Arbeitsgedächtnisses beansprucht, kann reduzierte Lesbarkeit offensichtlich nur dann ihre positive Wirkung entfalten, wenn dafür ausreichend Ressourcen des Arbeitsgedächtnisses vorhanden sind.

Andere Studien zum DFE konnten entweder keinen oder sogar einen negativen Effekt einer niedrigen Lesbarkeitsstufe auf die kognitiven Leistungen bzw. Lernperformance messen. So konnten Meyer et al. 2015 in einer Metastudie keinen DFE im Rahmen des Cognitive-Reflection-Tests feststellen. Rhodes und Castel 2008 manipulierten den Lesfluss anhand der Schriftgröße und fanden keine Unterschiede in der Anzahl der erinnerten Wörter zwischen den Teilnehmergruppen. Ein neutrales Ergebnis im Hinblick auf die generelle Lernperformance kam auch in der Untersuchung von Eitel et al. 2014 (außer Experiment 1) heraus. In der Studie von Yue et al. 2013 wiesen die Teilnehmer in der Gruppe mit einer niedrigen Lesbarkeitsstufe zum Teil sogar eine geringere Erinnerungsleistung auf als die anderen Teilnehmer. Diese Ergebnisse ließen bei einigen Forschern Zweifel an der generellen Generalisierbarkeit des DFE aufkommen.¹⁷² Solch divergierende Ergebnisse lassen sich aber teilweise auf die individuellen Unterschiede der Probanden (z. B. Arbeitsgedächtniskapazitäten) und die unterschiedlichen Manipulationsmethoden der Lesbarkeit in den einzelnen Studien zurückführen. Seufert et al. 2017 kontrollierten in ihrer Studie die individuellen Arbeitsgedächtniskapazitäten der Probanden und untersuchten die Wirkung von drei unterschiedlichen Lesbarkeitsstufen, die nach einer einheitlichen Methode generiert wurden. Im Ergebnis zeigte sich, dass ein gewisser Grad an reduzierter Lesbarkeit tatsächlich förderlich für die kognitiven (Lern-)Leistungen ist. Wird dieser Grad jedoch überschritten, verschlechtert sich die Performance.¹⁷³

Der *Cognitive-Load-Effekt* (CLE) steht für die negativen kognitiven Effekte einer zu niedrigen Lesbarkeitsstufe. Gemäß der Cognitive-Load-Theorie¹⁷⁴ existieren im Lernkontext drei Typen der kognitiven Belastung:¹⁷⁵

- **Intrinsische kognitive Belastung:** Diese Art der kognitiven Belastung wird von der Komplexität der zugrunde liegenden Aufgabe selbst ausgelöst. Je schwieriger der betrachtete Sachverhalt ist, desto höher ist die damit verbundene kognitive Anstrengung. In diesem Zusammenhang spielt die Elementinteraktivität eine zentrale Rolle. Bei einer niedrigen Elementinteraktivität sind die relevanten Informationselemente unabhängig voneinander und müssen folglich nicht gleichzeitig im Arbeitsgedächtnis behalten werden. Bei einer hohen Elementinteraktivität sind die relevanten In-

¹⁷² Vgl. Eitel et al. 2014, S. 488.

¹⁷³ Vgl. Seufert et al. 2017, S. 221.

¹⁷⁴ Die Cognitive-Load-Theorie geht auf Sweller 1994; Sweller et al. 1998 zurück. Im Lehrbuchformat kann diese Theorie bei Sweller et al. 2011 nachgeschlagen werden.

¹⁷⁵ Vgl. im Folgenden Sweller et al. 1998, S. 259ff.

formationselemente über Wechselwirkungen höherer Ordnung miteinander verknüpft. Die Verarbeitung von solchen Informationen erfordert große kognitive Anstrengung, weil viele Informationselemente simultan verarbeitet werden müssen und die limitierten Ressourcen des Arbeitsgedächtnisses dadurch stark beansprucht werden. So ist bspw. das reine Auswendiglernen von Vokabeln in einer Fremdsprache durch einen niedrigen Grad der Elementinteraktivität gekennzeichnet. Sollen darüber hinaus Sätze mit den gelernten Vokabeln gebildet werden, steigt die Elementinteraktivität deutlich an, weil nun die einzelnen Wörter in einen übergeordneten Zusammenhang gebracht und in Übereinstimmung mit den Grammatikregeln kombiniert werden müssen. Die intrinsische kognitive Belastung kann nicht beeinflusst werden, ohne dabei gleichzeitig die eigentliche Aufgabe zu ändern. Eine Person, die über die erforderliche Expertise zur Lösung einer Aufgabe verfügt, ist jedoch diesem Typ der kognitiven Belastung weniger stark ausgesetzt.

- Extrinsische kognitive Belastung: Diese Art der kognitiven Belastung hängt unmittelbar mit der formalen Gestaltung der zugrunde liegenden Aufgabe zusammen. Eine schwieriger zugängliche Information löst demnach eine höhere kognitive Belastung aus. Im Gegensatz zur intrinsischen kann die extrinsische kognitive Belastung also ausschließlich über die äußeren Merkmale wie Gestaltung und Formatierung des Lernmaterials beeinflusst werden. Diese Quelle der kognitiven Belastung sollte reduziert werden, weil sie im Lernprozess ablenkend wirkt.
- Lernbezogene kognitive Belastung (engl. Germane Cognitive Load): Schließlich stellt die lernbezogene kognitive Belastung den individuellen auf die Bearbeitung der eigentlichen Aufgabe ausgerichteten Aufwand dar. Ein wichtiger Aspekt der lernbezogenen kognitiven Belastung besteht in der Aktivierung bestehender bzw. dem Aufbau neuer Problemlösungsansätze. Diese Art der kognitiven Belastung ist erwünscht und sollte z. B. mittels motivationsfördernder Maßnahmen gesteigert werden.

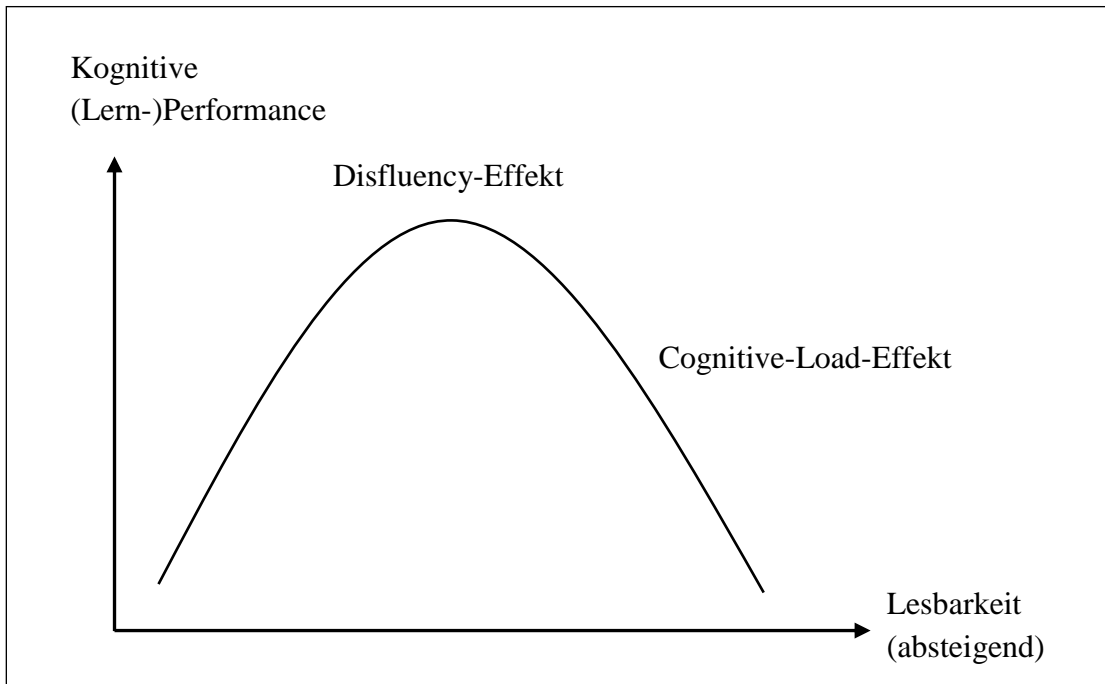
Die drei oben beschriebenen Typen der kognitiven Belastung sind additiv und stellen zusammengenommen den Gesamtaufwand für das Arbeitsgedächtnis dar. Da die intrinsische kognitive Belastung nicht unabhängig von der eigentlichen Aufgabe gesteuert werden kann, besteht gemäß der Cognitive-Load-Theorie das Ziel darin, bei gegebener intrinsischer Belastung die Lernmaterialien so zu gestalten, dass die extrinsische Belastung minimiert und die lernbezogene kognitive Belastung angeregt werden.¹⁷⁶

Gemäß der Cognitive-Load-Theorie gehört die Lesbarkeit zur Kategorie der extrinsischen kognitiven Belastung. Demnach nimmt die extrinsische Belastung zu, wenn die Lesbarkeit reduziert wird.¹⁷⁷ Eine streng reziproke Beziehung zwischen den beiden Größen konnte allerdings empirisch nicht nachgewiesen werden. In der Untersuchung

¹⁷⁶ Vgl. Lehmann et al. 2016, S. 91.

¹⁷⁷ Vgl. Eitel et al. 2014, S. 489.

Abbildung 8: Grad der Lesbarkeit und kognitive (Lern-)Performance



Quelle: In Anlehnung an Diemand-Yauman et al. 2011, S. 114; Seufert et al. 2017, S. 227.

von Seufert et al. 2017 machte sich die wahrgenommene extrinsische Belastung erst in der niedrigsten Lesbarkeitsstufe bemerkbar. Zugleich war die Lernperformance von Teilnehmern in der mittleren Lesbarkeitsstufe am größten.¹⁷⁸ Daraus kann geschlossen werden, dass der negative Effekt der niedrigen Lesbarkeit in Form des CLE erst ab einem gewissen Niveau niedriger Lesbarkeit auftritt.

Basierend auf den empirischen Beobachtungen kann also insgesamt ein umgekehrt verlaufender U-förmiger Zusammenhang zwischen der Lesbarkeit und der kognitiven (Lern-)Performance unterstellt werden (siehe Abbildung 8). Auf einem hohen Lesbarkeitsniveau können die Informationen flüssig verarbeitet werden, was standardmäßig im automatischen System-1-Denkmodus erfolgt. Wird die Lesbarkeit bis zu einem bestimmten Grad reduziert, löst der gehemmte Lesefluss ein Gefühl der metakognitiven Verarbeitungsschwierigkeit aus und System 2 wird aktiviert. Dabei ist wichtig, dass die extrinsische kognitive Belastung nicht alle verfügbaren Ressourcen des Arbeitsgedächtnisses okkupiert und noch genügend mentale Kapazitäten verbleiben, um die zugrundeliegende Aufgabe im analytischen System-2-Denkmodus zu bearbeiten. Wenn der Grad der reduzierten Lesbarkeit zu groß wird, überwiegt die extrinsische kognitive Belastung und es kommt zum CLE, d. h., eine Person muss die kompletten mentalen Ressourcen

¹⁷⁸ Vgl. Seufert et al. 2017, S. 231ff.

für die nebensächlichen Aktivitäten wie Textdekodierung, -navigation oder -suche aufwenden und kann sich nicht ausreichend der eigentlichen Aufgabe widmen.¹⁷⁹

2.3.2 Fremdsprache

Falls eine Information richtig verstanden wurde, sollte es für den Entscheidungsfindungsprozess irrelevant sein, ob diese Information in der eigenen Mutter- oder einer Fremdsprache zur Verfügung gestellt wurde. Tatsächlich zeigten die neueren psychologischen Untersuchungen, dass der sprachliche Kontext den Entscheidungsfindungsprozess beeinflussen kann.¹⁸⁰ Nachfolgend werden die möglichen negativen und positiven kognitiven Wirkungen einer Fremdsprache näher beleuchtet.

2.3.2.1 Brain-Drain-Modell

Das *Brain-Drain-Modell* (BDM) von Volk et al. 2014 bedient sich den DST und unterstreicht die negativen Wirkungen einer Fremdsprache auf die kognitiven Prozesse (siehe Abbildung 9).¹⁸¹ Gemäß dem BDM erfolgt die Informationsverarbeitung in der eigenen Muttersprache automatisch und mühelos.¹⁸² Im Gegensatz dazu erfordert die Verarbeitung von Informationen in einer Fremdsprache mehr an kognitiven Ressourcen und tangiert somit im höheren Maße das Arbeitsgedächtnis.¹⁸³ Ein Grund hierfür liegt in der Notwendigkeit einer kontrollierten Unterdrückung der eigenen Muttersprache, wenn eine Fremdsprache benutzt wird.¹⁸⁴ So muss bspw. ein deutschsprachiger Muttersprachler, der das Wort Apfel auf Englisch sagen möchte, aktiv darauf achten, dass er anstelle von „Apple“ nicht aus Versehen „Apfel“ sagt.¹⁸⁵ Der Einsatz einer fremden

¹⁷⁹ Auf die Gestaltungsprinzipien und kognitive Verhaltenswirkungen der Lesbarkeit speziell im betriebswirtschaftlichen Kontext wird noch ausführlicher in Kapitel 2.4 eingegangen.

¹⁸⁰ Vgl. Costa et al. 2017, S. 146; Hayakawa et al. 2016, S. 791.

¹⁸¹ Kognitive Prozesse in einer Fremdsprache betreffen neben der gesprochenen Sprache auch andere Formen der Informationsverarbeitung, wie das Hören, Lesen und Schreiben. Des Weiteren kann die Sprache sowohl nach außen (z. B. Gespräche mit und Textnachrichten an anderen Personen) als auch nach innen (z. B. Selbstgespräche, Notizen für den eigenen Gebrauch) gerichtet sein, vgl. Schrauf 2009, S. 5ff. Im Rahmen dieser Arbeit bildet die schriftliche Übermittlung von Informationen in Form eines Controllingberichts den Gegenstand der Untersuchung.

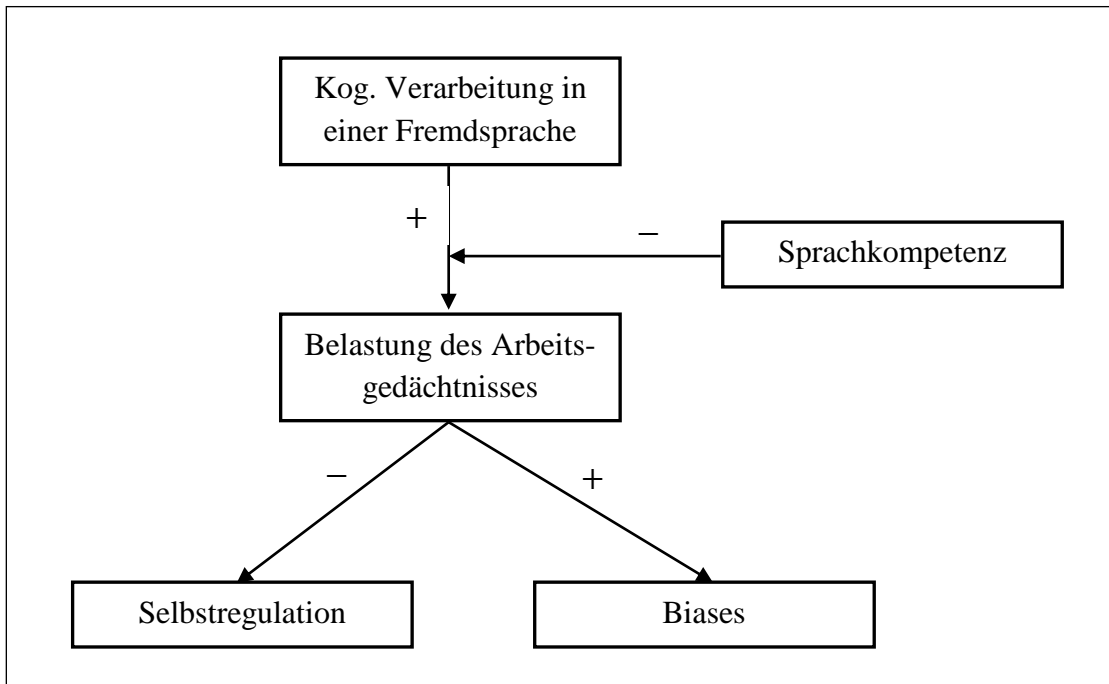
¹⁸² Das muttersprachliche Sprachniveau kann im Allgemeinen nur dann erreicht werden, wenn eine Sprache innerhalb eines gewissen jungen Alters – auch „kritische Periode“ genannt – erlernt wird. In der Literatur reichen die Angaben für dieses kritische Alter von sieben Jahren bis zur Pubertät, vgl. Costa und Sebastián-Gallés 2014, S. 340.

¹⁸³ Vgl. Abutalebi 2008, S. 470.

¹⁸⁴ Vgl. Abutalebi und Green 2016, S. 689.

¹⁸⁵ Vgl. Costa und Sebastián-Gallés 2014, S. 339f.

Abbildung 9: Brain-Drain-Modell



Quelle: In Anlehnung an Volk et al. 2014, S. 865.

Sprache führt also gemäß dem BDM zu höherer kognitiver Belastung und ist daher hinderlich für die Bearbeitung eines Problems im rationalen System-2-Denkmodus. Dies erhöht wiederum die Wahrscheinlichkeit dafür, dass Biases und Selbstregulationsprobleme auftreten.¹⁸⁶ Im Folgenden werden die beiden Bereiche näher beleuchtet.¹⁸⁷

- **Selbstregulation:** In der Psychologie sind unter dem Begriff Selbstregulation diejenigen Fähigkeiten gemeint, mit denen ein Individuum seine Aufmerksamkeit, Emotionen und Handlungen steuern kann. Die Forschung auf diesem Gebiet weist auf die besondere Rolle der Zielsetzung und Abweichungskontrolle im Rahmen eines effektiven Selbstregulationsprozesses hin.¹⁸⁸ Der Selbstregulationsprozess fängt mit der Bestimmung von klar definierten Zielen (z. B. mehr Geduld im Umgang mit Kindern) an. Diese Ziele müssen stets im Bewusstsein präsent sein, damit bei einer Abweichung vom gewünschten Zielzustand (z. B. einer zu lauten Stimme) korrigierend eingegriffen werden kann. Für einen erfolgreichen Selbstregulationsprozess sind also eine aktive Repräsentation und bewusste Zustandskontrolle von selbstgesetzten Zielen im Arbeitsgedächtnis erforderlich. Folglich wird der Selbstregulati-

¹⁸⁶ Vgl. Hadjichristidis et al. 2017b, S. 646.

¹⁸⁷ Vgl. im Folgenden Volk et al. 2014, S. 868ff.

¹⁸⁸ Vgl. Karoly 1993, S. 31ff.

onsprozess gestört, wenn die Kapazitäten des Arbeitsgedächtnisses von einer anderen kognitiv anspruchsvollen Tätigkeit beansprucht werden.¹⁸⁹ Das BDM postuliert, dass die Informationsverarbeitung in einer Fremdsprache so viele Ressourcen des Arbeitsgedächtnisses in Anspruch nimmt, dass die gesetzten Ziele im Rahmen eines Selbstregulationsprozesses unter Umständen nicht mehr bewusst aufrechterhalten bzw. kontrolliert werden können. Die gehemmte Fähigkeit zur emotionalen Regulation kann in Extremfällen sogar zu Wutausbrüchen und körperlicher Gewalt führen.¹⁹⁰ Auch die Fähigkeit der Aufmerksamkeitsregulation wird sehr wahrscheinlich unter dem Einsatz einer fremden Sprache leiden, weil Personen von diversen Störquellen leichter von der eigentlichen Aufgabe abgelenkt werden und damit den Fokus der Aufmerksamkeit verlieren. Schließlich kann sich ein gestörter Selbstregulationsprozess in einer geminderten Fähigkeit zur Handlungsregulation – d. h., zur Befolgung von sozial akzeptierten und erwünschten Verhaltensnormen – niederschlagen.

- Biases: Das logische Denken erfordert kognitive Ressourcen, die jedoch nur eingeschränkt zur Verfügung stehen, falls die Informationsverarbeitung in einer fremden Sprache erfolgt. Wenn logische Schlussfolgerungen nicht gezogen werden oder Menschen in einem Konflikt zu den eigenen Überzeugungen stehen, kann es zum Phänomen der Überzeugungsverzerrung (engl. Belief Bias) kommen.¹⁹¹ Die Wirkung der Überzeugungsverzerrung verstärkt sich insbesondere dann, wenn das logische Denken im Widerspruch zu bestehenden Dogmen steht und eine Überwindung der daraus resultierenden inneren Konflikte erforderlich ist. Ein Entscheider, der in einer Fremdsprache operiert, wird gemäß dem BDM weniger bereit sein, in radikal neue Technologien zu investieren. Des Weiteren können probabilistische Urteile wegen des Einsatzes einer Fremdsprache gestört werden, weil die beiden kognitiven Prozesse um die limitierten Ressourcen des Arbeitsgedächtnisses konkurrieren. An dieser Stelle wird auch auf die bereits im Zusammenhang mit der Repräsentativitätsheuristik (Kapitel 2.2.2.2) und der Verfügbarkeitsheuristik (Kapitel 2.2.2.4) erläuterten Biases verwiesen, die sich laut dem BDM bei der Informationsverarbeitung in einer fremden Sprache leichter manifestieren können. Schließlich argumentieren Volk et al. 2014 im Kontext der Biases, dass sich beim Einsatz einer fremden Sprache die Genauigkeit der sozialen Urteile reduzieren wird. Beispielhaft wird in diesem Zusammenhang das Denken in Stereotypen genannt.

Im Rahmen des BDM spielt die Sprachkompetenz eine moderierende Rolle. Bei einem hohen Kompetenzniveau können die Informationen ähnlich wie in der eigenen Muttersprache schnell und automatisch verarbeitet werden, was eine geringe Belastung des

¹⁸⁹ Empirische Evidenz für diese Hypothese liefern u. a. Hofmann et al. 2010, S. 212f.

¹⁹⁰ Vgl. Heatherton 2011, S. 374f.

¹⁹¹ Siehe Evans et al. 1983.

Arbeitsgedächtnisses bedeutet. Die Sprachkompetenz ist also mit der kognitiven Belastung negativ korreliert.¹⁹²

2.3.2.2 Foreign-Language-Effekt

Als Antwort auf das BDM plädieren Hadjichristidis et al. 2017b – gestützt auf empirische Befunde aus zahlreichen psychologischen Studien – für eine Erweiterung des Modells um mögliche positive kognitive Effekte einer Fremdsprache, die unter dem Sammelbegriff *Foreign-Language-Effekt* (FLE) subsumiert werden. Im weiteren Verlauf dieses Abschnitts wird zunächst ein Überblick über die Wirkungsbereiche, in denen der FLE festgestellt wurde, gegeben. Danach werden die in der Literatur am meisten diskutierten Erklärungsansätze für den FLE vorgestellt und die noch offenen Fragen thematisiert.

2.3.2.2.1 Wirkungsbereiche

Die beiden wichtigsten Wirkungsbereiche des FLE betreffen das Risikoverhalten und ethisches Verhalten sowie moralische Urteile.¹⁹³ Andere Wirkungsbereiche des FLE werden gesammelt in einer dritten Kategorie behandelt.

Viele empirische Untersuchungen demonstrierten, dass die Darstellung von Informationen in einer Fremdsprache das individuelle *Risikoverhalten* beeinflussen kann. Demnach sind die Individuen in einer Fremdsprache eher bereit, erwünschte bzw. „smarte“ Risiken¹⁹⁴ – d. h., Lotterien mit einem positiven Erwartungswert – einzugehen. Die meisten Individuen würden einen sicheren Gewinn von 1 € gegenüber einer Lotterie [2,50 € mit 50 % oder 0 € mit 50 %] bevorzugen. Aus normativer Sicht sollte die risikobehaftete Option gewählt werden, weil der entsprechende Erwartungswert über der sicheren Auszahlung liegt (1 € vs. 1,25 €). Keysar et al. 2012 untersuchten experimentell eine solche Situation zum einen mit fiktiven Geldeinsätzen (Experiment 2) und zum anderen mit einer echten Vergütung der Teilnehmer (Experiment 3) und fanden heraus, dass der Einsatz einer Fremdsprache die Bereitschaft der Teilnehmer, vorteilhafte Risiken einzugehen, signifikant erhöhte. Ein ähnliches Ergebnis konnte in der Studie von

¹⁹² Vgl. Service et al. 2002, S. 383ff.

¹⁹³ Vgl. Costa et al. 2017, S. 147; Hayakawa et al. 2016, S. 792. Für einen Überblick über die Wirkungsbereiche des FLE siehe u. a. Hadjichristidis et al. 2019, S. 258ff.

¹⁹⁴ Vgl. Hadjichristidis et al. 2017b, S. 647.

Costa et al. 2014a (Experiment 3a) im Zusammenhang mit dem Holt-Laury-Test festgestellt werden.¹⁹⁵ Die Teilnehmer dieser Studie orientierten sich in der fremdsprachlichen Gruppe stärker am Erwartungswert und zeigten darüber hinaus ein konsistenteres Präferenzverhalten.

In der Untersuchung von Hadjichristidis et al. 2015 wurden außerdem Anhaltspunkte dafür gegeben, dass eine Fremdsprache nicht nur das tatsächlich beobachtbare Risikoverhalten, sondern auch die generelle Risikowahrnehmung beeinflussen kann. In der genannten Studie wurden die Probanden gebeten, Sachverhalte, wie Atomkraft, Flugreisen oder Nanotechnologien, bezüglich ihrer Risiken und Chancen zu bewerten. Es stellte sich heraus, dass die wahrgenommenen Chancen überwogen, wenn die Probanden ihre Einschätzungen in einer fremden Sprache vornehmen mussten.¹⁹⁶

Des Weiteren deuten einige Studien darauf hin, dass Individuen bei der Informationsverarbeitung in einer fremden Sprache weniger für den Einrahmungseffekt (engl. Framing) anfällig sind.¹⁹⁷ In den entsprechenden Studien wird oft das berühmte Asian-Disease-Problem oder eine Abwandlung davon verwendet. Das klassische Asian-Disease-Problem lautet sinngemäß wie folgt:¹⁹⁸

Vor kurzem ist eine neue gefährliche Krankheit ausgebrochen. Ohne medizinische Behandlung werden infolge dieser Krankheit 600 000 Menschen sterben. Die nachfolgenden beiden Medikamente stehen Ihnen als Behandlungsalternativen zur Auswahl:

Gewinnframing:

Medikament A: 200 000 Menschen werden gerettet.

Medikament B: Mit einer Wahrscheinlichkeit von $1/3$ werden 600 000 Menschen gerettet und mit einer Wahrscheinlichkeit von $2/3$ wird keiner gerettet.

¹⁹⁵ Siehe Holt und Laury 2002. Dieser Test wurde entwickelt, um die individuelle Risikoeinstellung zu messen. Dabei werden den Probanden wiederholt zwei Lotterien A und B zur Auswahl gestellt. Die absoluten Auszahlungsbeträge der beiden Lotterien bleiben konstant und es werden lediglich die Eintrittswahrscheinlichkeiten variiert, sodass sich die Erwartungswerte verändern. Zu Beginn hat die Lotterie A einen höheren Erwartungswert im Vergleich zu B. Ab der fünften Runde verändert sich das Verhältnis der Erwartungswerte zugunsten der Lotterie B. Eine risikoneutrale Person sollte also bis einschließlich Runde vier die Lotterie A präferieren und anschließend die Lotterie B.

¹⁹⁶ Vgl. Hayakawa et al. 2016, S. 792.

¹⁹⁷ Der Einrahmungseffekt geht auf Tversky und Kahneman 1981 zurück und bedeutet, dass unterschiedliche Formulierungen ein und desselben Sachverhalts – z. B. als Gewinn oder Verlust – zu systematischen Unterschieden im Entscheidungsverhalten führen können. Dieses Phänomen kann mithilfe der PT erklärt werden. Wird ein Problem in einen Verlustkontext eingebettet, verhalten sich die Menschen tendenziell risikosuchend. Umgekehrt werden die Risiken eher gemieden, wenn ein Problem aus einer Gewinnperspektive dargestellt wird (siehe Kapitel 2.1.2).

¹⁹⁸ Vgl. im Folgenden Tversky und Kahneman 1981, S. 453.

Verlustframing:

Medikament A: 400 000 Menschen werden sterben.

Medikament B: Mit einer Wahrscheinlichkeit von 1/3 wird keiner sterben und mit einer Wahrscheinlichkeit von 2/3 werden 600 000 Menschen sterben.

Für welches Medikament würden Sie sich entscheiden?

Die Teilnehmer in den Studien von Keysar et al. 2012 (Experiment 1) und Costa et al. 2014a (Experiment 1) wiesen in ihrer Muttersprache ein gemäß dem Einrahmungseffekt zu erwartendes Antwortverhalten auf. In den Gruppen mit einer Fremdsprache waren dagegen keine signifikanten Unterschiede im Antwortverhalten der Teilnehmer festzustellen.¹⁹⁹

Ein weiterer Wirkungsbereich, in dem der FLE festgestellt werden konnte, betrifft *ethisches Verhalten* und *moralische Urteile*.²⁰⁰ Empirisch konnte vielfach nachgewiesen werden, dass Informationsverarbeitung in einer Fremdsprache zu stärker utilitaristisch ausgeprägten Verhalten und milderem moralischen Urteilen führt als in der eigenen Muttersprache.

In experimentellen Studien zum ethischen Verhalten wird häufig das Trolley-Problem²⁰¹ oder eine Abwandlung davon verwendet. Mithilfe des Trolley-Problems wird untersucht, ob sich Personen in einer ethischen Dilemmasituation eher an der utilitaristischen oder deontologischen Maxime orientieren.²⁰² Die „Footbridge“-Version des Trolley-Problems eignet sich besonders gut, um einen starken moralischen Konflikt zu erzeugen.²⁰³ Diese Version des Trolley-Problems lautet sinngemäß wie folgt:²⁰⁴

Ein Schienenfahrzeug hat sich auf einer Anhöhe gelöst und rast auf fünf Gleisarbeiter zu. Wenn nichts unternommen wird, werden alle fünf Gleisarbeiter umkommen. Sie befinden sich auf einer Brücke über dem ganzen Geschehen. Die einzige Möglichkeit die fünf Gleisarbeiter zu retten, besteht darin, eine fremde Person neben Ihnen von der Brücke zu stoßen, um den Wagen zu stoppen. Wenn Sie das machen, wird die gestoßene Person umkommen und die fünf Arbeiter werden unversehrt bleiben.

¹⁹⁹ Im Framingkontext ist der FLE jedoch nicht unumstritten. So konnten Oganian et al. 2016 keine nivellierende Wirkung einer Fremdsprache feststellen. Stattdessen führen die Autoren die festgestellten Effekte in den anderen Studien zum FLE darauf zurück, dass die Teilnehmer zwischen den Sprachen wechseln mussten, vgl. Oganian et al. 2016, S. 140ff.

²⁰⁰ Vgl. Corey et al. 2017, S. 1123.

²⁰¹ Vgl. Thomson 1985, S. 1395ff.

²⁰² Gemäß dem utilitaristischen Ansatz steht die Maximierung des Gesamtnutzens an oberster Stelle. In der deontologischen Ethik zählen nicht nur die Konsequenzen einer Handlung, sondern es kommt auch auf die Einhaltung von gewissen moralischen Ge- bzw. Verboten an.

²⁰³ Vgl. Costa et al. 2014b, S. 2.

²⁰⁴ Vgl. im Folgenden Cipolletti et al. 2016, S. 28.

Wäre es für Sie moralisch vertretbar, die Person neben Ihnen von der Brücke zu stoßen?

Aus utilitaristischer Sicht wäre es vertretbar, ein Menschenleben zu opfern, um fünf andere zu retten, weil dadurch das Gemeinwohl maximiert wird. Auf der anderen Seite steht gemäß dem deontologischen Ansatz das universelle Recht auf Leben und körperliche Unversehrtheit an höchster Stelle. In der eigenen Muttersprache entscheidet sich die überwiegende Mehrheit der Befragten in der vorgestellten Situation deontologisch. Wird aber der Text in einer Fremdsprache präsentiert, steigt die Bereitschaft zum utilitaristischen Verhalten deutlich an.²⁰⁵

Weitere Hinweise auf eine erhöhte Neigung zum konsequentialistischen Verhalten infolge des Einsatzes einer fremden Sprache liefert die experimentelle Untersuchung von Geipel et al. 2016. Die Teilnehmer im ersten Experiment bekamen entweder in ihrer Muttersprache oder einer Fremdsprache Szenarien vorgelegt, in denen fraglich motivierte Handlungen mit positiven Konsequenzen beschrieben wurden.²⁰⁶ Im zweiten Experiment verhielt es sich genau umgekehrt, die Handlungen hatten eine gute Intention aber einen schlechten Ausgang.²⁰⁷ Die Aufgabe der Probanden bestand darin, die beschriebenen Situationen im Hinblick auf die moralische Richtigkeit zu bewerten. Es zeigte sich, dass die Zustimmungswerte in den fremdsprachlichen Gruppen im ersten Experiment höher und im zweiten Experiment niedriger waren. Die Teilnehmer maßen also den Konsequenzen ein höheres Gewicht bei, wenn sie in einer fremden Sprache moralische Urteile fällen mussten. Aber auch generell fallen die moralischen Urteile milder aus, wenn eine Fremdsprache benutzt wird.²⁰⁸

Klesse et al. 2015 konnten einen positiven Einfluss einer Fremdsprache auf die *Selbstregulation* empirisch nachweisen. Die Autoren baten im Rahmen ihrer Untersuchung die Gäste in einem Restaurant, eine (mündliche) Bestellung eines Nachtisches entweder in der eigenen Muttersprache oder einer Fremdsprache zu tätigen. Der Einsatz einer Fremdsprache führte insgesamt zur gesünderen Dessertwahl.²⁰⁹ Offenbar konnten die Probanden in einer Fremdsprache der Versuchung eines leckeren aber ungesünderen Desserts besser widerstehen.

²⁰⁵ In den meisten Studien erhöhte sich die Bereitschaft zum utilitaristischen Verhalten um mehr als das Doppelte, vgl. Cipolletti et al. 2016, S. 31f.; Costa et al. 2014b, S. 4f.; Geipel et al. 2015b, S. 4ff.

²⁰⁶ Zum Beispiel: Ein Unternehmen spendet Geld an eine Hilfsorganisation, um dadurch den eigenen Bekanntheitsgrad zu steigern und höhere Umsätze zu generieren.

²⁰⁷ Zum Beispiel: Eine Person gibt einem armen Jungen etwas Geld, damit er sich Essen kaufen kann. Stattdessen verwendet der Junge das Geld, um sich Drogen zu kaufen, und stirbt später an einer Überdosis.

²⁰⁸ Siehe u. a. Geipel et al. 2015a. In dieser Studie wurden Fallstudien mit Verstößen gegen soziale Normen, wie geschütztem Sex zwischen Geschwistern, verwendet. Insgesamt verurteilten die Teilnehmer die Verstöße gegen die sozialen Normen weniger stark, wenn sie die Fallstudien in einer Fremdsprache erhielten.

²⁰⁹ Vgl. Klesse et al. 2015, S. 542f.

Außerdem neigen Menschen in einer Fremdsprache weniger stark dazu, *betrügerischen Handlungen* nachzugehen. In der Studie von Bereby-Meyer et al. 2018 wurden die Teilnehmer gebeten, unbeaufsichtigt einen Würfel zu werfen und das Ergebnis entweder in der eigenen Muttersprache oder einer Fremdsprache zu melden. Weil die Experimentvergütung von der Höhe der gemeldeten Zahl abhing, wurde den Probanden ein Anreiz geschaffen, mit überhöhten Angaben zu betrügen. Im Ergebnis zeigte sich, dass die durchschnittlichen Meldungen der fremdsprachlichen Gruppe weniger stark vom wahr-scheinlichkeitstheoretischen Erwartungswert abwichen und somit weniger überhöht waren.²¹⁰

Schließlich konnte empirisch nachgewiesen werden, dass Individuen in einer Fremd-sprache weniger anfällig für *abergläubisches Verhalten* und *illusorische Korrelationen* sind.²¹¹

2.3.2.2.2 Erklärungsansätze und Rolle der Sprachkompetenz

Die genaue Wirkungsweise des FLE ist noch nicht hinreichend geklärt. In der Literatur etablierten sich aber die folgenden drei Erklärungsansätze: Reduktion der emotionalen Resonanz, psychologische Distanz und Disfluency-Effekt (DFE).²¹² Bei allen diesen Erklärungsansätzen wird in Anlehnung an die DST davon ausgegangen, dass die Ver-arbeitung von Informationen in einer Fremdsprache entweder System 1 und/oder Sys-tem 2 beeinflusst. Im weiteren Verlauf dieses Abschnitts werden die drei Erklärungs-ansätze einzeln vorgestellt und anschließend wird der Einfluss der Sprachkompetenz im Rahmen des FLE diskutiert.

Der führende Erklärungsansatz für den FLE basiert auf der *Reduktion der emotionalen Resonanz*. Demnach induzieren affektive Stimuli weniger starke Emotionen in einer Fremdsprache im Vergleich zur eigenen Muttersprache.²¹³ Über reduzierte elektrodermale Aktivitäten, die Auskunft über den aktuellen emotionalen Zustand einer Person geben, konnte im Labor nachgewiesen werden, dass Menschen signifikant weniger stark aversiv auf Schimpf- und Tadelwörter aus der Kindheit reagieren, wenn sie sich im fremdsprachlichen Kontext befinden.²¹⁴ Aber auch die emotionale Wirkung von po-sitiv konnotierten Phrasen wie „ich liebe dich“ ist in einer Fremdsprache weniger stark ausgeprägt als in der eigenen Muttersprache.²¹⁵ Auf neuronaler Ebene konnte gezeigt werden, dass die Lektüre von emotional beladenen Textpassagen in einer Fremdsprache

²¹⁰ Vgl. Bereby-Meyer et al. 2018, S. 4ff.

²¹¹ Siehe Hadjichristidis et al. 2017a bzw. Gao et al. 2015.

²¹² Vgl. Costa et al. 2017, S. 147f.; Hayakawa et al. 2016, S. 792f.

²¹³ Vgl. Caldwell-Harris 2014, S. 1f.; Pavlenko 2012, S. 405ff.

²¹⁴ Vgl. Harris et al. 2003, S. 565ff.

²¹⁵ Vgl. Dewaele 2008, S. 1763ff.

mit einer geringeren Aktivierung von Gehirnregionen wie Amygdala, die für die emotionale Verarbeitung zuständig sind, verbunden ist.²¹⁶ Gemäß diesem Erklärungsansatz resultiert der FLE aus der Abschwächung von affektiven Reaktionen. Dies führt wiederum dazu, dass sich das elaborierte System 2 in einer Entscheidungssituation besser entfalten kann. Mithilfe dieses Erklärungsansatzes lassen sich vor alle höhere Risikobereitschaft und stärker utilitaristisch ausgeprägtes Verhalten erklären.²¹⁷

Speziell im Bereich des ethischen Verhaltens bzw. der moralischen Urteile ist der auf Emotionen basierende Erklärungsansatz nicht unumstritten. So konnten Geipel et al. 2015b zwar den positiven Einfluss einer fremden Sprache auf das utilitaristische Verhalten in der „Footbridge“-Version des Trolley-Problems messen, die Auswertung der emotionalen Ratings ergab jedoch, dass das Entscheidungsverhalten der Probanden nicht von den Emotionen mediiert war. Die Autoren führen den FLE stattdessen auf die Abschwächung des Einflusses von moralisch-sozialen Normen zurück.²¹⁸ Dieser Erklärungsansatz basiert auf dem Konzept des sprachabhängigen sozio-kulturellen Lernprozesses. Demnach sind unsere Erfahrungen zusammen mit dem linguistischen Kontext, in dem sie erstmalig auftraten, im Langzeitgedächtnis gespeichert.²¹⁹ Folglich kann sich eine Person leichter und schneller an die erlernten Inhalte erinnern, wenn diese in ein und derselben Sprache abgerufen und kodiert werden.²²⁰ Weil die moralisch-sozialen Normen i. d. R. über verbale Kommunikation in der eigenen Muttersprache übermittelt werden, sind diese Normen in einer Fremdsprache weniger stark präsent und üben somit geringeren Einfluss aus.²²¹

Ein anderer Erklärungsansatz für den FLE basiert auf dem bereits im Zusammenhang mit der Lesbarkeit eingeführten DFE.²²² Das Hören, Lesen, Schreiben sowie Sprechen in einer Fremdsprache erfolgt im Vergleich zur eigenen Muttersprache weniger flüssig und erfordert mehr an kognitiven Ressourcen.²²³ Eine Fremdsprache hemmt also den

²¹⁶ Vgl. Hsu et al. 2015, S. 282ff. Interessanterweise sind die gleichen Gehirnregionen auch in Gewinn- bzw. Verlustsituationen involviert, vgl. De Martino et al. 2006, S. 684ff.

²¹⁷ Vgl. Hayakawa et al. 2016, S. 792.

²¹⁸ Vgl. Geipel et al. 2015b, S. 5ff. Siehe auch Geipel et al. 2015a.

²¹⁹ Vgl. Schrauf und Rubin 2000, S. 616ff.

²²⁰ Vgl. Marian und Neisser 2000, S. 361.

²²¹ Vgl. Hadjichristidis et al. 2017b, S. 648.

²²² Siehe Kapitel 2.3.1.

²²³ Vgl. Costa et al. 2017, S. 147; Favreau und Segalowitz 1983, S. 565ff.

mentalen Fluss und sollte im Sinne der metakognitiven Verarbeitungsschwierigkeit das analytische System 2 aktivieren.²²⁴

Gemäß dem DFE kann also erwartet werden, dass der FLE auch in emotionsneutralen Entscheidungssituationen zu beobachten sein wird. Die Ergebnisse aus der Untersuchung von Costa et al. 2014a liefern jedoch wenig Hinweise zugunsten dieses Erklärungsansatzes. In einer Reihe von Experimenten konnten die Autoren den FLE nur in Entscheidungssituationen mit einer emotionalen Komponente nachweisen. Dagegen war der FLE bei abstrakten und rein auf Logik basierenden Aufgaben wie dem Cognitive-Reflection-Test abwesend.²²⁵ Zu ähnlichen Ergebnissen gelangen auch Vives et al. 2018 bei der Untersuchung der Repräsentativitätsheuristik und des Rückschaufehlers. Die Autoren konnten in diesem Zusammenhang keinen FLE feststellen und schlussfolgerten daraus, dass eine Fremdsprache nicht unmittelbar über den analytischen Denkmodus, sondern in erster Linie über reduzierte Emotionen operiert.²²⁶

Beim dritten Erklärungsansatz wird von höherer *psychologischer Distanz*²²⁷ ausgegangen. Erkenntnisse aus der Psychotherapie legen nahe, dass die distanzierende Wirkung einer Fremdsprache es den Menschen ermöglicht, leichter über unangenehme Themen und traumatische Erlebnisse zu sprechen.²²⁸ Im Zusammenhang mit höherer psychologischer Distanz wurden ähnliche Effekte auf den Einrahmungseffekt, das Risikoverhalten und ethische Entscheidungen festgestellt wie beim FLE.²²⁹ Der Einsatz einer fremden Sprache bedingt beim Entscheidungsträger eine globalere Sichtweise auf Probleme. Dies führt wiederum dazu, dass das konsequentialistische Verhalten gefördert wird. Außerdem fallen auf einer hohen Abstraktionsstufe irrelevante kontextuelle Details viel weniger ins Gewicht, womit bspw. die Abschwächung des Einrahmungseffekts erklärt werden kann.²³⁰ In diesem Kontext muss jedoch kritisch angemerkt werden, dass der

²²⁴ Vgl. Keysar et al. 2012, S. 661. Neurologische Studien zeigten, dass die Verarbeitung von Informationen in einer Fremdsprache diejenigen Gehirnregionen aktiviert, die mit kontrollierten Prozessen in Verbindung stehen, vgl. Branzi et al. 2016, S. 2367ff. Diese Gehirnregionen werden wiederum mit utilitaristischem Verhalten und reduzierter Risiko- bzw. Verlustaversion in Verbindung gebracht, vgl. Greene et al. 2004, S. 389ff.

²²⁵ Vgl. Costa et al. 2014a, S. 251f.

²²⁶ Vgl. Vives et al. 2018, S. 11.

²²⁷ Der Begriff psychologische Distanz stammt aus der Construal-Level-Theorie, siehe Trope und Liberman 2010. Höhere psychologische Distanz geht mit einem höheren Grad an mentaler Abstraktion einher und umgekehrt. Auf einer hohen Abstraktionsstufe werden Sachverhalte eher abstrakt, schematisch und kontextfrei repräsentiert. Dagegen ist eine niedrige Abstraktionsstufe durch konkrete, unstrukturierte und kontextabhängige Repräsentationen gekennzeichnet, vgl. Liberman und Trope 2008, S. 1201.

²²⁸ Vgl. Halberstadt 1994, S. 93ff.; Marcos 1976, S. 552ff.

²²⁹ Vgl. Aguilar et al. 2013, S. 449ff.; Lerner et al. 2015, S. 99ff.; Raue et al. 2015, S. 256ff.

²³⁰ Vgl. Hayakawa et al. 2016, S. 792f.

genaue Wirkungsmechanismus der psychologischen Distanz vor allem im Zusammenhang mit ethischem Verhalten und moralischen Urteilen noch nicht hinreichend geklärt ist.²³¹

Die drei oben vorgestellten Erklärungsansätze für den FLE können auch einander ergänzen. Metakognitive Verarbeitungsschwierigkeiten und/oder reduzierte Emotionen können bspw. gleichzeitig zu höherer psychologischer Distanz führen.²³² Aus diesem Grund ist es wichtig, bei der Untersuchung des FLE alle drei Erklärungsansätze gleichzeitig zu berücksichtigen.

Im Zusammenhang mit dem FLE spielt die Sprachkompetenz eine wichtige Rolle. Beim Einsatz einer Fremdsprache, die auf einem sehr hohen bzw. fast muttersprachlichen Niveau beherrscht wird, ist das Auftreten des FLE sehr unwahrscheinlich. Andererseits löst ein niedriges Sprachkompetenzniveau hohen kognitiven Aufwand aus, sodass das analytische System 2 von der extrinsischen kognitiven Belastung beansprucht und die eigentliche Entscheidungsaufgabe im heuristischen System-1-Denkmodus erfolgen wird. Insgesamt ist also der FLE bei einer mittelhohen Sprachkompetenz zu erwarten.²³³

2.3.2.2.3 Offene Fragen

Im vorangehenden Kapitel wurde thematisiert, dass bezüglich des genauen Erklärungsansatzes für den FLE immer noch Fragen bestehen. Je nach Entscheidungssituation und Kontext können unterschiedliche und/oder mehrere Erklärungsansätze gleichzeitig zum Tragen kommen. Das Ziel des vorliegenden Abschnitts ist es, weitere offene Fragen im Zusammenhang mit dem FLE zu adressieren.

Zunächst ist es im Zusammenhang mit dem Risikoverhalten unklar, ob eine Fremdsprache die generelle Risikobereitschaft erhöht oder eher eine strategische Sicht auf Risiken fördert. Im ersteren Fall wird nicht danach differenziert, ob es sich um vorteilhafte oder unerwünschte Risiken handelt. Beim strategischen Risikoverhalten werden dagegen nur die aus normativer Sicht vorteilhaften Risiken eingegangen und die unerwünschten Risiken gemieden.²³⁴

²³¹ Siehe z. B. Eyal et al. 2008; Gong und Medin 2012.

²³² Vgl. Alter und Oppenheimer 2008, S. 161ff.; Van Boven et al. 2010, S. 872ff.

²³³ Vgl. Costa et al. 2017, S. 148. In diesem Zusammenhang müssen auch weitere relevante Faktoren wie z. B. das Alter und der Kontext des Spracherwerbs beachtet werden, vgl. Hayakawa et al. 2016, S. 793. Auch der kulturelle Einfluss einer Fremdsprache und die Ähnlichkeit zur eigenen Muttersprache könnten ebenfalls für das Zustandekommen des FLE ausschlaggebend sein, vgl. Dylman und Champoux-Larsson 2020, S. 1ff.

²³⁴ Vgl. Hayakawa et al. 2017a, S. 2.

Tabelle 4: Generelle vs. strategische Risikobereitschaft beim FLE

| Studie | Domäne ^a | Cohens <i>d</i> (N) ^b | <i>p</i> -Wert |
|-------------------------------------|-------------------------|----------------------------------|----------------|
| Costa et al. 2014a; Experiment 3a | Gewinn – Vorteilhaft | 0,10 (300) | 0,542 |
| Costa et al. 2014a; Experiment 3a | Gewinn – Unvorteilhaft | -0,29 (300) | 0,027** |
| Keysar et al. 2012; Experiment 2 | Gewinn – Vorteilhaft | 0,44 (146) | 0,008*** |
| Keysar et al. 2012; Experiment 3 | Gewinn – Vorteilhaft | 0,55 (54) | 0,038** |
| Hayakawa et al. 2017a; Experiment 1 | Verlust – Vorteilhaft | 0,04 (117) | 0,821 |
| Hayakawa et al. 2017a; Experiment 1 | Verlust – Unvorteilhaft | -0,44 (121) | 0,017** |
| Hayakawa et al. 2017a; Experiment 2 | Verlust – Unvorteilhaft | 0,12 (155) | 0,330 |
| Hayakawa et al. 2017a; Experiment 3 | Verlust – Vorteilhaft | 0,03 (97) | 0,820 |
| Hayakawa et al. 2017a; Experiment 3 | Verlust – Unvorteilhaft | 0,08 (97) | 0,640 |
| Hayakawa et al. 2017a; Experiment 3 | Gewinn – Vorteilhaft | -0,07 (97) | 0,750 |
| Hayakawa et al. 2017a; Experiment 3 | Gewinn – Unvorteilhaft | -0,11 (97) | 0,580 |
| Hayakawa et al. 2017a; Experiment 4 | Verlust – Unvorteilhaft | 0,31 (197) | 0,07* |

*, **, *** repräsentieren eine Signifikanz auf den Niveaus 10 %, 5 % und 1 %.

^a Domäne steht für den Kontext des untersuchten Risikoverhaltens je nachdem, ob es sich um eine Verlust- oder Gewinnsituation und ein vorteilhaftes oder unvorteilhaftes Risiko handelt.

^b Cohens *d* ist die Effektgröße für Mittelwertunterschiede zwischen zwei Gruppen. Positive/negative Werte bedeuten, dass in der Fremdsprache mehr/weniger risikobehaftete Optionen gewählt wurden. Des Weiteren ist in Klammern die Gesamtzahl der Probanden angegeben.

Quelle: In Anlehnung an Hayakawa et al. 2017a, S. 10.

Mit Rückgriff auf bereits existierende Studien und anhand von vier eigenen Experimenten untersuchten Hayakawa et al. 2017a diese Fragestellung. Dabei wurde das Risikoverhalten mithilfe von Lotterien aus unterschiedlichen Domänen (Gewinn/Verlust und vorteilhaftes/unvorteilhaftes Risiko) analysiert (siehe Tabelle 4). Insgesamt wurden relativ divergente Ergebnisse gewonnen. Der FLE – positiver Wert für Cohens *d* zusammen mit einem signifikanten *p*-Wert – konnte in nur drei von 12 Experimenten festgestellt werden, womit die generelle Risikobereitschaft infolge des Einsatzes einer Fremdsprache angezweifelt werden kann. Ferner entfallen von den drei signifikanten Ergebnissen zwei auf vorteilhafte und eins auf unvorteilhafte Risikosituationen. Damit findet die These mit der strategischen Risikobereitschaft ebenfalls wenig Unterstützung. Wird nur die Gewinn-domäne betrachtet, fällt auf, dass die beiden signifikanten Ergebnisse auf vorteilhafte Risiken entfallen. Dies kann als ein Indiz dafür interpretiert werden, dass der FLE im Gewinnbereich vor allem in Verbindung mit vorteilhaften Risiken zu erwarten ist.

Weil die Erkenntnisse aus der Studie von Hayakawa et al. 2017a nicht besonders aufschlussreich waren, plädieren die Autoren für eine weitere Untersuchung der möglichen moderierenden Faktoren des FLE im Risikokontext. Ein solcher Faktor könnte z. B. die sprachliche Gestaltung sein. In der Untersuchung von Winkler et al. 2016 wurde bspw. der nivellierende Einfluss einer Fremdsprache auf den Einrahmungseffekt nur in Experiment 1 (Asian-Disease-Problem) jedoch nicht in Experiment 2 (Finanzentscheidung) beobachtet. Während das Asian-Disease-Problem in mehreren Sätzen formuliert war, handelte es sich bei der Finanzentscheidung in Experiment 2 nur um wenige einzelne Wörter. Offensichtlich üben linguistische Charakteristika einen wichtigen Einfluss im Zusammenhang mit dem FLE aus.²³⁵

In allen bisherigen Studien zum FLE wurden Situationen mit kalkulierbaren Risiken und bekannten Konsequenzen einer Handlung untersucht. Wie bereits in Kapitel 2.2.3 erläutert, ist unter solchen Prämissen eine analytische Herangehensweise an das Problem den heuristischen Entscheidungen überlegen. In komplexen und unzuverlässig prognostizierbaren Situationen starker Unsicherheit können einfache Problemlösungsansätze in Form der FF-Heuristiken manchmal sogar zu besseren und robusteren Resultaten führen als normative Optimierungsmodelle. Falls die Informationsverarbeitung in einer fremden Sprache das heuristische Denken konterkariert, stellt sich die Frage, ob sich daraus nachteilige Effekte für die Entscheidungen unter Unsicherheit oder Ungewissheit ergeben werden.²³⁶

Im Kontext der ethischen Entscheidungen bleibt es zunächst unklar, ob der FLE von der gestiegenen utilitaristischen oder reduzierten deontologischen Neigung bzw. beiden Merkmalen gleichzeitig verursacht wird. Mithilfe des Prozess-Dissoziation-Verfahrens lässt sich diese Frage klären.²³⁷ Hayakawa et al. 2017b und Muda et al. 2018 lieferten die erste empirische Evidenz dafür, dass eine Fremdsprache sowohl die rein deontologische als auch – überraschenderweise – utilitaristische Neigungen reduziert. Diese Befunde widersprechen dem DFE als Erklärungsansatz für den FLE, weil offenbar das konsequentialistische Denken infolge des Einsatzes einer fremden Sprache nicht gefördert wird. Dagegen wären sowohl die auf Emotionen als auch die auf sozialen Normen

²³⁵ Vgl. Winkler et al. 2016, S. 434.

²³⁶ Zum experimentellen Testen von FF-Heuristiken (z. B. TTB-Heuristik oder Rekognitionsheuristik) ist es keinesfalls erforderlich, die wirkliche (ungewisse) Welt im Labor zu simulieren. Stattdessen reicht es aus, wenn den Probanden nur ein Teil der relevanten Informationen vorenthalten wird, vgl. Polonioli 2018, S. 3.

²³⁷ Das Prozess-Dissoziation-Verfahren wurde ursprünglich von Jacoby 1991 entwickelt. Conway und Gawronski 2013 adaptierten dieses Verfahren auf ethische Dilemmata. Mithilfe der kongruenten (d. h., deontologische und utilitaristische Prinzipien stehen miteinander nicht in Konflikt) und inkongruenten (d. h., deontologische und utilitaristische Prinzipien stehen miteinander in Konflikt) Dilemmata lässt sich ein mathematischer Wert für die Neigung, die auf rein deontologische bzw. utilitaristische Motivation zurückzuführen ist, bestimmen.

basierenden Erklärungsansätze mit dem Befund der reduzierten deontologischen Neigung zu vereinbaren.²³⁸

Resümierend kann gesagt werden, dass der FLE ein facettenreiches linguistisches Phänomen ist. Der kognitive Einfluss einer Fremdsprache konnte in zahlreichen Verhaltensbereichen nachgewiesen und in unterschiedlichsten Sprachkonstellationen repliziert werden. Der FLE-Forschungszweig ist relativ jung und einige offene Fragen theoretischer und praktischer Natur müssen noch geklärt werden. Dies betrifft sowohl die Erklärungsansätze als auch die möglichen moderierenden Einflussfaktoren wie sprachliche Gestaltung.

2.4 Management Reporting

2.4.1 Begriffsabgrenzung und Aufgaben

Bereits in den 1960er-Jahren fanden die ersten intensiven theoretischen Auseinandersetzungen mit dem Thema Berichtswesen (engl. Reporting) statt.²³⁹ Bis heute ist diese Thematik nicht zuletzt aufgrund ihrer hohen praktischen Relevanz aktuell geblieben. Unter dem Begriff des betrieblichen Berichtswesens (Reporting i. w. S.) werden alle Institutionen und Vorgänge im Unternehmen verstanden, die zur Erarbeitung, Weiterleitung und Verarbeitung von steuerungsrelevanten Informationen²⁴⁰ dienen.²⁴¹ Dabei gehören alle Personenkreise innerhalb und außerhalb des Unternehmens zu den potenziellen Informationsadressaten.²⁴² Das innerbetriebliche Berichtswesen (Reporting

²³⁸ Vgl. Muda et al. 2018, S. 325.

²³⁹ Siehe u. a. Blohm und Heinrich 1967.

²⁴⁰ In der Literatur wird häufig zwischen den Begriffen Information und Daten unterschieden. Während sich die Informationen streng genommen nur auf die entscheidungsrelevanten Inhalte beziehen, können Daten auch irrelevante Aspekte enthalten, vgl. Iselin 1993, S. 249. Weil Berichte in der betrieblichen Praxis im Regelfall keine für die Entscheidung völlig überflüssigen Angaben aufweisen, wird im Rahmen dieser Arbeit auf eine konzeptionelle Unterscheidung zwischen den beiden Begriffen verzichtet.

²⁴¹ Vgl. Blohm 1982, S. 866. Eine ähnliche Begriffsabgrenzung wird auch von der International Group of Controlling (IGC) vorgenommen, vgl. IGC 2011, S. 33ff. Diese Sichtweise auf das Berichtswesen deckt also den gesamten Informationsprozess, angefangen bei der Informationsgewinnung bis einschließlich der Informationsnutzung, ab. Andere Autoren sehen hingegen die Phase der Informationsnutzung nicht mehr als Teil des Berichtswesens, vgl. Göpfert 2002, Sp. 144; Schmidt 1980, Sp. 321f. In der engsten Auslegung wird das Berichtswesen hauptsächlich auf die Phasen der Erstellung und Weiterleitung der Berichte beschränkt, vgl. Horvath et al. 2020, S. 329.

²⁴² Zu den externen Informationsadressaten gehören z. B. Investoren, Kunden oder Lieferanten. Die Adressaten des internen Berichtswesens – auch Controlling genannt – sind primär die Unternehmensleitung und andere Mitarbeiter auf unterschiedlichsten Hierarchieebenen. Im Gegensatz zum externen ist das interne Berichtswesen i. d. R. an keine handels- oder steuerrechtlichen Normen und Publizitätspflichten gebunden, vgl. Busse von Colbe 2011, S. 652ff.

i. e. S.) zielt hingegen nur auf die internen Informationsflüsse und Adressaten ab.²⁴³ Schließlich wird beim Management Reporting der Nutzerkreis der Informationen nur auf die mit Führungsaufgaben betrauten Mitarbeiter beschränkt.²⁴⁴

Die Kernaufgaben des Management Reporting umfassen die Aufbereitung, Darstellung und Übermittlung von steuerungsrelevanten Informationen. Diese Kernaufgaben können aber nicht losgelöst von den angrenzenden Phasen der Informationsgewinnung und Informationsnutzung betrachtet werden. Die Aufbereitung und Darstellung hängen zwangsläufig davon ab, welche Informationen zuvor gewonnen wurden. Ebenso beeinflusst die Nutzungssituation maßgeblich die Darstellungsform der übermittelten Inhalte. Ist außerdem der Informationsbedarf der Adressaten im Vorfeld unbekannt, muss dieser vor der eigentlichen Phase der Informationsgewinnung bestimmt werden. Insgesamt können die Aufgaben des Management Reporting im Phasenmodell des Informationsprozesses wie folgt konkretisiert werden (siehe Abbildung 10):²⁴⁵

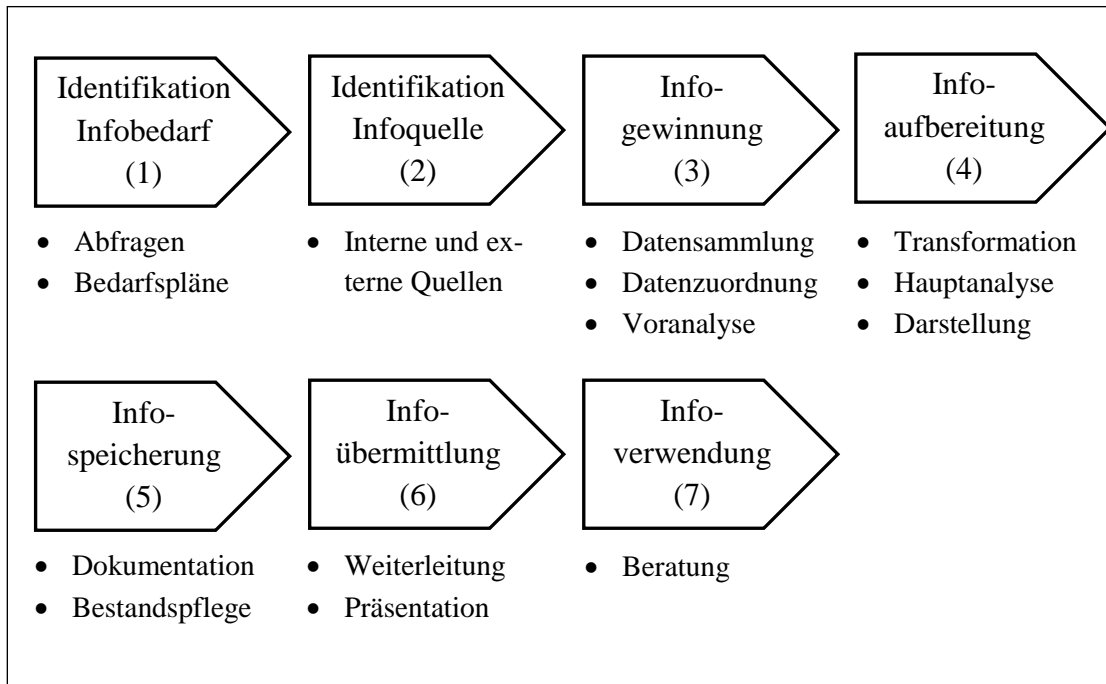
1. Identifikation Informationsbedarf: Ausgehend vom Informationsbedarf des Empfängers werden Anforderungen an den Bericht im Hinblick auf die Darstellungsform, Inhalte, Struktur und den Umfang abgeleitet. Im Management Reporting kann dies bspw. anhand von Vorlagen – auch Templates genannt – zur strukturierten Sammlung von Inputmaterial und Bedarfsplänen erfolgen.
2. Identifikation Informationsquelle: Nach der Ermittlung des Informationsbedarfs müssen geeignete Informationsquellen identifiziert und zugänglich gemacht werden. Bei unternehmensinternen Quellen (z. B. Daten des Rechnungswesens) sind die Zugriffsrechte zu prüfen. Bei externen Quellen müssen darüber hinaus die Zuverlässigkeit der Daten und etwaige Kosten der Nutzung beachtet werden.
3. Informationsgewinnung: In der Phase der Informationsgewinnung werden die Rohdaten gesammelt und nach einer ersten Prüfung (Aktualität, Plausibilität, Vollständigkeit usw.) den geplanten Berichtsinhalten zugeordnet.
4. Informationsaufbereitung: Um den individuellen Informationsbedürfnissen der Adressaten gerecht zu werden, müssen die gewonnenen Daten aufbereitet und transformiert werden. Diese Phase schließt eine Hauptanalyse (z. B. Soll-Ist-Vergleiche, Abweichungsanalysen, Zukunftsprognosen) und eine bedarfsgerechte Darstellung der Ergebnisse ein.

²⁴³ Vgl. Kammer 2005, S. 69f. In diesem Fall sind nur die zur operativen und strategischen Steuerung des Unternehmens geeignete Informationsflüsse – insbesondere in Form von Berichten – und keine Pressemitteilungen, Unternehmenszeitschriften o. Ä. gemeint, vgl. Welge 1988, S. 384.

²⁴⁴ Vgl. Taschner 2013a, S. 2. Empirische Untersuchungen zeigen regelmäßig, dass das Erstellen von Berichten für die Führungskräfte zu den wichtigsten und häufig auch zeitintensivsten Tätigkeiten des Controllings gehört, vgl. Taschner 2015, S. 1f.

²⁴⁵ Vgl. im Folgenden Gleich und Temmel 2008, S. 74ff.; Taschner 2013b, S. 40ff.

Abbildung 10: Aufgaben im Management-Reporting-Prozess



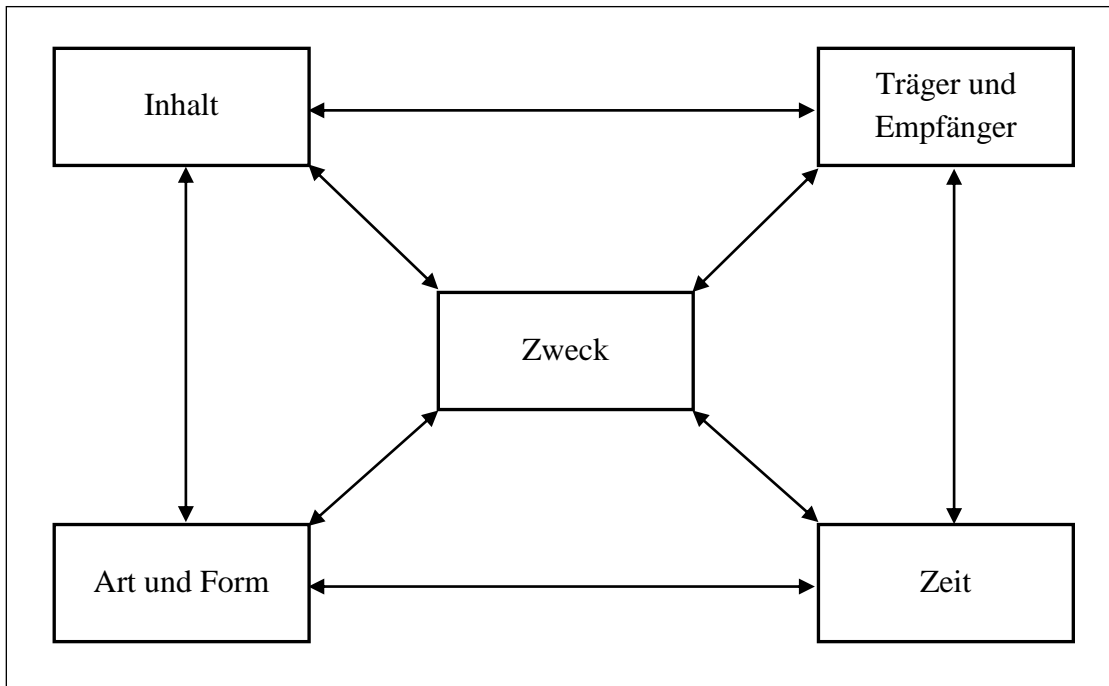
Quelle: In Anlehnung an Gleich und Temmel 2008, S. 77; Taschner 2013b, S. 41.

5. Informationsspeicherung: Obwohl eine Zwischenspeicherung prinzipiell nach jedem Schritt wichtig und erwähnenswert ist, stellen die Sicherung und Archivierung der fertig gestellten Berichte – sowie deren Zwischenstände – einen wichtigen Teil des Reportingprozesses dar. Ziel ist es, die Berichtsinhalte den autorisierten Personengruppen dauerhaft und zuverlässig zur Verfügung zu stellen.
6. Informationsübermittlung: Die Berichtsinhalte können erst zu (subjektiver) Information werden, wenn diese an den Berichtsempfänger übermittelt werden. Die Weiterleitung kann dabei über verschiedene Kanäle erfolgen. Häufig werden die übermittelten Inhalte zusätzlich von einer persönlichen Kommunikation (z. B. in Form von Präsentationen und Erklärungen) begleitet.
7. Informationsverwendung: In der Phase der Informationsverwendung geht es darum, den Berichtsempfänger beratend zu unterstützen. Dies umfasst sowohl Hilfe bei der Interpretation und Nutzung von Berichtsinhalten als auch Unterstützung bei der konkreten Umsetzung der aus dem Bericht abgeleiteten Maßnahmen.

Der Handlungsrahmen einer komplexen Organisation – etwa in Gestalt eines großen internationalen Unternehmens – erfordert einen hohen Grad der Arbeitsteilung.²⁴⁶ Aus diesem Grund müssen vor allem in großen Unternehmen die beschriebenen Phasen im Management-Reporting-Prozess in mehr oder weniger intensiver Form durchlaufen

²⁴⁶ Vgl. Horvath et al. 2020, S. 43.

Abbildung 11: Gestaltungsdimensionen des Berichtswesens



Quelle: In Anlehnung an Blohm 1982, S. 869; Weber und Schäffer 2016, S. 238.

werden. Die personelle Trennung zwischen dem Berichtsersteller und Berichtsnutzer kann außerdem zu verschiedenen intendierten und nicht intendierten Verhaltenswirkungen führen, die noch in Kapitel 2.4.3 ausführlich behandelt werden.

2.4.2 Gestaltungsdimensionen

Wenngleich der grundlegende Management-Reporting-Prozess immer nach dem gleichen Schema verläuft, so können sich die Ergebnisse dieses Prozesses in Form der fertig gestellten Berichte stark voneinander unterscheiden.²⁴⁷ Im Folgenden werden die fünf wesentlichen Gestaltungsdimensionen des Berichtswesens erläutert (siehe Abbildung 11).²⁴⁸

²⁴⁷ Vgl. Taschner 2013b, S. 43.

²⁴⁸ Neben den Gestaltungsdimensionen wurden in der Theorie unterschiedliche Anforderungskriterien an ein effizientes Berichtswesen formuliert, vgl. u. a. Blohm 1982, S. 870f. Kammer 2005 vergleicht verschiedene Anforderungskataloge und fasst die sechs aus seiner Sicht wichtigsten Kriterien wie folgt zusammen: Flexibilität, Objektivität, sachliche und zeitliche Relevanz, Transparenz sowie Wirtschaftlichkeit, vgl. Kammer 2005, S. 87.

In der Literatur herrscht weitgehend Einigkeit darüber, dass am Anfang aller Gestaltungsüberlegungen zunächst der Berichtszweck – sprich das „Wozu“ – steht.²⁴⁹ Der Berichtszweck wird vom Informationsbedarf des Berichtsempfängers determiniert.²⁵⁰ Aus dem Berichtszweck ergibt sich, über welche Inhalte zu berichten ist und auf welche Weise (Art, Form und Zyklus) dies geschehen soll. Zu beachten ist, dass alle Gestaltungsdimensionen miteinander in einer wechselseitigen Beziehung stehen.

2.4.2.1 Zweck

Die mit einem Bericht übermittelten Informationen erfüllen unterschiedliche Aufgaben und können vom Berichtsempfänger für unterschiedliche Zwecke genutzt werden. Also existieren mehrere Berichtszwecke nebeneinander, die sich wie folgt unterteilen lassen.²⁵¹

- **Dokumentation:** Vorgänge und Ereignisse mit den daraus resultierenden Konsequenzen sollen im Unternehmen dauerhaft festgehalten und für spätere Nutzung zugänglich gemacht werden. Auf die dokumentierten Informationen wird z. B. zurückgegriffen, um retrospektive Vergleiche und Zeitreihenanalysen zu erstellen. Außerdem kann die Dokumentation dazu dienen, die getroffenen Entscheidungen im Nachhinein zu rechtfertigen.
- **Kontrolle:** Anhand des Vergleichs von geplanten und realisierten Größen wird eine Kontrolle durchgeführt. Die im vorangehenden Punkt erwähnte Dokumentationsaufgabe ist somit eine unabdingbare Voraussetzung für Kontrollzwecke. Kontrollhandlungen können sich auf unterschiedlichste monetäre und nicht monetäre Größen sowie Zielobjekte (z. B. Produkte, Aufträge, Geschäftseinheiten) beziehen. In vielen regelmäßig erstellten Standardberichten ist der Kontrollzweck stark ausgeprägt. So dient bspw. der klassische Monatsbericht häufig primär dazu, die Einhaltung von Kosten-, Mengen- und Umsatzvorgaben zu überprüfen.
- **Planung- und Entscheidungsunterstützung:** Rationales und zielgerichtetes Handeln erfordert eine fundierte Informationsbasis. Die Bereitstellung von relevanten Informationen für Planungs- und Entscheidungsaufgaben ist dementsprechend ein wichtiger Berichtszweck. Hochrechnungen aus der Vergangenheit, Prognosen, Bench-

²⁴⁹ Vgl. Blohm 1982, S. 868; Koch 1994, S. 57; Küpper et al. 2013, S. 236; Weber und Schäffer 2016, S. 238; Welge 1988, S. 385.

²⁵⁰ Koch 1994 unterscheidet in diesem Kontext zwischen einem objektiven und subjektiven Informationsbedarf. Der objektive Informationsbedarf leitet sich unmittelbar aus den Unternehmenszielen sowie aus externen Umweltbedingungen ab. Gleichzeitig bestimmen die individuellen Eigenschaften des Informationsempfängers den subjektiven Informationsbedarf, vgl. Koch 1994, S. 57.

²⁵¹ Mit teils unterschiedlichen Bezeichnungen und Kategorisierungen vgl. im Folgenden Horvath et al. 2020, S. 329; Küpper et al. 2013, S. 230f.; Taschner 2013b, S. 46; Weber und Schäffer 2016, S. 238.

marks und Vergleiche zwischen Alternativen sind typische Instrumente zur Planungs- und Entscheidungsunterstützung. Werden darüber hinaus bedeutende Abweichungen im Laufe der Kontrolltätigkeit festgestellt, sollen im Rahmen der Steuerungsfunktion entsprechende Gegenmaßnahmen vorgeschlagen werden.²⁵²

Neben diesen Hauptzwecken wird in der Literatur noch eine Vielzahl weiterer Zwecke, wie z. B. die Auslösung von Bearbeitungsvorgängen oder schlicht Interessenweckung beim Entscheidungsträger, diskutiert.²⁵³ Letzten Endes soll ein Bericht in der einen oder anderen Form das Verhalten des Berichtnutzers beeinflussen. Die Verhaltensbeeinflussung könnte daher als der wichtigste Berichtszweck angesehen werden.²⁵⁴

2.4.2.2 Inhalt

Weil die Inhalte eines Berichts unmittelbar vom Abnehmer wahrgenommen werden, wird diese Gestaltungsdimension vielfach ins Zentrum der Betrachtung gerückt. Jedoch ist in diesem Zusammenhang zu beachten, dass die Berichtsinhalte – wie oben bereits erläutert – letztlich von dem Berichtszweck determiniert werden. Des Weiteren hängt der Nutzen der Berichtsinhalte wesentlich von den anderen Gestaltungsdimensionen ab. Insbesondere spielt es eine große Rolle, wann (Aktualität, zeitliche Abstände usw.) und wie (Darstellungsart und -form) die Berichtsinhalte übermittelt wurden. Eine losgelöste Priorisierung dieser Dimension wird somit nicht als gerechtfertigt erachtet.²⁵⁵

Bei der inhaltlichen Ausgestaltung muss eine Reihe von Festlegungen im Hinblick auf die folgenden Aspekte vorgenommen werden:

- Informationsgegenstand: Informationsgegenstände werden auch als Objekte des Reporting bezeichnet.²⁵⁶ Sie beziehen sich auf Sachverhalte in der Realität, die im Bericht abgebildet werden.²⁵⁷ Typische Informationsobjekte im Reporting sind bspw. Kunden, Produkte, Standorte, Tochtergesellschaften usw.
- Informationsart: Dabei handelt es sich um eine Kategorisierung des Aussagewerts. Die Berichtsinformation kann aus rein deskriptiven Aussagen bestehen, aber auch einen explanatorischen, normativen oder prognostischen Charakter haben. Im Bericht sollen zunächst die relevanten Fakten verständlich dargelegt werden. Danach wird möglicherweise eine Erläuterung und/oder Relativierung dieser Informationen

²⁵² Vgl. Schroeter 2002, S. 331.

²⁵³ Vgl. Asser 1974, S. 660; Baier 2008, S. 101.

²⁵⁴ Vgl. Koch 1994, S. 60.

²⁵⁵ Vgl. Taschner 2013b, S. 47

²⁵⁶ Vgl. Gräf und Nase 2008, S. 43ff.

²⁵⁷ Vgl. Koch 1994, S. 61.

erforderlich sein. Schließlich können eine Wertung und darauf aufbauend die Ableitung von zukünftigen Szenarien folgen.²⁵⁸

- Informationseigenschaft: Um das von dem Bericht realisierte Informationsangebot konkret beschreiben zu können, reicht es nicht aus, einzig die Informationsart und den Informationsgegenstand zu benennen. Es muss stets auch hinsichtlich der Informationseigenschaft unterschieden werden. Besonders wichtige Informationseigenschaften sind z. B. der Detaillierungs- und Genauigkeitsgrad, die Aktualität, Objektivität, Zuverlässigkeit sowie die Ausgewogenheit der Berichtsinformationen. Im Prinzip kann die Liste mit den möglichen Informationseigenschaften beliebig lang erweitert werden. Dies erklärt, warum es keine universelle Taxonomie der Informationseigenschaften geben kann.²⁵⁹
- Informationstransformation: Aussagekräftige Interpretationen können i. d. R. erst nach einem Verarbeitungsprozess von mehreren Informationsobjekten entstehen. Die Generierung von gewünschten Berichtsinhalten erfolgt meist, indem die Inputdaten mit logischen und/oder mathematischen Transformationen miteinander verknüpft, neu angeordnet und strukturiert werden.

Im Allgemeinen müssen die dargebotenen Informationen im Bericht für den jeweiligen Adressatenkreis transparent, verständlich und nachvollziehbar sein. In Bezug auf die Strukturierung von Berichtsinhalten hat sich eine Darstellung von zunächst allgemeinen hin zu spezielleren Informationen bewährt.²⁶⁰

2.4.2.3 Träger und Empfänger

Bei der Abgrenzung des Management Reporting in Kapitel 2.4.1 wurde bereits eine Einschränkung hinsichtlich des Adressatenkreises vorgenommen. Demnach handelt es sich bei den Empfängern von Management-Berichten um unternehmensinterne Mitarbeiter mit Führungsaufgaben. Insbesondere in Großunternehmen kann dieser Nutzerkreis sehr weit gefasst sein und von der obersten Geschäftsführung über Bereichsverantwortliche bis hin zu Projektleitern reichen. Somit müssen notwendigerweise die individuellen Informationsbedürfnisse dieser heterogenen Nutzer bei der Ausdifferenzierung des Berichtswesens berücksichtigt werden.²⁶¹

Das Erstellen von Management-Berichten obliegt in der Praxis primär dem Controller – d. h., er ist der Berichtsträger.²⁶² Im Zuge der zunehmenden Konvergenz zwischen

²⁵⁸ Vgl. Weber und Schäffer 2016, S. 240.

²⁵⁹ Vgl. Taschner 2013b, S. 48.

²⁶⁰ Vgl. Weber und Schäffer 2016, S. 240. Eine praktische Anleitung zur Strukturierung von Berichten geben u. a. Pollmann und Rühm 2007, S. 180ff.

²⁶¹ Vgl. Taschner 2013b, S. 51.

²⁶² Vgl. Weber und Schäffer 2016, S. 249.

dem internen und externen Berichtswesen sowie der Ausrichtung am Konzept der wertorientierten Unternehmensführung gewinnen auch andere Organisationseinheiten, wie etwa die Finanzbuchhaltung, im Prozess des Management Reporting zunehmend an Bedeutung.²⁶³ Funktionsstellen, wie die IT-Abteilung, sind ebenfalls regelmäßig mehr oder weniger intensiv an der Erstellung von Berichten beteiligt.²⁶⁴ Trotz vieler Beteiligter am Berichtserstellungsprozess sollte für den Manager im Idealfall bei allen Fragen eine klare Bezugsperson zur Verfügung stehen.²⁶⁵

2.4.2.4 Zeit

In der betrieblichen Praxis sind die Berichte meist terminlich determiniert, ihre Erstellung erfolgt häufig unter Zeitdruck.²⁶⁶ Dabei ist die Schnelligkeit, mit der ein Bericht dem Empfänger zur Verfügung gestellt wird, nur ein relevanter Gesichtspunkt unter vielen. Im Allgemeinen lässt sich die zeitliche Gestaltungsdimension wie folgt differenzieren:²⁶⁷

- **Rechtzeitigkeit:** Zunächst kann der Nutzen einer Berichtsinformation nur dann gewährleistet sein, wenn diese den Entscheidungsträger rechtzeitig erreicht. Aus diesem Grund ist es für den Berichtsersteller wichtig zu wissen, bis wann ein Bericht vorliegen soll bzw. wann dieser vom Berichtsempfänger benötigt wird.
- **Aktualität:** Nicht zu verwechseln mit der Rechtzeitigkeit ist die Aktualität als weitere zeitliche Eigenschaft der Berichtsinformation. Die Aktualität stellt den Bezug zur Gegenwart her, d. h., je mehr sich ein Informationsobjekt auf gegenwärtige Sachverhalte oder Ereignisse bezieht, desto aktueller ist es. Dieser Parameter ist stark situationsbezogen. So sind bspw. die Börsenkurse der letzten Woche unter Umständen nicht mehr entscheidungsrelevant. Wohingegen Berichte zum makroökonomischen Exportvolumen auch mehrere Monate alt sein dürfen, um nach wie vor als aktuell zu gelten.²⁶⁸
- **Zyklus und Zeitraum:** Der Berichtszyklus betrifft die zeitlichen Abstände, innerhalb derer ein Reportingprozess wiederholt wird. In der Praxis sind häufig wöchentliche,

²⁶³ Die Aufgaben des Controllings vor dem Hintergrund des Trends zur Integration der externen und internen Rechnungslegung werden ausführlich von Weißenberger 2011 behandelt.

²⁶⁴ Taschner 2013a, S. 9.

²⁶⁵ Eine solche Berichterstattung „aus einer Hand“ trägt positiv zur Zufriedenheit des Managers bei, vgl. Weber und Sandt 2001, S. 26.

²⁶⁶ Vgl. Taschner 2015, S. 5.

²⁶⁷ Vgl. im Folgenden Taschner 2013b, S. 49f.

²⁶⁸ Vgl. Taschner 2013a, S. 9.

monatliche, vierteljährliche und jährliche Berichtszyklen vorzufinden.²⁶⁹ Davon zu unterscheiden ist der Berichtszeitraum, der das Zeitfenster markiert, auf das sich die Information im Bericht bezieht. Häufig sind der Berichtszyklus und Berichtszeitraum insofern deckungsgleich, als sie die gleiche „Taktung“ aufweisen: In Monatszyklen erstellte Berichte beziehen sich auf einen Zeitraum von einem Monat, vierteljährlich erstellte Berichte auf ein Quartal usw. Zeitliche Überlappungen entstehen immer dann, wenn bspw. monatlich erstellte Berichte Informationen über mehr als einen Monat enthalten. Solche Redundanzen können im Berichtswesen durchaus sinnvoll und gewollt sein.

- Dauer: Schließlich spielt die Dauer des Reportingprozesses insgesamt eine wichtige Rolle. Je schneller der Berichtsprozess durchlaufen wird, desto weniger Zeit vergeht zwischen der Entstehung des Informationsbedarfs und dessen Befriedigung über die bereitgestellten Informationen.

Bei der Gestaltung der zeitlichen Dimension müssen eine Reihe von Abwägungen vorgenommen werden. So geht höhere Aktualität bei kürzeren Berichtszyklen mit potenzieller Informationsüberlastung des Berichtsempfängers²⁷⁰ sowie gesteigertem Arbeitsaufwand auf Seiten der Berichtsersteller einher.²⁷¹ Außerdem konnte in der jüngeren Vergangenheit eine Entwicklung hin zu Systemen, die jederzeit abrufbare Echtzeitdaten enthalten, beobachtet werden. Dies hat unmittelbar einen Einfluss auf alle oben genannten zeitlichen Kategorien. Insbesondere führten aber die besseren technischen Möglichkeiten zu deutlichen Verkürzungen der Wartezeiten.²⁷²

2.4.2.5 Art und Form

Weitere Gestaltungsmöglichkeiten ergeben sich im Hinblick auf die Berichtsart und die Darstellungsform der Berichtsinhalte. Die nachfolgenden drei Berichtsarten sind in der betriebswirtschaftlichen Praxis besonders wichtig:²⁷³

- Standardbericht: Ein Standardbericht – z. B. in Gestalt des klassischen Monatsberichts – zeichnet sich dadurch aus, dass seine wesentlichen Merkmale wie Inhalt, Form und Erscheinungstermin bereits im Vorfeld feststehen und er so „standardisiert“ ist. Der Erstellung von Standardberichten müssen zweckmäßigerweise eine

²⁶⁹ Vgl. Küpper et al. 2013, S. 239. Je nach Berichtszweck kann von diesen Intervallen abgewichen und ein anderer Zyklus gewählt werden. Dies ist bspw. bei Saisonal- oder Mehrjahresberichten der Fall.

²⁷⁰ In der Literatur ist ebenfalls die englische Bezeichnung Information Overload geläufig. Dieses Phänomen wird noch in Kapitel 2.4.3 thematisiert.

²⁷¹ Vgl. Volnhals 2010, S. 90.

²⁷² Vgl. Weber und Schäffer 2016, S. 248.

²⁷³ Vgl. im Folgenden Küpper et al. 2013, S. 231ff.; Horvath et al. 2020, S. 330; Weber und Schäffer 2016, S. 238ff.

sorgfältig ausgeführte Bedarfsanalyse und konzeptionelle Überlegungen vorangehen. Dieser Berichtstyp bildet den Kern des Berichtswesens²⁷⁴ und richtet sich i. d. R. an mehrere Empfänger gleichermaßen, was seine Universalität und Wirtschaftlichkeit im Vergleich zu den anderen Berichtsarten begründet. Zugleich eignen sich die Standardberichte wenig, um spezielle und aktuell auftretende Informationsbedürfnisse zu decken.²⁷⁵

- Abweichungsbericht: Im Gegensatz zu den Standardberichten werden Abweichungsberichte lediglich dann erstellt, wenn bestimmte Toleranzwerte über- bzw. unterschritten werden und der Sachverhalt die Aufmerksamkeit des Managers erfordert. Ein wesentliches Entscheidungsproblem bei der Gestaltung von Abweichungsberichten besteht in der Bestimmung der Toleranzwerte. Einerseits müssen die Anpassungsmaßnahmen frühzeitig ergriffen werden, andererseits soll mögliche Informationsüberlastung aufgrund zu häufigen Berichtens vermieden werden.
- Bedarfsbericht: Das konstitutive Merkmal von Bedarfsberichten ist die fallweise Anforderung seitens des Managements. Sie werden immer dann angefordert, wenn der Informationsbedarf von Standard- bzw. Abweichungsberichten nicht ausreichend abgedeckt wird. Da jeder Bedarfsbericht von den individuellen Informationsbedürfnissen des Empfängers abhängt, ist die Erstellung von Bedarfsberichten nur grob planbar. Dieser Berichtstyp trifft zwar unmittelbar den geäußerten Bedarf des Berichtsiniciators, dafür sind die Möglichkeiten seiner Verwendung für andere Adressatenkreise in der Regel stark eingeschränkt.²⁷⁶

Weil moderne Management-Informationssysteme einen schnellen und individualisierten Abruf von gewünschten Informationen aus einer umfassenden Datenbank ermöglichen, verwischen die Grenzen zwischen den Standard- und Bedarfsberichten immer mehr. Das Controlling muss auf die negativen Effekte einer möglichen Informationsüberlastung achten und bei Bedarf begrenzend eingreifen. Auch wenn die Technik eine zunehmende Individualisierung erlaubt, behalten die klassischen Standardberichte nach wie vor ihre Existenzberechtigung. Zum einen sind sie im Rahmen interaktiver Steuerungssysteme gut geeignet, um die Aufmerksamkeit der Manager auf bestimmte Sachverhalte zu lenken, zum anderen bilden sie bei allen an Entscheidungen beteiligten Akteuren im Unternehmen gewissermaßen eine gemeinsame Informationsgrundlage über den Geschäftsverlauf.²⁷⁷

²⁷⁴ Eine Befragungsstudie ergab, dass rund die Hälfte der gesamten Arbeitszeit von Controllern für die Erstellung von Standardberichten aufgewendet wird, vgl. Schäffer und Weber 2015, S. 16.

²⁷⁵ Eine fallstudienartige Auseinandersetzung mit einem exemplarischen Standardbericht findet sich u. a. bei Weißmann 2005, S. 57ff.

²⁷⁶ Ein exemplarischer Bedarfsbericht zusammen mit den Vor- und Nachteilen dieses Berichtstyps findet sich u. a. bei Jung 2011, S. 145f.

²⁷⁷ Vgl. Horvath et al. 2020, S. 331; Weber und Schäffer 2016, S. 240.

Unter die formalen Eigenschaften eines Berichts fallen die generelle Erscheinungsform sowie die Art und Weise der Informationsdarstellung.²⁷⁸ Dabei stehen die Fragen der textuellen, tabellarischen und grafischen Umsetzung der Berichtsinformationen, eines benutzerfreundlichen Layouts bis hin zur adäquaten Struktur und Gliederung der Berichtsinhalte im Fokus.²⁷⁹ Diese gestalterischen Aspekte sind für das Reporting neben rein ästhetischen Gründen auch insofern von großer Relevanz, als damit gewisse Verhaltenswirkungen beim Berichtsempfänger induziert werden können. Auf einige der Verhaltenswirkungen, die für den Forschungsschwerpunkt im Rahmen dieser Arbeit wichtig sind, wird in den beiden nachfolgenden Kapiteln näher eingegangen.

2.4.3 Störungen und Verhaltenswirkungen

Sowohl auf der Erstellerseite als auch auf der Empfängerseite wird der Reportingprozess letztendlich von Menschen bestimmt. Wie bereits in Kapitel 2.1.1 gezeigt wurde, zeichnen sich die Entscheider in der Realität durch beschränkte Rationalität und limitierte kognitive Ressourcen aus. Daraus können bestimmte Störungen²⁸⁰ und nicht intendierte Verhaltenswirkungen im Berichtsprozess resultieren, die den Schwerpunkt des vorliegenden Kapitels bilden.

Damit Berichtsinformationen ihren Zweck erfüllen können, müssen diese zunächst aufbereitet, danach übermittelt und schließlich vom Berichtsempfänger richtig verstanden bzw. interpretiert werden. Dabei kann es in diesem „Senden-Empfangen-Verstehens“-Prozess auf jeder Stufe zu Störungen kommen.²⁸¹ Solche Störungen lassen sich im Kontext des Management Reporting mithilfe der Semiotik wie folgt kategorisieren:²⁸²

- **Syntaktik:** Auf der syntaktischen Ebene liegt eine Störung vor, wenn die Zeichenfolgen einer Nachricht vom Berichtsempfänger nicht so wahrgenommen werden, wie vom Berichtsersteller beabsichtigt. So kann bspw. ein Kaffeefleck auf dem Bericht die Legende zur Grafik überdecken oder der Berichtsempfänger macht Fehler beim Lesen und entziffert ein Wort falsch. Die syntaktischen Fehler können ebenso aufseiten des Berichtserstellers auftreten, wenn ihm ein Tippfehler unterläuft oder er sich bei der Präsentation der Berichtsinhalte verspricht.
- **Semantik:** Störungen auf der semantischen Ebene beziehen sich auf das inhaltliche Verstehen der Zeichen. Weil die inhaltliche Bedeutung von Zeichen stets auf einer

²⁷⁸ Vgl. Koch 1994, S. 62.

²⁷⁹ Vgl. Taschner 2013b, S. 48.

²⁸⁰ Unter Störungen wird die Differenz zwischen dem normativ gebotenen und tatsächlich beobachtbaren Verhalten verstanden. Störungen rein technischer Natur (z. B. Computerabsturz) sollen an dieser Stelle ausgeklammert werden.

²⁸¹ Vgl. Taschner 2013a, S. 15.

²⁸² Vgl. Im Folgenden Koch 1994, S. 71ff.; Küpper et al. 2013, S. 240ff.; Taschner 2013b, S. 185.

Konvention beruht, ist es möglich, dass Sender und Empfänger den übermittelten Zeichen nicht die gleiche Bedeutung beimessen. In einem solchen Fall entsteht eine semantische Störung, die sowohl aufseiten des Senders als auch des Empfängers auftreten kann. So kann bspw. ein Fachausdruck im Report entweder vom Berichtsersteller falsch verwendet oder vom Berichtsempfänger falsch interpretiert werden.

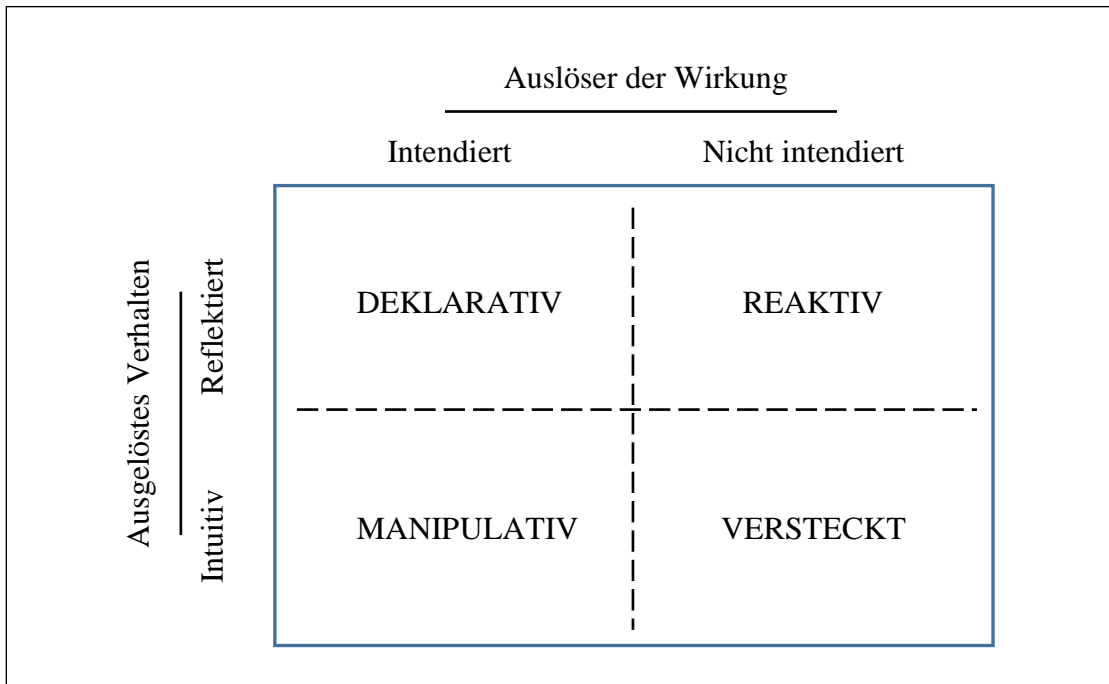
- **Pragmatik:** Pragmatische Störungen beziehen sich auf Situationen, in denen die Berichtsinformationen ihre Zweckbestimmung verfehlen und nicht das beabsichtigte Verhalten beim Berichtsempfänger auslösen. Wie bei den oben genannten Kategorien können sowohl der Berichtsersteller als auch der Berichtsempfänger für die pragmatischen Störungen verantwortlich sein. Einerseits kann der Berichtsersteller den Informationsbedarf des Adressaten falsch einschätzen und somit unpassende Informationen liefern. Andererseits kann es passieren, dass der Berichtsempfänger aus unterschiedlichsten Gründen (z. B. Misstrauen, fehlende Motivation) die Berichtsinhalte nicht akzeptiert und folglich diese als Entscheidungsgrundlage ablehnt.

Aus diesen Ausführungen geht deutlich hervor, dass die Abweichungen vom idealtypischen Informationsprozess vielfach auf die menschlichen „Unzulänglichkeiten“ zurückzuführen sind. Dabei kann es sich um Fehler, Missverständnisse sowie Fehlinterpretationen auf allen Prozessstufen und bei allen beteiligten Akteuren im Berichtswesen handeln. Eine besondere Beachtung verdient dabei die pragmatische Ebene, weil entsprechende verhaltensbezogene Störungen oft nur schwer zu identifizieren sind und gleichzeitig die Zweckerfüllung der Berichtsinformationen wesentlich beeinträchtigen.

Im Hinblick auf die verhaltensinduzierende Wirkung der Berichte kann zum einen danach differenziert werden, ob diese vom Berichtsersteller intendiert war oder nicht, zum anderen kann das von dem Bericht ausgelöste Verhalten eher intuitiv oder reflektiert sein. In Kombination stellen diese beiden Differenzierungsmerkmale die nachfolgenden vier Grundtypen der Verhaltenswirkungen von Berichten dar (siehe Abbildung 12).²⁸³

²⁸³ Vgl. im Folgenden Taschner 2015, S. 20ff.

Abbildung 12: Verhaltenswirkungen von Berichten



Quelle: In Anlehnung an Taschner 2015, S. 20.

- **Deklarative Wirkung:** Diese Wirkungsart stellt den typischen Hauptzweck von Berichten dar. Hierbei wird über den Bericht absichtlich ein bestimmtes Verhalten ausgelöst, das beim Berichtsempfänger reflektiert stattfindet. Eine bewusste Auswahlentscheidung des Managers zwischen zwei Investitionsprojekten anhand der Wirtschaftlichkeitsanalysen im Report wäre ein Beispiel für diesen Wirkungstyp. Die deklarative Wirkung stellt im Allgemeinen den erwartbaren Umgang mit Berichtsinformationen dar und ist somit für alle Beteiligten offenkundig.
- **Reaktive Wirkung:** Wirkungen dieses Typs sind dadurch gekennzeichnet, dass sie dem Berichtsnutzer unbeabsichtigt einen Anlass dazu geben, bestimmte Handlungen bewusst und reflektiert vorzunehmen. Vertrauliche Berichtsinhalte, die nicht entsprechend markiert wurden, könnten bspw. den Manager dazu veranlassen, diese Informationen an nicht autorisierte Stellen weiterzuleiten. Häufig – zumindest im Nachhinein – kann das beobachtbare Verhalten des Empfängers in einen unmittelbaren kausalen Zusammenhang mit dem Auslöser seines Verhaltens in Verbindung gebracht werden.
- **Manipulative Wirkung:** Dieser Wirkungstyp ist beabsichtigt und soll beim Berichtsempfänger eine unbewusst herbeigeführte Reaktion auslösen. So soll bspw. der Ein-

satz roter Farbe Dringlichkeit signalisieren und den Empfänger zum raschen Handeln veranlassen, eine absichtlich gewählte Formulierung bestimmte Interpretation nahelegen usw.

- Versteckte Wirkung: Versteckte Wirkungen sind weder intendiert noch für den Berichtsempfänger bewusst. Weil dieser Wirkungstyp für alle Beteiligten verborgen ist, können die daraus resultierenden Verhaltenseffekte lange Zeit unerkannt bleiben.

Häufig manifestieren sich die versteckten Wirkungen des Management Reporting in Form von Heuristiken und Biases.²⁸⁴ In der Regel sind Heuristiken effizient, d. h., sie liefern schnelle und (ausreichend) gute Entscheidungen unter Einsatz von wenigen kognitiven Ressourcen.²⁸⁵ Dagegen sind Heuristiken in Situation, in denen das analytische Denken gefragt ist, unerwünscht und führen systematisch zu Biases. Tabelle 5 enthält eine Auswahl von Biases, die einerseits in der betriebswirtschaftlichen Literatur oft genannt werden²⁸⁶ und andererseits speziell für das Controlling von besonderer Relevanz sind.²⁸⁷ Dabei können Biases grundsätzlich auf jeder Prozessstufe des Management Reporting und bei allen an der Erstellung des Berichts beteiligten Akteuren auftreten.²⁸⁸

²⁸⁴ Vgl. Taschner 2015, S. 29. Allgemein zu Heuristiken und Biases siehe Kapitel 2.2.2.

²⁸⁵ Vgl. Taschner 2013a, S. 19ff. Dieser Art von Heuristiken wurde in Kapitel 2.2.3 vorgestellt.

²⁸⁶ Siehe u. a. Bazerman und Moore 2013.

²⁸⁷ Vgl. Drews 2013, S. 49.

²⁸⁸ Für eine vertiefende Auseinandersetzung mit dem Thema Heuristiken und Biases speziell im Kontext des Management Reporting soll an dieser Stelle auf die weiterführende Literatur verwiesen werden. Basel 2012 beschäftigt sich mit dem Einsatz von Heuristiken aufseiten der Berichtsersteller. Taschner 2013b gibt eine kurze Beschreibung der relevanten Biases entlang des gesamten Management-Reporting-Prozesses. Dabei geht der Autor sowohl auf die Berichtsersteller und -empfänger als auch deren Interaktion ein. Weber et al. 2003 setzen sich mit verhaltensorientierten Problemen und insbesondere mit Biases am Beispiel von Kennzahlen und der Budgetierung auseinander. Potenzielle behavioristisch bedingte Fehlentscheidungen und mögliche Gegenmaßnahmen bei Investitionsprojekten werden u. a. von Scherpereel et al. 2015; Weber 2014 sowie Zayer und Hirsch 2006 ausführlich behandelt.

Tabelle 5: Ausgewählte Biases im Controlling

| Name oder Ursache des Bias ^a | Beschreibung | Beispiele für mögliche Auswirkungen |
|--|---|---|
| Ausgabeneffekt (Sunk Costs Effect) | Nicht mehr beeinflussbare, aufgrund früherer Entscheidungen entstandene Kosten wirken sich auf die aktuellen Entscheidungen aus | Nachschuss von Finanzmitteln, obwohl der Projekterfolg nicht mehr erreichbar ist |
| Einrahmungseffekt (Framing) | Beeinflussbarkeit von Entscheidungen über Art und Kontext der Problemformulierung | Fehlbeurteilung der Leistung eines Geschäftsbereichs aufgrund manipulativer Informationsdarstellung (z. B. aufgrund der Skalierung) |
| Herdenverhalten | Unreflektierte Übernahme der Meinung anderer | Innerhalb eines Meetings zu wenig kritisches Hinterfragen der Meinungen anderer |
| Reflexionseffekt | In Gewinnsituationen Suche nach Sicherheit und in Verlustsituationen Bevorzugung riskanter Alternativen | Präferenz von risikoärmeren Investitionsprojekten in Gewinnsituationen und riskanteren Alternativen in Verlustsituationen, auch dann, wenn dies ökonomisch gesehen irrational ist |
| Repräsentativitätsheuristik | Unterstellung von Trends, Überschätzung kausaler Zusammenhänge und seltener Ereignisse | Fehlbeurteilung der für ein Budget relevanten Wirkungszusammenhänge |
| Streben nach Bestätigung der eigenen Meinung (Confirmation Bias) | Ausblendung oder Unterbewertung von Informationen, die im Widerspruch zu den eigenen Überzeugungen stehen | Unzureichende Risikobeurteilung aufgrund der Vernachlässigung von Informationen, die gegen eine bereits vorhandene Überzeugung sprechen |
| Überoptimismus (Overconfidence) | Überschätzung von Erfolgsaussichten und eigenen Fähigkeiten | Zu niedriger Kostenansatz oder Unterschätzung von Risiken |
| Verankerungsheuristik (Anchoring) | Verzerrende Beeinflussung von Einschätzungen mittels Heranziehung eines Referenzwerts, der als Ausgangspunkt dient | Falsche Beurteilung der für eine Investition notwendigen Finanzmittel aufgrund eines Vergleichs mit einer sehr andersartigen Investition |
| Verfügbarkeitsheuristik | Überbewertung von Informationen, die besonders prägnant sind | Falsche Beurteilung von Risiken bedingt von einem Unglücksfall mit großem Medienecho |

^a Bei einigen Biases ist die englische Bezeichnung geläufiger und wird an den entsprechenden Stellen zusätzlich in Klammern angegeben.

Quelle: In Anlehnung an Drews und Friedrichsen 2012, S. 351; Drews 2013, S. 49.

Weiterhin sind für die vorliegende Arbeit die potenziellen versteckten Wirkungen der sprachlichen Gestaltung von Berichten von zentraler Bedeutung. Im Folgenden soll kurz auf die Studien zur Lesbarkeit speziell im betriebswirtschaftlichen Bereich eingegangen werden.

Es konnte gezeigt werden, dass die Darstellung und insbesondere die Lesbarkeit von Finanzberichten die Informationsaufnahme, -bewertung und -verarbeitung stark beeinflussen können.²⁸⁹ You und Zhang 2009 fokussierten sich in ihrer Untersuchung auf die Form-10-K-Berichte²⁹⁰ und fanden heraus, dass die in den Berichten enthaltenen Informationen von den Investoren desto weniger beachtet werden, je komplexer bzw. länger die Dokumente sind. Dies kann mit der reduzierten Fähigkeit und/oder Bereitschaft der Investoren zur Informationsaufnahme erklärt werden.²⁹¹ Vor allem nicht institutionelle Anleger scheinen von dieser Problematik besonders stark betroffen zu sein. So konnte bspw. Miller 2010 empirisch nachweisen, dass Aktien von Unternehmen, deren Finanzberichte länger und schwieriger zu lesen sind, von kleinen Privatanlegern weniger stark gehandelt werden. Aber auch professionelle Marktakteure können von der Komplexität der Finanzberichte durchaus tangiert werden. Für die Aktienanalysten konnte bspw. empirisch nachgewiesen werden, dass längere und/oder weniger lesbare Finanzberichte zu signifikant höheren Anstrengungsniveaus – definiert als die Zeit, die ein Analyst seit der Veröffentlichung der Berichtsinformationen für die Erstellung der Prognose braucht – und einer höheren Streuung der Gewinnprognosen führen.²⁹²

Die Verhaltensphänomene im Zusammenhang mit der Berichtslänge können gut mit der Informationsüberbelastung (engl. Information Overload) erklärt werden.²⁹³ Ein Überangebot an Daten verlängert die Informationsverarbeitungszeit und trägt dazu bei, dass die Suche und Identifikation von relevanten Informationen erschwert wird oder die Informationen teilweise vernachlässigt werden, was insgesamt der Entscheidungsqualität abträglich ist.²⁹⁴ Hirsch und Volnhals 2012 konnten speziell im Controlling-Kontext die negative Wirkung der Informationsüberlastung auf die Entscheidungsqualität der Berichtsempfänger empirisch nachweisen.

Aber auch wenn Informationen von Berichtsempfängern ohne Probleme verarbeitet werden konnten, können sich aufgrund der Lesbarkeit bestimmte kognitive Effekte er-

²⁸⁹ Vgl. Asay et al. 2017, S. 3f. und die dort zitierte Literatur.

²⁹⁰ Dabei handelt es sich um eine von der SEC standardisierte Form der Jahresberichterstattung, die für Unternehmen ab einer bestimmten Größe verpflichtend ist.

²⁹¹ Vgl. ähnlich Bloomfield 2002, S. 234.

²⁹² Vgl. Lehavay et al. 2011, S. 1106ff.; Loughran und McDonald 2014, S. 1663f.

²⁹³ Im Allgemeinen tritt die Informationsüberlastung ein, wenn die Anforderungen an die Informationsverarbeitung die zur Verfügung stehenden Kapazitäten zur Informationsverarbeitung übersteigen, vgl. Eppler und Mengis 2004, S. 326.

²⁹⁴ Vgl. Eppler und Mengis 2004, S. 333.

geben. So gelang Rennekamp 2012 experimentell der Nachweis, dass Aktienbewertungen tendenziell positiver (negativer) ausfallen, wenn die Unternehmensinformationen in einer gut lesbaren (schlecht lesbaren) Form präsentiert werden. Des Weiteren konnten Tan et al. 2014 empirisch demonstrieren, dass die Sprachwahl einen stärkeren Einfluss auf die Investoren ausübt, wenn die Lesbarkeit niedrig ist. Im Speziellen führte eine beschönigende Berichtssprache in Kombination mit einer niedrigen Lesbarkeit zu höheren Bewertungen seitens der weniger professionellen Investoren. Diese Befunde erklären die Autoren mithilfe der Dual-System-Theorie. Demnach führen die Informationsverarbeitungs-schwierigkeiten aufgrund niedriger Lesbarkeit dazu, dass sich die weniger professionellen Investoren mehr auf heuristische Urteile verlassen.²⁹⁵

2.4.4 Vermeidung von unerwünschten Verhaltenswirkungen

Störungen und unerwünschte Verhaltenswirkungen lassen sich im Management Reporting nicht gänzlich vermeiden. Bei vielen Arten von Biases hat sich nämlich gezeigt, dass das menschliche Entscheidungsverhalten selbst im Falle der Kenntnis über die Existenz möglicher Urteilsverzerrungen fehleranfällig bleibt.²⁹⁶

Das Auftreten bestimmter Biases wird von Faktoren bestimmt, die die Eigenschaften des Entscheiders (Charaktermerkmale, Erfahrungen, Wissensstand usw.), die vorliegende Situation (Ressourcenausstattung, Unternehmensziele, Umweltfaktoren usw.) sowie die Information (Darstellung, Detaillierungsgrad, Umfang usw.) betreffen.²⁹⁷ Dabei können sich diese Faktoren auf der einen Seite in ihrer Wirkung gegenseitig verstärken oder abschwächen. So wird sich ein fundiertes Expertenwissen i. d. R. mildernd auf die negativen Effekte der Informationsüberlastung und des Überoptimismus auswirken. Auf der anderen Seite sind Experten im Allgemeinen anfälliger für die Verfügbarkeitsheuristik und die Kontrollillusion. Den potenziellen Wechselwirkungen aller möglichen Einflussfaktoren sind praktisch keine Grenzen gesetzt und es ist daher unmöglich, ein „optimales Berichtsszenario“ zu definieren, das das Auftreten von unerwünschten versteckten Verhaltenswirkungen komplett verhindert. Aus diesem Grund kann es im Management Reporting nicht darum gehen, alle möglichen Störungen und Verzerrungen vollständig zu eliminieren. Vielmehr sollen etwaige Störungen und Verzerrungen mit individuellen Lösungsansätzen, die aus einer Vielzahl von Einzelmaßnahmen bestehen und jeweils an unterschiedlichen Aspekten der Berichtssituation ansetzen, angegangen werden.²⁹⁸

²⁹⁵ Vgl. Tan et al. 2014, S. 275f.

²⁹⁶ Vgl. Taschner 2013b, S. 212.

²⁹⁷ Vgl. Taschner 2013a, S. 60ff.

²⁹⁸ Vgl. Taschner 2015, S. 43.

Zur Verminderung von Biases kann das Controlling eine Reihe von präventiven Maßnahmen ergreifen oder im Nachhinein versuchen, die bereits aufgetretenen Biases zu erkennen und abzuschwächen. Im Folgenden werden vier ausgewählte Maßnahmen zur Rationalitätssicherung vorgestellt:²⁹⁹

- **Aufbau von Expertise:** Experten verfügen im Bereich ihrer Expertise im Allgemeinen über eine bessere Urteilsfähigkeit und Selbsteinschätzung. Dabei ist zu beachten, dass umfassende Erfahrungen oder bestimmtes Fachwissen allein keinen vollständigen Schutz vor Biases bieten. Vielmehr muss neben einschlägigen Erfahrungen ein konzeptionelles Verständnis der jeweiligen Art der Entscheidungssituation dazukommen. So sollte sich bspw. ein Manager bei der Beurteilung eines Investitionsprojekts nicht unreflektiert von Erfahrungen aus einem vergleichbaren Projekt beeinflussen lassen. Stattdessen sollte er aus einer kritischen Analyse vieler Projekte übergeordnetes Wissen bezüglich der entscheidenden Einflussfaktoren für den Projekterfolg aufbauen.
- **Debiasing:** Damit sind Trainingsmaßnahmen gemeint, mit deren Hilfe die Anfälligkeit für Biases reduziert werden soll.³⁰⁰ Zunächst geht es darum, den Entscheider für die möglichen Urteilsverzerrungen zu sensibilisieren und bei ihm ein Bewusstsein für die Notwendigkeit eines Debiasing-Trainings aufzubauen. Danach können Strategien für einen besseren Umgang mit Biases vermittelt und eingeübt werden. Schließlich müssen die vermittelten Strategien immer wieder aufgefrischt und angewendet werden.
- **Perspektive eines Außenstehenden:** Menschen schätzen oft den gleichen Sachverhalt aus einer Außensicht anders ein als aus einer Innensicht, wobei die Außensicht zumeist realistischere Einschätzungen liefert. So setzen Projektverantwortliche für die Dauer der eigenen Projekte tendenziell zu kurze Zeiträume an, während sie die Dauer ähnlicher Projekte bei anderen Personen besser einschätzen können. Die Einbringung einer Außensicht kann sowohl einen Beitrag zur Prävention von Biases leisten als auch das Erkennen von bereits aufgetretenen Urteilsverzerrungen erleichtern. Mit gezielten Fragen lässt sich eine distanziertere Perspektive auf zugrundeliegende Sachverhalte herstellen. Zur Identifizierung von Einrahmungseffekten bieten sich z. B. die folgenden allgemeinen Fragen an: Fördert die gewählte Darstellungsweise eine bestimmte Beurteilung? Lässt sich die getroffene Entscheidung allein mittels einer anderen Darstellungsweise verändern oder sogar umkehren?

²⁹⁹ Vgl. im Folgenden Drews und Friedrichsen 2012, S. 352f.; Drews 2013, S. 55f.; Taschner 2013b, S. 213f. Eine ausführliche Behandlung dieser und weiterer Maßnahmen mit Verweisen auf relevante empirische Studien finden sich u. a. bei Bazerman und Moore 2013.

³⁰⁰ Als konkrete Techniken könnten in diesem Zusammenhang z. B. die Pre- und Post-Mortem-Analyse oder die Devils's-Advocat-Methode genannt werden, vgl. Schäffer und Weber 2016, S. 10; Scherpereel et al. 2015, S. 36f.

- Entscheidungsunterstützungsmodelle: Hierbei kann es sich sowohl um simple Hilfsmittel wie Checklisten³⁰¹ als auch um komplexe technische Unterstützungssysteme, die den gesamten Informationsprozess abdecken, handeln. Speziell im Controlling kommen die Entscheidungsunterstützungsmodelle bspw. in Form von Investitions- und Kostenrechnungen oder Kennzahlensystemen zum Einsatz.

Um negativen Wirkungen des Cognitive-Load-Effekts oder gar einer möglichen Informationsüberlastung vorzubeugen, muss bei der Berichtsgestaltung auf die Lesbarkeit geachtet werden. In einem Bericht werden die Informationen typischerweise mit Tabellen, Grafiken oder Texten übermittelt,³⁰² daher bezieht sich die Lesbarkeit i. w. S. auf alle diese Präsentationsformate.³⁰³ Nachfolgend werden die Grundzüge dieser drei Grundformen der Informationsdarstellung beleuchtet.

Der Einsatz von *Tabellen* ermöglicht im Allgemeinen eine übersichtliche und gut strukturierte Darstellung einer relativ großen Anzahl von quantitativen Einzelinformationen und ist nicht zuletzt aus diesem Grund in der Praxis sehr beliebt.³⁰⁴ Unter dem Begriff *Tabelle* wird eine nach einem bestimmten Schema geordnete Zusammenstellung von Texten und Daten verstanden.³⁰⁵ In der „klassischen“ zweidimensionalen Tabelle sind die Werte in Zeilen und Spalten angeordnet, wobei nach einem Daten-, Beschriftungs- und Ergebnisbereich differenziert wird. Im Datenbereich werden die Merkmale der Berichtobjekte entweder als Wertgrößen (z. B. Kosten, Umsätze) oder Mengengrößen (z. B. Absatzzahlen, Produktionsmengen) dargestellt. Im Beschriftungsbereich stehen die Bezeichnungen der Berichtobjekte (z. B. Geschäftsbereiche, Produkte) mit ihren jeweiligen Ausprägungen. Schließlich stellt der Ergebnisbereich eine Verdichtung des Datenbereichs in Form von Summen, Mittelwerten o. ä. dar.³⁰⁶

³⁰¹ Zur Anwendung und zu konkreten Ausgestaltungsvorschlägen von Checklisten siehe Drews 2013, S. 56ff.; Kahneman et al. 2011, S. 54ff.

³⁰² Vgl. Ohlert und Weißenberger 2015, S. 57. Wobei es auch kombinierte Darstellungsformen wie z. B. Kennzahlensysteme, Management-Cockpits oder grafische Tabellen gibt. Außerdem werden sowohl Grafiken als auch Tabellen üblicherweise mithilfe textueller Komponenten in Form von Kommentaren, Überschriften oder Legenden ergänzt.

³⁰³ Auch in experimentellen Studien aus dem betriebswirtschaftlichen Bereich werden neben rein textuellen Manipulationen weitere gestalterische Mittel, die das Layout von Tabellen oder Grafiken betreffen, verwendet, siehe u. a. Ohlert und Weißenberger 2015; Rennekamp 2012; Tan et al. 2014. Ein solcher holistischer Ansatz findet sich auch in den von der SEC 1998 formulierten Grundsätzen für gut lesbare Finanzberichterstattung unter dem Titel „A plain English handbook: How to create clear SEC disclosure documents“ wieder.

³⁰⁴ So verglichen Hirsch et al. 2008 die Monatsberichte von sieben Konzernen und fanden heraus, dass der Einsatz von Tabellen, Grafiken und Textelementen zwischen den Unternehmen zwar starken Schwankungen unterlag, die tabellarische Darstellungsform jedoch am meisten verwendet wurde.

³⁰⁵ Vgl. Pollmann und Rühm 2007, S. 77.

³⁰⁶ Vgl. Taschner 2013b, S. 149.

Bei der Gestaltung von Tabellen sollten die einzelnen Tabellenbereiche klar und deutlich voneinander abgegrenzt werden. Dies kann mithilfe von Hervorhebungen (z. B. Fettdruck des Ergebnisbereichs), Farben (z. B. Schattierung des Beschriftungsbereichs) sowie mittels des Einsatzes von Trennlinien erreicht werden.³⁰⁷ Hierbei ist allerdings wichtig, dass diverse Gestaltungsinstrumente mit Bedacht eingesetzt werden und nicht zu einer „Überformatierung“ führen.³⁰⁸ Des Weiteren ist Vorsicht geboten, wenn logisch unterschiedliche Sachverhalte in ein und derselben Tabelle abgebildet werden. Die Unterschiede können dabei in den dargestellten Objektkategorien (z. B. Umsätze pro Monat und pro Quartal), Merkmalskategorien (z. B. Umsätze und Deckungsbeiträge) oder Ausprägungskategorien (z. B. Angabe der Umsätze in Euro und in Prozent) liegen. Um Fehlinterpretationen zu vermeiden, bietet sich bei vielen unterschiedlichen Informationen an, diese Informationen gemäß dem Ähnlichkeitsprinzip in getrennten Tabellen zu präsentieren. Alternativ kommt auch der Einsatz von sog. T-Tabellen oder Pivot-Tabellen infrage.³⁰⁹

Gemäß dem Spruch: „Ein Bild sagt mehr als tausend Worte“ lassen sich die folgenden Vorteile einer *Grafik* zur Vermittlung von Berichtsinformationen aufzählen:³¹⁰

- **Aufmerksamkeitsallokation:** Mit Bildern übermittelte Informationen erregen im Allgemeinen höhere Aufmerksamkeit. Dies führt dazu, dass bei mehreren konkurrierenden Darstellungsformaten die in Grafiken dargestellten Informationen i. d. R. als erste wahrgenommen werden.
- **Geringere Anstrengung:** Die kognitive Aufnahme und Verarbeitung von Bildinformationen ist mit einem geringeren Anstrengungsniveau verbunden. Eine Grafik kann zwar die quantitativen Informationen nicht ersetzen, mit dem Einsatz von Grafiken werden aber sowohl die Aufnahme der Zahlen als auch das Verstehen der inhärenten Datenzusammenhänge erleichtert.³¹¹
- **Höhere Informationsaufnahme:** Im Falle einer unvollständigen Informationsaufnahme, etwa infolge der Informationsüberlastung, fällt der nicht in den Entscheidungsprozess integrierte Informationsanteil bei Grafiken geringer aus als bei den anderen Präsentationsformaten.

³⁰⁷ An dieser Stelle soll angemerkt werden, dass Hervorhebungen stets die Gefahr einer Lenkung und Manipulation des Entscheidungsträgers durch Framingeffekte mit sich bringen, vgl. Drews 2013, S. 50.

³⁰⁸ Zu den visuellen Gestaltungsprinzipien von Tabellen siehe u. a. Few 2012, S. 155ff.

³⁰⁹ Vgl. Taschner 2013b, S. 151ff.

³¹⁰ Vgl. im Folgenden Kroeber-Riel 1987, S. 486.

³¹¹ Vgl. Koch 1994, S. 142ff.

Erleichterte Informationsaufnahme und besseres Verständnis der in visueller Form dargestellten Zusammenhänge können die Problemlösungsfähigkeiten insgesamt verbessern.³¹² Aus vielen empirischen Studien ging hervor, dass sich – unter bestimmten Voraussetzungen – die Entscheidungsqualität verbessert, wenn Informationen in grafischer Form präsentiert werden.³¹³ Speziell im Kontext der Biases konnten bspw. Ohlert und Weißenberger 2015 experimentell demonstrieren, dass der Einsatz von Kreisdiagrammen im Management Report den Basisratenfehler wesentlich reduziert.

Noch mehr als bei der tabellarischen Darstellungsform steht im Zusammenhang mit Grafiken eine Vielzahl von Gestaltungsoptionen zur Auswahl.³¹⁴ Insbesondere die Wahl des passenden Diagrammtyps ist für eine zweckadäquate Vermittlung von Informationen im Bericht von zentraler Bedeutung.³¹⁵ Neben der Wahl des passenden Diagrammtyps muss stets auch eine Reihe von weiteren Entscheidungen getroffen werden. Diese umfassen vor allem die Anzahl und Anordnung der Berichtsobjekte, Beschriftungen, Skalierungen usw.³¹⁶

Grafiken lassen sich besonders gut verhaltenssteuernd einsetzen. Dabei muss neben erwünschten und vom Berichtsersteller intendierten Verhaltenswirkungen immer auch auf mögliche versteckte und nicht intendierte Wirkungen geachtet werden.³¹⁷ In diesem Zusammenhang spielen etwaige Framingeffekte eine wichtige Rolle, die bspw. aufgrund der Auswahl der Ausgangsdaten, Skalierung, Beschriftung oder Formulierung der Schlussfolgerung in Kombination mit einer Grafik auftreten können.³¹⁸

Texte werden in der Praxis meist deutlich weniger eingesetzt als z. B. Tabellen.³¹⁹ Kaum ein Bericht ist jedoch ohne Beschriftungen, Erklärungen, Zusammenfassungen oder Kommentare vorstellbar. In manchen Situationen lässt sich der Informationsbedarf des

³¹² Vgl. Lurie und Mason 2007, S. 160.

³¹³ Siehe u. a. Cardinaels 2008; Hirsch et al. 2015.

³¹⁴ Vgl. Taschner 2013b, S. 154.

³¹⁵ In diesem Zusammenhang unterscheidet Zelazny 2015 fünf mögliche Aussage- bzw. Vergleichsarten (Struktur, Rangfolge, Zeitreihe, Häufigkeit und Korrelation) und schlägt dazu den jeweils passende Diagrammtyp vor, vgl. Zelazny 2015, S. 34ff.

³¹⁶ Vgl. Pollmann und Rühm 2007, S. 113ff.; Waniczek et al. 2018, S. 111ff.

³¹⁷ Siehe Kapitel 2.4.3.

³¹⁸ Vgl. Drews 2013, S. 50ff.

³¹⁹ In der bereits erwähnten Untersuchung von Hirsch et al. 2008 lag der Anteil der Textinformationen zwischen 0 % und 42 % der Berichtsinhalte.

Berichtsempfängers nicht oder nur sehr schwer mit quantitativen Angaben decken. Weber et al. 2005 empfehlen den Einsatz von Texten immer dann, wenn außergewöhnliche und gleichzeitig höchst entscheidungsrelevante Sachverhalte erklärt werden müssen.³²⁰

Weil Textinformationen tendenziell schwieriger aufzunehmen und zu verarbeiten sind als grafische oder tabellarische Informationen, muss bei der Verwendung von textuellen Elementen besonderes Augenmerk auf die Gewährleistung einer hohen Lesbarkeit gelegt werden.³²¹ Dabei setzt eine adäquate Gestaltung bereits bei der Typografie (Schriftart, -größe, -stärke u. ä.) an und reicht bis zur Verwendung eines leserfreundlichen Layouts und Schreibstils.³²²

2.5 Experimentelle Forschung

2.5.1 Einordnung

Empirische Forschung lässt sich im Allgemeinen in einen qualitativen und quantitativen Ansatz unterteilen. *Qualitative Forschung* zeichnet sich dadurch aus, dass Einzelfälle den Gegenstand der Untersuchung bilden. Ziel der qualitativen Forschung besteht in der Generierung neuer Theorien.

Im Gegensatz dazu kommt die *quantitative Forschung* zum Einsatz, wenn bereits einige Erkenntnisse über den Forschungsgegenstand vorliegen und es darum geht, theoretisch fundierte Hypothesen anhand von empirischen Daten und statistischen Verfahren zu validieren.³²³

In der quantitativen Forschung wird zwischen einer korrelativen und experimentellen Strategie unterschieden.³²⁴ Im Rahmen der *korrelativen Strategie* werden Beziehungsmuster zwischen den „im Feld“ beobachteten Variablen untersucht.³²⁵ Mithilfe von statistischen Verfahren wie Korrelations- oder Regressionsanalysen wird anschließend die Stärke des Zusammenhangs zwischen bestimmten Variablen gemessen.³²⁶ Dabei muss

³²⁰ Vgl. Weber et al. 2005, S. 40. Insbesondere die auf strategische Entscheidungen ausgerichteten Berichtsinformationen zeichnen sich dadurch aus, dass sie einmalig, zukunftsorientiert und wenig präzise sind. Solche Informationen werden in den Berichten oftmals mit qualitativen Aussagen übermittelt, vgl. Bloech 1994, S. 197.

³²¹ Vgl. Taschner 2013b, S. 162.

³²² Vgl. Koch 1994, S. 120f.; Wirth 2000, S. 82ff.

³²³ Eine ausführliche Behandlung und Abgrenzung der qualitativen und quantitativen Forschung finden sich u. a. bei Döring und Bortz 2016, S. 3ff.; Rubin und Babbie 2013, S. 39ff.

³²⁴ Vgl. Ray 1992, S. 27ff.

³²⁵ Vgl. Gliner et al. 2009, S. 89; Gravetter und Forzano 2018, S. 296.

³²⁶ Vgl. Gliner et al. 2009, S. 91.

beachtet werden, dass ein hoher Korrelationskoeffizient nicht zwangsläufig eine Ursache-Wirkungs-Beziehung impliziert.³²⁷ Basierend auf signifikanten Ergebnissen aus korrelativen Studien kann allenfalls unterstellt werden, dass die unabhängigen Variablen ursächlich für die Veränderungen der abhängigen Variablen sind. Es kann aber nicht mit letzter Sicherheit behauptet werden, dass tatsächlich ein valider kausaler Zusammenhang besteht.³²⁸

Die Verwendung der korrelativen Strategie erweist sich besonders dann als schwierig, wenn es darum geht, behavioristische Effekte entweder isoliert oder in Verbindung mit anderen Variablen zu bestimmen. Empirische Daten aus dem Feld sind häufig mit ökonomischen Problemen wie mangelnde Beschaffbarkeit oder Messungenauigkeit behaftet. Außerdem können die unabhängigen Variablen aus dem Feld „verschmutzt“ sein, weil die zu untersuchenden Effekte nicht von anderen Effekten getrennt werden können.³²⁹

Als methodische Alternative zur korrelativen Strategie kommt die *experimentelle Strategie* infrage. Mithilfe der experimentellen Strategie wird das Ziel verfolgt, einen kausalen Zusammenhang nachzuweisen. Außerdem können aufgrund der kontrollierten Bedingungen im Rahmen von Experimenten die oben aufgeführten Schwächen der korrelativen Forschung überwunden und damit Antworten auf Fragen geliefert werden, die mit einer anderen Methode nicht zu beantworten wären.³³⁰ Die methodischen Vorteile der experimentellen Forschung führten dazu, dass Experimente eine dominante Stellung in der angewandten verhaltenswissenschaftlichen Forschung einnehmen.³³¹

Im Folgenden wird die Methode der experimentellen Forschung zusammen mit den Anforderungskriterien, die bei der Konzeption und Durchführung von Experimenten wichtig sind, näher vorgestellt.

³²⁷ Vom kausalen Zusammenhang kann nur dann gesprochen werden, wenn neben der Korrelation die Ursache der Wirkung zeitlich vorausgeht und alle alternativen Erklärungen für die Ursache-Wirkungs-Beziehung ausgeschlossen werden können, vgl. Sedlmeier und Renkewitz 2013, S. 123.

³²⁸ Vgl. Ray 1992, S. 34f.

³²⁹ Vgl. Hirsch 2009, S. 169.

³³⁰ Vgl. Sprinkle und Williamson 2007, S. 416.

³³¹ Für einen strukturierten Überblick über experimentelle Studien im Bereich Behavioral Accounting siehe Trotman et al. 2011.

2.5.2 Merkmale und Designarten

Bei einem Experiment werden die unabhängigen Variablen planmäßig variiert, um ihre Effekte auf die abhängigen Variablen kontrolliert zu messen.³³² Hierbei wird das Ziel verfolgt, einen den Hypothesen entsprechenden Ursache-Wirkungs-Zusammenhang nachzuweisen.³³³ Die experimentelle Methode ist von den folgenden Prinzipien gekennzeichnet.³³⁴

- **Manipulation:** Unter Manipulation wird das aktive Eingreifen des Experimentators in die Untersuchungssituation verstanden, wobei die Ausprägungen der unabhängigen Variablen zweckdienlich variiert werden.
- **Messung und Vergleich:** Die Generierung von Daten erfolgt, indem die abhängigen Variablen in jeder Faktorstufe gemessen werden. Weil es sich bei den Faktorstufen i. d. R. um Ausprägungen einer Variablen auf dem nominalen Skalenniveau handelt, erfolgt die statistische Auswertung der Daten anhand von Gruppenvergleichen. Liegt die abhängige Variable in (quasi-)metrischer Form vor, werden – falls bestimmte statistische Voraussetzungen, wie etwa die Normalverteilung und Varianzhomogenität, erfüllt sind – vorzugsweise parametrische Verfahren verwendet. Wenn die abhängige Variable ein nominales Skalenniveau aufweist oder die statistischen Voraussetzungen für einen parametrischen Test nicht erfüllt sind, kommen nicht-parametrische Verfahren zum Einsatz.³³⁵
- **Kontrolle:** Damit von einem kausalen Zusammenhang gesprochen werden kann, muss die Kontrolle über die Störfaktoren gewährleistet sein,³³⁶ die ebenfalls einen systematischen Einfluss auf die abhängigen Variablen ausüben können. Zur Kontrolle von Störfaktoren stehen verschiedene Techniken zur Verfügung.³³⁷ Unerwünschte externe Einflussparameter wie Lärm lassen sich unter speziell eingerichteten Bedingungen im Labor relativ einfach mittels Elimination ausschließen. Ist die vollständige Elimination von störenden Einflüssen nicht möglich, kann mithilfe des Konstanthaltens zumindest versucht werden, die Wirkung der externen Störfaktoren

³³² Die abhängigen Variablen werden auch Zielvariablen genannt. Die unabhängigen Variablen werden als Faktoren bzw. Treatments und ihre Ausprägungen als Faktor- bzw. Treatmentstufen bezeichnet, vgl. Backhaus et al. 2018, S. 165.

³³³ Vgl. Picot 1975, S. 71. Streng genommen bezieht sich die obige Definition auf die Klasse der Prüfexperimente. Im Unterschied dazu werden Erkundungsexperimente durchgeführt, ohne zuvor (präzise) Hypothesen aufzustellen, vgl. Sarris und Reiß 2005, S. 40.

³³⁴ Vgl. im Folgenden Gravetter und Forzano 2018, S. 159.

³³⁵ Eine ausführliche Behandlung von Auswertungsmethoden in der experimentellen Forschung findet sich u. a. bei Koch et al. 2019, S. 179ff.; Sedlmeier und Renkewitz 2013, S. 175ff.

³³⁶ Einen Überblick über mögliche Störfaktoren geben u. a. Schnell et al. 2018, S. 190ff.; Stier 1999, S. 211ff.

³³⁷ Vgl. hier und im Folgenden Schnell et al. 2018, S. 194ff.; Stier 1999, S. 214ff.

über alle Treatments gleich zu halten. Beziehen sich die Störfaktoren auf die Merkmale der Experimententeilnehmer selbst,³³⁸ kommen die statistische Kontrolle, Parallelisierung oder Randomisierung als mögliche Kontrollmechanismen infrage. Bei der statistischen Kontrolle werden die zu kontrollierenden Eigenschaften der Teilnehmer im Rahmen des Experiments erhoben und in der späteren statistischen Analyse – z. B. als Kovariate – berücksichtigt. Das Ziel der Parallelisierung (auch Matching genannt) besteht darin, die Probanden so auf die Experimentgruppen zu verteilen, dass die Treatments hinsichtlich der Störfaktoren im Durchschnitt eine sehr ähnliche Struktur aufweisen und somit überall der gleiche Einfluss dieser Störvariablen unterstellt werden kann. Ein solches Vorgehen erweist sich allerdings aus zwei Gründen als wenig praktikabel: Zum einen kann die Identifikation von allen relevanten Störvariablen kaum sichergestellt werden, zum anderen führt die Berücksichtigung einer großen Anzahl von Störvariablen dazu, dass für diese Variablen eine adäquate Parallelisierung praktisch kaum möglich ist.³³⁹ Schließlich wird bei der Randomisierung basierend auf einem Zufallsprozess jeder Teilnehmer mit der gleichen Wahrscheinlichkeit einer Experimentgruppe zugeordnet, wodurch die systematischen Unterschiede zwischen den einzelnen Treatmentstufen sehr unwahrscheinlich werden. Die Randomisierung bietet sich vor allem dann an, wenn die Stichprobengrößen relativ groß sind. Der wesentliche Vorteil dieser Methode gegenüber der Parallelisierung besteht darin, dass die Störvariablen weder bekannt sein noch gemessen werden müssen.³⁴⁰

Eine wichtige Frage im Zusammenhang mit dem Experimentdesign betrifft die Zuordnung der Probanden zu den einzelnen Treatmentstufen. Dabei existieren die folgenden drei Möglichkeiten für das Design von Experimenten: Between-Subject-Design, Within-Subject-Design sowie gemischte Designs.³⁴¹

- **Between-Subject-Design:** Bei dieser Variante durchläuft jeder Teilnehmer nur eine Faktorstufe, sodass die Daten zwischen den Gruppen unabhängig voneinander erhoben werden. Diese Designart ist vergleichsweise einfach in der Umsetzung und der statistischen Auswertung.³⁴² Weil unterschiedliche Individuen an den einzelnen

³³⁸ Je nach Forschungsfrage können individuelle Charakteristika wie religiöse Einstellung, Vorurteile oder Vermögensausstattung für die Studie relevant sein, vgl. Bernard 2013, S. 92. Solche Variablen werden zusammen mit den soziodemografischen Merkmalen üblicherweise nach dem Experiment im Post-Test-Fragebogen (engl. Post Experimental Questionnaire) abgefragt.

³³⁹ Aus diesem Grund wird bei der Parallelisierung in der Praxis die Auswahl der Störvariablen nur auf diejenigen beschränkt, die am ehesten einen Einfluss auf die abhängigen Variablen ausüben, vgl. Sedlmeier und Renkewitz 2013, S. 133f.

³⁴⁰ Die Randomisierung wird im Allgemeinen als die effektivste Methode zur selektiven Kontrolle bezüglich störender Variablen erachtet, vgl. Schnell et al. 2018, S. 199.

³⁴¹ Eine ausführliche Behandlung von Experimentdesigns entlang von anschaulichen Beispielen erfolgt bei Koch et al. 2019, S. 67ff.

³⁴² Vgl. Keppel und Wickens 2004, S. 11.

Faktorstufen teilnehmen, besteht die Hauptherausforderung des Between-Subject-Designs darin, etwaige konfundierende Effekte aufgrund der individuellen Unterschiede zwischen den Gruppen zu kontrollieren. Mit der Parallelisierung bzw. der Randomisierung kann dieser Problematik begegnet werden.

- **Within-Subject-Design:** Diese Designart ist dadurch gekennzeichnet, dass jede Versuchsperson an allen Faktorstufen teilnimmt. Der Vergleich erfolgt also im Gegensatz zum Between-Subject-Design nicht interindividuell, sondern intraindividuell. Ein wesentlicher Vorteil des Within-Subject-Designs liegt darin, dass keine individuellen Unterschiede zwischen den Gruppen zu befürchten sind. Außerdem ist die Experimentdurchführung im Allgemeinen kostengünstiger, da zumeist weniger Probanden benötigt werden.³⁴³ Gleichzeitig resultiert aus den Messwiederholungen bei ein und derselben Person eine Reihe von designimmanenten Problemen. So kann es aufgrund der Reihenfolgeeffekte dazu kommen, dass sich die Leistung oder das Verhalten von Teilnehmern im Laufe des Experiments verändert, weil z. B. Lern-, Ermüdungs- oder Motivationseffekte auftreten. Während systematischen Effekten mit Kontrolltechniken der Balancierung oder Randomisierung begegnet werden kann,³⁴⁴ lässt sich der Carryover-Effekt im Rahmen des Within-Subject-Designs nicht effektiv umgehen.³⁴⁵
- **Gemischtes Design:** Bei dieser Variante handelt es sich um eine Kombination aus den beiden oben beschriebenen Designs. Das bedeutet, dass für einige unabhängigen Variablen ein und dieselbe Versuchsperson alle Treatmentstufen durchläuft (Within-Subject-Design). Andere unabhängigen Variablen wiederum werden nur zwischen Probandengruppen untersucht (Between-Subject-Design).

Nachdem die wesentlichen Merkmale und Designarten von Experimenten vorgestellt wurden,³⁴⁶ werden im nächsten Abschnitt die Gütekriterien der experimentellen Forschung erläutert.

2.5.3 Gütekriterien

Um die Modellgüte eines Experiments beurteilen zu können, wurde eine Reihe von *Gütekriterien* – auch Validitätskriterien genannt – ausgearbeitet. Nachfolgend werden die

³⁴³ Vgl. Koch et al. 2019, S. 72; Sedlmeier und Renkewitz 2013, S. 147.

³⁴⁴ Für eine ausführliche Behandlung der beiden Methoden siehe Elmes et al. 2006, S. 196ff.

³⁴⁵ Der Carryover-Effekt entsteht, wenn ein Treatment von einem früheren Treatment – z. B. aufgrund der Offenlegung bestimmter Informationen – inhaltlich beeinflusst wird, vgl. Sedlmeier und Renkewitz 2013, S. 158.

³⁴⁶ Für einen praktischen Leitfaden zur Durchführung von Experimenten angefangen bei der Generierung der Forschungsidee über die Entwicklung des Untersuchungsdesigns bis hin zur Datenauswertung sowie Interpretation der Ergebnisse siehe Koch et al. 2019, S. 245ff.; Sarris und Reiß 2005, S. 43ff.

vier wichtigsten Arten der Validität vorgestellt. Hierbei handelt es sich um die statistische Validität, die Konstruktvalidität sowie um die interne und externe Validität.³⁴⁷

Das Kriterium der *statistischen Validität* bezieht sich auf die korrekte Auswahl und Anwendung der statistischen Methoden, mit deren Hilfe ein korrelativer Zusammenhang zwischen den unabhängigen und abhängigen Variablen gemessen werden soll.

Die *Konstruktvalidität* ist ein Maß dafür, inwieweit die Operationalisierung des zu untersuchenden Konstrukts gelungen ist. Mit anderen Worten steht die Frage im Vordergrund, ob die allgemeine Versuchssituation und die Treatments tatsächlich dasjenige theoretische Konstrukt widerspiegeln, das gemäß der Forschungsfrage untersucht werden soll. Sofern ein Experiment mit der gebotenen Sorgfalt konzipiert und durchgeführt wurde, ist grundsätzlich davon auszugehen, dass eine hohe Konstruktvalidität gewährleistet ist.³⁴⁸

Die *interne Validität* liegt vor, wenn die Veränderung der abhängigen Variable eindeutig auf die Variation der unabhängigen Variable zurückgeführt werden kann. Kausalschlüsse lassen sich also nur aus Untersuchungsdesigns mit einer hohen internen Validität ziehen. Eine notwendige Voraussetzung besteht darin, dass alle alternativen Erklärungen für das Zustandekommen der Ursache-Wirkungs-Beziehung ausgeschlossen werden können. Demnach ist die interne Validität umso höher, je größer die Kontrolle über die Störfaktoren implementiert ist. Die experimentelle Forschung weist im Allgemeinen eine hohe interne Validität auf, weil alle Manipulationen im Experiment unter kontrollierten Bedingungen stattfinden.³⁴⁹

Die *externe Validität* bezieht sich auf die Generalisierbarkeit der Untersuchungsergebnisse. Dabei geht es konkret um die Frage, ob die identifizierten kausalen Beziehungen zwischen den Variablen auch bei Veränderung der Personen, des Kontextes oder der Messmethoden weiterhin bestehen bleiben.

An der experimentellen Forschung wird häufig kritisiert, dass lediglich eine künstliche „Laborsituation“ betrachtet wird und somit eine geringe externe Validität vorliegt.³⁵⁰ Die Ursache für diese Kritik sehen Swieringa und Weick 1982 in der Verwechslung von zwei konzeptionell verschiedenen Arten von Realismus: Der weltliche Realismus

³⁴⁷ Vgl. im Folgenden Shadish et al. 2002, S. 33ff. In der Literatur wird im Zusammenhang mit der empirischen Forschung ferner noch das Kriterium der Reliabilität genannt. Hierbei handelt es sich um die Stabilität bzw. Konsistenz der verwendeten Maße und die Replizierbarkeit der Studienergebnisse, vgl. Gravetter und Forzano 2018, S. 61f. Weil die experimentellen Untersuchungen unter kontrollierten Bedingungen stattfinden und die Versuchsabläufe i. d. R. gut dokumentiert werden, sind die Experimente von einer hohen Reliabilität gekennzeichnet.

³⁴⁸ Vgl. Birnberg et al. 1990, S. 40.

³⁴⁹ Vgl. Scandura und Williams 2000, S. 1252.

³⁵⁰ Sprinkle und Williamson 2007 bezeichnen die externe Validität sogar als „Achilles heel of experimentation“.

liegt vor, wenn die Umgebung und Abläufe in der Laborsituation so weit wie möglich denen in der natürlichen Umwelt entsprechen. Im Gegensatz dazu bezieht sich der experimentelle Realismus auf die Frage, ob die Untersuchungssituation von den Teilnehmern als glaubwürdig und ernst wahrgenommen wird.

Für bestimmte Probleme ist der weltliche Realismus im hohen Maße angebracht. Dies ist besonders dann der Fall, wenn komplexe Theorien und situative Bedeutungen im Fokus der Untersuchung stehen.³⁵¹ Vor dem Hintergrund der Zielsetzung experimenteller Forschung, einen kausalen Zusammenhang nachzuweisen, muss jedoch der weltliche Realismus kritisch betrachtet werden. Der weltliche Realismus kann bspw. die Individuen dazu verleiten, in „überlernte“ Routinehandlungen zu verfallen, die weniger sensibel für Manipulationen im Experiment sind. Des Weiteren können „realistischere“ Versuchsbedingungen diverse Signale beinhalten, die den Probanden wichtige vergangene Ereignisse in Erinnerung rufen. Diese Ereignisse aus der Vergangenheit können sich bei den einzelnen Probanden unterscheiden, was zur Folge hätte, dass bestimmten Situationen unterschiedliche Bedeutungen beigemessen werden und ein überbestimmtes Verhalten ausgelöst wird.³⁵² Solche Faktoren erhöhen die unerklärte Varianz im Experiment und laufen dem Ziel der experimentellen Kontrolle zuwider. Insgesamt kann also der weltliche Realismus wenig dazu dienen, Ursache-Wirkungs-Zusammenhänge nachzuweisen.³⁵³

Des Weiteren wird an der experimentellen Forschung bisweilen bemängelt, dass die Zusammensetzung des Teilnehmerpools und die Höhe der Experimentvergütung eine Übertragung der Ergebnisse auf die Realität nicht erlauben.³⁵⁴ Weil die Experimentteilnehmer häufig aus einem studentischen Personenkreis rekrutiert werden, ließen sich daraus keine Schlüsse auf Entscheidungen der Zielgruppe ziehen, lautet die Kritik. Empirische Untersuchungen zeigten jedoch, dass das Verhalten von Berufserfahrenen durchaus aus dem Verhalten von Studenten extrapoliert werden kann.³⁵⁵ Darüber hinaus weichen die studentischen Probanden von den normativen Rationalitätspostulaten tendenziell weniger stark ab als die übrige Population.³⁵⁶ Im Hinblick auf die Höhe der Teilnehmervergütung in Experimenten wird häufig moniert, dass diese verhältnismäßig zu gering sei, um den Gegebenheiten in der realen Welt zu entsprechen und aus diesem Grund eine höhere Vergütung andere Entscheidungen nach sich zöge. Diesem Einwand

³⁵¹ Vgl. Swieringa und Weick 1982, S. 85f.

³⁵² Vgl. Swieringa und Weick 1982, S. 80ff.

³⁵³ Vgl. Berkowitz und Donnerstein 1982, S. 247.

³⁵⁴ Vgl. Barr und Hitt 1986, S. 599ff.; Birnberg und Nath 1968, S. 40; Frey 1990, S. 170.

³⁵⁵ Vgl. u. a. Ball und Cech 1996, S. 257. Dies ist insbesondere dann der Fall, wenn Biases und intuitive Entscheidungen im Fokus der Untersuchung stehen, vgl. z. B. Shafir und LeBoeuf 2002, S. 505 und die dort zitierte Literatur.

³⁵⁶ Siehe u. a. Duch et al. 2017.

kann ebenfalls mit Verweis auf die einschlägigen experimentellen Studien widersprochen werden. Besonders dann, wenn es um die Untersuchung von Verhaltensanomalien geht, übt die Vergütungshöhe im Allgemeinen relativ wenig Einfluss auf die getroffenen Entscheidungen aus.³⁵⁷

Resümierend lässt sich festhalten, dass der experimentelle Forschungsansatz besonders dann geeignet ist, wenn es darum geht, einen Ursache-Wirkungs-Zusammenhang nachzuweisen. Der Hauptvorteil von Experimenten gegenüber anderen empirischen Methoden besteht darin, dass der Einfluss von gezielt manipulierten unabhängigen Variablen auf die abhängigen Variablen unter kontrollierten Bedingungen gemessen werden kann. Somit verfügen Experimente im Allgemeinen über einen hohen Grad der internen Validität. Da jedoch eine hohe interne Validität grundsätzlich zulasten der externen Validität geht – zumindest im Sinne des weltlichen Realismus –,³⁵⁸ sind andere Methoden der Feldforschung dann besser geeignet, wenn für die Beantwortung der Forschungsfrage eine Simulation realer Umweltbedingungen notwendig ist.

³⁵⁷ Camerer und Hogarth 1999 analysierten 74 Experimente mit unterschiedlich hohen finanziellen Anreizen und kamen zu dem Schluss, dass die Verletzungen der Rationalitätspostulate nicht verschwinden, wenn lediglich der finanzielle Anreiz erhöht wird. Im besten Fall kann eine höhere Vergütung dazu führen, dass die Verhaltensanomalien weniger häufig vorkommen und die Gesamtvarianz sinkt, vgl. Camerer und Hogarth 1999, S. 7; Frey 1990, S. 171.

³⁵⁸ Vgl. Koch et al. 2019, S. 52; Scandura und Williams 2000, S. 1252.

3. Experiment 1: Risikoverhalten³⁵⁹

3.1 Einleitung

Risikoentscheidungen von Managern haben oft weitreichende Konsequenzen und sind wichtig für den langfristigen Erfolg eines Unternehmens.³⁶⁰ Es wird angenommen, dass Manager Investitionsprojekte mit weniger volatilen Cashflows präferieren – auch dann, wenn diese geringere Nettokapitalwerte im Vergleich zu riskanteren Investitionsprojekten aufweisen –, um sich gegen das Downside-Risiko und einen möglichen Arbeitsplatzverlust infolgedessen abzusichern. Zugleich haben Unternehmenseigentümer die Möglichkeit, ihren Kapitaleinsatz breit zu diversifizieren, so dass diversifizierbare Risiken für sie keine Rolle spielen, wohingegen Manager einen Großteil ihres Einkommens aus dem Arbeitsverhältnis mit nur einem Unternehmen beziehen. Daher werden die Manager nicht bereit sein, das aus Sicht der Eigentümer optimale Maß an Risiken einzugehen.³⁶¹ Dieses Problem kann bspw. mittels aktienbasierter Managementvergütung zum gewissen Teil behoben werden.³⁶² In der vorliegenden Arbeit wird auf diese Problematik aus der behavioristischen Perspektive geschaut, indem sprachbezogene Phänomene untersucht werden.

Gemäß den Dual-System-Theorien (DST) führt ein niedriges Lesbarkeitsniveau zur stärkeren Beanspruchung von System 2 infolge von intensiveren Arbeitsgedächtnisaktivitäten und höherer kognitiver Belastung. Unter kognitiver Belastung verhalten sich die Menschen tendenziell risikoaverser. Dies kann damit erklärt werden, dass unter kognitiver Belastung nicht ausreichend mentale Ressourcen übrigbleiben, um Entscheidungen, die vom affektiven und für Biases anfälligen System 1 generiert werden, zu kontrollieren und zu korrigieren.³⁶³

Der Foreign-Language-Effekt (FLE) steht für den positiven Einfluss einer Fremdsprache auf die Förderung rationaler Entscheidungen. Zahlreiche kognitiv-psychologische Studien zeigten, dass Menschen sich deutlich weniger verlust- bzw. risikoavers verhalten, wenn sie die Informationen in einer fremden Sprache verarbeiten müssen.³⁶⁴

Die vorliegende Studie wurde konzipiert, um den negativen Effekt der kognitiven Belastung (infolge niedriger Lesbarkeit) und den positiven Effekt einer Fremdsprache auf die Bereitschaft, „smarte“ – d. h. aus normativer Sicht vorteilhafte – Risiken einzugehen, im Controlling-Kontext zu untersuchen. Weil einige Probanden der zusätzlichen kognitiven Belastung infolge niedriger Lesbarkeit ausgesetzt wurden, kann erwartet

³⁵⁹ Dieses Kapitel basiert auf der Veröffentlichung von Besuglov und Crasselt 2021.

³⁶⁰ Vgl. Hoskisson et al. 2016, S. 137.

³⁶¹ Vgl. Hoskisson et al. 2016, S. 140; Low 2009, S. 472.

³⁶² Siehe u. a. Jensen und Meckling 1976.

³⁶³ Ausführlich zu den DST und kognitiven Effekten der Lesbarkeit siehe Kapitel 2.2 bzw. Kapitel 2.3.1.

³⁶⁴ Ausführlich zum FLE und den Erklärungsansätzen für dieses Phänomen siehe Kapitel 2.3.2.2.

werden, dass der FLE nur bei einem hohen Lesbarkeitsniveau zu beobachten sein wird. Darüber hinaus wird im Rahmen dieser Studie auf die Wirkmechanismen hinter dem FLE geschaut, indem die Rolle der Sprachkompetenz und die in der Literatur am meisten diskutierten Erklärungsansätze für den FLE untersucht werden.

Die experimentelle Untersuchung ist als ein 2×2-Between-Subject-Design mit den wichtigsten unabhängigen Variablen: Sprache (Fremd- vs. Mutter-) und Lesbarkeit (hoch vs. niedrig) aufgebaut. Die Experimentteilnehmer nahmen dabei die Rolle eines Managers ein und erhielten entweder in ihrer Muttersprache (Deutsch) oder Fremdsprache (Englisch) einen fiktiven Managementbericht. Der Bericht umfasste zwei alternative Kaufangebote für die Produktion des Unternehmens. Ein Angebot war eine sichere Option und das andere Angebot stellte eine risikobehaftete, aber hinsichtlich des Erwartungswerts der Erlöse vorteilhaftere Option dar. Die Lesbarkeit wurde mithilfe ausgewählter linguistischer und formaler Kriterien gemäß dem „Plain English Handbook“ der US-amerikanischen Securities Exchange Commission (SEC 1998) manipuliert. Die Probanden wurden gebeten, sich zunächst zwischen den beiden Kaufangeboten zu entscheiden und danach ihre individuellen Sicherheitsäquivalente für die risikobehaftete Alternative anzugeben.

Die Experimentergebnisse ergaben einen hochsignifikanten und robusten Effekt der Lesbarkeit. In Übereinstimmung mit den Erwartungen waren die Teilnehmer unter niedriger Lesbarkeit deutlich weniger bereit, „smarte“ Risiken zu akzeptieren, als ihre Pendants unter hoher Lesbarkeit. Außerdem konnte infolge des Einsatzes einer Fremdsprache in Kombination mit hoher Lesbarkeit ein positiver Effekt auf die Bereitschaft, „smarte“ Risiken einzugehen, festgestellt werden. Dieser Effekt war jedoch nur schwach signifikant und nicht besonders robust. Ergänzende Analysen zeigten, dass die Teilnehmer mit einem mittelhohen Sprachkompetenzniveau und unter der Bedingung einer hohen Lesbarkeit im Vergleich zu den anderen Sprachkompetenzniveaus wesentlich häufiger bereit waren, „smarte“ Risiken einzugehen. Des Weiteren wiesen diese Probanden weniger stark ausgeprägte positive Gefühle gegenüber der Option mit dem sicheren Erlös auf. Somit konnte in der vorliegenden Untersuchung derjenige Erklärungsansatz für den FLE unterstützt werden, der auf der Reduktion der emotionalen Resonanz basiert.

Der Rest dieses Kapitels ist wie folgt aufgebaut: In Kapitel 3.2 werden die Hypothesen abgeleitet. Das Experimentdesign sowie die Teilnehmer werden in Kapitel 3.3 und die Experimentergebnisse in Kapitel 3.4 näher beleuchtet. Schließlich erfolgt die Diskussion der Ergebnisse in Kapitel 3.5.

3.2 Hypothesen

3.2.1 Hypothese 1

Bei der Hypothese 1 stehen die Auswirkungen von unterschiedlichen Lesbarkeitsstufen auf die Bereitschaft der Entscheidungsträger, „smarte“ Risiken zu akzeptieren, im Zentrum der Betrachtung.

Wie bereits in Kapitel 2.3.1 erläutert, kehrt der positive Disfluency-Effekt (DFE) ab einem bestimmten Grad der reduzierten Lesbarkeit um und es kommt zum negativen Cognitive-Load-Effekt (CLE). Zu viel an extrinsischer kognitiver Belastung infolge einer zu niedrigen Lesbarkeitsstufe führt zu einer schlechteren kognitiven Performance, weil nebensächliche Aktivitäten (wie z. B. semantische Dekodierung) die kognitiven Kapazitäten von der eigentlichen Entscheidungsaufgabe ablenken.

Weniger lesbare Texte beanspruchen stärker das limitierte Arbeitsgedächtnis aufgrund höherer extrinsischer kognitiver Belastung. Da die Arbeitsgedächtnisprozesse entscheidend für einen effektiven Einsatz von System 2 sind,³⁶⁵ sollten sich niedrige Lesbarkeitsstufen negativ auf den analytischen Denkmodus auswirken. Beispielsweise kann insbesondere bei weniger erfahrenen Investoren bei höheren Informationsverarbeitungsschwierigkeiten aufgrund niedriger Lesbarkeit die Entscheidungsqualität beeinträchtigt werden. Der Grund dafür ist, dass sich die Investoren in einem solchen Fall anstatt auf das analytische System 2 auf ihr heuristisches System 1 verlassen müssen und so die finanzielle Berichterstattung weniger effektiv verarbeiten.³⁶⁶

In Bezug auf das Risikoverhalten existiert überzeugende empirische Evidenz dafür, dass die Beanspruchung des Arbeitsgedächtnisses einen höheren Grad der Risikoaversion induziert. So zeigten Deck und Jahedi 2015 mithilfe einer ablenkenden Erinnerungsaufgabe, dass die Beanspruchung der Arbeitsgedächtniskapazitäten zum deutlich stärker ausgeprägten risikoaversen Verhalten führt. Ein ähnliches Ergebnis ist zu erwarten, wenn die Arbeitsgedächtnisressourcen aufgrund von extrinsischer kognitiver Belastung infolge niedriger Lesbarkeit okkupiert werden, weil so eine Entscheidungsfindung im analytischen System-2-Denkmodus verhindert wird. In Übereinstimmung mit dieser Argumentation lautet die erste Hypothese wie folgt:

H1: Entscheidungsträger werden weniger bereit sein, „smarte“ Risiken zu akzeptieren, wenn sie einen Controlling-Bericht mit einem niedrigen im Vergleich zu einem hohen Lesbarkeitsniveau verarbeiten müssen.

³⁶⁵ Vgl. Whitney et al. 2008, S. 1180.

³⁶⁶ Vgl. Tan et al. 2014, S. 279f.

3.2.2 Hypothese 2

Bei der zweiten Hypothese geht es um die Frage, welchen Effekt ein Controlling-Bericht entweder in der Muttersprache oder einer Fremdsprache des Entscheidungsträgers auf seine Bereitschaft, „smarte“ Risiken einzugehen, ausübt. Im Folgenden wird der Einfluss der Sprachwahl auf System 1 und System 2 näher betrachtet und es wird begründet, warum der FLE im Kontext der Übernahme von vorteilhaften Risiken abhängig von einer hohen Lesbarkeitsstufe sein sollte.

Es existieren überzeugende empirische Belege dafür, dass Menschen weniger emotional reagieren, wenn sie Informationen in einer Fremdsprache im Vergleich zu ihrer Muttersprache verarbeiten. Gemäß der „Risk-as-Feeling“-Hypothese³⁶⁷ lösen Risikosituationen negative Gefühle aus und fallen daher in die Einflussosphäre des FLE. Darüber hinaus können Risiken auch dann ein gewisses Niveau an Emotionen hervorrufen, wenn keine potenziellen Verluste involviert sind. Diese Auffassung wird von Costa et al. 2014a unterstützt, die mithilfe des Holt-Laury-Tests – bei dem alle Lotteriepaaare einen positiven Erwartungswert aufweisen – zeigten, dass Individuen, die den Test in einer fremden Sprache absolvierten, weniger risikoavers waren als ihre Pendants in der muttersprachlichen Gruppe. Die Autoren schlussfolgern: „Indeed, one could argue that the poor choices, in terms of expected value, prompted by risk aversion stem from the emotional reaction to risk itself.“³⁶⁸ Zusammengefasst fördert also der Einsatz einer Fremdsprache stärker das analytische System 2 relativ zum affektiven System 1, was zur höheren Bereitschaft, vorteilhafte Risiken einzugehen, führen soll.

In Übereinstimmung mit dem Erklärungsansatz für den FLE basierend auf der psychologischen Distanz ist die Risikowahrnehmung davon abhängig, ob die Individuen eher auf einer hohen oder niedrigen Abstraktionsstufe denken.³⁶⁹ Der Einsatz einer fremden Sprache könnte im Vergleich zur eigenen Muttersprache einen höheren mentalen Abstraktionsgrad bedingen, sodass der gewünschte Endzustand mehr und die Schwierigkeiten auf dem Weg dorthin weniger in den Blickwinkel rücken.³⁷⁰ Dies sollte einen positiven Effekt auf die Bereitschaft zur Akzeptanz von vorteilhaften Risiken ausüben.

Schließlich können gemäß dem DFE die Schwierigkeiten bei der Informationsverarbeitung als metakognitives Signal dazu dienen, das Entscheidungstempo zu reduzieren und den analytischen Denkmodus zu aktivieren. Weil die Informationsverarbeitung in einer Fremdsprache kognitiv anspruchsvoller ist als in der eigenen Muttersprache, könnte der Einsatz einer fremden Sprache System 2 aktivieren, was wiederum zur höheren Akzeptanz von vorteilhaften Risiken führen sollte. Der DFE erfordert sprachlich angereicherte Texte, die eine gewisse (nicht kritische) kognitive Belastung darstellen. Da typische

³⁶⁷ Siehe Loewenstein et al. 2001.

³⁶⁸ Costa et al. 2014a, S. 245

³⁶⁹ Vgl. Sagristano et al. 2002, S. 364ff.

³⁷⁰ Vgl. Hayakawa et al. 2016, S. 792f.

Controlling-Berichte sowohl Diagramme, Tabellen als auch textuelle Komponenten enthalten, kann davon ausgegangen werden, dass der FLE auch bei einem Controlling-Bericht auftreten kann.

Es konnte empirisch demonstriert werden, dass die sprachliche Verarbeitung in einer Fremdsprache und eine weitere parallel dazu auszuführende Denkaufgabe miteinander in Konflikt stehen. Weil linguistische Verarbeitung eine notwendige Voraussetzung für jeden daran anschließenden Denkprozess bildet, wird zuerst die konkurrierende Denkaufgabe „geopfert“, wenn es nicht ausreichend kognitive Ressourcen gibt, um beide Aufgaben gleichzeitig auszuführen.³⁷¹ Bei einem niedrigen Lesbarkeitsniveau werden kognitive Ressourcen, die für einen effektiven Gebrauch von System 2 notwendig sind, von extrinsischer kognitiver Belastung beansprucht. Dagegen besteht bei hoher Lesbarkeit die Möglichkeit, dass sich der FLE aufgrund der gestärkten Rolle von System 2 im Entscheidungsfindungsprozess entfalten kann.

Empirische Belege dafür, dass der FLE von hoher Lesbarkeit abhängig ist, können den existierenden psychologischen Studien entnommen werden. In solchen Studien werden die Aufgaben entweder in wenigen Worten oder in überschaubaren einfachen Sätzen formuliert. Oft wird das Asian-Disease-Problem oder eine Abwandlung davon verwendet.³⁷² Die ursprüngliche Version des Asian-Disease-Problems besteht im Gewinnframing aus lediglich sechs Sätzen mit 88 Wörtern und weist gemäß dem Gunning/Fog-Index (kurz Fog-Index) einen Lesbarkeitswert von 9,5 auf.³⁷³ Insgesamt lautet Hypothese 2 wie folgt (siehe auch Abbildung 13):

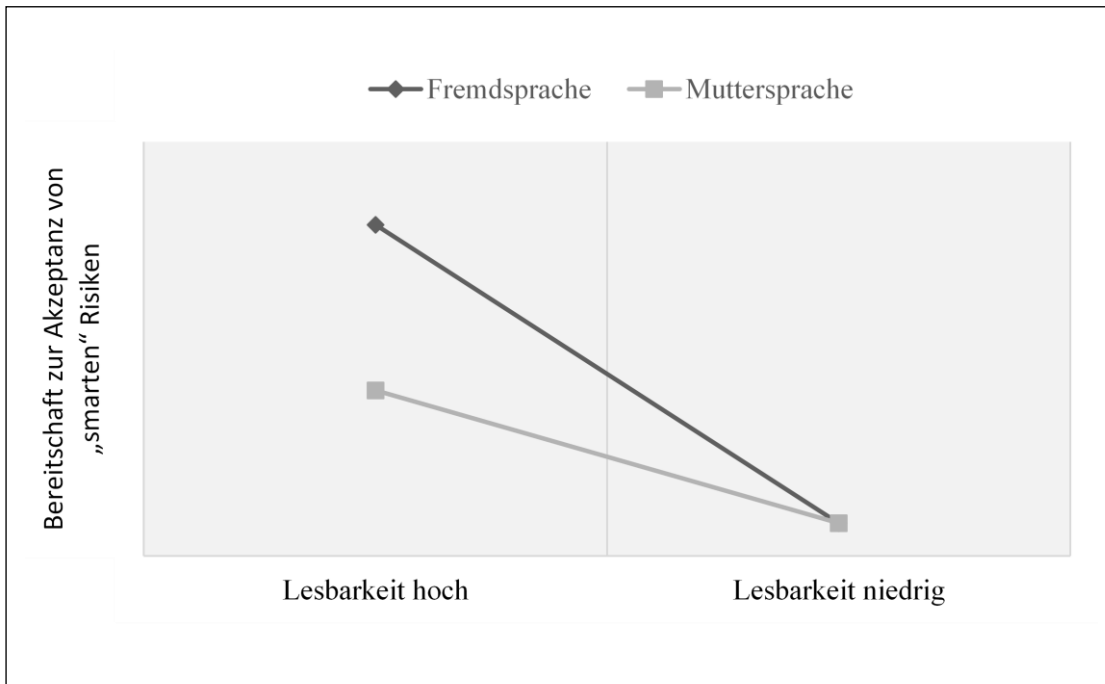
H2: Nur bei einem hohen Lesbarkeitsniveau werden die Entscheidungsträger eher bereit sein, „smarte“ Risiken zu akzeptieren, wenn sie einen Controlling-Bericht in einer Fremdsprache, statt in ihrer Muttersprache verarbeiten müssen.

³⁷¹ Vgl. Takano und Noda 1993, S. 446.

³⁷² Siehe stellvertretend für diese Studien Costa et al. 2014a oder Keysar et al. 2012.

³⁷³ Dieser Wert bedeutet, dass der Text selbst von einem Kind verstanden werden kann, vgl. Li 2008, S. 225. Für das Verlustframing gelten annähernd die gleichen Werte.

Abbildung 13: Visualisierung von H1 und H2 (Experiment 1)



Quelle: Eigene Darstellung.

3.2.3 Hypothese 3

Der Einfluss der Sprachwahl auf das Risikoverhalten sollte von der Sprachkompetenz abhängig sein.³⁷⁴ Einerseits ist es plausibel anzunehmen, dass ein annähernd muttersprachliches Kompetenzniveau nur einen minimalen Effekt auf das Risikoverhalten ausüben wird, weil es – wenn überhaupt – nur marginale Unterschiede in der Art der Informationsverarbeitung relativ zur eigenen Muttersprache geben wird. Andererseits werden beim geringen Sprachkompetenzniveau die kognitiven Ressourcen aufgrund der Schwierigkeiten bei der Informationsverarbeitung stärker beansprucht und dies wird einen negativen Einfluss auf System 2 ausüben. Der FLE sollte also am wahrscheinlichsten bei einem mittelhohen Sprachkompetenzniveau zu beobachten sein.³⁷⁵ Gemäß dieser Argumentation kann Folgendes prognostiziert werden:

H3: Der FLE – wie unter H2 formuliert – gilt nur für mittelhohe Sprachkompetenzstufen.

³⁷⁴ Vgl. Hayakawa et al. 2016, S. 793.

³⁷⁵ Vgl. Costa et al. 2017, S. 148.

3.3 Experimentdesign und Teilnehmer

3.3.1 Experimentdesign

3.3.1.1 Ablauf und Aufgabe

Bei dem Experiment handelt es sich um ein 2×2-Between-Subject-Design mit einer randomisierten Zuweisung von Teilnehmern in die verschiedenen Treatment-Gruppen. Das Experiment wurde zeitgleich in vier separaten Räumen im Sommersemester 2018 durchgeführt. Nach der Ankunft im jeweiligen Raum erhielt jeder Teilnehmer eine kurze schriftliche und mündliche Anweisung in der jeweiligen Sprache der Experimentgruppe. Um den Language-Switching-Effekt³⁷⁶ zu vermeiden, wurden die Teilnehmer in den fremdsprachlichen Gruppen gebeten, sich bei Bedarf ausschließlich auf Englisch zu verständigen.

Die Teilnehmer erhielten im ersten Teil des Experiments eine Fallstudie und einen Fragebogen in ausgedruckter Form. Nachdem Teil eins von allen Teilnehmern bearbeitet worden war, wurde dieser eingesammelt und Teil zwei wurde ausgeteilt. Der zweite Teil bestand aus Verständnisfragen in Bezug auf die Fallstudie und demografischen bzw. sprachbezogenen Abfragen.³⁷⁷ Schließlich mussten die Teilnehmer in den fremdsprachlichen Gruppen einen kurzen Englischtest³⁷⁸ innerhalb einer Zeitfrist von 15 Minuten absolvieren.

Im Rahmen der Experimentaufgabe erhielten die Teilnehmer eine Fallstudie, die aus einem Controlling-Bericht eines fiktiven Unternehmens bestand und worin zwei alternative Kaufangebote für die Produktion des Unternehmens präsentiert wurden. Das eine Kaufangebot stellte eine sichere Option mit einem festen Erlös dar und das andere Kaufangebot repräsentierte eine risikobehaftete Option mit einem höheren Erwartungswert der Erlöse im Vergleich zur sicheren Option.³⁷⁹

Die Probanden im Experiment mussten die Rolle des Managers einnehmen und sich zwischen den beiden Kaufangeboten entscheiden. Den Teilnehmern wurde mitgeteilt,

³⁷⁶ Siehe Oganian et al. 2016.

³⁷⁷ Sprachbezogene Informationen wurden in Anlehnung an den „language history questionnaire“ von Li et al. 2006 erhoben.

³⁷⁸ Siehe Kapitel 3.3.1.2 und Anhang I.

³⁷⁹ Das risikobehaftete Kaufangebot implizierte zwei mögliche Erlöse – 200.000 \$ bzw. 100.000 \$ – deren Eintrittswahrscheinlichkeiten mit 50 % gleich waren. Die Selbstkosten der Produktion betragen 100.000 \$ und damit belief sich der Erwartungswert des Gewinns auf 50.000 \$. Die anreizkompatible Höhe des Erlöses für das sichere Kaufangebot wurde anhand der Prospect-Theorie (siehe Kapitel 2.1.2) bestimmt, wobei für die Berechnungen die Wert- bzw. Entscheidungsgewichtungsfunktion gemäß Tversky und Kahneman 1992 zugrunde gelegt wurde. Daraus resultierte ein Erlös i. H. v. 138.000 \$ für das sichere Kaufangebot. Um jegliche Situation der Unsicherheit auszuschließen, wurden die Teilnehmer darüber informiert, dass es sich bei den Unternehmen, die die Kaufpreise anbieten, um langfristige Kunden handelt, die ihre Rechnungen immer fristgerecht bezahlen.

dass ihre Entscheidung Auswirkungen auf die finanzielle Performance des Unternehmens haben wird und dass ihre angenommene Managervergütung zum Teil von dieser Performance abhängig ist. Dieses Setting wurde gewählt, um die Entscheidung in einen Kontext zu bringen, der realen Entscheidungssituation in den Unternehmen ähnelt. Manager agieren als Adressaten und Benutzer von Controlling-Berichten üblicherweise im Namen von Eigentümern des Unternehmens und ihre Vergütung beinhaltet variable Bestandteile, die auf der Performance des Unternehmens basieren, um die Prinzipal-Agent-Problematik abzumildern.³⁸⁰ Da die Performance als Bemessungsgrundlage für die variable Managementvergütung auf dem aggregierten Ergebnis basiert – z. B. der Aktienkursanstieg oder Unternehmensgewinn – ist jedoch die Verbindung zwischen jeder einzelnen Entscheidung und der tatsächlich erhaltenen variablen Vergütung weit weniger ausgeprägt, als dies bspw. bei einem Verkäufer mit einem an die Verkäufe gebundenen Bonus der Fall wäre. Um die Verbindung zwischen der Entscheidung und den persönlichen finanziellen Konsequenzen nicht überzustrapazieren, wurde auf die tatsächliche Teilnehmervergütung im Experiment verzichtet.

Die Entscheidung, die Teilnehmer nicht zu vergüten, und die Art des Framings der Entscheidungssituation sind insofern von großer Bedeutung, als empirisch festgestellt werden konnte, dass der Grad der Verantwortung für die Entscheidung und das daraus resultierende Ergebnis einen Einfluss auf die Bereitschaft des Entscheidungsträgers ausübt, Risiken einzugehen.³⁸¹ Je nach Kontext sollte ein höherer Grad der Verantwortung risikoaverses Verhalten verstärken, ähnlich den Situationen, in denen die Individuen für sich selbst Entscheidungen treffen.³⁸² Es kann einerseits erwartet werden, dass sich das Gesamtniveau der Risikoaversion erhöhen wird, falls die Teilnehmer eine Experimentvergütung erhalten, weil so die Konsequenzen für sich selbst betont werden (ähnlich wie bei Lotterien mit kleinen Auszahlungsbeträgen). Andererseits sollte die Risikoaversion sinken, wenn die Experimentvergütung nicht vorgesehen ist, weil auf diese Weise die Aufmerksamkeit stärker auf den Verantwortungsaspekt gegenüber dem Eigentümer gelenkt wird. Mit der Entscheidung, die Experimentvergütung nur hypothetisch jedoch nicht tatsächlich zu berücksichtigen, wurde das Ziel verfolgt, eine betriebswirtschaftliche Situation nachzubilden, die den tatsächlich vorzufindenden Organisationsstrukturen in der Praxis möglichst weit entspricht. Die Auswirkung einer Variation des Verantwortungsaspekts in Verbindung mit den Einflüssen der Lesbarkeit und des FLE würden den Rahmen dieser Studie übersteigen.

³⁸⁰ Vgl. Hoskisson et al. 2016, S. 140.

³⁸¹ Siehe Losecaat Vermeer et al. 2020.

³⁸² Vgl. Pollmann et al. 2014, S. 389.

3.3.1.2 Unabhängige Variablen und Kontrollvariablen

Die erste unabhängige Variable ist die Sprache – Mutter- vs. Fremdsprache. Für die muttersprachlichen Gruppen wurden die Experimentinstruktion sowie die Fallstudie auf Deutsch verfasst und für die fremdsprachlichen Gruppen wurden die entsprechenden Unterlagen in englischer Sprache zur Verfügung gestellt. Um sicherzustellen, dass die in beiden Sprachbedingungen vermittelten Inhalte bedeutungsgleich sind, wurden die ursprünglich auf Deutsch geschriebenen Texte von zwei unabhängigen bilingualen Sprechern ins Englische übersetzt und anschließend ins Deutsche rückübersetzt.³⁸³ Der Vergleich der Originaltexte mit den rückübersetzten Texten ergab keine wesentlichen Abweichungen.

Die Sprachkompetenz als weitere Variable zur Messung des FLE wurde mithilfe des C-Tests erhoben.³⁸⁴ Im Kontext eines C-Tests ist es wichtig zu beachten, dass die individuellen Ergebnisse relativ zur Leistung der gesamten Testgruppe betrachtet werden müssen und die alleinige Interpretation der erreichten absoluten Werte für die Sprachkompetenz wenig aussagekräftig ist.³⁸⁵ Diesem Grundsatz folgend wurden die Teilnehmer auf die drei folgenden Sprachkompetenzgruppen aufgeteilt: (1) alle Probanden in den muttersprachlichen Gruppen, die sich als deutschsprachige Muttersprachler (oder vergleichbare Sprecher)³⁸⁶ identifiziert haben, wurden als Muttersprachler klassifiziert; (2) die besten 25 % der Probanden im C-Test wurden dem mittelhohen Sprachkompetenzniveau zugeordnet³⁸⁷ und (3) alle übrigen Teilnehmer wurden in die Gruppe mit dem geringen Sprachkompetenzniveau eingeordnet. Probanden in der Gruppe mit dem geringen Sprachkompetenzniveau erzielten durchschnittlich 56 % der Punkte im C-Test und diejenigen, die in die Gruppe mit dem mittelhohen Sprachkompetenzniveau eingruppiert wurden, erreichten ein Ergebnis von 78 %. Der Unterschied zwischen den beiden Gruppen ist höchst signifikant ($p < 0,010$). Da die drei Texte aus englischsprachigen Lehrbüchern für das mittelhohe Sprachkompetenzniveau entnommen wurden,

³⁸³ Entsprechend dem Verfahren von Brislin 1970.

³⁸⁴ Allgemein zum C-Test siehe Raatz und Klein-Braley 2002. Ein konkretes Beispiel zum hier verwendeten C-Test befindet sich in Anhang I. Um die Modellgüte des verwendeten C-Tests zu evaluieren, wurde eine Itemanalyse gemäß Grotjahn 2002 durchgeführt. Für alle drei Texte sind die Itemschwierigkeit und die Trennschärfe ausreichend hoch mit Werten über 0,45 bzw. 0,70. Die interne Konsistenz ist mit einem Cronbachs α von 0,86 auf einem sehr hohen Niveau. Schließlich bestätigte die explorative Faktorenanalyse die Existenz nur eines Faktors, was darauf schließen lässt, dass die zugrunde liegenden Texte dieselbe unbeobachtete latente Variable messen, nämlich die Sprachkompetenz.

³⁸⁵ Vgl. Grotjahn 2002, S. 218f.

³⁸⁶ Diese Bezeichnung umfasst in der vorliegenden Arbeit diejenigen Teilnehmer, die vor dem sog. kritischen Alter (siehe Fn. 182) nach Deutschland einwanderten und die deutsche Sprache vor Ort erlernten.

³⁸⁷ Es wurde sichergestellt, dass sich keine englischsprachigen (Beinahe-)Muttersprachler in dieser Gruppe befanden. Die entsprechenden sprachlichen Ausschlusskriterien wurden in Anlehnung an Hayakawa et al. 2017b, S. 1389 und die dort zitierte Literatur durchgeführt.

kann geschlussfolgert werden, dass die getroffene Kategorisierung im Durchschnitt die entsprechenden Sprachkompetenzstufen tatsächlich widerspiegelt.

Die Lesbarkeit (hoch vs. niedrig) dient als die zweite unabhängige Variable. In den Gruppen mit hoher Lesbarkeit sollten die in der Fallstudie enthaltenen Informationen nur eine geringe extrinsische kognitive Belastung ausüben. In den Treatmentstufen mit niedriger Lesbarkeit wurde genau das Gegenteil verfolgt. Manche Studien verwenden den Fog-Index zur Messung der Lesbarkeit.³⁸⁸ Das Problem mit dem Fog-Index besteht darin, dass einige mehrsilbige Wörter, die als komplex identifiziert werden (wie z. B. das englische Wort „company“), selbst von weniger versierten Inverstoren mühelos verstanden werden können.³⁸⁹ Bonsall IV et al. 2017 argumentieren, dass traditionelle Lesbarkeitsindizes wie der Fog-Index nicht alle relevanten Lesbarkeitsaspekte der finanziellen Berichterstattung erfassen können, und plädieren für die Verwendung eines umfassenderen Lesbarkeitsmaßes. Zur Erhöhung der Klarheit und Verständlichkeit der finanziellen Berichterstattung gibt die SEC konkrete Empfehlungen, die im „Plain English Handbook“ formuliert sind.³⁹⁰ Weil sich die Lesbarkeitsmanipulationen basierend auf diesen Empfehlungen als effektiv bei der Auslösung veränderter kognitiver Prozesse aufgrund der Leseflusswahrnehmung und kognitiver Belastung erwiesen,³⁹¹ wurden auch in der vorliegenden Studie einige ausgewählte sprachliche Empfehlungen der SEC verwendet. Um den Informationsgehalt unverändert zu halten, kamen – wie von Rennekamp 2012 vorgeschlagen – nur die folgenden linguistischen Mittel zum Einsatz: kurze Sätze, Verwendung von Sätzen in Aktivform, keine versteckten Verben, keine überflüssigen Wörter, affirmative Sprache, einfache Synonyme, Personalpronomen sowie Sätze, bei denen Subjekt, Prädikat und Objekt nah beieinander stehen. Hinsichtlich der Formatierungsmerkmale wurden klare Überschriften, Aufzählungszeichen und ein adäquates Layout der Tabellen verwendet. In der hohen Lesbarkeitsstufe des Controlling-Berichts wurden die linguistischen und formalen Merkmale in Übereinstimmung mit den oben genannten Prinzipien gestaltet. Im Gegensatz dazu wurden diese Prinzipien in der niedrigen Lesbarkeitsstufe verletzt. In Anhang II sind einige Beispiele für die Lesbarkeitsmanipulationen aufgeführt.

Bei der Untersuchung der Risikobereitschaft müssen stets bestimmte Kontrollvariablen berücksichtigt werden. So sind bspw. geschlechtsspezifische Unterschiede bei der Risikobereitschaft gut dokumentiert, wobei Männer in vielen Fällen risikobereiter sind als

³⁸⁸ Siehe u. a. Lehavey et al. 2011; Li 2008; Miller 2010.

³⁸⁹ Vgl. Loughran und McDonald 2014, S. 1645.

³⁹⁰ Siehe SEC 1998. Im Allgemeinen umfassen die „Plain-English“-Prinzipien folgende Aspekte: (1) kurze Sätze; (2) eindeutige, konkrete und alltägliche Sprache; (3) Verwendung von Sätzen in Aktivform; (4) tabellarische Darstellungsform oder Aufzählungslisten für komplexe Sachverhalte, wenn möglich; (5) Vermeidung eines sehr fachspezifischen Jargons und (6) Vermeidung von Mehrfachverneinungen.

³⁹¹ Siehe u. a. Rennekamp 2012; Tan et al. 2014.

Frauen.³⁹² Ein weiterer wichtiger Faktor ist das Alter. Figner et al. 2009 argumentieren, dass die Risikobereitschaft in der Jugend am höchsten ist und sich im Erwachsenenalter verringert. Figner und Weber 2011 stellten zusätzlich fest, dass das Risikoverhalten domänenspezifisch ist, d. h., dieselbe Person kann in einem Bereich (z. B. Fallschirmspringen) das Eingehen von Risiken genießen und in einem anderen Bereich (z. B. Geldanlage) eher risikoavers handeln. In der vorliegenden Studie wurde die Subskala „Gambling and Investing“ aus der Domain-Specific Risk-Taking Scale von Blais und Weber 2006 verwendet, um die generelle Risikoeinstellung im Investitionsbereich zu messen. Da die Risikoaversion eher mit System 1 als mit System 2 in Verbindung steht, besteht die Möglichkeit, dass der kognitive Stil das Risikoverhalten beeinflusst. Die Tendenz, sich gemäß System 1 mehr auf Bauchgefühle und Intuitionen zu verlassen, kann mit dem Konstrukt „Faith in Intuition“ erfasst werden. Gleichzeitig werden Menschen, die das logische Denken im System-2-Modus genießen, erwartungsgemäß höhere Werte auf dem Konstrukt „Need for Cognition“ erreichen. Die (verkürzte) Skala „Faith in Intuition“ wurde aus der Studie von Epstein et al. 1996 entnommen. Das Konstrukt „Need for Cognition“ wurde mithilfe der Skala von Beißert et al. 2015 gemessen.

3.3.1.3 Abhängige Variablen

Die wichtigste abhängige Variable im Experiment – die Bereitschaft zur Akzeptanz von „smarten“ Risiken – wurde in zwei Schritten gemessen. Zunächst mussten sich die Testpersonen zwischen den beiden Kaufangeboten in der Fallstudie entscheiden. Anschließend wurden die Teilnehmer gebeten, ihre individuellen Sicherheitsäquivalente (d. h., der erforderliche Erlös der sicheren Option, mit dem der Teilnehmer indifferent zwischen der sicheren und der risikobehafteten Option wird) anzugeben. Es gab keine vorgeformulierten Antwortmöglichkeiten, stattdessen mussten die Probanden ihre Antworten in Form eines freien Eintrags abgeben. Die Entscheidung fiel auf diese Alternative, um sicherzugehen, dass die Testpersonen das Experiment ernstnehmen, da akzeptable Antworten nicht einfach mit dem Ankreuzen eines zufälligen Kästchens abgegeben werden konnten. Aus diesem Grund basiert die statistische Analyse in dieser Studie auf den individuellen Sicherheitsäquivalenten.³⁹³

³⁹² Vgl. Byrnes et al. 1999, S. 377.

³⁹³ Neben den individuellen Sicherheitsäquivalenten wurden die Teilnehmer gebeten, ihre Präferenzstärken für das gewählte Kaufangebot anzugeben. Ähnlich wie bei der Auswahlentscheidung mussten die Testpersonen ihre Präferenzstärke mit dem Ankreuzen eines Kästchens angeben. Im Vergleich zu den individuellen Sicherheitsäquivalenten bieten die Präferenzstärken somit keinen Mehrwert in Bezug auf die Designwahl. Deswegen werden die Präferenzstärken in der statistischen Analyse in Kapitel 3.4 nicht weiter berücksichtigt. An dieser Stelle soll aber betont werden, dass sich die Ergebnisse der beiden Variablen nur in einem Punkt unterscheiden: Während der FLE für die Sicherheitsäquivalente nur knapp das Signifikanzniveau erreichte, war dieser für die Präferenzstärken als abhängige Variable nicht signifikant.

3.3.2 Teilnehmer

Die Mehrheit der Teilnehmer waren Studierende der Wirtschaftswissenschaften³⁹⁴ aus einer mittelgroßen Universität in Deutschland. Alle Probanden nahmen auf freiwilliger Basis teil und die Anonymität der Teilnehmer wurde stets gewährleistet. Deskriptive Statistiken zu den Teilnehmern und ihrer Verteilung zwischen den vier Experimentgruppen sind in Tabelle 6 aufgeführt. Die Versuchspersonen waren im Durchschnitt 23,4 Jahre alt, 27,3 % waren weiblich, Studierende im Bachelor- und Masterstudium waren annähernd gleich häufig vertreten. In Bezug auf die Kontrollvariablen (Geschlecht, Alter, domänenspezifische Risikoeinstellung, Need for Cognition und Faith in Intuition) gab es keine wesentlichen Unterschiede zwischen den Experimentgruppen.

Einige der Teilnehmer mussten aus verschiedenen Gründen aus der statistischen Hauptanalyse ausgeschlossen werden. Zunächst wurden Multiple-Choice-Fragen eingesetzt, um sicherzustellen, dass der Aspekt mit dem höheren Erwartungswert der Erlöse der risikobehafteten Option relativ zur sicheren Option verstanden wurde. Außerdem sollte das Verständnis aller Teilnehmer gewährleistet werden, dass die risikobehaftete Alternative keine potenzielle Verlustsituation impliziert.³⁹⁵

Zusätzlich zu den Manipulation-Checks in Form von Multiple-Choice-Fragen wurden die Teilnehmer gebeten, ihr selbstberichtetes Verständnis der Fallstudie anzugeben. Teilnehmer, die die Multiple-Choice-Fragen nicht richtig beantworteten und/oder ein selbstberichtetes Verständnis von weniger als 50 %³⁹⁶ angaben, wurden aus der statistischen Hauptanalyse ausgeschlossen.

³⁹⁴ Daneben gab es eine kleine Anzahl von Studierenden, die in mathematischen und ingenieurwissenschaftlichen Fächern eingeschrieben waren. Für diese Studierenden wurde sichergestellt, dass sie die für das Verständnis der Fallstudie relevanten betriebswirtschaftlichen Module absolviert hatten.

³⁹⁵ Dies ist relevant, da der Gewinn- bzw. Verlustbereich unterschiedliche Muster in Bezug auf das Risikoverhalten aufweisen, vgl. Tversky und Kahneman 1981, S. 454.

³⁹⁶ Derselbe Grenzwert für das selbstberichtete Verständnis wird typischerweise in den Studien zum FLE verwendet, siehe u. a. Costa et al. 2014a. Ein selbstberichtetes Verständnis von mindestens 50 % zusammen mit den richtig beantworteten Multiple-Choice-Fragen sollte ausreichen, um verwertbare Antworten zu gewährleisten. Nichtsdestotrotz werden für Zwecke der Robustheitsprüfung auch Gesamtergebnisse für die Grenzwerte von 60 % bzw. 70 % für das selbstberichtete Verständnis in den Fußnoten angegeben.

Tabelle 6: Deskriptive Statistiken zu den Teilnehmern (Experiment 1)

| | Lesbarkeit | | | | Gesamt |
|--|-------------|-------------|-------------|-------------|-------------|
| | Hoch | | Niedrig | | |
| | Sprache | | Sprache | | |
| | Fremd- | Mutter- | Fremd- | Mutter- | |
| Geschlecht | | | | | |
| Weiblich | 9 (32,1 %) | 11 (37,9 %) | 6 (20,7 %) | 7 (20,0 %) | 33 (27,3 %) |
| Männlich | 19 (67,9 %) | 18 (62,1 %) | 23 (79,3 %) | 28 (80,0 %) | 88 (72,7 %) |
| Studiengang^a | | | | | |
| Bachelor | 13 (50,0 %) | 14 (48,3 %) | 17 (58,6 %) | 13 (37,1 %) | 57 (47,9 %) |
| Master | 13 (50,0 %) | 15 (51,7 %) | 12 (41,4 %) | 22 (62,9 %) | 62 (52,1 %) |
| Alter (<i>M</i>) | 23,7 | 23,5 | 22,6 | 23,8 | 23,4 |
| Risikoeinstellung ^b (<i>M</i>) | 2,9 | 3,1 | 3,1 | 3,2 | 3,1 |
| Faith in Intuition ^c (<i>M</i>) | 4,5 | 4,7 | 4,8 | 4,8 | 4,7 |
| Need for Cognition ^c (<i>M</i>) | 5,1 | 5,0 | 4,9 | 4,7 | 4,9 |

^a Zwei Teilnehmer in Gruppe 1 gaben nicht an, ob sie sich im Bachelor- oder Masterstudium befinden.

^b Die Risikoeinstellung steht für die generelle Einstellung der Teilnehmer gegenüber Risiken im Bereich „Gambling and Investing“ gemäß Blais und Weber 2006.

^c Faith in Intuition und Need for Cognition sind Konstrukte zur Messung der Tendenz der Teilnehmer, sich mehr auf Intuitionen (System 1) bzw. das logische Denken (System 2) zu verlassen. Die (verkürzte) Version der Skala Faith in Intuition wurde der Studie von Epstein et al. 1996 entnommen und Need for Cognition wurde mithilfe der Skala von Beißert et al. 2015 gemessen.

Quelle: Eigene Darstellung.

Darüber hinaus wurden die Probanden aus der statistischen Hauptanalyse herausgenommen, wenn sie nicht plausible Sicherheitsäquivalente (d. h., die Wahl der risikobehafteten Option in Kombination mit der Angabe eines Sicherheitsäquivalents, das unter dem vorgegebenen Erlös der sicheren Option i. H. v. 138.000 \$ liegt, et vice versa) angaben. Schließlich wurden spezifische, für die Studien zum FLE übliche Screening- und Ausschlusskriterien in Bezug auf den sprachlichen Hintergrund der Teilnehmer eingesetzt.³⁹⁷

Eine deskriptive Statistik zu den oben genannten Ausschlusskriterien sowie deren Verteilung zwischen den Experimentgruppen kann Tabelle 16 in Anhang II entnommen

³⁹⁷ Es wurden bspw. diejenigen Testpersonen aus den muttersprachlichen Gruppen ausgeschlossen, die angaben, keine deutschsprachigen Muttersprachler oder vergleichbare Sprecher zu sein, siehe Fn. 386. Außerdem wurde kontrolliert, dass sich keine englischsprachigen (Beinahe-)Muttersprachler in den fremdsprachlichen Gruppen befanden, siehe Fn. 387.

werden. Die Screeningverfahren führten zu einem relativ großen Rückgang der Gruppengrößen, was insbesondere für die erste Gruppe gilt. Auf diese Problematik wird noch ausführlicher in Kapitel 3.4.2.1 eingegangen.

3.4 Ergebnisse

3.4.1 Deskriptive Statistiken

Tabelle 7 enthält deskriptive Statistiken zu den beiden abhängigen Variablen, die die Bereitschaft der Teilnehmer, „smarte“ Risiken einzugehen, messen. Konkret handelt es sich um (1) die Auswahlentscheidung zwischen der sicheren und der risikobehafteten Option – d. h. dem Kaufangebot in der Fallstudie – und (2) das individuelle Sicherheitsäquivalent für die risikobehaftete Option. Da die Auswahlentscheidung eine kategoriale Variable darstellt, werden in Panel A nur absolute Zahlen und entsprechende Prozentangaben für jede Experimentgruppe dargestellt. Die Sicherheitsäquivalente sind metrische Größen und ermöglichen die Berechnung von deskriptiven Statistiken. Für diese Variable werden für jede Experimentgruppe der Mittelwert und die Standardabweichung in Panel B angezeigt.

Deskriptive Ergebnisse deuten darauf hin, dass eine niedrige Lesbarkeit die Bereitschaft zur Akzeptanz von „smarten“ Risiken senkt. Zusammengefasst bevorzugten die Teilnehmer in den Gruppen mit hoher Lesbarkeit die risikobehaftete Option gegenüber der sicheren Option viel häufiger (80,7 %) als ihre Pendanten in den Gruppen mit niedriger Lesbarkeit (57,8 %). Auch die Werte für die durchschnittlichen Sicherheitsäquivalente sprechen für H1 (hohe Lesbarkeit: 145 901 vs. niedrige Lesbarkeit: 135 158).

In Hypothese H2 wird postuliert, dass Entscheidungsträger nur bei einem hohen Lesbarkeitsniveau eher bereit sein werden, „smarte“ Risiken zu akzeptieren, wenn sie einen Controlling-Bericht in einer Fremdsprache im Vergleich zu ihrer Muttersprache verarbeiten müssen. Die deskriptiven Ergebnisse liefern erste Hinweise zugunsten dieser Hypothese. Bei dem hohen Lesbarkeitsniveau wurde die risikobehaftete Option gegenüber der sicheren Option in der fremdsprachlichen Gruppe häufiger gewählt (89,3 %) als in der muttersprachlichen Gruppe (72,4 %). Dasselbe gilt für die Sicherheitsäquivalente innerhalb der hohen Lesbarkeitsstufe (Fremdsprache: 149 768 vs. Muttersprache: 142 034). Im Gegensatz dazu unterscheidet sich das Risikoverhalten zwischen den fremd- und muttersprachlichen Gruppen deutlich weniger, wenn das niedrige Lesbarkeitsniveau betrachtet wird. Alles in allem kann H2 auf deskriptiver Ebene unterstützt werden.

Tabelle 7: Deskriptive Statistiken zu den abhängigen Variablen (Experiment 1)

| Panel A: Auswahlentscheidung ^a (absolute Zahl [Prozentangabe]) | | | | | |
|---|----------------|----------------|----------------|----------------|----------------|
| | Lesbarkeit | | | | Gesamt |
| | Hoch | | Niedrig | | |
| | Sprache | | Sprache | | |
| | Fremd- | Mutter- | Fremd- | Mutter- | |
| Risikobehaftete Option | 25 [89,3 %] | 21 [72,4 %] | 15 [51,7 %] | 22 [62,9 %] | 83 [68,6 %] |
| Sichere Option | 3 [10,7 %] | 8 [27,6 %] | 14 [48,3 %] | 13 [37,1 %] | 38 [31,4 %] |

| Panel B: Sicherheitsäquivalent ^b (Mittelwert [Standardabweichung]) | | | | | |
|---|---------------------|---------------------|---------------------|---------------------|---------------------|
| | Lesbarkeit | | | | Gesamt |
| | Hoch | | Niedrig | | |
| | Sprache | | Sprache | | |
| | Fremd- | Mutter- | Fremd- | Mutter- | |
| Sicherheitsäquivalent | 149 768 [18 847] | 142 034 [16 879] | 132 673 [17 668] | 137 643 [21 447] | 140 310 [19 680] |

^a Die Auswahlentscheidung ist eine Dummy-Variable und kann in Abhängigkeit davon, ob die sichere oder risikobehaftete Option – d. h. das Kaufangebot in der Fallstudie – gewählt wurde, zwei mögliche Zustände einnehmen.

^b Das Sicherheitsäquivalent ist die wichtigste abhängige Variable und repräsentiert den individuell geforderten Erlös der sicheren Option, mit dem der Teilnehmer indifferent zwischen der sicheren und der risikobehafteten Option wird.

Quelle: Eigene Darstellung.

Schließlich besagt die dritte Hypothese H3, dass der FLE – wie unter H2 formuliert – nur für mittelhohe Sprachkompetenzstufen gilt. Die Ergebnisse auf deskriptiver Ebene liefern einige Hinweise zugunsten dieser Hypothese (Zahlen sind nicht abgebildet). Innerhalb der hohen Lesbarkeitsstufe entschieden sich alle Teilnehmer mit einem mittelhohen Sprachkompetenzniveau für die risikobehaftete Option, Probanden mit einem geringen Sprachkompetenzniveau präferierten die risikobehaftete Option am zweithäufigsten (82,4 %) und die niedrigste Präferenzrate zugunsten der risikobehafteten Option wiesen die Muttersprachler auf (72,4 %). Das gleiche Ergebnismuster lässt sich für die Sicherheitsäquivalente beobachten. Innerhalb der hohen Lesbarkeitsstufe gaben die Probanden mit dem mittelhohen Sprachkompetenzniveau im Durchschnitt die höchsten

Sicherheitsäquivalente an, gefolgt von Teilnehmern mit dem geringen Sprachkompetenzniveau und den Muttersprachlern (161 273; 142 324 und 142 034 in entsprechender Reihenfolge).³⁹⁸

3.4.2 Hypothesentests

3.4.2.1 H1 und H2

Statistische Tests der Hypothesen H1 und H2 wurden mithilfe der Kovarianzanalyse (ANCOVA) durchgeführt und als Follow-up-Tests zur Überprüfung der genauen Gruppenunterschiede wurde die Kontrastanalyse verwendet.³⁹⁹ Die Risikoeinstellung diente als Kovariate und die folgenden erklärenden Variablen wurden herangezogen: Sprache, Lesbarkeit und ihre Interaktion. Die individuellen Sicherheitsäquivalente stellten die abhängige Variable dar.⁴⁰⁰

Die Ergebnisse der ANCOVA zusammen mit den Ergebnissen der Kontrastanalyse sind in Tabelle 8 abgebildet. Der Effekt der Lesbarkeit befindet sich auf dem höchsten Signifikanzniveau ($F = 9,33$; $p = 0,003$; Panel A), was Hypothese H1 vollständig bestätigt.⁴⁰¹ Der signifikante Interaktionsterm ($F = 3,47$; $p = 0,065$, Panel A) erlaubt es, den FLE – wie unter H2 formuliert – weiter zu untersuchen. Tabelle 8, Panel B zusammen mit Abbildung 14 bezieht sich auf die Interaktion zwischen der Sprache und Lesbarkeit. Wie zu entnehmen ist, gibt es innerhalb der hohen Lesbarkeitsbedingung eine positive und leicht signifikante Differenz zwischen der fremd- und muttersprachlichen Gruppe ($t = 1,67$; $p = 0,097$).

³⁹⁸ Innerhalb der niedrigen Lesbarkeitsstufe war der Unterschied der Teilnehmer mit dem mittelhohen Sprachkompetenzniveau im Vergleich zu den anderen Teilnehmern weniger stark ausgeprägt. Die Auswahl der risikobehafteten Option ist für jedes Sprachkompetenzniveau wie folgt verteilt: 62,9 % (muttersprachlich); 35,7 % (mittelhoch) und 66,7 % (gering). Und für die Sicherheitsäquivalente gilt: 135 223 (gering); 129 940 (mittelhoch) und 137 643 (muttersprachlich).

³⁹⁹ Zu den beiden Analysemethoden im Rahmen von Experimenten siehe Sedlmeier und Renkewitz 2013.

⁴⁰⁰ Weil eine relativ große Anzahl von Teilnehmern aussortiert werden musste und die Ausschlussraten zwischen den Experimentgruppen nicht gleichmäßig verteilt sind, wurde im Rahmen der Voranalyse eine Dummy-Variable aufgenommen, um dem Umstand Rechnung zu tragen, ob ein Proband ausgeschlossen wurde oder nicht. Die Ergebnisse des vollen ANCOVA-Modells mit der Rohstichprobe (siehe Tabelle 17, Anhang II) zeigen, dass von allen Kontrollvariablen nur die Risikoeinstellung signifikant ist ($F = 3,85$; $p = 0,051$). Noch bedeutender ist, dass die oben erwähnte Dummy-Variable nicht signifikant ist ($F = 0,26$; $p = 0,612$). Daraus kann geschlussfolgert werden, dass die Tatsache, ob ein Teilnehmer ausgeschlossen wurde oder nicht, keinen systematischen Einfluss auf die Wahl des individuellen Sicherheitsäquivalents ausübte. Daher werden in der statistischen Hauptanalyse die um die Ausschlusskriterien bereinigte Stichprobe und ein ANCOVA-Modell mit Risikoeinstellung als Kontrollvariable verwendet.

⁴⁰¹ Höchstsignifikante Ergebnisse zur Lesbarkeit werden ebenfalls generiert, wenn statistische Modelle mit der Rohstichprobe und nur die Teilnehmer mit einem selbstberichteten Verständnis der Fallstudie von 60 % bzw. 70 % verwendet werden (Zahlen sind nicht abgebildet).

Tabelle 8: Statistische Analyseergebnisse – Teil 1 (Experiment 1)

| Panel A: Haupteffekte und Interaktion der ANCOVA | | | | |
|---|----|----------------------|--------|----------|
| Prädiktorvariable (Sicherheitsäquivalent ^a) | Fg | Quadratsumme | F-Wert | p-Wert |
| Sprache ^b | 1 | 6,23×10 ⁷ | 0,18 | 0,674 |
| Lesbarkeit ^c | 1 | 3,26×10 ⁹ | 9,33 | 0,003*** |
| Risikoeinstellung ^d | 1 | 1,44×10 ⁹ | 4,11 | 0,045** |
| Sprache × Lesbarkeit | 1 | 1,21×10 ⁹ | 3,47 | 0,065* |

| Panel B: Ergebnisse der Kontrastanalyse mit ANCOVA | | | | |
|--|-----|-----------|--------|--------|
| Kontrast (Sicherheitsäquivalent) | Fg | Differenz | t-Wert | p-Wert |
| Lesbarkeit hoch: Sprache: Fremd- vs. Mutter | 116 | 8 303 | 1,67 | 0,097* |
| Lesbarkeit niedrig: Sprache: Fremd- vs. Mutter | 116 | -4 397 | -0,94 | 0,352 |

Alle *p*-Werte sind zweiseitig. *, **, *** repräsentieren eine Signifikanz auf den Niveaus 10 %, 5 % und 1 %.

^aDas Sicherheitsäquivalent repräsentiert den individuell geforderten Erlös der sicheren Option, mit dem der Teilnehmer indifferent zwischen der sicheren und der risikobehafteten Option wird.

^bDie Sprache ist eine Dummy-Variable mit den beiden Ausprägungen: Muttersprache (Deutsch) und Fremdsprache (Englisch).

^cDie Lesbarkeit ist eine Dummy-Variable mit den beiden Ausprägungen hohe und niedrige Lesbarkeit. Die Lesbarkeit wurde mithilfe des „Plain English Handbook“ von SEC 1998 mittels ausgewählter linguistischer und formaler Gestaltung manipuliert.

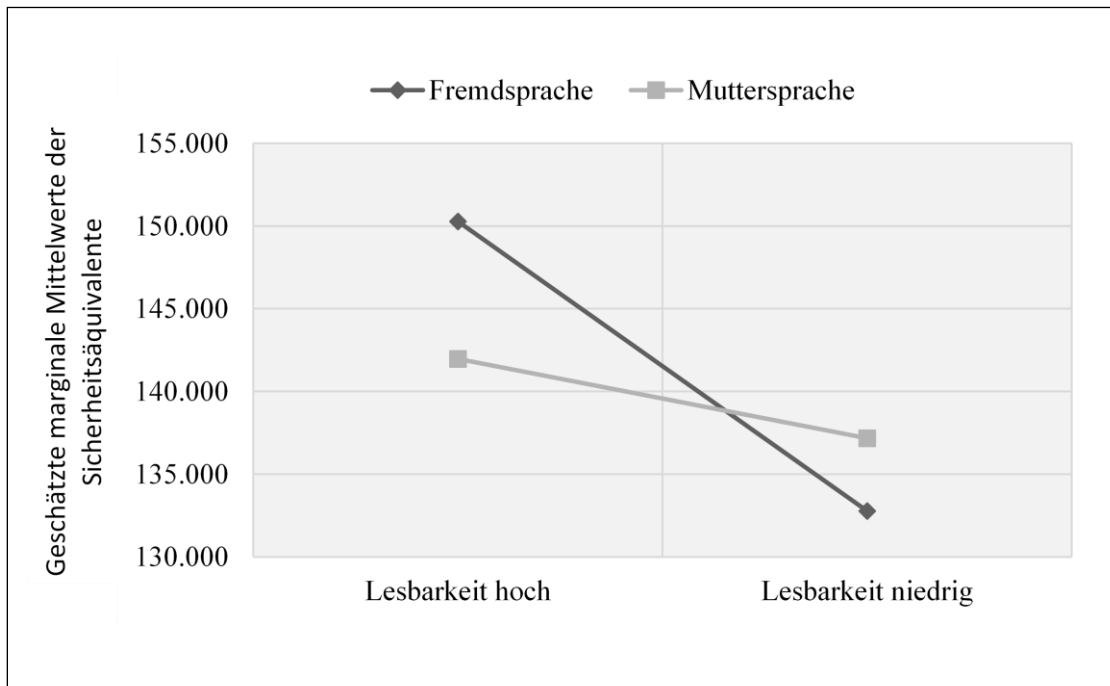
^dDie Risikoeinstellung steht für die generelle Einstellung der Teilnehmer gegenüber Risiken im Bereich „Gambling and Investing“ gemäß Blais und Weber 2006.

Quelle: Eigene Darstellung.

Im Gegensatz dazu verschwindet der signifikante Unterschied zwischen diesen Gruppen, wenn das niedrige Lesbarkeitsniveau betrachtet wird ($t = -0,94$; $p = 0,352$). Dieses Ergebnismuster spiegelt den in Hypothese H2 formulierten Sachverhalt wider. Alle Schlussfolgerungen müssen allerdings mit Vorsicht gezogen werden, da der Unterschied zwischen der fremd- und muttersprachlichen Gruppe innerhalb der hohen Lesbarkeitsbedingung nur marginal signifikant und nicht besonders robust gegenüber Änderungen der Stichprobenzusammensetzung aufgrund der Anwendung verschiedener Ausschlusskriterien ist.⁴⁰²

⁴⁰² Mit der Rohstichprobe ist der Interaktionsterm nicht signifikant. Wenn die Teilnehmer mit einem selbstberichteten Verständnis der Fallstudie von 60 % bzw. 70 % berücksichtigt werden, schwanken die Ergebnisse um das 10 %-Signifikanzniveau herum (Zahlen sind nicht abgebildet).

Abbildung 14: Interaktion zwischen Sprache und Lesbarkeit



Quelle: Eigene Darstellung.

Zusammengenommen wird die Hypothese H1, die auf dem CLE aufgrund niedriger Lesbarkeit aufbaut, stark unterstützt. Mit anderen Worten: Probanden innerhalb der niedrigen Lesbarkeitsbedingung gingen signifikant weniger „smarte“ Risiken ein – relativ zu ihren Pendanten in den Gruppen mit hoher Lesbarkeit. Darüber hinaus liefern die Ergebnisse erste (mit Vorsicht zu interpretierende) Hinweise zugunsten von H2. Demnach kann der Einsatz einer Fremdsprache im Vergleich zur eigenen Muttersprache die Bereitschaft zur Übernahme von „smarten“ Risiken fördern, aber nur, wenn eine hohe Lesbarkeit gewährleistet ist.

3.4.2.2 H3

Das Ziel dieses Abschnitts besteht darin, den FLE weiter zu untersuchen, indem die Sprachkompetenz der Experimentteilnehmer berücksichtigt wird. In Hypothese H3 wird postuliert, dass der FLE – wie unter H2 formuliert – nur für mittelhohe Sprachkompetenzstufen gilt.

Um Hypothese H3 zu testen, wurde – ähnlich dem Vorgehen von Oganian et al. 2016 – im ursprünglichen ANCOVA-Modell die unabhängige Variable Sprache durch Sprachkompetenz ersetzt (siehe Tabelle 9, Panel A). Es liegen ein höchst signifikanter Haupteffekt der Lesbarkeit ($F = 10,16$; $p = 0,002$) sowie ein signifikanter Interaktionsterm ($F = 4,57$; $p = 0,012$) vor.

Tabelle 9: Statistische Analyseergebnisse – Teil 2 (Experiment 1)

| Panel A: Haupteffekte und Interaktion der ANCOVA | | | | |
|---|----|----------------------|--------|----------|
| Prädiktorvariable (Sicherheitsäquivalent ^a) | Fg | Quadratsumme | F-Wert | p-Wert |
| Sprachkompetenz ^b | 2 | 3,77×10 ⁸ | 0,56 | 0,571 |
| Lesbarkeit ^c | 1 | 3,40×10 ⁹ | 10,16 | 0,002*** |
| Risikoeinstellung ^d | 1 | 1,52×10 ⁹ | 4,57 | 0,035** |
| Sprachkompetenz × Lesbarkeit | 2 | 3,06×10 ⁹ | 4,57 | 0,012** |

| Panel B: Ergebnisse der Kontrastanalyse mit ANCOVA | | | | |
|--|-----|-----------|--------|---------|
| Kontrast (Sicherheitsäquivalent) | Fg | Differenz | t-Wert | p-Wert |
| Lesbarkeit hoch: Sprachkompetenz: | | | | |
| Mittelhoch vs. gering | 114 | 18 507 | 2,61 | 0,010** |
| Mittelhoch vs. muttersprachlich | 114 | 19 499 | 3,01 | 0,003** |
| Gering vs. muttersprachlich | 114 | 992 | 0,18 | 0,860 |
| Lesbarkeit niedrig: Sprachkompetenz: | | | | |
| Mittelhoch vs. gering | 114 | -3 910 | -0,57 | 0,568 |
| Mittelhoch vs. muttersprachlich | 114 | -6 460 | -1,11 | 0,269 |
| Gering vs. muttersprachlich | 114 | -2 550 | -0,45 | 0,652 |

| Panel C: Geschätzte marginale Mittelwerte der Sicherheitsäquivalente | | | |
|--|-----------------|------------|---------|
| | Sprachkompetenz | | |
| | Mutterspr. | Mittelhoch | Gering |
| Lesbarkeit hoch | 141 982 | 161 481 | 142 974 |
| Lesbarkeit niedrig | 137 217 | 130 757 | 134 667 |

Alle p-Werte sind zweiseitig. *, **, *** repräsentieren eine Signifikanz auf den Niveaus 10 %, 5 % und 1 %.

^a Das Sicherheitsäquivalent repräsentiert den individuell geforderten Erlös der sicheren Option, mit dem der Teilnehmer indifferent zwischen der sicheren und der risikobehafteten Option wird.

^b Die Sprachkompetenz wurde anhand eines C-Tests (siehe Kapitel 3.3.1.2 und Anhang I) gemessen. Die Teilnehmer wurden in die drei Sprachkompetenzgruppen eingeteilt: muttersprachlich; mittelhoch; gering.

^c Die Lesbarkeit ist eine Dummy-Variable mit den beiden Ausprägungen hohe und niedrige Lesbarkeit. Die Lesbarkeit wurde mithilfe des „Plain English Handbook“ von SEC 1998 mittels ausgewählter linguistischer und formaler Gestaltung manipuliert.

^d Die Risikoeinstellung steht für die generelle Einstellung der Teilnehmer gegenüber Risiken im Bereich „Gambling and Investing“ gemäß Blais und Weber 2006.

Quelle: Eigene Darstellung.

Des Weiteren zeigt die Kontrastanalyse (siehe Tabelle 9, Panel B), dass innerhalb der hohen Lesbarkeitsbedingung ein signifikant positiver Unterschied zwischen den folgenden beiden Gruppenpaaren besteht: mittelhoch vs. gering ($t = 2,61$; $p = 0,010$) sowie mittelhoch vs. muttersprachlich ($t = 3,01$; $p = 0,003$). Daneben ist die Differenz zwischen dem Gruppenpaar gering vs. muttersprachlich nicht signifikant ($t = 0,18$; $p = 0,860$), was gleichermaßen für die übrigen Gruppenpaarvergleiche innerhalb niedriger Lesbarkeit gilt (alle p -Werte $> 0,269$).⁴⁰³

Diese Ergebnisse untermauern die Prognose aus H3 und liefern die erste Evidenz dafür, dass der FLE tatsächlich nur für mittelhohe Sprachkompetenzstufen gilt. Dieser Sachverhalt wird auch von einem umgekehrt verlaufenden U-förmigen Muster der geschätzten marginalen Mittelwerte der Sicherheitsäquivalente mit Sprachkompetenz bei der hohen Lesbarkeitsstufe deutlich unterstrichen (siehe Tabelle 9, Panel C).

3.4.3 Ergänzende Analysen

3.4.3.1 Manipulation-Checks und alternative Erklärungen

In diesem Kapitel werden die Validität der Lesbarkeitsmanipulationen überprüft und potenzielle Alternativerklärungen für die gewonnenen Ergebnisse adressiert.

Zunächst soll überprüft werden, ob die Lesbarkeitsmanipulationen erfolgreich waren und mit H1 konform sind. Basierend auf den Ratingskalen von Elliott et al. 2015 wurde die wahrgenommene Lesbarkeit gemessen, indem die Teilnehmer gebeten wurden, auf einer siebenstufigen Likert-Skala von 1 (= sehr leicht) bis 7 (= sehr schwer) anzugeben, wie „schwer zu lesen“, „schwer zu verstehen“ und „schwer zu verarbeiten“ sie die präsentierten Informationen in der Fallstudie empfanden. Da alle Fragen zu den Lesbarkeitsmanipulationen stark korrelieren (alle Pearson-Korrelationskoeffizienten $r > 0,61$), wurde aus diesen Fragen ein Mittelwert gebildet, der dann als Lesbarkeitsmaß verwendet wurde. Ergebnisse der ANOVA (Zahlen sind nicht abgebildet) mit dem Lesbarkeitsmaß als abhängige Variable und Sprache, Lesbarkeit sowie deren Interaktion als unabhängige Variablen zeigten, dass Sprache und der Interaktionsterm nicht signifikant sind ($F = 1,55$ und $F = 2,03$; $p = 0,215$ und $p = 0,157$, jeweils zweiseitig), wohingegen Lesbarkeit höchst signifikant ist ($F = 7,91$, $p < 0,006$, zweiseitig) – mit dem durchschnittlichen Lesbarkeitsmaß von 2,22 für die Gruppen mit hoher Lesbarkeit und 2,77 für die Gruppen mit niedriger Lesbarkeit. Daraus lässt sich schließen, dass die Lesbarkeitsmanipulationen erfolgreich waren und das Antwortverhalten der Teilnehmer mit H1 konform ist.

⁴⁰³ Annähernd identische Ergebnisse werden generiert, wenn nur die Teilnehmer mit einem selbstberichteten Verständnis der Fallstudie von 60 % bzw. 70 % berücksichtigt werden. Allerdings verliert der Interaktionsterm seine Signifikanz, wenn die Rohstichprobe verwendet wird (Zahlen sind nicht abgebildet).

Ein weiterer potenzieller Schlüsselfaktor für die Ergebnisse könnte ein zu geringes Verständnis einiger Teilnehmer im Experiment aufgrund der niedrigen Lesbarkeit und/oder des Einsatzes einer Fremdsprache sein. Um diesen potenziellen Erklärungsfaktor auszuschließen, wurde eine ANOVA mit dem selbstberichteten Verständnis der Fallstudie der Teilnehmer in Abhängigkeit von Sprache, Lesbarkeit sowie deren Interaktion durchgeführt (Zahlen sind nicht abgebildet). Gemäß den gewonnenen Ergebnissen ist keine der abhängigen Variablen signifikant (alle p -Werte $> 0,500$). Des Weiteren machten Probanden, die ein Verständnis von weniger als 50 % angaben, nur 8,6 % aller Teilnehmer im Experiment aus. Letzten Endes kann daraus geschlussfolgert werden, dass mangelndes Verständnis nicht als Erklärung für die Ergebnisse im Experiment dienen kann.

Schließlich wurden die meisten Teilnehmer aufgrund von falsch beantworteten Manipulation-Checks-Fragen ausgeschlossen, obwohl konstatiert werden kann, dass alle Probanden die für das Verständnis der Fallstudie relevanten Module besucht haben.⁴⁰⁴ Bei der Gestaltung des Experiments wurde großer Wert auf die Klarheit der relevanten Aspekte der Fallstudie gelegt – insbesondere die Erlöse und Wahrscheinlichkeiten waren fett abgedruckt – und eine kurze Definition des Erwartungswertkonzepts befand sich in Klammern neben den Manipulation-Checks-Fragen. Darüber hinaus gab es keine besonderen Vorkommnisse bei der Durchführung des Experiments. Über den genauen Grund für die hohe Ausschlussrate kann nur spekuliert werden. Da die Probanden absichtlich nicht dazu animiert wurden, sich bei ihrer Entscheidung auf formale Rechenschemata (wie das Erwartungswertkonzept) zu stützen, könnten einige Individuen bei der Entscheidungsfindung andere Kriterien zugrunde gelegt haben, ohne dabei die gegebenen Wahrscheinlichkeiten und Erlöse in der Fallstudie zu beachten.

3.4.3.2 Überprüfung der Erklärungsansätze für den FLE

In diesem Kapitel wird untersucht, ob der vorgefundene FLE auf die im theoretischen Teil der Arbeit diskutierten Erklärungsansätze zurückgeführt werden kann. Im Einzelnen werden die folgenden Erklärungsansätze getestet: Reduktion der emotionalen Resonanz, psychologische Distanz und DFE.

Laut dem ersten Erklärungsansatz führt reduzierte emotionale Resonanz dazu, dass die Individuen weniger anfällig für affektives Verhalten werden. Folglich verlassen sich Entscheidungsträger mehr auf den analytischen Denkmodus von System 2, der rationale Entscheidungen fördert – im vorliegenden Fall ist es höhere Bereitschaft zur Akzeptanz von „smarten“ Risiken. Bei einem geringen Sprachkompetenzniveau und/oder einer niedrigen Lesbarkeitsstufe ist jedoch System 2 so sehr mit den Informationsverarbeitungsschwierigkeiten okkupiert, dass der Entscheidungsträger sehr wahrscheinlich von System 1 geleitet wird. Es kann daher angenommen werden, dass der positive Effekt

⁴⁰⁴ Eine Frage umfasste betriebswirtschaftliches Grundwissen zur Bestimmung des Erwartungswerts.

der reduzierten emotionalen Resonanz nur für mittelhohe Sprachkompetenzniveaus und bei einer hohen Lesbarkeitsstufe gilt.

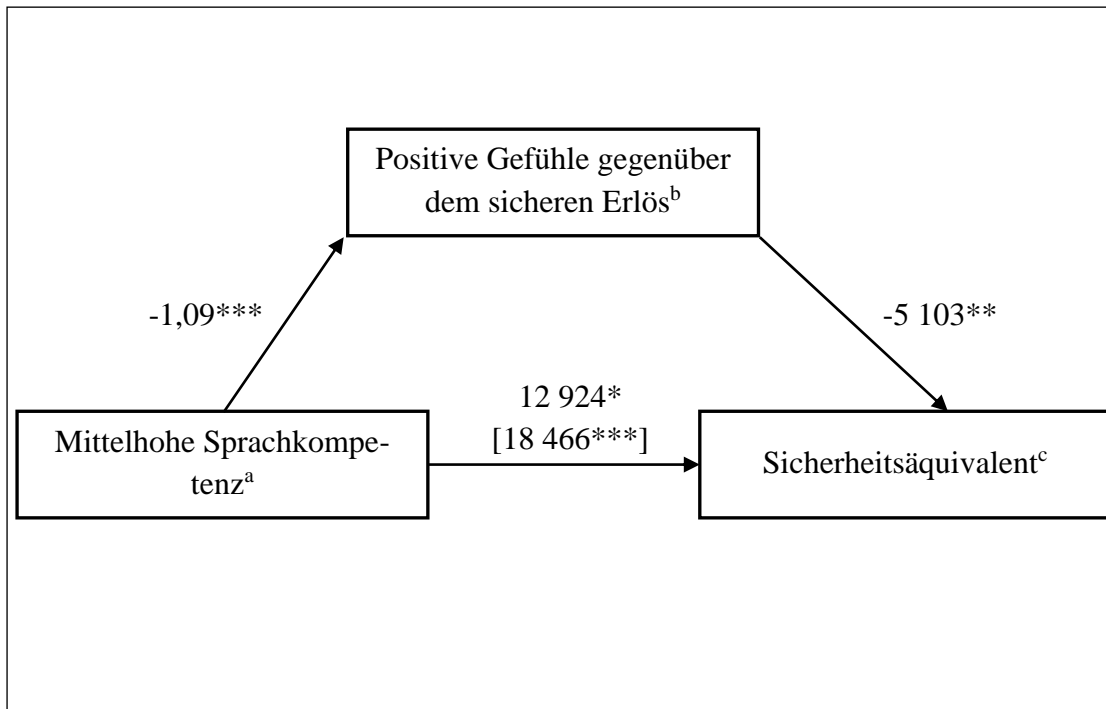
Um herauszufinden, in welchem Ausmaß die Entscheidungen der Teilnehmer in der Fallstudie von emotionalen Einflüssen geleitet waren, gaben die Probanden auf einer Skala von 1 (= weder glücklich, noch unglücklich) bis 7 (= sehr glücklich) ihre positiven Gefühle gegenüber dem sicheren Erlös des risikolosen Kaufangebots an. Um zusätzlich den Grad der negativen Gefühle gegenüber dem potenziellen niedrigen Erlös, der aus dem risikobehafteten Kaufangebot resultieren könnte, zu messen, wurden die Teilnehmer gebeten, auf einer Skala von 1 (= überhaupt nicht traurig) bis 7 (= sehr traurig) anzugeben, wie sie sich im Falle des niedrigen Erlöses fühlen würden.

Im ersten Schritt wurde ein ANOVA-Modell mit dem Einfluss der Sprache, Lesbarkeit und deren Interaktion auf positive/negative Gefühle aufgestellt, um den Erklärungsansatz basierend auf der reduzierten emotionalen Resonanz zu testen (Zahlen sind nicht abgebildet). Weder für positive noch für negative Gefühle sind die Ergebnisse signifikant. Im zweiten Schritt wurde ein sensibleres Maß für den FLE in Form von Sprachkompetenzniveaus benutzt. Der Interaktionsterm weist dann im Fall von positiven Gefühlen ein signifikantes Ergebnis auf ($F = 3,30$; $p = 0,040$). Bei dem hohen Lesbarkeitsniveau gaben die Teilnehmer aus den unterschiedlichen Sprachkompetenzgruppen im Durchschnitt folgende positive Gefühle an: 2,76 (muttersprachlich); 1,91 (mittelhoch) und 3,00 (gering). Für das niedrige Lesbarkeitsniveau sind die Werte wie folgt verteilt: 2,57 (muttersprachlich); 2,93 (mittelhoch) und 2,60 (gering). Follow-up-Tests anhand einer Kontrastanalyse zeigten, dass Probanden mit einem mittelhohen Sprachkompetenzniveau und innerhalb der hohen Lesbarkeitsbedingung signifikant weniger von positiven Gefühlen als die anderen Sprachkompetenzgruppen tangiert wurden (Zahlen sind nicht abgebildet).

Schließlich wurde ein Kausalmodell aufgestellt, um festzustellen, ob die Wirkung des mittelhohen Sprachkompetenzniveaus auf die erhöhte Bereitschaft, „smarte“ Risiken zu akzeptieren, von reduzierten positiven Gefühle gegenüber dem sicheren Erlös in der hohen Lesbarkeitsbedingung stammt (siehe Abbildung 15). Zunächst wurde eine Regressionsanalyse mit individuellen Sicherheitsäquivalenten als abhängige Variable und Sprachkompetenzniveaus als unabhängige Variable sowie Risikoeinstellung als Kontrollvariable durchgeführt. Erwartungskonform gaben Probanden mit dem mittelhohen Sprachkompetenzniveau verglichen mit anderen Sprachkompetenzstufen signifikant höhere Sicherheitsäquivalente an ($b = 18\,466$; $p = 0,006$). Ebenso besteht ein höchst signifikanter negativer Zusammenhang zwischen dem mittelhohen Sprachkompetenzniveau und positiven Gefühlen ($b = -1,09$; $p = 0,009$).

Wird schließlich ein Regressionsmodell mit Sicherheitsäquivalenten als abhängige Variable und Sprachkompetenz, Risikoeinstellung und positiven Gefühlen (Mediatorvariable) als unabhängige Variablen verwendet, verliert der Einfluss des mittelhohen

Abbildung 15: Test des Kausalmodells (hohe Lesbarkeitsbedingung)



Alle p -Werte sind zweiseitig. *, **, *** repräsentieren eine Signifikanz auf den Niveaus 10 %, 5 % und 1 %.

^a Die Sprachkompetenz wurde anhand eines C-Tests (siehe Kapitel 3.3.1.2 und Anhang I) gemessen. Die Teilnehmer wurden in die drei Sprachkompetenzgruppen eingeteilt: muttersprachlich; mittelhoch; gering.

^b Die positiven Gefühle gegenüber dem sicheren Erlös geben das Ausmaß der Attraktion des sicheren Kaufangebots in der Fallstudie an.

^c Das Sicherheitsäquivalent repräsentiert den individuell geforderten Erlös der sicheren Option, mit dem der Teilnehmer indifferent zwischen der sicheren und der risikobehafteten Option wird.

Quelle: Eigene Darstellung.

Sprachkompetenzniveaus auf die Sicherheitsäquivalente deutlich an Signifikanz ($b = 12\,924$; $p = 0,053$). Um zu überprüfen, ob die beobachtete Mediation statistisch signifikant ist, wurde eine bootstrapping-basierte Mediationsanalyse⁴⁰⁵ mit 5 000 Simulationen durchgeführt und es stellte sich heraus, dass der Mediationseffekt statistisch signifikant ist ($p = 0,005$). Insgesamt wird ausgehend von den vorliegenden Ergebnissen der auf der Reduktion der emotionalen Resonanz basierende Erklärungsansatz – der auch der führende Erklärungsansatz für den FLE im Risikokontext ist – unterstützt. Anders formuliert waren die Teilnehmer mit einem mittelhohen Sprachkompetenzniveau im Experiment von dem sicheren, aber weniger profitablen finanziellen

⁴⁰⁵ Siehe Preacher und Hayes 2004.

Ergebnis signifikant weniger „angezogen“ und zeigten daher innerhalb der hohen Lesbarkeitsbedingung eine höhere Bereitschaft, „smarte“ Risiken einzugehen.⁴⁰⁶

Ein anderer Erklärungsansatz für den FLE beruht auf dem Mechanismus der psychologischen Distanz. Übertragen auf das Risikoverhalten würde dieser Erklärungsansatz bedeuten, dass der Einsatz einer Fremdsprache die Aufmerksamkeit des Entscheidungsträgers mehr auf den erwünschten Endzustand (d. h. die Aussicht auf den höheren Erlös) über eine höhere Abstraktionsstufe lenkt. Demgegenüber wird in der eigenen Muttersprache aufgrund einer niedrigeren Abstraktionsstufe die Realisierbarkeit (d. h. die Wahrscheinlichkeit für das Zustandekommen von Erlösen) in den Vordergrund gerückt.⁴⁰⁷ Dieser Logik folgend wurden die Teilnehmer gebeten anzugeben, inwiefern ihre Entscheidung von folgenden Faktoren beeinflusst wurde: (1) höchster potenzieller Erlös und (2) Wahrscheinlichkeiten für die verschiedenen Erlöse. Für diese Faktoren als abhängigen Variablen lieferte keines der ANOVA-Modelle mit Sprache/Sprachkompetenz, Lesbarkeit und deren Interaktion als unabhängigen Variablen signifikante Ergebnisse (Zahlen sind nicht abgebildet). Zusammengenommen kann ausgehend von den vorliegenden Resultaten der auf der psychologischen Distanz basierende Erklärungsansatz verworfen werden.

Schließlich sollte gemäß dem DFE der Einsatz einer Fremdsprache mittels Generierung von metakognitiven Verarbeitungsschwierigkeiten den analytischen Denkmodus von System 2 aktivieren und damit auch die höhere Bereitschaft zur Akzeptanz von „smarten“ Risiken fördern. Dieser Erklärungsansatz wurde mithilfe der gleichen Fragen, die zur Überprüfung der Lesbarkeitsmanipulation dienten, untersucht. Wie bereits im vorherigen Kapitel beschrieben, ist für den zusammengesetzten Wert dieser Fragen nur die Lesbarkeit signifikant.⁴⁰⁸ Aus diesem Grund wird ausgehend von den vorliegenden Ergebnissen der DFE als Erklärung für den FLE nicht unterstützt.

3.5 Diskussion der Ergebnisse

Die vorliegende Studie wurde konzipiert, um zu untersuchen, ob und wie bestimmte linguistische Merkmale von Controlling-Berichten das Risikoverhalten im Hinblick auf „smarte“ Risiken beeinflussen. Im Speziellen wurden die Effekte der Lesbarkeit und des Einsatzes einer Fremdsprache betrachtet. Außerdem wurde der FLE tiefergehend untersucht, indem die Rolle der Sprachkompetenz berücksichtigt und die potenziellen Erklärungsansätze für den FLE getestet wurden.

⁴⁰⁶ Annähernd identische Ergebnisse werden generiert, wenn nur die Teilnehmer mit einem selbstberichteten Verständnis der Fallstudie von 60 % bzw. 70 % berücksichtigt werden (Zahlen sind nicht abgebildet).

⁴⁰⁷ Vgl. Sagristano et al. 2002, S. 365.

⁴⁰⁸ Gleiches gilt mit der Sprachkompetenz als unabhängige Variable.

Zur Überprüfung der Hypothesen wurde ein Between-Subject-Experiment durchgeführt. Teilnehmern, die die Rolle eines Managers einnahmen, wurde ein fiktiver Controlling-Bericht vorgelegt. Auf Grundlage der Informationen im Controlling-Bericht mussten sich die Probanden zwischen einer sicheren und einer risikobehafteten – dafür aber vorteilhafteren – Option entscheiden. Anschließend gaben die Teilnehmer ihre individuellen Sicherheitsäquivalente an, die die wichtigste abhängige Variable zur Messung der Bereitschaft, „smarte“ Risiken einzugehen, sind. Eine hohe Lesbarkeit wurde gewährleistet, indem bestimmte linguistische und formale Gestaltungsprinzipien gemäß SEC 1998 eingehalten wurden. Im Gegensatz dazu wurden diese Prinzipien in der niedrigen Lesbarkeitsbedingung verletzt. In Bezug auf die Sprache erhielt ein Teil der Testpersonen den Controlling-Bericht in ihrer Muttersprache und der andere Teil in einer Fremdsprache. Die fremdsprachliche Sprachkompetenz wurde mithilfe eines standardisierten C-Tests gemessen.

Die wichtigste Erkenntnis ist, dass die Bereitschaft, „smarte“ Risiken einzugehen, signifikant reduziert wird, wenn die Lesbarkeit niedrig ist. Dieses Phänomen kann mithilfe des CLE erklärt werden, der postuliert, dass ein schwierig zu lesender Text eine kognitive Belastung darstellt, die den Entscheidungsträger gemäß den DST daran hindert, eine Entscheidung im analytischen Denkmodus von System 2 zu treffen. Stattdessen wird der Entscheidungsträger vom affektiven System-1-Denkmodus geleitet.

Laut dem FLE steigert die Verarbeitung von Informationen in einer Fremdsprache den Beitrag von System 2 relativ zu System 1 im Entscheidungsfindungsprozess und infolgedessen – aufgrund reduzierter emotionaler Resonanz, psychologischer Distanz und/oder des DFE – die Bereitschaft, „smarte“ Risiken einzugehen. Da aber ein niedriges Lesbarkeitsniveau die beschränkten kognitiven Ressourcen, die für ein effektives Funktionieren von System 2 benötigt werden, beansprucht, ist das Auftreten des FLE nur bei hoher Lesbarkeit zu erwarten. Der FLE kann weiterführend untersucht werden, indem die Sprachkompetenz berücksichtigt wird. Ein annähernd muttersprachliches Kompetenzniveau wird das Entscheidungsergebnis nur marginal ändern. Zugleich induzieren niedrige Sprachkompetenzniveaus aufgrund von Verarbeitungsschwierigkeiten von Informationen erwartungsgemäß gewisse kognitive Belastung, was wiederum ein effektives Funktionieren von System 2 behindert. Also ist zu erwarten, dass der FLE nur für mittelhohe Sprachkompetenzstufen zum Tragen kommt.

Es konnte nur ein marginal signifikanter und nicht besonders robuster FLE mit „aggregierter“ Sprachvariable festgestellt werden. Unter Verwendung von Sprachkompetenzstufen verbesserte sich das Signifikanzergebnis und der FLE war nur für das mittelhohe Sprachkompetenzniveau und bei hoher Lesbarkeit zu beobachten.

Schließlich zeigten ergänzende Analysen in Bezug auf die potenziellen Erklärungsansätze für den FLE, dass Probanden mit dem mittelhohen Sprachkompetenzniveau und bei hoher Lesbarkeit signifikant weniger von positiven Gefühlen gegenüber dem sicheren Erlös geleitet waren. Dieses Ergebnismuster spricht zugunsten reduzierter emotionaler Resonanz als Erklärung für den FLE. Ausgehend von den gewonnenen Daten konnten die anderen Erklärungsansätze für den FLE nicht unterstützt werden.

Ein möglicher Grund für die wenig signifikanten Ergebnisse in Bezug auf den FLE könnte im Experimentdesign liegen, bei dem der Manager eine Risikoentscheidung im Namen anderer trifft. Gemäß der Empathy-Gap-Hypothese⁴⁰⁹ ziehen Individuen ungenaue Rückschlüsse auf die Risikopräferenzen anderer Personen und stützen ihre Entscheidungen darauf. Wenn also ein Agent eine Risikoentscheidung im Namen eines Prinzipals trifft, beruht diese Entscheidung zu einem Teil auf seiner eigenen Risikopräferenz und zu einem anderen Teil auf der vermuteten Risikoeinstellung des Prinzipals,⁴¹⁰ was unter der typischen Annahme eines weniger risikoaversen Prinzipals in der Summe zu höherer Risikoakzeptanz führen sollte.⁴¹¹ Im Kontext des vorliegenden Experiments, wenn die Risikoakzeptanz im Allgemeinen erhöht ist, wird die Wahrscheinlichkeit für die Messung eines ausgeprägten FLE minimiert. Die Teilnehmer im Experiment wurden angewiesen, davon auszugehen, dass ihre Vergütung als verantwortlicher Manager vom finanziellen Ergebnis abhängt. Es bleibt unklar, in welchem Ausmaß diese Kontextbeschreibung half, den zuvor beschriebenen Effekt zu nivellieren. Außerdem könnte die Verwendung von realen Auszahlungen den FLE verstärken, weil damit stärkere Emotionen zu erwarten sind.⁴¹²

⁴⁰⁹ Siehe Loewenstein 1996.

⁴¹⁰ Vgl. Desmoulins-Lebeault und Meunier 2018, S. 247.

⁴¹¹ Darüber hinaus werden fremde Gewinne nicht in das mentale Konto des Entscheidungsträgers selbst integriert, was zu höherer Risikoakzeptanz führen sollte, vgl. Losecaat Vermeer et al. 2020, S. 8.

⁴¹² Im Durchschnitt gaben die Experimentteilnehmer auf einer siebenstufigen Skala Werte von nur 2,66 für positive und 2,79 für negative Gefühle an. Ergebnisse aus psychologischen Studien belegen, dass solche Werte ausreichend sind, um den FLE feststellen zu können. So untersuchten bspw. Hadjichristidis et al. 2017a (Studie 2 und 3) abergläubisches Verhalten und stellten einen starken FLE auch für diejenigen Teilnehmer fest, die für ihre Gefühle auf einer neunstufigen Skala Werte von unter drei angaben. Ferner konnte der FLE im Kontext der Risikobereitschaft auch ohne Experimentvergütung nachgewiesen werden, siehe u. a. Keysar et al. 2012 (Studie 2).

4. Experiment 2: Ethisches Verhalten⁴¹³

4.1 Einleitung

In Anknüpfung an die Untersuchung der Lesbarkeit und des Foreign-Language-Effekts (FLE) in Kapitel 3 werden in diesem Abschnitt die Ergebnisse eines weiteren Experiments präsentiert. Anstelle des Risikoverhaltens liegt der Schwerpunkt des Experiments jedoch auf dem ethischen Verhalten im Controlling-Kontext.

Ethische Entscheidungen spielen nicht selten eine immense gesellschaftliche Rolle mit weitreichenden Konsequenzen. Es vergeht kaum ein Jahr, ohne dass mindestens ein größerer Unternehmensskandal an die Öffentlichkeit gelangt.⁴¹⁴ Als jüngste Beispiele aus Deutschland können in diesem Zusammenhang der „Dieselgate“-Skandal von Volkswagen⁴¹⁵ und die Bilanzmanipulationen von Wirecard⁴¹⁶ genannt werden.

Zwei fundamentale Systeme dominieren die Ethiktheorie: Utilitarismus und Deontologie.⁴¹⁷ Die utilitaristische Perspektive verlangt, dass diejenige Aktion gewählt wird, die den Gesamtnutzen maximiert. Im Gegensatz dazu wird im Rahmen des deontologischen Ansatzes nicht primär auf die Konsequenzen geschaut. Stattdessen liegt der Fokus auf der Richtigkeit einer Handlung selbst ausgehend von bestimmten unverrückbaren Normen, Pflichten, Prinzipien o. ä. So sind die Ziele von gewinnorientierten Unternehmen am ehesten mit dem utilitaristischen Ansatz zu vereinbaren. Den Nonprofit-Organisationen könnte es hingegen wichtig sein, dass ihre Mitarbeiter im deontologischen Sinne bestimmte Handlungsprinzipien befolgen.

In der Literatur wird gestützt auf die Dual-System-Theorien (DST) argumentiert, dass die utilitaristischen Urteile vom kontrollierten System 2 geleitet werden und für die deontologischen Urteile das affektive System 1 verantwortlich ist.⁴¹⁸ In Übereinstimmung mit dieser Sichtweise konnte im Zusammengang mit ethischen Entscheidungen experimentell nachgewiesen werden, dass eine sehr hohe kognitive Belastung – die zur Auslösung des Cognitive-Load-Effekts (CLE) geeignet ist – die Wahrscheinlichkeit für die Wahl der utilitaristischen Alternative signifikant reduziert.⁴¹⁹ Ein ähnlicher Effekt sollte

⁴¹³ Das ethische Verhalten im Rahmen dieser Untersuchung bezieht sich auf Situationen, in denen die utilitaristische und die deontologische Maxime miteinander in Konflikt stehen (ethische Dilemmata).

⁴¹⁴ Vgl. Schwartz 2017, S. 6f. mit konkreten Beispielen für Unternehmensskandale in unterschiedlichen Ländern.

⁴¹⁵ Siehe Meyer-Fünffinger 2020.

⁴¹⁶ Siehe Bartz und Hesse 2020.

⁴¹⁷ Siehe u. a. Cottell und Perlin 1990, S. 3ff.

⁴¹⁸ Vgl. Cipolletti et al. 2016, S. 26 mit Verweis auf Greene 2008, S. 40f.

⁴¹⁹ Siehe Trémolière et al. 2012 (Experiment 2). Die Effektivität der Erzeugung des CLE durch verschiedene kognitive Belastungsaufgaben im Rahmen von Untersuchungen des ethischen Verhaltens wird ausführlich in Kapitel 4.5 diskutiert.

sich aufgrund höherer kognitiver Belastung infolge niedriger Lesbarkeit und damit einer stärkeren Beanspruchung von System 2 einstellen.⁴²⁰

Wie bereits im theoretischen Teil dieser Arbeit eingehend erläutert wurde, erstreckt sich der FLE vor allem auf ethisches Verhalten und moralische Urteile. Es konnte empirisch gezeigt werden, dass sich die Menschen tendenziell utilitaristischer entscheiden, wenn sie Informationen in einer fremden Sprache, statt in der eigenen Muttersprache verarbeiten.⁴²¹

Mit der zweiten Untersuchung im Rahmen dieser Arbeit wurde das Ziel verfolgt zu überprüfen, ob sich systematische Unterschiede im ethischen Verhalten der Teilnehmer aufgrund kognitiver Belastung (infolge niedriger Lesbarkeit) und/oder aufgrund des Einsatzes einer Fremdsprache im Controlling-Kontext ergeben. Weil einige Probanden der zusätzlichen kognitiven Belastung infolge niedriger Lesbarkeit ausgesetzt wurden, kann erwartet werden, dass der FLE nur bei einem hohen Lesbarkeitsniveau zu beobachten sein wird. Darüber hinaus wird im Rahmen dieser Studie näher auf die Wirkmechanismen hinter dem FLE eingegangen, indem die Rolle der Sprachkompetenz und die in der Literatur am meisten diskutierten Erklärungsansätze für den FLE untersucht werden.

Das Experiment ist als ein 2×2-Between-Subject-Design mit den wichtigsten unabhängigen Variablen Sprache (Fremd- vs. Mutter-) und Lesbarkeit (hoch vs. niedrig) aufgebaut. Die Experimentteilnehmer nahmen die Rolle eines Managers ein und erhielten entweder in ihrer Muttersprache (Deutsch) oder Fremdsprache (Englisch) einen fiktiven Managementbericht mit einer Fallstudie. In der Fallstudie ging es um eine Entscheidung für (utilitaristische Alternative) oder gegen (deontologische Alternative) die Schließung eines Tochterunternehmens. Ferner wurden die Teilnehmer gebeten, ihre Präferenzstärke zugunsten der Unternehmensschließung anzugeben.

Die Ergebnisse brachten keine systematischen Unterschiede im Entscheidungsverhalten der Teilnehmer in Abhängigkeit von der Lesbarkeit und/oder der Sprache hervor. Bei hoher im Vergleich zur niedrigen Lesbarkeit waren zwar die utilitaristischen Entscheidungen etwas häufiger vertreten, der Unterschied war jedoch nicht statistisch signifikant. Auch die Sprachwahl übte in aggregierter Form keinen statistisch bedeutenden Einfluss aus. Einzig für die Sprachkompetenz als unabhängige Variable konnten lesbarkeitsniveauübergreifend statistisch signifikante Ergebnisse gemessen werden. Dabei sprachen sich die Teilnehmer mit der mittelhohen Sprachkompetenzstufe wesentlich häufiger zugunsten der Unternehmensschließung aus als die übrigen Testpersonen. In Bezug auf die möglichen Erklärungsansätze für den FLE waren die Resultate nur mit demjenigen Erklärungsansatz konform, der auf der Abschwächung des Einflusses von moralisch-sozialen Normen beruht.

⁴²⁰ Ausführlich zu den DST und kognitiven Effekten der Lesbarkeit siehe Kapitel 2.2 bzw. Kapitel 2.3.1.

⁴²¹ Ausführlich zum FLE und den Erklärungsansätzen für dieses Phänomen siehe Kapitel 2.3.2.2.

Der Rest dieses Kapitel ist wie folgt aufgebaut: In Kapitel 4.2 werden die Hypothesen abgeleitet. Das Experimentdesign sowie die Teilnehmer werden in Kapitel 4.3 und die Experimentergebnisse in Kapitel 4.4 näher beleuchtet. Die Diskussion der Ergebnisse erfolgt schließlich in Kapitel 4.5.

4.2 Hypothesen

4.2.1 Hypothese 1

Bei Hypothese 1 steht die Auswirkung niedriger Lesbarkeit auf das ethische Verhalten im Zentrum der Betrachtung.

Die Annahme, dass ein utilitaristisches Entscheidungsverhalten im Kontext der ethischen Dilemmata auf kontrollierten und kognitiv anstrengenden Denkprozessen beruht, ist zentral in der modernen Ethikforschung.⁴²² Falls utilitaristische Entscheidungen ein höheres Niveau an Kontrolle erfordern und weniger automatisch sind, sollte die Häufigkeit dieser Entscheidungen gemäß dem CLE abnehmen, wenn kognitive Kapazitäten – z. B. aufgrund von konkurrierenden Denkaufgaben oder Zeitdruck – temporär beansprucht sind. Ein solcher Effekt konnte bspw. von Conway und Gawronski 2013 oder Trémolière et al. 2012 gemessen werden. Suter und Hertwig 2011 zeigten außerdem, dass sich die utilitaristischen Entscheidungen – zumindest für einige ihrer Dilemmata – verringern, wenn Individuen unter Zeitdruck gesetzt werden. Da ein niedriges Lesbarkeitsniveau die limitierten Ressourcen des Arbeitsgedächtnisses auf eine ähnliche Weise okkupiert und damit zur stärkeren Beanspruchung von System 2 beiträgt, lautet Hypothese 1 wie folgt:

H1: Entscheidungsträger werden ein weniger ausgeprägtes utilitaristisches Entscheidungsverhalten zeigen, wenn sie einen Controlling-Bericht mit einem niedrigen im Vergleich zu einem hohen Lesbarkeitsniveau verarbeiten müssen.

4.2.2 Hypothese 2

Im Folgenden wird der Einfluss der Sprachwahl auf das ethische Verhalten näher betrachtet und es wird begründet, warum vermutet wird, dass der FLE in diesem Kontext abhängig von einer hohen Lesbarkeitsstufe ist.

Es existieren überzeugende empirische Belege dafür, dass Menschen weniger emotional reagieren, wenn sie Informationen in einer Fremdsprache im Vergleich zu ihrer Muttersprache verarbeiten. Demnach könnte eine fremde Sprache gemäß dem „blunted-deontology account“⁴²³ das deontologische Entscheidungsverhalten aufgrund emotionaler

⁴²² Vgl. Trémolière und Bonnefon 2014, S. 923.

⁴²³ Vgl. Hayakawa et al. 2017b, S. 1388.

und psychologischer Distanz abschwächen.⁴²⁴ Im Kontext (fiktiver) ethischer Dilemmata zeigten z. B. Aguilar et al. 2013, dass eine höhere Abstraktionsstufe zur höheren Bereitschaft, ein Menschenleben zu opfern, um mehrere andere Menschenleben zu retten, führt. Auf eine ähnliche Weise sollte eine Fremdsprache über die Schaffung emotionaler und psychologischer Distanz den Fokus der Entscheidung stärker auf die Konsequenzen legen, weil das affektive System 1 dann weniger stark am Entscheidungsprozess beteiligt ist.

Alternativ kann das deontologische Verhalten auch abgeschwächt werden, indem infolge des Einsatzes einer Fremdsprache die moralisch-sozialen Normen für den Entscheidungsträger weniger stark präsent sind.

Gemäß dem „heightened-utilitarianism account“⁴²⁵ fördert eine Fremdsprache das utilitaristische Verhalten mittels Aktivierung von System 2. Dieser Erklärungsansatz basiert auf dem Disfluency-Effekt (DFE). Demnach hemmt eine Fremdsprache den mentalen Fluss und sollte im Sinne einer metakognitiven Verarbeitungsschwierigkeit das analytische System 2, das mit dem utilitaristischen Verhalten assoziiert wird, aktivieren.

Empirische Belege dafür, dass der FLE von einer hohen Lesbarkeit abhängig ist, können den existierenden psychologischen Studien entnommen werden. In den Studien zum ethischen Verhalten wird üblicherweise die „Footbridge“-Version des Trolley-Problems⁴²⁶ oder eine Abwandlung davon verwendet. Diese Version des Trolley-Problems besteht lediglich aus sechs Sätzen mit 102 Wörtern und weist gemäß dem Fog-Index einen Lesbarkeitswert von 9,2 auf.⁴²⁷ Insgesamt lautet Hypothese 2 wie folgt (siehe auch Abbildung 16):

H2: Nur bei einem hohen Lesbarkeitsniveau werden die Entscheidungsträger stärker utilitaristisch geprägte Entscheidungen treffen, wenn sie einen Controlling-Bericht in einer Fremdsprache, statt in ihrer Muttersprache verarbeiten müssen.

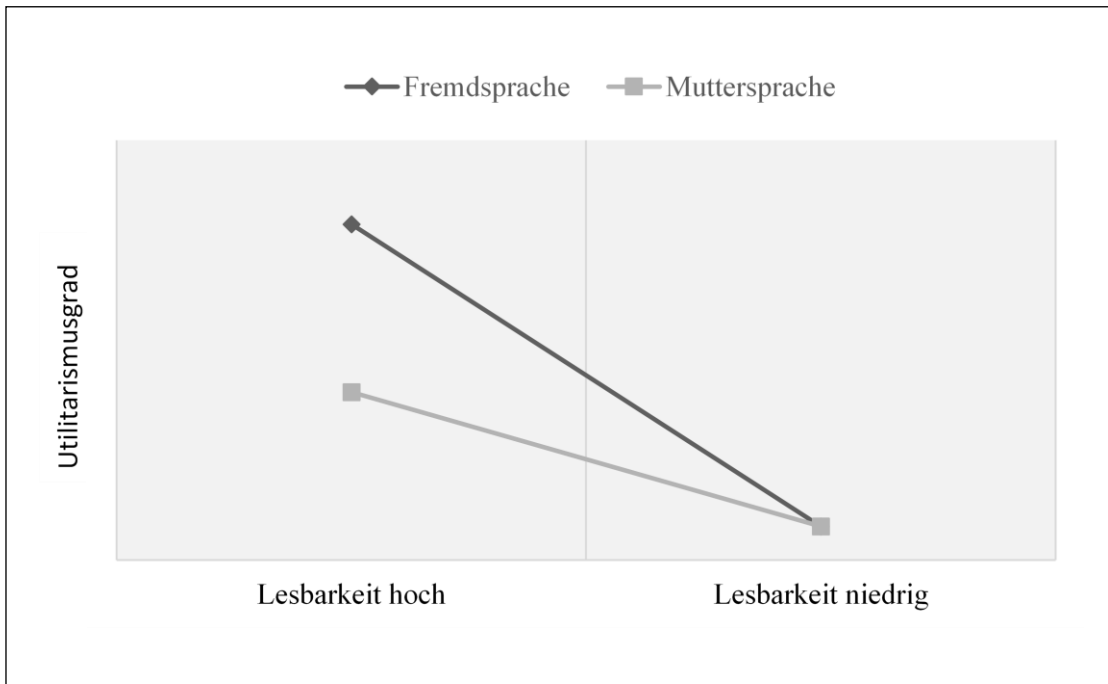
⁴²⁴ Vgl. Čavar und Tytus 2018, S. 18 mit Verweis auf Costa et al. 2014b.

⁴²⁵ Vgl. Hayakawa et al. 2017b, S. 1388.

⁴²⁶ Siehe Cicolletti et al. 2016, S. 28. Die Übersetzung dieser Version des Trolley-Problems ins Deutsche befindet sich in Kapitel 2.3.2.2.1.

⁴²⁷ Dieser Wert bedeutet, dass der Text selbst von einem Kind verstanden werden kann, vgl. Li 2008, S. 225.

Abbildung 16: Visualisierung von H1 und H2 (Experiment 2)



Quelle: Eigene Darstellung.

4.2.3 Hypothese 3

Der Einfluss der Sprachwahl auf das ethische Verhalten müsste von der Sprachkompetenz abhängig sein.⁴²⁸ In Übereinstimmung mit Kapitel 3.2.3 wird Hypothese 3 wie folgt formuliert:

H3: *Der FLE – wie unter H2 formuliert – gilt nur für mittelhohe Sprachkompetenzstufen.*

4.3 Experimentdesign und Teilnehmer

4.3.1 Experimentdesign

4.3.1.1 Ablauf und Aufgabe

Bei dem Experiment handelt es sich um ein 2×2-Between-Subject-Design mit einer randomisierten Zuweisung von Teilnehmern in die verschiedenen Treatment-Gruppen. Das Experiment wurde zeitgleich in vier separaten Räumen im Sommersemester 2019

⁴²⁸ Vgl. Costa et al. 2014b, S. 4.

durchgeführt. Der Ablauf des Experiments stimmt mit demjenigen aus Experiment 1 überein.⁴²⁹

Der erste Teil des Experiments bestand aus zwei Fallstudien – einer Hauptfallstudie und einer Kontrollfallstudie –, die in randomisierter Reihenfolge ausgeteilt wurden,⁴³⁰ und einem Fragebogen in ausgedruckter Form. Nachdem Teil eins von allen Teilnehmern bearbeitet worden war, wurde dieser eingesammelt und Teil zwei wurde ausgeteilt. Der zweite Teil bestand aus Verständnisfragen in Bezug auf die Hauptfallstudie und demografischen bzw. sprachbezogenen Abfragen.⁴³¹ Schließlich mussten die Teilnehmer in den fremdsprachlichen Gruppen einen kurzen Englischtest innerhalb einer Zeitfrist von 15 Minute absolvieren.⁴³²

In der Hauptfallstudie ging es um eine Entscheidung für oder gegen die Schließung eines Tochterunternehmens, um mit den so freiwerdenden finanziellen Mitteln ein anderes aufgrund eines Fabrikbrands in eine Notlage geratenes Tochterunternehmen zu retten. Weil das notleidende Tochterunternehmen c. p. deutlich mehr Mitarbeiter als das andere Tochterunternehmen beschäftigt, ist die Unternehmensschließung gemäß der utilitaristischen Maxime angebracht.⁴³³ Gleichzeitig würde die Unternehmensschließung implizieren, dass die Mitarbeiter dieses Tochterunternehmens unverschuldet ihre Arbeit verlieren, was dem deontologischen Ansatz zuwiderläuft.⁴³⁴ Die Probanden wurden gebeten, sich zum einen für oder gegen die Unternehmensschließung zu entscheiden und zum anderen ihre Präferenzstärke für die Unternehmensschließung anzugeben.⁴³⁵

⁴²⁹ Siehe 3.3.1.1.

⁴³⁰ Die Reihenfolge, in der die Fallstudien präsentiert wurden, hatte keinen systematischen Einfluss auf das ethische Entscheidungsverhalten der Teilnehmer im Experiment ($p > 0,663$).

⁴³¹ Die sprachbezogenen Informationen wurden in Anlehnung an den „language history questionnaire“ von Li et al. 2006 erhoben.

⁴³² Siehe Kapitel 3.3.1.2 und Anhang I.

⁴³³ In den Studien zum ethischen Verhalten steht der Gesamtnutzen in der utilitaristischen Alternative i. d. R. im Verhältnis fünf zu eins, siehe Trémolière und Bonnefon 2014. Aus diesem Grund wurde auch in der vorliegenden Untersuchung festgelegt, dass mit jedem entlassenen Mitarbeiter infolge der Unternehmensschließung fünf andere Mitarbeiter des notleidenden Tochterunternehmens ihre Arbeit behalten können.

⁴³⁴ Die deontologische Perspektive wurde mit dem zusätzlichen Hinweis verstärkt, dass die entlassenen Mitarbeiter mit ihrem Arbeitsplatz auch ihre Krankenversicherung verlieren werden. Zudem wurde erwähnt, dass die Entscheidung über die Schließung des Tochterunternehmens persönlich im Rahmen einer Mitarbeiterversammlung verkündet werden muss.

⁴³⁵ Um alle ökonomischen Faktoren auszuklammern und das zugrunde liegende Entscheidungsproblem auf eine ethisch-moralische Überlegung zu reduzieren, wurde in der Fallstudie deutlich hervorgehoben, dass die beiden möglichen Entscheidungsalternativen ökonomisch betrachtet äquivalent sind.

Des Weiteren beinhaltete die Experimentaufgabe eine Kontrollfallstudie, in der eine alltägliche Situation beschrieben wurde, wobei die utilitaristische und die deontologische Maxime nicht miteinander in Konflikt standen und es eine dominante Entscheidungsalternative gab.⁴³⁶

Die Experimentteilnehmer erhielten keine (variable) Experimentvergütung, weil es in der Hauptfallstudie keine normativ richtige Antwort gab und somit eine Anknüpfung der Entscheidungen an ein Vergütungssystem zu nicht intendierten Anreizen führen könnte.

4.3.1.2 Unabhängige Variablen und Kontrollvariablen

Die wichtigsten unabhängigen Variablen Sprache, Sprachkompetenz und Lesbarkeit wurden identisch zu Experiment 1 konstruiert.⁴³⁷ In Anhang III sind einige Beispiele für die Lesbarkeitsmanipulationen aufgeführt.

Das ethische Verhalten kann von bestimmten individuellen Faktoren bestimmt werden, die als Kontrollvariablen in der vorliegenden Untersuchung berücksichtigt werden sollen. Obwohl das Geschlecht in der ethischen Entscheidungsforschung bereits sehr ausführlich untersucht wurde, sind die Ergebnisse gemischt und nicht eindeutig.⁴³⁸ In einigen Studien wurde herausgefunden, dass sich Frauen im Durchschnitt ethischer verhalten als Männer. Dagegen sind Männer strikter und konsistenter in ihrem ethischen Entscheidungsverhalten.⁴³⁹ In der Studie von Beekun et al. 2010 verließen sich Frauen mehr auf das Prinzip der Gerechtigkeit und den Utilitarismus bei ethischen Entscheidungen, während Männer nur vom Gerechtigkeitsgedanken geleitet waren. Demgegenüber verhielten sich in der Untersuchung von Borkowski und Ugras 1992 die männlichen Probanden utilitaristischer als weibliche. Andere Studien wiederum fanden keine geschlechtsspezifischen Unterschiede.⁴⁴⁰ Borkowski und Ugras 1992 zeigten, dass sich ethische Einstellungen mit dem Alter ändern können und dass MBA-Absolventen utili-

⁴³⁶ In der vorliegenden Studie wurde die Kontrollfallstudie „Train or Bus“ von Geipel et al. 2015b verwendet und für die muttersprachlichen Gruppen ins Deutsche übersetzt. Bis auf drei Teilnehmer haben sich alle Probanden zugunsten der dominanten Alternative entschieden. Von den drei Teilnehmern, die sich gegen die dominante Alternative entschieden haben, wurde eine Person aufgrund der Ausschlusskriterien (siehe Kapitel 4.3.2) aus der Hauptanalyse aussortiert und die anderen beiden Probanden übten keinen signifikanten Einfluss auf die statistischen Gesamtergebnisse aus.

⁴³⁷ Siehe Kapitel 3.3.1.2. Die Itemanalyse ergab vergleichbare Werte für die Modellgüte des C-Tests wie in Experiment 1.

⁴³⁸ Vgl. Loe et al. 2000, S. 187.

⁴³⁹ Vgl. Craft 2013, S. 230f.

⁴⁴⁰ Siehe u. a. Chan und Leung 2006; Sweeney und Costello 2009.

taristischer an ethische Probleme herangehen als Bachelor-Studenten aus dem Fach Betriebswirtschaftslehre. Schließlich ist utilitaristisches Entscheidungsverhalten eher von Individuen mit einem rationalen kognitiven Stil zu erwarten.⁴⁴¹

4.3.1.3 Abhängige Variablen

Das ethische Verhalten der Teilnehmer wurde in zwei Schritten gemessen. Zunächst mussten sich die Probanden für (utilitaristische Alternative) oder gegen (deontologische Alternative) die Unternehmensschließung entscheiden. Um ein metrisches Maß in Form der Präferenzstärken zugunsten der utilitaristischen Alternative zu erhalten, wurden die Testpersonen des Weiteren gebeten, auf einer Skala von 1 (= zwingend) bis 7 (= keinesfalls) anzugeben, inwiefern die aktive Unternehmensschließung aus ihrer Sicht geboten ist. Weil die Präferenzstärken zugunsten der utilitaristischen Alternative den Einsatz von parametrischen Testverfahren – falls bestimmte statistische Voraussetzungen, wie etwa die Normalverteilung und Varianzhomogenität, erfüllt sind – erlauben, basieren die Hypothesentest in Kapitel 4.4.2 auf den Präferenzstärken als abhängige Variable.

4.3.2 Teilnehmer

Die Mehrheit der Teilnehmer waren Studierende der Wirtschaftswissenschaften⁴⁴² aus einer mittelgroßen Universität in Deutschland. Alle Probanden nahmen auf freiwilliger Basis teil und die Anonymität der Teilnehmer wurde stets gewährleistet. Deskriptive Statistiken zu den Teilnehmern und ihrer Verteilung zwischen den vier Experimentgruppen sind in Tabelle 10 aufgeführt. Die Versuchspersonen waren im Durchschnitt 24,6 Jahre alt, 37,3 % waren weiblich und 45,7 % befanden sich zum Zeitpunkt des Experiments im Bachelorstudium. In Bezug auf die Kontrollvariablen Geschlecht und Faith in Intuition gab es keine wesentlichen Unterschiede zwischen den Experimentgruppen. Das Alter war zwar zwischen den Experimentgruppen nicht uniform verteilt, sein Einfluss auf die abhängige Hauptvariable im Experiment war jedoch nicht feststellbar.⁴⁴³

⁴⁴¹ Siehe u. a. Bartels 2008. Conway und Gawronski 2013 fanden insbesondere heraus, dass utilitaristische Präferenzen mit der persönlichen Eigenschaft Need for Cognition korrelieren. Analog zu Experiment 1 wurde das Konstrukt Need for Cognition mithilfe der Skala von Beißert et al. 2015 gemessen und die (verkürzte) Version der Skala Faith in Intuition wurde aus der Studie von Epstein et al. 1996 entnommen.

⁴⁴² Daneben gab es eine kleine Anzahl von Studierenden, die in mathematischen und ingenieurwissenschaftlichen Fächern eingeschrieben waren.

⁴⁴³ Siehe Tabelle 20, Anhang III.

Tabelle 10: Deskriptive Statistiken zu den Teilnehmern (Experiment 2)

| | Lesbarkeit | | | | Gesamt |
|--|-------------|-------------|-------------|-------------|-------------|
| | Hoch | | Niedrig | | |
| | Sprache | | Sprache | | |
| | Fremd- | Mutter- | Fremd- | Mutter- | |
| Geschlecht^a | | | | | |
| Weiblich | 13 (41,9 %) | 13 (39,4 %) | 7 (24,1 %) | 14 (42,4 %) | 47 (37,3 %) |
| Männlich | 18 (58,1 %) | 20 (60,6 %) | 22 (75,9 %) | 19 (57,6 %) | 79 (62,7 %) |
| Studiengang | | | | | |
| Bachelor | 17 (53,1 %) | 12 (36,4 %) | 11 (37,9 %) | 18 (54,5 %) | 58 (45,7 %) |
| Master | 15 (46,9 %) | 21 (63,6 %) | 18 (62,1 %) | 15 (45,5 %) | 69 (54,3 %) |
| Alter (<i>M</i>) | 23,5 | 24,7 | 25,9 | 24,4 | 24,6 |
| Faith in Intuition ^b (<i>M</i>) | 4,6 | 4,7 | 4,8 | 4,7 | 4,7 |
| Need for Cognition ^b (<i>M</i>) | 4,9 | 5,1 | 4,7 | 4,6 | 4,8 |

^a Ein Teilnehmer in Gruppe 1 gab das Geschlecht nicht an.

^b Faith in Intuition und Need for Cognition sind Konstrukte zur Messung der Tendenz der Teilnehmer, sich mehr auf Intuitionen (System 1) bzw. das logische Denken (System 2) zu verlassen. Die (verkürzte) Version der Skala Faith in Intuition wurde aus der Studie von Epstein et al. 1996 entnommen und Need for Cognition wurde mithilfe der Skala von Beißert et al. 2015 gemessen.

Quelle: Eigene Darstellung.

Need for Cognition ist zwischen den Experimentgruppen nicht gleich verteilt, übt einen signifikanten Einfluss auf die Hauptvariable im Experiment aus und wird daher in der statistischen Hauptanalyse als Kovariate berücksichtigt.⁴⁴⁴

Einige der Teilnehmer mussten aus verschiedenen Gründen aus der statistischen Hauptanalyse ausgeschlossen werden. Zunächst wurden Multiple-Choice-Fragen eingesetzt, um sicherzustellen, dass die Teilnehmer zum einen die schwierige finanzielle Situation des aufgrund eines Fabrikbrands in eine Notlage geratenen Tochterunternehmens erkannten und zum anderen die Anzahl der Mitarbeiter bei beiden Tochterunternehmen angeben konnten.

⁴⁴⁴ Die unterschiedliche Verteilung der Need-for-Cognition-Werte zwischen den Experimentgruppen als mögliche Alternativklärung für das Zustandekommen der Studienergebnisse wird außerdem in Kapitel 4.4.3.1 diskutiert.

Zusätzlich zu den Manipulation-Checks in Form von Multiple-Choice-Fragen wurden, identisch zu Experiment 1, weitere Ausschlusskriterien basierend auf dem selbstberichteten Verständnis der Hauptfallstudie sowie dem sprachlichen Hintergrund der Teilnehmer eingesetzt.⁴⁴⁵

Deskriptive Statistiken zu den oben genannten Ausschlusskriterien sowie deren Verteilung zwischen den Experimentgruppen kann Tabelle 19 in Anhang III entnommen werden. In allen Gruppen ist die Ausschlussquote annähernd gleich und beträgt circa 15 %.

4.4 Ergebnisse

4.4.1 Deskriptive Statistiken

Tabelle 11 enthält deskriptive Statistiken zu den beiden abhängigen Variablen, die das ethische Verhalten der Teilnehmer messen. Konkret handelt es sich um (1) die Auswahlentscheidung für (utilitaristische Alternative) oder gegen (deontologische Alternative) die Unternehmensschließung und (2) die Präferenzstärke zugunsten der utilitaristischen Alternative.⁴⁴⁶ Da die Auswahlentscheidung eine kategoriale Variable darstellt, werden in Panel A nur absolute Zahlen und entsprechende Prozentangaben für jede Experimentgruppe dargestellt. Die Präferenzstärken, welche die wichtigste abhängige Variable repräsentieren, sind metrisch und ermöglichen die Berechnung von deskriptiven Statistiken. Für diese Variable werden für jede Experimentgruppe der Mittelwert und die Standardabweichung in Panel B gezeigt.

Anhand der deskriptiven Ergebnisse in Bezug auf die Auswahlentscheidung ist eine leichte Tendenz zum stärker utilitaristisch ausgeprägten Verhalten bei hoher im Vergleich zur niedrigen Lesbarkeit zu beobachten. Zusammengenommen bevorzugten die Probanden in den Gruppen mit hoher Lesbarkeit die utilitaristische Alternative etwas häufiger (67,7 %) als ihre Pendants in den Gruppen mit niedriger Lesbarkeit (58,1 %). Die durchschnittlichen Präferenzstärken waren annähernd gleich (hohe Lesbarkeit: 3,96 vs. niedrige Lesbarkeit: 3,94). Insgesamt liefern die deskriptiven Ergebnisse keine überzeugende Evidenz zugunsten von H1.

In H2 wird postuliert, dass die Entscheidungsträger nur bei hoher Lesbarkeit utilitaristischere Entscheidungen treffen werden, wenn sie einen Controlling-Bericht in einer Fremdsprache im Vergleich zu ihrer Muttersprache verarbeiten müssen. Diese Hypothese wird auf Basis der vorliegenden Daten nicht unterstützt und es ist sogar ein umgekehrtes Verhältnis festzustellen.

⁴⁴⁵ Siehe Kapitel 3.3.2.

⁴⁴⁶ Aufgrund der verwendeten Messskala gilt: Je niedriger der Wert, desto größer ist die Präferenzstärke zugunsten der utilitaristischen Alternative.

Tabelle 11: Deskriptive Statistiken zu den abhängigen Variablen (Experiment 2)

| Panel A: Auswahlentscheidung ^a (absolute Zahl [Prozentangabe]) | | | | | |
|---|----------------|----------------|----------------|----------------|----------------|
| | Lesbarkeit | | | | Gesamt |
| | Hoch | | Niedrig | | |
| | Sprache | | Sprache | | |
| | Fremd- | Mutter- | Fremd- | Mutter- | |
| Utilitaristische Option | 19 [59,4 %] | 25 [75,8 %] | 15 [51,7 %] | 21 [63,6 %] | 80 [63,0 %] |
| Deontologische Option | 13 [40,6 %] | 8 [24,2 %] | 14 [48,3 %] | 12 [36,4 %] | 47 [37,0 %] |

| Panel B: Präferenzstärke ^b (Mittelwert [Standardabweichung]) | | | | | |
|---|----------------|----------------|----------------|----------------|----------------|
| | Lesbarkeit | | | | Gesamt |
| | Hoch | | Niedrig | | |
| | Sprache | | Sprache | | |
| | Fremd- | Mutter- | Fremd- | Mutter- | |
| Präferenzstärke | 3,97 [1,47] | 3,94 [1,14] | 3,79 [1,57] | 4,09 [1,16] | 3,95 [1,33] |

^a Die Auswahlentscheidung ist eine Dummy-Variable und kann in Abhängigkeit davon, ob sich der Teilnehmer für (utilitaristische Alternative) oder gegen (deontologische Alternative) die Unternehmensschließung entschied, zwei mögliche Zustände einnehmen.

^b Die Präferenzstärke zugunsten der utilitaristischen Alternative repräsentiert die Bereitschaft der Testperson, die Unternehmensschließung durchzuführen. Je niedriger der Wert, desto größer ist die Präferenzstärke zugunsten der utilitaristischen Alternative.

Quelle: Eigene Darstellung.

Bei dem hohen Lesbarkeitsniveau wurde die utilitaristische Alternative in der fremdsprachlichen Gruppe weniger häufig gewählt (59,4 %) als in der muttersprachlichen Gruppe (75,8 %). Dasselbe gilt für die Präferenzstärken innerhalb der hohen Lesbarkeitsstufe (Fremdsprache: 3,97 vs. Muttersprache: 3,94).

Werden lesbarkeitsniveauübergreifend die Sprachgruppen verglichen, stellt sich das folgende Ergebnis ein: Bei der Auswahlentscheidung wählen die Teilnehmer in der fremdsprachlichen Gruppe die utilitaristische Alternative wieder weniger häufig als ihre Pendants in der muttersprachlichen Gruppe (55,7 % vs. 69,7 %). Das sensiblere Maß für das utilitaristische Verhalten in Form der Präferenzstärken deutet hingegen auf den FLE hin (Fremdsprache: 3,88 vs. Muttersprache: 4,02). Eine mögliche Erklärung für dieses Ergebnismuster könnte darin begründet liegen, dass die Wirkung der Lesbarkeitsmanipulationen im Experiment deutlich schwächer ausgefallen war als erwartet.⁴⁴⁷

⁴⁴⁷ Die nicht ausreichende Wirkung der Lesbarkeitsmanipulationen als mögliche Begründung für das Zustandekommen der Studienergebnisse wird noch ausführlich in Kapitel 4.5 diskutiert.

Schließlich besagt H3, dass der FLE – wie unter H2 formuliert – nur für mittelhohe Sprachkompetenzstufen gilt. Die Ergebnisse auf deskriptiver Ebene liefern einige Hinweise zugunsten dieser Hypothese (Zahlen sind nicht abgebildet). Innerhalb der hohen Lesbarkeitsstufe entschied sich die überwiegende Mehrheit der Teilnehmer mit dem mittelhohen Sprachkompetenzniveau für die Unternehmensschließung (81,8 %), muttersprachliche Probanden präferierten die utilitaristische Alternative am zweithäufigsten (75,8 %) und die niedrigste Präferenzrate wiesen die Teilnehmer mit einem geringen Sprachkompetenzniveau auf (47,6 %). Das gleiche Ergebnismuster lässt sich für die Präferenzstärken beobachten. Innerhalb der hohen Lesbarkeitsstufe gaben die Probanden mit dem mittelhohen Sprachkompetenzniveau im Durchschnitt die niedrigsten Präferenzstärken an,⁴⁴⁸ gefolgt von muttersprachlichen Teilnehmern und Probanden mit einem geringen Sprachkompetenzniveau (3,18; 3,94 und 4,38 in entsprechender Reihenfolge).

Werden lesbarkeitsniveauübergreifend die Sprachkompetenzgruppen verglichen, verstärken sich die oben beschriebenen Unterschiede (3,28; 4,02 und 4,14 in entsprechender Reihenfolge). Dieses Resultat deutet ebenfalls darauf hin, dass die Wirkung der Lesbarkeitsmanipulationen im Experiment nicht ausreichend war.⁴⁴⁹

4.4.2 Hypothesentests

4.4.2.1 H1 und H2

Die statistischen Tests der Hypothesen H1 und H2 wurden mithilfe der ANCOVA durchgeführt und als Follow-up-Tests zur Überprüfung der genauen Gruppenunterschiede wurde die Kontrastanalyse verwendet.⁴⁵⁰ Need for Cognition diente als Kovariate und die folgenden erklärenden Variablen wurden herangezogen: Sprache, Lesbarkeit und ihre Interaktion. Die Präferenzstärken stellten die abhängige Variable dar.

Die Ergebnisse der ANCOVA zusammen mit den Ergebnissen der Kontrastanalyse sind in Tabelle 12 aufgeführt. Keine der unabhängigen Variablen und insbesondere Sprache, Lesbarkeit und deren Interaktionsterm weisen einen signifikanten Effekt auf die Präferenzstärken auf (alle $p > 0,450$; Panel A).

⁴⁴⁸ Zur Erinnerung: Ein niedriger Wert bedeutet eine stärkere Präferenz zugunsten der utilitaristischen Alternative.

⁴⁴⁹ Mit der Auswahlentscheidung als abhängige Variable weisen die muttersprachlichen Teilnehmer lesbarkeitsniveauübergreifend ein am stärksten utilitaristisch ausgeprägtes Verhalten auf. Dabei wurde die Unternehmensschließung in den einzelnen Sprachkompetenzstufen wie folgt gewählt: 69,7 % (muttersprachlich); 61,1 % (mittelhoch) und 53,5 % (gering).

⁴⁵⁰ Zu den beiden Analysemethoden im Rahmen von Experimenten siehe Sedlmeier und Renkewitz 2013.

Tabelle 12: Statistische Analyseergebnisse – Teil 1 (Experiment 2)

| Panel A: Haupteffekte und Interaktion der ANCOVA | | | | |
|---|----|---------------|--------|--------|
| Prädiktorvariable (Präferenzstärke ^a) | Fg | Quadrat-summe | F-Wert | p-Wert |
| Sprache ^b | 1 | 0,53 | 0,30 | 0,586 |
| Lesbarkeit ^c | 1 | 0,00 | 0,00 | 0,982 |
| Need for Cognition ^d | 1 | 1,01 | 0,57 | 0,454 |
| Sprache × Lesbarkeit | 1 | 1,00 | 0,56 | 0,457 |

| Panel B: Ergebnisse der Kontrastanalyse mit ANCOVA | | | | |
|--|-----|-----------|--------|--------|
| Kontrast (Präferenzstärke) | Fg | Differenz | t-Wert | p-Wert |
| Lesbarkeit hoch: Sprache: Fremd- vs. Mutter | 122 | 0,05 | 0,15 | 0,882 |
| Lesbarkeit niedrig: Sprache: Fremd- vs. Mutter | 122 | -0,31 | -0,90 | 0,371 |

Alle *p*-Werte sind zweiseitig. *, **, *** repräsentieren eine Signifikanz auf den Niveaus 10 %, 5 % und 1 %.

^aDie Präferenzstärke zugunsten der utilitaristischen Alternative repräsentiert die Bereitschaft der Testperson, die Unternehmensschließung durchzuführen.

^bDie Sprache ist eine Dummy-Variable mit den beiden Ausprägungen: Muttersprache (Deutsch) und Fremdsprache (Englisch).

^cDie Lesbarkeit ist eine Dummy-Variable mit den beiden Ausprägungen hohe und niedrige Lesbarkeit. Die Lesbarkeit wurde mithilfe von „Plain English Handbook“ von SEC 1998 mittels ausgewählter linguistischer und formaler Gestaltung manipuliert.

^dNeed for Cognition ist ein Konstrukt zur Messung der Tendenz der Teilnehmer, sich mehr auf das logische Denken (System 2) zu verlassen. Diese Variable wurde mithilfe der Skala von Beißert et al. 2015 gemessen.

Quelle: Eigene Darstellung.

In Übereinstimmung dazu ergab die Kontrastanalyse keinen signifikanten Unterschied zwischen den Sprachgruppen bei hoher Lesbarkeit ($p = 0,882$, Panel B). Offensichtlich übten die Sprache (in aggregierter Form) und Lesbarkeit keinen Einfluss auf das ethische Entscheidungsverhalten der Teilnehmer aus, weshalb die Hypothesen H1 und H2 abgelehnt werden müssen.

4.4.2.2 H3

Das Ziel dieses Kapitels besteht darin, den FLE weiter zu untersuchen, indem die Sprachkompetenz der Experimentteilnehmer berücksichtigt wird. In Hypothese H3 wird postuliert, dass der FLE – wie unter H2 formuliert – nur für mittelhohe Sprachkompetenzstufen gilt.

Tabelle 13: Statistische Analyseergebnisse – Teil 2 (Experiment 2)

| Panel A: Haupteffekte und Interaktion der ANCOVA | | | | |
|---|----|--------------|--------|--------|
| Prädiktorvariable (Präferenzstärke ^a) | Fg | Quadratsumme | F-Wert | p-Wert |
| Sprachkompetenz ^b | 2 | 9,96 | 2,91 | 0,058* |
| Lesbarkeit ^c | 1 | 0,07 | 0,04 | 0,837 |
| Need for Cognition ^d | 1 | 3,02 | 1,77 | 0,186 |
| Sprachkompetenz × Lesbarkeit | 2 | 3,52 | 1,03 | 0,360 |

| Panel B: Ergebnisse der Kontrastanalyse mit ANCOVA | | | | |
|--|-----|-----------|--------|---------|
| Kontrast (Präferenzstärke) | Fg | Differenz | t-Wert | p-Wert |
| Sprachkompetenz: | | | | |
| Mittelhoch vs. gering | 120 | 0,96 | 2,52 | 0,013** |
| Mittelhoch vs. muttersprachlich | 120 | 0,79 | 2,19 | 0,030** |
| Gering vs. muttersprachlich | 120 | -0,18 | -0,69 | 0,493 |

| Panel C: Geschätzte marginale Mittelwerte der Präferenzstärke | | | |
|---|-----------------|------------|--------|
| | Sprachkompetenz | | |
| | Mutterspr. | Mittelhoch | Gering |
| Präferenzstärke | 4,01 | 3,22 | 4,19 |

Alle *p*-Werte sind zweiseitig. *, **, *** repräsentieren eine Signifikanz auf den Niveaus 10 %, 5 % und 1 %.

^a Die Präferenzstärke zugunsten der utilitaristischen Alternative repräsentiert die Bereitschaft der Testperson, die Unternehmensschließung durchzuführen. Je niedriger der Wert, desto größer ist die Präferenzstärke zugunsten der utilitaristischen Alternative.

^b Die Sprachkompetenz wurde anhand eines C-Tests (siehe Kapitel 3.3.1.2 und Anhang I) gemessen. Die Teilnehmer wurden in die drei Sprachkompetenzgruppen eingeteilt: muttersprachlich; mittelhoch; gering.

^c Die Lesbarkeit ist eine Dummy-Variable mit den beiden Ausprägungen hohe und niedrige Lesbarkeit. Die Lesbarkeit wurde mithilfe von „Plain English Handbook“ von SEC 1998 mittels ausgewählter linguistischer und formaler Gestaltung manipuliert.

^d Need for Cognition ist ein Konstrukt zur Messung der Tendenz der Teilnehmer, sich mehr auf das logische Denken (System 2) zu verlassen. Diese Variable wurde mithilfe der Skala von Beißert et al. 2015 gemessen.

Quelle: Eigene Darstellung.

Um die Hypothese H3 zu testen, wurde – ähnlich dem Vorgehen von Oganian et al. 2016 – im ursprünglichen ANCOVA-Modell die unabhängige Variable Sprache durch Sprachkompetenz ersetzt (siehe Tabelle 13, Panel A). Der Interaktionsterm ist nicht signifikant ($F = 1,03$; $p = 0,360$), was die Ablehnung von H3 zur Folge hat. Das insignifikante Ergebnis der Lesbarkeit in Kombination mit dem signifikanten Haupteffekt

der Sprachkompetenz ($F = 2,91$; $p = 0,058$) lässt jedoch auf einen FLE schließen, weil gleichzeitig die Präferenzstärken lesbarkeitsniveauübergreifend bei der mittelhohen Sprachkompetenzstufe im Vergleich zu den anderen Sprachkompetenzgruppen im Durchschnitt signifikant geringer ausfielen (siehe Tabelle 13, Panel B und C).⁴⁵¹ Es ist plausibel anzunehmen, dass das niedrige Lesbarkeitsniveau im Experiment keinen ausreichenden CLE ausübte und somit der FLE in Form eines Haupteffekts zu beobachten ist.

4.4.3 Ergänzende Analysen

4.4.3.1 Manipulation-Checks und alternative Erklärungen

In diesem Kapitel werden die Validität der Lesbarkeitsmanipulationen überprüft und potenzielle Alternativerklärungen für die gewonnenen Ergebnisse adressiert.

Die Lesbarkeitsmanipulationen wurden, identisch wie in Experiment 1, überprüft.⁴⁵² Da alle Fragen zu den Lesbarkeitsmanipulationen stark korrelieren (alle Pearson-Korrelationskoeffizienten $r > 0,60$), wurde aus diesen Fragen ein Mittelwert gebildet, der dann als Lesbarkeitsmaß verwendet wurde. Ergebnisse der ANOVA (Zahlen sind nicht abgebildet) mit dem Lesbarkeitsmaß als abhängige Variable und Sprache, Lesbarkeit sowie deren Interaktion als unabhängige Variablen zeigten einen insignifikanten Interaktionsterm ($F = 0,07$; $p = 0,789$; zweiseitig). Die Sprache und Lesbarkeit weisen signifikante Haupteffekte auf ($F = 5,09$ und $F = 11,04$; $p = 0,026$ und $p = 0,001$: jeweils zweiseitig).⁴⁵³ Der signifikante Einfluss der Sprache auf das Lesbarkeitsmaß lässt einen möglichen FLE aufgrund des DFE, der im nachfolgenden Abschnitt genauer untersucht wird, vermuten.

Ein potenzieller Erklärungsfaktor für den signifikanten Haupteffekt der Sprachkompetenz könnte ein zu geringes Verständnis einiger Teilnehmer – insbesondere solcher mit dem geringen Sprachkompetenzniveau – sein. Um diese potenzielle Erklärung auszuschließen, wurde eine ANOVA mit dem selbstberichteten Verständnis der Fallstudie in Abhängigkeit von Sprachkompetenz, Lesbarkeit sowie deren Interaktion durchgeführt (Zahlen sind nicht abgebildet). Gemäß den gewonnenen Ergebnissen ist zwar die Sprachfähigkeit signifikant ($F = 12,71$; $p = 0,000$; zweiseitig), durchschnittlich gaben jedoch die Probanden mit dem geringen Sprachkompetenzniveau mit 81,7 % einen ausreichend

⁴⁵¹ Ähnliche Ergebnisse in Bezug auf die Sprachkompetenz werden generiert, wenn statistische Modelle mit der Rohstichprobe und nur die Teilnehmer mit einem selbstberichteten Verständnis der Fallstudie von 60 % bzw. 70 % verwendet werden (Zahlen sind nicht abgebildet).

⁴⁵² Siehe Kapitel 3.4.3.1.

⁴⁵³ Die durchschnittlichen Werte für das Lesbarkeitsmaß sind zwischen den Sprachgruppen wie folgt verteilt: 2,63 (Fremdsprache) und 2,21 (Muttersprache). Für die Lesbarkeit sieht die Verteilung folgendermaßen aus: 2,12 (hohe Lesbarkeit) und 2,72 (niedrige Lesbarkeit).

hohen Wert für das Verständnis der Fallstudie an.⁴⁵⁴ Des Weiteren machten die Testpersonen mit einem selbstberichteten Verständnis der Fallstudie von weniger als 50 % nur 3,3 % aller Teilnehmer im Experiment aus. Insgesamt kann daraus geschlussfolgert werden, dass mangelndes Verständnis nicht als Erklärung für die Ergebnisse dienen kann.

Wie in Kapitel 4.3.2 erwähnt, sind die Werte für Need for Cognition nicht gleich zwischen den Experimentgruppen verteilt. Des Weiteren übt diese Variable unter Verwendung der Rohstichprobe mit allen Kontrollvariablen einen signifikanten Einfluss auf die Präferenzstärken aus. Somit könnte das ethische Verhalten auch mit der unterschiedlichen Verteilung der Need-for-Cognition-Werte in den einzelnen Sprachkompetenzgruppen erklärt werden. Gegen diese Erklärung spricht jedoch, dass die Regressionskonstante für die Variable Need for Cognition im Modell mit Präferenzstärken als abhängige Variable positiv ist ($b = 0,23$; $p = 0,050$). Ein höherer Wert für Need for Cognition impliziert also im vorliegenden Experiment, dass sich die Teilnehmer nicht mehr, sondern weniger utilitaristisch entschieden.⁴⁵⁵ Des Weiteren büßte die Variable Need for Cognition als Kovariate in der Hauptanalyse mit der finalen Stichprobe ihr Signifikanzniveau ein.

4.4.3.2 Überprüfung der Erklärungsansätze für den FLE

Das Ziel dieses Kapitels ist es, zu untersuchen, ob der festgestellte Haupteffekt der Sprachkompetenz auf die im theoretischen Teil der Arbeit diskutierten Erklärungsansätze für den FLE zurückgeführt werden kann. Im Einzelnen werden die folgenden Erklärungsansätze betrachtet: höhere emotionale und psychologische Distanz, DFE und Abschwächung des Einflusses der moralisch-sozialen Normen.

Gemäß dem „blunted-deontology account“⁴⁵⁶ könnte eine Fremdsprache das deontologische Entscheidungsverhalten aufgrund emotionaler und psychologischer Distanz abschwächen. Demnach hilft eine Fremdsprache dabei, den Fokus der Entscheidung stärker auf die Konsequenzen zu legen, weil das affektive System 1 dann aufgrund emotionaler und psychologischer Distanz weniger stark am Entscheidungsprozess beteiligt ist. Um diesen Erklärungsansatz zu überprüfen, wurden die Probanden gebeten, ihre Gefühle im Falle der Unternehmensschließung auf einer Skala von 1 (= sehr schlechte)

⁴⁵⁴ Im Vergleich dazu berichteten die Teilnehmer mit dem muttersprachlichen Kompetenzniveau im Durchschnitt einen Wert von 91,1 % und die Teilnehmer mit dem mittelhohen Sprachkompetenzniveau 93,1 %.

⁴⁵⁵ Für die einzelnen Sprachkompetenzniveaus sind die Need-for-Cognition-Werte wie folgt verteilt: 4,76 (muttersprachlich); 5,28 (mittelhoch) und 4,58 (gering). Trotz des gegenläufigen Effekts der Variable Need for Cognition waren die Entscheidungen der Teilnehmer mit dem mittelhohen Sprachkompetenzniveau bekanntlich am stärksten utilitaristisch ausgeprägt.

⁴⁵⁶ Vgl. Hayakawa et al. 2017b, S. 1388.

bis 7 (= sehr gute) anzugeben. Ein ANOVA-Modell (Zahlen sind nicht abgebildet) mit Gefühlsangaben als abhängige Variable und Sprachkompetenz, Lesbarkeit und deren Interaktion als unabhängige Variablen ergab keine statistisch signifikanten Ergebnisse (alle p -Werte $> 0,291$). Somit spielte der auf emotionaler und psychologischer Distanz basierende Erklärungsansatz keine Rolle im vorliegenden Experiment.

Das deontologische Verhalten kann infolge des Einsatzes einer fremden Sprache auch dadurch abgeschwächt werden, dass die moralisch-sozialen Normen für den Entscheider weniger stark präsent sind. Folglich sind die Entscheidungsträger in einer Fremdsprache eher bereit, gegen diese Normen zu verstoßen.

Um diesen Erklärungsansatz (indirekt) zu testen, wurden die Dimensionen Deontologie und Utilitarismus gemäß der Multidimensional Ethics Scale⁴⁵⁷ gemessen. Diese Variablen wurden für jede Sprachkompetenzgruppe im Rahmen eines Regressionsmodells mit den Präferenzstärken als abhängige Variable in eine Beziehung gesetzt, um die ethische Orientierung der Experimentteilnehmer in Bezug auf die Unternehmensschließung zu bestimmen.⁴⁵⁸ Die statistischen Ergebnisse sind in Tabelle 14 abgebildet. Es fällt auf, dass die deontologische Orientierung nur in der Gruppe mit dem mittelhohen Sprachkompetenzniveau keine Rolle spielte.⁴⁵⁹ Zudem übte die utilitaristische Orientierung lediglich bei den Teilnehmern mit dem geringen Sprachkompetenzniveau keinen Einfluss auf die Präferenzstärken aus.⁴⁶⁰ Diese Ergebnisse erlauben zwar keine Rückschlüsse auf eine Ursache-Wirkungs-Beziehung, daraus können aber Hinweise zugunsten desjenigen Erklärungsansatzes für den FLE, der auf der Abschwächung des Einflusses der moralisch-sozialen Normen beruht, abgeleitet werden. Wahrscheinlich waren die Probanden mit dem mittelhohen Sprachkompetenzniveau weniger stark von den moralisch-sozialen Normen geleitet und hatten gleichzeitig ausreichend kognitive Kapazitäten, um im utilitaristischen System-2-Denkmodus zu entscheiden. Personen mit dem geringen Sprachkompetenzniveau hatten dagegen aufgrund der Informationsverarbeitungsschwierigkeiten in einer Fremdsprache nicht genügend freie kognitive Ressourcen und fällten ihre Entscheidung im deontologischen System-1-Denkmodus.

⁴⁵⁷ Siehe Cohen et al. 1996; Cohen et al. 2001.

⁴⁵⁸ Dieses Verfahren wurde von Cohen et al. 2001 übernommen.

⁴⁵⁹ In der Gruppe mit dem geringen Sprachkompetenzniveau weist die entsprechende Regressionskonstante ein negatives Vorzeichen auf. Inhaltlich bedeutet dies, dass höhere deontologische Orientierung der Testpersonen aus dieser Sprachkategorie mit einem höheren Utilitarismusgrad einherging. Ein möglicher Grund für dieses inkonsistente Verhalten könnte sein, dass die Probanden mit dem geringen Sprachkompetenzniveau das Experiment weniger ernstnahmen.

⁴⁶⁰ Ähnliche Ergebnisse werden generiert, wenn nur die Teilnehmer mit einem selbstberichteten Verständnis der Fallstudie von 60 % bzw. 70 % verwendet werden. Wenn die Rohstichprobe betrachtet wird, verliert der Regressionskoeffizient für die deontologische Orientierung in der Gruppe mit dem geringen Sprachkompetenzniveau die Signifikanz (Zahlen sind nicht abgebildet).

Tabelle 14: Regressionskoeffizienten zur Messung ethischer Orientierung

| Prädiktorvariable (Präferenzstärke ^a) | Sprachkompetenz ^b | | |
|---|------------------------------|------------|---------|
| | Mutterspr. | Mittelhoch | Gering |
| Deontologie ^c | 0,25*** | -0,05 | -0,46** |
| Utilitarismus ^c | -0,38*** | -0,41** | -0,11 |

Alle *p*-Werte sind zweiseitig. *, **, *** repräsentieren eine Signifikanz auf den Niveaus 10 %, 5 % und 1 %.

^a Die Präferenzstärke zugunsten der utilitaristischen Alternative repräsentiert die Bereitschaft der Testperson, die Unternehmensschließung durchzuführen. Je niedriger der Wert, desto größer ist die Präferenzstärke zugunsten der utilitaristischen Alternative.

^b Die Sprachkompetenz wurde anhand eines C-Tests (siehe Kapitel 3.3.1.2 und Anhang I) gemessen. Die Teilnehmer wurden in die drei Sprachkompetenzgruppen eingeteilt: muttersprachlich; mittelhoch; gering.

^c Deontologie und Utilitarismus sind zwei Dimensionen der ethischen Urteilsfindung und wurden aus der Multidimensional Ethics Scale entnommen, siehe Cohen et al. 1996; Cohen et al. 2001.

Quelle: Eigene Darstellung.

Gemäß dem „heightened-utilitarianism account“⁴⁶¹ fördert eine Fremdsprache das utilitaristische Verhalten mittels Aktivierung von System 2. Dieser Erklärungsansatz basiert auf dem DFE. Demnach hemmt eine Fremdsprache den mentalen Fluss und sollte im Sinne einer metakognitiven Verarbeitungsschwierigkeit System 2 aktivieren. Dieser Erklärungsansatz wurde mithilfe der gleichen Fragen, die zur Überprüfung der Lesbarkeitsmanipulation dienten, untersucht. Wie bereits im vorherigen Abschnitt beschrieben, ist für den zusammengesetzten Wert dieser Fragen (im Folgenden Fluency) die Variable Sprache signifikant. Das Signifikanzniveau erhöht sich weiter, wenn Sprache durch Sprachkompetenz ersetzt wird ($F = 10,61$; $p = 0,000$; zweiseitig). Gegen diesen Erklärungsansatz spricht jedoch, dass der festgestellte Einfluss der Sprachkompetenz auf das utilitaristische Verhalten nicht durch Fluency mediiert wird.⁴⁶² Aus diesem Grund wird ausgehend von den vorliegenden Ergebnissen der DFE als Erklärung für den FLE nicht unterstützt.

⁴⁶¹ Vgl. Hayakawa et al. 2017b, S. 1388.

⁴⁶² Im Regressionsmodell mit Präferenzstärken als abhängige Variable und Sprachkompetenz sowie Fluency als unabhängigen Variablen weist Fluency als Mediator keinen signifikanten Effekt auf ($t = -0,15$; $p = 0,884$; zweiseitig).

4.5 Diskussion der Ergebnisse

Die vorliegende Studie wurde konzipiert, um zu untersuchen, ob und wie bestimmte linguistische Merkmale von Controlling-Berichten das ethische Verhalten beeinflussen. Im Speziellen wurden die Effekte der Lesbarkeit und des Einsatzes einer Fremdsprache betrachtet. Außerdem wurde der FLE tiefergehend untersucht, indem die Rolle der Sprachkompetenz berücksichtigt und die potenziellen Erklärungsansätze für den FLE getestet wurden.

Zur Überprüfung der Hypothesen wurde ein Between-Subject-Experiment durchgeführt. Teilnehmer, welche die Rolle eines Managers einnahmen, wurden mit einem fiktiven Controlling-Bericht ausgestattet. Auf Grundlage der Informationen im Controlling-Bericht mussten sich die Probanden für (utilitaristische Alternative) oder gegen (deontologische Alternative) eine Unternehmensschließung entscheiden. Anschließend gaben die Teilnehmer ihre Präferenzstärken zugunsten der Unternehmensschließung an, welche die wichtigste abhängige Variable im Experiment darstellten. Die unabhängigen Variablen Sprache, Sprachkompetenz und Lesbarkeit wurden identisch wie in Experiment 1 konstruiert.

Alle Hypothesen mussten ausgehend von den Ergebnissen der statistischen Analyse im vorliegenden Experiment abgelehnt werden. Es konnte lediglich ein signifikanter Haupteffekt der Sprachkompetenz festgestellt werden. Unter der Annahme, dass die Lesbarkeitsmanipulationen keine ausreichende Wirkung gemäß dem CLE entfalteten, ist dieser Befund im Sinne des FLE interpretierbar. Demnach gaben Teilnehmer mit dem mittelhohen Sprachkompetenzniveau durchschnittlich die am stärksten utilitaristisch ausgeprägten Antworten an. Ergänzende Analysen ergaben außerdem, dass Personen mit der mittelhohen Sprachkompetenzstufe am wenigsten deontologisch orientiert waren, was mit demjenigen Erklärungsansatz für den FLE konform ist, der auf der Abschwächung des Einflusses der moralisch-sozialen Normen beruht. Die anderen potenziellen Erklärungsansätze für den FLE konnten in der vorliegenden Untersuchung nicht unterstützt werden. Im Folgenden werden mögliche Gründe, welche die Ablehnung der Hypothesen bedingen könnten, diskutiert.

Obwohl ein signifikanter Unterschied im Lesbarkeitsmaß zwischen den hohen und niedrigen Lesbarkeitsgruppen bestand, gibt es gute Gründe anzunehmen, dass die Lesbarkeitsmanipulationen nicht ausreichend waren, um den CLE bei den Testpersonen auszulösen. In der Untersuchung von Trémolière et al. 2012 (Experiment 2) konnte der Einfluss einer ablenkenden Aufgabe im Zusammenhang mit ethischen Dilemmata nur bei einer sehr hoher kognitiver Belastung gemessen werden. Conway und Gawronski 2013 verwendeten zur Erzeugung der kognitiven Belastung im Rahmen ihrer Studie zum ethischen Verhalten siebenstellige passwortähnliche Kombinationen aus Zahlen und Zeichen. Vor dem Hintergrund, dass im menschlichen Arbeitsgedächtnis in etwa

maximal sieben Informationseinheiten gleichzeitig gespeichert werden können,⁴⁶³ fällt eine solche Erinnerungsaufgabe ebenfalls in die Kategorie einer sehr hohen kognitiven Belastung. Alles in allem scheint im Kontext der ethischen Dilemmata die kognitive Belastung aufgrund von Lesbarkeitsmanipulationen nicht ausreichend zu sein, um den CLE und damit ein reduziertes utilitaristisches Verhalten zu induzieren.⁴⁶⁴

Ein möglicher Grund für den schwachen FLE könnte sein, dass das ethische Dilemma im vorliegenden Experiment die deontologische relativ zur utilitaristischen Perspektive weniger stark förderte. In den Studien zum ethischen Entscheidungsverhalten fallen die Probanden i. d. R. eine (fiktive) Entscheidung über Leben und Tod von anderen Menschen. Wenn auch der ökonomische Entscheidungsrahmen es grundsätzlich zulässt, Konfliktsituation zwischen der deontologischen und utilitaristischen Maxime zu modellieren,⁴⁶⁵ scheint in der vorliegenden Untersuchung der utilitaristische Aspekt einen viel stärkeren Einfluss auf die Entscheidungen der Testpersonen ausgeübt zu haben. Dafür spricht die Tatsache, dass die utilitaristische Orientierung der Teilnehmer insgesamt signifikant höher ausgeprägt war als die deontologische Orientierung (4,66 vs. 4,13; $t = -3,22$; $p = 0,001$; zweiseitig). Alternativ ist es auch vorstellbar, dass die Probanden aus dem betriebswirtschaftlichen (Teil-)Studiengang generell utilitaristischer eingestellt sind als die übrige Population.⁴⁶⁶

⁴⁶³ Siehe Cowan 2000.

⁴⁶⁴ Im Experiment wurden die Teilnehmer auch nach ihrer geistigen Anstrengung in Bezug auf die Fallstudie gefragt. Auf einer Skala von 1 (= sehr wenig) bis 7 (= sehr viel) lag der Mittelwert für alle Testpersonen bei nur 3,22.

⁴⁶⁵ So konnten Gold et al. 2013 charakteristische Entscheidungsmuster sowohl in „klassischen“ ethischen Dilemmata als auch im ökonomischen ethischen Dilemma feststellen. Der wesentliche Unterschied zwischen dem ökonomischen Dilemma in der Untersuchung von Gold et al. 2013 und demjenigen in dieser Studie besteht darin, dass im vorliegenden Experiment eine Situation modelliert wurde, in der die Probanden keine physische Einwirkung auf andere Personen ausüben müssten. Physische Einflussnahme (d. h., muskuläre Einwirkung auf andere Personen) ist ein wichtiger Faktor bei ethischen Dilemmata, der dazu führt, dass sich die Testpersonen weniger utilitaristisch und stärker deontologisch entscheiden, siehe Greene et al. 2009. Die Berücksichtigung physischer Einflussnahme in einem auf den Controlling-Kontext ausgerichteten Experiment erscheint jedoch nicht sinnvoll zu sein.

⁴⁶⁶ Die Studie von Groves et al. 2008 ergab bspw., dass Manager – Studierende aus dem Fach Betriebswirtschaftslehre und Manager weisen im Allgemeinen vergleichbares Verhalten in experimentellen Untersuchungen auf, siehe Kapitel 2.5.3 – überwiegend utilitaristische Perspektive bei ethischen Entscheidungen einnehmen.

5. Zusammenfassung und Ausblick

Ausgangspunkt der vorliegenden Arbeit war die Frage, welche Verhaltenswirkungen die Lesbarkeit von und der Einsatz einer fremden Sprache in Controlling-Berichten auslösen können. Dabei standen der Einfluss dieser linguistischen Merkmale auf das Risikoverhalten und das ethische Verhalten eines Entscheidungsträgers im Zentrum der Untersuchung.

Um die Forschungsfrage beantworten zu können, wurden zunächst relevante theoretische Grundlagen aufgezeigt. Im ersten Schritt wurde der Unterschied zwischen dem auf bestimmten Rationalitätsaxiomen basierenden normativen und dem deskriptiven Ansatz im Rahmen der betriebswirtschaftlichen Entscheidungsforschung erläutert und die vorliegende Arbeit dem letzteren Ansatz zugeordnet. Im weiteren Schritt wurde das Konzept des Denkens in zwei sog. Systemen beleuchtet, das zentral für das Verständnis von Heuristiken und die Ableitung von Hypothesen im empirischen Teil dieser Arbeit ist. Im weiteren Verlauf wurden ausgewählte kognitive Effekte der Lesbarkeit und Fremdsprache behandelt, bevor auf die Gestaltungsdimensionen und Verhaltenswirkungen des Management Reporting eingegangen wurde. Die wesentlichen Merkmale der experimentellen Forschungsmethode rundeten den theoretischen Teil dieser Arbeit ab.

Den empirischen Teil der vorliegenden Arbeit bilden zwei experimentelle Untersuchungen. In ihnen wurden im ersten Experiment die Effekte der Lesbarkeit und einer Fremdsprache auf die Bereitschaft, „smarte“ Risiken einzugehen, untersucht. Im zweiten Experiment wurden die gleichen linguistischen Merkmale in Bezug auf das ethische Verhalten anhand eines Dilemmas, in dem die utilitaristische und die deontologische Maxime miteinander in Konflikt standen, untersucht.

Das Design und die Ergebnisse von Experiment 1 können thesenförmig wie folgt zusammengefasst werden:

- Bei dem Design handelt es sich um ein 2×2-Between-Subject-Experiment mit den wichtigsten unabhängigen Variablen: Sprache (Fremd- vs. Mutter-) und Lesbarkeit (hoch vs. niedrig). Die Lesbarkeit wurde mithilfe ausgewählter linguistischer und formaler Kriterien gemäß dem „Plain English Handbook“ von SEC 1998 manipuliert. Die Sprachkompetenz der Teilnehmer wurde anhand eines C-Tests gemessen.
- Die Experimentteilnehmer nahmen – im Rahmen der Prinzipal-Agent-Beziehung mit hypothetischen finanziellen Konsequenzen⁴⁶⁷ – die Rolle eines Managers ein und erhielten einen fiktiven Managementbericht. Der Bericht umfasste zwei alternative Kaufangebote für die Produktion des Unternehmens. Ein Angebot war eine sichere Option und das andere Angebot stellte eine risikobehaftete, aber hinsichtlich des Erwartungswerts der Erlöse vorteilhaftere Option dar. Die Probanden wurden

⁴⁶⁷ Die Entscheidung, die Teilnehmer nicht zu vergüten, wird in Kapitel 3.3.1.1 begründet.

gebeten, sich zunächst zwischen den beiden Kaufangeboten zu entscheiden und danach ihre individuellen Sicherheitsäquivalente für die risikobehaftete Alternative anzugeben. Zur statistischen Analyse wurden die Sicherheitsäquivalente aufgrund des metrischen Skalenniveaus als abhängige Variable herangezogen.

- In Übereinstimmung mit H1 waren die Testpersonen gemäß dem CLE signifikant weniger bereit, „smarte“ Risiken einzugehen, wenn sie einen Controlling-Bericht mit einem niedrigen im Vergleich zu einem hohen Lesbarkeitsniveau verarbeiten mussten. Dieses Ergebnis war sowohl auf deskriptiver als auch statistischer Ebene deutlich festzustellen.
- In H2 wird gemäß dem FLE postuliert, dass die Entscheidungsträger nur bei einem hohen Lesbarkeitsniveau eher bereit sein werden, „smarte“ Risiken zu akzeptieren, wenn sie einen Controlling-Bericht in einer Fremdsprache im Vergleich zu ihrer Muttersprache verarbeiten müssen. Auch diese Hypothese konnte deskriptiv und statistisch bestätigt werden. Allerdings war der FLE nur schwach signifikant und nicht besonders robust.
- Schließlich besagt H3, dass der FLE – wie unter H2 formuliert – nur für mittelhohe Sprachkompetenzstufen gilt. Diese Hypothese konnte ebenfalls verifiziert werden.
- Ergänzende Analysen in Bezug auf die potenziellen Erklärungsansätze für den FLE zeigten, dass Probanden mit dem mittelhohen Sprachkompetenzniveau und unter der Bedingung hoher Lesbarkeit signifikant weniger von positiven Gefühlen gegenüber dem sicheren Erlös geleitet waren. Dieses Ergebnismuster spricht zugunsten reduzierter emotionaler Resonanz als Erklärung für den FLE. Die anderen Erklärungsansätze für den FLE (psychologische Distanz und DFE) konnten nicht unterstützt werden.
- Ein möglicher Grund für die wenig signifikanten Ergebnisse in Bezug auf den FLE könnte im Experimentdesign liegen, bei dem die Teilnehmer eine Risikoentscheidung im Namen anderer treffen sollten. Die Empathy-Gap-Hypothese sagt voraus, dass eine solche Konstellation zu höherer Risikoakzeptanz seitens der Entscheider führen wird, was das Potenzial für den FLE reduzieren sollte. Außerdem könnte die Verwendung von realen Auszahlungen den FLE verstärken, weil so stärkere Emotionen in den Entscheidungsprozess involviert werden.

In Experiment 2 wurde das gleiche Design wie in Experiment 1 verwendet, jedoch auf der Grundlage einer anderen Fallstudie eine andere Fragestellung untersucht. In der Fallstudie von Experiment 2 ging es um eine Entscheidung für (utilitaristisch Alternative) oder gegen (deontologisch Alternative) die Schließung eines Tochterunternehmens. Ferner wurden die Teilnehmer gebeten, ihre Präferenzstärke zugunsten der Unternehmensschließung anzugeben. Zur statistischen Analyse wurden die Präferenzstärken aufgrund des metrischen Skalenniveaus als abhängige Variable herangezogen. Die

Hypothesen in der zweiten Untersuchung wurden identisch wie in Experiment 1 formuliert, mit dem einzigen Unterschied, dass sie sich, statt auf die Akzeptanz von „smarten“ Risiken, auf das utilitaristische Verhalten bezogen. Im Laufe der Untersuchung wurden im Einzelnen folgende Ergebnisse erzielt:

- Weder Sprache noch die Interaktion zwischen Sprache und Lesbarkeit übten einen signifikanten Einfluss auf das ethische Verhalten der Teilnehmer aus; somit mussten alle Hypothesen abgelehnt werden. Es konnte lediglich ein signifikanter Haupteffekt der Sprachkompetenz festgestellt werden. Unter der Annahme, dass die Lesbarkeitsmanipulationen keine ausreichende Wirkung gemäß dem CLE entfalteten, ist dieser Befund im Sinne des FLE interpretierbar. Demnach gaben Teilnehmer mit dem mittelhohen Sprachkompetenzniveau durchschnittlich die am stärksten utilitaristisch ausgeprägten Antworten an.
- Ergänzende Analysen ergaben außerdem, dass Personen mit der mittelhohen Sprachkompetenzstufe am wenigsten deontologisch orientiert waren, was für denjenigen Erklärungsansatz für den FLE spricht, der auf der Abschwächung des Einflusses der moralisch-sozialen Normen beruht. Die anderen potenziellen Erklärungsansätze für den FLE (emotionale und psychologische Distanz sowie DFE) konnten nicht untermauert werden.
- Vor dem Hintergrund der Generierung kognitiver Belastung in anderen Studien zum ethischen Verhalten kann vermutet werden, dass die vorgenommenen Lesbarkeitsmanipulationen nicht ausreichend sind, um den CLE und damit reduziertes utilitaristisches Verhalten zu induzieren.
- Ein möglicher Grund für den schwachen FLE könnte darin liegen, dass das modellierte ethische Dilemma die deontologische relativ zur utilitaristischen Perspektive weniger stark förderte. Des Weiteren ist es auch vorstellbar, dass die Probanden aus dem betriebswirtschaftlichen (Teil-)Studiengang generell utilitaristischer eingestellt sind als die übrige Population.

Zusammenfassend beleuchtet die vorliegende Arbeit verhaltensbezogene Einflüsse der Lesbarkeit und des Einsatzes einer Fremdsprache einerseits auf die Akzeptanz von „smarten“ Risiken und andererseits auf das ethische Verhalten im Controlling-Kontext. Controller, die die Informationen vorbereiten und diese den Entscheidungsträgern zur Verfügung stellen, sollten wissen, dass manche linguistische Charakteristika den Entscheidungsfindungsprozess beeinflussen können, selbst dann, wenn die Informationen von den Adressaten richtig verstanden wurden. Im Rahmen des ersten Experiments konnte gezeigt werden, dass ein weniger gut lesbarer Controlling-Bericht der Bereitschaft, „smarte“ Risiken einzugehen, tendenziell abträglich ist. Darüber hinaus wurde die erste Evidenz dafür geliefert, dass sich der Einsatz einer Fremdsprache im Management Reporting positiv auf die Akzeptanz von „smarten“ Risiken auswirken kann. Notwendige Voraussetzungen für diesen Effekt sind jedoch: ein mittelhohes Sprachkompetenzniveau des Entscheidungsträgers und ein gut lesbarer Controlling-Bericht. Im Unterschied zum ersten Experiment konnten im zweiten Experiment kein signifikanter

Effekt der Lesbarkeit und nur ein schwacher FLE festgestellt werden. Demnach erhöht sich die Neigung zum utilitaristischen Verhalten, wenn ein Controlling-Bericht in einer Fremdsprache verfasst ist und der Entscheidungsträger über ein mittelhohes Sprachkompetenzniveau verfügt. Diese Erkenntnisse verdeutlichen, wie unterschiedlich sich die Wirkungen der untersuchten linguistischen Charakteristika in Abhängigkeit von der Entscheidungsart entfalten können, und dienen gleichzeitig als Ausgangspunkt für weitere Forschungsbemühungen.

Es gibt noch vielversprechende Richtungen für zukünftige Forschungsbemühungen. Zunächst wurde die Entscheidungsregel in Experiment 1 auf eine einfache Kalkulation des Erwartungswerts reduziert. Im Gegensatz dazu sind reale betriebswirtschaftliche Entscheidungssituationen i. d. R. von komplexen Optimierungsstrategien gekennzeichnet. Schedlinsky et al. 2016 untersuchten z. B. das Risikoverhalten im Kontext wettbewerbsbasierter Vergütungssysteme, bei denen das Verhalten anderer Teilnehmer gemäß der Spieltheorie zu berücksichtigen war. Die Autoren fanden heraus, dass die Individuen in solchen komplexen Entscheidungssituationen aufgrund vereinfachter Entscheidungsregeln mit hoher Wahrscheinlichkeit suboptimale Entscheidungen treffen – d. h., übermäßig viele Risiken eingehen. Es stellt sich die Frage, ob der Einsatz einer Fremdsprache in Verbindung mit komplexen Entscheidungsstrategien und unerwünschten Risiken ebenfalls zur erhöhten Risikobereitschaft führen wird. Außerdem sollten die gewonnenen Ergebnisse mithilfe von Experimenten mit realer Vergütung validiert werden. In Bezug auf Experiment 2 könnte versucht werden, ethische Dilemmata zu modellieren, in denen die deontologische und die utilitaristische Maxime stärker miteinander in Konflikt stehen. Aufschlussreich wäre auch die Frage, ob Studierende aus dem Fach Betriebswirtschaftslehre bei Dilemmata im Controlling-Kontext tatsächlich utilitaristischer eingestellt sind als andere Zielgruppen. Schließlich könnten weitere Anwendungsfelder des FLE – wie z. B. betrügerische Handlungen, illusorische Korrelationen oder FF-Heuristiken – auf für das Controlling relevante Entscheidungssituationen übertragen und untersucht werden. Wie die behavioristische Perspektive und Erkenntnisse aus unterschiedlichen Fachgebieten, wie der Psychologie, der Soziologie oder der Lernforschung, weiterhin Eingang in die betriebswirtschaftliche Forschungswelt finden, bleibt mit großem Interesse abzuwarten.

Anhang I

C-Test

Ein C-Test ist ein zuverlässiges und schnelles Verfahren, mit dessen Hilfe die allgemeine Sprachkompetenz eruiert werden kann.⁴⁶⁸ Der Test besteht aus mehreren kurzen Texten und basiert auf dem Konzept der reduzierten Redundanz. Bei den Texten verbleiben der erste und letzte Satz unverändert. Auf die übrigen Sätze wird die „Regel von zwei“ angewendet: Beginnend mit dem zweiten Satz wird die zweite Hälfte jedes zweiten Wortes getilgt.⁴⁶⁹ Zahlen und Eigennamen werden nicht manipuliert. Dieser Vorgang wird fortgesetzt, bis circa 25 Lücken erreicht sind.

Die Entwicklung und Evaluierung des im Rahmen dieser Arbeit verwendeten C-Tests basiert auf den Instruktionen von Grotjahn 2002. Um eine heterogene Stichprobe von Texten zu erhalten, stammten die drei verwendeten Texte aus englischsprachigen Lehrbüchern, die für die folgenden Sprachkompetenzstufen konzipiert wurden: B1, B2 und C1. Für den kompletten Test galt die Zeitbeschränkung von 15 Minuten. Die folgende Textpassage ist ein Auszug aus dem ersten Text und soll einen Eindruck davon vermitteln, wie der im Rahmen dieser Arbeit eingesetzte C-Test aussah:

Many people think that we laugh because we see or hear something funny, but most of the time this isn't true.

In a st_____, a prof_____ of psych_____ and h_____ students list_____ to a_____ made no_____ on hund_____ of conver_____ in pub_____ places. [...]

Um die Modellgüte des verwendeten C-Tests zu evaluieren, wurde ein Pretest mit 107 Studierenden einige Monate vor den eigentlichen Experimenten durchgeführt. Die Itemanalyse zeigte annähernd identische Ergebnisse wie die Itemanalysen aus Experimenten 1 und 2.⁴⁷⁰

⁴⁶⁸ Vgl. Grotjahn 2002, S. 211.

⁴⁶⁹ Vgl. Raatz und Klein-Braley 2002, S. 75.

⁴⁷⁰ Siehe Kapitel 3.3.1.2.

Anhang II

Tabelle 15: Beispiele für Lesbarkeitsmanipulationen (Experiment 1)

| Linguistisches Mittel ^a | Niedrige Lesbarkeit | Hohe Lesbarkeit |
|---|---|---|
| Kurze Sätze | Da auch die Nachfrageprognosen keine schwankenden Steigerungen zeigten, hat sich das Unternehmen letztes Jahr entschieden, mit umfassenden Restrukturierungsmaßnahmen nicht abzuwarten. | Die damaligen Nachfrageprognosen zeigten stabile Steigerungen. Deshalb haben wir entschieden, umfassende Umbaumaßnahmen durchzuführen. |
| Aktivformen | Es wird uns ein Grundpreis von 100.000 \$ angeboten. | Sie bieten uns einen Grundpreis von 100.000 \$ an. |
| Sätze mit Verben | Die Überholung des stärksten Konkurrenten der MICRO-PC AG wurde mithilfe eines Jahresüberschusses in Höhe von 10,25 Mio. \$ im letzten Geschäftsturnus ermöglicht. | Mit einem Jahresüberschuss von 10,25 Mio. \$ im Vorjahr konnten wir unseren stärksten Wettbewerber überholen. |
| Keine überflüssigen Wörter | Nun freue ich mich Ihnen mitteilen zu können, dass die Restrukturierungsmaßnahmen erfolgreich zu Ende gebracht wurden [...] | Die Umbaumaßnahmen sind jetzt abgeschlossen [...] |
| Affirmative Sprache | Diese Führungsposition gilt es nun nicht zu verlieren. | Diese Führungsposition im Markt gilt es nun zu halten. |
| Personalpronomen | Beide Unternehmen sind langjährige Kunden der MICRO-PC AG [...] | Beide Unternehmen sind unsere langjährige Kunden [...] |
| Sätze bei denen das Subjekt, Prädikat und Objekt nah beieinander stehen | Die Kostenübersicht der zusätzlichen Produktion von 1.000 Computern, welche von der Controlling-Abteilung der MICRO-PC AG zusammengestellt wurde, finden Sie auf der nächsten Seite. | Unsere Controlling-Abteilung hat die Kostenübersicht der zusätzlichen Produktion von 1.000 Computern auf der nächsten Seite zusammengestellt. |

^a Linguistische Mittel wurden selektiv aus dem „Plain English Handbook“ von SEC 1998 entnommen. Quelle: Eigene Darstellung.

Tabelle 16: Ausschlusskriterien (Experiment 1)

| | Lesbarkeit | | | | Gesamt |
|----------------------------------|-------------|-------------|-------------|-------------|--------------|
| | Hoch | | Niedrig | | |
| | Sprache | | Sprache | | |
| | Fremd- | Mutter- | Fremd- | Mutter- | |
| Stichprobe (unbereinigt) | 69 (100 %) | 49 (100 %) | 52 (100 %) | 52 (100 %) | 222 (100 %) |
| Manipulation-Checks ^a | 36 (52,2 %) | 15 (30,6 %) | 18 (34,6 %) | 16 (30,8 %) | 85 (38,3 %) |
| Verständnis ^b | 1 (1,5 %) | 2 (4,1 %) | 1 (1,9 %) | 0 (0 %) | 4 (1,8 %) |
| Andere ^c | 4 (5,8 %) | 3 (6,1 %) | 4 (7,7 %) | 1 (1,9 %) | 12 (5,4 %) |
| Stichprobe (final) | 28 (40,6 %) | 29 (59,2 %) | 29 (55,8 %) | 35 (67,3 %) | 121 (54,5 %) |

^aDie Manipulation-Checks bestanden aus zwei Multiple-Choice-Fragen. Die erste Frage bezog sich auf den Erwartungswert der Erlöse der risikobehafteten Option. Die zweite Frage sollte ausschließen, dass die Teilnehmer von einer möglichen Verlustsituation ausgingen. Wenn eine der beiden Fragen falsch beantwortet wurde, wurden die entsprechenden Probanden ausgeschlossen.

^bDas Verständnis stellte nach den Manipulation-Checks das zweite Ausschlusskriterium dar. Die verbleibenden Teilnehmer nach den Multiple-Choice-Fragen, die ein selbstberichtetes Verständnis der Fallstudie von weniger als 50 % angaben, wurden ausgeschlossen. Folgende Teilnehmer beantworteten mindestens eine Multiple-Choice-Frage falsch und berichteten gleichzeitig ein Verständnis der Fallstudie von weniger als 50 %: 8 in Gruppe 1; 1 in Gruppe 2; 3 in Gruppe 3 und 3 in Gruppe 4.

^cAndere Ausschlusskriterien umfassten fehlende oder inkorrekt angegebene individuelle Sicherheitsäquivalente (d. h., die Wahl der risikobehafteten Option in Kombination mit der Angabe eines Sicherheitsäquivalents, das unter dem vorgegebenen Erlös der sicheren Option i. H. v. 138.000 \$ liegt, et vice versa) und sprachbezogene Angelegenheiten (z. B. wurden Testpersonen aussortiert, wenn sie angaben, kein deutschsprachiger Muttersprachler – oder vergleichbarer Sprecher – in der muttersprachlichen Gruppe zu sein). Diese Ausschlusskriterien wurden auf die nach den ersten beiden Screening-Stufen verbleibenden Probanden angewandt. Folgende Teilnehmer sind bei allen drei Ausschlusskategorien gleichzeitig durchgefallen: 2 in Gruppe 1; 0 in Gruppe 2; 1 in Gruppe 3 und 1 in Gruppe 4.

Quelle: Eigene Darstellung.

Tabelle 17: ANCOVA – Volles Modell mit Rohstichprobe (Experiment 1)

| Prädiktorvariable (Sicherheitsäquivalent ^a) | Fg | Quadrat-summe | F-Wert | p-Wert |
|---|----|----------------------|--------|---------|
| Sprache ^b | 1 | 3,01×10 ⁷ | 0,05 | 0,820 |
| Lesbarkeit ^c | 1 | 3,64×10 ⁹ | 6,28 | 0,013** |
| Geschlecht | 1 | 2,35×10 ⁵ | 0,00 | 0,984 |
| Alter | 1 | 6,23×10 ⁷ | 0,11 | 0,744 |
| Risikoeinstellung ^d | 1 | 2,24×10 ⁹ | 3,85 | 0,051* |
| Faith in Intuition ^e | 1 | 5,26×10 ⁷ | 0,09 | 0,764 |
| Need for Cognition ^e | 1 | 6,35×10 ⁸ | 1,09 | 0,297 |
| Ausschlusskriterien ^f | 1 | 1,50×10 ⁸ | 0,26 | 0,612 |
| Sprache × Lesbarkeit | 1 | 1,07×10 ⁸ | 0,18 | 0,668 |

Alle *p*-Werte sind zweiseitig. *, **, *** repräsentieren eine Signifikanz auf den Niveaus 10 %, 5 % und 1 %.

^aDas Sicherheitsäquivalent repräsentiert den individuell geforderten Erlös der sicheren Option, mit dem der Teilnehmer indifferent zwischen der sicheren und der risikobehafteten Option wird.

^bDie Sprache ist eine Dummy-Variable mit den beiden Ausprägungen: Muttersprache (Deutsch) und Fremdsprache (Englisch).

^cDie Lesbarkeit ist eine Dummy-Variable mit den beiden Ausprägungen hohe und niedrige Lesbarkeit. Die Lesbarkeit wurde mithilfe des „Plain English Handbook“ von SEC 1998 mittels ausgewählter linguistischer und formaler Gestaltung manipuliert.

^dDie Risikoeinstellung steht für die generelle Einstellung der Teilnehmer gegenüber Risiken im Bereich „Gambling and Investing“ gemäß Blais und Weber 2006.

^eFaith in Intuition und Need for Cognition sind Konstrukte zur Messung der Tendenz der Teilnehmer, sich mehr auf Intuitionen (System 1) bzw. das logische Denken (System 2) zu verlassen. Die (verkürzte) Version der Skala Faith in Intuition wurde der Studie von Epstein et al. 1996 entnommen und Need for Cognition wurde mithilfe der Skala von Beißert et al. 2015 gemessen.

^fDie Ausschlusskriterien werden durch eine Dummy-Variable mit den folgenden beiden Ausprägungen repräsentiert: Der Teilnehmer wurde aus der statistischen Hauptanalyse ausgeschlossen oder er wurde nicht ausgeschlossen. Detaillierte Informationen zu den Ausschlusskriterien können aus Tabelle 16 entnommen werden.

Quelle: Eigene Darstellung.

Anhang III

Tabelle 18: Beispiele für Lesbarkeitsmanipulationen (Experiment 2)

| Linguistisches Mittel ^a | Niedrige Lesbarkeit | Hohe Lesbarkeit |
|---|---|--|
| Kurze Sätze | Das Geschäftsjahr 2018 konnte von der ZANGEN GMBH mit einem konsolidierten Gesamtergebnis von 5,7 Mio. € abgeschlossen werden, nachdem die Umsätze im Vergleich zum Vorjahr eine Steigerung von 1,3 % erfahren haben. | Wir konnten das Geschäftsjahr 2018 mit einem konsolidierten Gesamtergebnis von 5,7 Mio. € abschließen. Die Umsätze sind im Vergleich zum Vorjahr um 1,3 % gestiegen. |
| Aktivformen / Sätze bei denen das Subjekt, Prädikat und Objekt nah beieinander stehen | Die nachfolgende Tabelle, welche von unserer Controlling-Abteilung vorbereitet wurde [...] | Unsere Controlling-Abteilung hat die nachfolgende Tabelle vorbereitet [...] |
| Sätze mit Verben | Die Entwicklung der SUPPL LLC ist gut. | Die SUPPL LLC entwickelt sich gut. |
| Keine überflüssigen Wörter | Leider und zu unserem großen Bedauern haben wir dafür keine Entschädigung von der Versicherung bekommen. | Leider haben wir dafür keine Entschädigung von der Versicherung bekommen. |
| Affirmative Sprache / Personalpronomen | Die Geschäfte der ZANGEN GMBH in den USA haben sich nicht gleich entwickelt. | Unsere Geschäfte in den USA haben sich unterschiedlich entwickelt. |

^aLinguistische Mittel wurden selektiv aus dem „Plain English Handbook“ von SEC 1998 entnommen. Quelle: Eigene Darstellung.

Tabelle 19: Ausschlusskriterien (Experiment 2)

| | Lesbarkeit | | | | Gesamt |
|----------------------------------|-------------|-------------|-------------|-------------|--------------|
| | Hoch | | Niedrig | | |
| | Sprache | | Sprache | | |
| | Fremd- | Mutter- | Fremd- | Mutter- | |
| Stichprobe (unbereinigt) | 38 (100 %) | 39 (100 %) | 34 (100 %) | 39 (100 %) | 150 (100 %) |
| Manipulation-Checks ^a | 5 (13,2 %) | 2 (5,1 %) | 3 (8,8 %) | 4 (10,3 %) | 14 (9,3 %) |
| Verständnis ^b | 1 (2,6 %) | 0 (0 %) | 1 (2,9 %) | 0 (0 %) | 2 (1,3 %) |
| Sprache ^c | 0 (0 %) | 4 (10,3 %) | 1 (2,9 %) | 2 (5,1 %) | 7 (4,7 %) |
| Stichprobe (final) | 32 (84,2 %) | 33 (84,6 %) | 29 (85,3 %) | 33 (84,6 %) | 127 (84,7 %) |

^aDie Manipulation-Checks bestanden aus zwei Multiple-Choice-Fragen. Die erste Frage sollte sicherstellen, dass die Teilnehmer die schwierige finanzielle Situation des aufgrund eines Fabrikbrands in eine Notlage geratenen Tochterunternehmens erkannten. In der zweiten Frage wurde die Anzahl der Mitarbeiter der beiden Tochterunternehmen abgefragt.

^bDas Verständnis stellte nach den Manipulation-Checks das zweite Ausschlusskriterium dar. Die verbleibenden Teilnehmer nach den Multiple-Choice-Fragen, die ein selbstberichtetes Verständnis der Fallstudie von weniger als 50 % angaben, wurden ausgeschlossen. Folgende Teilnehmer beantworteten mindestens eine Multiple-Choice-Frage falsch und berichteten gleichzeitig ein Verständnis der Fallstudie von weniger als 50 %: 2 in Gruppe 1 und 1 in Gruppe 4.

^cDie Sprache umfasste sprachbezogene Ausschlusskriterien (z. B. wurden die Testpersonen aussortiert, wenn sie angaben, kein deutschsprachiger Muttersprachler – oder vergleichbarer Sprecher – in der muttersprachlichen Gruppe zu sein). Diese Ausschlusskriterien wurden auf die nach den ersten beiden Screening-Stufen verbleibenden Probanden angewandt. Kein Teilnehmer ist bei allen drei Ausschlusskategorien gleichzeitig durchgefallen.

Quelle: Eigene Darstellung.

Tabelle 20: ANCOVA – Volles Modell mit Rohstichprobe (Experiment 2)

| Prädiktorvariable (Präferenzstärke ^a) | Fg | Quadrat- summe | F-Wert | p-Wert |
|---|----|-------------------|--------|--------|
| Sprache ^b | 1 | 2,59 | 1,40 | 0,242 |
| Lesbarkeit ^c | 1 | 0,04 | 0,02 | 0,883 |
| Geschlecht | 1 | 0,08 | 0,04 | 0,840 |
| Alter | 1 | 0,81 | 0,44 | 0,511 |
| Faith in Intuition ^d | 1 | 0,23 | 0,12 | 0,726 |
| Need for Cognition ^d | 1 | 5,74 | 3,07 | 0,082* |
| Ausschlusskriterien ^e | 1 | 0,27 | 0,15 | 0,704 |
| Sprache × Lesbarkeit | 1 | 0,64 | 0,34 | 0,559 |

Alle *p*-Werte sind zweiseitig. *, **, *** repräsentieren eine Signifikanz auf den Niveaus 10 %, 5 % und 1 %.

^a Die Präferenzstärke zugunsten der utilitaristischen Alternative repräsentiert die Bereitschaft der Testperson, die Unternehmensschließung durchzuführen.

^b Die Sprache ist eine Dummy-Variable mit den beiden Ausprägungen: Muttersprache (Deutsch) und Fremdsprache (Englisch).

^c Die Lesbarkeit ist eine Dummy-Variable mit den beiden Ausprägungen hohe und niedrige Lesbarkeit. Die Lesbarkeit wurde mithilfe des „Plain English Handbook“ von SEC 1998 mittels ausgewählter linguistischer und formaler Gestaltung manipuliert.

^d Faith in Intuition und Need for Cognition sind Konstrukte zur Messung der Tendenz der Teilnehmer, sich mehr auf Intuitionen (System 1) bzw. das logische Denken (System 2) zu verlassen. Die (verkürzte) Version der Skala Faith in Intuition wurde der Studie von Epstein et al. 1996 entnommen und Need for Cognition wurde mithilfe der Skala von Beißert et al. 2015 gemessen.

^e Die Ausschlusskriterien werden durch eine Dummy-Variable mit den folgenden beiden Ausprägungen repräsentiert: Der Teilnehmer wurde aus der statistischen Hauptanalyse ausgeschlossen oder er wurde nicht ausgeschlossen. Detaillierte Informationen zu den Ausschlusskriterien können aus Tabelle 19 entnommen werden.

Quelle: Eigene Darstellung.

Literaturverzeichnis

- Abdellaoui, Mohammed (2000): Parameter-free elicitation of utility and probability weighting functions. In: *Management Science* 46 (11), S. 1497–1512.
- Abdellaoui, Mohammed; Bleichrodt, Han; L'Haridor, Olivier (2008): A tractable method to measure utility and loss aversion under prospect theory. In: *Journal of Risk and Uncertainty* 36 (3), S. 245–266.
- Abdellaoui, Mohammed; L'Haridon, Olivier; Paraschiv, Corina (2011): Experienced vs. described uncertainty: Do we need two prospect theory specifications? In: *Management Science* 57 (10), S. 1879–1895.
- Abdellaoui, Mohammed; Vossman, Frank; Weber, Martin (2005): Choice-based elicitation and decomposition of decision weights for gains and losses under uncertainty. In: *Management Science* 51 (9), S. 1384–1399.
- Abutalebi, Jubin (2008): Neural aspects of second language representation and language control. In: *Acta psychologica* 128 (3), S. 466–478.
- Abutalebi, Jubin; Green, David W. (2016): Neuroimaging of language control in bilinguals: Neural adaptation and reserve. In: *Bilingualism: Language and cognition* 19 (4), S. 689–698.
- Aguiar, Pilar; Brussino, Silvina; Fernández-Dols, José-Miguel (2013): Psychological distance increases uncompromising consequentialism. In: *Journal of Experimental Social Psychology* 49 (3), S. 449–452.
- Allais, Maurice (1953): Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école américaine. In: *Econometrica* 21 (4), S. 503–546.
- Alter, Adam L.; Oppenheimer, Daniel M. (2006): Predicting short-term stock fluctuations by using processing fluency. In: *Proceedings of the National Academy of Sciences* 103 (24), S. 9369–9372.
- Alter, Adam L.; Oppenheimer, Daniel M. (2008): Effects of fluency on psychological distance and mental construal (or why New York is a large city, but New York is a civilized jungle). In: *Psychological Science* 19 (2), S. 161–167.
- Alter, Adam L.; Oppenheimer, Daniel M.; Epley, Nicholas; Eyre, Rebecca N. (2007): Overcoming intuition: Metacognitive difficulty activates analytic reasoning. In: *Journal of Experimental Psychology: General* 136 (4), S. 569–576.
- Anderson, John R. (1983): *The architecture of cognition*. Cambridge: Harvard University Press (Cognitive science series, 5).
- Anderson, John R. (2013): *Kognitive Psychologie*. 7. Aufl. Berlin: Springer.

Ariely, Dan (2008): Predictably Irrational: The hidden forces that shape our decisions. New York: HarperCollins.

Asay, H. Scott; Elliott, W. Brooke; Rennekamp, Kristina M. (2017): Disclosure readability and the sensitivity of investors' valuation judgments to outside information.

In: *The Accounting Review* 92 (4), S. 1–25.

Asser, Günter (1974): Das Berichtswesen (Analyse - Aufbau - Kontrolle). In: Robert Bobsin (Hg.): Handbuch der Kostenrechnung. 2. Aufl. München: Moderne Industrie, S. 653–678.

Backhaus, Klaus; Erichson, Bernd; Plinke, Wulff; Weiber, Rolf (2018): Multivariate analysemethoden: Eine anwendungsorientierte Einführung. 15. Aufl. Berlin: Springer Gabler.

Baddeley, Alan D. (2003): Working memory: Looking back and looking forward.

In: *Nature Reviews Neuroscience* 4 (10), S. 829–839.

Bahník, Stepan; Englich, Birte; Strack, Fritz (2016): Anchoring effect. In: Rüdiger F. Pohl (Hg.): Cognitive illusions: Intriguing phenomena in judgement, thinking and memory. 2. Aufl. New York: Routledge, S. 223–241.

Baier, Peter (2008): Praxishandbuch Controlling: Controllinginstrumente, Unternehmensplanung und Reporting. 2. Aufl. München: mi-Wirtschaftsbuch.

Balaz, Vladimír; Bačová, Viera; Drobná, Eva; Dudeková, Katarína; Adamik, Kamil (2013): Testing prospect theory parameters. In: *Ekonomický Časopis (Journal of Economics)* 61 (7), S. 655–671.

Ball, Sheryl B.; Cech, Paula-Ann (1996): Subject pool choice and treatment effects in economic laboratory research. In: *Research in Experimental Economics* 6 (3), S. 239–292.

Barberis, Nicholas C. (2013): Thirty years of prospect theory in economics: A review and assessment. In: *Journal of Economic perspectives* 27 (1), S. 173–196.

Bar-Hillel, Maya; Neter, Efrat (1993): How alike is it versus how likely is it: A disjunction fallacy in probability judgments. In: *Journal of Personality and Social Psychology* 65 (6), S. 1119–1131.

Barr, Steve H.; Hitt, Michael A. (1986): A comparison of selection decision models in manager versus student samples. In: *Personnel Psychology* 39 (3), S. 599–617.

Bartels, Daniel M. (2008): Principled moral sentiment and the flexibility of moral judgment and decision making. In: *Cognition* 108 (2), S. 381–417.

Bartz, Tim; Hesse, Martin (2020): Worum es im Wirecard-Skandal geht: Der Total Schaden. Hg. v. Der Spiegel. Online verfügbar unter <https://www.spiegel.de/wirtschaft/unternehmen/worum-es-im-wirecard-skandal-geht-a-96c88437-bdd5-4613-87dc-6f6360868fcc>, zuletzt geprüft am 05.08.2021.

- Basel, Jörn (2012): Heuristic reasoning in management accounting: A mixed methods analysis. Lohmar: Josef Eul (zugl. Dissertation, ESCP Business School Berlin).
- Bazerman, Max H.; Moore, Don A. (2013): Judgment in managerial decision making. 8. Aufl. New York: Wiley.
- Beekun, Rafik I.; Stedham, Yvonne; Westerman, James W.; Yamamura, Jeanne H. (2010): Effects of justice and utilitarianism on ethical decision making: A cross-cultural examination of gender similarities and differences. In: *Business Ethics: A European Review* 19 (4), S. 309–325.
- Beißert, Hanna; Köhler, Meike; Rempel, Marina; Beierlein, Constanze (2015): Kurzsкала Need for Cognition NFC-K. Hg. v. GESIS – Leibniz-Institut für Sozialwissenschaften. Online verfügbar unter https://zis.gesis.org/pdfFiles/Dokumentation/Beisert_%2B_Kurzskala_Need_for_Cognition_NFC-K_c.pdf, zuletzt geprüft am 05.08.2021.
- Bell, David E. (1982): Regret in decision making under uncertainty. In: *Operations Research* 30 (5), S. 961–981.
- Bell, David E. (1985): Disappointment in decision making under uncertainty. In: *Operations Research* 33 (1), S. 1–27.
- Bereby-Meyer, Yoella; Hayakawa, Sayuri L.; Shalvi, Shaul; Corey, Joanna D.; Costa, Albert; Keysar, Boaz (2018): Honesty speaks a second language. In: *Topics in Cognitive Science*, S. 1–12.
- Berkowitz, Leonard; Donnerstein, Edward (1982): External validity is more than skin deep: Some answers to criticisms of laboratory experiments. In: *American psychologist* 37 (3), S. 245–257.
- Bernard, H. Russell (2013): Social research methods: Qualitative and quantitative approaches. 2. Aufl. Los Angeles: Sage.
- Bernoulli, Daniel (1954): Exposition of a new theory on the measurement. In: *Econometrica* 22 (1), S. 23–36.
- Besuglov, Ewgenij; Crasselt, Nils (2021): The effect of readability and language choice in management accounting reports on risk-taking: An experimental study. In: *Journal of Business Economics* 91 (1), S. 5–33.
- Birnberg, Jacob G.; Nath, Raghu (1968): Laboratory experimentation in accounting research. In: *The Accounting Review* 43 (1), S. 38–45.
- Birnberg, Jacob G.; Shields, Jeffrey F. (1989): Three decades of behavioral accounting research: A search for order. In: *Behavioral Research in Accounting* 1, S. 23–74.
- Birnberg, Jacob G.; Shields, Michael D.; Young, S. Mark (1990): The case for multiple methods in empirical management accounting research (with an illustration from budget setting). In: *Journal of management accounting research* 2 (1), S. 33–66.

- Bjork, Robert A. (1994): Memory and metamemory considerations in the training of human beings. In: Janet Metcalfe und Arthur P. Shimamura (Hg.): *Metacognition: Knowing about knowing*. Cambridge: MIT Press, S. 185–205.
- Blais, Ann-Renée; Weber, Elke U. (2006): A domain-specific risk-taking (DOSPERT) scale for adult populations. In: *Judgment and Decision Making* 1 (1), S. 33–47.
- Bleichrodt, Han; Pinto, Jose L. (2000): A parameter-free elicitation of the probability weighting function in medical decision analysis. In: *Management Science* 46 (11), S. 1485–1496.
- Bloech, Jürgen (1994): Berichte in der strategischen Planung. In: Jürgen Bloech, Uwe Götz, Burrkhard Huch, Wolfgang Lücke und Friedhelm Rudolph (Hg.): *Strategische Planung Instrumente, Vorgehensweisen und Informationssysteme*. Heidelberg: Physica, S. 194–208.
- Blohm, Hans (1982): Betriebliches Berichtswesen. In: O. V. (Hg.): *Management Enzyklopädie: Das Managementwissen unserer Zeit*. 2. Aufl. München: Moderne Industrie (1), S. 866–876.
- Blohm, Hans; Heinrich, Lutz J. (1967): *Wie erstellt man einen Bericht? Eine Vier-Stufen-Methode*. Esslingen am Neckar: Friedrich Baierl.
- Bloomfield, Robert J. (2002): The “incomplete revelation hypothesis” and financial reporting. In: *Accounting Horizons* 16 (3), S. 233–243.
- Bonsall IV, Samuel B.; Leone, Andrew J.; Miller, Brian P.; Rennekamp, Kristina M. (2017): A plain English measure of financial reporting readability. In: *Journal of Accounting and Economics* 63 (2-3), S. 329–357.
- Borkowski, Susan C.; Ugras, Yusuf J. (1992): The ethical attitudes of students as a function of age, sex and experience. In: *Journal of Business Ethics* 11 (12), S. 961–979.
- Borzykowski, Bryan (2017): The international companies using only English: A growing number of global firms are using English as their main language - even if they are based in Japan or France. Hg. v. BBC. Online verfügbar unter <http://www.bbc.com/capital/story/20170317-the-international-companies-using-only-english>, zuletzt geprüft am 05.08.2021.
- Branzi, Francesca M.; Della Rosa, Pasquale A.; Canini, Matteo; Costa, Albert; Abutalebi, Jubin (2016): Language control in bilinguals: Monitoring and response selection. In: *Cerebral Cortex* 26 (6), S. 2367–2380.
- Brislin, Richard W. (1970): Back-translation for cross-cultural research. In: *Journal of Cross-Cultural Psychology* 1 (3), S. 185–216.

- Bröder, Arndt (2000): Assessing the empirical validity of the "Take-the-best" heuristic as a model of human probabilistic inference. In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26 (5), S. 1332–1346.
- Busse von Colbe, Walther (2011): Rechnungswesen. In: Walther Busse von Colbe, Nils Crasselt und Bernhard Pellens (Hg.): *Lexikon des Rechnungswesens: Handbuch der Bilanzierung und Prüfung, der Erlös-, Finanz-, Investitions- und Kostenrechnung*. 5. Aufl. München: Oldenbourg, S. 652–656.
- Byrnes, James P.; Miller, David C.; Schafer, William D. (1999): Gender differences in risk taking: A meta-analysis. In: *Psychological Bulletin* 125 (3), S. 367–383.
- Caldwell-Harris, Catherine L. (2014): Emotionality differences between a native and foreign language: Theoretical implications. In: *Frontiers in Psychology* 5, S. 1–4.
- Camerer, Colin F.; Hogarth, Robin M. (1999): The effects of financial incentives in experiments: A review and capital-labor-production framework. In: *Journal of Risk and Uncertainty* 19 (1-3), S. 7–42.
- Cardinaels, Eddy (2008): The interplay between cost accounting knowledge and presentation formats in cost-based decision-making. In: *Accounting, Organizations and Society* 33 (6), S. 582–602.
- Čavar, Franziska; Tytus, Agnieszka E. (2018): Moral judgement and foreign language effect: When the foreign language becomes the second language. In: *Journal of multi-lingual and multicultural development* 39 (1), S. 17–28.
- Chaiken, Shelly; Trope, Yaacov (Hg.) (1999): *Dual-process theories in social psychology*. New York: The Guilford Press.
- Chan, Samuel Y. S.; Leung, Philomena (2006): The effects of accounting students' ethical reasoning and personal factors on their ethical sensitivity. In: *Managerial auditing journal* 21 (4), S. 436–457.
- Cipolletti, Heather; McFarlane, Steven; Weissglass, Christine (2016): The moral foreign-language effect. In: *Philosophical Psychology* 29 (1), S. 23–40.
- Cohen, Jeffrey R.; Pant, Laurie W.; Sharp, David J. (1996): Measuring the ethical awareness and ethical orientation of Canadian auditors. In: *Behavioral Research in Accounting* 8, S. 98–119.
- Cohen, Jeffrey R.; Pant, Laurie W.; Sharp, David J. (2001): An examination of differences in ethical decision-making between Canadian business students and accounting professionals. In: *Journal of Business Ethics* 30 (4), S. 319–336.
- Conway, Paul; Gawronski, Bertram (2013): Deontological and utilitarian inclinations in moral decision making: A process dissociation approach. In: *Journal of Personality and Social Psychology* 104 (2), S. 216–235.

- Corey, Joanna D.; Hayakawa, Sayuri L.; Foucart, Alice; Aparici, Melina; Botella, Juan; Costa, Albert; Keysar, Boaz (2017): Our moral choices are foreign to us. In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 43 (7), S. 1109–1128.
- Costa, Albert; Foucart, Alice; Arnon, Inbal; Aparici, Melina; Apesteguia, Jose (2014a): “Piensa” twice: On the foreign language effect in decision making. In: *Cognition* 130 (2), S. 236–254.
- Costa, Albert; Foucart, Alice; Hayakawa, Sayuri L.; Aparici, Melina; Apesteguia, Jose; Heafner, Joy; Keysar, Boaz (2014b): Your morals depend on language. In: *PloS One* 9 (4), S. 1–7.
- Costa, Albert; Sebastián-Gallés, Núria (2014): How does the bilingual experience sculpt the brain? In: *Nature Reviews Neuroscience* 15 (5), S. 336–345.
- Costa, Albert; Vives, Marc-Lluís; Corey, Joanna D. (2017): On language processing shaping decision making. In: *Current Directions in Psychological Science* 26 (2), S. 146–151.
- Cottell, Philip G.; Perlin, Terry M. (1990): *Accounting ethics: A practical guide for professionals*. New York: Quorum Books.
- Cowan, Nelson (2000): The magical number 4 in short-term memory: A reconsideration of mental storage capacity. In: *Behavioral and brain sciences* 24 (1), S. 87–185.
- Craft, Jana L. (2013): A review of the empirical ethical decision-making literature: 2004–2011. In: *Journal of Business Ethics* 117 (2), S. 221–259.
- De Martino, Benedetto; Kumaran, Dharshan; Seymour, Ben; Dolan, Raymond J. (2006): Frames, biases, and rational decision-making in the human brain. In: *Science* 313 (5787), S. 684–687.
- Deck, Cary; Jahedi, Salar (2015): The effect of cognitive load on economic decision making: A survey and new experiments. In: *European Economic Review* 78, S. 97–119.
- Demski, Joel S.; Feltham, Gerald A. (1976): *Cost determination: A conceptual approach*. Ames: Iowa State University Press.
- Desmoulin-Lebeault, François; Meunier, Luc (2018): Moment risks: Investment for self and for a firm. In: *Decision Analysis* 15 (4), S. 242–266.
- Dewaele, Jean-Marc (2008): The emotional weight of I love you in multilinguals’ languages. In: *Journal of Pragmatics* 40 (10), S. 1753–1780.
- Diemand-Yauman, Connor; Oppenheimer, Daniel M.; Vaughan, Erikka B. (2011): Fortune favors the Bold (and the Italicized): Effects of disfluency on educational outcomes. In: *Cognition* 118 (1), S. 111–115.

- Döring, Nicola; Bortz, Jürgen (2016): *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. 5. Aufl. Berlin: Springer.
- Drews, Hanno (2013): Psychologie für Controller: Wie Bias zu Fehlentscheidungen führen können. In: Andreas Klein (Hg.): *Controller's Performance: Mit welchen Skills Sie überzeugen*. Freiburg: Haufe, S. 45–58.
- Drews, Hanno; Friedrichsen, Morten (2012): Rationalitätssicherung bei betrieblichen Entscheidungen: Bias erkennen und vermindern. In: *Wirtschaftswissenschaftliches Studium* 41 (7), S. 350–354.
- Duch, Raymond M.; Galvez, Tomas; Torres, Felipe (2017): Treatment effects and subject pool diversity in the experimental lab. Working Paper.
- Dylman, Alexandra S.; Champoux-Larsson, Marie-France (2020): It's (not) all Greek to me: Boundaries of the foreign language effect. In: *Cognition* 196, S. 1–8.
- Edwards, Ward (1954): The theory of decision making. In: *Psychological Bulletin* 51 (4), S. 380–417.
- Eichenberger, Reinhard (1992): *Verhaltensanomalien und Wirtschaftswissenschaft: Herausforderung, Reaktionen, Perspektiven*. Wiesbaden: Deutscher Universitäts-Verlag.
- Eisenführ, Franz; Weber, Martin; Langer, Thomas (2010): *Rationales Entscheiden*. 5. Aufl. Berlin: Springer.
- Eitel, Alexander; Kühl, Tim; Scheiter, Katharina; Gerjets, Peter (2014): Disfluency meets cognitive load in multimedia learning: Does harder-to-read mean better-to-understand? In: *Applied Cognitive Psychology* 28 (4), S. 488–501.
- Elliott, W. Brooke; Rennekamp, Kristina M.; White, Brian J. (2015): Does concrete language in disclosures increase willingness to invest? In: *Review of Accounting Studies* 20 (2), S. 839–865.
- Elmes, David G.; Kantowitz, Barry H.; Roediger, Henry L., III (2006): *Research methods in psychology*. 8. Aufl. Belmont: Thomson Wadsworth.
- Epley, Nicholas; Gilovich, Thomas (2006): The anchoring-and-adjustment heuristic: Why the adjustments are insufficient. In: *Psychological Science* 17 (4), S. 311–318.
- Eppler, Martin J.; Mengis, Jeanne (2004): The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines. In: *The Information Society* 20 (5), S. 325–344.
- Epstein, Seymour; Pacini, Rosemary; Denes-Raj, Veronika; Heier, Harriet (1996): Individual differences in intuitive–experiential and analytical–rational thinking styles. In: *Journal of Personality and Social Psychology* 71 (2), S. 390–405.
- Evans, Jonathan St. B. T. (2008): Dual-processing accounts of reasoning, judgment, and social cognition. In: *Annual Review of Psychology* 59, S. 255–278.

- Evans, Jonathan St. B. T.; Barston, Julie L.; Pollard, Paul (1983): On the conflict between logic and belief in syllogistic reasoning. In: *Memory and Cognition* 11 (3), S. 295–306.
- Eyal, Tal; Liberman, Nira; Trope, Yaacov (2008): Judging near and distant virtue and vice. In: *Journal of Experimental Social Psychology* 44 (4), S. 1204–1209.
- Favreau, Micheline; Segalowitz, Norman S. (1983): Automatic and controlled processes in the first-and second-language reading of fluent bilinguals. In: *Memory and Cognition* 11 (6), S. 565–574.
- Few, Stephen (2012): Show me the numbers: Designing tables and graphs to enlighten. 2. Aufl. Burlingame: Analytics Press.
- Figner, Bernd; Mackinlay, Rachael J.; Wilkening, Friedrich; Weber, Elke U. (2009): Affective and deliberative processes in risky choice: Age differences in risk taking in the Columbia Card Task. In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 35 (3), S. 709–730.
- Figner, Bernd; Weber, Elke U. (2011): Who takes risks when and why? Determinants of risk taking. In: *Current Directions in Psychological Science* 20 (4), S. 211–216.
- Franz, Stephan (2004): Grundlagen des ökonomischen Ansatzes: Das Erklärungskonzept des Homo Oeconomicus. Working Paper (2).
- Frederick, Shane (2005): Cognitive reflection and decision making. In: *Journal of Economic perspectives* 19 (4), S. 25–42.
- French, Matthew M. J.; Blood, Arabella; Bright, Naomi D.; Futak, Dez; Grohmann, M. J.; Hasthorpe, Alex et al. (2013): Changing fonts in education: How the benefits vary with ability and dyslexia. In: *The Journal of Educational Research* 106 (4), S. 301–304.
- Frey, Bruno S. (1990): Ökonomie ist Sozialwissenschaft: Die Anwendung der Ökonomie auf neue Gebiete. München: Franz Vahlen.
- Friedman, Milton; Savage, Leonard J. (1952): The expected-utility hypothesis and the measurability of utility. In: *Journal of Political Economy* 60 (6), S. 463–474.
- Gaissmaier, Wolfgang; Neth, Hansjörg (2016): Die Intelligenz einfacher Entscheidungsregeln in einer ungewissen Welt. In: *Controller Magazin* 41 (2), S. 19–26.
- Gao, Shan; Zika, Ondrej; Rogers, Robert D.; Thierry, Guillaume (2015): Second language feedback abolishes the “hot hand” effect during even-probability gambling. In: *The Journal of Neuroscience* 35 (15), S. 5983–5989.
- Geipel, Janet; Hadjichristidis, Constantinos; Surian, Luca (2015a): How foreign language shapes moral judgment. In: *Journal of Experimental Social Psychology* 59, S. 8–17.

- Geipel, Janet; Hadjichristidis, Constantinos; Surian, Luca (2015b): The foreign language effect on moral judgment: The role of emotions and norms. In: *PloS One* 10 (7), S. 1–17.
- Geipel, Janet; Hadjichristidis, Constantinos; Surian, Luca (2016): Foreign language affects the contribution of intentions and outcomes to moral judgment. In: *Cognition* 154, S. 34–39.
- Gervais, Will M.; Norenzayan, Ara (2012): Analytic thinking promotes religious disbelief. In: *Science* 336 (6080), S. 493–496.
- Gigerenzer, Gerd (2004): Fast and frugal heuristics: The tools of bounded rationality. In: Derek J. Koehler und Nigel Harvey (Hg.): *Blackwell handbook of judgment and decision making*. Malden: Blackwell Publishing, S. 62–88.
- Gigerenzer, Gerd (2007): *Bauchentscheidungen: Die Intelligenz des Unbewussten und die Macht der Intuition*. 2. Aufl. München: C. Bertelsmann.
- Gigerenzer, Gerd (2013): *Risiko: Wie man die richtigen Entscheidungen trifft*. 5. Aufl. München: C. Bertelsmann.
- Gigerenzer, Gerd; Gaissmaier, Wolfgang (2006): Denken und Urteilen unter Unsicherheit: Kognitive Heuristiken. In: Joachim Funke (Hg.): *Denken und Problemlösen*. Göttingen: Hogrefe, S. 330–374.
- Gigerenzer, Gerd; Gaissmaier, Wolfgang (2011): Heuristic decision making. In: *Annual Review of Psychology* 62, S. 451–482.
- Gigerenzer, Gerd; Goldstein, Daniel G. (1996): Reasoning the fast and frugal way: Models of bounded rationality. In: *Psychological Review* 103 (4), S. 650–669.
- Gigerenzer, Gerd; Goldstein, Daniel G. (2011): Reasoning the fast and frugal way: Models of bounded rationality. In: Gerd Gigerenzer, Ralph Hertwig und Thorsten Pachur (Hg.): *Heuristics: The foundations of adaptive behavior*. Oxford: Oxford University Press, S. 33–57.
- Gigerenzer, Gerd; Hertwig, Ralph; Pachur, Thorsten (Hg.) (2011): *Heuristics: The foundations of adaptive behavior*. Oxford: Oxford University Press.
- Gigerenzer, Gerd; Todd, Peter M. (1999a): Fast and frugal heuristics: The adaptive toolbox. In: Gerd Gigerenzer und Peter M. Todd (Hg.): *Simple heuristics that make us smart*. Oxford: Oxford University Press, S. 3–34.
- Gigerenzer, Gerd; Todd, Peter M. (Hg.) (1999b): *Simple heuristics that make us smart*. Oxford: Oxford University Press.
- Gillenkirch, Robert M.; Arnold, Markus C. (2008): State of the Art des Behavioral Accounting. In: *Wirtschaftswissenschaftliches Studium* 37 (3), S. 128–134.
- Gilovich, Thomas; Griffin, Dale; Kahneman, Daniel (Hg.) (2002): *Heuristics and biases: The psychology of intuitive judgment*. Cambridge: Cambridge University Press.

- Gilovich, Thomas; Vallone, Robert; Tversky, Amos (1985): The hot hand in basketball: On the misperception of random sequences. In: *Cognitive Psychology* 17 (3), S. 295–314.
- Gleich, Ronald; Temmel, Philipp (2008): Die Rolle und Organisation des Controllings im Management Reporting. In: Ronald Gleich, Peter Horvath und Uwe Michel (Hg.): *Management Reporting: Grundlagen, Praxis und Perspektiven*. Freiburg: Haufe, S. 63–91.
- Gliner, Jeffrey A.; Morgan, George A.; Leech, Nancy L. (2009): *Research methods in applied settings: An integrated approach to design and analysis*. 2. Aufl. New York: Routledge.
- Gold, Natalie; Pulford, Briony D.; Colman, Andrew M. (2013): Your money or your life: Comparing judgements in trolley problems involving economic and emotional harms, injury and death. In: *Economics and Philosophy* 29 (2), S. 213–233.
- Goldstein, Daniel G.; Gigerenzer, Gerd (2002): Models of ecological rationality: The recognition heuristic. In: *Psychological Review* 109 (1), S. 75–90.
- Gong, Han; Medin, Douglas L. (2012): Construal levels and moral judgment: Some complications. In: *Judgment and Decision Making* 7 (5), S. 628–638.
- Gonzalez, Richard; Wu, George (1999): On the shape of the probability weighting function. In: *Cognitive Psychology* 38 (1), S. 129–166.
- Göpfert, Ingrid (2002): Berichtswesen. In: Hans-Ulrich Küpper und Alfred Wagenhofer (Hg.): *Handwörterbuch Unternehmensrechnung und Controlling*. 4. Aufl. Stuttgart: Schäffer-Poeschel, Sp. 143–156.
- Gräf, Jens; Nase, Daniela (2008): Objekte des Management Reportings im Überblick. In: Ronald Gleich, Peter Horvath und Uwe Michel (Hg.): *Management Reporting: Grundlagen, Praxis und Perspektiven*. Freiburg: Haufe, S. 43–61.
- Gravetter, Frederick J.; Forzano, Lori-Ann B. (2018): *Research methods for the behavioral sciences*. 6. Aufl. Boston: Cengage.
- Grayot, James D. (2020): Dual process theories in behavioral economics and neuroeconomics: A critical review. In: *Review of Philosophy and Psychology* 11 (1), S. 105–136.
- Greene, Joshua D. (2008): The secret joke of Kant's soul. In: Walter Sinnott-Armstrong (Hg.): *Moral psychology: The neuroscience of morality: Emotion, brain disorders, and development*. Cambridge: MIT Press (3), S. 35–79.
- Greene, Joshua D.; Cushman, Fiery A.; Stewart, Lisa E.; Lowenberg, Kelly; Nystrom, Leigh E.; Cohen, Jonathan D. (2009): Pushing moral buttons: The interaction between personal force and intention in moral judgment. In: *Cognition* 111 (3), S. 364–371.

- Greene, Joshua D.; Nystrom, Leigh E.; Engell, Andrew D.; Darley, John M.; Cohen, Jonathan D. (2004): The neural bases of cognitive conflict and control in moral judgment. In: *Neuron* 44 (2), S. 389–400.
- Gröschner, Christian; Raab, Markus (2006): Vorhersagen im Fußball: Deskriptive und normative Aspekte von Vorhersagemodellen im Sport. In: *Zeitschrift für Sportpsychologie* 13 (1), S. 23–36.
- Grotjahn, Rüdiger (2002): Konstruktion und Einsatz von C-Tests: Ein Leitfaden für die Praxis. In: Rüdiger Grotjahn (Hg.): *Der C-Test: Theoretische Grundlagen und praktische Anwendungen*. Bochum: AKS-Verlag, S. 211–225.
- Groves, Kevin; Vance, Charles; Paik, Yongsun (2008): Linking linear/nonlinear thinking style balance and managerial ethical decision-making. In: *Journal of Business Ethics* 80 (2), S. 305–325.
- Güth, Werner; Schmittberger, Rolf; Schwarze, Bernd (1982): An experimental analysis of ultimatum bargaining. In: *Journal of Economic Behavior and Organization* 3 (4), S. 367–388.
- Hadjichristidis, Constantinos; Geipel, Janet; Keysar, Boaz (2019): The influence of native language in shaping judgment and choice. In: *Progress in brain research* 247, S. 253–272.
- Hadjichristidis, Constantinos; Geipel, Janet; Savadori, Lucia (2015): The effect of foreign language in judgments of risk and benefit: The role of affect. In: *Journal of Experimental Psychology: Applied* 21 (2), S. 117–129.
- Hadjichristidis, Constantinos; Geipel, Janet; Surian, Luca (2017a): Breaking magic: Foreign language suppresses superstition. In: *The Quarterly Journal of Experimental Psychology*, S. 1–11.
- Hadjichristidis, Constantinos; Geipel, Janet; Surian, Luca (2017b): How foreign language affects decisions: Rethinking the brain-drain model. In: *Journal of International Business Studies* 48 (5), S. 645–651.
- Halberstadt, Anna (1994): On the use of language in psychotherapy with émigrés from the former Soviet Union. In: *Journal of Jewish Communal Service* 71, S. 93–99.
- Harris, Catherine L.; Aycicegi, Ayse; Gleason, Jean B. (2003): Taboo words and reprimands elicit greater autonomic reactivity in a first language than in a second language. In: *Applied Psycholinguistics* 24 (4), S. 561–579.
- Hayakawa, Sayuri L.; Costa, Albert; Foucart, Alice; Keysar, Boaz (2016): Using a foreign language changes our choices. In: *Trends in Cognitive Sciences* 20 (11), S. 791–793.

Hayakawa, Sayuri L.; Lau, Becky K. Y.; Holtzmann, Sophie; Costa, Albert; Keysar, Boaz (2017a): On the reliability of the foreign language effect on risk-taking. In: *The Quarterly Journal of Experimental Psychology*, S. 1–12.

Hayakawa, Sayuri L.; Tannenbaum, David; Costa, Albert; Corey, Joanna D.; Keysar, Boaz (2017b): Thinking more or feeling less? Explaining the foreign-language effect on moral judgment. In: *Psychological Science* 28 (10), S. 1387–1397.

Heatherston, Todd F. (2011): Neuroscience of self and self-regulation. In: *Annual Review of Psychology* 62, S. 363–390.

Hertwig, Ralph; Gigerenzer, Gerd (1999): The ‘conjunction fallacy’ revisited: How intelligent inferences look like reasoning errors. In: *Journal of Behavioral Decision Making* 12 (4), S. 275–305.

Hilbig, Benjamin E. (2010): Reconsidering “evidence” for fast-and-frugal heuristics. In: *Psychonomic Bulletin & Review* 17 (6), S. 923–930.

Hirsch, Bernhard (2007): Controlling und Entscheidungen: Zur verhaltenswissenschaftlichen Fundierung des Controllings. Tübingen: Mohr Siebeck.

Hirsch, Bernhard (2008): Zur Integration psychologischen Wissens in betriebswirtschaftliche Controlling-Konzeptionen: Stand der Literatur und Forschungsbedarf. In: *Controlling & Management* 52 (1), S. 40–49.

Hirsch, Bernhard (2009): Controlling und experimentelle Forschung. In: Andreas G. Scherer, Ina M. Kaufmann und Moritz Patzer (Hg.): *Methoden in der Betriebswirtschaftslehre*. Wiesbaden: Gabler, S. 167–186.

Hirsch, Bernhard; Paefgen, Anne; Schaier, Sven (2008): Gestaltung von Monatsberichten in deutschen Großunternehmen. In: *Controlling & Management* 52 (5), S. 326–332.

Hirsch, Bernhard; Seubert, Anna; Sohn, Matthias (2015): Visualisation of data in management accounting reports: How supplementary graphs improve every-day management judgments. In: *Journal of Applied Accounting Research* 16 (2), S. 221–239.

Hirsch, Bernhard; Volnhals, Martina (2012): Information Overload im betrieblichen Berichtswesen: Ein unterschätztes Phänomen. In: *Die Betriebswirtschaft* 72 (1), S. 23–55.

Hochhold, Stefanie; Rudolph, Bernd (2009): Principal-Agent-Theorie. In: Manfred Schwaiger und Anton Meyer (Hg.): *Theorien und Methoden der Betriebswirtschaft: Handbuch für Wissenschaftler und Studierende*. München: Vahlen, S. 132–145.

Hofmann, Wilhelm; Friese, Malte; Schmeichel, Brandon J.; Baddeley, Alan D. (2010): Working memory and self-regulation. In: Kathleen D. Vohs und Roy F. Baumeister (Hg.): *Handbook of self-regulation: Research, theory, and applications*. New York: Guilford Press, S. 204–225.

- Holt, Charles A.; Laury, Susan K. (2002): Risk aversion and incentive effects. In: *American economic review* 92 (5), S. 1644–1655.
- Homann, Karl; Suchanek, Andreas (2005): *Ökonomik: Eine Einführung*. 2. Aufl. Tübingen: Mohr Siebeck.
- Horvath, Peter; Gleich, Ronald; Seiter, Mischa (2020): *Controlling*. 14. Aufl. München: Vahlen.
- Hoskisson, Robert E.; Chirico, Francesco; Zyung, Jinyong; Gambeta, Eni (2016): Managerial risk taking: A multitheoretical review and future research agenda. In: *Journal of Management* 43 (1), S. 137–169.
- Hsu, Chun-Ting; Jacobs, Arthur M.; Conrad, Markus (2015): Can Harry Potter still put a spell on us in a second language? An fMRI study on reading emotion-laden literature in late bilinguals. In: *Cortex* 63, S. 282–295.
- International Group of Controlling (2011): *Controlling-Prozessmodell: Ein Leitfaden für die Beschreibung und Gestaltung von Controlling-Prozessen*. Online verfügbar unter https://www.igc-controlling.org/fileadmin/downloads/Standards/Controlling_Prozessmodell.pdf, zuletzt geprüft am 05.08.2021.
- Iselin, Errol R. (1993): The effects of the information and data properties of financial ratios and statements on managerial decision quality. In: *Journal of Business Finance & Accounting* 20 (2), S. 249–266.
- Jacoby, Larry L. (1991): A process dissociation framework: Separating automatic from intentional uses of memory. In: *Journal of Memory and Language* 30 (5), S. 513–541.
- Jensen, Michael C.; Meckling, William H. (1976): Theory of the firm: managerial behavior, agency costs and ownership structure. In: *Journal of Financial Economics* 3 (4), S. 305–360.
- Jung, Hans (2011): *Controlling*. 2. Aufl. München: Oldenbourg.
- Kahneman, Daniel (2003): A perspective on judgment and choice: Mapping bounded rationality. In: *American psychologist* 58 (9), S. 697–720.
- Kahneman, Daniel (2014): *Schnelles Denken, langsames Denken*. 5. Aufl. München: Pantheon.
- Kahneman, Daniel; Frederick, Shane (2002): Representativeness revisited: Attribute substitution in intuitive judgment. In: Thomas Gilovich, Dale Griffin und Daniel Kahneman (Hg.): *Heuristics and biases: The psychology of intuitive judgment*. Cambridge: Cambridge University Press, S. 49–81.
- Kahneman, Daniel; Lovallo, Dan; Sibony, Olivier (2011): Before you make that big decision. In: *Harvard Business Review* 89 (6), S. 50–60.

- Kahneman, Daniel; Slovic, Paul; Tversky, Amos (Hg.) (1982): Judgment under uncertainty: Heuristics and biases. Cambridge: Cambridge University Press.
- Kammer, Karsten (2005): Reporting internationaler Unternehmen: Auswirkungen der Harmonisierung und der Konvergenz des Rechnungswesens in Europa. Wiesbaden: Gabler.
- Karoly, Paul (1993): Mechanisms of self-regulation: A systems view. In: *Annual Review of Psychology* 44 (1), S. 23–52.
- Keppel, Geoffrey; Wickens, Thomas D. (2004): Design and analysis: A researcher's handbook. 4. Aufl. London: Pearson Prentice Hall.
- Keysar, Boaz; Hayakawa, Sayuri L.; An, Sun G. (2012): The foreign-language effect: Thinking in a foreign tongue reduces decision biases. In: *Psychological Science* 23 (6), S. 661–668.
- Kilka, Michael; Weber, Martin (2001): What determines the shape of the probability weighting function under uncertainty? In: *Management Science* 47 (12), S. 1712–1726.
- Kirchgässner, Gebhard (2008): Homo oeconomicus: The economic model of behaviour and its applications in economics and other social sciences. New York: Springer.
- Kirsch, Werner (1970): Entscheidungsprozesse. Wiesbaden: Gabler (Verhaltenswissenschaftliche Ansätze der Entscheidungstheorie, 1).
- Kirsch, Werner (1994): Die Handhabung von Entscheidungsproblemen: Einführung in die Theorie der Entscheidungsprozesse. 4. Aufl. Herrsching: Barbara Kirsch.
- Klesse, Anne-Kathrin; Levav, Jonathan; Goukens, Caroline (2015): The effect of preference expression modality on self-control. In: *Journal of Consumer Research* 42 (4), S. 535–550.
- Klibanoff, Peter; Marinacci, Massimo; Mukerji, Sujoy (2005): A smooth model of decision making under ambiguity. In: *Econometrica* 73 (6), S. 1849–1892.
- Koch, Rembert (1994): Betriebliches Berichtswesen als Informations- und Steuerungsinstrument. Frankfurt am Main: Peter Lang.
- Koch, Thomas; Peter, Christina; Müller, Philipp (2019): Das Experiment in der Kommunikations- und Medienwissenschaft: Grundlagen, Durchführung und Auswertung experimenteller Forschung. Wiesbaden: Springer VS.
- Kroeber-Riel, Werner (1987): Weniger Informationsüberlastung durch Bildkommunikation: Zur Verwendung von Bildern in kommerziellen Kommunikationssystemen. In: *Wirtschaftswissenschaftliches Studium* 16 (10), S. 485–489.
- Kühberger, Anton (1994): Risiko und Unsicherheit: Zum Nutzen des Subjective Expected Utility-Modells. In: *Psychologische Rundschau* 45 (1), S. 3–23.

- Kühl, Tim; Eitel, Alexander; Damnik, Gregor; Körndle, Hermann (2014): The impact of disfluency, pacing, and students' need for cognition on learning with multimedia. In: *Computers in Human Behavior* 35, S. 189–198.
- Küpper, Hans-Ulrich; Friedl, Gunther; Hofmann, Christian; Hofmann, Yvette; Pedell, Burkhard (2013): *Controlling: Konzeption, Aufgaben, Instrumente*. 6. Aufl. Stuttgart: Schäffer-Poeschel.
- Lang, Matthias (2009): Normative Entscheidungstheorie. In: Manfred Schwaiger und Anton Meyer (Hg.): *Theorien und Methoden der Betriebswirtschaft: Handbuch für Wissenschaftler und Studierende*. München: Vahlen, S. 161–176.
- Langer, Astrid; Rogowski, Wolf (2009): Deskriptive Entscheidungstheorie. In: Manfred Schwaiger und Anton Meyer (Hg.): *Theorien und Methoden der Betriebswirtschaft: Handbuch für Wissenschaftler und Studierende*. München: Vahlen, S. 179–191.
- Laux, Helmut; Gillenkirch, Robert M.; Schenk-Mathes, Heike Y. (2014): *Entscheidungstheorie*. 9. Aufl. Berlin: Springer.
- Lee, Wayne (1977): *Psychologische Entscheidungstheorie*. Weinheim: Beltz.
- Lehavy, Reuven; Li, Feng; Merkley, Kenneth (2011): The effect of annual report readability on analyst following and the properties of their earnings forecasts. In: *The Accounting Review* 86 (3), S. 1087–1115.
- Lehmann, Janina; Goussios, Christina; Seufert, Tina (2016): Working memory capacity and disfluency effect: An aptitude-treatment-interaction study. In: *Metacognition and Learning* 11 (1), S. 89–105.
- Lerner, Eva; Streicher, Bernhard; Sachs, Rainer; Raue, Martina; Frey, Dieter (2015): The effect of construal level on risk-taking. In: *European Journal of Social Psychology* 45 (1), S. 99–109.
- Li, Feng (2008): Annual report readability, current earnings, and earnings persistence. In: *Journal of Accounting and Economics* 45 (2-3), S. 221–247.
- Li, Ping; Sepanski, Sara; Zhao, Xiaowei (2006): Language history questionnaire: A Web-based interface for bilingual research. In: *Behavior Research Methods* 38 (2), S. 202–210.
- Lieberman, Nira; Trope, Yaacov (2008): The psychology of transcending the here and now. In: *Science* 322 (5905), S. 1201–1205.
- Loe, Terry W.; Ferrell, Linda; Mansfield, Phylis (2000): A review of empirical studies assessing ethical decision making in business. In: *Journal of Business Ethics* 25 (3), S. 185–204.
- Loewenstein, George F. (1996): Out of control: Visceral influences on behavior. In: *Organizational Behavior and Human Decision Processes* 65 (3), S. 272–292.

- Loewenstein, George F.; Weber, Elke U.; Hsee, Christopher K.; Welch, Ned (2001): Risk as feelings. In: *Psychological Bulletin* 127 (2), S. 267–286.
- Losecaat Vermeer, Annabel B.; Boksem, Maarten A. S.; Sanfey, Alan G. (2020): Third-party decision-making under risk as a function of prior gains and losses. In: *Journal of Economic Psychology* 77, S. 1–12.
- Loughran, Tim; McDonald, Bill (2014): Measuring readability in financial disclosures. In: *The Journal of Finance* 69 (4), S. 1643–1671.
- Low, Angie (2009): Managerial risk-taking behavior and equity-based compensation. In: *Journal of Financial Economics* 92 (3), S. 470–490.
- Luce, Robert D.; Raiffa, Howard (1957): *Games and decisions: Introduction and critical survey*. New York: Wiley-VCH.
- Lurie, Nicholas H.; Mason, Charlotte H. (2007): Visual representation: Implications for decision making. In: *Journal of Marketing* 71 (1), S. 160–177.
- Marcos, Luis R. (1976): Bilinguals in psychotherapy: Language as an emotional barrier. In: *American journal of Psychotherapy* 30 (4), S. 552–560.
- Marian, Viorica; Neisser, Ulric (2000): Language-dependent recall of autobiographical memories. In: *Journal of Experimental Psychology: General* 129 (3), S. 361–368.
- McGlone, Matthew S.; Tofiqbakhsh, Jessica (2000): Birds of a feather flock conjointly (?): Rhyme as reason in aphorisms. In: *Psychological Science* 11 (5), S. 424–428.
- McLeod, Peter; Dienes, Zoltan (1996): Do fielders know where to go to catch the ball or only how to get there? In: *Journal of Experimental Psychology: Human Perception and Performance* 22 (3), S. 531–543.
- Meyer, Andrew; Frederick, Shane; Burnham, Terence C.; Guevara Pinto, Juan D.; Boyer, Ty W.; Ball, Linden J. et al. (2015): Disfluent fonts don't help people solve math problems. In: *Journal of Experimental Psychology: General* 144 (2), S. 16–30.
- Meyer, Florian (2009): Spieltheorie und ihre Anwendung in der BWL. In: Manfred Schwaiger und Anton Meyer (Hg.): *Theorien und Methoden der Betriebswirtschaft: Handbuch für Wissenschaftler und Studierende*. München: Vahlen, S. 207–224.
- Meyer-Fünffinger, Arne (2020): Skandal bei Volkswagen: Fünf Jahre "Dieselgate" - und heute? Hg. v. Tagesschau. Online verfügbar unter <https://www.tagesschau.de/wirtschaft/fuenf-jahre-dieselgate-101.html>, zuletzt geprüft am 05.08.2021.
- Miller, Brian P. (2010): The effects of reporting complexity on small and large investor trading. In: *The Accounting Review* 85 (6), S. 2107–2143.
- Morewedge, Carey K.; Kahneman, Daniel (2010): Associative processes in intuitive judgment. In: *Trends in Cognitive Sciences* 14 (10), S. 435–440.

- Mousavi, Shabnam; Gigerenzer, Gerd (2014): Risk, uncertainty, and heuristics. In: *Journal of Business Research* 67 (8), S. 1671–1678.
- Muda, Rafal; Niszczoła, Paweł; Bialek, Michał; Conway, Paul (2018): Reading dilemmas in a foreign language reduces both deontological and utilitarian response tendencies. In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 44 (2), S. 321–326.
- Neth, Hansjörg (2014): Wenn weniger mehr ist: Das Potenzial einfacher Heuristiken in Controlling und Management Reporting. In: Andreas Klein und Jens Gräf (Hg.): *Reporting und Business Intelligence: Berichte klar gestalten, effizient erstellen, prägnant kommentieren*. 2. Aufl. Freiburg: Haufe, S. 43–57.
- Neumann, John von; Morgenstern, Oskar (1947): *Theory of games and economic behavior*. 2. Aufl. Princeton: Princeton University Press.
- Newell, Ben R.; Weston, Nicola J.; Shanks, David R. (2003): Empirical tests of a fast-and-frugal heuristic: Not everyone “takes-the-best”. In: *Organizational Behavior and Human Decision Processes* 91 (1), S. 82–96.
- Norman, Elisabeth (2010): “The unconscious” in current psychology. In: *European Psychologist* 15 (3), S. 193–201.
- Oganian, Yulia; Korn, Christoph W.; Heekeren, Hauke R. (2016): Language switching – but not foreign language use per se – reduces the framing effect. In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 42 (1), S. 140–148.
- Ohlert, Christine R.; Weißenberger, Barbara E. (2015): Beating the base-rate fallacy: An experimental approach on the effectiveness of different information presentation formats. In: *Journal of Management Control* 26 (1), S. 51–80.
- Oppenheimer, Daniel M. (2006): Consequences of erudite vernacular utilized irrespective of necessity: Problems with using long words needlessly. In: *Applied Cognitive Psychology* 20 (2), S. 139–156.
- Oppenheimer, Daniel M. (2008): The secret life of fluency. In: *Trends in Cognitive Sciences* 12 (6), S. 237–241.
- Pavlenko, Aneta (2012): Affective processing in bilingual speakers: Disembodied cognition? In: *International Journal of Psychology* 47 (6), S. 405–428.
- Pfister, Hans-Rüdiger (1994): Noch immer von Nutzen – das SEU-Modell: Kommentar zu Kühbergers Artikel über den Stellenwert des Subjective Expected Utility-Modells (SEU). In: *Psychologische Rundschau* 45 (3), S. 157–160.
- Pfister, Hans-Rüdiger; Jungermann, Helmut; Fischer, Katrin (2017): *Die Psychologie der Entscheidung: Eine Einführung*. 4. Aufl. Berlin: Springer.
- Picot, Arnold (1975): *Experimentelle Organisationsforschung: Methodische und wissenschaftstheoretische Grundlagen*. Wiesbaden: Gabler.

- Plous, Scott (1993): *The psychology of judgment and decision making*. New York: McGraw-Hill.
- Pollmann, Monique M. H.; Potters, Jan; Trautmann, Stefan T. (2014): Risk taking by agents: The role of ex-ante and ex-post accountability. In: *Economics Letters* 123 (3), S. 387–390.
- Pollmann, Rainer; Rühm, Peter (2007): *Controlling-Berichte professionell gestalten*. Freiburg: Haufe.
- Polonioli, Andrea (2018): A blind spot in research on foreign language effects in judgment and decision-making. In: *Frontiers in Psychology* 9 (227), S. 1–5.
- Preacher, Kristopher J.; Hayes, Andrew F. (2004): SPSS and SAS procedures for estimating indirect effects in simple mediation models. In: *Behavior Research Methods, Instruments, & Computers* 36 (4), S. 717–731.
- Raab, Markus; Gula, Bartosz; Gigerenzer, Gerd (2012): The hot hand exists in volleyball and is used for allocation decisions. In: *Journal of Experimental Psychology: Applied* 18 (1), S. 81–94.
- Raatz, Ulrich; Klein-Braley, Christine (2002): Introduction to language testing and to C-Test. In: James A. Coleman, Rüdiger Grotjahn und Ulrich Raatz (Hg.): *University language testing and the C-test*. Bochum: AKS-Verlag, S. 75–91.
- Raue, Martina; Streicher, Bernhard; Lermer, Eva; Frey, Dieter (2015): How far does it feel? Construal level and decisions under risk. In: *Journal of Applied Research in Memory and Cognition* 4 (3), S. 256–264.
- Ray, William (1992): *Methods toward a science of behavior and experience*. 4. Aufl. Pacific Grove: Brooks/Cole Publishing Company.
- Reber, Rolf (2016): Availability. In: Rüdiger F. Pohl (Hg.): *Cognitive illusions: Intriguing phenomena in judgement, thinking and memory*. 2. Aufl. New York: Routledge, S. 185–203.
- Rennekamp, Kristina M. (2012): Processing fluency and investors' reactions to disclosure readability. In: *Journal of Accounting Research* 50 (5), S. 1319–1354.
- Rhodes, Matthew G.; Castel, Alan D. (2008): Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. In: *Journal of Experimental Psychology: General* 137 (4), S. 615–625.
- Ross, Michael; Sicoly, Fiore (1979): Egocentric biases in availability and attribution. In: *Journal of Personality and Social Psychology* 37 (3), S. 322–336.
- Rubin, Allen; Babbie, Earl R. (2013): *Essential research methods for social work*. 3. Aufl. Boston: Cengage.
- Russo, J. Edward; Schoemaker, Paul J. H. (1989): *Decision traps: The ten barriers to brilliant decision-making and how to overcome them*. New York: Fireside.

- Sagrignano, Michael D.; Trope, Yaacov; Liberman, Nira (2002): Time-dependent gambling: Odds now, money later. In: *Journal of Experimental Psychology: General* 131 (3), S. 364–376.
- Sarris, Viktor; Reiß, Siegbert (2005): Kurzer Leitfaden der Experimentalpsychologie. München: Pearson Studium.
- Scandura, Terri A.; Williams, Ethlyn A. (2000): Research methodology in management: Current practices, trends, and implications for future research. In: *Academy of Management Journal* 43 (6), S. 1248–1264.
- Schäffer, Utz; Weber, Jürgen (2015): Controlling: Trends & Benchmarks. Vallendar: WHU-Otto Beisheim School of Management.
- Schäffer, Utz; Weber, Jürgen (2016): Wirklich rationale Entscheidungen: Die nächste Herausforderung für das Controlling. In: *Controller Magazin* (2), S. 8–13.
- Schank, Roger C.; Ableson, Robert P. (1995): Knowledge and memory: The real story. In: Robert S. Wyer (Hg.): *Advances in social cognition*, Bd. 8. Hillsdale: Erlbaum, S. 1–85.
- Schedlinsky, Ivo; Sommer, Friedrich; Wöhrmann, Arnt (2016): Risk-taking in tournaments: An experimental analysis. In: *Journal of Business Economics* 86 (8), S. 837–866.
- Scherpereel, Peter; Gaul, Julian; Muhr, Martin (2015): Entscheidungsverhalten bei Investitionen steuern. In: *Controlling & Management Review* 59 (2), S. 32–39.
- Schmeidler, David (1989): Subjective probability and expected utility without additivity. In: *Econometrica* 57 (3), S. 571–587.
- Schmidt, Horst (1980): Organisation des Berichtswesen. In: Erwin Grochla (Hg.): *Handwörterbuch der Organisation*. 2. Aufl. Stuttgart: Poeschel, Sp. 320–330.
- Schnell, Rainer; Hill, Paul B.; Esser, Elke (2018): *Methoden der empirischen Sozialforschung*. 11. Aufl. Berlin: De Gruyter Oldenbourg.
- Scholl, Wolfgang (1992): Informationspathologien. In: Erich Frese (Hg.): *Handwörterbuch der Organisation*. 3. Aufl. Stuttgart: Poeschel, Sp. 900–912.
- Schrauf, Robert W. (2009): English use among older bilingual immigrants in linguistically concentrated neighborhoods: Social proficiency and internal speech as intracultural variation. In: *Journal of Cross-Cultural Gerontology* 24 (2), S. 157–179.
- Schrauf, Robert W.; Rubin, David C. (2000): Internal languages of retrieval: The bilingual encoding of memories for the personal past. In: *Memory and Cognition* 28 (4), S. 616–623.
- Schroeter, Bernhard (2002): *Operatives Controlling: Aufgaben, Objekte, Instrumente*. Wiesbaden: Gabler.

- Schwartz, Mark S. (2017): *Business ethics: An ethical decision-making approach*. Malden: Wiley Blackwell.
- Schwarz, Norbert; Bless, Herbert; Strack, Fritz; Klumpp, Gisela; Rittenauer-Schatka, Helga; Simons, Annette (1991): Ease of retrieval as information: Another look at the availability heuristic. In: *Journal of Personality and Social Psychology* 61 (2), S. 195–202.
- Schwering, Anja; Uepping, Florian (2017): Behavioral Management Accounting: Was der Controller vom Psychologen lernen kann. In: Andreas Hoffjan, Thorsten Knauer und Andreas Wömpener (Hg.): *Controlling: Konzeptionen – Instrumente – Anwendungen*. Stuttgart: Schäffer-Poeschel, S. 155–167.
- Securities and Exchange Commission (1998): A plain English handbook: How to create clear SEC disclosure documents. Online verfügbar unter <https://www.sec.gov/pdf/handbook.pdf>, zuletzt geprüft am 05.08.2021.
- Sedlmeier, Peter; Renkewitz, Frank (2013): *Forschungsmethoden und Statistik für Psychologen und Sozialwissenschaftler*. 2. Aufl. München: Pearson.
- Service, Elisabet; Simola, Marjut; Metsanheimo, Oili; Maury, Sini (2002): Bilingual working memory span is affected by language skill. In: *European Journal of Cognitive Psychology* 14 (3), S. 383–408.
- Serwe, Sascha; Frings, Christian (2006): Who will win Wimbledon? The recognition heuristic in predicting sports events. In: *Journal of Behavioral Decision Making* 19 (4), S. 321–332.
- Seufert, Tina; Wagner, Felix; Westphal, Julia (2017): The effects of different levels of disfluency on learning outcomes and cognitive load. In: *Instructional Science* 45 (2), S. 221–238.
- Shadish, William R.; Cook, Thomas D.; Campbell, Donald T. (2002): *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin Company.
- Shafir, Eldar; LeBoeuf, Robyn A. (2002): Rationality. In: *Annual Review of Psychology* 53 (1), S. 491–517.
- Simon, Herbert A. (1955): A behavioral model of rational choice. In: *The Quarterly Journal of Economics* 69 (1), S. 99–118.
- Simon, Herbert A. (1956): Rational choice and the structure of the environment. In: *Psychological Review* 63 (2), S. 129–138.
- Sloman, Steven A. (1996): The empirical case for two systems of reasoning. In: *Psychological Bulletin* 119 (1), S. 3–22.

- Song, Hyunjin; Schwarz, Norbert (2008): Fluency and the detection of misleading questions: Low processing fluency attenuates the Moses illusion. In: *Social Cognition* 26 (6), S. 791–799.
- Sprinkle, Geoffrey B.; Williamson, Michael G. (2007): Experimental research in managerial accounting. In: Christopher S. Chapman, Anthony G. Hopwood und Michael D. Shields (Hg.): *Handbook of management accounting research*. Amsterdam: Elsevier Science (1), S. 415–444.
- Standing, Lionel (1973): Learning 10000 pictures. In: *The Quarterly Journal of Experimental Psychology* 25 (2), S. 207–222.
- Stanovich, Keith E. (2011): *Rationality and the reflective mind*. Oxford: Oxford University Press.
- Stanovich, Keith E.; West, Richard F. (2000): Individual differences in reasoning: Implications for the rationality debate? In: *Behavioral and brain sciences* 23 (5), S. 645–665.
- Stier, Winfried (1999): *Empirische Forschungsmethoden*. 2. Aufl. Berlin: Springer.
- Stott, Henry P. (2006): Cumulative prospect theory's functional menagerie. In: *Journal of Risk and Uncertainty* 32 (2), S. 101–130.
- Suchanek, Andreas; Kerscher, Klaus-Jürgen (2007): Der Homo oeconomicus: Verfehltes Menschenbild oder leistungsfähiges Analyseinstrument? In: Rainhart Lang und Annett Schmidt (Hg.): *Individuum und Organisation: Neue Trends eines organisationalwissenschaftlichen Forschungsfeldes*. Wiesbaden: Deutscher Universitäts-Verlag, S. 251–275.
- Sungkhasettee, Victor W.; Friedman, Michael C.; Castel, Alan D. (2011): Memory and metamemory for inverted words: Illusions of competency and desirable difficulties. In: *Psychonomic Bulletin & Review* 18 (5), S. 973–978.
- Suter, Renata S.; Hertwig, Ralph (2011): Time and moral judgment. In: *Cognition* 119 (3), S. 454–458.
- Sweeney, Breda; Costello, Fiona (2009): Moral intensity and ethical decision-making: An empirical examination of undergraduate accounting and business students. In: *Accounting Education: an international journal* 18 (1), S. 75–97.
- Sweller, John (1994): Cognitive load theory, learning difficulty, and instructional design. In: *Learning and instruction* 4 (4), S. 295–312.
- Sweller, John; Ayres, Paul; Kalyuga, Slava (2011): *Cognitive load theory*. New York: Springer.
- Sweller, John; van Merriënboer, Jeroen J. G.; Paas, Fred G. W. C. (1998): Cognitive architecture and instructional design. In: *Educational Psychology Review* 10 (3), S. 251–296.

- Swieringa, Robert J.; Weick, Karl E. (1982): An assessment of laboratory experiments in accounting. In: *Journal of Accounting Research* 20, S. 56–101.
- Takano, Yohtaro; Noda, Akiko (1993): A temporary decline of thinking ability during foreign language processing. In: *Journal of Cross-Cultural Psychology* 24 (4), S. 445–462.
- Tan, Hun-Tong; Wang, Elaine; Zhou, Bo (2014): When the use of positive language backfires: The joint effect of tone, readability, and investor sophistication on earnings judgments. In: *Journal of Accounting Research* 52 (1), S. 273–302.
- Taschner, Andreas (2013a): Management Reporting für Praktiker: Lösungsorientierte Kompaktedition. Wiesbaden: Springer.
- Taschner, Andreas (2013b): Management Reporting: Erfolgsfaktor internes Berichtswesen. Wiesbaden: Springer.
- Taschner, Andreas (2015): Management Reporting und Behavioral Accounting: Verhaltenswirkungen des Berichtswesens im Unternehmen. Wiesbaden: Springer.
- Teigen, Karl H. (2016): Judgments by representativeness. In: Rüdiger F. Pohl (Hg.): *Cognitive illusions: Intriguing phenomena in judgement, thinking and memory*. 2. Aufl. New York: Routledge, S. 204–222.
- Thompson, Valerie A. (2009): Dual-process theories: A metacognitive perspective. In: Jonathan St. B. T. Evans und Keith Frankish (Hg.): *In two minds: Dual processes and beyond*. Oxford: Oxford University Press, S. 171–195.
- Thomson, Judith J. (1985): The trolley problem. In: *The Yale Law Journal* 94 (6), S. 1395–1415.
- Tietze, Susanne (2008): *International management and language*. New York: Routledge.
- Tietzel, Manfred (1981): Die Rationalitätsannahme in den Wirtschaftswissenschaften: Der homo oeconomicus und seine Verwandten. In: *Jahrbuch für Sozialwissenschaft* 32 (2), S. 115–138.
- Todd, Peter M. (2001): Fast and frugal heuristics for environmentally bounded minds. In: Gerd Gigerenzer und Reinhard Selten (Hg.): *Bounded rationality: The adaptive toolbox*. Cambridge: MIT Press, S. 51–70.
- Trémolière, Bastien; Bonnefon, Jean-François (2014): Efficient kill–save ratios ease up the cognitive demands on counterintuitive moral utilitarianism. In: *Personality and Social Psychology Bulletin* 40 (7), S. 923–930.
- Trémolière, Bastien; De Neys, Wim; Bonnefon, Jean-François (2012): Mortality salience and morality: Thinking about death makes people less utilitarian. In: *Cognition* 124 (3), S. 379–384.

- Trope, Yaacov; Liberman, Nira (2010): Construal-level theory of psychological distance. In: *Psychological Review* 117 (2), S. 440–463.
- Trotman, Ken T.; Tan, Hwee C.; Ang, Nicole (2011): Fifty-year overview of judgment and decision-making research in accounting. In: *Accounting & Finance* 51 (1), S. 278–360.
- Tversky, Amos (1969): Intransitivity of preferences. In: *Psychological Review* 76 (1), S. 31–48.
- Tversky, Amos; Fox, Craig R. (1995): Weighing risk and uncertainty. In: *Psychological Review* 102 (2), S. 269–283.
- Tversky, Amos; Kahneman, Daniel (1973): Availability: A heuristic for judging frequency and probability. In: *Cognitive Psychology* 5 (2), S. 207–232.
- Tversky, Amos; Kahneman, Daniel (1974): Judgment under uncertainty: Heuristics and biases. In: *Science* 185 (4157), S. 1124–1131.
- Tversky, Amos; Kahneman, Daniel (1979): Prospect theory: An analysis of decision under risk. In: *Econometrica* 47 (2), S. 263–291.
- Tversky, Amos; Kahneman, Daniel (1981): The framing of decisions and the psychology of choice. In: *Science* 211 (4481), S. 453–458.
- Tversky, Amos; Kahneman, Daniel (1982): Judgment under uncertainty: Heuristics and biases. In: Daniel Kahneman, Paul Slovic und Amos Tversky (Hg.): *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press, S. 3–20.
- Tversky, Amos; Kahneman, Daniel (1983): Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. In: *Psychological Review* 90 (4), S. 293–315.
- Tversky, Amos; Kahneman, Daniel (1992): Advances in prospect theory: Cumulative representation of uncertainty. In: *Journal of Risk and Uncertainty* 5 (4), S. 297–323.
- Van Boven, Leaf; Kane, Joanne; McGraw, A. Peter; Dale, Jeannette (2010): Feeling close: Emotional intensity reduces perceived psychological distance. In: *Journal of Personality and Social Psychology* 98 (6), S. 872–885.
- Vives, Marc-Lluís; Aparici, Melina; Costa, Albert (2018): The limits of the foreign language effect on decision-making: The case of the outcome bias and the representativeness heuristic. In: *PloS One* 13 (9), S. 1–14.
- Volk, Stefan; Köhler, Tine; Pudelko, Markus (2014): Brain drain: The cognitive neuroscience of foreign language processing in multinational corporations. In: *Journal of International Business Studies* 45 (7), S. 862–885.

- Volnhals, Martina (2010): Information Overload und Controlling: Analyse kognitiver Restriktionen bei der Wahrnehmung von Berichtsinformationen. Hamburg: Verlag Dr. Kovač (zugl. Dissertation, Universität der Bundeswehr München).
- Volz, Kirsten G.; Gigerenzer, Gerd (2012): Cognitive processes in decisions under risk are not the same as in decisions under uncertainty. In: *Frontiers in Neuroscience* 6 (105), S. 1–6.
- Waniczek, Mirko; Feichter, Andreas; Schwarzl, Patrick; Eisl, Christoph (2018): Management Reporting: Berichte wirksam und adressatengerecht gestalten. Wien: Linde.
- Warrington, Elizabeth K.; McCarthy, Rosaleen A. (1988): The fractionation of retrograde amnesia. In: *Brain and Cognition* 7 (2), S. 184–200.
- Weber, Jürgen (2013): Verhaltensorientiertes Controlling: Plädoyer für eine (nicht ganz) neue Sicht auf das Controlling. In: *Controlling* 25 (4/5), S. 217–222.
- Weber, Jürgen (2014): Verhaltensanalyse im Controlling: Durch psychologische Erkenntnisse den Unternehmenserfolg steigern. Weinheim: Wiley-VCH (Advanced Controlling, 91).
- Weber, Jürgen; Hirsch, Bernhard; Linder, Stefan; Zayer, Eric (2003): Verhaltensorientiertes Controlling: Der Mensch im Mittelpunkt. Vallendar: WHU, Otto-Beisheim-Hochschule (Advanced Controlling, 34).
- Weber, Jürgen; Sandt, Joachim (2001): Erfolg durch Kennzahlen: neue empirische Erkenntnisse. Vallendar: WHU, Otto-Beisheim-Hochschule (Advanced Controlling, 21).
- Weber, Jürgen; Schäffer, Utz (2016): Einführung in das Controlling. 15. Aufl. Stuttgart: Schäffer-Poeschel.
- Weber, Jürgen; Schaiër, Sven; Strangfeld, Oliver (2005): Berichte für das Top-Management: Ergebnisse einer Benchmarking-Studie. Weinheim: Wiley-VCH (Advanced Controlling, 43).
- Weißberger, Barbara E. (2011): IFRS für Controller: Einführung, Anwendung, Fallbeispiele. 2. Aufl. Freiburg: Haufe.
- Weißmann, Fritz (2005): Unternehmen steuern mit Controlling: Leitfaden und Toolbox für die Praxis. Berlin: Springer.
- Welge, Martin (1988): Unternehmensführung. Stuttgart: Poeschel (Controlling, 3).
- Wenig, Christina (2009): Die Prospect-Theorie. In: Manfred Schwaiger und Anton Meyer (Hg.): Theorien und Methoden der Betriebswirtschaft: Handbuch für Wissenschaftler und Studierende. München: Vahlen, S. 193–206.
- Whitney, Paul; Rinehart, Christa A.; Hinson, John M. (2008): Framing effects under cognitive load: The role of working memory in risky decisions. In: *Psychonomic Bulletin & Review* 15 (6), S. 1179–1184.

- Winkel, Heather; Ratitamkul, Theeraporn; Brambley, Victoria; Nagarachinda, Tulaya; Tiencharoen, Sutheemar (2016): Decision-making and the framing effect in a foreign and native language. In: *Journal of Cognitive Psychology* 28 (4), S. 427–436.
- Wirth, Thomas (2000): Leserorientierte Gestaltung von Managementberichten: Hinweise aus der angewandten Psychologie. In: *Controlling & Management* 44 (2), S. 79–85.
- Witt, Peter (2018): Besser entscheiden in unsicheren Situationen: Ziele bestimmen – Alternativen bewerten – Entschlüsse durchsetzen. Berlin: Erich Schmidt Verlag.
- Wu, George; Gonzalez, Richard (1996): Curvature of the probability weighting function. In: *Management Science* 42 (12), S. 1676–1690.
- Wu, George; Gonzalez, Richard (1999): Nonlinear decision weights in choice under uncertainty. In: *Management Science* 45 (1), S. 74–85.
- You, Haifeng; Zhang, Xiao-jun (2009): Financial reporting complexity and investor underreaction to 10-K information. In: *Review of Accounting Studies* 14 (4), S. 559–586.
- Yue, Carole L.; Castel, Alan D.; Bjork, Robert A. (2013): When disfluency is – and is not – a desirable difficulty: The influence of typeface clarity on metacognitive judgments and memory. In: *Memory and Cognition* 41 (2), S. 229–241.
- Zayer, Eric; Hirsch, Bernhard (2006): Fehlentscheidungen bei Investitionsprojekten: Fehlerquellen und Gegenmaßnahmen. In: *Controlling* 18 (12), S. 647–658.
- Zelazny, Gene (2015): Wie aus Zahlen Bilder werden: Der Weg zur visuellen Kommunikation. 7. Aufl. Wiesbaden: Springer.
- Zimmerman, Jerold L. (2017): Accounting for decision making and control. 9. Aufl. New York: McGraw-Hill.